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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 October 2007 (04.10.2007)

PCT

(10) International Publication Number
WO 2007/110619 A1

(51) International Patent Classification:
C12N 5/10 (2006.01) *C12N 9/38* (2006.01)
C12N 15/63 (2006.01)

(21) International Application Number:
PCT/GB2007/001081

(22) International Filing Date: 27 March 2007 (27.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0606112.1 28 March 2006 (28.03.2006) GB

(71) Applicant *(for all designated States except US)*: **MCDONALD, Sarah, Caroline** [GB/ZA]; 2060 Voltaire Drive, Dainfern Valley, Extension 19, 2055 (ZA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TZORTZIS, Georgios** [GR/GB]; 23 Blenheim Road, Reading, Berkshire RG1 5NG (GB). **GOULAS, Athanasios, K.** [GR/GB]; Food Microbial Sciences Unit, School of Food Biosciences, The University of Reading, Reading, RG6 6AP (GB). **GOULAS, Theodoros** [GR/GB]; Food Microbial Sciences Unit, School of Food Biosciences, The University of Reading, Reading, RG6 6AP (GB).

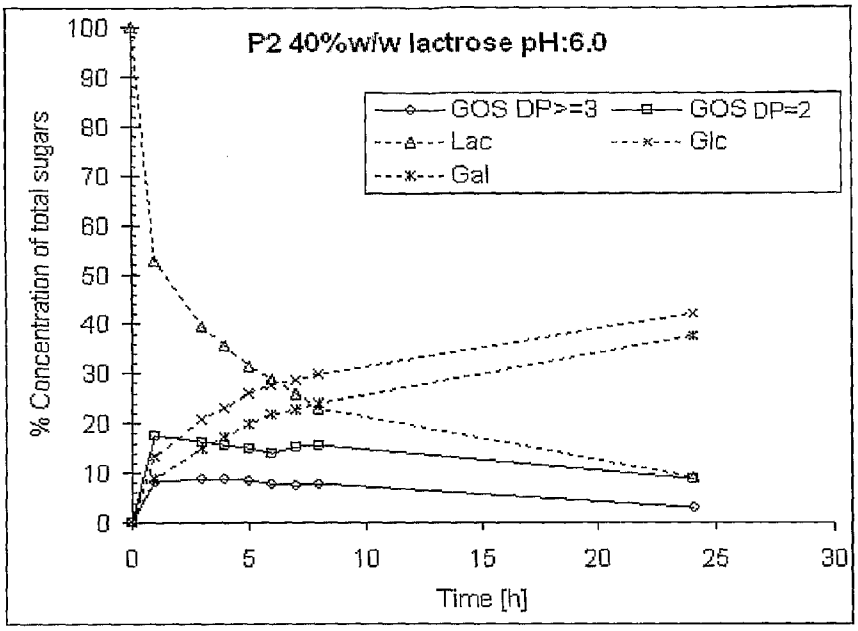
(74) Agent: **SAUNDERS & DOLLEYMORE**; 9 Rickmansworth Road, Watford, Hertfordshire WD18 0JU (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,

[Continued on next page]

(54) Title: BETA-GALACTOSIDASE WITH TRANSGALACTOSYLATING ACTIVITY



(57) Abstract: The present invention concerns a new β-galactosidase with transgalactosylating activity isolated from *Bifidobacterium bifidum*. The β-galactosidase is capable of converting lactose to a mixture of galactooligosaccharides which are β-linked and unexpectedly produces the α-linked disaccharide galactobiose. The mixture may be incorporated into numerous food products or animal feeds for improving gut health by promoting the growth of bifidobacteria in the gut, and repressing the growth of the pathogenic microflora.

WO 2007/110619 A1

ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

BETA-GALACTOSIDASE WITH TRANSGALACTOSYLATING ACTIVITY

The present invention relates to a new β-galactosidase with transgalactosylating activity capable of converting lactose to a mixture of galactooligosaccharides. In particular 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 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.

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.

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.

