

The present invention relates to a new β -galactosidase with transgalactosylating activity capable of converting lactose to a mixture of galactooligosaccharides. In particular
5 it relates to a β -galactosidase isolated from a recently discovered strain of *Bifidobacterium bifidum*.

The invention particularly relates to DNA sequences encoding the isolated new β -galactosidase enzyme, to the enzyme encoded by such a DNA sequence and to a host cell
10 comprising the DNA sequence or containing a recombinant vector incorporating the DNA sequence. The invention also relates to the use of the enzyme encoded by a DNA sequence, or of the host cell containing a DNA sequence or recombinant vector, to produce galactooligosaccharides.

15 Bifidobacteria naturally colonise the lower intestinal tract, an environment which is poor in mono and disaccharides since such sugars are preferentially consumed by the host and microbes present in the upper intestinal tract. In order to survive in the lower intestinal tract bifidobacteria produce various kinds of exo- and endoglycosidases in surface bound and/or extracellular forms, by which they can utilise diverse carbohydrates.

20

Besides hydrolase activity, some enzymes from bifidobacteria show transferase activity. This transglycosylation activity of glycosidases is extensively used for the enzymatic synthesis of various oligosaccharides, which have proven to act as bifidobacteria growth promoting factors.

25

It is known that members of bifidobacteria produce β -galactosidase enzymes that are involved in the bacterial metabolism of lactose. Møller, P.L. et al in Appl & Environ. Microbial., (2001), 62, (5), 2276-2283 describe the isolation and characterisation of three β -galactosidase genes from a strain of *Bifidobacterium bifidum*. They found that all three β -
30 galactosidases were able to catalyse the formation of beta-linked galactooligosaccharides

by transgalactosylation.

Dumortier et al in Carbohydrate Research, 201, (1990), 115-123 described the formation of beta-linked oligosaccharides by a transgalactosylation reaction during lactose
 5 hydrolysis with *Bifidobacterium bifidum* DSM 20456. Their analysis of the structure of the mixture of oligosaccharides produced showed that the linkages were β -(1 $\rightarrow$ 3), β -(1 $\rightarrow$ 6) and β -(1 $\rightarrow$ 4)-D-galactosyl linkages. Dumortier suggested that compounds produced by *Bifidobacterium bifidum* are involved in the adherence of bacteria in the large intestine.

10 WO 01/90317 describes a new β -galactosidase from *Bifidobacterium bifidum*, in particular a truncated version of the enzyme that has a high transgalactosylating activity.

A strain of *Bifidobacterium bifidum* has been found that is capable of producing a galactosidase enzyme activity that converts lactose to a novel mixture of
 15 galactooligosaccharides which unexpectedly contains up to 35% of disaccharides including galabiose (Gal (α 1-6) - Gal). This disaccharide is known (see Paton, J C and Paton, A W (1998), Clin. Microbiol. Revs., 11, 450-479; Carlsson, K A (1989), Ann. Reviews Biochem., 58, 309-350) to be an antiadhesive capable of preventing the adhesion of toxins, eg Shiga toxin and pathogens such as *E. coli*, to the wall of the gut.

20

This strain of *B bifidum* was deposited under accession number NCIMB 41171 at the National Collection of Industrial & Marine Bacteria, Aberdeen, UK on 31 March 2003. It is also described in UK Patent No 2 412 380.

25 It has now been found that this strain of *B bifidum* produces several β -galactosidases, including a novel β -galactosidase. This enzyme produces a number of different oligosaccharides which are β -linked.

30 According to the invention there is provided a DNA sequence which encodes a protein with an amino acid sequence as given in SEQ. ID NO: 2 or hybridises under

stringent conditions to the DNA sequence which encodes this protein. The DNA sequence is given in SEQ. ID NO: 1 or may comprise a fragment or degenerative thereof.

The phrase "degenerative" is construed to mean a DNA sequence which is at least
 5 50% homologous to SEQ ID NO : 1, preferably from 50 to 98% homologous, most preferably from 75 to 95% homologous.

Such a DNA sequence may comprise nucleotide substitutions, additions or deletions which result in less than 60%, preferably less than 45%, more preferably less than 25%
 10 change in the amino acid sequence shown in SEQ. ID NO: 2. Nucleotide substitutions may result in conservative amino acid substitutions.

According to a second aspect of the invention there is provided an enzyme encoded by a DNA sequence as defined above. Such an enzyme may comprise the amino acid
 15 sequence given in SEQ. ID NO: 2 or a fragment thereof.

According to a third aspect of the invention there is provided a recombinant vector, preferably an expression vector, comprising a DNA sequence as defined above. Such a vector may be incorporated into a host cell such as a bacterial, yeast or fungal cell.
 20 Alternatively, the DNA sequence may be incorporated into such a host cell. A suitable host cell may be selected from the group comprising Bifidobacterium, Lactococcus, Lactobacillus, Bacillus for example Bacillus subtilis or Bacillus circulans, Escherichia and Aspergillus for example Aspergillus niger.

