PRODUCTION OF HIGH MANNOSE PROTEINS IN PLANT CULTURE

Field of the Invention

The present invention relates to transformed host cells for the production of high mannose proteins and a method and system for producing these proteins, particularly in plant culture.

Background of the Invention

Gaucher's disease is the most prevalent lysosomal storage disorder. It is caused by a recessive genetic disorder (chromosome 1 q21-q31) resulting in deficiency of glucocerebrosidase, also known as glucosylceramidase, which is a membrane-bound lysosomal enzyme that catalyzes the hydrolysis of the glycosphingolipid glucocerebroside (glucosylceramide, GlcCer) to glucose and ceramide. Gaucher disease is caused by point mutations in the hGCD (human glucocerebrosidase) gene (GBA), which result in accumulation of GlcCer in the lysosomes of macrophages. The characteristic storage cells, called Gaucher cells, are found in liver, spleen and bone marrow. The associated clinical symptoms include severe hepatosplenomegaly, anemia, thrombocytopenia and skeletal deterioration.

The gene encoding human GCD was first sequenced in 1985 (6) The protein consists of 497 amino acids derived from a 536-mer pro-peptide. The mature hGCD contains five N-glycosylation amino acid consensus sequences (Asn-X-Ser/Thr). Four of these sites are normally glycosylated. Glycosylation of the first site is essential for the production of active protein. Both high-mannose and complex oligosaccharide chains have been identified (7). hGCD from placenta contains 7% carbohydrate, 20% of which is of the high-mannose type (8). Biochemical and site-directed mutagenesis studies have provided an initial map of regions and residues important to folding, activator interaction, and active site location (9).

Treatment of placental hGCD with neuraminidase (yielding an asialo enzyme) results in increased clearance and uptake rates by rat liver cells with a concomitant increase in hepatic enzymatic activity (Furbish et al., 1981, Biochim. Biophys. Acta 673:425-434). This glycan-modified placental hGC is currently used as a therapeutic agent in the treatment of Gaucher's disease. Biochemical and site-directed mutagenesis

studies have provided an initial map of regions and residues important to folding, activator interaction, and active site location [Grace et al., J. Biol. Chem. 269:2283-2291 (1994)].

There are three different types of Gaucher disease, each determined by the level of hGC activity. The major cells affected by the disease are the macrophages, which are highly enlarged due to GlcCer accumulation, and are thus referred to as "Gaucher cells".

The identification of a defect in GCD as the primary cause of Gaucher's disease led to the development of enzyme replacement therapy as a therapeutic strategy for this disorder.

Another well characterized lysosomal storage disorder is Fabry disease. Fabry disease is an X-linked lysosomal storage disease that is caused by deficient activity of lysosomal enzyme α-galactosidase A (α-Gal A). Patients with classic Fabry disease typically have α-Gal A activity of less than 1% and often demonstrate the full spectrum of symptoms, including severe pain in the extremities (acroparesthesias), hypohidrosis, corneal and lenticular changes, skin lesions (angiokeratoma), renal failure, cardiovascular disease, pulmonary failure, neurological symptoms and stroke. In atypical Fabry disease, individuals with residual enzyme activity demonstrate symptoms later in life, and the symptoms are usually limited to one or a few organs. Clinical manifestations in female carriers vary greatly because of random X-chromosome inactivation. Although carriers commonly remain asymptomatic throughout life, many demonstrate clinical symptoms as variable and severe as those of affected males.

De Duve first suggested that replacement of the missing lysosomal enzyme with exogenous biologically active enzyme might be a viable approach to treatment of lysosomal storage diseases [Fed Proc. 23:1045 (1964)].

Since that time, various studies have suggested that enzyme replacement therapy may be beneficial for treating various lysosomal storage diseases. The best success has been shown with individuals with type I Gaucher disease, who were treated with exogenous enzyme (ß-glucocerebrosidase), prepared from placenta (Ceredase™) or, more recently, recombinantly (Cerezyme™).

Unmodified glucocerebrosidase derived from natural sources is a glycoprotein with four carbohydrate chains. This protein does not target the phagocytic cells in the body and is therefore of limited therapeutic value. In developing the current therapy for

Gaucher's disease, the terminal sugars on the carbohydrate chains of glucocerebrosidase are sequentially removed by treatment with three different glycosidases. This glycosidase treatment results in a glycoprotein whose terminal sugars consist of mannose residues. Since phagocytes have mannose receptors that recognize glycoproteins and glycopeptides with oligosaccharide chains that terminate in mannose residues, the carbohydrate remodeling of glucocerebrosidase has improved the targeting of the enzyme to these cells [Furbish *et al.*, Biochem. Biophys. Acta 673:425, (1981)].

As indicated herein, glycosylation plays a crucial role in hGCD activity, therefore deglycosylation of hGCD expressed in cell lines using either tunicamycin (*Sf9* cells) or point mutations abolishing all glycosylation sites (both *Sf9* and COS-1 cells), results in complete loss of enzymatic activity. In addition, hGCD expressed in *E. coli* was found to be inactive. Further research indicated the significance of the various glycosylation sites for protein activity. In addition to the role of glycosylation in the actual protein activity, the commercially produced enzyme contains glycan sequence modifications that facilitate specific drug delivery. The glycosylated proteins are remodeled following extraction to include only mannose containing glycan sequences.

The human GCD enzyme contains 4 glycosylation sites and 22 lysines. The recombinantly produced enzyme (Cerezyme^{Tm}) differs from the placental enzyme (Ceredase^{TM}) in position 495 where an arginine has been substituted with a histidine. Furthermore, the oligosaccharide composition differs between the recombinant and the placental GCD as the former has more fucose and N-acetyl-glucosamine residues while the latter retains one high mannose chain. As mentioned above, both types of GCDs are treated with three different glycosidases (neuraminidase, galactosidase, and P-N acetyl-glucosaminidase) to expose terminal mannoses, which enables targeting of phagocytic cells. A pharmaceutical preparation comprising the recombinantly produced enzyme is described in US 5,549,892. It should be noted that all references mentioned are hereby incorporated by reference as if fully set forth herein.

Recombinant α-Galactosidase A for enzyme replacement therapy has been produced in insect (sf9) cells (see US 7,011,831) in human fibroblasts (see US 6,395,884) and in plant cells (see US 6,846,968). Clinical trials with recombinant a-Gal A (agalsidase beta [Fabrazyme]: Genzyme Corporation, Cambridge, Mass; agalsidase

alfa [Replagal]: TKT Corporation, Cambridge, Mass) have been performed, and both drugs have been approved for clinical use.

One drawback associated with existing lysosomal enzyme replacement therapy treatment is that the *in vivo* bioactivity of the enzyme is undesirably low, e.g. because of low uptake, reduced targeting to lysosomes of the specific cells where the substrate is accumulated, and a short functional *in vivo* half-life in the lysosomes.

Another major drawback of the existing GCD recombinant enzymes is their expense, which can place a heavy economic burden on health care systems. The high cost of these recombinant enzymes results from a complex purification protocol, and the relatively large amounts of the therapeutic required for existing treatments. There is therefore, an urgent need to reduce the cost of GCD so that this life saving therapy can be provided to all who require it more affordably.

Proteins for pharmaceutical use have been traditionally produced in mammalian or bacterial expression systems. In the past decade a new expression system has been developed in plants. This methodology utilizes Agrobacterium, a bacteria capable of inserting single stranded DNA molecules (T-DNA) into the plant genome. Due to the relative simplicity of introducing genes for mass production of proteins and peptides, this methodology is becoming increasingly popular as an alternative protein expression system (1).

While post translational modifications do not exist in bacterial expression systems, plant derived expression systems do facilitate these modifications known to be crucial for protein expression and activity. One of the major differences between mammalian and plant protein expression system is the variation of protein sugar side chains, caused by the differences in biosynthetic pathways. Glycosylation was shown to have a profound effect on activity, folding, stability, solubility, susceptibility to proteases, blood clearance rate and antigenic potential of proteins. Hence, any protein production in plants should take into consideration the potential ramifications of plant glycosylation.

Protein glycosylation is divided into two categories: N-linked and O-linked modifications (2). The two types differ in amino acid to which the glycan moiety is attached to – N-linked are attached to Asn residues, while O-linked are attached to Ser or Thr residues. In addition, the glycan sequences of each type bears unique

distinguishing features. Of the two types, N-linked glycosylation is the more abundant, and its effect on protein function has been extensively studied. O-linked glycans, on the other hand are relatively scarce, and less information is available regarding their affect on proteins.

Summary of the Invention

In a first aspect, there is provided a pharmaceutical composition comprising a human α-galactosidase protein having a
molecular weight of 48.6 kDa, as determined by mass spectrometry, said α-galactosidase protein being further characterized by α-galactosidase catalytic activity and being contiguously linked to a C-terminal KDEL amino acid sequence, and a pharmaceutically acceptable carrier.

In a second aspect, there is provided the use of the pharmaceutical composition as defined above for the manufacture of a medicament for the treatment of Fabry's disease in a subject.

The background art does not teach or suggest a device, system or method for selectively producing glycosylated proteins in plant culture. The background art also does not teach or suggest such a device, system or method for producing high mannose proteins in plant culture. The background art also does not teach or suggest a device, system or method for producing proteins in plant culture through the endoplasmic reticulum (ER). The background art also does not teach or suggest such a device, system or method for producing proteins in plant culture through the endoplasmic reticulum (ER) while by-passing the Golgi body. The background art also does not teach or suggest such a device, system or method for producing proteins in plant culture by using an ER signal to by-pass the Golgi body.

The present disclosure overcomes these disadvantages of the background art by providing a device, system and method for producing glycosylated proteins in plant culture, particularly proteins having a high mannose glycosylation, while optionally and preferably targeting (and/or otherwise manipulating processing of) such proteins with an ER signal. Without wishing to be limited by a single hypothesis, it is believed that such targeting causes the proteins to by-pass the Golgi body and thereby to retain the desired glycosylation, particularly high mannose glycosylation. It should be noted that the term "plant culture" as used herein includes any type of transgenic and/or otherwise genetically engineered plant cell that is grown in culture. The genetic engineering may optionally be permanent or transient. Preferably, the culture features cells that are not assembled to form a complete plant, such that at least one biological structure of a plant is not present. Optionally and preferably, the culture may feature a plurality of different types of plant cells, but preferably the culture features a particular type of plant cell. It should be noted that optionally plant cultures featuring a particular type of plant cell may be originally derived from a plurality of different types of such plant cells.

The plant cells may be grown according to any type of suitable culturing method, including but not limited to, culture on a solid surface (such as a plastic culturing vessel or plate for example) or in suspension.

The disclosure further relates to vectors and methods for expression and production of enzymatically active high mannose lysosomal enzymes using transgenic plant root, particularly carrot cells. More particularly, the disclosure relates to host cells, particularly transgenic suspended carrot cells, vectors and methods for high yield expression and production of biologically active high mannose Glucocerebrosidase (GCD) and α-galactosidase A. The disclosure further provides for compositions and methods for the treatment of lysosomal storage diseases.

The present disclosure is also of a device, system and method for providing sufficient quantities of biologically active lysosomal enzymes, and particularly, human GCD and α-galactosidase A, to deficient cells. The present disclosure is also of host cells comprising new vector compositions that allow for efficient production of genes encoding lysosomal enzymes, such as GCD and α-galactosidase A.

The present disclosure therefore solves a long-felt need for an economically viable technology to produce proteins having particular glycosylation requirements, such as the high mannose glycosylation of lysosomal enzymes such as GCD and α-galactosidase A for example. The present disclosure is able to solve this long felt need by using plant cell culture.

In order to further explain the present disclosure, a brief explanation is now provided of the biosynthetic pathway of high-mannose proteins. The basic biosynthesis pathway of high-mannose and complex N-linked glycans is highly conserved among all eukaryotes. Biosynthesis begins in the Endoplasmic Reticulum (ER) with the transfer of the glycan precursor from a dolichol lipid carrier to a specific Asn residue on the protein by the oligosaccharyl transferase. The precursor is subsequently modified in the ER by glycosidases I and II and a hypothetical mannosidase to yield the high mannose structures, similar to the process occurring in mammals.

Further modifications of the glycan sequence to complex and hybrid structures occur in the Golgi. Such modifications include removal of one of the four mannose residues by α-mannosidase I, addition of an *N*-acetylglucosamine residue, removal of the two additional mannose residues by α-mannosidase II, addition of *N*-

acetylglucosamine and optionally, at this stage, xylose and fucose residues may be added to yield plant specific N-linked glycans. After the transfer of xylose and fucose to the core, complex type N-glycans can be further processed via the addition of terminal fucose and galactose. Further modifications may take place during the glycoprotein transport.

Several approaches are currently used in the background art to control and tailor protein glycosylation in plants, all of which have significant deficiencies, particularly in comparison to the present disclosure. Gross modifications, such as complete inhibition of glycosylation or the removal of glycosylation sites from the peptide chain is one strategy. However, this approach can result in structural defects. An additional approach involves knock-out and introduction of specific carbohydrate processing enzymes. Again, this approach is difficult and may also have detrimental effects on the plant cells themselves.

The present disclosure overcomes these deficiencies of the background art approaches by using an ER signal and/or by blocking secretion from the ER to the Golgi body. Without wishing to be limited by a single hypothesis, since a high mannose structure of lysosomal enzymes is preferred, if secretion can be blocked and the protein can be maintained in the ER, naturally occurring high mannose structures are obtained without the need for remodeling.

As indicated above, proteins transported via the endomembrane system first pass into the endoplasmic reticulum. The necessary transport signal for this step is represented by a signal sequence at the N-terminal end of the molecule, the so-called signal peptide. As soon as this signal peptide has fulfilled its function, which is to insert the precursor protein attached to it into the endoplasmic reticulum, it is split off proteolytically from the precursor protein. By virtue of its specific function, this type of signal peptide sequence has been conserved to a high degree during evolution in all living cells, irrespective of whether they are bacteria, yeasts, fungi, animals or plants.

Many plant proteins, which are inserted into the endoplasmic reticulum by virtue of the signal peptide do not reside in the ER, but are transported from the endoplasmic reticulum to the Golgi and continue trafficking from the Golgi to the vacuoles. One class of such sorting signals for this traffic resides are signals that reside on the C-terminal part of the precursor protein [Neuhaus and Rogers, (1998) Plant Mol. Biol.

38:127-144]. Proteins containing both an N-terminal signal peptide for insertion into the endoplasmic reticulum and a C-terminal vacuolar targeting signal are expected to contain complex glycans, which is attached to them in the Golgi [Lerouge et al., (1998) Plant Mol. Biol. 38:31-48]. The nature of such C- terminal sorting signals can vary very widely. US 6,054,637 describes peptide fragments obtained from the region of tobacco basic chitinase, which is a vacuolar protein that act as vacuolar targeting peptides. An example for a vacuolar protein containing a C-terminal targeting signal and complex glycans is the phaseolin storage protein from bean seeds [Frigerio et al., (1998) Plant Cell 10:1031-1042; Frigerio et al., (2001) Plant Cell 13:1109-1126.].

The paradigm is that in all eukaryotic cells vacuolar proteins pass via the ER and the Golgi before sequestering in the vacuole as their final destination. Surprisingly, the transformed plant root cells of the present disclosure produced an unexpected high mannose GCD and α-galactosidase A. Advantageously, this high mannose product was found to be biologically active and therefore no further steps were needed for its activation. Without wishing to be limited by a single hypothesis, it would appear that the use of an ER signal with the recombinant protein being produced in plant cell culture was able to overcome transportation to the Golgi, and hence to retain the desired high mannose glycosylation. Optionally, any type of mechanism which is capable to produce high mannose glycosylation, including any type of mechanism to by-pass the Golgi, may be used in accordance with the present disclosure.

In one embodiment, the present disclosure relates to a host cell producing a high mannose recombinant protein of interest. This cell may be transformed or transfected with a recombinant nucleic acid molecule encoding a protein of interest or with an expression vector comprising the nucleic acid molecule. Such nucleic acid molecule comprises a first nucleic acid sequence encoding the protein of interest operably linked to a second nucleic acid sequence encoding a vacuolar targeting signal peptide. The first nucleic acid sequence may be optionally further operably linked to a third nucleic acid sequence encoding an ER (endoplasmic reticulum) targeting signal peptide. The host cell of the disclosure is characterized in that the protein of interest is produced by the cell in a highly mannosylated form.

The host cell of the disclosure may be a eukaryotic or prokaryotic cell.

In one embodiment, the host cell of the disclosure is a prokaryotic cell, preferably, a bacterial cell, most preferably, an *Agrobacterium tumefaciens* cell. These cells are used for infecting the preferred plant host cells described below.

In another preferred embodiment, the host cell of the disclosure may be a eukaryotic cell, preferably, a plant cell, and most preferably, a plant root cell selected from the group consisting of *Agrobacterium rihzogenes* transformed root cell, celery cell, ginger cell, horseradish cell and carrot cell.

In a preferred embodiment, the plant root cell is a carrot cell. It should be noted that the transformed carrot cells of the disclosure are grown in suspension. As mentioned above and described in the Examples, these cells were transformed with the *Agrobacterium tumefaciens* cells.

In another embodiment, the recombinant nucleic acid molecule comprised within the host cell of the disclosure, comprises a first nucleic acid sequence encoding a lysosomal enzyme that is in operable linkage with a second nucleic acid sequence encoding a vacuolar targeting signal peptide derived from the basic tobacco chitinase A gene. This vacuolar signal peptide has the amino acid sequence as denoted by SEQ ID NO: 2. The first nucleic acid sequence may be optionally further linked in an operable linkage with a third nucleic acid sequence encoding an ER (endoplasmic reticulum) targeting signal peptide as denoted by SEQ ID NO: 1. In one embodiment, the recombinant nucleic acid molecule comprised within the host cell of the disclosure further comprises a promoter that is functional in plant cells. This promoter should be operably linked to the recombinant molecule of the disclosure.

In another embodiment, this recombinant nucleic acid molecule may optionally further comprise an operably linked terminator which is preferably functional in plant cells. The recombinant nucleic acid molecule of the disclosure may optionally further comprise additional control, promoting and regulatory elements and/or selectable markers. It should be noted that these regulatory elements are operably linked to the recombinant molecule.

In a preferred embodiment, the high mannose protein of interest produced by the host cell of the disclosure may be a high mannose glycoprotein having exposed mannose terminal residues.

Such high mannose protein may be according to another preferred embodiment, a lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-*N*-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase. In a preferred embodiment, the lysosomal enzyme may be the human glucocerebrosidase (GCD) or α-galactosidase A. Hereinafter recombinant GCD, rGCD, rhGCD all refer to various forms of recombinant human GCD unless otherwise indicated. Henceforth A-gal, A-gal A, recombinant A-gal, rA-gal, rhA-gal all refer to various forms of recombinant human α-galactosidase A [Genbank accession numbers NM000169 (coding sequence) and CAA29232 (amino acid sequence)] unless otherwise indicated.

As previously described, Gaucher's disease, the most prevalent lysosomal storage disorder, is caused by point mutations in the hGCD (human glucocerebrosidase) gene (GBA), which result in accumulation of GlcCer in the lysosomes of macrophages. The identification of GCD deficiency as the primary cause of Gaucher's disease led to the development of enzyme replacement therapy as a therapeutic strategy for this disorder. However, glycosylation plays a crucial role in hGCD activity and uptake to target cells.

Therefore, according to other preferred embodiments of the present disclosure, suitably glycosylated hGCD or α-galactosidase A is preferably provided by controlling the expression of hGCD or hα-galactosidase A in plant cell culture, optionally and more preferably by providing an ER signal and/or otherwise by optionally and more preferably blocking transportation to the Golgi.

Optionally and preferably, the hGCD or α-galactosidase A has at least one oligosaccharide chain comprising an exposed mannose residue for the treatment or prevention of Gaucher's disease.

Still further, in a particular embodiment, this preferred host cell is transformed or transfected by a recombinant nucleic acid molecule which further comprises an ^{35}S promoter from Cauliflower Mosaic Virus, an octopine synthase terminator of *Agrobacterium tumefaciens* and TMV (Tobacco Mosaic Virus) omega translational enhancer element. According to a preferred embodiment, this recombinant nucleic acid molecule comprises the nucleic acid sequence substantially as denoted by SEQ ID NO:

13 and encodes a high mannose GCD having the amino acid sequence substantially as denoted by SEQ ID NOs: 14 or 15.

It should be appreciated that the present disclosure further provides for an expression vector comprising a nucleic acid molecule encoding a biologically active lysosomal enzyme.

In one preferred embodiment, the expression vector of the disclosure comprises a nucleic acid molecule encoding a biologically active high mannose human glucocerebrosidase (GCD) or α-galactosidase A. Preferably, this preferred expression vector comprises a nucleic recombinant nucleic acid molecule which having the nucleic acid sequence substantially as denoted by SEQ ID NOs: 13, 17 or 19.

In another embodiment, the present disclosure relates to a recombinant high mannose protein produced by the host cell of the disclosure.

In a preferred embodiment, this high mannose protein may be a biologically active high mannose lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase. Most preferably, this lysosomal enzyme may be human glucocerebrosidase (GCD).

Still further, the disclosure provides for a recombinant biologically active high mannose lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue.

According to a preferred embodiment, the recombinant lysosomal enzyme of the disclosure can bind to a mannose receptor on a target cell in a target site. Preferably, this site may be within a subject suffering from a lysosomal storage disease.

It should be noted that the recombinant lysosomal enzyme has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme for the target cell. In a specific embodiment, the target cell at the target site may be a Kupffer cell in the liver of the subject.

In a preferred embodiment, the recombinant lysosomal enzyme may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase,

hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase or sialidase.

Most preferably, this recombinant lysosomal enzyme is glucocerebrosidase (GCD).

In another embodiment, the disclosure relates to a method of producing a high mannose protein. Accordingly, the method of the disclosure comprises the steps of: (a) preparing a culture of recombinant host cells transformed or transfected with a recombinant nucleic acid molecules encoding a recombinant protein of interest or with an expression vector comprising the recombinant nucleic acid molecules; (b) culturing these host cell culture prepared by step (a) under conditions permitting the expression of the protein, wherein the host cells produce the protein in a highly mannosylated form; (c) recovering the protein from the cells and harvesting the cells from the culture provided in (a); and (d) purifying the protein of step (c) by a suitable protein purification method.

According to a preferred embodiment, the host cell used by this method is the host cell of the disclosure.

In another preferred embodiment, the high mannose protein produced by the method of the disclosure may be a biologically active high mannose lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue.

