

BETA-HEXOSYL-TRANSFERASES AND USES THEREOF

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of US Provisional Appn. No. 61/734,742 filed December 7, 2012, Bruno-Barcena *et al.*, entitled “Beta-hexosyl-transferases and Uses Thereof” having Atty. Dkt. No. NS12009USV which is hereby incorporated by reference in its entirety.

1. FIELD OF THE INVENTION

[0002] This invention relates generally to the discovery of novel recombinant forms of β -hexosyl-transferases (BHT) and uses thereof to produce galacto-oligosaccharides (GOS).

2. BACKGROUND OF THE INVENTION

2.1. Introduction

[0003] The complex interaction between diet, normal intestinal microbiota, and wellbeing has encouraged the development of strategies to promote the selective proliferation of beneficial microorganisms into the gastrointestinal track of humans. Probiotics are microorganisms that positively affect human health with attributed *powerful antipathogenic and anti-inflammatory properties* (27) (Table1).

[0004] **TABLE 1.** Health Benefits of Probiotics

Intestinal	Immunity	Reduce disease risk
Helicobacter pylori infection Lactose intolerance Irritable bowel syndrome Ulcerative colitis Crohn's disease Diarrhea Constipation Stimulate mineral adsorption	Reducing allergic reactions Reducing opportunity of infection by pathogens	Coronary heart disease High blood pressure Upper respiratory tract infections Urinary tract disease Reduced cholesterol and lipids Aid in prevention of colon cancer

[0005] Also, years of probiotic research indicate that a selective modification of the intestinal microbiota and its associated biochemical activities can be promoted by selective prebiotics. Osborn DA, Sinn JK. Prebiotics in infants for prevention of allergic disease and food hypersensitivity. Cochrane Database of Systematic Reviews 2007. Prebiotics are non-digestible oligosaccharides (NDOs) that have a dual ability. First they reduce the intestinal

colonizing efficiency of harmful bacteria and second they act as selective substrate to promote the growth and thereby increasing the number of specific probiotic bacteria.

[0006] In addition, an increasing number of studies have shown that probiotics work best when combined with prebiotics. Mayer *et al.* 2003 Research for creation of functional foods with Bifidobacteria. *Acta Alimentaria* 32 27-39. This combined form of delivery is known as a synbiotic. Gibson GR, Roberfroid MB. 1995 Dietary Modulation of the Human Colonic Microbiota - Introducing the Concept of Prebiotics. *Journal of Nutrition* 125:1401-12.

[0007] Galacto-oligosaccharides (GOS) are considered one of the preferred choices of prebiotics and in the gastrointestinal tract, GOS are resistant to enzymes and transit through the small intestine without being digested, but in the large intestine GOS are fermented and can activate growth of intestinal bifidobacteria such as *Lactobacillus acidophilus* and *L. casei*, hence acting as a prebiotic(26, 27, 37).

[0008] GOS are non-digestible oligosaccharides owing to the conformation of their anomeric C atom (C₁ or C₂), which allows their glycosidic bonds to evade hydrolysis by digestive enzymes in the stomach or small intestine. Free oligosaccharides are found in the milk of all placental mammals, providing a natural example of prebiotic feeding during infancy. According to the latest definition by the International Scientific Association for Probiotics and Prebiotics (ISAPP) “a dietary prebiotic is a selectively fermented ingredient that results in specific changes in the composition and/or activity of the gastrointestinal microbiota, thus conferring benefit(s) upon host health” (30). The composition of human milk oligosaccharides (HMO) is very complex, which makes it unlikely to find alternative sources containing oligosaccharides of analogous composition. Improved colonic health among breastfed infants has been attributed to the presence of GOS in the mother’s milk (2). In fact, infant formula with added GOS replicated the bifidogenic effect of the human milk with respect to metabolic activity of colonic microbiota and bacterial numbers (6,21). Among non-milk oligosaccharides, GOS are of special interest as their structure resembles the core molecules of HMOs (3). However, GOS concentration and composition vary with the method and the enzyme utilized for their generation, which in turn may influence their prebiotic effects and the proliferation of colonic probiotic strains (29). Traditionally, GOS have been produced using β -galactosidases from mesophilic microorganisms. Mesophilic β -galactosidases require high initial concentrations of lactose to drive the reaction away from lactose hydrolysis and towards GOS synthesis. Since lactose is more soluble at elevated temperatures, thermostable β -galactosidases exhibiting high initial velocities and increased

half-lives have been utilized to reach a favorable equilibrium for the transgalactosylation reaction (27,37). However, competitive inhibition by glucose and/or galactose is another obstacle that remains which may be overcome by incorporating cells in the reaction (16,20,25,27,35).

[0009] The basidiomycete yeast *Sporobolomyces singularis* (formerly *Bullera singularis*) cannot utilize galactose to grow but proliferates on lactose due to the activity of its β -hexosyl-transferase (BHT, EC 3.2.1.21). Studies have shown that the BHT has transgalactosylation activity even at low lactose concentrations and very limited lactose hydrolysis. In addition, the enzyme does not appear to be inhibited by lactose concentrations above 20% and has the potential for conversions into GOS close to the maximum theoretical of 75% (1,9,10,28). Unlike β -galactosidases, the BHT from *S. singularis* simultaneously carries out glycosyl-hydrolase and β -hexosyl-transferase activities, converting lactose to GOS without extracellular accumulation of galactose. Two molecules of lactose are required during the transgalactosylation event: one molecule is hydrolyzed and the second acts as galactose acceptor, generating the trisaccharide galactosyl-lactose (β -D-Gal(1-4)- β -D-Gal(1-4)- β -D-Glc) and residual glucose. Galactosyl-lactose can also act as acceptor of a new galactose to generate the tetrasaccharide galactosylgalactosyl-lactose (β -D-Gal(1-4)- β -D-Gal(1-4)- β -D-Gal(1-4)- β -D-Glc), and similarly for the tetrasaccharide and subsequent products. The tri, tetra, and penta saccharides accumulating in *S. singularis* have been collectively designated GOS (9,10).

[0010] For practical interests, a recombinant secreted BHT could have several advantages over the native enzyme, including improved large scale production and purification. Currently, purification of active enzyme from *S. singularis* requires cell lysis followed by multiple chromatography steps (1,4,16). Previous attempts to express recombinant β -hexosyl-transferase in *E. coli* BL21 have resulted in high levels of production, but the enzyme was inactive and insoluble (16).

3. SUMMARY

[0011] The present disclosure provides an isolated DNA coding for a recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20. Disclosed is an isolated DNA coding for a recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 4, 6, 8, 10, 12, 14, or 20. The isolated DNA coding for the recombinant β -hexosyl-transferase (rBHT) protein may have the nucleic acid sequence set forth in SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17 or 19. In one aspect, the present invention provides an isolated DNA coding for a recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 12, 14, or 20, preferably having the amino acid sequence set forth in SEQ ID NO. 12 or 14

[0012] The disclosure also provides, an enzymatically active recombinant β -hexosyl-transferase (rBHT) protein wherein the protein has the amino acid sequence set forth in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20. Also disclosed is an enzymatically active recombinant β -hexosyl-transferase (rBHT) protein wherein the protein has the amino acid sequence set forth in SEQ ID NO. 4, 6, 8, 10, 12, 14, or 20. In another aspect, the present invention provides an enzymatically active recombinant β -hexosyl-transferase (rBHT) protein wherein the protein has the amino acid sequence set forth in SEQ ID NO. 12, 14, or 20. The rBHT protein may be membrane bound or may be a soluble enzyme.

[0013] The enzymatically active rBHT producing GOS may, or may not be, inhibited by galactose.

[0014] The disclosure also provides a method for producing enzymatically active recombinant β -hexosyl-transferase (rBHT) protein in a eukaryotic host cell which comprises transforming the eukaryotic host cell with a plasmid under the control of a suitable promotor wherein the plasmid contains an isolated DNA coding for an rBHT protein having the amino acid sequence set forth in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20. In yet another aspect, there is provided a method for producing enzymatically active recombinant β -hexosyl-transferase (rBHT) protein in a yeast host cell which comprises transforming the eukaryotic host cell with a plasmid under the control of a suitable promotor wherein the plasmid contains an isolated DNA coding for a rBHT protein as described herein.

[0015] In some embodiments, the isolated DNA is linked to a DNA coding for a signal peptide. The signal peptide may be an *S. cerevisiae* α -factor signal peptide and the suitable promotor may be an alcohol oxidase promotor.

[0016] In some embodiments, the enzymatically active rBHT protein has a specific activity of about 8 U.mg⁻¹ at 20°C.

[0017] The eukaryotic host cell may be a yeast cell such as *Pichia pastoris*.

[0018] The disclosure also provides a method for producing galacto-oligosaccharides (GOS) which comprises reacting lactose with an enzymatically active recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 under suitable conditions so as to produce GOS. Also disclosed is a method for producing galacto-oligosaccharides (GOS) which comprises reacting lactose with an enzymatically active recombinant β -hexosyl-transferase (rBHT) protein produced under suitable conditions so as to produce GOS.

[0019] The enzymatically active rBHT protein may be immobilized on a solid support. The solid support may be in a batch or continuous stirred-tank reactor, a packed-bed reactor, or an ultrafiltration membrane reactor. Alternatively, the enzymatically active rBHT protein may be used directly in a batch or continuous stirred-tank reactor, a packed-bed reactor, or an ultrafiltration membrane reactor. The method may further comprise a glucose removal system to avoid competitive glucose inhibition such as a generally recognized as safe (GRAS) organism. The glucose removal system may be used simultaneously with the enzymatically active rBHT protein.

[0020] The disclosure is also directed to a modified lactose-containing foodstuff or food byproduct comprising a recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20. Also disclosed is a modified lactose-containing foodstuff or food byproduct comprising a recombinant β -hexosyl-transferase (rBHT) protein produced as described herein. The lactose-containing foodstuff or food byproduct may be a dairy product or byproduct such as whey. In

some embodiments, the foodstuff or food byproduct is a baby dessert, a baby juice, a baby snack, a baby yoghurt drink, an energy drink, a fitness water or thirst quencher, a frozen dairy dessert, a fruit drink (vitamin/mineral fortified), a fruit pie filling, a fruit preparation, an infant formula, an infant meal replacement drink, a jelly jam, a meal replacement drink, a milk, a milk-based drink, a milk substitute, a syrup flavoring for milk, a yoghurt, or a whey.

4. BRIEF DESCRIPTION OF THE FIGURES

[0021] FIGURE 1A-1C. Gel electrophoresis of purified rBHT expressed in *P. pastoris*. (1A) SDS-PAGE of purified rBHT silver stained: Lane 1, 0.5 μg rBHT; lane 2, 0.1 μg rBHT. (1B) Western blot analysis with anti-rBHT antiserum. Lane M indicates the molecular masses (kDa) of the marker proteins are shown to the right of Panel A. (1C) Zymogram of rBHT. Lane 1, purified rBHT-6XHIS expressed in *E. coli* BLR cultures; Lane 2, broth from untransformed methanol induced GS115; Lane 3, broth from methanol induced recombinant GS115/*rBHT*. Activity was visualized by the formation of a blue precipitate resulting from enzymatic cleavage of X-GAL.

[0022] FIGURE 2A-2D. rBHT relative activity dependence on (2A) pH. (2B) Temperatures from 20°C to 80°C. (2C) Concentration of galactose (solid circle) and glucose (solid square). (2D) Thermal stability at 20°C to 50°C. Samples were removed every 12 h and assayed for activity at 20°C. Enzyme activity assays were conducted in 50 mM sodium phosphate (pH 5.0) containing 1.3 mM ONP-Glu at 20°C, except for (A) which used sodium phosphate (pH 5-11) or citrate (pH 2-5) buffers. Enzyme activities were calculated relative to the value taken at time zero (100%). The initial concentration of tested enzyme was 0.2 $\text{U}\cdot\text{mL}^{-1}$ assayed at 20°C ($K_m=0.37$ mM and $V_{\max}=0.09$ $\text{mM}\cdot\text{min}^{-1}$). Data represents the means of two experiments with a reliability of $\pm 5\%$.

[0023] FIGURE 3. Synthesis of galactosyl-lactose from lactose (2%) using partially purified rBHT (0.5 $\text{U}\cdot\text{g}^{-1}$ lactose) in 5 mM sodium phosphate buffer pH 5.0 incubated at 42°C. Concentrations of lactose, glucose, galactose and galactosyl-lactose are shown in $\text{g}\cdot\text{L}^{-1}$. The residual non-quantified GOS species are shown as signal intensity readings from the refractive-index detector. Data represents the means of two experiments with a reliability of $\pm 5\%$.

[0024] FIGURE 4A-4B. Synthesis of galacto-oligosaccharides from lactose (20%) using *P. pastoris* resting cells (harboring membrane-bound rBHT at 0.5 $\text{U}\cdot\text{g}^{-1}$ lactose) in 5 mM sodium phosphate buffer pH 5.0 incubated at 42°C (4A) or 30°C (4B). Concentrations of

lactose, glucose, galactose and galactosyl-lactose are shown in g.l⁻¹. The residual non-quantified GOS species are shown as signal intensity readings from the refractive-index detector. Data represents the means of two experiments with a reliability of $\pm 5\%$.

[0025] FIGURE 5. Shows a comparison of the enzymatic activities and the sugars including GOS produced by the rBHT described herein (GOS NCSU) and two commercially available enzymes Vivinal GOS and Oligomate 55NP from Yakalt Pharmaceuticals, Inc. at 30°C. See the procedure in Section 6.3.

[0026] FIGURE 6A-6B. Kyte and Doolittle hydropathy plot for BHT. The amino acid sequence for the predicted membrane anchor/signal sequence has been amplified. The transmembrane region prediction algorithm (www.ch.embnet.org/software/TMPRED_form.html) also forecasted a stretch of hydrophobic residues 1-17 in BHT typical for integral membrane spanning proteins and the SignalP algorithm (www.cbs.dtu.dk/services/SignalP/) predicted a possible signal sequence for the same residues and possible cleavage site between residues 17 and 18 and between 22 and 23. The signal sequence was retained in constructions 1, 2, 3, 4 and 7 to serve as a natural membrane anchor.

[0027] FIGURE 7A-7C. Sequence analysis of BHT. (7A) Schematic representation of the BHT polypeptide from *S. singularis* determined by the web-based SMART program (<http://smart.embl-heidelberg.de>). The leader peptide represented by a solid square was determined by the *SignalP* program. Segment of low compositional complexity represented by a solid circle was determined by the *SEG* program (<http://mendel.imp.ac.at/METHODS/seg.server.html>). Hits only found by BLAST are indicated by the Glyco hydrolase domain. Solid triangles indicate the positions of the three putative N-glycosylation sites determined by NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc/>). (7B) Schematic representation of the leader peptide calculated by the RHYTHM transmembrane prediction method. The amino acid sequence for the predicted membrane anchor sequence has been amplified. Membrane contact amino acids are in large bold type and the helix contact amino acid is in large type and underlined. Also indicated are the positions of predicted cytoplasmic and extracellular regions. Arrows indicate possible cleavage sites. (7C) Hydropathy plot. The plot was generated using Kyte-Doolittle method of calculating hydrophilicity over a window length of nine amino acids. The number of amino acids is shown below the X-axis. Zero on the Y-axis separates hydrophobic and hydrophilic amino acids.

[0028] FIGURE 8A-8D. Concentration of secreted protein by each recombinant *P. pastoris* strain. Graphic representations of recombinant strains containing *rBht* and *scFv13R4* are shown to the left (8A) and right (8D) of the plots, respectively. (8B) rBHT-HIS secreted. (8C) scFv13R4-HIS secreted. The values of secreted protein were normalized for OD₆₀₀ and represented the mean $\pm$ SE. Secreted proteins were analyzed from the following recombinant strains: (8A and 8B) row 1, GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS; row 2, GS115::*rBht*-HIS; row 3, GS115::*rBht*₍₂₃₋₅₉₄₎-HIS; row 4, GS115::*αMF-rBht*-HIS. (8C and 8D) row 1, GS115::*αMF-scFv13R4*-HIS; row 2, GS115::*rBht*₍₁₋₁₁₀₎-*scFv13R4*-HIS; row 3, GS115::*rBht*₍₂₃₋₁₁₀₎-*scFv13R4*-HIS; row 4, GS115::*scFv13R4*-HIS.

[0029] FIGURE 9A-9C. SDS-PAGE (8%) separation and Western blots revealed anti-His antiserum showing secreted, cell associated, or purified rBHT-HIS expressed by different *P. pastoris* GS115 recombinants. (9A) Protein cell free extracts (secreted proteins) generated by all recombinants were concentrated 20 fold. (9B) Cell associated proteins were obtained from five OD₆₀₀ of recombinant cells disrupted with glass beads in 1X Laemmli buffer. Lane 1, GS115::*αMF-rBht*-HIS; lane 2, GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS; lane 3, GS115::*rBht*-HIS; lane 4, GS115::*rBht*₍₂₃₋₅₉₄₎-HIS; and lane 5, GS115 control. (9C) Silver stain of rBHT-HIS expressed by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS purified using nickel affinity chromatography and resolved in SDS-PAGE(8%). M indicates marker lane. The molecular masses (kDa) of the marker proteins are shown to the left of the panels.

[0030] FIGURE 10A-10B. Time course of galactosyl-lactose synthesis using soluble rBHT or *P. pastoris* resting cells containing membrane bound rBHT (0.5 U rBHT.g⁻¹ lactose). (10A) Synthesis by secreted rBHT-HIS expressed by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS (solid line) and rBHT by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎ (dashed line). (10B) The rate of galactosyl-lactose synthesis by resting cells GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS (solid square) and GS115::*αMF-rBht*₍₂₃₋₅₉₄₎ (open square). Assays contained 200 g.L⁻¹ lactose and purified enzyme or resting cells of *P. pastoris* in 5 mM sodium phosphate buffer pH 5.0 and incubated at 30°C. Samples were removed periodically and analyzed by HPLC. Concentrations of lactose, glucose, galactose and galactosyl-lactose are shown in g.L⁻¹. The residual non-quantified GOS species are shown as signal intensity readings from the refractive-index detector. Data represents the means of two independent experiments.

5. DETAILED DESCRIPTION OF THE INVENTION

[0031] This invention reports several methods for the expression of the *S. singularis* BHT including, but not limited to, a method using a codon optimized, synthetic *rBht* gene (GenBank accession number JF29828) expressed in *Pichia pastoris*. We investigated the kinetics of GOS production from lactose by the secreted recombinant β -hexosyl-transferase (rBHT) as compared to *P. pastoris* resting cells harboring membrane-bound rBHT.

[0032] "rBHT proteins," as meant herein, includes full length rBHT proteins and fragments and/or variants thereof, which includes proteins encoded by naturally-occurring allelic variants of the rBHT gene, as well as recombinantly-produced rBHT proteins, which may contain some sequence changes relative to naturally-occurring rBHT proteins.

[0033] A "recombinant" protein is one resulting from the process of genetic engineering. The term "genetic engineering" refers to a recombinant DNA or nucleic acid method used to create a cell that expresses a gene at elevated levels or at lowered levels, or expresses a mutant form of the gene. In other words, the cell has been transfected, transformed or transduced with a recombinant polynucleotide molecule, and thereby altered so as to cause the cell to alter expression of a desired polypeptide.

[0034] "galacto-oligosaccharide" or "GOS" means a galactose-containing polysaccharide with two or more sugar units such as Gal-Gal or $[\text{Gal}]_n\text{-Glc}$ ($1 \leq n \leq 8$), including $\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Glc}$, $\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Glc}$, and $\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Gal}(1 \rightarrow 4)\text{-}\beta\text{-D-Glc}$.

5.1. Signal Sequences

[0035] Soluble secreted proteins and proteins expressed on the cell surface often comprise an N-terminal "signal sequence," which is a hydrophobic sequence that mediates insertion of the protein through the membrane bounding the endoplasmic reticulum (ER) in a eukaryotic cell. Type 1 transmembrane proteins also comprise signal sequences. "Signal sequences," as meant herein are amino-terminal hydrophobic sequences which are usually enzymatically removed following the insertion of part or all of the protein through the ER membrane into the lumen of the ER. Thus, it is known in the art that a signal sequence can be present as part of a precursor form of a secreted or transmembrane protein, but will generally be absent from the mature form of the protein. When a protein is said to comprise a signal sequence, it is to be understood that, although a precursor form of the protein does contain the signal sequence, a mature form of the protein will likely not contain the signal sequence. Signal sequences may contain a residue adjacent to and immediately upstream from the

cleavage site (position -1) and another residue at position -3, which are important for this enzymatic cleavage. Nielsen *et al.* 1997 *Protein Eng* 10(1) 1-6; von Heijne 1983 *Eur J Biochem* 133(1) 7-21; von Heijne 1985 *J Mol Biol* 184 99-105, the portions of which describe signal sequences and how to identify them are incorporated herein by reference. Signal sequences can be identified as described by Nielsen *et al.* (supra). Examples of signal peptides or sequences that are functional in mammalian host cells include the following: the *Saccharomyces cerevisiae* pre-pro-alpha-mating factor signal sequence the signal sequence for interleukin-7 (IL-7) described in US Patent 4,965,195; the signal sequence for interleukin-2 receptor described in Cosman *et al.* 1984 *Nature* 312 768-771); the interleukin-4 receptor signal peptide described in EP Patent 0 367 566; the type 1 interleukin-1 receptor signal sequence described in US Patent 4,968,607; the type II interleukin-1 receptor signal peptide described in EP Patent 0 460 846. Many other signal sequences are known in the art.

5.2. rBHT Protein

[0036] The instant invention encompasses secreted, soluble versions of rBHT, as well as versions comprising a transmembrane domain that can be expressed on a cell surface. Such proteins can be isolated, that is, be part of a purified protein preparation in which the rBHT protein constitutes at least 80% or at least 90% of the protein present in the preparation. The invention further includes rBHT proteins encoded by the rBHT nucleic acids described below. An rBHT protein, as meant herein, encompasses a protein comprising the amino acid sequence of SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20, as well as fragments, derivatives, and variants thereof, including fusion proteins. The amino acid sequence of SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 includes a signal sequence.

5.3. Conservative Substitutions

[0037] In some embodiments the substitutions can be conservative amino acid substitutions. Examples of conservative amino acid substitutions, unlikely to affect biological activity, include the following: alanine for serine, valine for isoleucine, aspartate for glutamate, threonine for serine, alanine for glycine, alanine for threonine, serine for asparagine, alanine for valine, serine for glycine, tyrosine for phenylalanine, alanine for proline, lysine for arginine, aspartate for asparagine, leucine for isoleucine, leucine for valine, alanine for glutamate, aspartate for glycine, and these changes in the reverse. See e.g. Neurath *et al.*, The Proteins, Academic Press, New York (1979), the relevant portions of

which are incorporated herein by reference. Further, an exchange of one amino acid within a group for another amino acid within the same group is a conservative substitution, where the groups are the following: (1) alanine, valine, leucine, isoleucine, methionine, norleucine, and phenylalanine; (2) histidine, arginine, lysine, glutamine, and asparagine; (3) aspartate and glutamate; (4) serine, threonine, alanine, tyrosine, phenylalanine, tryptophan, and cysteine; and (5) glycine, proline, and alanine.