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-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 hydrolysis with *Bifidobacterium bifidum* DSM 20456. Their analysis of the structure of the mixture of oligosaccharides produced showed that the linkages were β-(1→3), β-(1→6) and β-(1→4)-D-galactosyl linkages. Dumortier suggested that compounds produced by *Bifidobacterium bifidum* are involved in the adherence of bacteria in the large intestine.

WO 01/90317 describes a new β-galactosidase from *Bifidobacterium bifidium*, 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 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.

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.

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.

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 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% 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 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. 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 subtilus or Bacillus circulans, Escherichia and Aspergillus for example Aspergillus niger.

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 (β1-3) Glc, Gal (β1-3) Gal, Gal (β1-6) Gal and Gal (α1-6) Gal, trisaccharides and tetrasaccharides such as Gal (β1-6) 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

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, 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.

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.

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.

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.

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 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.

Figure 3 is a graph showing the time course reaction during galactooligosaccharide synthesis with β-galactosidase and 40% (w/w) lactose in O.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 NCIMB 41171 using 40% (w/w) lactose in O.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. 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 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 Education, Inc., San Francisco, 187-189) using the BigDye Terminator V.3.O 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 (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 example.

Example 1

Materials and Methods

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

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 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. (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York) and when necessary was supplemented with antibiotics (100μg/ml ampicillin and/or 15μg/ml chloramphenicol) and 40μl of 2% X-β-Gal, 7μl of 20% (isopropyl-β-D-thiogalactoside) IPTG which were applied on the surface of a pre-made 90mm agar plate.

E. coli DH5a strain (Invitrogen, Paisley, UK) (genotype: F^- φ80*lac*ZΔM Δ(*lac*ZYA-argF)U169 recA1 endA1 hsdR17(r_k^-, m_k^-) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1λ^-) is an α-

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

Genomic DNA Extraction from *Bifidobacterium bifidum*

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 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 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 phasphatase. 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 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 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 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, *Hind*III, *Sma*I, *Kpn*I and *Eco*RI according to the manufacturer instructions using a tenfold excess of enzyme over DNA (enzyme units :μgr DNA equal to ten units of enzyme per one μgr 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 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 (100μgr/ml) was an indicator of the undigested 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 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.

Genomic DNA Library Construction

Genomic DNA was partially digested with six restriction enzymes that recognise frequently occurring hexa-nucleotide sequences within prokaryotic DNA. *EcoRI, BamHI, PstI, KpnI, SmaI* and *HindIII* are type II restriction endonucleases specifically recognizing the sequences 5'G/AATTC'3, 5'G/GATCC'3, 5'CTGCA/G'3 , 5'GGTAC/C3', 5'CCC/GGG3' and 5'A/AGCTT3' 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.

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.

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.

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 ~29.7ng DNA) plus insert (foreign 15 fmoles ~ 69.3ng)
C	pUC control (0.056 fmoles [~100 pg])

The molar ratio of plasmid DNA vector to insert DNA fragment should be ~1:1 in the ligation reaction. The final DNA concentration should be ~10ng/µl.

Table 1: Ligation mixtures. Tube A shows the number of self-ligated vector DNA which must be subtracted form 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.

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.

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.

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).

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 (100μgr/ml), X-β-Gal (40μl of 2% X-β-Gal) and IPTG (7μl of 20%IPTG).

The number of transformants from each ligation reaction was measured. The number of transformants commonly obtained from tube C was $2x10^5$-$1x10^6$cfu/μg whereas from tube A was 500-600 cfu/μg. The number of transformants in tube A was an indication

of the efficient treatment of the vector DNA. The number of transformants in tube B was in a range from 2-4x10^4cfu/µg.