This recombinant enzyme can bind to a mannose receptor on a target cell in a target site. More particularly, the recombinant enzyme produced by the method of the disclosure has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell. Accordingly, the target cell at the target site may be Kupffer cell in the liver of the subject.

In a specific embodiment, this lysosomal enzyme may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase. Most preferably, this lysosomal enzyme may be glucocerebrosidase (GCD) or an α-galactosidase A.

In another preferred embodiment, the host cell used by the method of the disclosure may be a plant root cell selected from the group consisting of *Agrobacterium rihzogenes* transformed root cell, celery cell, ginger cell, horseradish cell and carrot cell.

Most preferably, the plant root cell is a carrot cell. It should be particularly noted that in the method of the disclosure, the transformed host carrot cells are grown in suspension.

In a further embodiment, the present disclosure relates to a method for treating a subject having lysosomal storage disease using exogenous recombinant lysosomal enzyme, comprising: (a) providing a recombinant biologically active form of lysosomal enzyme purified from transformed plant root cells, and capable of efficiently targeting cells abnormally deficient in the lysosomal enzyme. This recombinant biologically active enzyme has exposed terminal mannose residues on appended oligosaccharides; and (b) administering a therapeutically effective amount of the recombinant biologically active lysosomal enzyme to the subject. In a preferred embodiment, the recombinant high mannose lysosomal enzyme used by the method of the disclosure may be produced by the host cell of the disclosure. Preferably, this host cell is a carrot cell.

In another preferred embodiment, the lysosomal enzyme used by the method of the disclosure may be a high mannose enzyme comprising at least one oligosaccharide chain having an exposed mannose residue. This recombinant enzyme can bind to a mannose receptor on a target cell in a target site within a subject. More preferably, this recombinant lysosomal enzyme has increased affinity for these target cells, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell.

More specifically, the lysosomal enzyme used by the method of the disclosure may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase or sialidase. Preferably, this lysosomal enzyme is glucocerebrosidase (GCD).

According to a preferred embodiment, the method of the disclosure is therefore intended for the treatment of a lysosomal storage disease, particularly Gaucher's disease.

In such case the target cell at the target site may be a Kupffer cell in the liver of the subject.

The disclosure further provides for a pharmaceutical composition for the treatment of a lysosomal storage disease comprising as an active ingredient a recombinant biologically active high mannose lysosomal enzyme as defined by the

disclosure. The composition of the disclosure may optionally further comprise pharmaceutically acceptable dilluent, carrier or excipient.

In a specific embodiment, the composition of the disclosure is intended for the treatment of Gaucher's disease. Such composition may preferably comprise as an effective ingredient a biologically active high mannose human glucocerebrosidase (GCD), as defined by the disclosure.

The disclosure further relates to the use of a recombinant biologically active high mannose lysosomal enzyme of the disclosure in the manufacture of a medicament for the treatment or prevention of a lysosomal storage disease. More particularly, such disease may be Gaucher's disease.

Accordingly, this biologically active lysososomal enzyme is a biologically active high mannose human glucocerebrosidase (GCD), as defined by the disclosure.

According to the present disclosure, there is provided a host cell producing a high mannose recombinant protein, comprising a polynucleotide encoding the recombinant protein and a signal for causing the recombinant protein to be produced as a high mannose protein. Preferably, the polynucleotide comprises a first nucleic acid sequence encoding the protein of interest operably linked to a second nucleic acid sequence encoding a signal peptide. Optionally, the signal peptide comprises an ER (endoplasmic reticulum) targeting signal peptide. Preferably, the polynucleotide further comprises a third nucleic acid sequence for encoding a vacuolar targeting signal peptide.

Preferably, the signal causes the recombinant protein to be targeted to the ER. More preferably, the signal comprises a signal peptide for causing the recombinant protein to be targeted to the ER. Most preferably, the polynucleotide comprises a nucleic acid segment for encoding the signal peptide.

Optionally and preferably, the signal causes the recombinant protein to by-pass the Golgi. Preferably, the signal comprises a signal peptide for causing the recombinant protein to not be targeted to the Golgi. More preferably, the polynucleotide comprises a nucleic acid segment for encoding the signal peptide.

Optionally and preferably, the host cell is any one of a eukaryotic and a prokaryotic cell. Optionally, the prokaryotic cell is a bacterial cell, preferably an *Agrobacterium tumefaciens* cell. Preferably, the eukaryotic cell is a plant cell. More preferably, the plant cell is a plant root cell selected from the group consisting of

Agrobacterium rihzogenes transformed root cell, celery cell, ginger cell, horseradish cell and carrot cell. Most preferably, the plant root cell is a carrot cell.

Preferably, the recombinant polynucleotide comprises a first nucleic acid sequence encoding the protein of interest that is in operable link with a second nucleic acid sequence encoding a vacuolar targeting signal peptide derived from the basic tobacco chitinase A gene, which vacuolar signal peptide has the amino acid sequence as denoted by SEQ ID NO: 2, wherein the first nucleic acid sequence is optionally further operably linked to a third nucleic acid sequence encoding an ER (endoplasmic reticulum) targeting signal peptide as denoted by SEQ ID NO: 1.

More preferably, the recombinant polynucleotide further comprises a promoter that is functional in plant cells, wherein the promoter is operably linked to the recombinant molecule.

Most preferably, the recombinant polynucleotide further comprises a terminator that is functional in plant cells, wherein the terminator is operably linked to the recombinant molecule.

Also most preferably, the recombinant polynucleotide optionally further comprises additional control, promoting and regulatory elements and/or selectable markers, wherein the regulatory elements are operably linked to the recombinant molecule.

Preferably, the high mannose protein is a high mannose glycoprotein having glycosylation with at least one exposed mannose residue. More preferably, the high mannose protein is a biologically active high mannose lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase

Most preferably, the lysosomal enzyme is human glucocerebrosidase (GCD).

Preferably, the GCD comprises the amino acid sequence substantially as denoted by SEQ ID NO: 8, encoded by the nucleic acid sequence as denoted by SEQ ID NO: 7.

More preferably, the cell is transformed or transfected with a recombinant polynucleotide or with an expression vector comprising the molecule, which recombinant polynucleotide further comprises an ^{35}S promoter from Cauliflower Mosaic Virus, an octopine synthase terminator of *Agrobacterium tumefaciens*, and the

regulatory element is the TMV (Tobacco Mosaic Virus) omega translational enhancer element, and having the nucleic acid sequence substantially as denoted by SEQ ID NO: 13 encoding GCD having the amino acid sequence substantially as denoted by SEQ ID NOs: 14 or 15.

According to preferred embodiments, there is provided a recombinant high mannose protein produced by the host cell described above.

Preferably, the high mannose protein is a biologically active high mannose lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase.

More preferably, the lysosomal enzyme is human glucocerebrosidase (GCD).

According to other preferred embodiments of the present disclosure, there is provided a recombinant biologically active high mannose lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue.

According to still other preferred embodiments, there is provided a recombinant protein, comprising a first portion having signal peptide activity and a second portion having lysosomal enzyme activity, the first portion causing the second portion to be processed in a plant cell with at least one oligosaccharide chain comprising an exposed mannose residue.

Preferably, the lysosomal enzyme comprises a protein for the treatment or prevention of Gaucher's disease.

More preferably, the protein comprises hGCD.

In another embodient, the lysosomal enzyme comprises a protein for the treatment or prevention of Fabry disease.

More preferably, the protein comprises α-galactosidase A.

Preferably, the first portion comprises a plant cell ER targeting signal peptide. More preferably, the recombinant enzyme can bind to a mannose receptor on a target cell in a target site within a subject suffering from a lysosomal storage disease. Most preferably, the recombinant lysosomal enzyme has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal

enzyme for the target cell. The target cell can be a fibroblast, macrophage, and the like which having mannose receptors.

Also most preferably, the recombinant lysosomal enzyme is selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase or sialidase.

Preferably, the recombinant lysosomal enzyme is glucocerebrosidase (GCD).

Also preferably, the target cell at the target site is a Kupffer cell in the liver of the subject.

According to still other preferred embodiments there is provided a recombinant high mannose protein, produced in plant cell culture. Preferably, the protein features a plant signal peptide for targeting a protein to the ER.

More preferably, the plant signal peptide comprises a peptide for targeting the protein to the ER in a root plant cell culture. Most preferably, the root plant cell culture comprises carrot cells.

According to yet other preferred embodiments there is provided a recombinant mannose-rich hGCD or α-galactosidase A protein, produced in plant cell culture.

According to still other preferred embodiments there is provided use of a plant cell culture for producing a high mannose protein.

According to other preferred embodiments there is provided a method of producing a high mannose protein comprising: preparing a culture of recombinant host cells transformed or transfected with a recombinant polynucleotide encoding for a recombinant protein; culturing the host cell culture under conditions permitting the expression of the protein, wherein the host cells produce the protein in a highly mannosylated form.

Preferably, the host cell culture is cultured in suspension. More preferably, the method further comprises purifying the protein.

According to other preferred embodiments, the method is performed with the host cell as previously described. Preferably, the high mannose protein is a biologically active high mannose lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue. More preferably, the recombinant enzyme binds to a mannose receptor on a target cell in a target site. Most preferably, the

recombinant enzyme has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell.

Preferably, the lysosomal enzyme is selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase.

More preferably, the lysosomal enzyme is glucocerebrosidase (GCD) or α-galactosidase A. Most preferably, the target cell at the target site is a fibroblast or a Kupffer cell in the liver of the subject.

Preferably, the host cell is a plant root cell selected from the group consisting of *Agrobacterium rihzogenes* transformed root cell, celery cell, ginger cell, horseradish cell and carrot cell. Inanother embodiment, the host cell is a tobacco cell.

More preferably, the plant root cell is a carrot cell.

Most preferably, the transformed host carrot cells are grown in suspension.

According to still other preferred embodiments there is provided a method for treating a subject having lysosomal storage disease using exogenous recombinant lysosomal enzyme, comprising: providing a recombinant biologically active form of lysosomal enzyme purified from transformed plant root cells, and capable of efficiently targeting cells abnormally deficient in the lysosomal enzyme, wherein the recombinant biologically active enzyme has exposed terminal mannose residues on appended oligosaccharides; and administering a therapeutically effective amount of the recombinant biologically active lysosomal enzyme to the subject. This method may optionally be performed with any host cell and/or protein as previous described.

Preferably, the recombinant enzyme can bind to a mannose receptor on a target cell in a target site within a subject. More preferably, the recombinant lysosomal enzyme has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell. Most preferably, the lysosomal enzyme is selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-

mannosidase or sialidase. Also most preferably, the lysosomal enzyme is glucocerebrosidase (GCD).

Also most preferably, the lysosomal storage disease is Gaucher's disease. Also most preferably, the target cell at the target site is a Kupffer cell in the liver of the subject.

In another embodiment, the storage disease is Fabry's disease, the lysosomal enzyme is α-galactosidase A, and the target cell is a fibroblast.

According to still other preferred embodiments there is provided a pharmaceutical composition for the treatment of a lysosomal storage disease comprising as an active ingredient a recombinant biologically active high mannose lysosomal enzyme as described above, which composition optionally further comprises pharmaceutically acceptable dilluent, carrier or excipient. Preferably, the lysosomal storage disease is Gaucher's disease. More preferably, the recombinant lysosomal enzyme is a biologically active high mannose human glucocerebrosidase (GCD).

According to still other preferred embodiments there is provided the use of a recombinant biologically active mannose-rich lysosomal enzyme as described above, in the manufacture of a medicament for the treatment or prevention of a lysosomal storage disease. Preferably, the disease is Gaucher's disease. More preferably, the biologically active lysososomal enzyme is a biologically active high mannose human glucocerebrosidase (GCD).

In another embodiment, the disease is Fabry's disease and the biologically active lysosomal enzyme is α-galactosidase A.

The disclosure will be further described on the hand of the following figures, which are illustrative only and do not limit the scope of the disclosure which is also defined by the appended claims.

Brief Description of the Figures

The invention is herein described, by way of example only, with reference to the accompanying drawings, wherein:

Figure 1A-1B

1A shows the resulting expression cassette comprising ^{35}S promoter from Cauliflower Mosaic Virus, TMV (Tobacco Mosaic Virus) omega translational enhancer element, ER targeting signal, the human GCD sequence (also denoted by SEQ ID NO: 7), vacuolar signal and octopine synthase terminator sequence from *Agrobacterium tumefaciens*.

1B shows a schematic map of pGreenII plasmid backbone.

Figure 2 shows Western blot analysis of hGCD transformed cell extracts using anti hGCD specific antibody. Standard Cerezyme (lane 1) was used as a positive control, untransformed callus was used as negative control (lane 2), various selected calli extracts are shown in lanes 3-8.

Figure 3A-3C shows the first step of purification of rhGCD on a strong cation exchange resin (Macro-Prep high-S support, Bio-Rad), packed in a XK column (2.6x20cm). The column was integrated with an AKTA prime system (Amersham Pharmacia Biotech) that allows conductivity monitoring, pH and absorbency at 280nm. Elution of the rh-GCD was obtained with equilibration buffer containing 600mM NaCl. Fig 3A represents a standard run of this purification step. The fractions collected during the run were monitored by enzyme activity assay, as shown by Fig 3B, and tubes exhibiting enzymatic activity (in the elution peak) were pooled. Fig 3C shows coomassie-blue stain of elution fractions assayed for activity.

Figures 3D-3F show corresponding graphs as for figures 3A-3C but for the second column.

Figure 4A-C : shows the final purification step of the recombinant hGCD on a hydrophobic interaction resin (TSK gel, Toyopearl Phenyl-650C, Tosoh Corp.), packed in a XK column (2.6x20cm). The column was integrated with an AKTA prime system (Amersham Pharmacia Biotech) that allows conductivity monitoring, pH and absorbency at 280nm. The GCD elution pool from the previous column was loaded at 6ml/min followed by washing with equilibration buffer until the UV absorbance reach the baseline. The pure GCD was eluted by 10mM citric buffer containing 50% ethanol.

Fig 4A represents a standard run of this purification step.

Fig 4B shows the fractions collected during the run that were monitored by enzyme activity assay.

Fig 4C shows coomassie-blue stain of elution fractions assayed for activity.

Figure 5 shows activity of recombinant hGCD following uptake by peritoneal macrophages (Figures 5A-5C), while Figure 5D shows a Western blot of recombinant GCD according to the present invention.

Figure 6 shows comparative glycosylation structures for rGCD according to the present invention and that of Cerezyme™.

Figure 7 shows glycosylation structures for rGCD according to the present invention.

Figure 8a-8d shows additional N-glycan glycosylation structures for rGCD according to the present invention.

Figures 9a-9b show the antigenic and electrophoretic identity of purified recombinant human GCD of the present invention and a commercial human GCD (Cerezyme ®) recombinantly produced in mammalian CHO cells. Fig. 9a is a Coomassie blue stained SDS-PAGE analysis of the plant produced hGCD of the invention (lanes 1 and 2, 5 and 10 μg of protein, respectively) and Cerezyme ®, (lanes 3 and 4, 5 and 10 μg protein, respectively). Fig. 9b is a Western blot analysis of SDS-PAGE separated recombinant human GCD (lanes 1 and 2, 50 and10 ng respectively) of the present invention compared to the commercial Cerezyme ® enzyme. SDS-PAGE-separated proteins were blotted onto nitrocellulose (lanes 3 and 4, 50 and 100 ng antigen, respectively), and immunodetected using a polyclonal anti-GCD antibody and peroxidase-conjugated goat anti-rabbit HRP secondary antibody. Note the consistency of size and immune reactivity between the plant recombinant GCD of the present invention and the mammalian-cell (CHO) prepared enzyme (Cerezyme®). MW= molecular weight standard markers;

Figures 10a-10b are schematic representations of the glycan structures of the recombinant human GCD of the present invention. Fig. 10a shows the results of a major glycan structure analysis of the GCD, indicating all structures and their relative amounts based on HPLC, enzyme array digests and MALDI. Retention time of individual glycans is compared to the retention times of a standard partial hydrolysis of dextran giving a ladder of glucose units (GU). Fig. 10b shows the glycan structures of the mammalian-cell (CHO) prepared enzyme (Cerezyme®), before and after the *in-vitro*

modification process. Note the predominance of the xylose and exposed mannose glycosides in the recombinant human GCD of the present invention;

Figure 11 is a HP-anion exchange chromatography analysis of the gycan profile of the recombinant human GCD of the present invention, showing the consistent and reproducible glycan structure of recombinant human GCD from batch to batch;

Figure 12 is a kinetic analysis showing the identical catalytic kinetics characteristic of both recombinant human GCD of the invention (open triangles) and the mammalian-cell (CHO) prepared enzyme (Cerezyme®) (closed squares). Recombinant human GCD of the invention and Cerezyme® (0.2 μg) were assayed using C6-NBDGlcCer (5 min, 37 °C) in MES buffer (50 mM, pH 5.5). Michaelis-Menten kinetics was analyzed using GraphPad Prism software. Data are means of two independent experiments;

Figures 13A and 13B are plots of the results of Molecular Weight analysis of human recombinant α-galactosidase A expressed in tobacco plants. Figure 13A shows molecular weight as determined by gel filtration as described herein. Figure 13B shows molecular weight as determined by mass spectrometry (MALDI-Tof). Note the prominent peak (on MS) at 48.6 kDa, corresponding to the MW of native human α-galactosidase A;

Figures 14A and 14B are a PAGE analysis and amino acid sequence of human recombinant α-galactosidase A expressed in tobacco plants. Figure 14A shows two distinct bands of human recombinant α-galactosidase A, corresponding to 62 kDa and 47.6 kDa, resolved in the PAGE. Figure 14B shows the amino acid sequence derived from each of the two bands (labeled "Upper band" and "Lower band"). The portion of the polypeptides available for sequencing in each band is indicated in red. Regions unable to provide sequence data (possibly masked by glycan structures) are indicated in black. Note the complete agreement of sequenced regions between upper and lower bands, indicating identical polypeptides with possible distinctions in glycan structure;

Figure 15 is photograph of a Western blot showing immunoreactivity of human recombinant α-galactosidase A expressed in tobacco plants. Protein extracted from tobacco plants expressing either the human α-galactosidase A targeted to the vacuole (α-gal-vac, lane "vac") or human α-galactosidase with ER retention signal (α-gal-KDEL, lane "KDEL") was separated on PAGE, blotted onto nitrocellulose, and reacted with

anti- α-galactosidase A antibody (against amino acids 326-429 of human α-galactosidase), and visualized with HRP second antibody. Note the strong, specific reaction in both α-gal-vac and α-gal-KDEL expressed proteins, while protein extracted from transgenic control plants (GFP) was unreactive;

Figures 16A and 16B are graphs representing kinetic analysis of catalytic properties of human recombinant α-galactosidase expressed in plants. Figure 16A is a Michaelis-Menten plot comparing plant expressed human recombinant α-galactosidase A and commercially available recombinant α-galactosidase A preparations. Figure 16B is a Lineweaver-Burke plot of the enzyme kinetics derived from Figure 16A, showing Km and Vmax (detailed in table inset). Green indicates plant-expressed human recombinant α-galactosidase; black indicates Fabrazyme and blue indicates Replagal. Note the close correspondence in enzyme kinetics between plant-expressed human recombinant α-galactosidase and the commercially available preparations;

Figures 17A and 17B are photographs of SDS-PAGE showing the stability of human recombinant α-galactosidase expressed in plants in a range of temperatures. Human recombinant α-galactosidase expressed in plants (Plant a-Gal) and commercial human recombinant α-galactosidase (Replagal) were incubated at the indicated temperatures for 2 hours in activity buffer (Figure 17A) or cell media (Figure 17B), separated on SDS-PAGE, and visualized as described herein;

Figure 18 is a photograph of Western blot analysis of fibroblast cell lysate showing uptake and retention of plant-expressed human recombinant α-galactosidase in Fabry fibroblasts. Lanes "Plant αGalA" are fibroblast lysate from human Fabry (α-galactosidase-deficient) fibroblasts incubated with plant expressed human recombinant α-galactosidase for 2 hours, washed and lysed. Rightmost lane is Fabrazyme®, inbetween are molecular weight ladders;

Figure 19 is a plot of NP-HPLC profile showing peaks of characteristic glycan structures, and the schematics of the glycans themselves.

Detailed Description of the Invention

Proteins for pharmaceutical use have been traditionally produced in mammalian or bacterial expression systems. In the past few years a promising new expression system was found in plants. Due to the relative simplicity of introducing new genes and

potential for mass production of proteins and peptides, 'molecular pharming' is becoming increasingly popular as a protein expression system.

One of the major differences between mammalian and plant protein expression system is the variation of protein glycosylation sequences, caused by the differences in biosynthetic pathways. Glycosylation was shown to have a profound effect on activity, folding, stability, solubility, susceptibility to proteases, blood clearance rate and antigenic potential of proteins. Hence, any protein production in plants should take into consideration the potential ramifications of plant glycosylation.

This is well illustrated by the difficulties encountered in previous attempts to produce biologically active mammalian proteins in plants. For example, US Patent No. 5,929,304, to Radin et al (Crop Tech, Inc) discloses the production, in tobacco plants, of a human α-L-iduronase (IDUA) and a glucocerebrosidase (hGC), by insertion of the relevant human lysosomal enzyme coding sequences into an expression cassette for binary plasmid for *A. tumefaciens*- mediated transformation of tobacco plants. Despite demonstration of recombinant human lysosomal protein production in the transgenic plants, and the detection of catalytic activity in the recombinant protein, no binding to or uptake into target cells was disclosed, and the lysosomal enzyme compositions remained unsuitable for therapeutic applications, presumably due to the absence of accurate glycosylation of the protein, and subsequent inability of the polypeptides to interact efficiently with their target cells/tissue though a specific receptor.

Carbohydrate moiety is one of the most common post-translational modifications of proteins. Protein glycosylation is divided into two categories: N-linked and O-linked. The two types differ in amino acid to which the glycan moity is attached on protein – N-linked are attached to Asn residues, while O-linked are attached to Ser or Thr residues. In addition, the glycan sequences of each type bears unique distinguishing features. Of the two types, N-linked glycosylation is the more abundant, and its effect on proteins has been extensively studied. O-linked glycans, on other hand are relatively scarce, and less information is available regarding their influence on proteins. The majority of data available on protein glycosylation in plants focuses on N-linked, rather than O-linked glycans.

The present invention describes herein a plant expression system based on transgenic plant cells, which are preferably root cells, optionally and preferably grown

in suspension. This expression system is particularly designed for efficient production of a high mannose protein of interest. The term "high mannose" includes glycosylation having at least one exposed mannose residue.