5.4. Glycosylation

[0038] rBHT proteins may be glycosylated to varying degrees or not glycosylated. As an illustration, an rBHT protein of the invention may comprise one or more N- or O-linked glycosylation sites in addition to those already found in a protein comprising SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20. One of skill in the art would be aware that asparagine residues that are part of the sequence Asn Xxx Ser/Thr (where Xxx is any amino acid except proline) can serve as sites of addition for N-glycans. In addition, there are many serine and threonine residues that may serve as O-linked glycosylation sites. Glycosylation may increase *in vivo* half-life or alter biological activity. Variants of rBHT proteins also include proteins comprising one, two, three, four, five, six, seven, eight, nine, or ten more N- and/or O-linked glycosylation sites than are present in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 as long as the resulting protein can act as a glycosyl hydrolase and a β -hexosyl-transferase. Variant rBHT proteins also include those that have one, two, three, four, or five fewer N- and/or O-linked glycosylation sites than are present in SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 as long as they can act as a glycosyl hydrolase and a β -hexosyl-transferase.

[0039] rBHT proteins, as meant herein, can be fusion proteins comprising at least one rBHT polypeptide, which can comprise an amino acid sequence that is a variant and/or a fragment of SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 (as explained above), and at least one other moiety. The other moiety can also be a non-protein moiety such as, for example, a polyethylene glycol (PEG) moiety or a cytotoxic, cytostatic, luminescent, and/or radioactive moiety. Attachment of PEG has been shown to increase the *in vivo* half-life of at least some proteins. Moreover, cytotoxic, cytostatic, luminescent, and/or radioactive moieties have been fused to antibodies for diagnostic or therapeutic purposes.

[0040] A variety of polypeptides other than rBHT can be fused to an rBHT polypeptide for a variety of purposes such as, for example, to increase *in vivo* half-life of the protein, to facilitate identification, isolation and/or purification of the protein, to increase the activity of the protein, and to promote oligomerization of the protein.

[0041] Many polypeptides can facilitate identification and/or purification of a recombinant fusion protein of which they are a part. Examples include polyarginine, polyhistidine, or HAT™ (Clontech), which is a naturally-occurring sequence of non-adjacent histidine residues that possess a high affinity for immobilized metal ions. rBHT proteins comprising these polypeptides can be purified by, for example, affinity chromatography using immobilized nickel or TALON™ resin (Clontech), which comprises immobilized cobalt ions. See e.g. Knol *et al.* 1996 *J Biol Chem* 27(26) 15358-15366. Polypeptides comprising polyarginine allow effective purification by ion exchange chromatography. Other useful polypeptides include, for example, the antigenic identification peptides described in U.S. Patent 5,011,912 and in Hopp *et al.* 1988 *Bio/Technology* 6 1204. One such peptide is the FLAG™ peptide, which is highly antigenic and provides an epitope reversibly bound by a specific monoclonal antibody, enabling rapid assay and facile purification of expressed recombinant fusion protein. A murine hybridoma designated 4E11 produces a monoclonal antibody that binds the FLAG peptide in the presence of certain divalent metal cations, as described in U.S. Patent 5,011,912. The 4E11 hybridoma cell line has been deposited with the American Type Culture Collection under Accession No. HB 9259. Monoclonal antibodies that bind the FLAG peptide can be used as affinity reagents to recover a polypeptide purification reagent that comprises the FLAG peptide. Other suitable protein tags and affinity reagents are: 1) those described in GST-Bind™ system (Novagen), which utilizes the affinity of glutathione-S-transferase fusion proteins for immobilized glutathione; 2) those described in the T7-TAG® affinity purification kit, which utilizes the affinity of the amino terminal 11 amino acids of the T7 gene 10 protein for a monoclonal antibody; or 3) those described in the STREP-TAG® system (Novagen), which utilizes the affinity of an engineered form of streptavidin for a protein tag. Some of the above-mentioned protein tags, as well as others, are described in Sassenfeld 1990 *TIBTECH* 8: 88-93, Brewer *et al.*, in Purification and Analysis of Recombinant Proteins, pp.239-266, Seetharam and Sharma (eds.), Marcel Dekker, Inc. (1991), and Brewer and Sassenfeld, in Protein Purification Applications, pp. 91 - 111, Harris and Angal (eds.), Press, Inc., Oxford England (1990). The portions of these references that describe protein tags are incorporated herein by reference. Further, fusions of two or more of the tags described above, such as, for example, a fusion of a FLAG tag and a polyhistidine tag, can be fused to an rBHT protein of the invention.

5.5. rBHT Nucleic Acids

[0042] The invention encompasses isolated nucleic acids, including, for example DNAs and RNAs, that encode the rBHT proteins described herein, which include proteins comprising the amino acid sequence of SEQ ID NO. 2, 4, 6, 8, 10, 12, 14, 16, 18 or 20 and fragments and/or variants thereof. Preferably, the proteins have the amino acid sequence of SEQ ID NO. 12 or 14. These nucleic acids are useful for, *inter alia*, producing recombinant proteins having glycosyl hydrolase and a β -hexosyl-transferase activity. Such nucleic acids can be modified genomic DNA or cDNA. Preferably, the nucleic acids can comprise an uninterrupted open reading frame encoding an rBHT protein. Nucleic acid molecules of the invention include DNA and RNA in both single-stranded and double-stranded form, as well as the corresponding complementary sequences. An "isolated nucleic acid" is a nucleic acid that has been separated from adjacent genetic sequences present in the genome of the organism from which the nucleic acid was isolated, in the case of nucleic acids isolated from naturally-occurring sources, in the case of nucleic acids synthesized chemically, such as oligonucleotides, or enzymatically from a template, such as polymerase chain reaction (PCR) products or cDNAs, it is understood that the nucleic acids resulting from such processes are isolated nucleic acids. An isolated nucleic acid molecule refers to a nucleic acid molecule in the form of a separate fragment or as a component of a larger nucleic acid construct.

[0043] The present invention also includes nucleic acids comprising the sequence of SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, or a fragment thereof or nucleic acids that hybridize under moderately stringent conditions, and optionally highly stringent conditions, to nucleic acids comprising the nucleotide sequence of SEQ ID NO 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, which is the nucleotide sequence of the full length rBHT cDNA, wherein the nucleic acid encodes a protein that can act as a glycosyl hydrolase and a β -hexosyl-transferase. Hybridization techniques are well known in the art and are described by Sambrook, J., E. F. Fritsch, and T. Maniatis (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., chapters 9 and 11, 1989) and Current Protocols in Molecular Biology (F. M. Ausubel *et al.*, eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4 1995), the relevant portions of which are incorporated by reference herein. Moderately stringent conditions for filter hybridizations include hybridization in about 50% formamide, 6 x SSC at a temperature from about 42°C to 55°C and washing at about 60°C in 0.5 x SSC, 0.1 % SDS. Highly stringent conditions are defined as hybridization conditions as above, but with washing at approximately 68°C in 0.2 x SSC, 0.1 % SDS. SSPE (1 xSSPE is 0.15 M NaCl, 10 mM NaH₂P0₄, and 1.26 mM EDTA, pH 7.4) can be substituted for SSC (1

xSSC is 0.15 M NaCl and 15 mM sodium citrate) in the hybridization and wash buffers; washes, optionally at least two washes, are performed for 15 minutes after hybridization is complete.

[0044] It should be understood that the wash temperature and wash salt concentration can be adjusted as necessary to achieve a desired degree of stringency by applying the basic principles that govern hybridization reactions and duplex stability, as known to those skilled in the art and described further below (see e.g., Sambrook *et al.*, *supra*). When nucleic acids of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the nucleic acids (for example, using GAP) and identifying the region or regions of optimal sequence complementarity. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5°C to 10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m (degrees C) = $2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$. For hybrids above 18 base pairs in length, T_m (degrees C) = $81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41 (\% \text{ G} + \text{C}) - (600/N)$, where N is the number of bases in the hybrid, and $[\text{Na}^+]$ is the concentration of sodium ions in the hybridization buffer. Each such hybridizing nucleic acid has a length that is at least 15 nucleotides (or at least 18 nucleotides, or at least 20, or at least 25, or at least 30, or at least 40, or at least 50, or at least 100. Sambrook *et al.*, *supra*).

[0045] rBHT nucleic acids include nucleic acids comprising the following polynucleotides: (1) all or a fragment of SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, wherein the fragment encodes an rBHT protein that can act as a glycosyl hydrolase and a β -hexosyl-transferase; (2) a polynucleotide including nucleotide sequences at least 80%, 85%, 90%, 95%, 97%, 98%, 99%, 99.5%, or 99.7% identical to SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, wherein the alignment window is at least 100, 125, 150, 175, 200, 225, 250, 300, 400, 500, 600, 800, 1000. or 1200 nucleotides long and wherein the sequence encodes an rBHT protein that can act as a glycosyl hydrolase and a β -hexosyl-transferase; (3) a polynucleotide that comprises not more than 1, 2, 3, 4, 6, 8, 10, 15, 20, 25, or 30 alteration(s) of a single nucleotide relative to SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, wherein an alteration can be an insertion, deletion or substitution of a single nucleotide, and wherein the polynucleotide encodes an rBHT protein can act as a glycosyl hydrolase and a β -hexosyl-transferase; and (4) a polynucleotide that encodes an rBHT protein as described herein, which includes fragments, derivatives and variants of a rBHT protein. In a preferred embodiment, the rBHT

protein is produced by replacing the leader sequence with a heterologous secretion signal peptide.

5.6. Methods of Making rBHT Proteins

[0046] rBHT proteins can be made as follows. A nucleic acid that encodes an rBHT protein, as described herein, can be introduced into a vector, which can be introduced into a host cell. Vectors and host cells comprising nucleic acids encoding an rBHT protein are encompassed by the invention. The host cell containing the nucleic acids encoding an rBHT protein can be cultured under conditions such that the rBHT protein can be expressed. The expressed rBHT protein can then be obtained from the medium in which the cells are cultured or from the cells and purified by any of the many appropriate means known in the art. In addition, genetic engineering methods for the production of rBHT proteins include the expression of the polynucleotide molecules in cell free expression systems, in cellular hosts, in tissues, and in animal models, according to known methods.

[0047] The vector can include a selectable marker and an origin of replication, for propagation in a host. The vector can further include suitable transcriptional or translational regulatory sequences, such as those derived from mammalian, microbial, viral, or insect genes, operably linked to the nucleic acid encoding the rBHT protein. Examples of such regulatory sequences include transcriptional promoters, operators, or enhancers, mRNA ribosomal binding sites, and appropriate sequences that control transcription and translation. Nucleotide sequences are operably linked when the regulatory sequence functionally relates to the DNA encoding the target protein. Thus, a promoter nucleotide sequence is operably linked to an rBHT nucleic sequence if the promoter nucleotide sequence directs the transcription of the rBHT protein-encoding sequence. If the rBHT protein is a fusion protein, a nucleic acid sequence encoding a portion of the fusion protein, for example, a signal sequence, can be part of a vector, and a nucleic acid encoding an rBHT protein can be inserted into the vector such that a protein comprising the added signal sequence plus the rBHT protein is encoded by the vector.

[0048] Suitable host cells for expression of rBHT proteins include prokaryotic cells, yeast cells, plant cells, insect cells, and higher eukaryotic cells. The regulatory sequences in the vector will be chosen such that they are operable in the host cell. Suitable prokaryotic host cells include bacteria of the genera *Escherichia*, *Bacillus*, and *Salmonella*, as well as members of the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. For expression in

prokaryotic cells, for example, in *E. coli* the polynucleotide molecule encoding an rBHT protein preferably includes an N-terminal methionine residue to facilitate expression of the recombinant polypeptide. The N-terminal methionine may optionally be cleaved from the expressed polypeptide. Suitable yeast host cells include cells from genera including *Saccharomyces*, *Pichia*, and *Kluyveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. A suitable system for expression in an insect host cell is described, for example, in the review by Luckow and Summers (1988 *BioTechnology* 6 47-55), the relevant portions of which are incorporated herein by reference. Suitable mammalian host cells include the COS-7 line of monkey kidney cells (Gluzman *et al.* 1981 *Cell* 23 175-182), baby hamster kidney (BHK) cells, Chinese hamster ovary (CHO) cells (Puck *et al.* 1958 *PNAS USA* 60 1275-1281), CV-1 (Fischer *et al.* 1970 *Int J Cancer* 5 21-27), 293 cells from human kidney (American Type Culture Collection (ATCC®) catalog no. CRL-10852™), and human cervical carcinoma cells (HELA) (ATCC® CCL 2). The relevant portions of the references referred to in this paragraph are incorporated herein by reference.

[0049] Expression vectors for use in cellular hosts generally comprise one or more phenotypic selectable marker genes. Such genes encode, for example, a protein that confers antibiotic resistance or that supplies an auxotrophic requirement. A wide variety of such vectors are readily available from commercial sources. Examples include pGEM vectors (Promega), pSPORT vectors, and pPROEX vectors (Invitrogen, Life Technologies, Carlsbad, CA), Bluescript vectors (Stratagene), and pQE vectors (Qiagen). Yeast vectors will often contain an origin of replication sequence from a yeast plasmid, an autonomously replicating sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Vectors replicable in both yeast and *E. coli* (termed shuttle vectors) may also be used. In addition to the above-mentioned features of yeast vectors, a shuttle vector will also include sequences for replication and selection in *E. coli*. Direct secretion of the target polypeptides expressed in yeast hosts may be accomplished by the inclusion of nucleotide sequence encoding the yeast α -factor leader sequence at the 5' end of the rBHT-encoding nucleotide sequence. Brake 1989 *Biotechnology* 13 269-280.

[0050] Examples of suitable expression vectors for use in mammalian host cells include pcD A3.1/Hygro (Invitrogen), pDC409 (McMahan *et al.* 1991 *EMBO J* 10: 2821 -2832), and pSVL (Pharmacia Biotech). Expression vectors for use in mammalian host cells can include transcriptional and translational control sequences derived from viral genomes.

[0051] Commonly used promoter sequences and enhancer sequences that can be used to express rBHT RNA include, but are not limited to, those derived from human cytomegalovirus (CMV). Adenovirus 2, Polyomavirus, and Simian virus 40 (SV40). Methods for the construction of mammalian expression vectors are disclosed, for example, in Okayama and Berg (1982 *Mol Cell Biol* 2: 161-170), Cosman *et al.* (1986 *Mol Immunol* 23:935-941), Cosman *et al.* (1984 *Nature* 312: 768-771), EP-A-0367566, and WO 91/18982. The relevant portions of these references are incorporated herein by reference.

5.7. Purification Tags

[0052] In addition to the 6XHIS tag described herein a variety of purification methods may be used such as affinity tags, such as antigenic tags (e.g., FLAG (Sigma-Aldrich, Hopp *et al.* 1988 *Nat Biotech* 6:1204–1210), hemagglutinin (HA) (Wilson *et al.*, 1984 *Cell* 37:767), Intein fusion expression systems (New England Biolabs, USA) Chong *et al.* 1997 *Gene* 192(2), 271-281, or maltose-binding protein (MBP)), glutathione S transferase (GST)/glutathione, poly His/Ni or Co (Gentz *et al.*, 1989 *PNAS USA* 86:821-824),. Fusion proteins containing GST-tags at the N-terminus of the protein are also described in U.S. Pat. No. 5,654,176 (Smith). Magnetic separation techniques may also be used such as Streptavidin-DynaBeads® (Life Technologies, USA). Alternatively, photo-cleavable linkers may be used, *e.g.*, U.S. Pat. No. 7,595,198 (Olejnuk & Rothchild). Many other systems are known in the art and are suitable for use with the present invention.

5.8. Methods of Making Galacto-Oligosaccharides (GOS)

[0053] In one embodiment of the invention, the galacto-oligosaccharides (GOS) are produced by incubating the cell expressing the rBHT in a medium that comprises a disaccharide substrate such as for example lactose or cellobiose. In one embodiment, the GOS is produced from lactose simultaneously with a glucose removal system. The glucose removal system may be a generally recognized as safe (GRAS) organism.

5.9. Formulations

[0054] Another aspect of the invention concerns use of the rBHT protein or cells expressing rBHT to produce a foodstuff or a dietary supplement containing galacto-oligosaccharides (GOS). The foodstuff may be dairy foodstuff such as yogurt, cheese or fermented dairy products. The rBHT or cell expressing rBHT may be part added to the

foodstuff or dietary supplements. The rBHT may be dried using Spray Dry; a quick and gentle method of obtaining even the smallest quantities of temperature sensitive substances in powder form. The dried rBHT also may be encapsulated form using the Spray dryer's ability to coat particles, immobilize solid material in a matrix and manufacture microcapsules (www.buchi.com/Mini_Spray_Dryer_B-290.179.0.html). Other drug delivery applications using functional GRAS encapsulating agents and technologies may be used. The dried rBHT tablet and powder forms may be analysed for rBHT rate of activity once rehydrated in buffer containing lactose and in milk products.

[0055] Examples of the foodstuffs include, but are not limited to, infant formula, dairy products, beverages, and dietary supplements. See Table 2 below.

TABLE 2.

Food group	Food group category
Infant formulas for term infants and baby foods	Infant formula
	Infant meal replacement drinks
	Baby juice
	Baby yoghurt drinks
	Baby dessert
	Baby snack
Dairy products	Yoghurt
	Frozen dairy desserts
Milk beverages	Milk
	Milk drinks
	Syrup flavoring for milk
	Meal replacement drinks
	Milk substitutes
Fruit drinks and water quenchers	Fruit drinks (vitamin/mineral fortified) and energy drinks
	Fitness waters and thirst quenchers
Fruit preparations	Fruit pie filling
	Fruit prep
	Jelly jam

[0056] Any of the above-described rBHT proteins may be delivered in the form of a composition, that is, with one or more additional components such as a physiologically acceptable carrier, excipient, or diluent. For example, a composition may comprise a soluble rBHT protein as described herein plus a buffer, an antioxidant such as ascorbic acid, a low molecular weight polypeptide (such as those having less than 10 amino acids), a protein, amino acids, carbohydrates such as glucose, sucrose, or dextrin, chelating agent such as

EDTA, glutathione, and/or other stabilizers, excipients, and/or preservatives. The composition may be formulated as a liquid or a freeze-dried powder. Further examples of components that may be employed in pharmaceutical formulations are presented in Remington's Pharmaceutical Sciences, 16th Ed., Mack Publishing Company, Easton, Pa., (1980), the relevant portions of which are incorporated herein by reference.

[0057] Compositions comprising therapeutic molecules described above can be administered by any appropriate means including, but not limited to, parenteral, topical, oral, nasal, vaginal, rectal, or pulmonary (by inhalation) administration. If injected, the composition(s) can be administered intra-articularly, intravenously, intraarterially, intramuscularly, intraperitoneally or subcutaneously by bolus injection or continuous infusion. Localized administration, that is, at the site of disease, is contemplated, as are transdermal delivery and sustained release from implants, skin patches, or suppositories. Delivery by inhalation includes, for example, nasal or oral inhalation, use of a nebulizer, inhalation in aerosol form, and the like. Administration via a suppository inserted into a body cavity can be accomplished, for example, by inserting a solid form of the composition in a chosen body cavity and allowing it to dissolve. Other alternatives include eye drops, oral preparations such as pills, lozenges, syrups, and chewing gum, and topical preparations such as lotions, gels, sprays, and ointments. In most cases, therapeutic molecules that are polypeptides can be administered topically or by injection or inhalation.

[0058] The therapeutic molecules described above can be administered at any dosage, frequency, and duration that can be effective to treat the condition being treated. The dosage depends on the molecular nature of the therapeutic molecule and the nature of the disorder being treated. Treatment may be continued as long as necessary to achieve the desired results. The periodicity of treatment may or may not be constant throughout the duration of the treatment. For example, treatment may initially occur at weekly intervals and later occur every other week. Treatments having durations of days, weeks, months, or years are encompassed by the invention. Treatment may be discontinued and then restarted.

[0059] Maintenance doses may be administered after an initial treatment. Dosage may be measured as milligrams per kilogram of body weight (mg/kg) or as milligrams per square meter of skin surface (mg/m²) or as a fixed dose, irrespective of height or weight. These are standard dosage units in the art. A person's skin surface area is calculated from her height and weight using a standard formula. For example, a therapeutic rBHT protein can be administered at a dose of from about 0.05 mg/kg to about 10 mg/kg or from about 0.1 mg/kg

to about 1.0 mg kg. Alternatively, a dose of from about 1 mg to about 500 mg can be administered. Or a dose of about 5 mg, 10 mg, 15 mg, 20 mg, 25 mg, 30 mg, 35 mg, 40 mg, 45 mg, 50 mg, 55 mg, 60 mg, 100 mg, 200 mg, or 300 mg can be administered.

[0060] The article “a” and “an” are used herein to refer to one or more than one (i.e., to at least one) of the grammatical objects of the article. By way of example, “an element” means one or more elements.

[0061] Throughout the specification the word “comprising,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps. The present invention may suitably “comprise”, “consist of”, or “consist essentially of”, the steps, elements, and/or reagents described in the claims.

[0062] The following Examples further illustrate the invention and are not intended to limit the scope of the invention.

6. EXAMPLES

6.1. MATERIALS AND METHODS

[0063] **Design of a codon optimized β -hexosyl-transferase gene.** The DNA coding sequence for the *S. singularis* β -hexosyl-transferase gene (16) (GenBank accession number AB126324; 1,782 bp) was assembled by joining exons using Clone Manager software (Cary, NC). The coding sequence was redesigned based on *P. pastoris* and *E. coli* preferred codons, optimized for minimum free energy (-619.9), specific restriction sites (5' NcoI and 3' NotI), and GC content (48.89%). This redesigned version of the gene was labeled *rBht* (GenBank accession number JF29828), synthesized and inserted into pGS21a and pUC57 to generate pJB100 and pJB107, respectively (Table 1). The DNA sequence of *rBht* was confirmed (GenScript, Piscataway, NJ).

[0064] **Construction of plasmids containing *rBht* for expression in *E. coli*.** Cloning procedures were carried out as previously described (32). *E. coli* strains used for cloning and expression of rBHT are listed in Table 3. Restriction endonucleases and T4 DNA ligase were obtained from New England Biolabs (Beverly, MA). The plasmids pJB100 and pJB107 were digested with NcoI and NotI, and the fragment containing the *rBht* gene was inserted into

Novagen plasmids to generate the expression plasmids pJB101, pJB103, pJB104, pJB105, and pJB106 (see Table 3 for a description of the constructions).