25 Using lactose as a substrate, the enzyme encoded by the DNA sequence as defined above produces a mixture of oligosaccharides, comprising disaccharides, such as Gal (β 1-3) Glc, Gal (β 1-3) Gal, Gal (β 1-6) Gal and Gal (α 1-6) Gal, trisaccharides and tetrasaccharides such as Gal (β 1-6) Gal (β 1-4) Glc, Gal (β 1-3) Gal (β 1-4) Glc and Gal (β 1-6) Gal (β 1-6) Gal (β 1-4) Glc
 30

The enzyme or the host cell as described above may be used to produce a mixture of galactooligosaccharides which may form part of a product for improving gut health. Such a

product may be selected from the group consisting of dairy products (for example liquid milk, dried milk powder such as whole milk powder, skimmed milk powder, fat filled milk powders, whey powders, baby milks, baby formula, ice cream, yoghurt, cheese, fermented dairy products), beverages such as fruit juice, infant foods, cereals, bread, biscuits,
 5 confectionery, cakes, food supplements, dietary supplements, animal feeds, poultry feeds or indeed any other food or beverage. The presence of galactooligosaccharides in such products has the advantage of enhancing the growth of health-promoting Bifidobacterium in the product or in the intestinal flora of the consumer after intake of the product or both.

10 Alternatively, the oligosaccharides so produced may be used for the preparation of a medicament, for example in tablet or capsule form, for preventing the adhesion of pathogens or toxins produced by pathogens to the gut wall. The medicament may be administered to a patient, for example following a course of antibiotic treatment, which often alters or even destroys the normal healthy gut flora.

15 According to yet a further aspect of the invention there is provided a process for producing an enzyme as defined above which comprises culturing a host cell as defined above in a suitable culture medium under conditions permitting expression of the enzyme and recovering the resulting enzyme or enzyme products from the culture.

20 The invention is also directed to a process for producing the mixture of galactooligosaccharides which comprises contacting the enzyme as defined above with a lactose-containing material under conditions that lead to the formation of the galactooligosaccharide mixture.

25 Suitable lactose containing material may be selected from commercially available lactose, whole milk, semi-skimmed milk, skimmed milk, whey, fat-filled milk and whey permeate. Such milk products may be obtained from cows, buffaloes, sheep or goats. Fat-filled milk is defined as whole milk that has been skimmed to remove the dairy fat, which is
 30 subsequently replaced by the addition of vegetable fat or oil.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence (SEQ. ID NO: 1) of *Bifidobacterium bifidum* β -galactosidase of the invention; and

Figure 2 shows the amino acid sequence (SEQ. ID NO: 2) corresponding to the nucleotide sequence of Figure 1.

5 Figure 3 is a graph showing the time course reaction during galactooligosaccharide synthesis with β -galactosidase and 40% (w/w) lactose in 0.1M phosphate buffer at pH 6.0 as substrate; and

Figure 4 shows a high performance anion exchange chromatogram of the galactooligosaccharide mixture synthesized by the β -galactosidase from *B. bifidum*
 10 NCIMB 41171 using 40% (w/w) lactose in 0.1M phosphate buffer at pH 6.0 as substrate (Glc = glucose, Gal = galactose, Lac = lactose, α (1-6) = galactobiose, DP = degree of polymerisation).

Genomic DNA was isolated from the *Bifidobacterium bifidum* strain (NCIMB 41171) using the method of Lawson *et al.* (1989) Fems Microbiol Letters, 65, (1-2), 41-45.
 15 The DNA was digested with restriction enzymes and fragments having a maximum size of 15 kbp were ligated with pSP72 vector which had been digested with the same restriction enzymes. *E. coli* cells were transformed with a vector containing insertions consisting of PstI, Eco RI, Bam HI, KpnI, SmaI or HindIII digested chromosomal DNA from the *B. bifidum*. Clones with β -galactosidase activity were selected on Luria Bertani agar plates
 20 containing p-nitrophenyl, X- β -Gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) and isopropyl- β -D-thiogalactoside (IPTG). Ligation mixtures with Pst I chromosomal DNA gave rise to thirteen β -galactosidase positive clones, one of which is identified as pP2.

DNA sequencing of the inserted DNA fragment P2 was performed using the dideoxy chain-termination method of Sanger (Russel P., 2002 iGenetics, Pearson
 25 Education, Inc., San Francisco, 187-189) using the BigDye Terminator V.3.0 cycle sequencing kit (Applied Biosystems, USA). The DNA sequence of P2 is shown in Figure 1 (SEQ. ID NO: 1).

The open reading frame (ORF) was located by using the ORF finder from NCBI
 30 (National Center of Biotechnology Information). The nucleotide sequence of Figure 1 was translated in all six possible reading frames and one open reading frame of 738 amino acids

encoding a putative β -galactosidase was identified. The translation is shown in Figure 2 (SEQ. ID NO: 2).

The present invention will be further described by way of reference to the following
5 example.

Example 1

Materials and Methods

10 All chemicals and media preparations used throughout this study were obtained from Sigma (Dorset, UK), Invitrogen (Paisley, UK), Oxoid (Basingstoke, UK), Qiagen (West Sussex, UK) and Promega (Southampton, UK).