Thus, in a first aspect, the present invention relates to a host cell producing a high mannose recombinant protein of interest. Preferably, the recombinant protein features an ER (endoplasmic reticulum) signal peptide, more preferably an ER targeting signal peptide. Alternatively or additionally, the recombinant protein features a signal that causes the protein to by-pass the Golgi. The signal preferably enables the recombinant protein to feature high mannose glycosylation, more preferably by retaining such glycosylation, and most preferably by targeting the ER and/or by by-passing the Golgi. As described in greater detail herein, such a signal is preferably implemented as a signal peptide, which more preferably forms part of the protein sequence, optionally and more preferably through engineering the protein to also feature the signal peptide as part of the protein. It should be noted that the signal may optionally be a targeting signal, a retention signal, an avoidance (by-pass) signal, or any combination thereof, or any other type of signal capable of providing the desired high mannose glycosylation structure.

Without wishing to be limited by a single hypothesis, it would appear that the use of an ER targeting signal with the recombinant protein being produced in plant cell culture was able to overcome transportation to the Golgi, and hence to retain the desired high mannose glycosylation. Optionally, any type of mechanism which is capable to produce high mannose glycosylation, including any type of mechanism to by-pass the Golgi, may be used in accordance with the present invention. ER targeting signal peptides are well known in the art; they are N-terminal signal peptides. Optionally any suitable ER targeting signal peptide may be used with the present invention.

A host cell according to the present invention may optionally be transformed or transfected (permanently and/or transiently) with a recombinant nucleic acid molecule encoding a protein of interest or with an expression vector comprising the nucleic acid molecule. Such nucleic acid molecule comprises a first nucleic acid sequence encoding the protein of interest, optionally and preferably operably linked to a second nucleic acid sequence encoding a vacuolar targeting signal peptide. It should be noted that as used herein, the term "operably" linked does not necessarily refer to physical linkage.

The first nucleic acid sequence may optionally and preferably further be operably linked to a third nucleic acid sequence encoding an ER (endoplasmic reticulum) targeting signal peptide. In one embodiment, the cell of the invention is characterized in that the protein of interest is produced by the cell in a form that includes at least one exposed mannose residue, but is preferably a highly mannosylated form. In a more preferred embodiment, the cell of the protein of interest is produced by the cell in a form that includes an exposed mannose and at least one xylose residue, in yet a more preferred embodiment, in a form that further includes an exposed mannose and at least one fucose residue. In a most preferred embodiment, the protein is produced by the cell in a form that includes an exposed mannose, a core α (1,2) xylose residue and a core α-(1,3) fucose residue.

"Cells", "host cells" or "recombinant host cells" are terms used interchangeably herein. It is understood that such terms refer not only to the particular subject cells but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generation due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. "Cell" or "host cell" as used herein refers to cells which can be transformed with naked DNA or expression vectors constructed using recombinant DNA techniques. As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., naked DNA or an expression vector, into a recipient cells by nucleic acid-mediated gene transfer. "Transformation", as used herein, refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of the desired protein.

It should be appreciated that a drug resistance or other selectable marker is intended in part to facilitate the selection of the transformants. Additionally, the presence of a selectable marker, such as drug resistance marker may be of use in keeping contaminating microorganisms from multiplying in the culture medium. Such a pure culture of the transformed host cell would be obtained by culturing the cells under conditions which are required for the induced phenotype's survival.

As indicated above, the host cells of the invention may be transfected or transformed with a nucleic acid molecule. As used herein, the term "nucleic acid" refers

to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The terms should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

In yet another embodiment, the cell of the invention may be transfected or transformed with an expression vector comprising the recombinant nucleic acid molecule. "Expression Vectors", as used herein, encompass vectors such as plasmids, viruses, bacteriophage, integratable DNA fragments, and other vehicles, which enable the integration of DNA fragments into the genome of the host. Expression vectors are typically self-replicating DNA or RNA constructs containing the desired gene or its fragments, and operably linked genetic control elements that are recognized in a suitable host cell and effect expression of the desired genes. These control elements are capable of effecting expression within a suitable host. Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system. Such system typically includes a transcriptional promoter, an optional operator to control the onset of transcription, transcription enhancers to elevate the level of RNA expression, a sequence that encodes a suitable ribosome binding site, RNA splice junctions, sequences that terminate transcription and translation and so forth. Expression vectors usually contain an origin of replication that allows the vector to replicate independently of the host cell.

Plasmids are the most commonly used form of vector but other forms of vectors which serves an equivalent function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels et al. Cloning Vectors: a Laboratory Manual (1985 and supplements), Elsevier, N.Y.; and Rodriquez, et al. (eds.) Vectors: a Survey of Molecular Cloning Vectors and their Uses, Buttersworth, Boston, Mass (1988), which are incorporated herein by reference.

In general, such vectors contain, in addition, specific genes which are capable of providing phenotypic selection in transformed cells. The use of prokaryotic and eukaryotic viral expression vectors to express the genes coding for the polypeptides of the present invention are also contemplated.

Optionally, the vector may be a general plant vector (as described with regard to the Examples below). Alternatively, the vector may optionally be specific for root cells.

In one preferred embodiment, the cell of the invention may be a eukaryotic or prokaryotic cell.

In a specific embodiment, the cell of the invention is a prokaryotic cell, preferably, a bacterial cell, most preferably, an *Agrobacterium tumefaciens* cell. These cells are used for infecting the preferred plant host cells described below.

In another preferred embodiment, the cell of the invention may be a eukaryotic cell, preferably, a plant cell, and most preferably, a plant root cell selected from the group consisting of *Agrobacterium rihzogenes* transformed plant root cell, celery cell, ginger cell, horseradish cell and carrot cell.

In a preferred embodiment, the plant root cell is a carrot cell. It should be noted that the transformed carrot cells of the invention are grown in suspension. As mentioned above and described in the Examples, these cells were transformed with the *Agrobacterium tumefaciens* cells of the invention.

The expression vectors or recombinant nucleic acid molecules used for transfecting or transforming the host cells of the invention may be further modified according to methods known to those skilled in the art to add, remove, or otherwise modify peptide signal sequences to alter signal peptide cleavage or to increase or change the targeting of the expressed lysosomal enzyme through the plant endomembrane system. For example, but not by way of limitation, the expression construct can be specifically engineered to target the lysosomal enzyme for secretion, or vacuolar localization, or retention in the endoplasmic reticulum (ER).

In one embodiment, the expression vector or recombinant nucleic acid molecule, can be engineered to incorporate a nucleotide sequence that encodes a signal targeting the lysosomal enzyme to the plant vacuole. For example, and not by way of limitation, the recombinant nucleic acid molecule comprised within the host cell of the invention, comprises a first nucleic acid sequence encoding a lysosomal enzyme that is in operable linkage with a second nucleic acid sequence encoding a vacuolar targeting signal peptide derived from the basic tobacco chitinase A gene. This vacuolar signal peptide has the amino acid sequence as denoted by SEQ ID NO: 2. The first nucleic acid

sequence may be optionally further linked in an operable linkage with a third nucleic acid sequence encoding an ER (endoplasmic reticulum) targeting signal peptide as denoted by SEQ ID NO: 1. In one embodiment, the recombinant nucleic acid molecule comprised within the host cell of the invention further comprises a promoter that is functional in plant cells. This promoter should be operably linked to the recombinant molecule of the invention.

The term "operably linked" is used herein for indicating that a first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Optionally and preferably, operably linked DNA sequences are contiguous (e.g. physically linked) and, where necessary to join two protein-coding regions, in the same reading frame. Thus, a DNA sequence and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the regulatory sequence(s).

In another embodiment, this recombinant nucleic acid molecule may optionally further comprise an operably linked terminator which is preferably functional in plant cells. The recombinant nucleic acid molecule of the invention may optionally further comprise additional control, promoting and regulatory elements and/or selectable markers. It should be noted that these regulatory elements are operably linked to the recombinant molecule.

Regulatory elements that may be used in the expression constructs include promoters which may be either heterologous or homologous to the plant cell. The promoter may be a plant promoter or a non-plant promoter which is capable of driving high levels transcription of a linked sequence in plant cells and plants. Non-limiting examples of plant promoters that may be used effectively in practicing the invention include cauliflower mosaic virus (CaMV) ^{35}S, rbcS, the promoter for the chlorophyll a/b binding protein, AdhI, NOS and HMG2, or modifications or derivatives thereof. The promoter may be either constitutive or inducible. For example, and not by way of limitation, an inducible promoter can be a promoter that promotes expression or

increased expression of the lysosomal enzyme nucleotide sequence after mechanical gene activation (MGA) of the plant, plant tissue or plant cell.

The expression vectors used for transfecting or transforming the host cells of the invention can be additionally modified according to methods known to those skilled in the art to enhance or optimize heterologous gene expression in plants and plant cells. Such modifications include but are not limited to mutating DNA regulatory elements to increase promoter strength or to alter the protein of interest.

In a preferred embodiment, the high mannose protein of interest produced by the host cell of the invention may be a mannose-rich glycoprotein having at least one exposed mannose residue (at least one terminal mannose residue). In another embodiment, in the glycoprotein of the invention has most (greater than 75%) of the mannose residues are terminal, exposed mannose residues.

Such high mannose protein may be according to another preferred embodiment, a lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase

The term "lysosomal enzyme", as used herein with respect to any such enzyme and product produced in a plant expression system described by the invention, refers to a recombinant peptide expressed in a transgenic plant cell from a nucleotide sequence encoding a human or animal lysosomal enzyme, a modified human or animal lysosomal enzyme, or a fragment, derivative or modification of such enzyme. Useful modified human or animal lysosomal enzymes include but are not limited to human or animal lysosomal enzymes having one or several naturally occurring or artificially introduced amino acid additions, deletions and/or substitutions.

Soluble lysosomal enzymes share initial steps of biosynthesis with secretory proteins, i.e., synthesis on the ribosome, binding of the N-terminal signal peptide to the surface of the rough endoplasmic reticulum (ER), transport into the lumen of the ER where the signal peptide is cleaved, and addition of oligosaccharides to specific asparagine residues (N-linked), followed by further modifications of the nascent protein in the Golgi apparatus [von Figura and Hasilik, Annu. Rev. Biochem. 55:167-193 (1986)]. The N-linked oligosaccharides can be complex, diverse and heterogeneous, and

may contain high-mannose residues. The proteins undergo further processing in a post-ER, pre-Golgi compartment and in the cis-Golgi to form either an N-linked mannose 6-phosphate (M-6-P) oligosaccharide-dependent or N-linked M-6-P oligosaccharide-independent recognition signal for lysosomal localized enzymes [Kornfeld & Mellman, Ann. Rev. Cell Biol., 5:483-525 (1989); Kaplan et al., Proc. Natl. Acad. Sci. USA 74:2026 (1977)]. The presence of the M-6-P recognition signal results in the binding of the enzyme to M-6-P receptors (MPR). These bound enzymes remain in the cell, are eventually packaged into lysosomes, and are thus segregated from proteins targeted for secretion or to the plasma membrane.

In a preferred embodiment, the lysosomal enzyme may be the human glucocerebrosidase (GCD) or human α-galactosidase A.

Still further, in a particular embodiment, this preferred host cell is transformed or transfected by a recombinant nucleic acid molecule which further comprises an ^{35}S promoter from Cauliflower Mosaic Virus, preferably, having the nucleic acid sequence as denoted by SEQ ID NO: 9, an octopine synthase terminator of *Agrobacterium tumefaciens*, preferably, having the nucleic acid sequence as denoted by SEQ ID NO: 12 and TMV (Tobacco Mosaic Virus) omega translational enhancer element. According to a preferred embodiment, this recombinant nucleic acid molecule comprises the nucleic acid sequence substantially as denoted by SEQ ID NO: 13 and encodes a high mannose GCD having the amino acid sequence substantially as denoted by SEQ ID NOs: 14 or 15.

It should be appreciated that the present invention further provides for an expression vector comprising a nucleic acid molecule encoding a biologically active high mannose lysosomal enzyme.

In one preferred embodiment of the aspect, the expression vector of the invention comprises a nucleic acid molecule encoding a biologically active high mannose human glucocerebrosidase (GCD). Preferably, this preferred expression vector comprises a recombinant nucleic acid molecule which having the nucleic acid sequence substantially as denoted by SEQ ID NO: 13. According to a specific embodiment, a preferred expression vector utilizes the pGREEN II plasmid as described by the following Example 1. In another embodiment of the invention, the expression vector comprises a nucleic acid molecule encoding a biologically active high mannose human

α-galactosidase (α-gal-A). Preferably, this preferred expression vector comprises a recombinant nucleic acid molecule which having the nucleic acid sequence substantially as denoted by SEQ ID NO: 17 or 19. According to a specific embodiment, a preferred expression vector utilizes the pICH19170 plasmid as described by the following Example 5a.

It should be further noted, that the invention provides for an expression cassette comprised within the expression vector described above.

In a second aspect, the present invention relates to a recombinant high mannose protein produced by the host cell of the invention.

In a preferred embodiment, this high mannose protein may be a biologically active high mannose lysosomal enzyme selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase. Most preferably, this lysosomal enzyme may be human glucocerebrosidase (GCD).

The term "biologically active" is used herein with respect to any recombinant lysosomal enzyme produced in a plant expression system to mean that the recombinant lysosomal enzyme is able to hydrolyze either the natural substrate, or an analogue or synthetic substrate of the corresponding human or animal lysosomal enzyme, at detectable levels.

Still further, the invention provides for a recombinant biologically active mannose-rich lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue.

According to a preferred embodiment, the recombinant lysosomal enzyme of the invention can bind to a mannose receptor on a target cell in a target site. Preferably, this site may be within a subject suffering from a lysosomal storage disease.

Optionally and more preferably, the recombinant lysosomal enzyme has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme for the target cell. In a specific embodiment, the target cell at the target site may be a Kupffer cell in the liver of the subject.

In a preferred embodiment, the recombinant lysosomal enzyme may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase,

hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase or sialidase.

Most preferably, this recombinant lysosomal enzyme is glucocerebrosidase (GCD) or α-galactosidase A.

In a third aspect, the invention relates to a method of producing a high mannose protein. Accordingly, the method of the invention comprises the steps of: (a) preparing a culture of recombinant host cells transformed or transfected with a recombinant nucleic acid molecules encoding for a recombinant protein of interest or with an expression vector comprising the recombinant nucleic acid molecules; (b) culturing the host cell culture prepared by step (a) in suspension under conditions permitting the expression of the high mannose protein, wherein the host cells produce the protein in a highly mannosylated form; (c) harvesting the cells from the culture provided in (a) and recovering the protein from the cells; and (d) purifying the protein of step (c) by a suitable protein purification method.

Optionally and preferably, the recombinant protein may be produced by plant cells according to the present invention by culturing in a device described with regard to US Patent No. 6,391,638, issued on May 21 2002 and hereby incorporated by reference as if fully set forth herein. Conditions for culturing plant cells in suspension with this device are described with regard to the US patent application entitled "CELL/TISSUE CULTURING DEVICE, SYSTEM AND METHOD" by one of the present inventors and owned in common with the present application, which is hereby incorporated by reference as if fully set forth herein and which was filed on the same day as the present application.

According to a further aspect of an embodiment of the invention, the recombinant protein can be expressed in a whole plant, or a part thereof. Accordingly, the method of the invention comprises the steps of: (a) transforming or transfecting a plant or plant cells with a recombinant nucleic acid molecules encoding for a recombinant protein of interest or with an expression vector comprising the recombinant nucleic acid molecules; (b) growing the transformed or transfected plant or cells prepared by step (a) under conditions permitting the expression of the mannose-rich protein, wherein the plant cells produce the protein in mannosylated form having exposed and terminal mannose residues; (c) harvesting the plants or tissues from the

plant or plant tissues provided in (a) and recovering the protein from the cells; and (d) purifying the protein of step (c) by a suitable protein purification method. In another embodiment of the invention, transformation of the plants with the vector is stable transformation, and step (b) is followed by selection of plants expressing the recombinant protein of interest, and propagation of the selected transgenic plants, before harvesting and recovering the recombinant protein. Transforming plants or plant tissues (including, but not limited to callus, immature embryo, pollen, seed, shoot apex parts in culture as well as *in planta*) with recombinant expression vectors, for constitutive or conditional expression of desired mammalian polypeptide, is well known in the art, for example, using a binary plasmid for *A. tumefaciens*- mediated transformation of tobacco plants (as described in US Patent No: 5,763,748), using an co-integrated vector, or using a mobilization vector.

A particular and non limiting example for recovering and purification of a high mannose protein of interest produced by the method of the invention may be found in the following Examples. The Examples show that a recombinant h-GCD produced by the invention was unexpectedly bound to internal membrane of the transformed carrot cells of the invention and not secreted to the medium. The soluble rh-GCD may be separated from cell debris and other insoluble component according to means known in the art such as filtration or precipitation. For Example, following a freeze-thaw cycle, the cells undergo breakage and release of intracellular soluble proteins, whereas the h-GCD remains bound to insoluble membrane debris. This soluble and insoluble membrane debris mixture was next centrifuged and the soluble fraction was removed thus simplifying the purification. The membrane bound h-GCD can then be dissolved by mechanical disruption in the presence of a mild detergent, protease inhibitors and neutralizing oxidation reagent. The soluble enzyme may be further purified using chromatography techniques, such as cation exchange and hydrophobic interaction chromatography columns. During rh-GCD production in the bio-reactor and the purification process the h-GCD identity, yield, purity and enzyme activity can be determined by one or more biochemical assays. Including but not limited to detecting hydrolysis of the enzyme's substrate or a substrate analogue, SDS-polyacrylamide gel electrophoresis analysis and immunological analyses such as ELISA and Western blot.

Yield, purity, enzyme activity, antigenic character, biological activity and glycan

profile of recombinant proteins expressed in whole plants and plant tissues, as described in the following Examples, can be assessed by one or more biochemical assays, including but not limited to detecting hydrolysis of the enzyme's substrate or a substrate analogue, SDS-polyacrylamide gel electrophoresis analysis, immunological analyses such as ELISA and Western blot, glycan analysis by glycosidase enzymes and chromatography.

According to a preferred embodiment, the host cell used by this method comprises the host cell of the invention.

In another preferred embodiment, the high mannose protein produced by the method of the invention may be a biologically active high mannose lysosomal enzyme having at least one oligosaccharide chain comprising an exposed mannose residue.

This recombinant enzyme can bind to a mannose receptor on a target cell in a target site. More particularly, the recombinant enzyme produced by the method of the invention has increased affinity for the target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell. Accordingly, the target cell at the target site may be Kupffer cell in the liver of the subject.

In a specific embodiment, this lysosomal enzyme may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase and sialidase. Most preferably, this lysosomal enzyme may be glucocerebrosidase (GCD).

In another preferred embodiment, the host cell used by the method of the invention may be a plant root cell selected from the group consisting of *Agrobacterium rihzogenes* transformed root cell, celery cell, ginger cell, horseradish cell and carrot cell. Most preferably, the plant root cell is a carrot cell. It should be particularly noted that the transformed host carrot cells are grown in suspension.

In a further aspect, the present invention relates to a method for treating a subject, preferably a mammalian subject, having lysosomal storage disease by using exogenous recombinant lysosomal enzyme.

Disclosed and described, it is to be understood that this invention is not limited to the particular examples, process steps, and materials disclosed herein as such process steps and materials may vary somewhat. It is also to be understood that the terminology

used herein is used for the purpose of describing particular embodiments only and not intended to be limiting since the scope of the present invention will be limited only by the appended claims and equivalents thereof.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise.

The following examples are representative of techniques employed by the inventors in carrying out aspects of the present invention. It should be appreciated that while these techniques are exemplary of preferred embodiments for the practice of the invention, those of skill in the art, in light of the present disclosure, will recognize that numerous modifications can be made without departing from the spirit and intended scope of the invention.

Examples

Experimental procedures:

Plasmid vectors

* CE-T – Was constructed from plasmid CE obtained from Prof. Galili [United States Patent *5,367,110* November 22, (1994)]. Plasmid CE was digested with SalI.

The SalI cohesive end was made blunt-ended using the large fragment of DNA polymerase I. Then the plasmid was digested with PstI and ligated to a DNA fragment coding for the ER targeting signal from the basic endochitinase gene [*Arabidopsis thaliana*] ATGAAGACTAATCTTTTTCTCTTTCTCATCTTTTCA

CTTCTCCTATCATTATCCTCGGCCGAATTC, and vacuolar targeting signal from Tobacco chitinase A: GATCTTTTAGTCGATACTATG digested with SmaI and PstI.

* pGREENII - obtained from Dr. P. Mullineaux [Roger P. Hellens et al., (2000) Plant Mol. Bio. 42:819-832]. Expression from the pGREEN II vector is controlled by the 35S promoter from Cauliflower Mosaic Virus, the TMV (Tobacco Mosaic Virus) omega translational enhancer element and the octopine synthase terminator sequence from *Agrobacterium tumefaciens*.

cDNA

hGCD – obtained from ATCC (Accession No. 65696), GC-2.2 [GCS-2kb; lambda-EZZ-gamma3 *Homo sapiens*] containing glucosidase beta acid [glucocerebrosidase]. Insert lengths (kb): 2.20; Tissue: fibroblast WI-38 cell.

Construction of expression plasmid

The cDNA coding for hGCD (ATTC clone number 65696) was amplified using the forward: 5' CAGAATTCGCCCGCCCCTGCA 3' and the reverse: 5' CTCAGATCTTGGCGATGCCACA 3' primers. The purified PCR DNA product was digested with endonucleases EcoRI and BglII (see recognition sequences underlined in the primers) and ligated into an intermediate vector having an expression cassette E-T digested with the same enzymes. The expression cassette was cut and eluted from the intermediate vector and ligated into the binary vector pGREENII using restriction enzymes SmaI and XbaI, forming the final expression vector. Kanamycine resistance is conferred by the NPTII gene driven by the nos promoter obtained together with the pGREEN vector (Fig. 1B). The resulting expression cassette is presented by Fig. 1A.

The resulting plasmid was sequenced to ensure correct in-frame fusion of the signals using the following sequencing primers: 5' 35S promoter: 5' CTCAGAAGACCAGAGGGC 3', and the 3' terminator: 5' CAAAGCGGCCATCGTGC 3'.

Establishment of carrot callus and cell suspension cultures

Establishment of carrot callus and cell suspension cultures we preformed as described previously by Torres K.C. (Tissue culture techniques for horticular crops, p.p. 111, 169).

Transformation of carrot cells and isolation of transformed cells.