Number of Transformants

Ligation mixtures with *Pst*I chromosomal DNA gave rise to 13 β-galactosidase positive clones out of ~2500 screeened 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
1st Digestion	pB1, pB2, pB3, pB4, pB5, pB6, pB7	*Bam*HI
2nd Digestion	pP1, pP2, pP3, pP4, pP5, pP6, pP7, pP8, pP9, pP10, pP11	*Pst*I
3rd Digestion	pP12, pP13, pP14	*Pst*I
4th Digestion	pE1, pE2, pE3	*Eco*RI
5th Digestion	pP1, pP12, pB1, pP2, pE1, pE2, pE3.....	*Pst*I and *Eco*RI
6th Digestion	pS1, pS2, pS3	*Sma*I
7th Digestion	pP1, pP12, pB1, pP2, pS1, pS2, pS3	*Pst*I and *Sma*I
8th Digestion	pK1, pK2, pK3, pK4, pK5, pK6, pK7	*Kpn*I
9th Digestion	pP1, pP12, pB1, pP2, pK1, pK2, pK3, pK4, pK5, pK6, pK7	*Pst*I and *Kpn*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 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 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.

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 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.

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, 3.2pmol of primer solution and 4μl 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

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

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 Figure3. 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≥3	GOS DP=2	Lac	Glc	Gal
% (w/w)	Concentration (% of total sugars)				
40	8.82	16.25	39.40	20.76	14.85

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

Claims

1. A DNA sequence which

a) encodes a protein with an amino acid sequence as given in SEQ. ID. NO: 2, or

b) hybridises under stringent hybridization conditions to the sequence of a), or

c) is a degenerative of the sequence of a) or b).

2. The DNA sequence according to Claim 1, wherein the sequence is given in SEQ. ID NO: 1 or a fragment thereof.

3. The DNA sequence according to Claim 1 or Claim 2, wherein said sequence comprises nucleotide substitutions, additions or deletions which result in less than 60%, preferably less than 45%, more preferably less than 25% change in the amino acid sequence according to SEQ. ID NO: 2 or a fragment thereof.

4. A DNA sequence according to Claim 1 or Claim 2, wherein said sequence comprises nucleotide substitutions which result in conservative amino acid substitutions.

5. An enzyme encoded by a DNA sequence of any one of Claims 1 to 4.

6. An enzyme comprising an amino acid sequence according to SEQ. ID NO: 2 or a fragment thereof.

7. A β-galactosidase having the sequence as defined in SEQ. ID NO: 2.

8. A recombinant vector comprising a DNA sequence of any one of Claims 1 to 4.

9. The vector according to Claim 8, wherein said vector is an expression vector.

10. A host cell comprising a DNA sequence of any one of Claims 1 to 4.

11. A host cell comprising the vector of Claim 8 or Claim 9.

12. The host cell according to Claim 10 or Claim 11, wherein said cell is a bacterial cell, a yeast cell or a fungal cell.

13. The host cell according to Claim 12, wherein said cell is selected from the group consisting of Bifidobacterium, Lactococcus, Lactobacillus, Escherichia, Bacillus and Aspergillus.

14. The host cell according to Claim 13, wherein the cell is selected from the group consisting of Bifidobacterium bifidum, Bacillus subtilis, Bacillus circulans and Aspergillus niger.

15. Use of an enzyme of any one of Claims 5 to 7, or a cell of any one of Claims 10 to 14, for producing a mixture of oligosaccharides.

16. Use of an enzyme of any one of Claims 5 to 7, or a cell of any one of Claims 10 to 14, for producing a mixture of oligosaccharides to be part of a product selected from the group consisting of diary 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 products, prebiotic comestible products, animal feeds, poultry feeds and medicaments.

17. Use of a host cell of any one of Claims 10 to 14, 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 supplement, probiotic comestible product, prebiotic comestible product, animal feeds, poultry feeds and medicaments.

18. A process for producing an enzyme of any one of Claim 5 to 7 comprising culturing a host cell of any one of Claims 10 to 14 in a suitable culture medium under conditions permitting expression of said enzyme and recovering the resulting enzyme from the culture.