Bacterial Strains

15 The *Bifidobacterium bifidum* strain (NCIMB 41171) was maintained on cryogenic beads in Microbank tubes at -70°C . For later experiments, the strain was revived on Wilkinson Chalgren (WC) agar (Oxoid, UK) and TPY medium (trypticase phytone yeast extract medium) and grown anaerobically (CO_2 and N_2 composition 80% and 20% respectively) at 37°C for 48 hours. The colony morphology and the absence of
20 contamination were tested by gram staining.

E. coli strains

Escherichia coli strain DH5a used in this study was commonly incubated under aerobic conditions at 37°C in Luria Bertani (LB) agar or broth (Sambrook J. and Russell W. D.
25 (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York) and when necessary was supplemented with antibiotics (100 $\mu\text{g}/\text{ml}$ ampicillin and/or 15 $\mu\text{g}/\text{ml}$ chloramphenicol) and 40 μl of 2% X- β -Gal, 7 μl of 20% (isopropyl- β -D-thiogalactoside) IPTG which were applied on the surface of a pre-made 90mm agar plate.

30 *E. coli* DH5a strain (Invitrogen, Paisley, UK) (genotype: $\text{F}^- \phi 80\text{lacZ}\Delta\text{M} \Delta(\text{lacZYA-argF})\text{U169 recA1 endA1 hsdR17}(\text{r}_k^-, \text{m}_k^-) \text{phoA supE44 thi-1 gyrA96 relA1}\lambda^-)$ is an α -

galactosidase positive strain and was used in expression experiments and for other genetic manipulations.

Genomic DNA Extraction from *Bifidobacterium bifidum*

5 Genomic DNA was isolated from the *Bifidobacterium bifidum* strain (NCIMB 41171) using the following method in which chromosomal DNA was prepared from cell pellet harvested from 100 ml of WC anaerobe broth. The cells were resuspended in 10 ml TES buffer (10 mM Tris-HCl, 10 mM EDTA, 10 mM NaCl, pH8) and treated with 200 µl of lysozyme/mutanolysin mixture (4:1, lysozyme 10 mg/ml mutanolysin 1 mg/ml) for 30
10 minutes at 37°C. The cells were then treated with 200 µl of proteinase K (at 20 mg/ml) and 200 µl of RNase (both 10 mg/ml) and incubated for one hour at 65°C. Finally the cells were treated with 2 ml of 10% SDS and incubated for 15 minutes at 65°C. 12 ml of phenol/chloroform were added and the extraction was repeated until the water phase could easily be separated from the interphase. The genomic DNA was precipitated with
15 isopropanol and resuspended in 10 mM Tris-HCl – 1 mM EDTA (pH8). The genomic DNA was then digested with restriction enzymes, ligated into pSP72 digested with the same enzymes and treated with alkaline phosphatase. Digestion of *B. bifidum* genomic DNA was performed using *EcoRI*, *PstI*, *BamHI*, *SmaI* and *KpnI*. Ligation mixtures were used to transform *E. coli* DH5a and β-galactosidase positive clones were identified as blue
20 colonies on X-Gal-containing plates.

Vector DNA Preparation

 The vector used for cloning and expression throughout this study was the pSP72 (Promega, UK) (Krieg, P.A. and Melton, D.A. (1987). In vitro RNA synthesis with SP6
25 RNA polymerase. *Methods in Enzymology*. 155: 397-415).

 This vector was chosen due to the lack of complementing activity of the α-fragment of β-galactosidase which is not encoded in pSP72. This vector does not carry the short segment of *E. coli* DNA containing the regulatory sequence and the coding information for the first
30 146 amino acids of β-galactosidase which in combination with *E. coli* strains (ie DH5a)

which express the carboxy-terminal portion of this β -galactosidase is giving an active β -galactosidase (α -complementation).

The vector was digested with the following restriction enzymes: *Pst*I, *Bam*HI,
 5 *Hind*III, *Sma*I, *Kpn*I and *Eco*RI according to the manufacturer instructions using a tenfold excess of enzyme over DNA (enzyme units : μ gr DNA equal to ten units of enzyme per one μ gr of plasmid DNA or ten enzyme units per 0.5 pmol of plasmid DNA). After enzyme heat inactivation (20 min at 65°C) the restriction patterns were analysed by horizontal gel electrophoresis analysis. The presence of a single fragment in the gel indicated the
 10 complete vector digestion and the single restriction digestion of it.

The sufficient digestion of the vector was tested also by transforming unligated molecules into competent *E. coli* DH5a cells. The number of formed colonies on LB agar plates supplemented with ampicillin (100 μ gr/ml) was an indicator of the undigested
 15 molecules and the expected background during the subsequent experiments.

The vectors were further dephosphorylated with calf intestinal alkaline phosphatase, CIAP (Promega, Southampton, UK) according to the manufacturer instructions. The efficiency of the treatment was tested by ligation (with Bacteriophage T4 DNA ligase
 20 according to manufacturer instructions) following transformation into DH5a cells. The number of formed colonies showed the number of recircularised molecules (non cloned vector) and a subtraction of the above with the formed colonies without CIAP vector treatment showed the number of non dephosphorylated vectors.