Transformation of carrot cells was preformed using *Agrobacterium* transformation by an adaptation of a method described previously [Wurtele, E.S. and Bulka, K. Plant Sci. 61:253-262 (1989)]. Cells growing in liquid media were used throughout the process instead of calli. Incubation and growth times were adapted for transformation of cells in liquid culture. Briefly, *Agrobacteria* were transformed with the pGREEN II vector by electroporation [den Dulk-Ra, A. and Hooykaas, P.J. (1995) Methods Mol. Biol. 55:63-72] and then selected using 30 mg/ml paromomycine antibiotic. Carrot cells were transformed with *Agrobacteria* and selected using 60 mg/ml of paromomycine antibiotics in liquid media.

Screening of transformed carrot cells for isolation of calli expressing high levels of GCD

14 days following transformation, cells from culture were plated on solid media at dilution of 3% packed cell volume for the formation of calli from individual clusters of cells. When individual calli reached 1-2 cm in diameter, the cells were homogenized in SDS sample buffer and the resulting protein extracts were separated on SDS-PAGE [Laemmli U., (1970) Nature 227:680-685] and transferred to nitrocellulose membrane (hybond C nitrocellulose, 0.45 micron. Catalog No: RPN203C From Amersham Life Science). Western blot for detection of GCD was performed using polyclonal anti hGCD antibodies (described herein below). Calli expressing significant levels of GCD were expanded and transferred to growth in liquid media for scale up, protein purification and analysis.

Preparation of polyclonal antibodies

75 micrograms recombinant GCD (Cerezyme™) were suspended in 3 ml complete Freund's adjuvant and injected to each of two rabbits. Each rabbit was given a

booster injection after two weeks. The rabbits were bled about 10 days after the booster injection and again at one week intervals until the antibody titer began to drop. After removal of the clot the serum was divided into aliquots and stored at -20°C.

Upscale culture growth in a bioreactor

An about 1cm (in diameter) callus of genetically modified carrot cells containing the rh-GCD gene was plated onto Murashige and Skoog (MS) 9cm diameter agar medium plate containing 4.4gr/l MSD medium (Duchefa), 9.9mg/l thiamin HCl (Duchefa), 0.5mg folic acid (Sigma) 0.5mg/l biotin (Duchefa), 0.8g/l Casein hydrolisate (Ducifa), sugar 30g/l and hormones 2-4 D (Sigma). The callus was grown for 14 days at 25°C.

Suspension cell culture was prepared by sub-culturing the transformed callus in a MSD liquid medium (Murashige & Skoog (1962) containing 0.2 mg/l 2,4-dicloroacetic acid), as is well known in the art. The suspension cells were cultivated in 250ml Erlenmeyer flask (working volume starts with 25ml and after 7 days increases to 50ml) at 25°C with shaking speed of 60rpm. Subsequently, cell culture volume was increased to 1L Erlenmeyer by addition of working volume up to 300ml under the same conditions. Inoculum of the small bio-reactor (10L) [see WO98/13469] containing 4L MSD medium, was obtained by addition of 400ml suspension cells derived from two 1L Erlenmeyer that were cultivated for seven days. After week of cultivation at 25°C with 1Lpm airflow, MDS medium was added up to 10L and the cultivation continued under the same conditions. After additional five days of cultivation, most of the cells were harvested and collected by passing the cell media through 80µ net. The extra medium was squeezed out and the packed cell cake was store at –70°C.

Further details of the bioreactor device may be found with regard to US Patent No. 6,391,638, issued on May 21 2002 and previously incorporated by reference.

Protein purification

In order to separate the medium from the insoluble GCD, frozen cell cake containing about 100g wet weight cells was thawed, followed by centrifugation of the thawed cells at 17000xg for 20min at 4°C. The insoluble materials and intact cells were washed by re-suspension in 100ml washing buffer (20mM sodium phosphate pH 7.2,

20mM EDTA), and then precipitated by centrifugation at 17000g for 20min at 4^0C. The rh-GCD (recombinant human GCD) was extracted and solubilized by homogenization of the pellet in 200ml extraction buffer (20mM sodium phosphate pH 7.2, 20mM EDTA, 1mM PMSF, 20mM ascorbic acid, 3.8g polyvinylpolypyrrolidone (PVPP), 1mM DTT and 1% Triton-x-100). The homogenate was then shaken for 30min at room temperature and clarified by centrifugation at 17000xg for 20min at 4^0C. The pellet was discarded and the pH of the supernatant was adjusted to pH 5.5 by addition of concentrated citric acid. Turbidity generated after pH adjustment was clarified by centrifugation under the same conditions described above.

Further purification was performed by chromatography columns procedure as follows: 200ml of clarified medium were loaded on 20ml strong cation exchange resin (Macro-Prep high-S support, Bio-Rad) equilibrated in 25mM sodium citrate buffer pH 5.5, packed in a XK column (2.6x20cm). The column was integrated with an AKTA (prime system (Amersham Pharmacia Biotech) that allowed to monitor the conductivity, pH and absorbency at 280nm. The sample was loaded at 20ml/min, afterwards the column was washed with equilibration buffer (25mM sodium citrate buffer pH 5.5) at flow rate of 12ml/min until UV absorbency reached the base line. Pre-elution of the rh-GCD was performed with equilibration buffer containing 200mM NaCl and the elution was obtained with equilibration buffer containing 600mM NaCl. Fractions collected during the run were monitored by enzyme activity assay, and tubes exhibiting enzymatic activity (in the elution peak) were pooled. Pooled samples were diluted (1:5) in water containing 5% ethanol and pH adjusted to 6.0 with NaOH. Sample containing the rh-GCD was applied on the second XK column (1.6x20cm) packed with 10ml of the same resin as in the previous column. The resin in this column was equilibrate with 20mM citrate buffer pH 6.0 containing 5% ethanol. Following the sample load the column was washed with the equilibration buffer and the GCD was eluted from the column by elution buffer (20mM citrate buffer pH 6.0, 5% ethanol and 1M NaCl). The fractions of the absorbent peak in the elution step were pooled and applied on a third column.

The final purification step was performed on a XK column (1.6x20cm) packed with 8ml hydrophobic interaction resin (TSK gel, Toyopearl Phenyl-650C, Tosoh Corp.). The resin was equilibrated in 10mM citrate buffer pH 6.0 containing 5%

ethanol. The GCD elution pool from the previous column was loaded at 6ml/min followed by washing with equilibration buffer until the UV absorbent reach the baseline. The pure GCD was eluted by 10mM citric buffer containing 50% ethanol, pooled and stored at -20^{0}C.

Determination of protein concentration

Protein concentrations in cell extracts and fractions were assayed by the method of Lowry/Bradford (Bio Rad protein assay) [Bradford, M., Anal. Biochem. (1976) 72:248] using a bovine serum albumin standard (fraction V Sigma). Alternatively, concentration of homogenous protein samples was determined by absorption at 280 nm, 1mg/ml=1.4 O.D$_{280}$. Purity was determined by 280/260nm ratio.

GCD enzyme activity assay

Enzymatic activity of GCD was determined using p-nitrophenyl-β-D-glucopyranoside (Sigma) as a substrate. Assay buffer contained 60mM phosphate-citrate buffer pH=6, 4mM β-mercaptoethanol, 1.3mM EDTA, 0.15% Triton X-100, 0.125% sodium taurocholate. Assay was preformed in 96 well ELISA plates, 0-50 microliter of sample were incubated with 250 microliter assay buffer and substrate was added to final concentration of 4mM. The reaction was incubated at 37^{0}C for 60min. Product (p-nitrophenyl; pNP) formation was detected by absorbance at 405nm. Absorbance at 405nm was monitored at t=0 and at the end point. After 60 min, 6 microliter of 5N NaOH were added to each well and absorbance at 405 nm was monitored again. Reference standard curve assayed in parallel, was used to quantitate concentrations of GCD in the tested samples [Friedman *et al.*, (1999) Blood, 93(9):2807-16].

Kinetic studies:

For kinetic studies, GCD activity was assayed as described by hereinabove with some modifications, using a fluorescent short-acyl-chain analogue of glucosylceramide, N-[6-[(7-nitrobenzo-2-oxa-1,3-diazol-4-yl)amino]hexanoyl]-*D erythro*-glucosylsphingosine (C6-NBD-*D-erythro*-GlcCer). C_6-NBD-GlcCer was synthesized by N-acylation of glucosylsphingosine using succinimidyl 6-7-nitrobenzo-2-oxa-1,3-diazol-4-yl) aminohexanoate as described by Schwarzmann and Sandhoff (1987). The assay was

performed using 0.2 μg of either Cerezyme® or plant GCD of the invention in a final volume of 200 μl MES buffer (50 mM, pH 5.5). Concentrations of C_6-NBD-GlcCer ranged from 0.25 to 100 μM. Reactions were allowed to proceed for 5 min at 37 ^{0}C, and were stopped by addition of 1.5 ml of chloroform/methanol (1:2, v/v) prior to extraction and analysis of the fluorescent lipids.

Biochemical analyses:

In gel proteolysis and mass spectrometry analysis

The stained protein bands in the gel were cut with a clean razor blade and the proteins in the gel were reduced with 10mM DTT and modified with 100 mM iodoacetamide in 10mM ammonium bicarbonate. The gel pieces were treated with 50% acetonitrile in 10 mM ammonium bicarbonate to remove the stain from the proteins following by drying the gel pieces. The dried gel pieces were rehydrated with 10% acetonitrile in 10 mM ammonium bicarbonate containing about 0.1 μg trypsin per sample. The gel pieces were incubated overnight at 37°C and the resulting peptides were recovered with 60% acetonitrile with 0.1% trifluoroacetate.

The tryptic peptides were resolved by reverse-phase chromatography on 0.1 X 300-mm fused silica capillaries (J&W, 100 micrometer ID) home-filled with porous R2 (Persepective). The peptides were eluted using a 80-min linear gradient of 5 to 95% acetonitrile with 0.1% acetic acid in water at flow rate of about 1 μl/min. The liquid from the column was electrosprayed into an ion-trap mass spectrometer (LCQ, Finnegan, San Jose, CA). Mass spectrometry was performed in the positive ion mode using repetitively full MS scan followed by collision induces dissociation (CID) of the most dominant ion selected from the first MS scan. The mass spectrometry data was compared to simulated proteolysis and CID of the proteins in the NR-NCBI database using the Sequest software [J. Eng and J. Yates, University of Washington and Finnegan, San Jose].

The amino terminal of the protein was sequenced on Peptide Sequencer 494A (Perkin Elmer) according to manufacture instructions.

GCD Uptake of peritoneal macrophages

Targeting and uptake of GCD to macrophages is known to be mediated by the Mannose/N-acetylglucosmine receptor and can be determined using thioglycolate-elicited peritoneal macrophages obtained from mice, as described by Stahl P. *and Gordon S.* [J. Cell Biol. (1982) 93(1):49-56]. Briefly, mice (female, strain C57-B6) were injected intraperitoneally with 2.5 ml of 2.4% Bacto-thioglycolate medium w/o dextrose (Difco Cat. No. 0363-17-2). After 4-5 days, treated mice were sacrificed by cervical dislocation and the peritoneal cavity rinsed with phosphate buffered saline. Cells were pelleted by centrifugation (1000xg 10 min) and were resuspended in DMEM (Beit Haemek, Israel) containing 10% fetal calf serum. Cells were then plated at 1-2x10^5 cell/well in 96-well tissue culture plates and incubated at 37^0C. After 90 minutes, non-adherent cells were washed out three times using PBS, and the adherent macrophages were incubated for 90 min at 37°C, in culture medium containing specified quantities of rhGCD, ranging from 0 to 40 micrograms in 200 microliter final volume, in the absence and presence of yeast mannan (2-10, 5 mg/ml). After incubation, medium containing excess rGCD was removed, and cells were washed three times with PBS and then lysed with lysis buffer (10mM Tris pH=7.3, 1mM $MgCl_2$, 0.5% NP-40 and protease inhibitors). The activity of rGCD taken up by the cells was determined by subjecting the cell lysates to *in vitro* glycosidase assay as described above.

EXAMPLE 1

CONSTRUCTION OF EXPRESSION PLASMID

This Example describes the construction of an exemplary expression plasmid, used with regard to the Examples below, in more detail.

The cDNA coding for hGCD (ATTC clone number 65696) was amplified using the forward: 5' CAGAATTCGCCCGCCCCTGCA 3' (also denoted by SEQ ID NO: 3) and the reverse: 5' CTCAGATCTTGGCGATGCCACA 3' (also denoted by SEQ ID NO: 4) primers.

The purified PCR DNA product was digested with endonucleases EcoRI and BglII (see recognition sequences underlined in the primers) and ligated into an intermediate vector having an expression cassette CE-T digested with the same enzymes. CE-T includes ER targeting signal MKTNLFLFLIFSLLLSLSSAEF (also

denoted by SEQ ID NO: 1) from the basic endochitinase gene [*Arabidopsis thaliana*], and vacuolar targeting signal from Tobacco chitinase A: DLLVDTM* (also denoted by SEQ ID NO: 4).

The expression cassette was cut and eluted from the intermediate vector and ligated into the binary vector pGREENII using restriction enzymes SmaI and XbaI, forming the final expression vector. Kanamycine resistance is conferred by the NPTII gene driven by the nos promoter together with the pGREEN vector (Fig. 1B). The resulting expression cassette is presented by Fig. 1A.

The resulting plasmid was sequenced to ensure correct in-frame fusion of the signals using the following sequencing primers:

Primer from the 5' 35S promoter: 5' CTCAGAAGACCAGAGGGC 3' (also denoted by SEQ ID NO: 5), and the 3' terminator: 5' CAAAGCGGCCATCGTGC 3' (also denoted by SEQ ID NO: 6). The verified cloned hGCD coding sequence is denoted by SEQ ID NO: 7.

EXAMPLE 2

TRANSFORMATION OF CARROT CELLS AND SCREENING FOR TRANSFORMED CELLS EXPRESSING rhGCD

This Example describes an exemplary method for transforming carrot cells according to the present invention, as used in the Examples below.

Transformation of carrot cells was performed by Agrobacterium transformation as described previously by [Wurtele and Bulka (1989) *ibid.*]. Genetically modified carrot cells were plated onto Murashige and Skoog (MS) agar medium with antibiotics for selection of transformants. As shown by Fig. 2, extracts prepared from arising calli were tested for expression of GCD by Western blot analysis using anti hGCD antibody, and were compared to Cerezyme standard (positive control) and extracts of non-transformed cells (negative control). Of the various calli tested, one callus (number 22) was selected for scale-up growth and protein purification.

The Western blot was performed as follows.

For this assay, proteins from the obtained sample were separated in SDS polyacrylamide gel electrophoresis and transferred to nitrocellulose. For this purpose, SDS polyacrylamide gels were prepared as follows. The SDS gels consist of a stacking

gel and a resolving gel (in accordance with Laemmli, UK 1970, Cleavage of structural proteins during assembly of the head of bacteriphage T4, Nature 227, 680-685). The composition of the resolving gels was as follows: 12% acrylamide (Bio-Rad), 4 microliters of TEMED (N,N,N',N'-tetramethylethylenediamine; Sigma catalog number T9281) per 10ml of gel solution, 0.1% SDS, 375 mM Tris-HCl, pH 8.8 and ammonium persulfate (APS), 0.1%. TEMED and ammonium persulfate were used in this context as free radical starters for the polymerization. About 20 minutes after the initiation of polymerization, the stacking gel (3% acrylamide, 0.1% SDS, 126 mM Tris-HCl, pH 6.8, 0.1% APS and 5 microliters of TEMED per 5ml of stacking gel solution) was poured above the resolving gel, and a 12 or 18 space comb was inserted to create the wells for samples.

The anode and cathode chambers were filled with identical buffer solution: Tris glycine buffer containing SDS (Biorad, catalog number 161-0772), pH 8.3. The antigen-containing material was treated with 0.5 volume of sample loading buffer (30ml glycerol (Sigma catalog number G9012), 9% SDS, 15 ml mercaptoethanol (Sigma catalog number M6250), 187.5 mM Tris-HCl, pH 6.8, 500 microliters bromophenol blue, all volumes per 100 ml sample buffer), and the mixture was then heated at 100 °C for 5 minutes and loaded onto the stacking gel.

The electrophoresis was performed at room temperature for a suitable time period, for example 45-60 minutes using a constant current strength of 50-70 volts followed by 45-60 min at 180-200 Volt for gels of 13 by 9 cm in size. The antigens were then transferred to nitrocellulose (Schleicher and Schuell, Dassel).

Protein transfer was performed substantially as described herein. The gel was located, together with the adjacent nitrocellulose, between Whatmann 3 MM filter paper, conductive, 0.5 cm-thick foamed material and wire electrodes which conduct the current by way of platinum electrodes. The filter paper, the foamed material and the nitrocellulose were soaked thoroughly with transfer buffer (TG buffer from Biorad, catalog number 161-0771, diluted 10 times with methanol and water buffer (20% methanol)). The transfer was performed at 100 volts for 90 minutes at 4°C.

After the transfer, free binding sites on the nitrocellulose were saturated, at 4 °C over-night with blocking buffer containing 1% dry milk (Dairy America), and 0.1%

Tween 20 (Sigma Cat P1379) diluted with phosphate buffer (Riedel deHaen, catalog number 30435). The blot strips were incubated with an antibody (dilution, 1:6500 in phosphate buffer containing 1% dry milk and 0.1% Tween 20 as above, pH 7.5) at 37 °C for 1 hour.

After incubation with the antibody, the blot was washed three times for in each case 10 minutes with PBS (phosphate buffered sodium phosphate buffer (Riedel deHaen, catalog number 30435)). The blot strips were then incubated, at room temperature for 1 h, with a suitable secondary antibody (Goat anti rabbit (whole molecule) HRP (Sigma cat # A-4914)), dilution 1:3000 in buffer containing 1% dry milk (Dairy America), and 0.1% Tween 20 (Sigma Cat P1379) diluted with phosphate buffer (Riedel deHaen, catalog number 30435)). After having been washed several times with PBS, the blot strips were stained with ECL developer reagents (Amersham RPN 2209).

After immersing the blots in the ECL reagents the blots were exposed to X-ray film FUJI Super RX 18x24, and developed with FUJI-ANATOMIX developer and fixer (FUJI-X fix cat# FIXRTU 1 out of 2). The bands featuring proteins that were bound by the antibody became visible after this treatment.

Upscale culture growth in bioreactors

Suspension cultures of callus 22 were obtained by sub-culturing of transformed callus in a liquid medium. Cells were cultivated in shaking Erlenmeyer flasks, until total volume was sufficient for inoculating the bioreactor (as described in Experimental procedures). The genetically modified transgenic carrot cells can be cultivated over months, and cell harvest can be obtained in cycling of 5 to 7 days (data not shown). At the seventh cultivation day, when the amount of rh-GCD production in carrot cell is at the peak, cells were harvested by passing of culture through 100mesh nets. It should be noted that cells may be harvested by means known in the art such as filtration or centrifugation. The packed cell cake, which provides the material for purification of h-GCD to homogeneity, can be stored at freezing temperature.

EXAMPLE 3

PURIFICATION OF RECOMBINANT ACTIVE hGCD PROTEIN FROM TRANSFORMED CARROT CELLS

Recombinant h-GCD expressed in transformed carrot cells was found to be bound to internal membranes of the cells and not secreted to the medium. Mechanically cell disruption leaves the rGCD bound to insoluble membrane debris (data not shown). rGCD was then dissolved using mild detergents, and separated from cell debris and other insoluble components. The soluble enzyme was further purified using chromatography techniques, including cation exchange and hydrophobic interaction chromatography columns as described in Experimental procedures.

In order to separate the medium from the insoluble GCD, frozen cell cake containing about 100g wet weight cells was thawed, followed by centrifugation at 17000xg for 20min at 4^0C. The insoluble materials and intact cells were washed by re-suspension in 100ml washing buffer (20mM sodium phosphate pH 7.2, 20mM EDTA), and precipitated by centrifugation at 17000g for 20min at 4^0C. The rGCD was extracted and solubilized by homogenization of the pellet in 200ml extraction buffer (20mM sodium phosphate pH 7.2, 20mM EDTA, 1mM PMSF, 20mM ascorbic acid, 3.8g polyvinylpolypyrrolidone (PVPP), 1mM DTT, 1% Triton-x-100 (Sigma)). The homogenate was shaken for 30min at room temperature and clarified by centrifugation at 17000g for 20min at 4^0C. The pellet was discarded and the pH of the supernatant was adjusted to pH 5.5 by addition of concentrated citric acid. Turbidity generated after pH adjustment was clarified by centrifugation under the same conditions described above.

Further purification was performed by chromatography columns as follows: in a first stage, 200ml of clarified extract were loaded on 20ml strong cation exchange resin (Macro-Prep high-S support, Bio-Rad) equilibrated in 25mM sodium citrate buffer pH 5.5, packed in a XK column (2.6x20cm). The column was integrated with an AKTA prime system (Amersham Pharmacia Biotech) that allowed to monitor the conductivity, pH and absorbency at 280nm. The sample was loaded at 20ml/min, afterwards the column was washed with equilibration buffer (25mM sodium citrate buffer pH 5.5) at flow rate of 12ml/min until UV absorbency reached the base line. Pre-elution of the rh-GCD was performed with equilibration buffer containing 200mM NaCl and the elution was obtained with equilibration buffer containing 600mM NaCl. Fractions collected

during the run were monitored by enzyme activity assay, and tubes exhibiting enzymatic activity (in the elution peak) were pooled. Pooled samples were diluted (1:5) in water containing 5% ethanol and pH adjusted to 6.0 with NaOH.

Fig. 3A represents a standard run of this purification stage. The fractions collected during the run were monitored by enzyme activity assay, as shown by Fig. 3B, and Fig. 3C shows coomassie-blue stain of elution fractions assayed for activity.

Elution fractions containing the rGCD was applied on a second XK column (1.6x20cm) packed with 10ml of the same resin as in the previous column, for a second purification stage. The resin in this column was equilibrated with 20mM citrate buffer pH 6.0 containing 5% ethanol. Following the sample load the column was washed with the equilibration buffer and the rGCD was eluted from the column by elution buffer (20mM citrate buffer pH 6.0, 5% ethanol and 1M NaCl). Fig. 3D represents a standard run of this purification stage. The fractions collected during the run were monitored by enzyme activity assay, as shown by Fig. 3E, and Fig. 3F shows a coomassie-blue stain of elution fractions assayed for activity.