5

1/5

Figure 1

```
   1 tcctttagat tggattgaaa acttcctttt tttatgtaaa aggatattcg gggaaattcc
  61 ctttttgaaa atttcatgac caaaaatccc cttaacgtga gttttcgttc cactgagcgt
 121 cagaccccgt agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct
 181 gctgcttgca aacaaaaaaa cccccgctac cagcggtggt ttgtttgccg gatcaagagc
 241 taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc
 301 ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc
 361 tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg
 421 ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt
 481 cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg
 541 agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg
 601 gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt
 661 atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag
 721 gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt
 781 gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta
 841 ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt
 901 cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc
 961 cgattcatta atgcaggtta acctggctta tcgaaattaa tacgactcac tatagggaga
1021 ccggcagatc tgatatcatc gatgaattcg agctcggtac ccggggatcc tctagagtcg
1081 acctgcagaa gtactggcag tccggcctcg ccgccggcgc cgtcaaggag tgatctcgcg
1141 ggagggccag tgtccgcttg cagcatgcgt cacgttcgtc tgacggtctt gacgtgcgcc
1201 tagcggctgg ccctccctac ggaacatgta tatgatatgg agggacgctc accgtgtgcg
1261 atgcgggctg gcgcgatatg gcaccagccc gcatcgcatg tcacagttga tttttcgtcg
1321 atgcgacggt tcggccgggg cgtcagacaa tgaacaagta atggagcgga atatgagtaa
1381 acgcagaaag cacagttggc cgcagccgct gaagggcgcc gaatcccgtc tctggtatgg
1441 cggtgattac aatcccgacc aatggccgga ggaggtctgg gacgatgata tccgtctgat
1501 gaagaaggcc ggcgtcaacc ttgtatcagt cggcatcttc tcatgggcga agatcgagcc
1561 ggaggaaggc aagtacgatt tcgactggct ggaccgtgcc atcgataagc tcggcaaggc
1621 gggcatcgct gtagacctcg cctccgccac cgcctccccg ccgatgtggc tgacccaggc
1681 ccatccggag gtgctgtgga aggacgagcg cggcgacacc gtgtggccgg gtgcccgtga
1741 gcattggcgt ccgacgagcc ctgtgttccg cgagtacgcg ctgaacctgt gccgtcgtat
1801 ggccgagcac tataagggca acccatacgt ggtggcctgg cacgtgagca acgaatacgg
1861 ctgccacaac cgcttcgact actcggacga cgccatgcgc gccttccaga agtggtgcaa
1921 gaagcgctac aagaccatcg acgcggtgaa cgaggcctgg ggaacagcgt tctgggccca
1981 gcacatgaac gatttctccg agatcatccc gcctcgctac atcggtgacg gcaacttcat
2041 gaacccgggc aagctgctcg actacaagcg gttcagctcc gacgcgctca aggagctcta
2101 catcgccgaa cgtgacgtgc tggagtccat cacgccgggt ctgccgctga ccaccaattt
2161 catggtgtct gccggcggct cgatgcttga ctatgacgac tggggtgcgg aggtcgattt
2221 cgtctccaac gaccactact tcacgccggg cgaggaccat ttcgacgagg ttgcgtacgc
2281 cgcgtcgctg atggacggca tttcacgcaa ggaaccgtgg ttccagatgg agcattccac
2341 ctccgcggtg aattggcgtc cgatcaacta ccgtgcggag ccgggctcgg tcgtgcgcga
2401 ctccctcgcg caggtcgcta tgggcgccga tgccatctgc tatttccagt ggcgacagtc
2461 caaggccggc gcggagaagt ggcactcctc gatggtgccg catgcgggcg aggattcgca
2521 gatcttccgt gatgtgtgcg agctgggcgc cgatctggga cgtctgtccg acgagggtct
```