25 Genomic DNA Library Construction

Genomic DNA was partially digested with six restriction enzymes that recognise frequently occurring hexa-nucleotide sequences within prokaryotic DNA. *Eco*RI, *Bam*HI, *Pst*I, *Kpn*I, *Sma*I and *Hind*III are type II restriction endonucleases specifically recognizing the sequences 5'G/AATTC'3, 5'G/GATCC'3, 5'CTGCA/G'3, 5'GGTAC/C'3,
 30 5'CCC/GGG'3 and 5'A/AGCTT'3 respectively, and make double-strand breaks within these sequences generating 5'overhangs of four nucleotides, AATT, GATC, AGCT for

EcoRI, *BamHI* and *Hind III* respectively, and 3' overhangs, ACGT, GTAC for *PstI* and *KpnI* respectively and blunt ends for *SmaI*.

- 5 All these enzymes were active and able to cleave DNA only in the presence of divalent magnesium ions. These ions were the only required cofactor.

Restriction Digestion of DNA.

- 10 All restriction digestions of the genomic DNA samples were incubated for 2 hours at 37°C and finally heat inactivated at 65°C for 20 minutes. The reactions were then cooled at room temperature and the appropriate amount of loading buffer was added, followed by gentle mixing with a sealed glass capillary. The solutions then were loaded into wells of a 0.8% agarose gel (power supply 4-5volts/cm for 14-16 hours) and the size of the digested DNA was estimated with that of 1kbp DNA standards (Promega, UK) (Sambrook J. Molecular Cloning: A Laboratory Manual (2002)).

Purification of the fragments generated after restriction digestion.

Fragment purification from the reaction mixtures and the agarose gels was done by using the QIAEX gel extraction kit from Qiagen (West Sussex, UK). Protocols are described with details in the manufacturer's manual.

20 DNA Ligation and Transformation

After purification of the DNA fragments with the Qiaex gel extraction kit, they were ligated with CIAP-treated pSP72 vector. For ligation, appropriate amounts of DNA were transferred to sterile 0.5 ml microfuge tubes as shown in Table 1.

Tube	DNA
A	Vector (15 fmoles [~ 29.7 ng])
B	Vector (15 fmoles ~ 29.7 ng DNA) plus insert (foreign 15 fmoles ~ 69.3 ng)
C	pUC control (0.056 fmoles [~ 100 pg])
The molar ratio of plasmid DNA vector to insert DNA fragment should be $\sim 1:1$ in the ligation reaction. The final DNA concentration should be ~ 10 ng/ μ l.	

Table 1: Ligation mixtures. Tube A shows the number of self-ligated vector DNA which must be subtracted from the total number of transformants after transformation. Tube B shows the ligation of the vector with the DNA fragments and tube C shows the control in order the transformation efficiency to be calculated.

5

Before each ligation the DNA fragments were warmed at 45°C for 5 minutes to melt any cohesive termini that reannealed during fragment preparation. A molar ratio of vector:insert DNA of 1:1 was chosen for all ligation reactions and the reaction assembly was done according to Promega's instructions.

10

To tubes A and B 1.0 µl of 10x ligation buffer and 0.5 Weiss units of T4 DNA ligase (Promega, UK) were added and the ligation volume was adjusted to 10 µl with molecular biology grade water. To tubes C 1.0 µl of 10x ligation buffer were added and the ligation volume was adjusted to 10 µl with molecular biology grade water.

15

DNA fragments were added to the tubes together with the water and then warmed to 45°C for 5 minutes to melt any cohesive termini that were reannealed during preparation. The DNA was chilled to 0°C before the remainder of the ligation reagents were added and the reaction mixtures were incubated overnight at 16°C (Sambrook and Russell, 2001).

20

After ethanol precipitation and purification of the ligated fragments (in order to remove the ligation mixture which causes a reduction of the transformation efficiency) transformations were performed according to Hanahan instructions. ~50ng of ligated DNA in 5µl solution was added to 100µl of competent DH5a cells. After heat treatment and expression of the ampicillin resistance gene the cells were spread over the surface of LB plates containing ampicillin (100µg/ml), X-β-Gal (40µl of 2% X-β-Gal) and IPTG (7µl of 20%IPTG).

25

The number of transformants from each ligation reaction was measured. The number of transformants commonly obtained from tube C was 2×10^5 - 1×10^6 cfu/µg whereas from tube A was 500-600 cfu/µg. The number of transformants in tube A was an indication

30

of the efficient treatment of the vector DNA. The number of transformants in tube B was in a range from $2-4 \times 10^4$ cfu/ μ g.

Number of Transformants

Ligation mixtures with *Pst*I chromosomal DNA gave rise to 13 β -galactosidase positive clones out of ~2500 screened transformants whereas with *Bam*HI gave rise to 7 positive clones (~1500 scr. transformants), *Eco*RI gave rise to 3 positive clones (~1300 scr. transformants), *Kpn*I gave rise to 7 positive clones (~2000 scr. transformants), *Sma*I gave rise to 3 positive clones (~1600 scr. transformants) and *Hind*III gave rise to 2 positive clones (~1200 scr. transformants).