The fractions of the absorbent peak in the elution step were pooled and applied on a third column, for a third purification stage. The third purification stage was performed on a XK column (1.6x20cm) packed with 8ml hydrophobic interaction resin (TSK gel, Toyopearl Phenyl-650C, Tosoh Corp.). The resin was equilibrated in 10mM citrate buffer pH 6.0 containing 5% ethanol. The GCD elution pool from the previous column was loaded at 6ml/min followed by washing with equilibration buffer until the UV absorbance reached the baseline. The pure GCD was eluted by 10mM citric buffer containing 50% ethanol, pooled and stored at -20^{0}C.

Fig. 4A represents a standard run of this purification stage. The fractions collected during the run were monitored by enzyme activity assay (Fig. 4B), and Fig. 4C shows coomassie-blue stain of elution fractions assayed for activity.

In a batch purification of cells that were processed, rGCD protein was purified to a level greater than 95%; if only the first and third stages are performed, purity is achieved at a level of about 80% (results not shown).

Biochemical analysis

To validate the identity of purified rhGCD, Mass-Spec Mass-Spec (MSMS) analysis was preformed. Results obtained showed 49% coverage of protein sequence that matched the predicted amino acid sequence, based on the DNA of the expression cassette, including the leader peptide and targeting sequences.

Characterization and Sequencing of prGCD: To further characterize the plant produced human recombinant GCD of the invention, the rhGCD was solubilized using Triton X-100, in the presence of an antioxidant, and purified to homogeneity by cation exchange and hydrophobic chromatography (Fig. 9a). Amino-acid sequencing of the plant produced human recombinant GCD of the invention demonstrated that the rhGCD sequence (SEQ ID NO: 15) corresponds to that of the human GCD (Swiss Prot P04062, protein ID AAA35873), and includes two additional amino acids (EF) at the N-terminus (designated -2 and -1 accordingly), derived from the linker used for fusion of the signal peptide, and an additional 7 amino acids at the C-terminus (designated 497-503) derived from the vacuolar targeting signal.

Immunodetection of the purified plant produced human recombinant GCD of the invention with anti-GCD polyclonal antibody was performed by Western blotting of the SDS-PAGE separated protein, along with Cerezyme ® protein (Fig. 9b), confirming antigenic identity of the plant produced and CHO-produced proteins.

Enzymatic activity of recombinant hGCD:

The activity of plant produced human recombinant GCD of the present invention was compared to that of Cerezyme®, using a fluorescent GlcCer analogue. Figure 12 shows that similar specific activities were obtained, with V_{max} values of 0.47 ± 0.08 Kmol C6-NBD-ceramide formed/min/mg protein for prGCD and 0.43 ± 0.06 for Cerezyme®, and similar K_m values (20.7 ± 0.7 KM for the GCD of the invention and 15.2 ± 4.8 KM for Cerezyme®). Thus, these kinetic studies show that the activity of the plant produced human recombinant GCD of the present invention is similar to that of the CHO expressed enzyme.

Uptake and activity of recombinant hGCD in peritoneal macrophages

To determine whether the rhGCD produced in carrot has been correctly glycosylated and can undergo uptake by target cells, and thus be useful for treatment of Gaucher's disease, the ability of the rhGCD to bind to and be taken up by macrophages was next assayed. Targeting of rhGCD to macrophages is mediated by the Mannose/N-acetylglucosamine (Man/GlcNAc) receptor and can be determined using thioglycolate-elicited peritoneal macrophages. As shown by Fig. 5, rGCD undergoes uptake by cells at a high level. Figure 5A shows uptake by cells of rGCD according to the present invention with regard to mannan concentration.

Figure 5A shows uptake at comparable levels with Cerezyme™ (this preparation was prepared to 80% purity with only the first and third stages of the purification process described above).

Figures 5B and 5C show that rGCD uptake is at a higher level than Cerezyme™, as this preparation was prepared to greater than 95% purity with all three stages of the purification process described above.

With regard to Figure 5C, clearly the percent of specific activity from total activity, inhibited by 4mg/ml mannan, is higher for the GCD of the present invention (rGCD or recombinant human GCD) than for the currently available product in the market as follows: GCD (CB-mix1, which is the rGCD of the present invention) – 75% Cerezyme – 65%.Furthermore, as shown by the figures, addition of mannan clearly inhibited binding of rGCD by the cells. At concentration of 2mg/ml of mannan, the binding of rGCD was inhibited by 50%.

These results show that even without remodeling of glycan structures, rhGCD expressed and purified from transformed carrot cells can undergo uptake to target macrophage cells specifically through Man/GlcNAc receptors. Moreover, this recombinant rhGCD is enzymatically active.

Figure 5D shows that the rhGCD is also recognized by an anti-GCD antibody in a Western blot; rGCD refers to the protein according to the present invention, while GCD standard (shown at 5, 10 and 25 ng per lane) is commercially purchased GCD (Cerezyme®).

EXAMPLE 4

TOXICOLOGY TESTING

The material obtained according to the above purification procedure was tested according to standard toxicology testing protocols (Guidance for Industry on Single Dose Acute Toxicity Testing for Pharmaceuticals, Center for Drug Evaluation and Research (CDER) PT 1 (61 FR 43934, August 26, 1996) and by ICH M3(M) Non-clinical Safety Studies for the Conduct of Human Clinical Trials for Pharmaceuticals CPMP/ICH/286/95 modification, Nov 16 2000).

Mice were injected as follows: An initial dose of 1.8 mg/kg (clinical dose) was followed by doses of 9 and 18 mg/kg. Testing groups included six mice (ICR CD-1; 3 males and 3 females) for receiving rGCD (in a liquid carrier featuring 25 mM citrate buffer, 150 mM NaCl, 0.01% Tween 80, 5% ethanol) according to the present invention, and another six mice for being treated with the carrier alone as a control group. The mice were then observed for 14 days and were euthanized.

In another study, vehicle solution alone, or doses of prGCD in multiples of 1, 5, or 10 times the standard clinical dose (60units/kg) were given to ICR (CD-1®) mice. The animals (6 per group, 3 males and 3 females), received the drug intravenously in a 10 ml/kg volume.

Both toxicity studies revealed no obvious treatment-related adverse reactions, no gross pathological findings, no changes in body weight and no mortality incidences observed even at the highest dose administered. Furthermore, blood samples taken from animals in the high-dose group, which had been administered with 10-fold the clinical dose, were tested for hematology and clinical chemistry. All hematology and clinical chemistry values were in normal ranges. In addition, the animals treated with the high dose were subjected to histopathological examination of the liver, spleen and kidney, and there were no macro or micro histopathological findings.

EXAMPLE 5

GLYCOSYLATION ANALYSIS

Analysis of glycan structures present on rGCD produced as described with regard to the previous Examples was performed. As described in greater detail below, results indicate that the majority of glycans contain terminal mannose residues as well

as high mannose structures. Advantageously, this high mannose product was found to be biologically active, and therefore no further steps were needed for its activation.

The following methods were used to determine the glycosylation structure of the recombinant hGCD produced according to the Examples given above. Briefly, the monosaccharide linkages for both N- and O-glycans were determined by using a hydrolysis and GC-MS strategy. This method estimates the linkage type of the carbohydrates to the peptide and the general monosaccharide composition of a glycoprotein. Based on prior knowledge and also the ratios between various monosaccharides, this method may suggest the types of glycans on the glycoprotein. This information is important to estimate the possible glycan structures present on the protein.

Another method featured oligosaccharide analysis of the N-glycan population. FAB-MS and MALDI-TOF MS were performed, following digestion of aliquots of the samples with trypsin and peptide N-glycosidase F (PNGaseF) and permethylation of the glycans. This method is used to detach and isolate N-linked carbohydrates from the enzymatically digested glycoprotein. The masses of the glycan populations in the isolated glycan mix are determined and their masses are compared with those of known structures from databases and in light of the monosaccharide composition analysis. The proposed structures are based also on the glycosylation patterns of the source organism.

Another method included analyzing the O-glycan population following reductive elimination of the tryptic and PNGase F treated glycopeptides, desalting and permethylation. O-glycans are not released by PNGase F, therefore, glycans remaining linked to peptides are most likely O-linked glycans. These glycans are then released by reductive elimination and their mass analyzed.

Monosaccharide composition analysis (summarized below) revealed a characteristic distribution of hexoses, hexosamines and pentoses characteristic of plant glycosylation. The ratios between GlcNac and Mannose, suggest that characteristic N-linked structures are the predominant glycan population.

Mass Spectrometric analysis of the N-glycans from hGCD produced as described above indicates that the predominant N-glycan population has the monosaccharide composition Pent.deoxyHex.Hex3.HexNAc2.

Materials and Methods

Analysis was performed using a combination of Gas Chromatography-Mass Spectrometry (GC-MS), Fast Atom Bombardment-Mass Spectrometry (FAB-MS) and Delayed Extraction-Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass-Spectrometry (DE-MALDI-TOF MS).

For oligosaccharide analysis, the N-glycan population was analysed by FAB-MS and MALDI-TOF MS following digestion of aliquots of the samples with trypsin and peptide N-glycosidase F (PNGaseF) and permethylation of the glycans. The O-glycan population was analysed following reductive elimination of the tryptic and PNGase F treated glycopeptides, desalting and permethylation.

The monosaccharide linkages for both N- and O-glycans were determined using a hydrolysis, derivatisation GC-MS strategy.

EXPERIMENTAL DESCRIPTION

Sample

The sample vials were received were given the unique sample numbers as follows (Table 1):

Table 1

Product	reference number
Glucocerebrosidase. Four tubes containing	62995
1ml of sample each at a stated	62996
concentration of 0.8mg/ml in 25mM	62997
Citrate Buffer pH6.0, 0.01% Tween 80	62998

The samples were stored between -10 and -30°C until required.

Protein chemistry

Dialysis of intact samples

One vial (containing 1ml of protein at a stated concentration of 0.8mg/ml) was injected into a Slide-A-Lyzer dialysis cassette (10kDa molecular weight cutoff) and

dialysed at 4°C over a period of 24 hours against water, the water being changed 3 times. Following dialysis the sample was removed form the cassette and lyophilised.

Trypsin digestion of the intact samples for oligosaccharide screening

The dialysed, lyophilised sample was resuspended in 50mM ammonium bicarbonate buffer adjusted to pH 8.4 with 10% aq. ammonia and digested with TPCK treated trypsin for 4 hours at 37°C according to SOPs B001 and B003. The reaction was terminated by placing in a heating block at 95°C for 2 minutes followed by lyophilisation.

Carbohydrate chemistry

Peptide N-Glycosidase A Digestion

The tryptically cleaved peptide/glycopeptide mixtures from the glycoprotein sample was treated with the enzyme peptide N-glycosidase A (PNGaseA) in ammonium acetate buffer, pH 5.5 at 37°C for 15 hours. The reaction was stopped by freeze-drying. The resulting products were purified using a C_{18} Sep-Pak cartridge.

Reductive elimination

The Sep-Pak fraction containing potential O-linked glycopeptides was dissolved in a solution of 10mg/ml sodium borohydride in 0.05M sodium hydroxide and incubated at 45°C for 16 hours. The reaction was terminated by the addition of glacial acetic acid.

Desalting of reductively eliminated material

Desalting using Dowex beads was performed according to SOP B022. The sample was loaded onto the column and eluted using 4ml of 5% *aq*. acetic acid. The collected fraction was lyophilised.

Permethylation of released carbohydrates

N-linked carbohydrates eluting in the 5% aq. acetic acid Sep-Pak fraction and potential O-linked glycans released by reductive elimination, were permethylated using

the sodium hydroxide (NaOH)/methyl iodide (MeI) procedure (SOP B018). A portion of the permethylated N-linked glycan mixture was analysed by FAB-MS and MALDI-TOF MS and the remainder was subjected to linkage analysis.

Linkage Analysis of the N-linked Carbohydrate

Derivatisation

The permethylated glycan sample mixtures obtained following tryptic and PNGase A digestion or reductive elimination were hydrolysed (2M TFA, 2 hours at 120°C) and reduced (sodium borodeuteride ($NaBD_4$) in 2M NH_4OH, 2 hours at room temperature, SOP B025). The borate produced on the decomposition of the borodeuteride was removed by 3 additions of a mixture of methanol in glacial acetic acid (90:10) followed by lyophilisation. The samples were then acetylated using acetic anhydride (1 hour at 100°C). The acetylated samples were purified by extraction into chloroform. The partially methylated alditol acetates were then examined by gas chromatography/mass spectrometry (GC/MS). Standard mixtures of partially methylated alditol acetates and a blank were also run under the same conditions.

Gas Liquid Chromatography/Mass Spectrometry (GC/MS)

An aliquot (1µl) of the derivatised carbohydrate samples dissolved in hexane, were analysed by GC/MS using a Perkin Elmer Turbomass Gold mass spectrometer with an Autosystem XL gas chromatograph and a Dell data system under the following conditions:

Gas Chromatography

Column: DB5

Injection: On-column

Injector Temperature: 40°C

Programme: 1 minute at 40°C then 70°C/minute to 100°C, held at 100°C for 1 minute, then 8°C/minute to 290°C, finally held at 290°C for 5 minutes.

Carrier Gas: Helium

Mass Spectrometry

Ionisation Voltage: 70eV

Acquisition Mode: Scanning

Mass Range: 35-450 Daltons

MS Resolution: Unit

Sugar analysis of intact glucocerebrosidase

Derivatisation

An aliquot equivalent to 500µg of glucocerebrosidase was lyophilised with 10µg of Arabitol as internal standard. This was then methanolysed overnight at 80°C and dried under nitrogen. Released monosaccharides were re-N-acetylated using a solution of methanol, pyridine and acetic anhydride, dried under nitrogen again and converted to their trimethylsilyl (TMS) derivatives according to SOP B023. The TMS derivatives were reduced in volume under nitrogen, dissolved in 2ml of hexane and sonicated for 3 minutes. The samples were then allowed to equilibrate at 4°C overnight. A blank containing 10µg of Arabitol and a standard monosaccharide mixture containing 10µg each of Fucose, Xylose, Mannose, Galactose, Glucose, N-acetylgalactosamine, N-acetylglucosamine, N-acetylneuraminic acid and Arabitol were prepared in parallel. The TMS derivatives were then examined by gas chromatography/mass spectrometry (GC/MS).

Gas Liquid Chromatography/Mass Spectrometry

(GC/MS)

An aliquot (1µl) of the derivatised carbohydrate sample dissolved in hexane, was analysed by GC/MS using a Perkin Elmer Turbomass Gold mass spectrometer with an Autosystem XL gas chromatograph and a Dell data system under the following conditions:

Gas Chromatography

Column: DB5

Injection: On-column

Injector Temperature: 40°C

Programme: 1 minute at 90°C then 25°C/minute to 140°C, 5°C/minute to 220°C, finally 10°C/minute to 300°C and held at 300°C for 5 minutes.

Carrier Gas: Helium

Mass Spectrometry

Ionisation Voltage: 70eV

Acquisition Mode: Scanning

Mass Range: 50-620 Daltons

MS Resolution: Unit

Delayed Extraction Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (DE-MALDI-MS) and Fast Atom Bombardment-Mass Spectrometry (FAB-MS)

MALDI-TOF mass spectrometry was performed using a Voyager STR Biospectrometry Research Station Laser-Desorption Mass Spectrometer coupled with Delayed Extraction (DE).

Dried permethylated glycans were redissolved in methanol:water (80:20) and analysed using a matrix of 2,5-dihydroxybenzoic acid. Bradykinin, Angiotensin and ACTH were used as external calibrants.

Positive Ion Fast Atom Bombardment mass spectrometric analyses were carried out on M-Scan's VG AutoSpecE mass spectrometer operating at Vacc = 8kV for 4500 mass range at full sensitivity with a resolution of approximately 2500. A Caesium Ion Gun was used to generate spectra operating at 30kV. Spectra were recorded on a VAX data system 3100 M76 using Opus software.

Dried permethylated glycans were dissolved in methanol and loaded onto a target previously smeared with 2-4µl of thioglycerol as matrix prior to insertion into the source.

In a second set of glycosylation analysis, similar methods were used to determine the glycosylation patterns, and to identify the major glycosylated products produced by the carrot cell suspension culture of the present invention:

Glycosylation patterns were analyzed by the Glycobiology Center of the National Institute for Biotechnology (Ben Gurion University, Beer Sheba, Israel) to

determine glycan structure and relative amounts using sequential digestion with various exoglycosidases. The plant GCD samples of the invention were run on SDS-PAGE and a 61KDa band was cut out and incubated with either PNGase A, or with trypsin followed by PNGase A to release the N-linked glycans. The glycans were fluorescently labeled with anthranilamide (2AB) and run on normal phase HPLC.

Sequencing of the labeled glycan pool was achieved by sequential digestion with various exoglycosidases followed by HPLC analysis. Retention times of individual glycans were compared to those of a standard partial hydrolysate of dextran giving a ladder of glucose units (GU). Unlabeled glycans were further purified and analyzed by MALDI mass spectrometry. Exoglycosidases used: Bovine kidney _-fucosidase (digests _1-6 and _1-3 core fucose, Prozyme), Jack bean mannosidase (removes _1-2, 6>3 mannose, Prozyme), Xanthomonas beta1,2-xylosidase (removes _ 1-2 xylose only after removal of _-linked mannose, Calbiochem).

Bovine testes -galactosidase (hydrolyses non-reducing terminal galactose _ 1-3 and _ 1-4 linkages, Prozyme), Streptococcus pneumoniae hexosaminidase (digest _1—2,3,4,6 GalNAc and GlcNAc, Prozyme). Glycosylation was further analyzed by M-Scan (Berkshire, England) using gas chromatography mass spectrometry (GC-MS), fast atom bombardment-mass spectrometry (FAB-MS), and delayed extraction-matrix assisted laser desorption ionization - time of flight mass-spectrometry (DE-MALDI-TOF MS). For oligosaccharide determination, the N-glycan population was analyzed by FAB-MS and MALDI-TOF MS, following digestion of samples with trypsin and PNGase A, and permethylation of the glycans. O-glycans were analyzed following reductive elimination of the tryptic and PNGase A-treated glycopeptides, desalting and permethylation.

The similarity of the N-glycans in different batches of prGCD was analyzed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD, a Dionex method) following digestion with trypsin and PNGase A, to obtain chromatographic profiles for oligosaccharides released from glycoproteins for the purpose of demonstrating consistency from batch to batch of prGCD. This procedure permits chromatographic comparison of oligosaccharide patterns in a qualitative and quantitative manner.

RESULTS AND DISCUSSION

TMS sugar analysis of Glucocerebrosidase

N-linked oligosaccharide screening

The intact glycoprotein was subjected to dialysis followed by trypsin digestion and the lyophilised products were digested using PNGase A and then purified using a C_{18} Sep-Pak. The 5% *aq*. acetic acid (N-linked oligosaccharide containing) fraction was permethylated and FAB mass spectra were obtained using a portion of the derivatised oligosaccharide in a low mass range for fragment ions and DE-MALDI-TOF mass spectra were obtained using a portion of the derivatised oligosaccharides in a high mass range for molecular ions.

Analysis of N-glycans from glucocerebrosidase

Table 1 lists the predominant fragment ions present in the FAB spectra and molecular ions present in the MALDI spectra. The molecular ion region (shown in Appendix III) contains a predominant signal at m/z 1505.8 (consistent with an $[M+Na]^+$ quasimolecular ion for a structure having the composition Pent.deoxyHex.Hex_3.$HexNAc_2$). A range of less intense quasimolecular ions were also detected consistent with complex and high mannose structures. The high mannose structures detected range in size from Hex_5.$HexNAc_2$ at m/z 1579.8 to Hex_8.$HexNAc_2$ at m/z 2193.0. The complex signals are produced from less extensively processed N-glycans such as m/z 1331.7 (consistent with an $[M+Na]^+$ quasimolecular ion for a structure having the composition Pent.Hex_3.$HexNAc_2$) or from larger N-glycans for example m/z 1751.0 (consistent with an $[M+Na]^+$ quasimolecular ion for a structure having the composition Pent.deoxyHex.Hex_3.$HexNAc_3$), m/z 2375.4 (consistent with an $[M+Na]^+$ quasimolecular ion for a structure having the composition Pent.deoxyHex_2.Hex_4.$HexNAc_4$) and m/z 2753.6 (consistent with an $[M+Na]^+$ quasimolecular ion for a structure having the composition Pent.deoxyHex_3.Hex_5.$HexNAc_4$).

The FAB mass spectrum provides information regarding antennae structures by virtual of fragment ions in the low mass region of the spectrum (data not shown).

Signals were detected identifying hexose (at m/z 219) and HexNAc (at m/z 260) as non-reducing terminal monosaccharides in the N-glycans.

Table 2: Masses observed in the permethylated spectra of Glucocerebrosidase (reference number 62996) following Tryptic and Peptide N-glycosidase A digestion

Signals observed (m/z)	Possible Assignment
Low Mass	
219	Hex^+
228	$HexNAc^+$ (- methanol)
260	$HexNAc^+$
High Mass	
1032.4	$Pent.Hex_3.HexNAc^+$
1171.5	$Hex_3.HexNAc_2OMe + Na^+$
1299.6	Elimination of fucose from m/z 1505.8
1331.6	$Pent.Hex_3.HexNAc_2OMe + Na^+$
1345.6	$deoxyHex.Hex_3.HexNAc_2OMe + Na^+$
1505.7	$Pent.deoxyHex.Hex_3.HexNAc_2OMe + Na^+$
1579.8	$Hex_5.HexNAc_2OMe + Na^+$
1709.9	$Pent.deoxyHex.Hex_4.HexNAc_2OMe + Na^+$
1750.9	$Pent.deoxyHex.Hex_3.HexNAc_3OMe + Na^+$
1783.9	$Hex_6.HexNAc_2OMe + Na^+$
1989.0	$Hex_7.HexNAc_2OMe + Na^+$
1997.0	$Pent.deoxyHex.Hex_3.HexNAc_4OMe + Na^+$
2027.0	Not assigned
2099.0	Not assigned
2130.0	$Pent.deoxyHex_2.Hex_4.HexNAc_3OMe + Na^+$
2193.1	$Hex_8.HexNAc_2OMe + Na^+$
2375.2	$Pent.deoxyHex_2.Hex_4.HexNAc_4OMe + Na^+$

Signals observed (m/z)	Possible Assignment
2753.4	Pent.deoxyHex$_3$.Hex$_5$.HexNAc$_4$OMe + Na^+

All masses in column one are monoisotopic unless otherwise stated. The mass numbers may not relate directly to the raw data as the software often assigns mass numbers to ^{13}C isotope peaks particularly for masses above 1700Da.

Linkage analysis of N-glycans from Glucocerebrosidase

Linkage analysis was performed on the N-linked carbohydrates released following PNGase A digestion, Sep-Pak purification and permethylation.