2581 gatgggcacc aagacggtca agtccaaggt cgccgtcgtg ttcgactacg agtcccagtg
2641 ggccaccgag tacaccgcga acccgaccca gcaggtcgat cattggaccg agccgttgga
2701 ttggttccgc gcgcttgccg acaacggcat caccgccgac gtggttccgg tgcggtccga
2761 ttgggattcc tacgagattg ccgtgctgcc gtgcgtgtat ttgctgtccg aagagacgag
2821 ccgcagggtt cgggagttcg tggcgaacgg cggcaagttg ttcgtgacgt actacaccgg
2881 attgagcgac gagaacgacc acatctggct tggcggctac ccgggctcga ttcgtgacgt
2941 cgtcggcgtg cgcgtcgaag agttcgcccc gatgggcaac gacatgccgg gcgcattgga
3001 tcacctcgac ctggacaatg gcacggtggc gcatgacttt gccgacgtga tcacgtccac
3061 ggccgacacg tccacggtac tggcctccta taaggcggaa cgctggaccg gcatgaacga
3121 ggtgccggcc atcgtggcca acgggtatgg cgacgggcgg accgtatatg tcggatgccg
3181 tctgggccgt caggggcttg cgaagagcct gccggcgatg ctgggttcca tggggctgtc
3241 ggacctcgcc ggtgatggcc gcgtgctgcg tgtcgaacgc gccgacgccg cggcggcgag
3301 ccgcttcgag ttcgtgttca accgcaccca tgaaccggtg accgtcgacg ttgaagggga
3361 ggccatcgcg gcttcgctcg cgcatgtcga cgacgggcgg gccaccatcg atccgacggg
3421 tgttgtcgtg ctcaggcgat aatcgttgga aacactgggc tgtaagggct taggaaaggc
3481 gtatgtttgc ggtgacacgc gacatacgcc ttatgggaaa gaaggcgctg gcgcttaccg
3541 ggctgcggcg atgctggtca gcgtcgctgc gtgcggcaac tcaagcagca gctccgctcc
3601 gaagcaggaa ggcgacgtca aggaaatcac cgtgtgggct tggggcctac gctgactcag
3661 gtggccaagg acttcaaaaa aggagaccgg catcaatgtc aacctggtga acaccggcca
3721 gggcgacaag acctgggacg agttctatca ggacgccaag aagattcaca ccttggcga
3781 caactactac atcacgtccg acaccggtgt cgccggcttc tacgactcga tgacctggct
3841 ggccagtgcg acgctgttct ccaccgaagg cgagacggtc accattaacc tgactggcgc
3901 cccgaaggtc aaggcccgcg gtatcttcgg cgactacctt ggcaagtcct acaccggcaa
3961 ccagaagctg agcgatggcg tcgccgcttt gggaacaggc tctgaaggac tacgcgaagg
4021 atcagggcta caccgtcaag taaccttcgc agtcaagcaa tctggcgtgg taatgaccgg
4081 aatacggtga ccttcggtca tcccttcctc gtgtgaaggc ccctcccctc aacagggagg
4141 ggccttcaca tatctgcccc tgttgcaacg cgcgtgtaaa ctctacgatg agcgaattct
4201 tcccgacaca tcgagcacgc taaggagatg acatgacgat atcggcacgg ttgtggcggc
4261 tgcacctgca tatctttgtg ttgctcaaga tctgtgagat gacactggca gcacgcctcc
4321 agcgcgccgc cggacacgcg caccctcatc accgaatgga cgggggacca tatcatgacg
4381 actctgatcg ccaat

Figure 2

MYMIWRDAHRVRCGLARYGTSPHRMSQLIFRRCDGSAGASDNEQVMERNMSK
RRKHSWPQPLKGAESRLWYGGDYNPDQWPEEVWDDDIRLMKKAGVNLVSV
GIFSWAKIEPEEGKYDFDWLDRAIDKLGKAGIAVDLASATASPPMWLTQAHP
EVLWKDERGDTVWPGAREHWRPTSPVFREYALNLCRRMAEHYKGNPYVVAWH
VSNEYGCHNRFDYSDDAMRAFQKWCKKRYKTIDAVNEAWGTAFWAQHMNDFS
EIIPPRYIGDGNFMNPGKLLDYKRFSSDALKELYIAERDVLESITPGLPLTT
NFMVSAGGSMLDYDDWGAEVDFVSNDHYFTPGEDHFDEVAYAASLMDGISRK
EPWFQMEHSTSAVNWRPINYRAEPGSVVRDSLAQVAMGADAICYFQWRQSKA
GAEKWHSSMVPHAGEDSQIFRDVCELGADLGRLSDEGLMGTKTVKSKVAVVF
DYESQWATEYTANPTQQVDHWTEPLDWFRALADNGITADVVPVRSDWDSYEI
AVLPCVYLLSEETSRRVREFVANGGKLFVTYYTGLSDENDHIWLGGYPGSIR
DVVGVRVEEFAPMGNDMPGALDHLDLDNGTVAHDFADVITSTADTSTVLASY
KAERWTGMNEVPAIVANGYGDGRTVYVGCRLGRQGLAKSLPAMLGSMGLSDL
AGDGRVLRVERADAAAASRFEFVFNRTHEPVTVDVEGEAIAASLAHVDDGRA
TIDPTGVVVLRR