[0065] Expression of rBHT fusion constructions in *E. coli* BLR. The expression was carried out as described in the pET system manual TB055 8th Edition 02/99 (Novagen). Briefly, the expression plasmids were transformed into *E. coli* BLR and after IPTG induction screened for rBHT activity. *In vivo* rBHT activity was assessed by incubating IPTG induced BLR cells in 50 mM sodium phosphate buffer at pH 4 and pH 6, and 50 $\mu\text{g}\cdot\text{ml}^{-1}$ cell penetrating chromogenic substrate X-GAL (Table 4). Cultures were incubated overnight at 37°C for visualization of the appearance of color. BL21 cells containing endogenous β -galactosidase activity were used as the positive control.

[0066] Production of anti-rBHT. Anti-rBHT antiserum was produced using rBHT-6XHIS expressed and purified from *E. coli* BLR cells harboring pJB101. Purification of rBHT-6XHIS present in the cell free extract fraction was performed by using nickel agarose gel chromatography according to the manufacturer's instructions (QIAGEN, Germany) and then further purified by electro-elution from gels according to manufacturer's instructions (Bio-Rad, Richmond, CA). The pure protein was used for rabbit immunization (Cocalico Biologics Reamstown, PA). Additional antibodies used in the study are listed in Table 4.

[0067] Electrophoresis and immunoblotting. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE-4-12%) was routinely carried out in the Laemmli system (23). Proteins were visualized by Coomassie blue (Bio-Rad, Richmond, CA) or silver staining (Bio-Rad, Richmond, CA). SeeBlue plus (Invitrogen, Carlsbad, CA) was used as a molecular mass marker. Immunoblot analysis on duplicate PAGE gels was carried out as previously described (7) except detection was performed using alkaline phosphatase conjugated goat anti-rabbit or goat anti-mouse antibodies (Rockland Immunochemicals, Gilbertsville, PA) and visualized using NBT (nitro-blue tetrazolium chloride) and BCIP (5-bromo-4-chloro-3'-indolylphosphate p-toluidine salt) premixed solution (Sigma-Aldrich, St. Louis, MO).

[0068] Construction of *P. pastoris* recombinant strain. The *rBht* gene was PCR amplified from pJB107 using primers JBB1 and JBB2 to add XhoI and NotI restriction sites in frame with *Saccharomyces cerevisiae* pre-pro-alpha-mating factor signal sequence contained in pPIC9, followed by a 6XHIS tag, and a modified TEV protease cleavage site (Table 2). The PCR product was ligated into pPIC9 XhoI and NotI restriction sites to

generate pJB108 (Table 3). Correct in-frame ligation was confirmed by sequencing (Sequatch, Mountain View, CA) using primers 5' AOX1 and 3' AOX1 (Table 4).

[0069] *P. pastoris* GS115 (Table 1) was electro-transformed with pJB108 linearized with SacI (Invitrogen's *Pichia* expression kit manual, version M) using a Bio-Rad Gene Pulser (Bio-Rad, Richmond, CA). Recombinants were selected on histidine deficient Regeneration Dextrose (RDB-agar plates) at 30°C. His⁺ colonies were randomly selected, and the genomic integration of the expression cassette was verified by PCR using primers 5' AOX1, 3' AOX1, and α -Factor (Table 4). The methanol utilization (mut⁺) phenotype of recombinant GS115/*rBht* was determined according to the procedure outlined in Invitrogen's *Pichia* expression kit manual, version M.

[0070] rBHT production in *P. pastoris*. To select a high-level producer of the recombinant rBHT, six His⁺ isolates were grown in yeast extract peptone dextrose medium (YPD) at 30°C and 250 rpm for 12 h and then used to inoculate 100 ml buffered glycerol complex medium (BMGY) to an initial OD₆₀₀=0.1. When the culture exceeded an OD₆₀₀=10, methanol was added to a final concentration of 0.5% at 24 h intervals until the culture exceeded an OD₆₀₀=50 after which methanol was added every 12 h. Media samples were analyzed for presence of BHT activity and by Western blot using rabbit anti-rBHT every 24 h to determine the optimal harvest time. The selected GS115/*rBht* recombinant was routinely grown in 0.5 L BMGY and induced with methanol for 6 days.

[0071] Purification of secreted rBHT. Culture supernatants (500 ml) were fractionated with ammonium sulfate. Precipitates between 60%-80% ammonium sulfate were resuspended in 50 mM sodium phosphate buffer (pH 6). After desalting and concentrating with an Amicon MWCO 15 membrane (Amicon Inc., Beverly, MA) the solution was applied to a 1/30 (5 ml) Mono Q pre-equilibrated column (Quarternary amino ethyl) (Amersham Biosciences). The column was then washed with 50 ml of buffer and eluted with 3 column volumes of a linear gradient of sodium chloride from 0 to 0.2 M in 50 mM sodium phosphate buffer (pH 6.0) at a flow rate of 1 ml.min⁻¹. The eluate was collected in 1 ml portions. The active fractions were pooled, concentrated and resuspended in 10 mM sodium phosphate (pH 6.8), then applied to a Bio-Gel HT hydroxyapatite column (Bio-Rad, Richmond, CA) (1/10 2 ml) pre-equilibrated with the same buffer, washed with 10 mM sodium phosphate (pH 6.8), and eluted with 50 mM sodium phosphate (pH 6.8). The fractions with the highest specific activity contained pure rBHT with specific activity of 8.2

U.mg⁻¹. Enzymatic activity was assayed (described below) on all chromatography fractions and purification steps were carried out at 25°C.

[0072] Determination of molecular mass. Culture medium concentrated 20 fold by ultrafiltration (0.5 ml) was applied to a size exclusion column Superdex 200 (Amersham Biosciences) 1/30 (18 ml) pre-equilibrated with 50 mM sodium phosphate buffer, pH 6.0, 0.1 M NaCl. Fractions of 0.5 ml were collected at a flow rate of 0.5 ml.min⁻¹ and assayed for rBHT activity using ONP-Glu as the substrate and by zymogram as described below. Elution of rBHT and molecular mass standards were monitored at 280 nm. The column was calibrated using the following molecules: Thyroglobulin, 669 kDa; Ferritin, 440 kDa; Catalase, 232 kDa; Lactate dehydrogenase, 140 kDa; Bovine Serum Albumin; 67 kDa (GE, Healthcare). The molecular mass of rBHT was extrapolated from a calibration plot of log molecular mass (Y axis) versus elution volume (X axis). All chromatographic steps were carried out at 25°C.

[0073] Enzymatic activity assays. The initial reaction rate of rBHT was measured by a modification of the Kuby's method (13,22) under the established conditions. Briefly the reactions were performed in a volume of 250 µl containing 1.3 mM ONP-Glu and 50 mM sodium phosphate buffer (pH 5). The assays were carried out for 10 min under the established conditions and stopped by adding 1 volume of 0.25 M Na₂CO₃. The reaction mixture containing boiled rBHT and substrate served as negative control. Assays were always performed in duplicate with a reliability of ±5%. Samples of cell-free broth, and protein concentrates were obtained as described above. Resting cells (harboring membrane-bound rBHT) prewashed with 50 mM sodium phosphate buffer (pH 5.0) were assayed, under established conditions. When X-GAL was the substrate of the reaction the concentration was 50 µg.ml⁻¹ in 50 mM sodium phosphate buffer (pH 4).

[0074] One unit (U) of enzyme activity equals 1 µmol of *o*-nitrophenol released per min under the assay conditions. Specific activity is defined as Units/mg protein. Molar extinction coefficients of *o*-nitrophenol were: ε= 0.033 mM⁻¹cm⁻¹, pH 4; ε=0.036 mM⁻¹cm⁻¹, pH 5; ε=0.038 mM⁻¹cm⁻¹, pH 6. The amount of *o*-nitrophenol released was extrapolated from a calibration plot of the *o*-nitrophenol absorbance at 405 nm (Y axis) versus *o*-nitrophenol concentration (X axis).

[0075] Enzymatic activities were also visualized by zymograms. Native PAGE were performed using a modification of the protocol described by Gallagher (8). Proteins from *E. coli* lysates or *P. pastoris* supernatants were solubilized in 5% (w/v) sucrose/10 µg.ml⁻¹

Bromophenol blue and separated in 6% native polyacrylamide gels, utilizing as running buffer 50 mM sodium phosphate buffer (pH 6). The gel was kept cool in a Mighty Small Hoefer electrophoresis apparatus where cold water was re-circulated during electrophoresis at 60 mA for 5 h. After electrophoresis, the gel was rinsed twice in wash buffer (50 mM sodium phosphate buffer, pH 4.0) for 10 min. The zymograms were developed for 24 h by laying filter paper soaked in wash buffer containing 50 $\mu\text{g}.\text{ml}^{-1}$ X-GAL at 20°C. A blue precipitate defined the location of the enzyme.

[0076] Enzyme kinetics. Series of enzyme dilutions ranging from 0 to 1.2 $\text{U}.\text{ml}^{-1}$ were assayed in 50 mM sodium phosphate (pH 5) at 42°C. The enzymatic activity assay was initiated by adding 1.3 mM ONP-Glu and the absorbance monitored at 405 nm for 1 min intervals for a total of 20 min. The experimental absorbance values were plotted against time showing linear proportionality up to 0.6 $\text{U}.\text{ml}^{-1}$ for at least 20 min while at enzyme concentrations above 1.0 $\text{U}.\text{ml}^{-1}$ the absorbance values plateau prior to 5 min.

[0077] The Michaelis-Menten constants (k_m and $V_{\max}$) of 0.2 $\text{U}.\text{ml}^{-1}$ rBHT (at 42°C) were determined by varying ONP-Glu from 0 to 10.4 mM in 50 mM sodium phosphate (pH 5) and measuring the initial reaction rate at 20°C, 30°C, 40°C, and 50°C. The kinetic constants at each temperature were determined with OriginPro 7.5 using nonlinear regression of the Hill equation with a Hill coefficient of 1. The values obtained under the established conditions were as follows (T, k_m , $V_{\max}$): (20°C, 0.37 mM, 0.09 $\text{mM}.\text{min}^{-1}$), (30°C, 0.48 mM, 0.12 $\text{mM}.\text{min}^{-1}$), (40°C, 0.71 mM, 0.23 $\text{mM}.\text{min}^{-1}$) and (50°C, 1.3 mM, 0.42 $\text{mM}.\text{min}^{-1}$). The fitting coefficients of regression (R^2) were 0.9869, 0.99065, 0.99115 and 0.98996 at 20°C, 30°C, 40°C and 50°C, respectively.

[0078] Characterization of rBHT. Enzymatic activity assays were performed under the established conditions described above. The influence of pH on enzyme activity was tested in buffer solutions including 50 mM sodium phosphate (pH 5.0 to 11.0), 50 mM citrate (pH 2.0 to 5.0) and 50 mM phosphate-citrate (pH 2 to 11) (Fig. 2A). Competitive inhibition by monosaccharides (glucose and galactose) was examined by varying their concentrations in the reaction mixture (Fig. 2C). Temperature and thermostability were determined by measuring residual activity at 20, 30, 40, 50, 60, 70 or 80°C (Figs. 2B and 2D). Similarly, enzymatic activity assays were used to evaluate additives as potential inhibitors/activators. The following additives up to 10 mM were tested: chelating agent (EDTA), reducing agents (dithiothreitol (DTT), 2-mercaptoethanol (2-ME), and copper (Cu^{2+})), and ions (monovalent cations; NH_4^+ , Cs^+ , K^+ , Na^+ , Li^+ , and Rb^+ ; divalent cations; Ba^{2+} , Ca^{2+} , Co^{2+} , Fe^{2+} , Mg^{2+} ,

Mn²⁺, Ni²⁺, and Zn²⁺; trivalent cation Ag³⁺). Heavy metals (Co²⁺, Fe²⁺, and Zn²⁺) were tested in 50 mM citrate buffer (pH 5.0) to avoid precipitation. Additionally, surfactants added to the reaction mixture at 1% (v/v) were: TritonX-100, Tween 20, Tween 80, and Sodium Dodecyl Sulfate (SDS). Solvents tested at 10% v/v included: ethanol, methanol, acetone, acetonitrile, PEG400, and glycerol.

[0079] GOS production and analysis. The standard transgalactosylation reactions, utilizing either purified rBHT or *P. pastoris* resting cells (harboring membrane-bound enzyme), were initiated by adding standardized amounts of enzyme (0.5 U g⁻¹ lactose) in 5 mM sodium phosphate buffer (pH 5.0) containing lactose (22 g L⁻¹ or 200 g L⁻¹) at 30°C or 42°C.

[0080] Products and substrates of the reactions were analyzed by high-performance liquid chromatography (HPLC) (Shimadzu Corporation, Kyoto, Japan) under isocratic conditions at 65°C and at 0.4 ml.min⁻¹ flow rate. The mobile phase was 5 mM sulfuric acid (H₂SO₄) using an Alltech IOA-1000 organic acids column (300 mm by 7.8 mm) (Alltech, IL) coupled to a refractive-index detector. The column was calibrated using; galactosyl-lactose (Carbosynth (Berkshire, UK)), lactose, glucose, and galactose (Sigma-Aldrich, St. Louis, MO). The residual non-quantified GOS species (tetrasaccharide and pentasaccharide) are reported as signal intensity readings from the refractive-index detector.

6.2. RESULTS

[0081] Expression of a recombinant β -hexosyl-transferase (rBHT) in *E. coli*. *E. coli* BL21 is the most widely used host for heterologous protein production. Unfortunately, this host strain contains an active endogenous β -galactosidase that interferes with the evaluation of the β -hexosyl-transferase, designated as rBHT (see materials and methods). After screening different *E. coli* strains appropriate for pET- based expression systems including BL21, BLR, NovaBlue, Origami, and Rosetta (Table 3), *E. coli* BLR was confirmed as lacking of endogenous β -galactosidase activity. *E. coli* CC118 (Δ lacZ) strain was used as a control (36).

[0082] The *rBht* gene was inserted into pET expression plasmids containing a C-terminal 6XHIS or one of the four solubility-enhancing co-expression partners (glutathione-S-transferase (GST), thioredoxin (Trx), the PelB leader, and DsbA) resulting in pJB101, pJB103, pJB104, pJB105, and pJB106 (Table 3). Transformation into *E. coli* BLR and induction with IPTG resulted in expression of inactive rBHT either in whole cells or in cell

free extracts. Immunoblotting analysis of the fusion proteins with rBHT-antiserum detected all rBHT fusion proteins at their predicted molecular masses, with the strongest reactivity observed in the insoluble fractions. To rule out possible host-dependent protein insolubility, rBHT expression was analyzed in *E. coli* CC118 harboring pJB102 (pJB101 with its T7 promoter replaced with a tetracycline (TET) inducible promoter) but also proved to be unsuccessful (data not shown).

[0083] Expression and purification of rBHT from *P. pastoris*. *P. pastoris* is able to introduce post-translational modifications and is well known for its ability to produce a number of active recombinant proteins (where *E. coli* fails) (14). Thus, we inserted the codon optimized *rBht* gene into pPIC9 under the control of the alcohol oxidase promoter (AOX1), in frame with the *S. cerevisiae* α -factor signal (sequence for protein secretion) and an N-terminal 6XHIS followed by a TEV protease cleavage site (pJB108, Table 3). *P. pastoris* GS115 was transformed with pJB108 (GS115/*rBht*) and the activity of rBHT was evaluated in six GS115/*rBht* recombinants. The recombinant strain secreting the highest concentration of bioactive protein was studied further. Zymograms confirmed the presence of an active rBHT: only GS115/*rBht* gave a positive signal, while cell extracts from *E. coli* BLR harboring pJB101 and culture supernatants from untransformed GS115 were negative (Fig. 1C). As expected, protein transmembrane regions in BHT also resulted in GS115/*rBht* cells displaying cell surface-associated rBHT activity, emulating the location of native BHT in *S. singularis* (15,16).

[0084] Purification of rBHT was attempted using nickel affinity chromatography, but the HIS-tag was not present, nor was the protein detected by Western blot analysis using anti-HIS antiserum, indicating possible processing of the N-terminal signal sequence at predicted cleavage sites (16). Subsequently, the rBHT enzyme was purified (specific activity of 8.2 U.mg⁻¹ at 20°C) using Mono Q and hydroxyapatite chromatography (Table 5).

6.3. Characterization of rBHT expressed in *P. pastoris*.

[0085] 1. Apparent molecular mass of rBHT. The estimated molecular mass for a non-glycosylated fully processed rBHT that included the 6XHIS and TEV protease site tag was 68 kDa. The enzyme has been previously purified as a dimer as well as a monomer with apparent molecular mass ranging from 53 to 146 kDa data that reflects variations in protein glycosylation (Table 5) (1,16). Here, the enzyme activity eluted as one monomeric peak with an experimental apparent molecular mass of 110 kDa. We surmised that the dimeric form

may predominate within the acidic range of the native enzyme's pH optimum (3.7 to 6) (Table 5). However, fractions from the Superdex 200 column at pH 4.0 depicted the same profile, confirming the stable monomeric form of the recombinant enzyme. No higher molecular mass aggregates were detected by enzyme activity assay, zymogram, or Western blot. Purification of the column fractions and immunoblot analysis using anti-rBHT verified that the enzyme migrated as a single band with an apparent molecular mass of 110 kDa (Figs. 1A and 1B).

[0086] 2. Substrate specificity. ONP-Gal has traditionally been used as substrate for β -galactosidases and ONP-Gal, PNP-Gal and PNP-Glu have all been used in previous studies for detection of native BHT activity (Table 5). The enzyme activities of rBHT (0.2 U.ml⁻¹) were compared between the above substrates at the same experimental conditions. The recombinant enzyme was equally active in response to ONP-Glu and PNP-Glu. The substrates with a galactose in the glycon moiety were hydrolyzed at a rate of approximately 41% (ONP-Gal) and 23% (PNP-Gal) of that for ONP-Glu. These results indicate that rBHT has a narrow specificity with respect to the sugar and more flexibility toward configuration of the carbon linkage position at C2 and C4 when glucose sugar derivatives are used as substrates.

[0087] 3. Optimum pH, temperature and thermostability of rBHT. rBHT was active within a broad pH range (from pH 3.5 to 6) displaying the highest values at pH 4.0 (Fig. 2A). The enzymatic activity profile showed a steep decline to less than 50% maximal enzyme activity at pH values greater than 7 or less than 3.5. These results were consistent with reported pH optimum (1) (see references from Table 5); suggesting that alkaline conditions may have a detrimental effect on enzyme activity and stability.

[0088] The initial reaction rate measured at different temperatures ranging from 20°C to 80°C indicated that the enzyme was active over a temperature range from 20°C to 50°C with the optimum occurring between 40°C and 50°C (Fig. 2B). At temperatures below 30°C a 50% reduction in the initial reaction rate is observed and temperatures above 50°C quickly and irreversibly inactivated the enzyme. The optimum temperature when maximizing the enzyme reaction rate can also be obtained from the highest value of $V_{\max}/K_m$ (33). $V_{\max}$ increased at a faster rate than K_m at temperatures between 20°C to 40°C, consequently the $V_{\max}/K_m$ values (0.242 min⁻¹, 20°C; 0.255 min⁻¹, 30°C; 0.324 min⁻¹, 40°C; 0.322 min⁻¹, 50°C) increased over this range and were constant at temperatures between 40°C and 50°C. Thus

the optimum temperature as determined by $V_{\max}/K_m$ was within 40°C-50°C, confirming the optimal values established using the initial reaction rate values at each temperature.

[0089] As shown in Fig. 2D, the thermostability of rBHT was evaluated from 20°C to 50°C. At 20°C and 30°C the enzyme retained at least 90% of the original activity for 6 days, confirming previously reported results for the native enzyme (4). Five independent batches of rBHT stored for 6 months at 4°C retained 95% of the initial activity (data not shown). Although the optimal temperature was found in the 40°C to 50°C range, incubation at temperatures above 40°C was deleterious to rBHT. At 40°C, the enzyme retained 70% of the initial activity by 12 h and this level of activity only persisted for an additional 36 h. In contrast, enzyme activity decreased sharply at 50°C within the first 12 h incubation period.

[0090] *4. Effect of metals, salts, surfactants, and solvents on rBHT activity.* Enzyme inhibition/activation was tested within a broad range of additives. rBHT did not exhibit a requirement for any of the ions tested (NH_4^+ , Ba^{2+} , Ca^{2+} , Cs^+ , Co^{2+} , Cu^{2+} , Li^+ , Rb^+ , Mg^{2+} , Ni^{2+} , Na^+ and Zn^{2+}) even though magnesium dramatically increases the enzyme activity of some β -galactosidases (26). Moreover, the recombinant enzyme was fully active in the presence of 1 and 5 mM of the ion-chelating agent EDTA, confirming the above findings and a previous report (1). Additionally, compounds proven to disrupt disulfide bridges, such as Cu^{2+} and the reducing agents dithiothreitol (DTT) and 2-mercaptoethanol (2ME) (1,18,19), had no negative consequences on the activity. The solvents methanol, ethanol, acetone and acetonitrile only partially inhibit the enzyme (retaining 66%-81% relative activity). In contrast, the addition of 10% glycerol or 1% of SDS (a non-ionic surfactant) almost completely inhibited the enzyme.

[0091] **GOS synthesis using purified rBHT.** Once the enzyme was characterized, the secreted rBHT was tested for biotransformation of 2% lactose into GOS. The conditions of the reaction were 0.5 U rBHT.g⁻¹ of lactose at 42°C in 5 mM sodium phosphate buffer (pH 5.0). Figure 3 shows lactose consumption and GOS accumulation over time. The highest rate of production was observed during the first 12 h and galactosyl-lactose and glucose were the main products. Galactose was not detected, indicating that it was being completely incorporated in the transgalactosylation reaction to form GOS. After 30 h, 4.2 g.L⁻¹ of glucose had accumulated and lactose utilization (54%) was at its maximum. Furthermore at this time point, 7.8 g.L⁻¹ of the trisaccharide had accumulated reaching an average of 67% conversion of the utilized lactose. As the reaction proceeded, galactose began to escape the enzymatic reaction and accumulate at trace concentrations. Since competitive enzyme

inhibition could reduce the efficiency of lactose biotransformation, we examined the effect of varying concentrations of glucose or galactose on enzyme activity in the reaction mixture. The presence of 5 g.L⁻¹ glucose reduced rBHT activity up to 90% whereas enzymatic activity was unaffected by up to 70 g.L⁻¹ galactose, under the established conditions (Fig. 2C).