A complex chromatogram was obtained with some impurity peaks originating from the derivatising reagents. Comparison of the retention time and the spectra with standard mixtures allowed provisional assignments of the sugar containing peaks listed in Table 3.

Table 3: Retention times of the variously linked monosaccharides detected as their partially methylated alditol acetates in the GC-MS analysis of Glucocerebrosidase (reference number 62996) following Tryptic and Peptide N-glycosidase A digestion

Compounds Observed	Retention time (mins) Glucocerebrosidase (62996)
Terminal Xylose	10.41
Terminal Fucose	10.84
Terminal Mannose	12.29 (major)
Terminal Galactose	12.55

2-linked Mannose	13.40
4-linked Glucose	13.58
2,6-linked Mannose	14.91
3,6-linked Mannose	15.08
2,3,6-linked Mannose	15.87
4-linked GlcNAc	16.73
3,4-linked GlcNAc	17.59

4.3 **O-linked oligosaccharide screening**

Reductive elimination was carried out on the 60% 2-propanol fraction (potential O-linked glycopeptide fraction) from the Sep-Pak purification of Glucocerebrosidase following trypsin and PNGase A digestions. The sample was desalted following termination of the reaction and, after borate removal, was permethylated. FAB mass spectra were obtained using a portion of the derivatised oligosaccharide in a low mass range for fragment ions and DE-MALDI-TOF mass spectra were obtained using a portion of the derivatised oligosaccharides in a high mass range for molecular ions. No signals consistent with the presence of O-linked glycans were observed (data not shown).

Linkage Analysis of O-glycans from glucocerebrosidase

Linkage analysis was carried out on the products of reductive elimination after permethylation. No signals consistent with the presence of typical O-linked glycans were observed (data not shown).

Figure 6 shows some exemplary glycan structures as a comparison between GCD obtained from CHO (Chinese hamster ovary) cells, which are mammalian cells (Cerezyme™) and the GCD of the present invention, from carrot cells. As shown, remodeling of these structures is required to obtain exposed mannose residues for Cerezyme™. By contrast, such exposed mannose residues are directly obtained for the GCD obtained from plant cells according to the present invention, without requiring further manipulation, for example with glycosylases.

Figure 7 represents the main glycan structure found in rGCD. Figure 7 shows proposed structures of: a) the predominant oligosaccharide population found on hGC expressed in carrot cell suspension (1505.7m/z); b) typical N-linked core; c) Fucosylated plant N-linked core. N-linked glycans are coupled to the protein via-Aspargine and through the reducing end of the GlcNac (GN) residue on the right hand of the diagrams. N plant glycosylation patterns, Fucose residues may be part of the core structure, bound to the first GlcNac using an alpha(1-3) glycosidic bond, while mammalian structures typically use the alpha(1-6) glycosidic bond.

Figures 8A-8D show all possible structures for the N-glycans detected on the rGCD protein according to the present invention.

The dominant glycan structure that was identified is the core glycan structure found in most plant glycoproteins from pea, rice, maize and other edible plants. This structure contains a core xylose residue as well as a core alpha-(1,3)-fucose. Work done by Bardor et al (33) shows that 50% of nonallergic blood donors have specific antibodies for core xylose in their sera, and 25% have specific antibodies to core alpha-(1,3)-fucose.. However it is still to be studied whether such antibodies might introduce limitations to the use of plant-derived biopharmaceutical glycoproteins.

The minor glycan populations of the hGCD produced as described above were mainly high mannose structures Hex4HexNAc2 to Hex8HexNAc2. Among the complex structures exhibited structures such as Pent.deoxyHex2.Hex4.HexNAc3 and Pent.deoxyHex3.Hex5.HexNAc3. Pent.Hex3.HexNAc2 was detected in smaller proportions.

The major terminal monosaccharides are hexose (Mannose or Galactose) and N-acetylhexosamine, which is consistent with the presence of high mannose structures and partially processed complex structures.

With regard to O-linked oligosaccharide screening, no signals that are consistent with typical O-linked glycans were observed. GCD is known in the art to not have O-linked oligosaccharides, such that these results are consistent with the known glycosylation of GCD from other cell systems, including native GCD and recombinant GCD produced in mammalian culture systems. However, in the monosaccharide composition, signals consistent with Arabinose were detected.

An important point with regard to the present invention is that the hGCD protein N-glycan composition analysis showed that the majority of the N-glycans terminate with mannose residues. This agrees with the requirement for mannose terminating N-glycans assisting the uptake of therapeutic hGCD by the macrophage mannose receptor. However, neither native GCD nor recombinant GCD produced in mammalian cells is high mannose. Therefore, the present invention overcomes a significant drawback of commercially produced hGCD proteins, which is that these proteins are *modified* to terminate with mannose sugars, unlike the protein produced as described above.

Further glycosylation analysis was performed on a purified human recombinant glucocerebrosidase prepared in plant cells. Glycosylation was analyzed (Glycobiology Center of the National Institute for Biotechnology (Ben Gurion University, Beer Sheba, Israel) to determine glycan structure and the glycan quantitative ratio using sequential digestion with various exoglycosidases (see Methods, above). In this analysis, it was found that the N-linked glycans have a main core of two GlcNAc residues and a 1-4 linked mannose, attached to two additional mannose residues in _1-3 and _1-6 linkages. The additional residues found are shown in Fig. 10a, which presents all structures and their relative amounts based upon HPLC, enzyme array digests and MALDI. Fig. 10b shows the glycan structure of Cerezyme® before and after *in vitro* enzymatic processing. Notably, analysis of the glycan structures of the GCD of the invention revealed that >90% of the glycans were mannose-rich, bearing exposed, terminal mannose residues (Fig. 10a), whereas in the case of Cerezyme®, mannose residues are exposed only after a complex *in-vitro* procedure (Fig. 10b). The dominant glycan in the GCD of the invention is the core structure found in most glycoproteins purified from pea, rice, maize and other edible plants. This structure contains a core _-(1,2)-xylose residue as well as a core _-(1,3)-fucose (Fig. 10a). The DE-MALDI-MS data contained no signals consistent with typical O-linked glycans. Further analysis of the glycan profiles for the

GCD of the invention obtained from different production batches was performed in order to asses the batch-to-batch reproducibility of the GCD produced in the carrot cell system. As presented in Fig. 11, the population of glycans on plant GCD of the invention is highly reproducible between batches.

EXAMPLE 5a

EXPRESSION OF BIOLOGICALLY ACTIVE α-GALACTOSIDASE IN PLANT CELLS

Human α-galactosidase A, the lesion in the X-linked lysosomal storage disorder Fabry disease, has been sequenced and cloned. In order to test whether α-galactosidase A suitable for therapeutic use can be produced in plant cells, vectors including the human α-galactosidase A coding sequence targeted to the plant endoplasmic reticulum were expressed in plant cells, and polypeptide sequence, biological activity and glycan structure of the plant-derived, recombinant human α-galactosidase A was evaluated.

Human α-galactosidase A expression vectors: Vectors containing Human α-galactosidase A coding sequence and an N-terminal apple pectinase leader peptide (SEQ ID NO: 16-MALKTQLLWSFVVVFVVSFSTTSCSG), for targeting the translated protein to the plant endoplasmic reticulum (ER) secretory system, were constructed. Two different constructs were cloned, with different C-terminal sequences designed to sustain the translated protein in a specific cellular compartment. One construct (α-gal-vac, SEQ ID NO: 17) contained a C-terminal vacuolar targeting signal (DLLVDTM, SEQ ID NO: 4) designed for transport of the protein from the ER to the plant vacuole, where the protein is retained. A second construct (α-gal-KDEL, SEQ ID NO: 19) lacked the C-terminal vacuolar targeting signal, and contained a C-terminal ER retention sequence (KDEL, SEQ ID NO:23) designed to allow retrograde transport from the *cis*-Golgi back to the ER where the protein is retained (see Rayon et al. Journal of Experimental Botany, Vol. 49, No. 326, pp. 1463–1472, 1998; and Evron et al. 2007 FASEB J).

The α-gal-vac clone:

The human α-gal coding sequence was artificially synthesized by GENEART AG. (Regensburg; Germany)(SEQ ID NO:17). The α-gal-vac sequence includes the

apple pectinase leader (SEQ ID NO:16) (MALKTQLLWSFVVVFVVSFSTTSCSG), mature alpha-galactosidase sequence (SEQ ID NO:24) and a vacuolar retention signal (SEQ ID NO: 4). The synthetic gene is surrounded by restriction sites to facilitate sub-cloning.

The gene was cloned using NCOI and HindIII, into a vector developed by ICON genetics (Halle, Germany) for transient expression in *Nicotiana benthamiama* plants.

The α-gal-KDEL clone:

For constructing the clone with the C-terminal ER retention signal, the vacuolar signal was replaced with an ER retention signal by adding a phosphorylated linker (SEQ ID NOs: 21 and 22) to replace the BglII- HindIII fragment. The phosphorylated linker codes for the ER retention signal and has sticky ends compatible with the ends generated by the enzymes BglII and HindIII:

S E K D E L * *

GATCTTAGTGAGAAGGACGAGCTCTGATAA (SEQ ID NO: 21)

AATCACTCT TCC TGCTCGAGACTAT TTCGA (SEQ ID NO:22)

Sac I

To produce the α-gal-KDEL construct, the α-gal-vac construct (SEQ ID NO:17) was digested with BglII and HindIII and ligated with the above linker. Insertion of the linker was verified by restriction with SacI. The resultant construct was then cloned into the ICON vector as described herein, for transient expression in *N. benthamiama*.

Transient expression system in *N.benthamiana*

The use of plant viral vectors was chosen in this case as an alternative to transgenic plants, allowing for the rapid, high level transient expression of proteins in whole plants.

The protein of interest is expressed from a strong duplicated viral promoter such as the coat protein sub-genomic promoter. The system relies on transient amplification (agroinfection) of viral vectors delivered to a plant by *agrobacterium*. In agroinfection a plant functional promoter and RNA virus cDNA are transferred as T-DNA from *agrobacterium* into plant cells. The T-DNA is transcribed *in-planta* to generate biologically active viral RNAs that can initiate self replication

This approach allows the rapid assembly and expression of arrays of proteins variants. This approach is not only very versatile but also provides milligram quantities of proteins in just a few days.

Transfection of whole plants-

N. Benthamiana plants are germinated and grown in commercial mix soil (Givaat Ada, IL) supplemented with granular slow release fertilizer (Scott Marysville, OH) under a long day (16h light / 8h dark) light regime (50μE) at 24°C-25°C.

For the transient expression a 3 vector recombination system developed by ICON genetics (Weinbergweg, Germany) was used as described (Gleba et al., Vaccine 23 2042-2048, 2005) one of the vectors was inserted with α-galactosidase cDNA and the two other vectors containing genes for construction of the whole viral replicon (RdRp and Integrase), thus generating the biologically active viral RNA that can initiate self replication

Agrobacteria were transformed with α-galactosidase vectors containing plasmids using electroporation (2500V, 5msec) [den Dulk-Ra, A. and Hooykaas, P.J. (1995) Methods Mol. Biol. 55:63-72]. Plants were infiltrated with Agrobacteria containing the 3 ICON plasmids by vacuum infiltration with standard methods known in the art. Briefly, *N. benthamiana* Plants, 5-6 week old were infiltrated by immersing all aerial plant organs into a bacterial suspension and were placed in a vacuum chamber. A -0.8 bar vacuum was applied for 1 minute, followed by a quick return to atmospheric pressure. Plants were returned to the greenhouse for additional 5-7 days under the same conditions.

Protein purification:

Tobacco leaves were frozen and thereafter ground with a mortar and pestle. The ground leaves were resuspended in extraction buffer containing 20mMTris 20mM EDTA, 20mM ascorbic acid 1mMDTT, 1mMPMSF pH 7.2 in a 1:1 volume to weight ratio. Thereafter, cells were disrupted and homogenized. The suspension was further homogenized using a knife homogenizer. The cell suspension was passed through a micro-fluid cell disruptor and the resulting preparation was centrifuged. The pellet was discarded and the supernatant was treated with ammonium sulfate and centrifuged. The

pellet was then dissolved in citrate buffer (20mM pH 6) and the solution was further acidified to pH 5.5, centrifuged, and filtered (0.45 μM). The filtrate was loaded on an hydrophobic interaction chromatography column and eluted fractions were pooled and loaded on a cation exchange chromatography column. Eluted fractions were pooled and analyzed for catalytic activity

Western Blotting:

Western blot was performed to identify the α-galactosidase molecules from transformed tobacco plants by using polyclonal rabbit anti α-gal A antibody.

Protein transfer was performed substantially as described herein. Briefly, transfer from the gel to nitrocellulose was performed at 100 volts for 90 minutes at 4°C. After the transfer, the blot was blocked with blocking buffer (1% dry milk, 0.1% Tween 20 (Sigma Cat P1379) in phosphate buffer). Blots were then immune detected by incubation with antibody, washed, and reacted with a suitable secondary antibody (Jackson-Labs HRP conjugated Goat anti Rabbit Ab). Blots were then developed with ECL developer reagents (Amersham RPN 2209), and autoradiography used for visualization.

Determination of active α-galactosidase enzyme:

The level of active plant α-galactosidase A was determined against a calibration curve of the activity of the commercial *α-galactosidase Fabrazyme* (Genzyme, Cambridge, Mass) plotted for the concentration range of 200-12.5 ng/ml. Activity was determined using *p*-nitrophenyl-α-D-galactopyranoside (Sigma) as a hydrolysis substrate. Assay buffer contained 20 mM citric acid 30mM sodium phosphate 0.1 % BSA and 0.67 % ethanol at pH 4.6. Assay was performed in 96 well ELISA plates (Greiner # 655061,96W), 50 microliter of sample were incubated with 150 microliter assay buffer and 30 microliter substrate was added to final concentration of 8mM. The reaction mixture was incubated at 37^0C for 90 min. and results were plotted against the calibration results. Product (p-nitrophenyl; pNP) formation was detected by absorbance at 405nm. Absorbance at 405nm was monitored at t=0 and at the end point. After 90 min, 100 microliter of 1.98 M Sodium carbonate were added to each well and absorbance at 405 nm was monitored again.

Kinetic studies:

To determine the Km, the concentration of p-nitrophenyl; pNP (Sigma) was varied in the range of 1000 μM to 45000 μM. Reaction mixtures containing 25 ng/mL of α-galactosidase and varying concentrations of the substrate were allowed to react for time periods ranging from 85 to 105 minutes at 37 °C. The reaction samples were quenched with saturated sodium carbonate and the absorbance of the p-nitrophenol product was detected at 430 nm.

Biochemical analyses

Tryptic digestion of protein bands from PAGE was effected by the Smoler Proteomics Center (Technion, Haifa, IL). Briefly, the stained protein bands in the gel were cut with a clean razor blade and the proteins in the gel were reduced with 10mM DTT and modified with 100 mM iodoacetamide in 10mM ammonium bicarbonate. The gel pieces were treated with 50% acetonitrile in 10 mM ammonium bicarbonate to remove the stain from the proteins following by drying the gel pieces. The dried gel pieces were rehydrated with 10% acetonitrile in 10 mM ammonium bicarbonate containing about 0.1 μg trypsin per sample. The gel pieces were incubated overnight at 37°C and the resulting peptides were recovered with 60% acetonitrile with 0.1% trifluoroacetate.

The tryptic peptides were resolved by reverse-phase chromatography on 0.1 X 300-mm fused silica capillaries (J&W, 100 micrometer ID) home-filled with porous R2 (Persepective). The peptides were eluted using a 80-min linear gradient of 5 to 95% acetonitrile with 0.1% acetic acid in water at flow rate of about 1 μl/min. The liquid from the column was electrosprayed into an ion-trap mass spectrometer (LTQ Orbitrap, Waltham, MA). Mass spectrometry was performed in the positive ion mode using repetitively full MS scan followed by collision induces dissociation (CID) of the most dominant ion selected from the first MS scan. The mass spectrometry data was compared to simulated proteolysis and CID of the proteins in the NR-NCBI database using the Sequest software [J. Eng and J. Yates, University of Washington and Finnegan, San Jose].

The amino terminal of the protein was sequenced on Peptide Sequencer 494A (Perkin Elmer) according to manufacture instructions.

MALDI-TOF:

MALDI-TOF mass spectrometry was performed using a MALDI TOF TOF 4700 (Applied Biosystems) according to methods known in the art in Smoler Proteomics Center (Technion, Haifa, IL).

Gel Filtration:

Gel filtration chromatography separates proteins on the basis of size. Molecules move through porous beads, diffusing into the beads when smaller molecules diffuse further into the pores of the beads and therefore move through the space more slowly, while larger molecules enter less or not at all and thus move through the space more quickly. Both molecular weight and three dimensional shape contribute to the degree of retention. Plant α-galactosidase samples were resuspended in analysis buffer (50mM Sodium Phosphate , pH=6.0) the flow rate was 0.5ml/min and 100 µg of sample were lowded onto the Column (-TSK Gel- 2000, Tosoh Bioscience ,San Francisco, CA)

α-galactosidase uptake in fibroblasts:

Targeting and uptake of α-galactosidase was tested on human fibroblasts originating from Fabry Patients (Cat. ID GM02775 Cornell Institute). Fibroblasts were cultured in DMEM medium (cat. D5546, Sigma) supplemented with 12 % FBS, 5ml L-Glutamine, 5 ml MEM Eagle vitamin solution 10 ml MEM amino acid solution 5 ml MEM Eagle non essential amino acid solution and 5 ml Pen-Strep solution, all supplements from Biological Indusries (Beit Haemek, IL). Cells were incubated with 300 ug/ml plant α-galactosidase A in PBS supplemented with 12 % FBS for 4 hrs, then washed and lysed (20 mM Phosphate buffer pH 6.8, 0.1 % Triton+ Protease inhibitors cocktail (Sigma P-2714), by two cycles of freeze-thawing). 20 ul of samples were loaded on 12 % SDS gels and analyzed by Western blotting (see above).

SDS-PAGE: Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE) separates proteins primarily by their molecular weight. In addition, this technique provides a large amount of information about the purity and composition of proteins. The molecular weight identity and the protein impurity pattern of α-galactosidase produced from tobacco plants were examined by SDS-PAGE analysis

using Coomassie Brilliant blue staining, according to standard gel separation protocols. Briefly, the SDS gels consist of a stacking gel (3%) and a resolving gel (12%). Running buffer was Tris/SDS, pH 8.3, loading buffer glycerol-Tris-mercaptoethanol, pH 6.8.

Results:

Figure 13A and 13B shows the characterization of molecular weight of plant-derived recombinant α-galactosidase, expressed and produced in tobacco plants by the method of the instant invention by gel filtration (Fig. 13A) and mass spectrometry (Fig. 13B). The mass spectrometry profile shows that the estimate of the molecular weight of plant-expressed α-galactosidase consists of several populations in the range of 48-52KDa. Since alpha Gal is a non-covalent dimer the energy of the MALDI-TOF causes the dissociation of the dimer to the monomers. This molecular weight reflects 407 amino acids contributing 46.3 kDa and the addition of the glycan structures for the remaining molecular weight. Mass spectrometry confirms the protein is 48.6 kDa, and these results are well in the range of the molecular weight of native human α-galactosidase (about 51 kDa). The gel filtration calibration curve shows the molecular weight corresponding to the retention time of the main peak of plant α-galactosidase (18.41 min.) is 76.56 kDa, suggesting a dimer. The very small peak at 20.403 min corresponds according to the calibration curve to a monomer (43.27 kDa). Since the gel filtration analysis is conducted at mild conditions it enables observing the protein at its dimer form.

Resolution of the recombinant α-galactosidase by PAGE analysis revealed two main bands (Fig. 14A), suggesting a difference in glycosylation of the mature recombinant enzyme. Sequencing of the isolated bands from PAGE indicated that this was indeed the case, as the regions of the polypeptide available for sequencing (i.e. not masked by glycans), although not located identically in the two bands, displayed 100% identity where overlapping (see Fig. 14B, sequenced portions in red). Fig 14B (Lower band) shows a complete sequence identification of a deglycosylated α-galactosidase [deglycosylation was effected using PNGase F (Sigma)]. The sequences in red indicate previously identified sequences. Green sequences indicate sequences available to sequencing following glycan removal. Black sequences indicate as yet unidentified sequences. Native glycosylation sites are depicted in yellow highlight. The C terminal

KDEL was verified using an anti KDEL antibody (Santa Cruz, CA). Taken together, these results show the α-galactosidase protein expressed by the methods of the present invention is identical to the expected cloned sequence.

Plant-expressed human recombinant α-galactosidase is antigenically identical to native recombinant α-galactosidase:

Further verification of the suitability of the constructs of the invention for accurate expression of human lysosomal enzymes in plant recombinant systems was provided by immune detection of the plant-expressed α-galactosidase on a Western blot. Fig. 15 shows that polypeptides expressed in tobacco plants from both the α-gal-vac (lane "vac") and α-gal-KDEL (lane "KDEL") constructs included a fraction detected by rabbit polyclonal antibody raised against the polypeptide fragment as set forth between amino acids 326 and 429 of native human α-galactosidase. Control plants, transformed with GFP (lane "GFP"), failed to produce any immune-reactive bands.

Kinetic analysis of recombinant α-galactosidase:

In order to evaluate the suitability of plant-expressed human recombinant α-galactosidase, purified recombinant α-galactosidase from tobacco plants was subjected to kinetic analysis, and Km and Vmax values were determined. Figs. 16A and 16B show the kinetics of recombinant α-galactosidase (red symbols) compared to those of commercially available recombinant human α-galactosidases Fabrazyme ® (black symbols) and Replagal ®. (blue symbols). Calculation of the Km and Vmax show that the recombinant α-galactosidase of the invention, targeted to and expressed in the ER, has very similar Km and Vmax parameters as the commercial enzymes, having a higher Vmax and lower Km than Replagal ®, and a greater V max and slightly higher Km than Fabrazyme ®, indicating accurate expression and processing of the polypeptide in the plant, and catalytic activity suitable for clinical applications.

Recombinant plant-expressed human α-galactosidase is stable in a wide range of temperatures:

In order to further evaluate the recombinant human α-galactosidase expressed and purified from plants according to one embodiment of the invention, stability of the

polypeptide at a range of temperatures (4 °C to 37 °C) was tested. Figs. 17A and 17B show that the plant-expressed recombinant human α-galactosidase (Plant a-Gal) did not undergo any alteration in electrophoretic mobility, and was as stable as, if not more stable than commercially available recombinant α-galactosidase from mammalian cells (Replagal®), at all temperatures tested. The stability of the recombinant α-galactosidase was evident whether incubated in activity buffer (Fig. 17A) or cell media buffer (Fig. 17B).