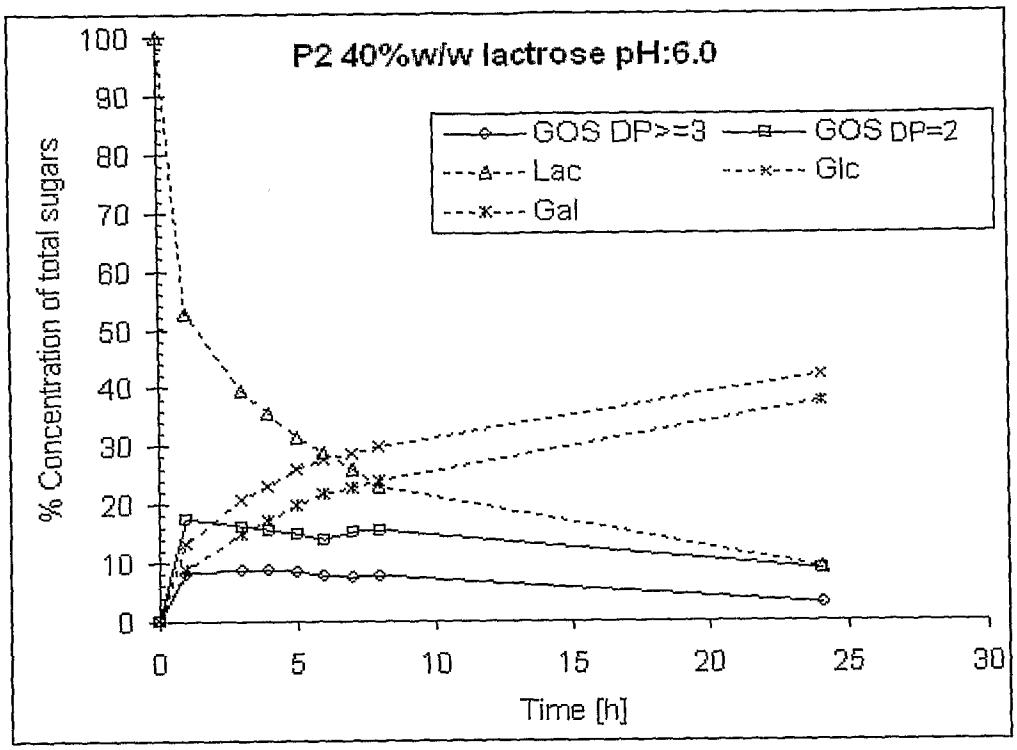

P2 40%w/w lactrose pH:6.0
GOS DP>=3
GOS DP=2
Lac
Glc
Gal
% Concentration of total sugars
Time [h]
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30