[0092] GOS synthesis by *P. pastoris* resting cells (harboring membrane-bound rBHT). To avoid competitive inhibition and confirm that conversion of lactose into GOS could be improved upon if glucose is simultaneously eliminated from the reaction mixture, we evaluated the biotransformation of 20% lactose by resting cells of *P. pastoris* GS115/*rBht*. *P. pastoris* GS115/*rBht* harboring membrane-bound rBHT were normalized to a cell density containing 0.5 U rBHT.g⁻¹ of lactose in 5 mM sodium phosphate buffer (pH 5). Reactions were conducted for 10 days at 30°C, the optimal temperature for growth of *P. pastoris*, and at 42°C, the temperature for which the initial reaction rate is at its maximum for the secreted rBHT. As expected, 90% of the initial lactose was converted into GOS with no secondary products at 30°C, as compared to 51% lactose utilization at 42°C. The results indicated that resting cells were physiologically active and able to consume the glucose byproduct of the reaction, thereby avoiding competitive inhibition. However, the initial reaction rate of GOS formation at 42°C was double the rate at 30°C during the first 48 h. A final concentration of 80 g.L⁻¹ galactosyl-lactose was reached, corresponding to a productivity of 1.6 g.L⁻¹.h⁻¹ at 42°C (Fig. 4A). At 30°C when the lactose utilization was at 63% the concentration of galactosyl-lactose was 100 g.L⁻¹ and a productivity of 0.8 g.L⁻¹.h⁻¹ was reached after 5 days (Fig. 4B).

6.4. DISCUSSION

[0093] In this study we optimized the DNA sequence of the β -hexosyl-transferase gene from *S. singularis* (acc. number AB126324) for expression in *E. coli* and *P. pastoris*. The resulting *rBht* gene was synthetically generated (acc. number JF29828) and expressed in *E. coli*. However, this bacterial host lacked the ability to incorporate post-translational modifications essential for producing a soluble and active rBHT as previously observed (16). Subsequently, the *rBht* gene under the control of the *AOX1* promoter was successfully integrated into the *P. pastoris* chromosome, resulting in the expression of a fully active enzyme that was detected in the culture broth as well as associated with the cell surface. Secretion of rBHT by *P. pastoris* GS115 allowed us to avoid the complex purification processes that are required to obtain pure BHT from *S. singularis*. Furthermore, since *P.*

pastoris naturally secretes only very low levels of native proteins, recovery of the extracellular rBHT was as simple as removal of whole cells from the medium by centrifugation or filtration (5).

[0094] The molecular mass of the recombinant enzyme corresponded to a single 110 kDa catalytically active polypeptide and no smaller polypeptides or rBHT aggregates were detected. Posttranslational modifications play a critical role in protein folding, structural stability, oligomer formation and substrate recognition (17,24), so it was not surprising that the molecular mass was higher than the 68 kDa protein expressed in *E. coli* and predicted by the amino acid sequence. Posttranslational glycosylation of the native BHT by *S. singularis* has been previously reported, and a shift in the molecular mass of the purified protein from 73.9 to 66.3 kDa was observed after treatment with chitinase and EndoHf (16). Future mutagenesis of the predicted glycosylation sites should aid in determining whether glycosylation is also the cause for the shift in rBHT molecular mass.

[0095] The data reported here are part of a larger study that compared the utilization of rBHT to documented data for the native *S. singularis* BHT (Table 5). Our study confirmed that the recombinant enzyme does not require cofactors or reducing agents often essential for β -galactosides. The enzyme showed better thermostability at lower temperatures (below 40°C) and optimal activities at temperatures from 40°C to 50°C and pH 3.5 to 6. Additionally, the enzyme was controlled by glucose inhibition though rBHT was not sensitive to galactose or Ag^{3+} as previously reported for native BHT (4,16,25,35).

[0096] Lactose utilization and initial lactose concentration are two key factors that contribute to maximizing the final GOS accumulation. Here, we demonstrated a process with improved lactose utilization (90%) employing physiologically active resting cells of *P. pastoris* GS115/rBht. Under these conditions, the cells consume residual glucose at 30°C, circumventing glucose inhibition and ensuring a significant process improvement and a higher degree of final prebiotic purity (Fig. 4B). In contrast, temperatures higher than 25°C were reported as preventing *S. singularis* resting cells containing membrane-bound native BHT from consuming residual glucose, which in turn limits final GOS concentration and purity (reviewed by Gosling *et al.* (11)).

[0097] GS115/rBht resting cells incubated at 42°C only converted 51% of the initial lactose into galactosyl-lactose and residual glucose (Fig. 4A). These data closely resembled galactosyl-lactose formation by secreted rBHT under the same conditions (Fig. 3). Furthermore, conversion and utilization values were comparable to previously reported

processes by *S. singularis* (Table 5). Typical lactose utilization values have been reported between 30% to 40% of initial lactose (11), with one study reporting 50% lactose utilization using 10.8 times more enzyme per gram of lactose compared to secreted rBHT (Table 5)(4).

[0098] The discovery of GOS synthesis by *S. singularis* in the mid-20th century has encouraged the exploration for superior β -galactosidases that more efficiently produce GOS (9,10,37). However, the enzymes studied showed lower lactose utilization values and higher final concentrations of undesirable byproducts when compared to the BHT from *S. singularis* (1,4,12,16,27,34,35). Nevertheless, advances in the industrial utilization of *S. singularis* BHT have been slower than desired due to the challenging multistep purification processes required to obtain pure native BHT (1,4,16). The bioactive rBHT either secreted or membrane bound enzyme from *P. pastoris* signifies a clear process advantage for producing GOS. Future studies will explore protein modification strategies to improve protein expression yield, protein stability, and enzyme activity.

[0099] Table 6 below summarizes the constructs that were prepared in connection with this invention. All the constructs were prepared in pPIC9 but Exist as chromosomal integrants in *P. pastoris*.

6.5. Galacto-oligosaccharides from lactose-rich whey

[00100] The global dairy market was \$299.7B in 2009, and is expected to grow to \$370.9B by 2014, an increase of 23.8%. Global. Dairy Industry Profile: Global. *Dairy Industry Profile: Global* [serial online]. October 2010:1. Available from: Business Source Complete, Ipswich, MA. Accessed February 24, 2012. *Global* (2012). Milk and cheese accounted for 35.2% and 28.3% of the market, respectively. Global cheese consumption is expected to reach 21 million tons in 2015. Lenoir-Wijnkoop I. & van Aalderen WM,B.G.K.D.S.A.N.MJ. Cost-effectiveness model for a specific mixture of prebiotics in The Netherlands. *Eur J Health Econ* (2010). The growth of the dairy market creates significant industrial waste treatment issues given the high lactose content of milk and byproducts produced during dairy processing. High lactose content waste fluids have an exceptionally high biological oxygen demand (BOD), which means that the amount of oxygen required to break down the lactose is high enough to rob other organisms of oxygen needed for survival. Therefore, many countries have enacted environmental protection laws that restrict the disposal of lactose containing fluids directly into bodies of water. Ganzle,M.G., Haase,G. & Jelen,P. Lactose: Crystallization, hydrolysis and value-added derivatives. *International Dairy Journal* **18**, 685-

694 (2008). The added burden to municipal water treatment processes can be especially costly and problematic for countries and states dependent on a dairy economy (4-6). Affertsholt-Allen T. Market developments and industry challenges for lactose and lactose derivatives. IDF Symposium "Lactose and its Derivatives." Moscow. (*lactose.ru/present/ITage_Affertsholt-Allen.pdf*. Accessed Sept 30, 2009) (2007); Markets and Markets. U.S. Digestive Health Ingredients Market Worth \$495.3 million in 2015. (*www.marketsandmarkets.com/PressReleases/us-digestive-health-ingredients-market.asp*) (2010); UBI consulting. THE WORLD GALACTO-OLIGOSACCHARIDE MARKET. (*www.ubic-consulting.com/template/fs/documents/Dairy-Ingredients/Galacto-Oligosaccharide-Ingredient-Market.pdf*) (2010). For example, cheese manufacturing generates two products: cheese and whey. For every pound of cheese made, nine pounds of whey are generated, creating a growing surplus of whey (186 million tons in 2008), which contains ~5% lactose. This lactose fraction has a BOD that is approximately 175-fold greater than typical sewage effluent, therefore the untreated waste cannot be directly disposed into bodies of water. Smithers, G.W. Whey and whey proteins. "From Glitter to Gold". *International Dairy Journal* **18**, 695-704 (2008). The traditional solution to this problem has been to bioremediate lactose-rich effluents by applying expensive processes to extract the lactose, which can then be sold as a commodity product at a ceiling value of \$1.50/kg. Only 50% of the cheese whey produced annually is recycled into useful products such as food ingredients and animal feed. The rest is considered waste either because critical volumes to allow for economical recycling are not reached or due to the high degree of technical difficulty involved in biotransformation. Therefore, there is a strategic need to convert lactose into commercially viable, high value products to reduce the overall process cost and improve the US dairy industry economy.

[00101] Here, we propose the simultaneous biotransformation/bioremediation of the commodity chemical lactose by applying a new food product development process as the ideal solution to this industrial problem. We have improved a method by which lactose can be converted into galactosyl-lactose derivatives called galacto-oligosaccharides (GOS) through an enzymatic reaction. GOS are classified as prebiotics, which stimulate the growth and activity of beneficial bacteria in the digestive system, and are widely used in food products such as infant formulas, nutritional supplements, yogurts, baked goods, and animal feed. Unlike the commodity product lactose, GOS are highly valued food ingredients, and the economic value of this transformation is easily demonstrated by comparing the current

market price of lactose at \$1.50/kg to the \$5.20-8.50/kg market price of GOS. GOS are a part of a trend in digestive health food ingredients valued at \$265.9 million in 2010 with an annual growth rate of 18.3% and expected to grow at a compound annual growth rate of 13.2% from 2010 to 2015.

[00102] Our lactose to GOS conversion method is vastly superior to existing processes because we can reduce the overall volume of reaction by 50 fold by utilizing an efficient host to produce the enzyme, increase the volume and purity of GOS produced, and potentially generate lactose-free products. Euromonitor. Lactose-free Foods Maintain Their Global Appeal, March 1, 2011. *Euromonitor* (2011). The lactose free product global market was \$3.4B in 2009, and is expected to grow as consumers continue to focus on health and wellness functional foods. While other dairy products are extremely price sensitive, functional foods such as lactose-free products can be sold at a premium.

[00103] By using the improved process described herein, US and global dairy industry and food supplement manufacturers can clearly benefit in three ways: 1) creation of large volumes of quality GOS, a health promoting food ingredient/dietary supplement with high market value, 2) simultaneous potential generation of valuable lactose free milk or whey products, and 3) cost effective reduction of environmental impact through the recycling of whey and milk byproducts.

TABLE 3. Strains and plasmids used in this study

<i>Strains/Plasmids</i>	<i>Description or genotype</i>	<i>Source or Reference</i>
<i>Strains</i>		
<i>E. coli</i>		
BL21	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm</i> (DE3)	Novagen
BLR	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcmΔ(srl-recA)306::Tn10(Tet^R)(DE3)</i>	Novagen
CC118	F2 <i>D(ara-leu)7697 araD139 Δ(lac)X74 phoAD20 galE galK thi rpsE rpoB argE(Am) recA1</i>	(36)
NovaBlue	<i>endA1 hsdR17(r_{K12}⁻ m_{K12}⁺) supE44 thi-1 recA1 gyrA96 relA1 lacF[proA⁺B⁺ lacI^q ZΔM15::Tn10] (Tet^R)</i> (DE3)	Novagen
Origami	Δ(<i>ara-leu</i>)7697 Δ <i>lacX74 ΔphoA PvuII phoR araD139 ahpC galE galK rpsLF[lac⁺ lacI^q pro] gor522::Tn10 trxB</i> (Kan ^R , Str ^R , Tet ^R) (DE3)	Novagen
Rosetta	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm pRARE2</i> (Cam ^R)(DE3)	Novagen
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI^q ZΔM15 Tn10</i> (Tet ^R)]	Stratagene
<i>P. pastoris</i>		
GS115	<i>his4</i> (his ⁻ mut ⁺)	Invitrogen
GS115/LacZ	GS115 <i>his4::LacZ</i> (<i>E. coli</i> β-galactosidase (117 kDa) intracellularly (his ⁺ mut ⁺))	Invitrogen

KM71	GS115 <i>arg4 his4 aox1::ARG4</i> (his ⁻ mut ^s)	Invitrogen
JB208	GS115 integrated with plasmid pJB108 (his ⁺ mut ⁺)	This study
(GS115/ <i>rBht</i>)		
Plasmids		
<i>E. coli</i>		
pJB100	pGS21a-rBHT	This study
pET24d	Optional C-terminal 6XHIS tag, T7lac promoter, Kan ^R	Novagen
pJB101	pET24d-rBHT-6XHIS	This study
pJB102	pET24d-TET promoter-rBHT-6XHIS	This study
pET41a	GST tag, T7lac promoter, Kan ^R	Novagen
pJB103	pET41a-GST-rBHT	This study
pET32a	Trx tag, T7lac promoter, Amp ^R	Novagen
pJB104	pET32a-Trx-rBHT	This study
pET22b	PelB coding sequence, T7lac promoter, Amp ^R	Novagen
pJB105	pET22b-pelB-rBHT	This study
pET39b	DsbA●tag TM coding sequence, T7lac promoter, Kan ^R	Novagen
pJB106	pET39b-DsbA-rBht	This study
pJB107	pUC57-rBHT	This study
<i>P. pastoris</i>		
pPIC9	<i>P. pastoris</i> expression plasmid carrying <i>AOX1</i> promoter and transcription terminator, <i>HIS4</i> , Amp ^r in <i>E. coli</i> , PBR322 ori, α -factor secretion signal from <i>S. cerevisiae</i>	Invitrogen
pJB108	pPIC9- α MF-6XHIS-TEV-rBHT	This study
α MF, <i>S. cerevisiae</i> α -mating factor secretion signal.		

TABLE 4. Primers, antibodies and substrates used in this study (SEQ ID NOS: 21-27)

Primers	Open Reading Frame	^a Sequence	Source
JBB1	6XHIS-TEV-rBHT forward primer	5'- ccgCTCGAGAAAAGAGAGGCTGAAG CTCACCACCACCACCA CCACGAAAACCTGTATTTTCAGATG ATGCTGCATGCTGCAC-3'	This study
JBB2	<i>rBHT</i> reverse primer	5'- aaggaaaaaaGCGGCCGCTTACAGA TGATTACGCCCAAATTG-3'	This study
JBB3	<i>rBHT</i> forward internal sequencing primer	5'- ATCACTATGCCAGCACGCAGTGT A-3'	This study
JBB4	<i>rBHT</i> reverse internal sequencing primer	5'- TTTAAAGCCGATTTCACCTGCC GC-3'	This study
5' AOX1 sequencing Primer	AOX1	5'-GACTGGTTCCAATTGACAAGC-3'	Invitrogen
3' AOX1	AOX1	5'-GCAAATGGCATTCTGACATCC-3'	Invitrogen

sequencing primer				
α -Factor	α -factor	5'-TACTATTGCCAGCATTGCTGC-3'	Invitrogen	
sequencing primer				
<i>Antibodies</i>	<i>Antigen</i>			
Mouse anti-HIS	6XHIS		Qiagen	
Rabbit anti-Bgal	<i>E coli</i> β -galactosidase		Sigma	
Rabbit anti-rBHT	β -hexosyl-transferase		This study	
<i>Substrate</i>	<i>Abbreviation</i>			
<i>o</i> NP- β -D-glucopyranoside	ONP-Glu		Sigma	
<i>o</i> NP- β -D-galactopyranoside	ONP-Gal		Sigma	
<i>p</i> NP- β -D-glucopyranoside	PNP-Glu		Sigma	
<i>p</i> NP- β -D-galactopyranoside	PNP-Gal		Sigma	
5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside	X-GAL		Sigma	

^a Coding regions are capitalized, restriction sites have been underlined.

TABLE 5. Reports evaluating BHT from *Sporobolomyces singularis* for the production of galactooligosaccharides (GOS)

<i>Enzyme Named</i>	<i>MM</i>	<i>SA</i>	<i>T</i>	<i>pH</i>	<i>U.g⁻¹ L⁻¹ lac</i>	<i>L_{int}</i>	<i>L_{util}</i>	<i>C_{max} (Y)</i>	<i>Conditions Used</i>	<i>Ref</i>
β -transglycosyl	—	—	22	3.75-4.0	—	10	—	50	STR. Growing cells	(9)
β -hexosyl-transferase	—	—	20	6.5	—	6	68	25(36)	Cell extract	(10)
β -hexosidase	140-145	41.2 ^a	45-50	6.5	—	5	—	—	Purified enzyme	(1)
β -galactosidase	—	—	45	3.7	—	30	73	54 ^e (51)	Batch IE, partially purified enzyme	(34)

β -galactosidase	—	—	45	4.8	—	10	70	55 ^e (53)	Cont. IE, PBR, partially purified enzyme	(34)
β -galactosidase	—	—	45	3.7	0.13	30	70	40(57)	Partially purified enzyme	(35)
β -galactosidase	53	56 ^b	50	5.0	5.4	18	71	50(70)	Purified enzyme	(4)
β -galactosidase like	146	8.6 9 ^c	40	6.0	0.8	20	—	—	Purified enzyme	(16)
β -galactosidase	—	—	55	5.0& 6.0	—	60	60	41.1(68)	Batch, resting cells	(31)
β -galactosidase	—	—	55	5.0& 6.0	—	60	60	40.4(67)	RB IE Alginate, resting cells	(31)
β -hexosyl-transferase	110	8.2 ^d	42	6.0	0.5	2	52	37 ^f (71) ^f	Purified enzyme	This study
β -hexosyl-transferase	110	—	42	6.0	0.5	20	52	36 ^f (70) ^f	Batch, recombinant resting cells	This study
β -hexosyl-transferase	110	—	30	6.0	0.5	20	69	51 ^f (74) ^f	Batch, recombinant resting cells	This study

MM, molecular mass (kDa); SA, specific activity (U.mg⁻¹ enzyme); T, temperature (°C); U.g⁻¹lac, units per gram initial lactose; L_{int}, initial lactose concentration (%); L_{util}, lactose utilized (%); C_{max}, maximum conversion of GOS (%) from initial lactose; (Y), conversion % (total GOS formed from utilized lactose); STR, stirred tank reactor; IE, immobilized enzyme; PBR, packed bed reactor; Cont. IE, immobilized enzyme (continuous); RB IE, immobilized enzyme (repeated batch). Substrates of the enzyme reaction ^aONP-Gal; ^bPNP-Gal; ^cPNP-Glu; ^dONP-Glu. ^eGOS include disaccharides; ^fGOS values reported were performed at the value of maximum accumulation of trisaccharide (galactosyl-lactose)

TABLE 6. *Pichia pastoris* rBHT Proteins

FP #	Protein Construct			Protein construct without aMF		Signal sequence	Plasmid
	Name	Amino Acid #	Mol. Weight	Amino acid # w/o aMF	Mol. weight w/o aMF		
1	aMF-6XHIS-TEV(Q/M)-rBHT	695	76599.3	606	67279.3	aMF and BHTss	pPIC9
2	aMF-6XHIS-TEV(Q/M)-aMF-rBHT-6XHIS	701	77422.2	612	68102.1	aMF and BHTss	pPIC9
3	aMF-rBHT-6XHIS	689	75805.8	600	66484.4	aMF and BHTss	pPIC9
4	aMF-rBHT	683	74981.6	594	65661.6	aMF and BHTss	pPIC9
5	aMF-rBHT(Δ 1-22)-6XHIS	667	73604.7	578	64284.6	aMF	pPIC9
6	aMF-rBHT(Δ 1-22)	661	72781.8	572	63461.7	aMF	pPIC9
7	rBHT-6XHIS	600	66483.3	600	66484.4	BHTss	pPIC9
8	rBHT(Δ 1-22)-6XHIS	579	64414.7	579	64415.8		pPIC9

9	aMF-rBHT(Δ 1-110)-6XHIS	584	65005.4	491	55146.5	aMF	pPIC9
---	---------------------------------	-----	---------	-----	---------	-----	-------

aMF=*Saccharomyces cerevisiae* alpha mating factor found in pPIC9 vector
 BHTss=BHT signal sequence found in amino acids 1-22

6.6. Secretion of β -hexosyltransferase is enhanced by replacing signal domain

6.6.1. ABSTRACT

[00104] The β -hexosyltransferase (BHT) from *Sporobolomyces singularis* is known for its ability to catalyze transgalactosylation reactions and synthesize galacto-oligosaccharides (GOS). We previously reported the heterologous expression of a bioactive full-length polypeptide (rBHT) by a recombinant strain of *Pichia pastoris* (GS115::aMF-HIS-TEV-rBht). This recombinant strain carries the full length *Bht* gene preceded by the *Saccharomyces cerevisiae* α mating factor pre pro signal (aMF), a histidine tag, and a TEV cleavage site. After methanol induction the rBHT generated by GS115::aMF-HIS-TEV-rBht was only partially secreted and most of the protein remained associated to the cell membrane. To increase the amount of secreted rBHT, this work examines the uncharacterized BHT amino-terminus region (amino acids 1-110) containing two putative endogenous structural domains. The amino terminus includes a domain (amino acids 1-22) which may serve as a classical secretion leader signal while the remaining 23-110 amino acids contain a putative non-classical secretion signal. Thus, we functionally evaluated these domains by generating recombinant *P. pastoris* GS115 strains expressing rBHT-HIS. The results show signal interference affecting protein secretion when aMF was followed by the rBht₍₁₋₂₂₎ classical leader signal (amino acids 1-22), while the substitution of the leader signal (amino acids 1-22) with the aMF (aMF-rBht₍₂₃₋₅₉₄₎) enhanced *P. pastoris* production of both secreted and membrane bound enzyme by as much as 50 and 14 fold, respectively. To validate the BHT amino-terminus domains role promoting protein secretion, we tested the domains with a non-secreted alternative protein, the anti- β -galactosidase single chain variable antibody fragment scFv13R4. Recombinant *P. pastoris* strains expressing combinations of the aMF and amino-terminus domains of rBHT, followed by the antibody scFv13R4 were able to generate results that correlate with the strength of secretion obtained by the recombinants expressing rBHT-HIS. Finally, active rBHT-HIS and rBHT proteins obtained from the more efficient recombinants (GS115::aMF-rBht₍₂₃₋₅₉₄₎-HIS and GS115::aMF-rBht₍₂₃₋₅₉₄₎) were purified to

homogeneity and evaluated for possible alterations in enzyme activity. The enzymatic activity of the 6XHIS tagged and the non-tagged secreted enzymes were comparable as shown by the rates of GOS generation.