Recombinant plant-expressed α-galactosidase is actively taken up in human fibroblasts:

To determine whether the recombinant α-galactosidase produced in tobacco can undergo uptake by target cells, and thus be useful for treatment of Fabry disease, the ability of the recombinant human α-galactosidase to bind to and be taken up by fibroblasts was next assayed.. As shown in Fig. 18, recombinant α-galactosidase undergoes uptake by cells (lanes "plant αGalA" show immunodetection of alpha Gal in 20 ul of samples taken from fibroblast lysate run together with 20ng of the commercial recombinant alpha Gal run as a standard (Replagal, first lane from the right). In between are molecular weight marker ladders.

These results show that even without remodeling of glycan structures, recombinant α-galactosidase expressed and purified from transformed tobacco plants can undergo uptake to target α-galactosidase- deficient fibroblast cells. Moreover, the recombinant α-galactosidase is enzymatically active.

Glycan profile of plant-expressed recombinant human α-galactosidase:

Analysis of glycan structures present on human α-galactosidase produced as described with regard to the previous Examples was performed. As described in greater detail below, results indicate that the majority of glycans contain terminal mannose residues as well as high mannose structures. Advantageously, this high mannose product was found to be biologically active, and therefore no further steps were needed for its activation.

When the PAGE-separated band identified as human α-galactosidase was sequenced following trypsin digest, fluorescent labeling of glycans and sequential

digestion with exoglycosidases BKF, JBM, XYL and JBH, followed by HPLC, a characteristic pattern of glycosylation is discerned (Fig. 19, 87 minutes). Glycan structures having exposed mannose predominate,

Monosaccharide composition analysis (see Fig. 19) revealed a distribution of hexoses, hexosamines and pentoses characteristic of plant glycosylation. The ratios between GlcNac and Mannose, suggest that characteristic N-linked structures are the predominant glycan population.

EXAMPLE 6

TREATMENT WITH THE PRESENT INVENTION

The recombinant protein produced according to the present invention preferably comprises a suitably glycosylated protein produced by a plant cell culture, which is preferably a lysosomal enzyme for example, and/or a high mannose glycosylated protein.

According to preferred embodiments herein, the protein produced according to the present invention is suitable for treatment of a lysosomal-associated disease, such as a lysosomal storage disease for example.

The method of treatment optionally and preferably comprises: (a) providing a recombinant biologically active form of lysosomal enzyme purified from transformed plant root cells, and capable of efficiently targeting cells abnormally deficient in the lysosomal enzyme. This recombinant biologically active enzyme has exposed terminal mannose residues on appended oligosaccharides; and (b) administering a therapeutically effective amount of the recombinant biologically active lysosomal enzyme, or of composition comprising the same to the subject. In a preferred embodiment, the recombinant high mannose lysosomal enzyme used by the method of the invention may be produced by the host cell of the invention. Preferably, this host cell is a carrot cell.

By "mammalian subject" or "mammalian patient" is meant any mammal for which gene therapy is desired, including human, bovine, equine, canine, and feline subjects, most preferably, a human subject.

It should be noted that the term "treatment" also includes amelioration or alleviation of a pathological condition and/or one or more symptoms thereof, curing such a condition, or preventing the genesis of such a condition.

In another preferred embodiment, the lysosomal enzyme used by the method of the invention may be a high mannose enzyme comprising at least one oligosaccharide chain having an exposed mannose residue. This recombinant enzyme can bind to a mannose receptor on a target cell in a target site within a subject. More preferably, this recombinant lysosomal enzyme has increased affinity for these target cell, in comparison with the corresponding affinity of a naturally occurring lysosomal enzyme to the target cell. Therefore, each dose is dependent on the effective targeting of cells abnormally deficient in GCD and each dose of such form of GCD is substantially less than the dose of naturally occurring GCD that would otherwise be administered in a similar manner to achieve the therapeutic effect.

According to preferred embodiments of the present invention, the protein is suitable for the treatment of lysosomal storage diseases, such that the present invention also comprises a method for treating such diseases. Lysosomal storage diseases are a group of over 40 disorders which are the result of defects in genes encoding enzymes that break down glycolipid or polysaccharide waste products within the lysosomes of cells. The enzymatic products, e.g., sugars and lipids, are then recycled into new products. Each of these disorders results from an inherited autosomal or X-linked recessive trait which affects the levels of enzymes in the lysosome. Generally, there is no biological or functional activity of the affected enzymes in the cells and tissues of affected individuals. In such diseases the deficiency in enzyme function creates a progressive systemic deposition of lipid or carbohydrate substrate in lysosomes in cells in the body, eventually causing loss of organ function and death. The genetic etiology, clinical manifestations, molecular biology and possibility of the lysosomal storage diseases are detailed in Scriver et al. [Scriver et al. eds., The Metabolic and Molecular Basis of Inherited Disease, 7th Ed., Vol. II, McGraw Hill, (1995)].

Examples of lysosomal storage diseases (and their associated deficient enzymes) include but are not limited to Fabry disease (α-galactosidase), Farber disease (ceramidase), Gaucher disease (glucocerebrosidase), G_{ml} gangliosidosis (ß-galactosidase), Tay-Sachs disease (ß-hexosaminidase), Niemann-Pick disease (sphingomyelinase), Schindler disease (α.-N-acetylgalactosaminidase), Hunter syndrome (iduronate-2-sulfatase), Sly syndrome (ß-glucuronidase), Hurler and

Hurler/Scheie syndromes (iduronidase), and I-Cell/San Filipo syndrome (mannose 6-phosphate transporter).

Gaucher disease is the most common lysosomal storage disease in humans, with the highest frequency encountered in the Ashkenazi Jewish population. About 5,000 to 10,000 people in the United States are afflicted with this disease [Grabowski, Adv. Hum. Genet. 21:377-441(1993)]. Gaucher disease results from a deficiency in glucocerebrosidase (hGCD; glucosylceramidase). This deficiency leads to an accumulation of the enzyme's substrate, glucocerebroside, in reticuloendothelial cells of the bone marrow, spleen and liver, resulting in significant skeletal complications such as bone marrow expansion and bone deterioration, and also hypersplenism, hepatomegaly, thrombocytopenia, anemia and lung complications [Grabowski, (1993) ibid.; Lee, Prog. Clin. Biol. Res. 95:177-217 (1982)].

More specifically, the lysosomal enzyme used by the method of the invention may be selected from the group consisting of glucocerebrosidase (GCD), acid sphingomyelinase, hexosaminidase, α-N-acetylgalactosaminidise, acid lipase, α-galactosidase, glucocerebrosidase, α-L-iduronidase, iduronate sulfatase, α-mannosidase or sialidase. Preferably, where the treated disease is Gaucher's disease, the lysosomal enzyme used by the method of the invention is glucocerebrosidase (GCD).

The protein of the present invention can be used to produce a pharmaceutical composition. Thus, according to another aspect of the present invention there is provided a pharmaceutical composition which includes, as an active ingredient thereof, a protein and a pharmaceutical acceptable carrier. As used herein a "pharmaceutical composition" refers to a preparation of one or more of the active ingredients described herein, such as a recombinant protein, with other chemical components such as traditional drugs, physiologically suitable carriers and excipients. The purpose of a pharmaceutical composition is to facilitate administration of a protein or cell to an organism. Pharmaceutical compositions of the present invention may be manufactured by processes well known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

In a preferred embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the

U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. Hereinafter, the phrases "physiologically suitable carrier" and "pharmaceutically acceptable carrier" are interchangeably used and refer to an approved carrier or a diluent that does not cause significant irritation to an organism and does not abrogate the biological activity and properties of the administered conjugate.

The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E. W. Martin. Such compositions will contain a therapeutically effective amount of the protein, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should be suitable for the mode of administration.

Herein the term "excipient" refers to an inert substance added to a pharmaceutical composition to further facilitate processes and administration of the active ingredients. Examples, without limitation, of excipients include calcium carbonate, calcium phosphate, various sugars and types of starch, cellulose derivatives, gelatin, vegetable oils and polyethylene glycols.

Further techniques for formulation and administration of active ingredients may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition, which is incorporated herein by reference as if fully set forth herein.

The pharmaceutical compositions herein described may also comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include, but are not limited to, calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin and polymers such as polyethylene glycols.

Suitable routes of administration may, for example, include oral, rectal, transmucosal, transdermal, intestinal or parenteral delivery, including intramuscular, subcutaneous and intramedullary injections as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more pharmaceutically acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active ingredients into preparations which, can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the active ingredients of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hank's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the active ingredients can be optionally formulated through administration of the whole cells producing a protein according to the present invention, such as GCD or α-galactosidase for example. The active ingredients can also be formulated by combining the active ingredients and/or the cells with pharmaceutically acceptable carriers well known in the art. Such carriers enable the active ingredients of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for oral ingestion by a patient. Pharmacological preparations for oral use can be made using a solid excipient, optionally grinding the resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol,

or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carbomethylcellulose; and/or physiologically acceptable polymers such as polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, titanium dioxide, lacquer solutions and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active ingredient doses.

Pharmaceutical compositions, which can be used orally, include push-fit capsules made of gelatin as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules may contain the active ingredients in admixture with filler such as lactose, binders such as starches, lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active ingredients may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for the chosen route of administration.

For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the active ingredients for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from a pressurized pack or a nebulizer with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichloro-tetrafluoroethane or carbon dioxide. In the case of a pressurized aerosol, the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the active ingredient and a suitable powder base such as lactose or starch.

The active ingredients described herein may be formulated for parenteral administration, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multidose containers with optionally, an added preservative. The compositions may be suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical compositions for parenteral administration include aqueous solutions of the active preparation in water-soluble form. Additionally, suspensions of the active ingredients may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acids esters such as ethyl oleate, triglycerides or liposomes. Aqueous injection suspensions may contain substances, which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the active ingredients to allow for the preparation of highly concentrated solutions.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, pharmaceutical compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The pharmaceutical compositions of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with anions such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The active ingredients of the present invention may also be formulated in rectal compositions such as suppositories or retention enemas, using, e.g., conventional suppository bases such as cocoa butter or other glycerides.

The pharmaceutical compositions herein described may also comprise suitable solid of gel phase carriers or excipients. Examples of such carriers or excipients include, but are not limited to, calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin and polymers such as polyethylene glycols.

The topical route is optionally performed, and is assisted by a topical carrier. The topical carrier is one which is generally suited for topical active ingredient administration and includes any such materials known in the art. The topical carrier is selected so as to provide the composition in the desired form, e.g., as a liquid or non-liquid carrier, lotion, cream, paste, gel, powder, ointment, solvent, liquid diluent, drops and the like, and may be comprised of a material of either naturally occurring or synthetic origin. It is essential, clearly, that the selected carrier does not adversely affect the active agent or other components of the topical formulation, and which is stable with respect to all components of the topical formulation. Examples of suitable topical carriers for use herein include water, alcohols and other nontoxic organic solvents, glycerin, mineral oil, silicone, petroleum jelly, lanolin, fatty acids, vegetable oils, parabens, waxes, and the like. Preferred formulations herein are colorless, odorless ointments, liquids, lotions, creams and gels.

Ointments are semisolid preparations, which are typically based on petrolatum or other petroleum derivatives. The specific ointment base to be used, as will be appreciated by those skilled in the art, is one that will provide for optimum active ingredients delivery, and, preferably, will provide for other desired characteristics as well, e.g., emolliency or the like. As with other carriers or vehicles, an ointment base should be inert, stable, nonirritating and nonsensitizing. As explained in Remington: The Science and Practice of Pharmacy, 19th Ed. (Easton, Pa.: Mack Publishing Co., 1995), at pages 1399-1404, ointment bases may be grouped in four classes: oleaginous bases; emulsifiable bases; emulsion bases; and water-soluble bases. Oleaginous ointment bases include, for example, vegetable oils, fats obtained from animals, and semisolid hydrocarbons obtained from petroleum. Emulsifiable ointment bases, also known as absorbent ointment bases, contain little or no water and include, for example,

hydroxystearin sulfate, anhydrous lanolin and hydrophilic petrolatum. Emulsion ointment bases are either water-in-oil (W/O) emulsions or oil-in-water (O/W) emulsions, and include, for example, cetyl alcohol, glyceryl monostearate, lanolin and stearic acid. Preferred water-soluble ointment bases are prepared from polyethylene glycols of varying molecular weight; again, reference may be made to Remington: The Science and Practice of Pharmacy for further information.

Lotions are preparations to be applied to the skin surface without friction, and are typically liquid or semiliquid preparations, in which solid particles, including the active agent, are present in a water or alcohol base. Lotions are usually suspensions of solids, and may comprise a liquid oily emulsion of the oil-in-water type. Lotions are preferred formulations herein for treating large body areas, because of the ease of applying a more fluid composition. It is generally necessary that the insoluble matter in a lotion be finely divided. Lotions will typically contain suspending agents to produce better dispersions as well as active ingredients useful for localizing and holding the active agent in contact with the skin, e.g., methylcellulose, sodium carboxymethylcellulose, or the like.

Creams containing the selected active ingredients are, as known in the art, viscous liquid or semisolid emulsions, either oil-in-water or water-in-oil. Cream bases are water-washable, and contain an oil phase, an emulsifier and an aqueous phase. The oil phase, also sometimes called the "internal" phase, is generally comprised of petrolatum and a fatty alcohol such as cetyl or stearyl alcohol; the aqueous phase usually, although not necessarily, exceeds the oil phase in volume, and generally contains a humectant. The emulsifier in a cream formulation, as explained in Remington, supra, is generally a nonionic, anionic, cationic or amphoteric surfactant.

Gel formulations are preferred for application to the scalp. As will be appreciated by those working in the field of topical active ingredients formulation, gels are semisolid, suspension-type systems. Single-phase gels contain organic macromolecules distributed substantially uniformly throughout the carrier liquid, which is typically aqueous, but also, preferably, contain an alcohol and, optionally, an oil.

Various additives, known to those skilled in the art, may be included in the topical formulations of the invention. For example, solvents may be used to solubilize certain active ingredients substances. Other optional additives include skin permeation

enhancers, opacifiers, anti-oxidants, gelling agents, thickening agents, stabilizers, and the like.

The topical compositions of the present invention may also be delivered to the skin using conventional dermal-type patches or articles, wherein the active ingredients composition is contained within a laminated structure, that serves as a drug delivery device to be affixed to the skin. In such a structure, the active ingredients composition is contained in a layer, or "reservoir", underlying an upper backing layer. The laminated structure may contain a single reservoir, or it may contain multiple reservoirs. In one embodiment, the reservoir comprises a polymeric matrix of a pharmaceutically acceptable contact adhesive material that serves to affix the system to the skin during active ingredients delivery. Examples of suitable skin contact adhesive materials include, but are not limited to, polyethylenes, polysiloxanes, polyisobutylenes, polyacrylates, polyurethanes, and the like. The particular polymeric adhesive selected will depend on the particular active ingredients, vehicle, etc., i.e., the adhesive must be compatible with all components of the active ingredients-containing composition. Alternatively, the active ingredients-containing reservoir and skin contact adhesive are present as separate and distinct layers, with the adhesive underlying the reservoir which, in this case, may be either a polymeric matrix as described above, or it may be a liquid or hydrogel reservoir, or may take some other form.

The backing layer in these laminates, which serves as the upper surface of the device, functions as the primary structural element of the laminated structure and provides the device with much of its flexibility. The material selected for the backing material should be selected so that it is substantially impermeable to the active ingredients and to any other components of the active ingredients-containing composition, thus preventing loss of any components through the upper surface of the device. The backing layer may be either occlusive or non-occlusive, depending on whether it is desired that the skin become hydrated during active ingredients delivery. The backing is preferably made of a sheet or film of a preferably flexible elastomeric material. Examples of polymers that are suitable for the backing layer include polyethylene, polypropylene, and polyesters.

During storage and prior to use, the laminated structure includes a release liner. Immediately prior to use, this layer is removed from the device to expose the basal

surface thereof, either the active ingredients reservoir or a separate contact adhesive layer, so that the system may be affixed to the skin. The release liner should be made from an active ingredients/vehicle impermeable material.

Such devices may be fabricated using conventional techniques, known in the art, for example by casting a fluid admixture of adhesive, active ingredients and vehicle onto the backing layer, followed by lamination of the release liner. Similarly, the adhesive mixture may be cast onto the release liner, followed by lamination of the backing layer. Alternatively, the active ingredients reservoir may be prepared in the absence of active ingredients or excipient, and then loaded by "soaking" in an active ingredients/vehicle mixture.

As with the topical formulations of the invention, the active ingredients composition contained within the active ingredients reservoirs of these laminated system may contain a number of components. In some cases, the active ingredients may be delivered "neat," i.e., in the absence of additional liquid. In most cases, however, the active ingredients will be dissolved, dispersed or suspended in a suitable pharmaceutically acceptable vehicle, typically a solvent or gel. Other components, which may be present, include preservatives, stabilizers, surfactants, and the like.

It should be noted that the protein of the invention, such as a high mannose lysosomal enzyme, is preferably administered to the patient in need in an effective amount. As used herein, "effective amount" means an amount necessary to achieve a selected result. For example, an effective amount of the composition of the invention may be selected for being useful for the treatment of a lysosomal storage disease.

Pharmaceutical compositions suitable for use in context of the present invention include compositions wherein the active ingredients are contained in an amount effective to achieve the intended purpose. More specifically, a therapeutically effective amount means an amount of active ingredient effective to prevent, alleviate or ameliorate symptoms of disease or prolong the survival of the subject being treated.

Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

For any active ingredient used in the methods of the invention, the therapeutically effective amount or dose can be estimated initially from activity assays

in animals. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined by activity assays.

Toxicity and therapeutic efficacy of the active ingredients described herein can be determined by standard pharmaceutical procedures in experimental animals, e.g., by determining the IC_{50} and the LD_{50} (lethal dose causing death in 50 % of the tested animals) for a subject active ingredient. The data obtained from these activity assays and animal studies can be used in formulating a range of dosage for use in human. For example, therapeutically effective doses suitable for treatment of genetic disorders can be determined from the experiments with animal models of these diseases.

The dosage may vary depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl, et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1).

Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the modulating effects, termed the minimal effective concentration (MEC). The MEC will vary for each preparation, but may optionally be estimated from whole animal data.

Dosage intervals can also be determined using the MEC value. Preparations may optionally be administered using a regimen, which maintains plasma levels above the MEC for 10-90 % of the time, preferable between 30-90 % and most preferably 50-90 %.

Depending on the severity and responsiveness of the condition to be treated, dosing can also be a single administration of a slow release composition described hereinabove, with course of treatment lasting from several days to several weeks or until cure is effected or diminution of the disease state is achieved.

Compositions of the present invention may, if desired, be presented in a pack or dispenser device, such as an FDA approved kit, which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. The pack or dispenser may also be accompanied by a notice associated with the container in a form prescribed by a

governmental agency regulating the manufacture, use or sale of pharmaceuticals, which notice is reflective of approval by the agency of the form of the compositions or human or veterinary administration. Such notice, for example, may be of labeling approved by the U.S. Food and Drug Administration for prescription drugs or of an approved product insert. Compositions comprising an active ingredient of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

As used herein, the term "modulate" includes substantially inhibiting, slowing or reversing the progression of a disease, substantially ameliorating clinical symptoms of a disease or condition, or substantially preventing the appearance of clinical symptoms of a disease or condition. A "modulator" therefore includes an agent which may modulate a disease or condition.

Other Embodiments

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, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

REFERENCES

1. Ma, J. K. C., Drake, P.M.W., and Christou, P. (2003) *Nature reviews* **4**, 794-805
2. Lerouge, P., Cabanes-Macheteau, M, Rayon, C, Fischette-Laine, A.C., Gomord, V, Faye, L. (1998) *Plant Mol Biol* **38**, 31-48
3. Lee, R. E. (1982) *Prog Clin Biol Res* **95**, 177-217
4. Grabowski, G. (1993) *Adv Hum Genet.* **21**, 377-441
5. Grabowski, G. A., and Hopkin, R. J. (2003) *Annual Review of Genomics and Human Genetics* **4**, 403-436
6. Sorge, J. W., C.,Westwood, B., Beutler, E. (1985) *Proc Natl Acad Sci USA.* **82**, 7289-7293
7. Berg-Fussman, A., Grace, M., Ioannou, Y., and Grabowski, G. (1993) *J. Biol. Chem.* **268**, 14861-14866
8. Grace, M., Grabowski, GA. (1990) *Biochem Biophys Res Commun* **168**, 771-777
9. Grace, M., Newman, K., Scheinker, V., Berg-Fussman, A., and Grabowski, G. (1994) *J. Biol. Chem.* **269**, 2283-2291
10. Barton, N. W., Brady, R.O., Dambrosia, J.M., Di Bisceglie, A.M., Doppelt, S.H., Hill, S.C., Mankin, H.J., Murray, G.J., Parker, R.I., Argoff, C.E., et al. (1991) *N Engl J Med.* **324**, 1464-1470
11. Grabowski, G. A., Barton, N. W., Pastores, G., Dambrosia, J. M., Banerjee, T. K., McKee, M. A., Parker, C., Schiffmann, R., Hill, S. C., and Brady, R. O. (1995) *Ann Intern Med* **122**, 33-39
12. Pastores, G. M., Sibille, A.R., Grabowski, G.A. (1993) *Blood* **82**, 408-416.
13. Weinreb, N. J., Charrow, J, Andersson, H.C., Kaplan, P, Kolodny, E.H., Mistry, P, Pastores, G, Rosenbloom, B.E., Scott, C.R., Wappner, R.S., Zimran, A. (2002) *Am J Med* **113**, 112-119
14. Bijsterbosch, M. K., Donker, W, van de Bilt, H, van Weely, S, van Berkel, T.J., Aerts, JM. (1996) *Eur J Biochem* **237**, 344-349
15. Friedman, B., Vaddi, K., Preston, C., Mahon, E., Cataldo, J. R., and McPherson, J. M. (1999) *Blood* **93**, 2807-2816
16. Furbish, F. S., Steer, C.J., Krett, N.L., Barranger, J.A. (1981) *Biochim Biophys Acta* **673**, 425-434

17. Doebber, T., Wu, M., Bugianesi, R., Ponpipom, M., Furbish, F., Barranger, J., Brady, R., and Shen, T. (1982) *J. Biol. Chem.* **257**, 2193-2199

18. Dwek, R. A., Butters, T.D., Platt, F.M., Zitzmann, N. (2002) *Nature reviews* **1**, 65-75

19. Neuhaus, J. M., Rogers, J.C. (1998) *Plant Mol Biol* **38**, 127-144

20. Vitale, A., and Galili, G. (2001) *Plant Physiol.* **125**, 115-118

21. Hellens, R., Edwards, EA., Leyland, N.R., Bean, S.,Mullineaux, P.M. (2000) *Plant Mol Biol* **42**, 819-832

22. Wurtele, E. S., Bulka, K. (1989) *Plant Sci* **61**, 253-262

23. den Dulk-Ras, A., Hooykaas, P.J. (1995) *Methods Mol Biol.* **55**, 63-72

24. Laemmli, U. K. (1970) *Nature reviews* **227**, 680-685

25. Bradford, M. M. (1976) *Anal Biochem* **72**, 248-254

26. Stahl, P. G. S. (1982) *J Cell Biol* **93**, 49-56

27. Takasaki, S., Murray, G., Furbish, F., Brady, R., Barranger, J., and Kobata, A. (1984) *J. Biol. Chem.* **259**, 10112-10117

28. Lerouge, P., Cabanes-Macheteau, M., Rayon, C., Fitchette-Laine, A. C., Gomord, V., and Faye, L. (1998) *Plant Mol. Biol.* **38**, 31-48

29. Frigerio, L., Pastres, A., Prada, A., and Vitale, A. (2001) *Plant Cell* **13**, 1109-1126

30. Frigerio, L., de Virgilio, M., Prada, A., Faoro, F., and Vitale, A. (1998) *Plant Cell* **10**, 1031-1042

31. Hadlington JL, D. J. (2000) Curr Opin Plant Biol. 3, 461-468.

32. Okamoto, T., Shimada, T., Hara-Nishimura, I., Nishimura, M., and Minamikawa, T. (2003) Plant Physiol. 132, 1892-1900

33. Bardor, M., Faveeuw, C., Fitchette, A.-C., Gilbert, D., Galas, L., Trottein, F., Faye, L., and Lerouge, P. (2003) Glycobiology 13, 427-434

CLAIMS:

1. A pharmaceutical composition comprising a human α-galactosidase protein having a molecular weight of 48.6 kDa, as determined by mass spectrometry, said α-galactosidase protein being further characterized by α-galactosidase catalytic activity and being contiguously linked to a C-terminal KDEL amino acid sequence, and a pharmaceutically acceptable carrier.