Positive Clone Digestion

In order to identify the different β -galactosidase genes, the plasmids isolated from the positive clones were digested according to the following table:

	Samples	Enzymes
1 st Digestion	pB1, pB2, pB3, pB4, pB5, pB6, pB7	<i>Bam</i> HI
2 nd Digestion	pP1, pP2, pP3, pP4, pP5, pP6, pP7, pP8, pP9, pP10, pP11	<i>Pst</i> I
3 rd Digestion	pP12, pP13, pP14	<i>Pst</i> I
4 th Digestion	pE1, pE2, pE3	<i>Eco</i> RI
5 th Digestion	pP1, pP12, pB1, pP2, pE1, pE2, pE3.....	<i>Pst</i> I and <i>Eco</i> RI
6 th Digestion	pS1, pS2, pS3	<i>Sma</i> I
7 th Digestion	pP1, pP12, pB1, pP2, pS1, pS2, pS3	<i>Pst</i> I and <i>Sma</i> I
8 th Digestion	pK1, pK2, pK3, pK4, pK5, pK6, pK7	<i>Kpn</i> I
9 th Digestion	pP1, pP12, pB1, pP2, pK1, pK2, pK3, pK4, pK5, pK6, pK7	<i>Pst</i> I and <i>Kpn</i> I
The first letter (p) indicates plasmid and the insert gene whereas the second letter (P,B,E,S,K) indicates the restriction enzyme that was used for isolation of the respective clone from the genomic DNA.		

Gel electrophoresis analysis of the generated fragments after digestion showed that plasmids pB1, pP1, pP2 and pP11 each have an insert which encodes a different β -galactosidase. The clones containing P2 were used for further analysis.

DNA Sequencing

DNA sequencing was performed with the dideoxy chain-termination method of Sanger by using the BigDye Terminator v.3.0 cycle sequencing kit (Applied Biosystems, USA) and analysed with the ABI Prism 3100, a fluorescence-based DNA analysis system
5 incorporating capillary electrophoresis.

The 5'- and 3'- ends of the insert DNA fragments were sequenced with vector specific primers. The inserts were further sequenced by using the Genome Priming System (GPS-1) (New England Biolabs, Uk). GPS-1 is a TN7 transposon-based in vitro system
10 which uses TnsABC Transposase to insert Transposon randomly into the DNA target. The donor: target DNA mass ratio of 1:4 was used according to the manufacturer instructions. The number of isolated plasmids for sequencing after insertion of the Transprimer into the target plasmid was 25. This number was calculated according to the manufacturer instructions and it assumes a 5-fold depth of coverage.

15

Due to the long nucleotide sequence of pP2 plasmid in which protein P2 was encoded, only the part which contained the β -galactosidase gene was chosen to be sequenced. The enzyme inactivated after insertion of the transposase insert at a relative position of 172bp of the sequenced fragment indicated that the start codon was upstream of
20 this position. Similarly, insertion of insert at position 2882bp completely eliminated the enzyme activity indicating that the stop codon existed downstream of this position. Moreover, the enzyme activity was eliminated completely with insertion of inserts at positions 262bp, 331bp, 375bp, 621bp, 866bp, 1348bp, 1358bp, 1394bp, 1513bp, 1704bp, 2128bp, 2519bp in respect of the first nucleotide that has been sequenced.

25

Analysis of the N-terminal domain with SignalIP and PSORT software, did not show any signal peptide, indicating that P2 is not secreted extracellularly.

The sequencing reaction mix contained approximately 400-600ng plasmid DNA,
30 3.2pmol of primer solution and 4 μ l of BigDye Terminator solution.

Open Reading Frame Identification

The open reading frame (ORF) of P2 was located by using the ORF finder from NCBI web address <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). The bacterial genetic code was used and the frame length was determined to be 100bp. The nucleotide sequence was translated in all six possible frames and an open reading frame of 738 amino acids encoding a putative β -galactosidase was identified (The translation is shown in Figure 2).

Example 2

10 Synthesis with the β -galactosidase cloned enzyme isolated from *Bifidobacterium bifidum* NCIMB 41171 in *E. coli* host (strainDH5a)

The following described synthesis, unless otherwise stated, was performed with the whole *E. coli* DH5a host cells after treatment of the *E. coli* biomass (collected by centrifugation at 10,000 g) with toluene at a concentration of 2000 ppm in order to increase cell permeability and also to render the cells non-viable by destroying their cytoplasmic membrane. The *E. coli* biomass was prepared as described in Example 1 under “*E. coli* strains”.

Synthesis with cloned enzyme

20 Synthesis with the β -galactosidase was performed at a substrate concentration of 40% (w/w) initial lactose concentration. The synthesis solution was prepared in 0.1 M phosphate buffer at pH 6.0 containing additional 1 g/l Tween 80 (polyoxyethylene (20) sorbiton monooleate). Synthesis was performed at 40°C in shaking waterbath at 150 rpm. The pH optimum for the specific enzyme was chosen based on activity measurements (using o-nitrophenyl- β -D-galactopyranoside as substrate) of a specific enzymatic preparation at varying pH values.

For galactooligosaccharide synthesis 2 ml of cell lysate supernatant were used (after disruption of the *E. coli* cells by French press) with 8g of 50% (w/w) lactose in order to give a final substrate concentration of 40% (w/w). This enzymatic preparation had an activity of 735 U/ml.