6.6.2. INTRODUCTION

[00105] There is an increasing interest in the use of enzymes for the production of functional foods, especially in the field of prebiotic production from lactose to obtain lactose derivatives. *Sporobolomyces singularis* can assimilate lactose and glucose but not galactose indicating that they are only capable of metabolizing the glucose portion of lactose. Moreover, the unutilized galactose monomer can be only found in the broth as a constituent of the galacto-oligosaccharides (GOS). This physiological feature led to the discovery of the β -hexosyltransferase (BHT) (Blakely and Mackenzi 1021-25; Phaff and Carmo-Sousa 193-207; Spencer, de Spencer, and Lalue 147-56; Gorin, Phaff, and Spencer 1341-44; Gorin, Spencer, and Phaff 2307-17). Prebiotics such as GOS synthesized by the BHT from *S. singularis* are recognized as GRAS and widely used as a functional food additive (Tzortzis and Vulevic 207-44).

[00106] In addition to the synthesis of GOS from lactose, BHT also catalyzes the hydrolysis of β -glycosidic linkages such as ONP-Glu and PNP-Glu (Blakely and Mackenzi 1021-25). BHT enzymatic capabilities are particularly appealing with respect to competing technologies, since it is one of few enzymes capable of catalyzing the production of GOS with industrial advantages including; catalysis occurring in the absence of additional ions and cofactors as well as its ability to perform transgalactosylation reactions independently of the initial lactose concentration (Gosling *et al.* 307-18; Blakely and Mackenzi 1021-25).

[00107] In *S. singularis*, *Bht* is an inducible gene that is repressed by glucose and when in the presence of an inductor such as lactose the generated enzyme (BHT) is mostly found associated to the cell membrane. Due to the cellular location, the purification of BHT requires multiple chromatography steps and has been recovered from *S. singularis* at very low yields ranging from 14% to 16% (Blakely and Mackenzi 1021-25; Cho, Shin, and Bucke 2107-11; Ishikawa *et al.* 331-39). Since, conventional protein purification protocols limit enzyme recovery and thereof its technological application, alternative strategies have been evaluated. The first approach consisted of exposing *S. singularis* to selection through mutagenesis. Applying this methodology a new strain was selected lacking glucose repression and able to generate a 10-fold increase in membrane bound BHT; however, there

was no reported increase in the production of secreted enzyme (Ishikawa *et al.* 331-39). Alternatively, we recently described that *P. pastoris* GS115 is capable of secreting a biologically active recombinant rBHT polypeptide when preceded by the α MF prepro secretion signal consisting of a 19 amino acid signal sequence (presequence) followed by a 66 amino acid prosequence and a dibasic Kex2 endopeptidase processing site (Kurjan and Herskowitz 933-43). In this study the analysis of the cell free extract and membrane-bound associated activity showed that the majority of the enzyme remained associated to the cell membrane. Hence showing that *P. pastoris* GS115 is an adequate host for production and secretion of the bioactive rBHT that will facilitate downstream processing and demonstrating the feasibility of production of both secreted and membrane-associated bioactive rBHT (Dagher *et al.*, 2013).

[00108] A further look into the native BHT protein sequence showed that it contains endogenous structural features at the amino-terminus, including amino terminal domains that may serve as suitable classical and non-classical secretion signals. Supporting these roles as leader signals, it has been shown that following treatment of *S. singularis* with cell wall lytic enzymes most of the released BHT was devoid of the amino terminal classical leader signal (Ishikawa *et al.* 331-39). Since the efficiency of gene expression and protein secretion may be affected by those protein structural elements participating in cell association and secretion, the same functions may also be extended to the protein cellular localization when expressed by *P. pastoris* GS115.

[00109] In this study, we tested the physiological role of the BHT amino-terminal domains and their relationship with protein secretion by *P. pastoris* GS115. Furthermore, the secretory roles of rBHT amino-terminal domains were validated using recombinant chimeras containing as carboxyl terminal the single chain anti- β -galactosidase antibody scFv13R4. The antibody scFv13R4 is an example of a non-secreted hyper-stable single chain protein that is independent of disulfide bridge formation for binding activity (Martineau, Jones, and Winter 117-27). As such, scFv13R4 has been heterologously expressed in *Escherichia coli*, *Saccharomyces cerevisiae*, and Chinese hamster ovary cells (CHO) (Visintin *et al.* 11723-28;Grage and Rehm 254-62;Bach *et al.* 79-93).

6.6.3. RESULTS

[00110] *In silico* protein sequence analysis of the β -hexosyl transferase (BHT). The carboxyl terminal portion (amino acids 111 to 594) of the BHT polypeptide (594 amino

acids) has noticeably homology to β -glucosidases. This glycohydrolase I (GHI) domain contains the putative catalytic acid/base, catalytic nucleophile, and three asparagine residues potentially required for protein N-glycosylation (Fig. 7A). Remarkably, the BHT amino terminus revealed a unique region that spans the first 110 amino acids. This region comprises an amino terminal classical leader signal domain (amino acids 1 to 22) followed by a predicted (<http://www.cbs.dtu.dk/services/SecretomeP>) non-classical secretion signal (NC) of low complexity (amino acids 72 to 83) (<http://mendel.imp.ac.at/METHODS/seg.server.html>). The amino terminal leader signal (amino acids 1 to 22) can be further broken down into the N-region (amino terminal; amino acids 1 to 5), H-region (hydrophobic; amino acids 6 to 17) and C-region (carboxyl terminal; amino acids 18 to 22) (Fig. 7B). Alternative algorithms such as the Phobius web-based program (<http://phobius.sbc.su.se/>) and the SignalP algorithm (<http://www.cbs.dtu.dk/services/SignalP/>) also predicted the putative classical leader signal and potential cleavage sites between residues 17 and 18 and between 22 and 23. Furthermore, the classical leader signal was predicted to contain five amino acids that contact the membrane within the H-region (RHYTHM, <http://proteininformatics.charite.de/rhythm/>) and a charge distribution that could facilitate localizing and secretion of membrane proteins (Boyd and Beckwith 1031-33). The trans-membrane region prediction algorithm (http://www.ch.embnet.org/software/TMPRED_form.html) also forecasted a stretch of hydrophobic residues from 1-17 and 177-199 in BHT typical for integral membrane spanning proteins and non-cytoplasmic region (amino acids 23-594) as depicted in the Kyte and Doolittle hydropathy plot (Fig. 7C). These structural features may also indicate that the amino terminal classical leader signal may act as a membrane anchor during passage through the yeast secretory pathway.

[00111] Amino terminal domains participate in protein secretion. To investigate the probable physiological roles of the two amino terminal secretion domains (classical and non-classical secretion signals), expression of the carboxyl terminal BHT domain (amino acids 111 to 594) and the single chain anti- β -galactosidase antibody (scFv13R4) were tested. Stable recombinant *P. pastoris* GS115 strains were obtained by chromosomal integration of the appropriate modified gene combinations preceded by the *rBht* amino terminal domains and/or the strong 9.3 kDa α MF prepro sequence. The *rBht* and *scFv13R4* gene combinations were inserted downstream of the AOX1 promoter and followed by carboxyl-terminal 6XHIS tag to assist detection and purification Fig. 8A and 8D (see Materials and Methods).

[00112] Replacement of the leader signal (amino acids 1 to 22) with the strong α MF prepro secretion signal (GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-HIS) increased protein secretion to more than 19 fold (9.80 $\mu\text{g.ml}^{-1}$) compared to expression of the full-length rBHT-HIS preceded by α MF secretion signal (GS115:: α MF-*rBht*-HIS) (0.49 $\mu\text{g.ml}^{-1}$) (Fig. 8B). Similarly, in the absence of the α MF, the leader signal was able to direct the heterologous protein for secretion (GS115::*rBht*-HIS) (6.35 $\mu\text{g.ml}^{-1}$). Additionally, protein secretion was detected in the absence of both α MF and leader signals (GS115::*rBht*₍₂₃₋₅₉₄₎-HIS) (4.65 $\mu\text{g.ml}^{-1}$), suggesting that both the classical leader and the non-classical secretion signals contain information targeting the protein for secretion.

[00113] To validate that the amino-terminal domains, as described above, target proteins to the secretory pathway, we choose the antibody scFv13R4, an intracellular protein depleted of signal sequences. Diagrams of the antibody *scFv13R4* chimeras are shown in Figure (8C). The scFv13R4-HIS when expressed by GS115::*scFv13R4*-HIS (lacking leader secretion signals) could not be detected in the culture broth by SDS-PAGE silver staining or Western blot analysis (data not shown). Secretion of scFv13R4-HIS by GS115::*rBht*₍₁₋₁₁₀₎-*scFv13R4*-HIS or GS115::*rBht*₍₂₃₋₁₁₀₎-*scFv13R4*-HIS manifested when either *scFv13R4* was fused to the *Bht* classical leader secretion signal (25.17 $\mu\text{g.ml}^{-1}$), or when fused to the *Bht* non-classical signal (7.03 $\mu\text{g.ml}^{-1}$). Likewise, as seen with BHT-HIS, secretion driven by α MF, (GS115:: α MF-*scFv13R4*-HIS) provided the highest level of secreted protein (91.02 $\mu\text{g.ml}^{-1}$).

[00114] Enzyme activity and Western blot analysis of rBHT-HIS expressed by *P. pastoris* GS115. To confirm that protein expression correlated with enzymatic activity, the secreted and the membrane bound rBHT-HIS activities were measured using ONP-Glu as the substrate (see Materials and Methods). All recombinant strains secreted rBHT-HIS in detectable amounts and the values of activity reflected increases in secreted protein. The protein secreted by GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-HIS displayed an enzymatic activity of 3.7 mU.OD⁻¹ that was 6-fold higher than the measured activity when secretion was driven by the complete amino terminal region (amino acids 1-110) (GS115::*rBht*-HIS (0.63 mU.OD⁻¹)). Similarly, the measured enzymatic activity of the protein secreted by GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-HIS was 53-fold higher than obtained from the recombinant containing both the α MF and the leader signals (GS115:: α MF-*rBht*-HIS (0.07 mU.OD⁻¹)). The recombinant GS115::*rBht*₍₂₃₋₅₉₄₎-HIS (0.26 mU.OD⁻¹) show a reduced amount of active secreted enzyme Table 8.

[00115] The activity of the membrane-bound enzyme displayed by resting cells of each recombinant was also tested. We found values of activity that correlate with total secreted protein showing an increase in membrane bound activity of 15-fold for the strain GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*-HIS (21.52 mU.OD⁻¹) and 1.3-fold increase for the strain GS115::*rBht*-HIS (1.94 mU.OD⁻¹) compared to GS115:: *α MF-rBht*-HIS (1.48 mU.OD⁻¹). The recombinant GS115::*rBht₍₂₃₋₅₉₄₎*-HIS (0.15 mU.OD⁻¹) show a reduced amount of membrane bound enzyme, confirming that this recombinant redirects the protein through the putative non-classical secretion pathway. Overall these results show that neither α MF nor the BHT leader secretion signal could fully complete the secretion of rBHT-HIS which may be related to the presence of a transmembrane region predicted between amino acids 177 to 199.

[00116] Western blot analysis of the cell-free extracts using anti-HIS antibody confirmed the secreted protein values by GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*-HIS, GS115::*rBht*-HIS, and GS115::*rBht₍₂₃₋₅₉₄₎*-HIS (Fig. 9A). In each case the prominent rBHT-HIS band corresponding to a molecular mass of approximately 110 kDa was present. These results are in agreement with previously reported SDS-PAGE and size exclusion chromatography migration patterns (Dagher, Azcarate-Peril, and Bruno-Bárcena). Western blot analysis of cell extracts obtained from GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*-HIS, GS115::*rBht*-HIS, and GS115:: *α MF-rBht*-HIS also exhibited a molecular mass of approximately 110 kDa, while GS115::*rBht₍₂₃₋₅₉₄₎*-HIS showed prominent bands between 98 and 64 kDa that may indicate intracellular degradation or alternative glycosylation patterns (Fig. 9B).

[00117] rBHT hydrolytic activity. Additionally, we tested the secreted enzymes for both HIS-tagged and non-HIS tagged protein from GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*-HIS and GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*, respectively. The enzymes delivered comparable results and were active in a wide range of temperatures (10 to 50°C) and at pH values (2.8 to 6). Maximum activity was observed from pH 3.6 to 5 (91 to 100% of maximum activity) followed by a steady decrease down to pH 2.6 (43% of maximum) and up to pH 6.8 (29% of maximum). Likewise, the optimum temperature was found in the range of 40 and 45°C (97 to 100% maximum activity) but rapidly decreased at temperatures above 50°C and below 20°C (less than 25% of maximum) (data not shown). The enzyme was stable in 50 mM sodium phosphate buffer pH 5 at 4°C for at least 6 months and the activity was unaffected by storage at -80°C. The values for the kinetic constants for the enzyme secreted by GS115:: *α MF-rBht₍₂₃₋₅₉₄₎*-HIS were obtained from the Hill equation (Km 0.79 mM and Vmax 3.97 mmol.min⁻¹ per mg⁻¹ of enzyme at 42°C pH 4). Those findings were in agreement with

previous reports by us and others (Blakely and Mackenzi 1021-25; Cho, Shin, and Bucke 2107-11; Gorin, Phaff, and Spencer 1341-44; Gorin, Spencer, and Phaff 2307-17; Ishikawa *et al.* 331-39; Sakai *et al.* 285-93; Shin, Park, and Yang 787-92; Shin and Yang 484-89; Dagher, Azcarate-Peril, and Bruno-Bárcena).

[00118] rBHT stability. To examine the long-term stability of the enzyme, all the freshly induced recombinant strains were incubated in buffer containing 2% glucose and the hydrolytic activity of membrane bound and secreted rBHT was measured over time. Secreted rBHT-HIS obtained from all recombinants through classical or non-classical secretion pathway remained stable over the one-week testing period and retained more than 95% of initial activity. The same stability was observed when resting cells containing membrane-associated enzyme was expressed by GS115::*αMF-rBht*-HIS, GS115::*rBht*-HIS, and GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS. However, when testing resting cells containing membrane-associated enzyme expressed by GS115::*rBht*₍₂₃₋₅₉₄₎-HIS, the activity began to decrease within 24 h pointing to the alternative non-classical secretion pathway.

[00119] Purification and characterization of rBHT-HIS generated by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS. The rBHT protein expressed by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS was purified using nickel affinity chromatography. The placement of the 6XHIS tag on the carboxyl-terminus successfully allowed the single step recovery of more than 73% of the original enzymatic activity and after SDS-PAGE a single polypeptide band of approximately 110 kDa was seen (Fig. 9C). The 6.54 fold protein purification from the culture supernatant recovered 7.24 mg of enzyme rendering a specific activity of 18.45 mU.mg⁻¹ at 42°C and pH 4 (Table 9). Moreover, following the same methodology we purified the rBHT-HIS secreted by the different recombinants and found comparable specific activities ranging from 18.45 to 18.65 mU.mg⁻¹. A determination of the amino-terminal sequences of the secreted polypeptide by GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS showed that the entire rBHT₍₂₃₋₅₉₄₎-HIS protein was present in the broth (V-X-Y-P-G residues) in addition to a product containing two additional amino-terminal amino acids (E-A-V-X-Y residues). Variability in the cleavage of amino acids A-E during secretion can be affected by the surrounding amino acid sequence and the tertiary structure (Cereghino and Cregg 45-66). The remaining non-classical sequence did not introduce a new cleavage site.

[00120] HIS tag impact on rBHT transferase activity. Recombinants GS115::*αMF-rBht*₍₂₃₋₅₉₄₎-HIS and GS115::*αMF-rBht*₍₂₃₋₅₉₄₎ were further employed to comparatively evaluate whether the presence of the HIS tag may impact GOS synthesis from lactose. GOS

accumulation was analyzed quantitatively by HPLC from reaction mixtures containing 220 gL⁻¹ initial lactose, 0.5 U rBHT g⁻¹ lactose and incubated at 30°C.

[00121] Fig. 10A shows comparative GOS accumulation and lactose consumption over time when the reaction was catalyzed by either the 6XHIS tagged or non-tagged secreted enzymes. In both cases the maximum rate of production was observed during the first 25 h with galactosyl-lactose as the main product. Confirming previously described enzymatic competitive glucose inhibition after 125 h, galactosyl-lactose (75 gL⁻¹) accumulation was stationary reaching an average of 67% conversion from the 60% initial lactose utilized (Dagher, Azcarate-Peril, and Bruno-Bárcena).

[00122] When using resting cells expressing membrane-associated HIS and non-HIS tagged rBHT, comparative GOS accumulation and lactose consumption over time was also confirmed. (Fig. 10B) shows that the presence of carboxyl-terminal HIS had no impact on the initial reaction rate of galactosyl-lactose formation (1.87 and 1.7 g.L⁻¹.h⁻¹). As previously reported, the glucose was consumed by resting cells of *P. pastoris* while the galactose was used to synthesize GOS (68% yield (g/g)) approaching the theoretical yield of 75% (Dagher, Azcarate-Peril, and Bruno-Bárcena).

6.6.4. DISCUSSION

[00123] Prebiotics are carbohydrate derivatives marketed as functional foods and actively promoted to improve consumer health that is intended to specifically stimulate the growth of beneficial bacteria in the gut. The fundamental force that drives development of prebiotics is the promise of more efficient production processes at lower operating costs. However, production or synthesis of specific carbohydrate derivatives by chemical methods is complex and requires protection and deprotection steps due to the presence of several hydroxyl groups of similar reactivity (Sears and Wong 2344-50). Therefore, the development of enzymatic approaches is of practical interest and genetic modification has been extensively used to modify enzymatic activity, to obtain a deeper knowledge of catalytic mechanisms, and to increase protein secretion. Proteins destined for secretion are usually preceded by amino-terminal leader signals of 20–30 amino acids and eventually processed by membrane bound signal peptidases (Von Heijne 17-21). Protein secretion by *P. pastoris* is influenced by the nature of the initial nucleotide sequences and occasionally requires codon optimization, as well as consideration of glycosylation patterns, final 3-dimensional structure, culture conditions, and medium composition (Damasceno, Huang, and Batt 31-39). Additionally, the

distribution of charged amino acids within leader domains plays an important role in facilitating the localization of membrane and secreted proteins (Boyd and Beckwith 1031-33).

[00124] As reported previously, *P. pastoris* offers advantages over *E. coli* for the expression of rBHT thanks to its ability to efficiently incorporate post-translational modifications that allowed for the heterologous production of small amounts of bioactive rBHT. *In silico* analyses of BHT suggested that this enzyme contains trans-membrane domains that needed to be studied to increase secretion of the enzyme by *P. pastoris*. The BHT unique protein region (1-110 amino acids) contains two domains predicted to function as classical leader signal (BHT₍₁₋₂₂₎) and non-classical secretion signal domains (BHT₍₂₃₋₁₁₀₎). The classical leader signal also targets the protein to perform its function at the cellular membrane (as predicted by the RHYTHM method and hydropathy plots, Fig. 1). In particular, the presence of basic amino acids such as arginine at position 17 could be implicated in secretion efficiency and protein orientation in the membrane (Fig 1).

[00125] The data presented here shows protein secretion interference by GS115:: α MF-*rBht-HIS* due to the simultaneous presence of α MF and leader signal (BHT₍₁₋₂₂₎). The levels of protein secretion were comparable to the previously reported values by GS115:: α MF-*HIS-TEV-rBht* (Dagher, Azcarate-Peril, and Bruno-Bárcena). On the other hand, higher accumulation of secreted protein was obtained by the recombinants GS115::*rBht-HIS* and GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-*HIS* lacking either α MF or BHT₍₁₋₂₂₎, respectively. Therefore, demonstrating that the leader signal domain (BHT₍₁₋₂₂₎) is involved in the signal peptide-mediated mechanism (classical secretory pathway) comparable to α MF. Both leader signals were individually able to increase protein expression of membrane-associated and secreted rBHT compared to GS115:: α MF-*rBht-HIS*, containing both leader sequences. The best values for secreted (50-fold increase) and membrane-bound (14-fold increase) bioactive rBHT protein were obtained by the recombinant GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-*HIS* (Table 3). The subsequent purification of the bioactive protein expressed from GS115:: α MF-*rBht*₍₂₃₋₅₉₄₎-*HIS* resulted in very pure protein by SDS-PAGE with a specific activity of 18.45 U.mg⁻¹ (Table 4). The molecular mass of rBHT (110 kDa) did not deviate between cell membrane-associated and secreted rBHT and displayed similar enzyme activity, thermostability, reusability and storage stability compared with rBHT from our previous study (Dagher *et al.*, 2013).

[00126] We expected that removal of both the leader domains (BHT₍₁₋₂₂₎) and α MF would hinder enzyme secretion. However, elimination of both α MF and BHT₍₁₋₂₂₎ still showed low amounts of secreted protein. Also confirming that the 110 amino acid unique region contains a dual function by which the leader domain BHT₍₁₋₂₂₎ acts as an efficient secretion signal (classical secretion pathway) and the predicted BHT₍₂₃₋₁₁₀₎ domain may operate as alternative secretion signal (non-classical secretion pathway). Western blot analyses show proteolysis indicating increased protein sensitivity in the cell. The majority of the measured hydrolytic activity was detected as multiple bands below the maximal mass of 110 kDa, likely affecting the amount of secreted enzyme (Fig. 9A-9C). We speculate that the signal sequences may act to protect the protein during secretion by maintaining the protein away from proteases within the secretory pathway.

[00127] Although greater amounts of protein were secreted by GS115:: α MF-rBht₍₂₃₋₅₉₄₎-HIS, significant quantities remained stably bound to the membrane, possibly due to the predicted limited mobility in the membrane by the presence of the trans-membrane domain within the protein carboxy-terminus domain (amino acids 177-199). Therefore, to confirm the physiological function of these domains as secretory signals we generated new protein chimeras by replacing the rBHT₍₂₃₋₅₉₄₎ domain with the antibody scFv13R4 protein. The BHT₍₁₋₁₁₀₎ classical leader followed by putative non-classical leader, the BHT₍₂₃₋₁₁₀₎ putative non-classical leader, and the α MF domain were placed in frame at the amino terminal position with scFv13R4. Analyses of these new recombinants were able to corroborate the leader secretory function by directing the antibody to secretion. Our results confirm that the choice of signal sequences has a strong impact on both production and secretion levels of proteins including the recombinant BHT and scFv13R4. Noteworthy is the fact that the new smaller leader signal (22 amino acids) has a size advantage compared to α MF (66 amino acids) and has been demonstrated here to be a new unique sequence able to direct secretion of heterologous proteins. This leader signal domain adds a new feature that can be built into intracellular enzymes that otherwise need to be extracted by disruption using mechanical means or permeabilization with chemical treatments (Panesar *et al.* 530-43).