Figure 3

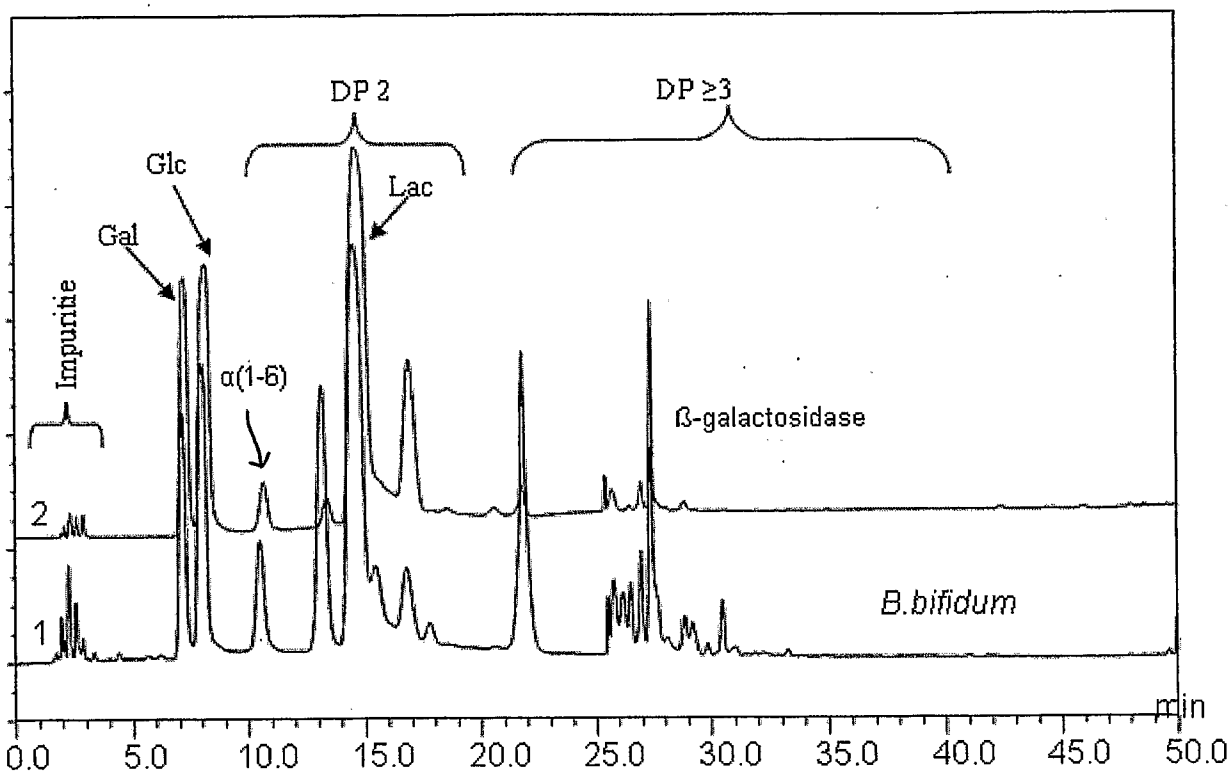

DP 2
DP ≥3
Glc
Lac
Gal
Impurities
α(1-6)
ß-galactosidase
2
1
B.bifidum
min
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
40.0
45.0
50.0

Figure 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2007/001081

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N5/10 C12N15/63 C12N9/38

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)
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)
EPO-Internal, WPI Data, EMBASE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2005/003329 A (CLASADO INC [PA]; WYNNE ANTHONY GRAHAM [GB]; GIBSON GLENN [GB]; SLUPIN) 13 January 2005 (2005-01-13) cited in the application page 4, paragraph 5 claim 5	1-18
X	DATABASE EMBL [Online] EBI; 29 November 1993 (1993-11-29), XP002435365 Database accession no. AAR38735 abstract -& DATABASE EMBL [Online] EBI; 29 November 1993 (1993-11-29), XP002435366 Database accession no. AAQ46333 abstract	1-6,8-18

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family 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.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 May 2007	04/07/2007

Name and mailing address of the ISA/	Authorized officer
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	Barnas, Christoph

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2007/001081

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	-& JP 05 146296 A (YAKULT HONSHA KK) 15 June 1993 (1993-06-15) abstract -----	1-6,8-18
X	DATABASE EMBL [Online] EBI; 23 August 2004 (2004-08-23), XP002435367 Database accession no. AY691690 abstract -----	1-6,8-18

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2007/001081

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2005003329 A	13-01-2005	AT 360682 T	15-05-2007
		AU 2004254110 A1	13-01-2005
		BR PI0408911 A	28-03-2006
		CA 2520043 A1	13-01-2005
		EP 1644482 A1	12-04-2006
		GB 2412380 A	28-09-2005
		HK 1077323 A1	28-07-2006
		KR 20050111624 A	25-11-2005
		MX PA05010349 A	08-03-2006
JP 5146296 A	15-06-1993	NONE	