The concentrations of the different sugars present in the mixture during synthesis are shown in Figure 3. High performance anion exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD) chromatograms of galactooligosaccharide mixtures synthesized by the β -galactosidase cloned from *B. bifidum* NCIMB 41171 are shown in Figure 4. The galactooligosaccharide mixture sugar concentrations at the optimum synthesis time point are shown in table 1.

Table 1. Carbohydrate composition of galactooligosaccharide synthesis at 40 % (w/w) initial lactose concentration at the time point where maximum oligosaccharide concentration was observed.

Synthesis Init. Subst.	GOS DP $\geq$ 3	GOS DP=2	Lac	Glc	Gal
% (w/w)	Concentration (% of total sugars)				
40	8.82	16.25	39.40	20.76	14.85

15

Lac: Lactose, Glc: glucose, Gal: galactose, DP: degree of polymerisation

CLAIMS

1. A DNA sequence encoding a β -galactosidase, wherein the sequence is selected from the sequence given in SEQ. ID NO: 1 or a fragment or a degenerative thereof, said degenerative being from 75% to 95% homologous.
2. The DNA sequence according to Claim 1, wherein said sequence comprises nucleotide substitutions, additions or deletions which result in less than 25% change in the amino acid sequence according to SEQ. ID NO: 2.
3. A DNA sequence according to Claim 1 or Claim 2, wherein said sequence comprises nucleotide substitutions which result in conservative amino acid substitutions.
4. A β -galactosidase enzyme encoded by a DNA sequence of any one of Claims 1 to 3.
5. A β -galactosidase enzyme comprising an amino acid sequence according to SEQ. ID NO: 2.
6. A recombinant vector comprising a DNA sequence of any one of Claims 1 to 3.
7. The vector according to Claim 6, wherein said vector is an expression vector.
8. A host cell which is a yeast cell, a fungal cell or a bacterial cell selected from the group consisting of Lactococcus, Lactobacillus, Escherichia, Bacillus and Aspergillus comprising a DNA sequence of any one of Claims 1 to 3.
9. A host cell comprising the vector of Claim 6 or Claim 7.
10. The host cell according to Claim 8 or Claim 9, wherein the cell is selected from the group consisting of Bacillus subtilis, Bacillus circulans and Aspergillus niger.
11. Use of an enzyme of Claim 4 or Claim 5, or a cell of any one of Claims 8 to 10, for producing a mixture of oligosaccharides.
12. Use of an enzyme of Claim 4 or Claim 5, or a cell of anyone of Claims 8 to 10, for producing a mixture of oligosaccharides to be part of a product selected from the group consisting of dairy products, such as liquid milk, dried milk powder, baby milks, baby formula, ice cream, yoghurt, cheese, fermented dairy products, beverages such as fruit juice, infant foods, cereals, bread, biscuits, confectionary, cakes, food supplements, dietary

supplements, probiotic comestible products, prebiotic comestible products, animal feeds, poultry feeds and medicaments.

13. Use of a host cell of any one of Claims 8 to 10, for producing a product selected from the group consisting of dairy products, such as liquid milk, dried milk powder, baby milks, baby formula, ice cream, yoghurt, cheese, fermented diary products, beverages such as fruit juice, infant foods, cereals, bread, biscuits, confectionary, cakes, food supplements, dietary supplements, probiotic comestible product, prebiotic comestible product, animal feeds, poultry feeds and medicaments.

14. A process for producing an enzyme of Claim 4 or Claim 5, comprising culturing a host cell of any one of Claims 8 to 10 in a suitable culture medium under conditions permitting expression of said enzyme and recovering the resulting enzyme from the culture.

Application (text form)
SEQUENCE LISTING

<110> McDONALD, Sarah C
<120> Product and Process
<130> 14011:KG
<150> GB 0606112.1
<151> 2006-03-28
<160> 2
<170> PatentIn version 3.4
<210> 1
<211> 4395
<212> DNA
<213> Bifidobacterium bifidum

<220>
<221> Beta-galactosidase
<222> (1)..(4395)

```

<400> 1
tccttttagat tggattgaaa acttcctttt tttatgtaaa aggatattcg gggaaattcc      60
ctttttgaaa atttcatgac caaaaatccc cttaacgtga gttttcgttc cactgagcgt      120
cagaccccggt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct      180
gctgcttgca aacaaaaaaa cccccgctac cagcggtggt ttgtttgccg gatcaagagc      240
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc      300
ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc      360
tcgctctgct aatcctgtta ccagtggttg ctgccagtgg cgataagtcg tgtcttaccg      420
ggttgggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt      480
cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg      540
agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggat ccggtaaagc      600
gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt      660
atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag      720
gggggcggag cctatggaaa aacgccagca acggggcctt tttacggttc ctggcctttt      780
gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta      840
ttaccgcctt tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt      900
cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc      960
cgattcatta atgcaggtta acctggctta tcgaaattaa tacgactcac tatagggaga     1020
ccggcagatc tgatatcatc gatgaattcg agctcggtag ccggggatcc tctagagtcg     1080
acctgcagaa gtactggcag tccggcctcg ccgcccggcg cgtaaggag tgatctcgcg     1140
ggagggccag tgtccgcttg cagcatgcgt cacgttcgtc tgacgggtct gacgtgcgcc     1200
tagcggctgg ccctccctac ggaacatgta tatgatatgg agggacgctc accgtgtgcg     1260