[00128] Continued molecular development of BHT will help address food industries problems for enzymes with novel properties such as thermo-activity, cold stability and synthesis of specific oligosaccharides. The present findings motivate further structural analysis to elucidate features that contribute to transglycosylation activity and substrate specificity. Mutagenesis of catalytic sites and rational mutagenesis based on the 3D structure

will pave the way for alterations in substrate specificity for production of novel GOS as prebiotic candidates.

Table 7. Strains and plasmids used in this study

<i>Strains/Plasmids</i>	<i>^aDescription or genotype</i>	<i>Source or Reference</i>
Strains		
<i>E. coli</i>		
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI^qΔM15 Tn10</i> (Tet ^R)]	Stratagene
<i>P. pastoris</i>		
GS115	<i>his4</i> (<i>his</i> ⁻ <i>mut</i> ⁺)	Invitrogen
JB210	GS115:: <i>αMF-rBht</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB212	GS115:: <i>αMF-rBht</i> ₍₂₃₋₅₉₄₎ -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB213	GS115:: <i>αMF-rBht</i> ₍₂₃₋₅₉₄₎ (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB214	GS115:: <i>rBht</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB215	GS115:: <i>rBht</i> ₍₂₃₋₅₉₄₎ -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB217	GS115:: <i>αMF-scFv13R4</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB220	GS115:: <i>rBht</i> ₍₁₋₁₁₀₎ - <i>scFv13R4</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB221	GS115:: <i>rBht</i> ₍₂₃₋₁₁₀₎ - <i>scFv13R4</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
JB222	GS115:: <i>scFv13R4</i> -HIS (<i>his</i> ⁺ <i>mut</i> ⁺)	This study
Plasmids		
<i>E. coli</i>		
pJB100	pGS21a- <i>rBht</i>	Dagher, 2013
pPM163R4	pPM160 containing the anti-β-galactosidase antibody gene <i>scFv13R4</i>	Martineau, 1998
<i>P. pastoris</i>		
pPIC9	<i>P. pastoris</i> integrative vector carrying <i>AOX1</i> promoter and transcription terminator, <i>HIS4</i> , Amp ^r in <i>E. coli</i> , pBR322 ori, α-mating factor secretion signal from <i>S. cerevisiae</i> (<i>αMF</i>)	Invitrogen
pJB110	pPIC9- <i>αMF-rBht</i> -HIS	This study
pJB112	pPIC9- <i>αMF-rBht</i> ₍₂₃₋₅₉₄₎ -HIS	This study
pJB113	pPIC9- <i>αMF-rBht</i> ₍₂₃₋₅₉₄₎	This study
pJB114	pPIC9- <i>rBht</i> -HIS	This study
pJB115	pPIC9- <i>rBht</i> ₍₂₃₋₅₉₄₎ -HIS	This study
pJB117	pPIC9- <i>αMF-scFv13R4</i> -HIS	This study
pJB120	pPIC9- <i>rBht</i> ₍₁₋₁₁₀₎ - <i>scFv13R4</i> -HIS	This study
pJB121	pPIC9- <i>rBht</i> ₍₂₃₋₁₁₀₎ - <i>scFv13R4</i> -HIS	This study
pJB122	pPIC9- <i>scFv13R4</i> -HIS	This study

^a*αMF*, *S. cerevisiae* α-mating factor secretion signal found in pPIC9 vector.

TABLE 8. Secreted and membrane bound rBHT-HIS enzyme activity by different

recombinants of *P. pastoris*

Enzyme Source	Mean activity (mU.OD ⁻¹) ± SD ^a		
	Secreted	Membrane bound	Ratio Secreted/Membrane Bound
GS115::αMF- <i>rBht</i> ₍₂₃₋₅₉₄₎ -HIS	3.70 ± 0.063	21.52 ± 1.38	0.172
GS115:: <i>rBht</i> -HIS	0.63 ± 0.018	1.94 ± 0.02	0.325
GS115:: <i>rBht</i> ₍₂₃₋₅₉₄₎ -HIS	0.26 ± 0.003	0.15 ± 0.02	1.606
GS115::αMF- <i>rBht</i> -HIS	0.07 ± 0.003	1.48 ± 0.02	0.046

TABLE 9. PURIFICATION OF RBHT-HIS SECRETED BY *P. PASTORIS* GS115::αMF-RBHT₍₂₃₋₅₉₄₎-HIS

Enzyme Source	Total activity in media (UL ⁻¹) ^a	Total protein in media (mg) ^b	Specific activity media (U.mg ⁻¹) ^c	Ni column n (U) ^d	Ni column n (mg) ^e	Specific activity Ni column (U.mg ⁻¹) ^f	Purification (fold) ^g	Recovery (%) ^h
GS115::αMF- <i>rBht</i> ₍₂₃₋₅₉₄₎ -HIS	180.67	64.00	2.82	133.54	7.24	18.45	6.54	73.91

^a, 1 Liter of culture was grown in BMGY broth at 28°C.

^b, Protein concentration determined by Bradford assay.

^c, Specific activity is expressed as the total activity (U) divided by the total protein (mg).

^d, Total units following nickel chromatography purification.

^e, Total protein (mg) following nickel chromatography purification.

^f, Specific activity expressed as total activity (U) divided by total yield (mg) following nickel chromatography purification.

^g, Increase in specific activity.

^h, Yield expressed as total activity following nickel column chromatography divided by total activity in the broth.

6.7. References for Section 6.6

Bach, Horacio, *et al.* "Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies." *Journal of Molecular Biology* 312.1 (2001): 79-93.

Boyd, Dana and Jon Beckwith. "The role of charged amino acids in the localization of secreted and membrane proteins." *Cell* 62.6 (1990): 1031-33.

Cereghino, J. L. and J. M. Cregg. "Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*." *FEMS Microbiol.Rev.* 24.1 (2000): 45-66.

- Damasceno, Leonardo, Chung Jr Huang, and Carl Batt. "Protein secretion in *Pichia pastoris* and advances in protein production." Applied Microbiology and Biotechnology 93.1 (2012): 31-39.
- Grage, Katrin and Bernd H. A. Rehm. "In Vivo Production of scFv-Displaying Biopolymer Beads Using a Self-Assembly-Promoting Fusion Partner." Bioconjugate Chemistry 19.1 (2007): 254-62.
- Kurjan, Janet and Ira Herskowitz. "Structure of a yeast pheromone gene (MFa): A putative α -factor precursor contains four tandem copies of mature α -factor." Cell 30.3 (1982): 933-43.
- Martineau, Pierre, Peter Jones, and Greg Winter. "Expression of an antibody fragment at high levels in the bacterial cytoplasm." Journal of Molecular Biology 280.1 (1998): 117-27.
- Panesar, Parmjit S., *et al.* "Microbial production, immobilization and applications of β -D-galactosidase." Journal of Chemical Technology & Biotechnology 81.4 (2006): 530-43.
- Sears, P. and C. H. Wong. "Toward automated synthesis of oligosaccharides and glycoproteins." Science 291.5512 (2001): 2344-50.
- Spencer, J. F. T., A. L. R. de Spencer, and C. Lalue. "Non-conventional yeasts." Appl.Microbiol.Biotechnol. 58.2 (2002): 147-56.
- Tzortzis, George and Jelena Vulevic. "Galacto-Oligosaccharide Prebiotics." Prebiotics and Probiotics Science and Technology. Ed. Dimitris Charalampopoulos and RobertA Rastall. Springer New York, 2009. 207-44.
- Visintin, Michela, *et al.* "Selection of antibodies for intracellular function using a two-hybrid *in vivo* system." Proceedings of the National Academy of Sciences 96.21 (1999): 11723-28.
- Von Heijne, Gunnar. "Patterns of Amino Acids near Signal-Sequence Cleavage Sites." European Journal of Biochemistry 133.1 (1983): 17-21.

7. REFERENCES

1. **Blakely, J. A. and S. L. Mackenzi.** 1969. Purification and properties of a b-hexosidase from *Sporobolomyces singularis*. *Can. J. Biochem.* **47**:1021-1025.
2. **Boehm, G., J. Jelinek, B. Stahl, K. van Laere, J. Knol, S. Fanaro, G. Moro, and V. Vigi.** 2004. Prebiotics in infant formulas. *J. Clin. Gastroenterol.* **38**:S76-S79.
3. **Boehm, G. and B. Stahl.** 2007. Oligosaccharides from milk. *J. Nutr.* **137**:847S-849S.
4. **Cho, Y. J., H. J. Shin, and C. Bucke.** 2003. Purification and biochemical properties of a galactooligosaccharide producing b-galactosidase from *Bullera singularis*. *Biotechnol. Lett.* **25**:2107-2111.

5. **Cregg, J. M., T. S. Vedvick, and W. C. Raschke.** 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. *BioTechnol.* **11**:905-910.
6. **Cuppa, G. V. and O. Gabrelli.** 2008. Human milk oligosaccharides as prebiotics, p. 131-146. *In* J. Versalovic and M. Byand Wilson (eds.), *Therapeutic microbiology. Probiotics and related strategies*. ASM Press Washington DC Chapter.
7. **Dagher, S. F., J. L. Wang, and R. J. Patterson.** 1995. Identification of Galectin-3 as a factor in pre-mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.* **92**:1213-1217.
8. **Gallagher, S. R.** 1995. *Current protocols in protein science*, p. 10.3.1-10.3.11.
9. **Gorin, P. A. J., H. J. Phaff, and J. F. T. Spencer.** 1964. Structures of Galactosyl-lactose and Galactobiosyl-lactose produced from lactose by *Sporobolomyces singularis*. *Can. J. Chem.* **42**:1341-1344.
10. **Gorin, P. A. J., J. F. T. Spencer, and H. J. Phaff.** 1964. Synthesis of b-galacto-b-glucopyranosyl disaccharides by *Sporobolomyces singularis*. *Can. J. Chem.* **42**:2307-2317.
11. **Gosling, A., G. W. Stevens, A. R. Barber, S. E. Kentish, and S. L. Gras.** 2010. Recent advances refining galactooligosaccharide production from lactose. *Food Chem.* **121**:307-318.
12. **Grosova', M. Rosenberg, and Rebroš.** 2008. Perspectives and applications of immobilised b-galactosidase in food industry - a review. *Czech J. Food Sci.* **26**:1-14.
13. **Hestrin, S., D. S. Feingold, and M. Schramm.** 1955. Hexoside hydrolases. *Meth. Enzymol.* **1**:231-257.
14. **Higgins, D. R.** 1998. Introduction to *Pichia pastoris*. *Methods Mol. Biol.* **103**:1-15.
15. **Hofmann.** 1993. TMbase - A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* **374**.
16. **Ishikawa, E., T. Sakai, H. Ikemura, K. Matsumoto, and H. Abe.** 2005. Identification, cloning, and characterization of a *Sporobolomyces singularis* b-galactosidase-like enzyme involved in galacto-oligosaccharide production. *J. Biosci. Bioeng.* **99**:331-339.
17. **Jenkins, N., R. B. Parekh, and D. C. James.** 1996. Getting the glycosylation right: Implications for the biotechnology industry. *Nat. Biotechnol.* **14**:975-981.
18. **Katayama, K.** 1984. Inhibition by copper ion of the activities of b-galactosidase and dehydrogenases of activated sludge. *Japan J. Wat. Pollut. Res.* **7**:100-107.
19. **Katayama, K.** 1986. Inhibition of the activities of b-galactosidase and dehydrogenases of activated sludge by heavy metals. *Water Res.* **20**:591-594.
20. **Kim, C. S., E. S. Ji, and D. K. Oh.** 2004. A new kinetic model of recombinant b-galactosidase from *Kluyveromyces lactis* for both hydrolysis and transgalactosylation reactions. *Biochem. Biophys. Res. Comm.* **316**:738-743.

21. **Knol, J., P. Scholtens, C. Kafka, J. Steenbakkers, S. Groß, K. Helm, M. Klarczyk, H. Schopfer, H. M. Bockler, and J. Wells.** 2005. Colon microflora in infants fed formula with galacto- and fructo-oligosaccharides: More like breast-fed infants. *J. Pediatr. Gastroenterol. Nutr.* **40**:36-42.
22. **Kuby, S. A. and H. A. Lardy.** 1953. Purification and kinetics of b-D-galactosidase from *Escherichia coli* strain-K-12. *J. Am. Chem. Soc.* **75**:890-896.
23. **Laemmli, U. K.** 1970. Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nat.* **227**:680-685.
24. **Li, P. Z., A. Anumanthan, X. G. Gao, K. Ilangoan, V. V. Suzara, N. Duzgunes, and V. Renugopalakrishnan.** 2007. Expression of recombinant proteins in *Pichia pastoris*. *Appl. Biochem. Biotechnol.* **142**:105-124.
25. **Neri, D. F. M., V. M. Balcão, R. S. Costa, I. C. A. P. Rocha, E. M. F. C. Ferreira, D. P. M. Torres, L. R. M. Rodrigues, L. B. Carvalho, and J. A. Teixeira.** 2009. Galacto-oligosaccharides production during lactose hydrolysis by free *Aspergillus oryzae* b-galactosidase and immobilized on magnetic polysiloxane-polyvinyl alcohol. *Food Chem.* **115**:92-99.
26. **Otieno, D. O.** 2010. Synthesis of b-galactooligosaccharides from lactose using microbial b-galactosidases. *Compr. Rev. Food Sci. Food Saf.* **9**:471-482.
27. **Park, A. R. and D. K. Oh.** 2010. Galacto-oligosaccharide production using microbial b-galactosidase: current state and perspectives. *Appl. Microbiol. Biotechnol.* **85**:1279-1286.
28. **Phaff, H. J. and L. D. Carmo-Sousa.** 1962. Four new species of yeast isolated from insect frass in bark of *Tsuga heterophylla* (Raf.) Sargent. *Antonie Van Leeuwenhoek.* **28**:193-207.
29. **Rastall, R. A. and V. Maitin.** 2002. Prebiotics and synbiotics: towards the next generation. *Curr. Opin. Biotechnol.* **13**:490-496.
30. **Roberfroid, M., G. R. Gibson, L. Hoyles, A. L. McCartney, R. Rastall, I. Rowland, D. Wolvers, B. Watzl, H. Szajewska, B. Stahl, F. Guarner, F. Respondek, K. Whelan, V. Coxam, M. J. Davicco, L. Leotoing, Y. Wittrant, N. M. Delzenne, P. D. Cani, A. M. Neyrinck, and A. Meheust.** 2010. Prebiotic effects: metabolic and health benefits. *Br. J. Nutr.* **104**:S1-S63.
31. **Sakai, T., H. Tsuji, S. Shibata, K. Hayakawa, and K. Matsumoto.** 2008. Repeated-batch production of galactooligosaccharides from lactose at high concentration by using alginate-immobilized cells of *Sporobolomyces singularis* YIT 10047. *J. Gen. Appl. Microbiol.* **54**:285-293.
32. **Sambrook J.** 2001. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.
33. **Scopes, R.** 1995. The effect of temperature on enzymes used in diagnostics. *Clin. Chim. Acta* **237**:17-23.
34. **Shin, H. J., J. M. Park, and J. W. Yang.** 1998. Continuous production of galacto-oligosaccharides from lactose by *Bullera singularis* b-galactosidase immobilized in chitosan beads. *Process Biochem.* **33**:787-792.

35. **Shin, H. J. and J. W. Yang.** 1998. Enzymatic production of galactooligosaccharide by *Bullera singularis* b-galactosidase. J. Microbiol. Biotechnol. **8**:484-489.
36. **Temple, L. M., A. A. Weiss, K. E. Walker, H. J. Barnes, V. L. Christensen, D. M. Miyamoto, C. B. Shelton, and P. E. Orndorff.** 1998. *Bordetella avium* virulence measured *in vivo* and *in vitro*. Infect. Immun. **66**:5244-5251.
37. **Torres, D. P. M., M. D. F. Gonçalves, J. A. Teixeira, and L. R. Rodrigues.** 2010. Galactooligosaccharides: Production, properties, applications, and significance as prebiotics. Compr. Rev. Food Sci. Food Saf. **9**:438-454.

[00129] It is to be understood that, while the invention has been described in conjunction with the detailed description, thereof, the foregoing description is intended to illustrate and not limit the scope of the invention. Other aspects, advantages, and modifications of the invention are within the scope of the claims set forth below. All publications, patents, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

8. SEQUENCE LISTING

SEQ ID NO. 1

```
>gi|345649663|gb|JF298281.1| Synthetic construct beta-hexosyl transferase
(bglA) gene, partial cds
ATGATGCTGCAIGCTGCACTGCTAGTAGCGCTGCCATGTGTTGTTTTGGCGCGCCCGGCCGGAGCGGTTA
CTTATCCGGGAGCCATTCTCTGTCCCTGACGAGCAATTACGAAACCCCAAGTCCGACAGCAATCCCGCT
GGAGCCAACACCGACGGCTACCGGTACAGCAGAATTAGATGCGCTGTGGAACCTAGTCGAAGCTCAGTAC
CCAGTTCAAACCTGCTGCAGTGACAACCTTTGGTGACAGTGCCCGATGATTATAAGTTTGAGGCAGATCCAC
CGAGTTATGCATTAGCAGGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAGGGGTTTAAGTT
TGGTGTGTCGGGGGCGAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGCCCAAGTACCTGG
GATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAACAATTATGATCCCGATATTACAACCAACCATT
ACTACCTGTACCCATTGGACTTTGCGCGCCTGCAACACCTAGGCATTAACTTACTCGTTTTCAATTTT
ATGGACGCGTATTTATCCATTGGGCGCAGGCTATGTTAATGAAGCAGGGTTAGCCCACTATGATGCCGTA
ATCCATAGTGCCAAGAAGTATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTC
TGATGCTGAAATACGGTGCCCTGGCAAGATACTGGTGACCAAAATTGTTAAGGACTTTGTTACCTATGCCAC
AACTGTGTTTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCACGGGTTTTCTGT
TCACAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATTAACAGCACCTCCGCTGTATTTT
GTTGCACCTACAATGTTCTGAAAGCTCATGGTCATGCTGTTAAAGTGTATCGGGATCTAGTTGCCTCCGG
GACCATTGCGGCAGGTGAAATCGGCTTTAAATCCGATGATAACTACCCAATCCCGGCCCGTCCAGGGAAC
GCCGATGACGAGGAATCAGCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCGGTTTATG
GTAATGGCGATTATCCAGATGTTGTTAAAGAAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAAGA
TAAAGGATACATTAAAGGTAGCGGAGATATTTTGGCATTGACGGGTATCGTACCGATATTTCCCATGCG
GCTCTGAACGGGATCGGAATTGTATTTCGCAACCAAGTGACCCGAATTGGCCAGTGTGTGAAGAAGGGT
CAGATCCTTTTGCTCATGTTTACCCATCCGGGTTTGCTATTGGTCAATCAGCCGATCCACTGTCTTCATG
GTTAGTCAACTCAGCCCCGTTTATCCGCGATCAACTGAAGTTTCTGACACAAACCTACCCTGCTAAGGGT
GGTATTTATTTCTCGGAATTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATTA
CCTGGGATGGTCTGCGTACGCAATACCTGACGGACTATCTGAGCCAGCTGCTGTTGGCTGTGCACAAAGA
CGGGATTAATCTGCGAGGCGCGCTGACGTGGAGTTTGTGCGATAATTGGGAGTGGGGTTTAGGGATGCAA
CAGAAATTCGGATTTTCAGTTTGTTAATCAATCAGATCCCGATCTGACACGCACGTTTAAACTGAGCGCTC
```

ACGCTTACGCCCAATTTGGGCGTAATCATCTG

SEQ ID NO. 2

>gi|345649664|gb|AE014215.1| beta-hexosyl transferase [synthetic construct]
 MMLHAALLVALPCVVLARPAGAVTYPGAIPLSLTSNYETPSPTAIPLEPTPTATGTAEALDNLWNLVEAQY
 PVQTAAVTTTLVTPDDYKFEADPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGAAKAEGRGPSTW
 DYLCCHYASTQCNNYDPDITTNHYLYPLDFARLQHLGINTYSFISISWTRIYPLGAGYVNEAGLAHYDAV
 IHSKKGYLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTVFVKRYGNEVKTWFTFNEPRVFC
 SQNSGLPYNLTYPEGINSTSAVFRCTYNVLKAHGHAVKVYRDLVASGTIAAGEIGFKSDDNYPIPARPGN
 ADDEESAKRHEAFRIGFAQPVYGNPDYDVVKETVGDMPLALTDKGYIKGSGDIFAIIDGYRDISHA
 ALNGIANCIRNQSDPNWVPCVEGSDPFVHVPYSGFAIGQSADPLSSWLVSAPFIRDQLKFLTQTYPAKG
 GIYFSEFGWAEDAEDYRQLLYQITWDGLRTQYLTIDYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGMO
 QKFGFQFVNQSDPDLTRTFKLSAHAYAQFGRNHL

SEQ ID NO. 3

FP#1 aMF-6XHIS-TEV(Q/M)-rBHT (XhoI-NotI)