```

Application (text form)

atgcgggctg gcgcgatatg gcaccagccc gcatcgcatg tcacagttga tttttcgctg	1320
atgcgacggg tgggcccggg cgtcagacaa tgaacaagta atggagcggg atatgagtaa	1380
acgcagaaaag cacagttggc cgcagccgct gaagggcgcc gaatcccgtc tctggtatgg	1440
cgggtgattac aatcccagacc aatggccgga ggagggtctgg gacgatgata tccgtctgat	1500
gaagaaggcc ggcgtcaacc ttgtatcagt cggcatcttc tcatgggcga agatcgagcc	1560
ggaggaaggc aagtacgatt tcgactggct ggaccgtgcc atcgataagc tcggcaaggc	1620
gggcatcgct gtagacctcg cctccgccac cgcctccccg ccgatgtggc tgaccaggc	1680
ccatccggag gtgctgtgga aggacgagcg cggcgacacc gtgtggccgg gtgcccgtga	1740
gcattggcgt cggacgagcc ctgtgttccg cgagtacgcy ctgaacctgt gccgtcgtat	1800
ggccgagcac tataagggca acccatacgt ggtggcctgg cacgtgagca acgaatacgg	1860
ctgccacaac cgcttcgact actcggacga cgccatgcgc gccttccaga agtggtgcaa	1920
gaagcgctac aagaccatcg acgcggtgaa cgaggcctgg ggaacagcgt tctgggcca	1980
gcacatgaac gatttctccg agatcatccc gcctcgctac atcggtgacg gcaacttcat	2040
gaaccggggc aagctgctcg actacaagcg gttcagctcc gacgcgctca aggagctcta	2100
catcgccgaa cgtgacgtgc tggagtccat cagccgggt ctgccgtga ccaccaattt	2160
catggtgtct gccggcggt cgatgcttga ctatgacgac tggggtgcgg aggtcgattt	2220
cgtctccaac gaccactact tcacgccggg cgaggaccat ttcgacgagg ttgcgtacgc	2280
cgcgtcgctg atggacggca tttcacgcaa ggaaccgtgg ttccagatgg agcattccac	2340
ctccgcggtg aattggcgct cgatcaacta ccgtgaggag ccgggctcgg tcgtgcgcga	2400
ctccctcggc caggctgcta tgggcgccga tgccatctgc tatttccagt ggcgacagtc	2460
caaggccggc gcggagaagt ggcactcctc gatggtgccg catgcgggcg aggattcgca	2520
gatcttccgt gatgtgtgcg agctgggcgc cgatctggga cgtctgtccg acgagggctt	2580
gatgggcacc aagacggtca agtccaaggt cgcgctcgtg ttcgactacg agtcccagtg	2640
ggccaccgag tacaccgcga acccgacca gcaggctgat cattggaccg agccgttgga	2700
ttggttccgc gcgcttgccg acaacggcat caccgccgac gtggttccgg tgcggtccga	2760
ttgggattcc tacgagattg ccgtgctgcc gtgcgtgtat ttgctgtccg aagagacgag	2820
ccgcagggtt cgggagttcg tggcgaacgg cggcaagttg ttcgtgacgt actacaccgg	2880
attgagcgac gagaacgacc acatctggct tggcggctac ccgggctcga ttcgtgacgt	2940
cgtcggcgtg cgcgtcgaag agttcgcccc gatgggcaac gacatgccgg gcgcattgga	3000
tcacctcgac ctggacaatg gcacggtggc gcatgacttt gccgacgtga tcacgtccac	3060
ggccgacacg tccacggtac tggcctccta taaggcgga cgtggaccg gcatgaacga	3120
ggtgccggcc atcgtggcca acgggtatgg cgacgggcgg accgtatatg tcggatgccg	3180
tctgggccgt caggggcttg cgaagagcct gccggcgatg ctgggttcca tggggctgtc	3240
ggacctcgcc ggtgatggcc gcgtgctgcg tgtcgaacgc gccgacgccg cggcggcgag	3300

Application (text form)

ccgcttcgag ttcgtgttca accgcaccca tgaaccggtg accgtcgacg ttgaagggga 3360
ggccatcgcg gcttcgctcg cgcattgtcga cgacggggcg gccaccatcg atccgacggg 3420
tggtgtcgtg ctcaggcgat aatcgttgga aacactgggc tgtaagggct taggaaaggc 3480
gtatgtttgc ggtgacacgc gacatacgcc ttatgggaaa gaaggcgctg gcgcttaccg 3540
ggctgcggcg atgctgggtc gcgtcgctgc gtgcggcaac tcaagcagca gctccgctcc 3600
gaagcaggaa ggcgacgtca aggaaatcac cgtgtgggct tggggcctac gctgactcag 3660
gtggccaagg acttcaaaaa aggagaccgg catcaatgtc aacctggtga acaccggcca 3720
gggcgacaag acctgggacg agttctatca ggacgccaag aagattcaca cccttggcga 3780
caactactac atcacgtccg acaccgggtg cgccggcttc tacgactcga tgacctggct 3840
ggccagtgcg acgctgttct ccaccgaagg cgagacggtc accattaacc tgactggcgc 3900
cccgaaggtc aaggcccgcg gtatcttcgg cgactacctt ggcaagtcct acaccggcaa 3960
ccagaagctg agcgatggcg tcgccgcttt gggaacaggc tctgaaggac tacgcgaagg 4020
atcagggcta caccgtcaag taaccttcgc agtcaagcaa tctggcggtg taatgaccgg 4080
aatacggtga ccttcggtca tcccttcctc gtgtgaaggc ccctcccctc aacagggagg 4140
ggccttcaca tatctgcccc tgttgcaacg cgcggtgaaa ctctacgatg agcgaattct 4200
tcccgacaca tcgagcacgc taaggagatg acatgacgat atcggcacgg ttgtggcgcg 4260
tgcacctgca tatctttgtg ttgctcaaga tctgtgagat gacactggca gcacgcctcc 4320
agcgcgccgc cggacacgcg caccctcatc accgaatgga cgggggacca tatcatgacg 4380
actctgatcg ccaat 4395

<210> 2
<211> 738
<212> PRT
<213> Bifidobacterium bifidum

<220>
<221> Beta-galactosidase
<222> (1)..(738)

<400> 2

Met Tyr Met Ile Trp Arg Asp Ala His Arg Val Arg Cys Gly Leu Ala
1 5 10 15

Arg Tyr Gly Thr Ser Pro His Arg Met Ser Gln Leu Ile Phe Arg Arg
20 25 30

Cys Asp Gly Ser Ala Gly Ala Ser Asp Asn Glu Gln Val Met Glu Arg
35 40 45

Asn Met Ser Lys Arg Arg Lys His Ser Trp Pro Gln Pro Leu Lys Gly
50 55 60

Ala Glu Ser Arg Leu Trp Tyr Gly Gly Asp Tyr Asn Pro Asp Gln Trp
Page 3

65	70	Application (text form)	75	80
----	----	-------------------------	----	----

Page 4

Application (text form)

340

345

350

Ala Ser Leu Met Asp Gly Ile Ser Arg Lys Glu Pro Trp Phe Gln Met
355 360 365

Glu His Ser Thr Ser Ala Val Asn Trp Arg Pro Ile Asn Tyr Arg Ala
370 375 380

Glu Pro Gly Ser Val Val Arg Asp Ser Leu Ala Gln Val Ala Met Gly
385 390 395 400

Ala Asp Ala Ile Cys Tyr Phe Gln Trp Arg Gln Ser Lys Ala Gly Ala
405 410 415

Glu Lys Trp His Ser Ser Met Val Pro His Ala Gly Glu Asp Ser Gln
420 425 430

Ile Phe Arg Asp Val Cys Glu Leu Gly Ala Asp Leu Gly Arg Leu Ser
435 440 445

Asp Glu Gly Leu Met Gly Thr Lys Thr Val Lys Ser Lys Val Ala Val
450 455 460

Val Phe Asp Tyr Glu Ser Gln Trp Ala Thr Glu Tyr Thr Ala Asn Pro
465 470 475 480

Thr Gln Gln Val Asp His Trp Thr Glu Pro Leu Asp Trp Phe Arg Ala
485 490 495

Leu Ala Asp Asn Gly Ile Thr Ala Asp Val Val Pro Val Arg Ser Asp
500 505 510

Trp Asp Ser Tyr Glu Ile Ala Val Leu Pro Cys Val Tyr Leu Leu Ser
515 520 525

Glu Glu Thr Ser Arg Arg Val Arg Glu Phe Val Ala Asn Gly Gly Lys
530 535 540

Leu Phe Val Thr Tyr Tyr Thr Gly Leu Ser Asp Glu Asn Asp His Ile
545 550 555 560

Trp Leu Gly Gly Tyr Pro Gly Ser Ile Arg Asp Val Val Gly Val Arg
565 570 575

Val Glu Glu Phe Ala Pro Met Gly Asn Asp Met Pro Gly Ala Leu Asp
580 585 590

His Leu Asp Leu Asp Asn Gly Thr Val Ala His Asp Phe Ala Asp Val
595 600 605

Ile Thr Ser Thr Ala Asp Thr Ser Thr Val Leu Ala Ser Tyr Lys Ala
Page 5

610

Application (text form)
615 620

Glu Arg Trp Thr Gly Met Asn Glu Val Pro Ala Ile Val Ala Asn Gly
625 630 635 640

Tyr Gly Asp Gly Arg Thr Val Tyr Val Gly Cys Arg Leu Gly Arg Gln
645 650 655

Gly Leu Ala Lys Ser Leu Pro Ala Met Leu Gly Ser Met Gly Leu Ser
660 665 670

Asp Leu Ala Gly Asp Gly Arg Val Leu Arg Val Glu Arg Ala Asp Ala
675 680 685

Ala Ala Ala Ser Arg Phe Glu Phe Val Phe Asn Arg Thr His Glu Pro
690 695 700

Val Thr Val Asp Val Glu Gly Glu Ala Ile Ala Ala Ser Leu Ala His
705 710 715 720

Val Asp Asp Gly Arg Ala Thr Ile Asp Pro Thr Gly Val Val Val Leu
725 730 735

Arg Arg