ATGAGATTTTCCTTCAATTTTTACTGCGATTTTATTCGCAGCATCCTCCGCATTAGCTGCT
 CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATCCGGCTGAAGCTGTCATCGGT
 TACTCAGATTTAGAAGGGGATTTTCATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
 TCTCTCGAGAAAAGAGAGGCTGAAGCT CACCACCACCACCACCAGAAACCTGTATTTT
CAGATGATGCTGCATGCTGCACCTGCTAGTAGCGCTGCCATGTGTTGTTTTGGCGCGCCCG
GCCGGAGCGGTTACTTATCCGGGAGCCATTCTCTGTCCCTGACGAGCAATTACGAAACC
CCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCGACGGCTACCGGTACAGCAGAATTA
GATGCGCTGTGGAACCTAGTCGAAGCTCAGTACCCAGTTCAAACCTGCTGCAGTGACAAC
TTGGTGACAGTCCCCGATGATTATAAGTTTGAGGCAGATCCACCGAGTTATGCATTAGCA
GGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAGGGGTTTAAAGTTTGGTGTT
GCGGGGGCAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGCCCAAGTACC
TGGGATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAACAATTATGATCCCGATATT
ACAACCAACCATTACTACCTGTACCCATTGGACTTTGCGCGCTGCAACACCTAGGCATT
AACACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCATTGGGCGCAGGCTATGTT
AATGAAGCAGGGTTAGCCCACTATGATGCCGTAATCCATAGTGCCAAGAAGTATGGTCTG
GAACCAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTCTGATGCTGAAATACGGT
GCCTGGCAAGATACTGGTGACCAATTTGTTAAGGACTTTGTTACCTATGCCACAACCTGTG
TTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCACGGGTTTTT
TGTTACAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATTAACAGCACC
TCCGCTGTATTTTCGTTGCACCTACAATGTTCTGAAAGCTCATGGTCATGCTGTTAAAGTG
TATCGGGATCTAGTTGCCTCCGGGACCATTTGCGGCAGGTGAAATCGGCTTTAAATCCGAT
GATAACTACCAATCCCGGCCCGTCCAGGGAACGCCGATGACGAGGAATCAGCCAAGCGT
CACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCGGTTTATGGTAATGGCGATTATCCA
GATGTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAAGATAAAAGGA
TACATTAAAGGTAGCGGAGATATTTTTGCGATTGACGGGTATCGTACCGATATTTCCCAT
GCGGCTCTGAACGGGATCGCGAATTGTATTGCAACCAAGTGACCCGAATTGGCCAGTG
TGTGAAGAAGGGTCAGATCCTTTTGCTCATGTTTACCCATCCGGGTTTGCTATTGGTCAA
TCAGCCGATCCACTGTCTTCATGGTTAGTCAACTCAGCCCCGTTTATCCGCGATCAACTG
AAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGTATTTATTTCTCGGAATTTGGTTGG
GCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATTACCTGGGATGGTCTGCGT
ACGCAATACCTGACGGACTATCTGAGCCAGCTGCTGTTGGCTGTGCACAAAGACGGGATT
AATCTGCGAGGCGCGCTGACGTGGAGTTTTTGTCGATAATTGGGAGTGGGGTTTAGGGATG
CAACAGAAATTCGGATTTTCAAGTTTGTAAATCAATCAGATCCCGATCTGACACGCACGTTT
AAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGTAATCATCTGTAA

SEQ ID NO. 4

FP#1 aMF-6XHIS-TEV(Q/M)-rBHT

MRFPSIFTAVLFAASSALAAPVNTTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNSTN
 NGLLFINTTIIASIAAKEEGVSLEKREAEA HHHHHHENLYFQMMLHAALLVALPCVVLARP
AGAVTYPGAIPLSLTSNYETPSPTAIPLEPTPTATGTAELDALWNLVEAQYPVQTAAVTT
LVTVPDDYKFEADPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGAAKAEGRGPST
WDYLCHHYASTQCNNYDPDITTNHYLYPLDFARLQHLGINTYSFSISWTRIYPLGAGYV
NEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTV
EKRYGNEVKTWFTFNEPRVFCSQNSGLPYNLTYPEGINSTSAVFRCTYNVLKAHGHAVKV
YRDLVASGTIAAGEIGFKSDDNYPPIPARPGNADDEESAKRHEAFRIGIFAQPVYNGNDYP
DVVKETVGDMLPALTDEDKGYIKSGDIFAIDGYRTDISHAALNGIANCIRNQSDPNWPV
CEEGSDPFAHVYPSPGFAIGQSADPLSSWLVSAPFIRDQLKFLTQTYPAKGGIYFSEFGW
AEDAEDRQLLYQITWDGLRTQYLTDYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGM
QQKFGFQFVNQSDPDLTRTFKLSAHAYAQFGRNHL

SEQ ID NO. 5

FP#2 aMF-6XHIS-TEV(Q/M)-aMF-rBHT-6XHIS (XhoI-NotI)

ATGAGATTTCCCTCAATTTTTACTGCGAGTTTATTCGCGCATCCTCCGCATTAGCTGCT
 CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTCATCGGT
 TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
 TCTCTCGAGAAAAGAGAGGCTGAAGCT CACCACCACCACCACCACGAAAACCTGTATTTT
CAGATGATGCTGCATGCTGCACTGCTAGTAGCGCTGCCATGTGTTGTTTTGGCGCGCCCG
GCCGGAGCGGTTACTTATCCGGGAGCCATTCTCTGTCCCTGACGAGCAATTACGAAACC
CCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCGACGGCTACCGGTACAGCAGAATTA
GATGCGCTGTGGAACCTAGTCGAAGCTCAGTACCCAGTTCAAACCTGCTGCAGTGACAAC
TTGGTGACAGTGCCCGATGATTATAAGTTTGAGGCAGATCCACCGAGTTATGCATTAGCA
GGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAGGGGTTTAAAGTTTGGTGTT
GCGGGGGCAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGCCAAGTACC
TGGGATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAACAATTATGATCCCGATATT
ACAACCAACCAATTACTACCTGTACCCATTGGACTTTGCGCGCCTGCAACACCTAGGCATT
AACACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCATTGGGCGCAGGCTATGTT
AATGAAGCAGGGTAGCCCACTATGATGCCGTAATCCATAGTGCCAAGAAGTATGGTCTG
GAACCAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTCTGATGCTGAAATACGGT
GCCTGGCAAGATACTGGTGACCAATTTGTTAAGGACTTTGTTACCTATGCCACAACCTGTG
TTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCACGGGTTTTCT
TGTTACAAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATTAACAGCACC
TCCGCTGTATTTCGTTGCACCTACAATGTTCTGAAAGCTCATGGTCATGCTGTTAAAGTG
TATCGGGATCTAGTTGCCTCCGGGACCATTGCGGCAGGTGAAATCGGCTTTAAATCCGAT
GATAACTACCCAATCCCGGCCCGTCCAGGGAACGCCGATGACGAGGAATCAGCCAAGCGT
CACGAGGCTTTTTCGATTGGGATTTTTGCGCAACCGGTTTATGGTAATGGCGATTATCCA
GATGTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAAGATAAAGGA
TACATTAAAGGTAGCGGAGATATTTTTGCGATTGACGGGTATCGTACCGATATTTCCCAT
GCGGCTCTGAACGGGATCGCGAATTGTATTCGCAACCAAAGTGACCCGAATTGGCCAGTG
TGTGAAGAAGGGTCAGATCCTTTTGTCTATGTTTACCCATCCGGGTTTGCTATTGGTCAA
TCAGCCGATCCACTGCTTTCATGGTTAGTCAACTCAGCCCCGTTTATCCGCGATCAACTG
AAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGTATTTATTTCTCGGAATTTGGTTGG
GCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATTACCTGGGATGGTCTGCGT
ACGCAATACCTGACGGAATCTGAGCCAGCTGCTGTTGGCTGTGCACAAAGACGGGATT
AATCTGCGAGGCGCGCTGACGTGGAGTTTTGTCGATAATTGGGAGTGGGGTTTAGGGATG
CAACAGAAATTCGGATTTTCAGTTTGTTAATCAATCAGATCCCGATCTGACACGCACGTTT
AAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGTAATCATCTGCACCACCACCACCAC
CACTAA

SEQ ID NO. 6

FP#2 aMF-6XHIS-TEV(Q/M)-aMF-rBHT-6XHIS

MRFPSIFTAVLFAASSALAAPVNTTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNSTN

NGLLFINTTIIASIAAKEEGVSLEKREAEA HHHHHHENLYFQMMLHAALLVALPCVVLARP
AGAVTYPGAIPLSLTSNYETPSPTAIPLEPTPTATGTAELDALWNLVEAQYPVQTAAVTT
LVTVPDDYKFEADPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGAAKAEGRGPST
WDYLCHHYASTQCNNYDPDITTNHYLYPLDFARLQHLGINTYSFSISWTRIYPLGAGYV
NEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTV
FKRYGNEVKTWFTFNEPRVFCQNSGLPYNLTYPEGINSTSAVFRCTYNVLKAHGHAVKV
YRDLVASGTIAAGEIGFKSDDNYPIPARPGNADDEESAKRHEAFRIGIFAQPVYGNPDYP
DVVKETVGDMPLALTDDEKGYIKSGDIFAIDGYRTDISHAALNGIANCIRNQSDPNWPV
CEEGSDPFAHVYPSGFAIGQSADPLSSWLVSAPFIRDQLKFLTQTYPAKGGIYFSEFGW
AEDAEDRQLLYQITWDGLRTQYLTDYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGM
QOKFGFQFVNQSDPDLTTRTFKLSAHAYAQFGRNHLHHHHHH

SEQ ID NO. 7

FP#3 aMF-rBHT-6XHIS (XhoI-NotI)

ATGAGATTTCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGTA
TCTCTCGAGAAAAGAGAGGCTGAAGCT ATGATGCTGCATGCTGCACTGCTAGTAGCGCTG
CCATGTGTTGTTTTGGCGCGCCCGCGGAGCGGTTACTTATCCGGGAGCCATTCTCTGTG
TCCCTGACGAGCAATTACGAAACCCCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCG
ACGGCTACCGGTACAGCAGAATTAGATGCGCTGTGGAACCTTAGTCGAAGCTCAGTACCCA
GTTCAAACCTGCTGCAGTGACAACCTTTGGTGACAGTGCCCGATGATTATAAGTTTGAGGCA
GATCCACCGAGTTATGCATTAGCAGGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTT
CCAAAGGGGTTAAGTTTGGTGTTCGGGGGCGAGCCATTCAAGTTGAAGGTGCAGCAAAA
GCCGAAGGGCGGGGCCCAAGTACCTGGGATTATCTGTGTCACTACTATGCCAGCAGCAG
TGTAACAATTATGATCCCGATATTACAACCAACCATTACTACCTGTACCCATTGGACTTT
GCGCGCCTGCAACACCTAGGCATTAACACTTACTCGTTTTCAATTTTCATGGACGCGTATT
TATCCATTGGGCGCAGGCTATGTTAATGAAGCAGGGTTAGCCCACTATGATGCCGTAATC
CATAGTGCCAAAGAAGTATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCA
CTGTCTCTGATGCTGAAATACGGTGCCTGGCAAGATACTGGTGACCAAAATTGTTAAGGAC
TTTGTTACCTATGCCACAACCTGTGTTTAAAGCGTTATGGTAATGAAGTCAAGACGTGGTTT
ACTTTCAATGAACCACGGGTTTTCTGTTTCAAAAATAGTGGTCTGCCATACAATCTGACG
TATCCAGAAGGTATTAACAGCACCTCCGCTGTATTTTCGTTGCACCTACAATGTTCTGAAA
GCTCATGGTCACTGCTGTTAAAGTGTATCGGGATCTAGTTGCCTCCGGGACCAATTGCGGCA
GGTGAATTCGGCTTTAAATCCGATGATAACTACCCAATCCCGGCCCGTCCAGGGAACGCC
GATGACGAGGAATCAGCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCG
GTTTATGGTAATGGCGATTATCCAGATGTTGTTAAAGAACTGTTGGAGATATGCTGCCG
GCCCTGACGGATGAAGATAAAGGATACATTAAAGGTAGCGGAGATATTTTTGCGATTGAC
GGGTATCGTACCGATATTTCCCATGCGGCTCTGAACGGGATCGCGAATTGTATTCGCAAC
CAAAGTGACCCGAATTGGCCAGTGTGTGAAGAAGGGTCAGATCCTTTTGCATGTTTAC
CCATCCGGGTTTGCTATTGGTCAATCAGCCGATCCACTGTCTTCATGGTTAGTCAACTCA
GCCCCGTTTATCCGCGATCAACTGAAGTTTCTGACACAAACCTACCTGCTAAGGGTGGT
ATTTATTTTCTCGGAATTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTAT
CAAAATTACCTGGGATGGTCTGCGTACGCAATACCTGACGGAATCTGAGCCAGCTGCTG
TTGGCTGTGCACAAAGACGGGATTAATCTGCGAGGCGCGCTGACGTGGAGTTTTGTGAT
AATTGGGAGTGGGGTTTAGGGATGCAACAGAAATTCGGATTTTCAAGTTTGTAAATCAATCA
GATCCCGATCTGACACGCACGTTTAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGT
AATCATCTGCACCACCACCACCACCCTAA

SEQ ID NO. 8

FP#3 aMF-rBHT-6XHIS

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSdleGDFDVAVLFPFSNSTN
NGLLFINTTIIASIAAKEEGVSLEKREAEA MMLHAALLVALPCVVLARPAGAVTYPGAIPL
SLTSNYETPSPTAIPLEPTPTATGTAELDALWNLVEAQYPVQTAAVTTT
LVTVPDDYKFEADPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGAAKAEGRGPSTWDYLCHHYASTQ

CNNYDPDITTNHYLYPLDFARLQHLGINTYSFISISWTRIYPLGAGYVNEAGLAHYDAVI
HS AKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTVFKRYGNEVKTW
TFNEPRVFC SQNSGLPYNLTYPEGINSTSAVFRTYNVLKAHGHAVKVYRDLVASGTIAA
GEIGFKSDDNYP IPARPGNADDEESAKRHEAFRIGIFAQPVYNGDYPDVVKETVGDMLP
ALTDDEKGYIKSGDIF AIDGYRTDISHAALNGIANCIRNQSDPNWPVCEE SDFFAHVY
PSGFAIGQSADPLSSWL VNSAPF IRDQLKFLTQTYPAKGGIYFSEFGWAEDA EYDRQLLY
QITWDGLRTQYLT DYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGMQQKFQGFVNQS
DPDLTRTFKLSAHAYA QFGRNHLHHHHHH

SEQ ID NO. 9

FP#4 aMF-rBHT (XhoI-NotI)

ATGAGATTTCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
TCTCTCGAGAAAAGAGAGGCTGAAGCT ATGATGCTGCATGCTGCACTGCTAGTAGCGCTG
CCATGTGTTGTTTTGGCGCGCCCGCGGAGCGGTTACTTATCCGGGAGCCATTCTCTG
TCCCTGACGAGCAATTACGAAACCCCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCG
ACGGCTACCGGTACAGCAGAATTAGATGCGCTGTGGAACCTAGTCGAAGCTCAGTACCCA
GTTCAAACCTGCTGCAGTGACAACCTTTGGTGACAGTGCCCGATGATTATAAGTTTGAGGCA
GATCCACCGAGTTATGCATTAGCAGGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTT
CCAAAGGGGTTTAAGTTTGGTGTGCGGGGCGAGCCATTCAAGTTGAAGGTGCAGCAAAA
GCCGAAGGGCGGGGCCCAAGTACCTGGGATTATCTGTGTATCATCTATGCCAGCACGCAG
TGTAACAATTATGATCCCGATATTACAACCAACCATTACTACCTGTACCCATTGGACTTT
GCGCGCTGCAACACCTAGGCATTAACACTTACTCGTTTTCAATTTTCATGGACGCGTATT
TATCCATTGGGCGCAGGCTATGTTAATGAAGCAGGGTAGCCCACTATGATGCCGTAATC
CATAGTGCCCAAGAAGTATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCA
CTGTCTCTGATGCTGAAATACGGTGCCTGGCAAGATACTGGTGACCAAATTTGTTAAGGAC
TTTGTTACCTATGCCACAACCTGTGTTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTT
ACTTTCAATGAACCACGGGTTTTCTGTTTACAAAATAGTGGTCTGCCATACAATCTGACG
TATCCAGAAGGTATTAACAGCACCTCCGCTGTATTCGTTGCACCTACAATGTTCTGAAA
GCTCATGGTCATGCTGTTAAAGTGTATCGGGATCTAGTTGCCTCCGGGACCATTCGCGCA
GGTGAATCGGCTTTAAATCCGATGATAACTACCCAATCCCGGCCGTCAGGGAACGCC
GATGACGAGGAATCAGCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCG
GTTTTATGGTAATGGCGATTATCCAGATGTTGTAAAGAAACGTTGGAGATATGCTGCCG
GCCCTGACGGATGAAGATAAAGGATACATTAAAGGTAGCGGAGATATTTTTGCGATTGAC
GGGTATCGTACCGATATTTCCCATGCGGCTCTGAACGGGATCGCGAATTGTATTGCGAAC
CAAAGTGACCCGAATTGGCCAGTGTGTGAAGAAGGGTCAGATCCTTTTGCTCATGTTTAC
CCATCCGGGTTTGCTATTGGTCAATCAGCCGATCCACTGCTTTCATGGTTAGTCAACTCA
GCCCCGTTTATCCGCGATCAACTGAAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGT
ATTTATTTCTCGGAATTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTAT
CAAATTACCTGGGATGGTCTGCGTACGCAATACCTGACGGACTATCTGAGCCAGCTGCTG
TTGGCTGTGCACAAAGACGGGATTAATCTGCGAGGCGCGCTGACGTGGAGTTTTGTGCGAT
AATTGGGAGTGGGTTTAGGGATGCAACAGAAATTCGGATTTTCAGTTTGTAAATCAATCA
GATCCCGATCTGACACGCACGTTTAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGT
AATCATCTGTAA

SEQ ID NO. 10

FP#4 aMF-rBHT

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYS DLEGDFDVAVL PFSNSTN
NGLLFINTTIIASIAAKEEGVSLEKREAE MMLHAALLVALPCVVLARPAGAVTYPGA IPL
SLTSNYETPSPTAIPLEPTPTATGTAELDALWNLVEAQYPVQTAAVTTLVTPDDYKFEA
DPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGA AKA EGRGPSTWDYLCHHYASTQ
CNNYDPDITTNHYLYPLDFARLQHLGINTYSFISISWTRIYPLGAGYVNEAGLAHYDAVI
HS AKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTVFKRYGNEVKTW
TFNEPRVFC SQNSGLPYNLTYPEGINSTSAVFRTYNVLKAHGHAVKVYRDLVASGTIAA

GEIGFKSDDNYP IPARPGNADDEESAKRHEAFRIGIFAQPVYGNPDYDPVVKETVGDMLP
ALTDKGYIKGSGDIFAIDGYRTDISHAALNGIANCIRNQSDPNWPVCEEGSDPFAHVY
PSGFAIGQSADPLSSWLVSAPFIRDQLKFLTQTYPAKGGIYFSEFGWAEDAHEYDRQLLY
QITWDGLRTQYLT DYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGMQQKFQFVNQS
DPDLTRTFKLSAHAYAQFGRNHL

SEQ ID NO. 11

FP#5 aMF-rBHT ($\Delta 1-22$)-6XHIS (XhoI-NotI)

ATGAGATTTCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCAATCGGT
TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
TCTCTCGAGAAAAGAGAGGCTGAAGCT GTTACTTATCCGGGAGCCATTCCCTCTGTCCCTG
ACGAGCAATTACGAAACCCCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCGACGGCT
ACCGGTACAGCAGAATTAGATGCGCTGTGGAAGCTTAGTCGAAGCTCAGTACCCAGTTCAA
ACTGCTGCAGTGACAACCTTTGGTGACAGTGCCGATGATTATAAGTTTGAGGCAGATCCA
CCGAGTTATGCATTAGCAGGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAG
GGGTTTTAAGTTTGGTGTTGCGGGGCGAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAA
GGGCGGGGCCCCAAGTACCTGGGATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAAC
AATTATGATCCCGATATTACAACCAACCATTACTACCTGTACCCATTGGACTTTGCGCGC
CTGCAACACCTAGGCATTAACACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCA
TTGGGCGCAGGCTATGTTAATGAAGCAGGGTTAGCCCACTATGATGCCGTAATCCATAGT
GCCAAGAAGTATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCT
CTGATGCTGAAATACGGTGCTTGGCAAGATACTGGTGACCAAATTGTTAAGGACTTTGTT
ACCTATGCCACAACCTGTGTTTAAAGCGTTATGGTAATGAAGTCAAGACGTGGTTACTTTC
AATGAACCACGGGTTTTCTGTTCACAAAATAGTGGTCTGCCATACAATCTGACGTATCCA
GAAGGTATTAAACAGCACCTCCGCTGTATTTCTGTTGCACCTACAATGTTCTGAAAGCTCAT
GGTCATGCTGTTAAAGTGTATCGGGATCTAGTTGCTCCGGGACCATTGCGGCAGGTGAA
ATCGGCTTTAAATCCGATGATAACTACCCAATCCCGGCCCGTCCAGGGAACGCCGATGAC
GAGGAATCAGCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCGGTTTAT
GGTAATGGCGATTATCCAGATGTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTG
ACGGATGAAGATAAAAGGATACATTAAAGGTAGCGGAGATATTTTTGCGATTGACGGGTAT
CGTACCGATATTTCCCATGCGGCTCTGAACGGGATCGCGAATTGTATTTCGCAACCAAAGT
GACCCGAATTGGCCAGTGTGTGAAGAAGGGTCAGATCCTTTTGCTCATGTTTACCCATCC
GGGTTTGCTATTTGGTCAATCAGCCGATCCACTGTCTTCATGGTTAGTCAACTCAGCCCCG
TTTATCCGCGATCAACTGAAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGTATTTAT
TTCTCGGAATTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATT
ACCTGGGATGGTCTGCGTACGCAATACCTGACGGACTATCTGAGCCAGCTGCTGTGGCT
GTGCACAAAGACGGGATTAATCTGCGAGGCGCGCTGACGTGGAGTTTTGTCGATAATTGG
GAGTGGGGTTTAGGGATGCAACAGAAATTCGGATTTCAGTTTTGTTAATCAATCAGATCCC
GATCTGACACGCACGTTTAAACTGAGCGCTCACGCTTACGCCAATTTGGGCGTAATCAT
CTGCACCACCACCACCACCCTAA

SEQ ID NO. 12

FP#5 aMF-rBHT ($\Delta 1-22$)-6XHIS

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLFPFSNSTN
NGLLFINTTIIASIAAKEEGVSLEKREAEA VTYPGAIPLSLTSNYETPSPTAIPLEPTPTA
TGTAELDALWNLVEAQYPVQTAAVTTTLVTVPDDYKFEADPPSYALAGYETSEIAGLKFPK
GFKFGVAGAAIQVEGAAKAEGRPSTWDYLCHHYASTQCNNYDPDITTNHYLYPLDFAR
LQHLGINTYSFSISWTRIYPLGAGYVNEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLS
LMLKYGAWQDTGDQIVKDFVTYATTTFKRYGNEVKTWFTFNEPRVFCSONSGLPYNLTYP
EGINSTSAVFRCTYNVLKAHGHAVKVYRDLVASGTIAAGEIGFKSDDNYP IPARPGNADD
EESAKRHEAFRIGIFAQPVYGNPDYDPVVKETVGDMLPALTDKGYIKGSGDIFAIDGY
RTDISHAALNGIANCIRNQSDPNWPVCEEGSDPFAHVYPSGFAIGQSADPLSSWLVSAP
FIRDQLKFLTQTYPAKGGIYFSEFGWAEDAHEYDRQLLYQITWDGLRTQYLT DYLSQLLLA
VHKDGINLRGALTWSFVDNWEWGLGMQQKFQFVNQSDPDLTRTFKLSAHAYAQFGRNH

LHHHHHH

SEQ ID NO. 13

FP#6 aMF-rBHT ($\Delta 1-22$) (XhoI-NotI)

ATGAGATTTCTTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
 CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCGGCTGAAGCTGTTCATCGGT
 TACTCAGATTTAGAAGGGGATTTTCGATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
 TCTCTCGAGAAAAGAGAGGCTGAAGCT GTTACTTATCCGGGAGCCATTCTCTGTCCCTG
ACGAGCAATTACGAAACCCCAAGTCCGACAGCAATCCCGCTGGAGCCAACACCGACGGCT
ACCGGTACAGCAGAATTAGATGCGCTGTGGAAGCTTAGTCAAGCTCAGTACCCAGTTCAA
ACTGCTGCAGTGACAACCTTTGGTGACAGTGCCCGATGATTATAAGTTTGAGGCAGATCCA
CCGAGTTATGCATTAGCAGGGTATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAG
GGGTTTTAAGTTTGGTGTTCGGGGGCGAGCCATCAAGTTGAAGGTGCAGCAAAAGCCGAA
GGGCGGGGCCCAAGTACCTGGGATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAAC
AATTATGATCCCGATATTACAACCAACCATTACTACCTGTACCCATTGGACTTTGCGCGC
CTGCAACACCTAGGCATTAACACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCA
TTGGGCGCAGGCTATGTTAATGAAGCAGGGTAGCCCACTATGATGCCGTAATCCATAGT
GCCAAGAAGTATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCT
CTGATGCTGAAATACGGTGCTGGCAAGATACTGGTGACCAAAATTGTTAAGGACTTTGTT
ACCTATGCCACAACCTGTGTTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTC
AATGAACCACGGGTTTTCTGTTTACAAAATAGTGGTCTGCCATACAATCTGACGTATCCA
GAAGGTATTAACAGCACCTCCGCTGTATTTCTGTTGCACCTACAATGTTCTGAAAGCTCAT
GGTCATGCTGTTAAAGTGTATCGGGATCTAGTTGCCCTCCGGGACCATTGCGGCAGGTGAA
ATCGGCTTTAAATCCGATGATAACTACCCAATCCCGGCCCGTCCAGGGAACGCCGATGAC
GAGGAATCAGCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTGCGCAACCGGTTTTAT
GGTAATGGCGATTATCCAGATGTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTG
ACGGATGAAGATAAAGGATACATTAAAGGTAGCGGAGATATTTTTGCGATTGACGGGTAT
CGTACCGATATTTCCCATGCGGCTCTGAACGGGATCGCGAATTGTATTGCAACCAAAAGT
GACCCGAATTGGCCAGTGTGTGAAGAAGGGTCAGATCCTTTTGCTCATGTTTATCCCATCC
GGGTTTGCTATTGGTCAATCAGCCGATCCACTGTCTTCATGGTTAGTCAACTCAGCCCCG
TTTATCCGCGATCAACTGAAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGTATTTAT
TTCTCGGAATTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATT
ACCTGGGATGGTCTGCGTACGCAATACCTGACGGACTATCTGAGCCAGCTGCTGTGGCT
GTGCACAAAGACGGGATTAACTCTGCGAGGCGCGCTGACGTGGAGTTTTGTCGATAATTGG
GAGTGGGTTTAGGGATGCAACAGAAATTCGATTTCAGTTTGTTAATCAATCAGATCCC
GATCTGACACGCAGTTTAAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGTAATCAT
CTGTAA

SEQ ID NO. 14

FP#6 aMF-rBHT ($\Delta 1-22$)

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLFPNSNSTN
 NGLLFINTTIIASIAAKEEGVSLEKREAEA VTYPGAIPLSLTSNYETPSPTAIPLEPTPTA
TGTAELDALWNLVEAQYPVQTAAVTTLVTPDDYKFADPPSYALAGYETSEIAGLKFPK
GFKFGVAGAAIQVEGAAKAEGRGPSTWDYLCCHYASTQCNNYDPDITTNHYLYPLDFAR
LQHLGINTYSFISISWTRIYPLGAGYVNEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLS
LMLKYGAWQDQIVKDFVTYATTVFVKRYGNEVKTWFTFNEPRVFCSQNSGLPYNLTYP
EGINSTSAVFRCTYNVLKAHGHAVKVYRDLVASGTIAAGEIGFKSDDNYPIPARPGNADD
EESAKRHEAFRIGIFAQPVYGNPDYDVVKETVGDMPLALTDKGYIKSGDIFAIDGY
RTDISHAALNGIANCIRNQSDPNWPVCEECSDFAHVYPSGFAIGQSADPLSSWLVSAP
FIRDQLKFLTQTPAKGGIYFSEFGWAEDA EYDRQLLYQITWDGLRTQYLT DYLSQLLLA
VHKDGINLRGALTWSFVDNWEWGLGMQKQFGFQFVNQSDPDLTRTFKLSAHAYAQFGRNH
L

Seq. ID No. 15

FP#7 rBHT-6XHIS (BamHI-NotI)

ATGATGCTGCATGCTGCACTGCTAGTAGCGCTGCCATGTGTTGTTTTGGCGCGCCCGGCC
 GGAGCGGTTACTTATCCGGGAGCCATTCTCTGTCCCTGACGAGCAATTACGAAACCCCA
 AGTCCGACAGCAATCCCGCTGGAGCCAACACCGACGGCTACCGGTACAGCAGAATTAGAT
 GCGCTGTGGAACCTAGTCGAAGCTCAGTACCCAGTTCAAACCTGCTGCAGTGACAACCTTTG
 GTGACAGTGCCCGATGATTATAAGTTTGAGGCAGATCCACCGAGTTATGCATTAGCAGGG
 TATGAAACAAGCGAGATTGCCGGACTGAAGTTTCCAAAGGGGTTTAAGTTTGGTGTGCG
 GGGGCAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGCCCAAGTACCTGG
 GATTATCTGTGTCATCACTATGCCAGCACGCAGTGTAACAATTATGATCCCGATATTACA
 ACCAACCATTACTACCTGTACCCATTGGACTTTGCGCGCCTGCAACACCTAGGCATTAACT
 ACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCATTGGGCGCAGGCTATGTTAAT
 GAAGCAGGGTTAGCCCACTATGATGCCGTAATCCATAGTGCCAAGAAGTATGGTCTGGAA
 CCAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTCTGATGCTGAAATACGGTGCC
 TGGCAAGATACTGGTGACCAAATTGTTAAGGACTTTGTTACCTATGCCACAACCTGTGTTT
 AAGCGTTATGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCACGGGTTTTCTGT
 TCACAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATTAACAGCACCTCC
 GCTGTATTTCTGTTGCACCTACAATGTTCTGAAAGCTCATGGTCATGCTGTTAAAGTGTAT
 CGGGATCTAGTTGCCCTCCGGGACCATTGCGGCAGGTGAAATCGGCTTTAAATCCGATGAT
 AACTACCCAATCCCGGCCCGTCCAGGGAACGCCGATGACGAGGAATCAGCCAAGCGTCAC
 GAGGCTTTTTTCGATTGGGATTTTTGCGCAACCGGTTTATGGTAATGGCGATTATCCAGAT
 GTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAAGATAAAGGATAC
 ATTAAAGGTAGCGGAGATATTTTTGCGATTGACGGGTATCGTACCGATATTTCCCATGCG
 GCTCTGAACGGGATCGCGAATTGTATTTCGCAACCAAAGTGACCCGAATTGGCCAGTGTGT
 GAAGAAGGGTCAGATCCTTTTGTCTATGTTTACCCATCCGGGTTTGTCTATTGGTCAATCA
 GCCGATCCACTGTCTTCATGGTTAGTCAACTCAGCCCCGTTTATCCGCGATCAACTGAAG
 TTTCTGACACAAACCTACCCTGCTAAGGTTGGTATTTATTTCTCGGAATTTGGTTGGGCT
 GAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATTACCTGGGATGGTCTGCGTACG
 CAATACCTGACGGACTATCTGAGCCAGCTGCTGTTGGCTGTGCACAAAGACGGGATTAAT
 CTGCGAGGCGCGCTGACGTGGAGTTTTGTGCGATAATTGGGAGTGGGGTTTAGGGATGCAA
 CAGAAATTCGGATTTTCAGTTTGTAAATCAATCAGATCCCGATCTGACACGCACGTTTAAAT
 CTGAGCGCTCAGCTTACGCCCAATTTGGGCGTAATCATCTGCACCACCACCACCACCAC
 TAA

SEQ ID NO. 16

FP#7 rBHT-6XHIS

MMLHAALLVALPCVVLARPAGAVTYPGAIPLSLTSNYETPSPTAIPLEPTPTATGTAELD
 ALWNLVEAQYPVQTAAVTTLVTVPDDYKFEADPPSYALAGYETSEIAGLKFPKGFKFGVA
 GAAIQVEGAAKAEGRGPSTWDYLCHHYASTQCNNYDPDITTNHYLYPLDFARLQHLGIN
 TYSFISISWTRIYPLGAGYVNEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLSLMLKYGA
 WQDTGDQIVKDFVTYATTVFVKRYGNEVKTWFTFNEPRVFCSONSGLPYNLTYPEGINSTS
 AVFRCTYNVLKAHGHAVKVYRDLVASGTIAAGEIGFKSDDNYPIPARPGNADDEESAKRH
 EAFRIGIFAQPVYNGDYPDVVKETVGDMLPALTDDEKGYIKSGDIFAIIDGYRTDISHA
 ALNGIANCIRNQSDPNWPVCEEGSDPFAHVYPSGFAIGQSADPLSSWLVSAPFIRDQLK
 FLTQTYPAKGGIYFSEFGWAEDAHEYDRQLLYQITWDGLRTQYLTIDYLSQLLLAVHKDGIN
 LRGALTWSFVDNWEWGLGMQKQKFGFQFVNQSDPDLTRTFKLSAHAYAQFGRNHLHHHHHH

SEQ ID NO. 17

FP#8 rBHT(Δ 1-22)-6XHIS (BamHI-NotI)

ATGGTTACTTATCCGGGAGCCATTCTCTGTCCCTGACGAGCAATTACGAAACCCCAAGT
 CCGACAGCAATCCCGCTGGAGCCAACACCGACGGCTACCGGTACAGCAGAATTAGATGCG
 CTGTGGAACCTAGTCGAAGCTCAGTACCCAGTTCAAACCTGCTGCAGTGACAACCTTTGGTG
 ACAGTGCCCGATGATTATAAGTTTGAGGCAGATCCACCGAGTTATGCATTAGCAGGGTAT

GAAACAAGCGAGATTGCCGGACTGAAGTTTCAAAGGGGTTTAAGTTTGGTGTGCGGGG
 GCAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGCCAAGTACCTGGGAT
 TATCTGTGTCACTATGCCAGCACGCAGTGTAAACAATTATGATCCCGATATTACAACC
 AACCATTACTACCTGTACCCATTGGACTTTGCGCGCCTGCAACACCTAGGCATTAACACT
 TACTCGTTTTCAATTCATGGACGCGTATTTATCCATTGGGCGCAGGCTATGTTAATGAA
 GCAGGGTTAGCCCACTATGATGCCGTAATCCATAGTGCCAAGAAGTATGGTCTGGAACCA
 GTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTCTGATGCTGAAATACGGTGCCTGG
 CAAGATACTGGTGACCAAATTGTTAAGGACTTTGTTACCTATGCCACAACCTGTGTTTAAG
 CGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCACGGGTTTTCTGTTCA
 CAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATTAACAGCACCTCCGCT
 GTATTTCTGTTGCACCTACAATGTTCTGAAAGCTCATGGTCAATGCTGTTAAAGTGTATCGG
 GATCTAGTTGCCCTCCGGGACCATTGCGGCAGGTGAAATCGGCTTTAAATCCGATGATAAC
 TACCAATCCCGGCCCGTCCAGGGAACGCCGATGACGAGGAATCAGCCAAGCGTCACGAG
 GCTTTTCGCAATTGGGATTTTTGCGCAACCGGTTTATGGTAATGGCGATTATCCAGATGTT
 GTTAAAGAAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAAGATAAAGGATACATT
 AAAGGTAGCGGAGATATTTTTGCGATTGACGGGTATCGTACCGATATTTCCCATGCGGCT
 CTGAACGGGATCGCGAATTGTATTGCAACCAAAGTGACCCGAATTGGCCAGTGTGTGAA
 GAAGGGTCAGATCCTTTTGCTCATGTTTACCCATCCGGGTTTGCTATTGGTCAATCAGCC
 GATCCACTGTCTCATGGTTAGTCAACTCAGCCCCGTTTATCCGCGATCAACTGAAGTTT
 CTGACACAAACCTACCTGCTAAGGGTGGTATTTATTTCTCGGAATTTGGTTGGGCTGAA
 GACGCCGAATATGATCGTCAACTGCTGTATCAAATTACCTGGGATGGTCTGCGTACGCAA
 TACCTGACGGACTATCTGAGCCAGCTGCTGTTGGCTGTGCACAAAGACGGGATTAATCTG
 CGAGGCGCGCTGACGTGGAGTTTTGTCGATAATTGGGAGTGGGGTTTAGGGATGCAACAG
 AAATTCGGATTTCAGTTTGTTAATCAATCAGATCCCGATCTGACACGCACGTTTAAACTG
 AGCGCTCACGCTTACGCCCAATTTGGGCGTAATCATCTGCACCACCACCACCACCTAA

SEQ ID NO. 18

FP#8 rBHT (Δ 1-22) -6XHIS

MVTPGAIPLSLTSNYETPSPTAIPLEPTPTATGTAELDALWNLVEAQYPVQTAAVTTLV
 TVPDDYKFADPPSYALAGYETSEIAGLKFPKGFKFGVAGAAIQVEGAAKAEGRGPSTWD
 YLCHHYASTQCNNYDPDITTNHYLYPLDFARLQHLGINTYSFISISWTRIYPLGAGYVNE
 AGLAHYDAVHSAKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYATTVFK
 RYGNVEKTFWTFNEPRVFCQNSGLPYNLTYPEGINSTSAVFRCTYNVLKAHGHAVKVYR
 DLVASGTIAAGEIGFKSDDNYPPIPARPGNADDEESAKRHEAFRIGIFAQPVYNGNDYDPDV
 VKETVGDMPLALTDDEDKGYIKGSGDIFAIDGYRTDISHAALNGIANCIRNQSDPNWVPCE
 EGSDPFAHVYPSGFAIGQSADPLSSWLVSAPFIRDQLKFLTQTYPAKGGIYFSEFGWAE
 DAEYDRQLLYQITWDGLRTQYLTLDYLSQLLLAVHKDGINLRGALTWSFVDNWEWGLGMQQ
 KFGFQFVNQSDPDLTRTFKLSAHAYAQFGRNHLHHHHHH

SEQ ID NO. 19

FP#9 aMF-rBHT (Δ 1-110) -6XHIS (EcoRI-NotI)

ATGAGATTTCCCTCAATTTTTACTGCAGTTTTATTTCGCAGCATCCTCCGCATTAGCTGCT
 CCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTCATCGGT
 TACTCAGATTAGAAAGGGGATTTTCATGTTGCTGTTTTGCCATTTTCCAACAGCACAAAT
 AACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTA
 TCTCTCGAGAAAAGAGAGGCTGAAGCTTACGTAGAATTC ATGTTTCAAAGGGGTTAAG
 TTTGGTGTTCGGGGGCAGCCATTCAAGTTGAAGGTGCAGCAAAAGCCGAAGGGCGGGGC
 CCAAGTACCTGGGATTATCTGTGTCATCACATATGCCAGCACGCAGTGTAAACAATTATGAT
 CCGATATTACAACCAACCATTACTACCTGTACCCATTGGACTTTGCGCGCCTGCAACAC
 CTAGGCATTAACACTTACTCGTTTTCAATTTTCATGGACGCGTATTTATCCATTGGGCGCA
 GGCTATGTTAATGAAGCAGGGTTAGCCCACTATGATGCCGTAATCCATAGTGCCAAGAAG
 TATGGTCTGGAACCAAGTGGGCACCGTTTTTCACTGGGATACGCCACTGTCTCTGATGCTG
 AAATACGGTGCTGGCAAGATACTGGTGACCAAATTGTTAAGGACTTTGTACCTATGCC
 ACAACTGTGTTAAGCGTTATGGTAATGAAGTCAAGACGTGGTTTACTTTCAATGAACCA
 CGGGTTTTCTGTTCAAAAATAGTGGTCTGCCATACAATCTGACGTATCCAGAAGGTATT
 AACAGCACCTCCGCTGTATTTCTGTTGCACCTACAATGTTCTGAAAGCTCATGGTCAATGCT

GTTAAAGTGTATCGGGATCTAGTTGCCTCCGGGACCATTGCGGCAGGTGAAATCGGCCTTT
AAATCCGATGATAACTACCCAATCCCGGCCCGTCCAGGGAACGCCGATGACGAGGAATCA
GCCAAGCGTCACGAGGCTTTTCGCATTGGGATTTTTCGCAACCGGTTTATGGTAATGGC
GATTATCCAGATGTTGTTAAAGAACTGTTGGAGATATGCTGCCGGCCCTGACGGATGAA
GATAAAGGATACATTAAAGGTAGCGGAGATATTTTTCGATTGACGGGTATCGTACCGAT
ATTTCCCATGCGGCTCTGAACGGGATCGCGAATTGTATTCGCAACCAAAGTGACCCGAAT
TGGCCAGTGTGTGAAGAAGGGTCAGATCCCTTTTGCTCATGTTTACCCATCCGGGTTTGCT
ATTGGTCAATCAGCCGATCCACTGTCTTCATGGTTAGTCAACTCAGCCCCGTTTATCCGC
GATCAACTGAAGTTTCTGACACAAACCTACCCTGCTAAGGGTGGTATTTATTTCTCGGAA
TTTGGTTGGGCTGAAGACGCCGAATATGATCGTCAACTGCTGTATCAAATTACCTGGGAT
GGTCTGCGTACGCAATACCTGACGGACTATCTGAGCCAGCTGCTGTTGGCTGTGCACAAA
GACGGGATTAATCTGCGAGGCGCGCTGACGTGGAGTTTTGTGATAATTGGGAGTGGGGT
TTAGGGATGCAACAGAAATTCCGATTTTCACTTTGTTAATCAATCAGATCCCGATCTGACA
CGCACGTTTAAACTGAGCGCTCACGCTTACGCCCAATTTGGGCGTAATCATCTGCACCAC
CACCACCACCACTAA

SEQ ID NO. 20

FP#9 aMF-rBHT ($\Delta 1-110$)-6XHIS

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPPFSNSTN
NGLLFINTTIIASIAAKEEGVSLEKREAEAYVEF MFPKGFKFGVAGAAIQVEGAAKAEGRG
PSTWDYLCCHYASTQCNNYDPDITTNHYLYPLDFARLQHLGINTYSFSISWTRIYPLGA
GYVNEAGLAHYDAVIHSAKKYGLEPVGTVFHWDTPLSLMLKYGAWQDTGDQIVKDFVTYA
TTVFKRYGNEVKTWFTFNEPRVFCSQNSGLPYNLTYPEGINSTSAVFRCTYNVLKAHGHA
VKVYRDLVASGTIAAGEIGFKSDDNYPIPARPGNADDEESAKRHEAFRIGIFAQPVYNGG
DYPDVKETVGDMLPALTDEDKGYIKSGDIFAIDGYRTDISHAALNGIANCIRNQSDPN
WPFVCEEGSDPFAHVYPGSGFAIGQSADPLSSWLVSAPFIRDQLKFLTQYPAKGGIYFSE
FGWAEDA EYDRQLLYQITWDGLRTQYLT DYLSQLLLAVHKDGINLRGALTWSFVDNWEWG
LGMQQKFQGFVNQSDPDLTRTFKLSAHAYAQFGRNHLHHHHHH

Claims

What is claimed is:

1. An isolated DNA coding for a recombinant β -hexosyl-transferase (rBHT) protein having the amino acid sequence set forth in SEQ ID NO. 12, 14, or 20, preferably having the amino acid sequence set forth in SEQ ID NO. 12 or 14.
2. The isolated DNA coding for the recombinant β -hexosyl-transferase (rBHT) protein of claim 1, wherein the DNA has at least 97% identity with the nucleic acid sequence set forth in SEQ ID NO. 11, 13, or 19, preferably having the nucleic acid sequence as set forth in SEQ ID NO. 11 or 13.
3. The isolated DNA coding for the recombinant β -hexosyl-transferase (rBHT) protein of claim 2, wherein the DNA has the nucleic acid sequence set forth in SEQ ID NO. 11, 13, or 19, preferably having the nucleic acid sequence as set forth in SEQ ID NO. 11 or 13.
4. An enzymatically active recombinant β -hexosyl-transferase (rBHT) protein wherein the protein has the amino acid sequence set forth in SEQ ID NO. 12, 14, or 20.
5. The enzymatically active rBHT protein of claim 4, wherein the protein is membrane bound or a soluble enzyme.
6. A method for producing enzymatically active recombinant β -hexosyl-transferase (rBHT) protein in a yeast host cell which comprises transforming the yeast host cell with a plasmid under the control of a suitable promotor wherein the plasmid contains an isolated DNA coding for a rBHT protein according to claim 1, 2, or 3.
7. The method of claim 6, wherein the suitable promotor is an alcohol oxidase promotor.
8. The method of claim 6, wherein the enzymatically active rBHT protein has a specific activity of about 8 U.mg⁻¹ at 20°C.
9. The method of claim 6, wherein the yeast host cell is a *Pichia pastoris* yeast cell.

10. The method of claim 6, further comprising using the enzymatically active recombinant β -hexosyl-transferase (rBHT) protein produced under suitable conditions to convert lactose to galacto-oligosaccharides (GOS).
11. The method of claim 10, wherein the enzymatically active rBHT protein is immobilized on a solid support.
12. The method of claim 10, wherein the suitable conditions are a batch or continuous stirred-tank reactor, a packed-bed reactor, or an ultrafiltration membrane reactor.
13. The method of claim 10, further comprising a generally recognized as safe (GRAS) organism as a glucose removal system to avoid competitive glucose inhibition, preferably the glucose removal system is used simultaneously with the enzymatically active rBHT protein.
14. The method of claim 6, further comprising using the recombinant β -hexosyl-transferase (rBHT) protein produced in a modified lactose-containing foodstuff or food byproduct, preferably the foodstuff is a dairy product or byproduct.
15. The method of claim 14, wherein the foodstuff or food byproduct is a baby dessert, a baby juice, a baby snack, a baby yoghurt drink, an energy drink, a fitness water or thirst quencher, a frozen dairy dessert, a fruit drink (vitamin/mineral fortified), a fruit pie filling, a fruit preparation, an infant formula, an infant meal replacement drink, a jelly jam, a meal replacement drink, a milk, a milk-based drink, a milk substitute, a syrup flavoring for milk, a yoghurt, or a whey.