2. The pharmaceutical composition of claim 1, wherein said α-galactosidase protein further comprises the amino acid sequence: SHINPTGTVLLQLENTMQMSLK located at the C-terminus portion of said protein.

3. The pharmaceutical composition of claim 1 or 2, wherein said α-galactosidase protein is immunologically reactive with a rabbit polyclonal antibody raised against the polypeptide fragment as set forth between amino acids 326 and 429 of native human α-galactosidase.

4. A pharmaceutical composition comprising a human α-galactosidase protein comprising the amino acid sequence: SHINPTGTVLLQLENTMQMSLK located at the C-terminus portion of said protein, said α-galactosidase protein being further characterized by α-galactosidase catalytic activity and being contiguously linked to a C-terminal KDEL amino acid sequence, and a pharmaceutically acceptable carrier.

5. The pharmaceutical composition of claim 4, wherein said α-galactosidase protein having a molecular weight of 48.6 kDa, as determined by mass spectrometry.

6. The pharmaceutical composition of claim 4 or 5, wherein said α-galactosidase protein is immunologically reactive with a rabbit polyclonal antibody raised against the polypeptide fragment as set forth between amino acids 326 and 429 of native human α-galactosidase.

7. A pharmaceutical composition comprising a human α-galactosidase protein, wherein said α-galactosidase is immunologically reactive with a rabbit polyclonal antibody raised against the polypeptide fragment as set forth between amino acids 326 and 429 of native human α-galactosidase, said α-galactosidase protein being further characterized by α-galactosidase catalytic activity and being contiguously linked to a C-terminal KDEL amino acid sequence, and a pharmaceutically acceptable carrier.

8. The pharmaceutical composition of claim 7, wherein said α-galactosidase protein has a molecular weight of 48.6 kDa, as determined by mass spectrometry.

9. The pharmaceutical composition of claim 7 or 8, wherein said α-galactosidase protein comprises the amino acid sequence: SHINPTGTVLLQLENTMQMSLK located at the C-terminus portion of said protein.

10. The pharmaceutical composition of any one of claims 1-9, wherein said α-galactosidase protein comprises at least one exposed mannose residue.

11. The pharmaceutical composition of claim 10, wherein said human lysosomal protein further comprises at least one xylose residue.

12. The pharmaceutical composition of claim 13, wherein said human lysosomal protein is further comprises at least one core α-(1,2) xylose and at least one core α-(1,3) fucose.

13. The pharmaceutical composition of any one of claim 1-12, wherein said human α-galactosidase protein has uptake into fibroblasts.

14. Use of the pharmaceutical composition of any one of claims 1-13 for the manufacture of a medicament for the treatment of Fabry's disease in a subject.

SEQUENCE LISTING

<110> PROTALIX LTD.
Shaaltiel, Yoseph
Baum, Gideon
Bartfeld, Daniel
Hashmueli, Sharon
Lewkowicz, Ayala

<120> PRODUCTION OF HIGH MANNOSE PROTEINS IN PLANT CULTURE

<130> 43186

<160> 24

<170> PatentIn version 3.5

<210> 1
<211> 22
<212> PRT
<213> Artificial sequence

<220>
<223> ER signal peptide

<400> 1

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1 5 10 15

Leu Ser Ser Ala Glu Phe
20

<210> 2
<211> 7
<212> PRT
<213> Artificial sequence

<220>
<223> Vacuolar targeting signal from Tobacco chitinase A

<400> 2

Asp Leu Leu Val Asp Thr Met
1 5

<210> 3
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 3
cagaattcgc ccgcccctgc a 21

<210> 4
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 4
ctcagatctt ggcgatgcca ca 22

<210> 5
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 5
ctcagaagac cagagggct 19

<210> 6
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<400> 6
caaagcggcc atcgtgc 17

<210> 7
<211> 1491
<212> DNA
<213> Homo sapiens

<400> 7
gcccgccccc gcatccctaa aagcttcggc tacagctcgg tggtgtgtgt ctgcaatgcc 60

acatactgtg actcctttga ccccccgacc tttcctgccc ttggtacctt cagccgctat 120

gagagtacac gcagtgggcg acggatggag ctgagtatgg ggcccatcca ggctaatcac 180

acgggcacag gcctgctact gaccctgcag ccagaacaga agttccagaa agtgaaggga 240

tttggagggg ccatgacaga tgctgctgct ctcaacatcc ttgccctgtc acccccctgcc 300

caaaatttgc tacttaaatc gtacttctct gaagaaggaa tcggatataa catcatccgg 360

gtacccatgg ccagctgtga cttctccatc cgcacctaca cctatgcaga cacccctgat 420

gatttccagt tgcacaactt cagcctccca gaggaagata ccaagctcaa gatacccctg 480

attcaccgag ccctgcagtt ggcccagcgt cccgtttcac tccttgccag cccctggaca 540

tcacccactt ggctcaagac caatggagcg gtgaatggga aggggtcact caagggacag 600

cccggagaca tctaccacca gacctgggcc agatactttg tgaagttcct ggatgcctat 660

gctgagcaca agttacagtt ctgggcagtg acagctgaaa atgagccttc tgctgggctg 720

ttgagtggat acccccttcca gtgcctgggc ttcacccctg aacatcagcg agacttcatt 780

gcccgtgacc taggtcctac cctcgccaac agtactcacc acaatgtccg cctactcatg 840

ctggatgacc aacgcttgct gctgccccac tgggcaaagg tggtactgac agacccagaa 900

gcagctaaat atgttcatgg cattgctgta cattggtacc tggactttct ggctccagcc 960

aaagccaccc taggggagac acaccgcctg ttccccaaca ccatgctctt tgcctcagag 1020

gcctgtgtgg gctccaagtt ctgggagcag agtgtgcggc taggctcctg ggatcgaggg 1080

atgcagtaca gccacagcat catcacgaac ctcctgtacc atgtggtcgg ctggaccgac 1140

tggaaccttg ccctgaaccc cgaaggagga cccaattggg tgcgtaactt tgtcgacagt 1200

cccatcattg tagacatcac caaggacacg ttttacaaac agcccatgtt ctaccacctt 1260

ggccacttca gcaagttcat tcctgagggc tcccagagag tggggctggt tgccagtcag 1320

aagaacgacc tggacgcagt ggcactgatg catcccgatg gctctgctgt tgtggtcgtg 1380

ctaaaccgct cctctaagga tgtgcctctt accatcaagg atcctgctgt gggcttcctg 1440

gagacaatct cacctggcta ctccattcac acctacctgt ggcatcgcca g 1491

<210> 8
<211> 497
<212> PRT
<213> Homo sapiens

<400> 8

```
Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser Val Val Cys
1               5                  10                  15

Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro Thr Phe Pro
            20                  25                  30

Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser Gly Arg Arg
        35                  40                  45

Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr Gly Thr Gly
    50                  55                  60

Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys Val Lys Gly
65                  70                  75                  80

Phe Gly Gly Ala Met Thr Asp Ala Ala Ala Leu Asn Ile Leu Ala Leu
                85                  90                  95

Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe Ser Glu Glu
            100                 105                 110

Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro Met Ala Ser Cys Asp Phe
        115                 120                 125

Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr Pro Asp Asp Phe Gln Leu
    130                 135                 140
```

```
His Asn Phe Ser Leu Pro Glu Glu Asp Thr Lys Leu Lys Ile Pro Leu
145                 150                 155                 160

Ile His Arg Ala Leu Gln Leu Ala Gln Arg Pro Val Ser Leu Leu Ala
                165                 170                 175

Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys Thr Asn Gly Ala Val Asn
            180                 185                 190

Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly Asp Ile Tyr His Gln Thr
        195                 200                 205

Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp Ala Tyr Ala Glu His Lys
    210                 215                 220

Leu Gln Phe Trp Ala Val Thr Ala Glu Asn Glu Pro Ser Ala Gly Leu
225                 230                 235                 240

Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly Phe Thr Pro Glu His Gln
                245                 250                 255

Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro Thr Leu Ala Asn Ser Thr
            260                 265                 270

His His Asn Val Arg Leu Leu Met Leu Asp Asp Gln Arg Leu Leu Leu
        275                 280                 285

Pro His Trp Ala Lys Val Val Leu Thr Asp Pro Glu Ala Ala Lys Tyr
    290                 295                 300

Val His Gly Ile Ala Val His Trp Tyr Leu Asp Phe Leu Ala Pro Ala
305                 310                 315                 320

Lys Ala Thr Leu Gly Glu Thr His Arg Leu Phe Pro Asn Thr Met Leu
                325                 330                 335

Phe Ala Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu Gln Ser Val
            340                 345                 350

Arg Leu Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His Ser Ile Ile
        355                 360                 365

Thr Asn Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp Asn Leu Ala
    370                 375                 380

Leu Asn Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe Val Asp Ser
385                 390                 395                 400

Pro Ile Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys Gln Pro Met
                405                 410                 415

Phe Tyr His Leu Gly His Phe Ser Lys Phe Ile Pro Glu Gly Ser Gln
            420                 425                 430
```

Arg Val Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp Ala Val Ala
435 440 445

Leu Met His Pro Asp Gly Ser Ala Val Val Val Val Leu Asn Arg Ser
450 455 460

Ser Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val Gly Phe Leu
465 470 475 480

Glu Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu Trp His Arg
485 490 495

Gln
497

<210> 9
<211> 338
<212> DNA
<213> Cauliflower mosaic virus

<400> 9
ttttcacaaa gggtaatatc gggaaacctc ctcggattcc attgcccagc tatctgtcac 60

ttcatcgaaa ggacagtaga aaaggaaggt ggctcctaca aatgccatca ttgcgataaa 120

ggaaaggcta tcgttcaaga tgcctctacc gacagtggtc ccaaagatgg acccccaccc 180

acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaagca agtggattga 240

tgtgatatct ccactgacgt aagggatgac gcacaatccc actatccttc gcaagaccct 300

tcctctatat aaggaagttc atttcatttg gagaggac 338

<210> 10
<211> 66
<212> DNA
<213> Artificial sequence

<220>
<223> Nucleic acid sequence encoding the ER signal peptide

<400> 10
atgaagacta atctttttct ctttctcatc ttttcacttc tcctatcatt atcctcggcc 60

gaattc 66

<210> 11
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Nucleic acid sequence encoding the vacuolar targeting sequence

<400> 11
gatcttttag tcgatactat g 21

<210> 12
<211> 167
<212> DNA
<213> Artificial sequence

<220>
<223> Nucleic acid sequence of the Agrobacterium tumefaciens terminator

<220>
<221> misc_feature
<222> (162)..(162)
<223> n is a, c, g, or t

<400> 12
taatttcatg atctgttttg ttgtattccc ttgcaatgca gggcctaggg ctatgaataa 60

agttaatgtg tgaatgtgtg aatgtgtgat tgtgacctga agggatcacg actataatcg 120

tttataataa acaaagactt tgtcccaaaa acccccccc cngcaga 167

<210> 13
<211> 2186
<212> DNA
<213> Artificial sequence

<220>
<223> nucleic acid sequence encoding high mannose human glucocerebrosidase (GCD)

<220>
<221> misc_feature
<222> (2181)..(2181)
<223> n is a, c, g, or t

<400> 13
ttttcacaaa gggtaatatc gggaaacctc ctcggattcc attgcccagc tatctgtcac 60

ttcatcgaaa ggacagtaga aaaggaaggt ggctcctaca aatgccatca ttgcgataaa 120

ggaaaggcta tcgttcaaga tgcctctacc gacagtggtc ccaaagatgg acccccaccc 180

acgaggaaca tcgtggaaaa agaagacgtt ccaaccacgt cttcaaagca agtggattga 240

tgtgatatct ccactgacgt aagggatgac gcacaatccc actatccttc gcaagaccct 300

tcctctatat aaggaagttc atttcatttg gagaggacag gcttcttgag atccttcaac 360

aattaccaac aacaacaaac aacaaacaac attacaatta ctatttacaa ttacagtcga 420

gggatccaag gagatataac aatgaagact aatctttttc tctttctcat cttttcactt 480

ctcctatcat tatcctcggc cgaattcgcc cgcccctgca tccctaaaag cttcggctac 540

agctcggtgg tgtgtgtctg caatgccaca tactgtgact cctttgaccc cccgaccttt 600

cctgcccttg gtaccttcag ccgctatgag agtacacgca gtgggcgacg gatggagctg 660

agtatggggc ccatccaggc taatcacacg ggcacaggcc tgctactgac cctgcagcca 720

gaacagaagt tccagaaagt gaagggattt ggaggggcca tgacagatgc tgctgctctc 780

aacatccttg ccctgtcacc ccctgcccaa aatttgctac ttaaatcgta cttctctgaa 840

gaaggaatcg gatataacat catccgggta cccatggcca gctgtgactt ctccatccgc 900

acctacacct atgcagacac ccctgatgat ttccagttgc acaacttcag cctcccagag 960

gaagatacca agctcaagat acccctgatt caccgagccc tgcagttggc ccagcgtccc 1020

gtttcactcc ttgccagccc ctggacatca cccacttggc tcaagaccaa tggagcggtg 1080

aatgggaagg ggtcactcaa gggacagccc ggagacatct accaccagac ctgggccaga 1140

tactttgtga agttcctgga tgcctatgct gagcacaagt tacagttctg ggcagtgaca 1200

gctgaaaatg agccttctgc tgggctgttg agtggatacc ccttccagtg cctgggcttc 1260

acccctgaac atcagcgaga cttcattgcc cgtgacctag gtcctaccct cgccaacagt 1320

actcaccaca atgtccgcct actcatgctg gatgaccaac gcttgctgct gccccactgg 1380

gcaaaggtgg tactgacaga cccagaagca gctaaatatg ttcatggcat tgctgtacat 1440

tggtacctgg actttctggc tccagccaaa gccaccctag gggagacaca ccgcctgttc 1500

cccaacacca tgctctttgc ctcagaggcc tgtgtgggct ccaagttctg ggagcagagt 1560

gtgcggctag gctcctggga tcgagggatg cagtacagcc acagcatcat cacgaacctc 1620

ctgtaccatg tggtcggctg gaccgactgg aaccttgccc tgaaccccga aggaggaccc 1680

aattgggtgc gtaactttgt cgacagtccc atcattgtag acatcaccaa ggacacgttt 1740

tacaaacagc ccatgttcta ccaccttggc cacttcagca agttcattcc tgagggctcc 1800

cagagagtgg ggctggttgc cagtcagaag aacgacctgg acgcagtggc actgatgcat 1860

cccgatggct ctgctgttgt ggtcgtgcta aaccgctcct ctaaggatgt gcctcttacc 1920

atcaaggatc ctgctgtggg cttcctggag acaatctcac ctggctactc cattcacacc 1980

tacctgtggc atcgccaaga tcttttagtc gatactatgt aatttcatga tctgttttgt 2040

tgtattccct tgcaatgcag ggcctagggc tatgaataaa gttaatgtgt gaatgtgtga 2100

atgtgtgatt gtgacctgaa gggatcacga ctataatcgt ttataataaa caaagacttt 2160

gtcccaaaaa ccccccccc ngcaga 2186

<210> 14
<211> 526
<212> PRT
<213> Artificial sequence

<220>
<223> High mannose human glucocerebrosidase (GCD)

<400> 14

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1 5 10 15

Leu Ser Ser Ala Glu Phe Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly
20 25 30

Tyr Ser Ser Val Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe
35 40 45

Asp Pro Pro Thr Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser
50 55 60

Thr Arg Ser Gly Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala
65 70 75 80

```
Asn His Thr Gly Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys
                85                  90                  95

Phe Gln Lys Val Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala Ala
            100                 105                 110

Leu Asn Ile Leu Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys
        115                 120                 125

Ser Tyr Phe Ser Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro
    130                 135                 140

Met Ala Ser Cys Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr
145                 150                 155                 160

Pro Asp Asp Phe Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp Thr
                165                 170                 175

Lys Leu Lys Ile Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln Arg
            180                 185                 190

Pro Val Ser Leu Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys
        195                 200                 205

Thr Asn Gly Ala Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly
    210                 215                 220

Asp Ile Tyr His Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp
225                 230                 235                 240

Ala Tyr Ala Glu His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu Asn
                245                 250                 255

Glu Pro Ser Ala Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly
            260                 265                 270

Phe Thr Pro Glu His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro
        275                 280                 285

Thr Leu Ala Asn Ser Thr His His Asn Val Arg Leu Leu Met Leu Asp
    290                 295                 300

Asp Gln Arg Leu Leu Leu Pro His Trp Ala Lys Val Val Leu Thr Asp
305                 310                 315                 320

Pro Glu Ala Ala Lys Tyr Val His Gly Ile Ala Val His Trp Tyr Leu
                325                 330                 335

Asp Phe Leu Ala Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg Leu
            340                 345                 350

Phe Pro Asn Thr Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys
        355                 360                 365
```

```
Phe Trp Glu Gln Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln
    370                 375                 380

Tyr Ser His Ser Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp
385                 390                 395                 400

Thr Asp Trp Asn Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val
                405                 410                 415

Arg Asn Phe Val Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr
            420                 425                 430

Phe Tyr Lys Gln Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe
        435                 440                 445

Ile Pro Glu Gly Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn
    450                 455                 460

Asp Leu Asp Ala Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val
465                 470                 475                 480

Val Val Leu Asn Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp
                485                 490                 495

Pro Ala Val Gly Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His
            500                 505                 510

Thr Tyr Leu Trp His Arg Gln Asp Leu Leu Val Asp Thr Met
        515                 520                 525
```

<210> 15
<211> 506
<212> PRT
<213> Artificial sequence

<220>
<223> Processed plant produced human recombinant GCD protein

<400> 15

```
Glu Phe Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser Val
1               5                   10                  15

Val Cys Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro Thr
            20                  25                  30

Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser Gly
        35                  40                  45

Arg Arg Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr Gly
    50                  55                  60

Thr Gly Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys Val
65                  70                  75                  80

Lys Gly Phe Gly Gly Ala Met Thr Asp Ala Ala Ala Leu Asn Ile Leu
```

85 90 95

Ala Leu Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe Ser
100 105 110

Glu Glu Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro Met Ala Ser Cys
115 120 125

Asp Phe Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr Pro Asp Asp Phe
130 135 140

Gln Leu His Asn Phe Ser Leu Pro Glu Glu Asp Thr Lys Leu Lys Ile
145 150 155 160

Pro Leu Ile His Arg Ala Leu Gln Leu Ala Gln Arg Pro Val Ser Leu
165 170 175

Leu Ala Ser Pro Trp Thr Ser Pro Thr Trp Leu Lys Thr Asn Gly Ala
180 185 190

Val Asn Gly Lys Gly Ser Leu Lys Gly Gln Pro Gly Asp Ile Tyr His
195 200 205

Gln Thr Trp Ala Arg Tyr Phe Val Lys Phe Leu Asp Ala Tyr Ala Glu
210 215 220

His Lys Leu Gln Phe Trp Ala Val Thr Ala Glu Asn Glu Pro Ser Ala
225 230 235 240

Gly Leu Leu Ser Gly Tyr Pro Phe Gln Cys Leu Gly Phe Thr Pro Glu
245 250 255

His Gln Arg Asp Phe Ile Ala Arg Asp Leu Gly Pro Thr Leu Ala Asn
260 265 270

Ser Thr His His Asn Val Arg Leu Leu Met Leu Asp Asp Gln Arg Leu
275 280 285

Leu Leu Pro His Trp Ala Lys Val Val Leu Thr Asp Pro Glu Ala Ala
290 295 300

Lys Tyr Val His Gly Ile Ala Val His Trp Tyr Leu Asp Phe Leu Ala
305 310 315 320

Pro Ala Lys Ala Thr Leu Gly Glu Thr His Arg Leu Phe Pro Asn Thr
325 330 335

Met Leu Phe Ala Ser Glu Ala Cys Val Gly Ser Lys Phe Trp Glu Gln
340 345 350

Ser Val Arg Leu Gly Ser Trp Asp Arg Gly Met Gln Tyr Ser His Ser
355 360 365

Ile Ile Thr Asn Leu Leu Tyr His Val Val Gly Trp Thr Asp Trp Asn

370 375 380

Leu Ala Leu Asn Pro Glu Gly Gly Pro Asn Trp Val Arg Asn Phe Val
385 390 395 400

Asp Ser Pro Ile Ile Val Asp Ile Thr Lys Asp Thr Phe Tyr Lys Gln
405 410 415

Pro Met Phe Tyr His Leu Gly His Phe Ser Lys Phe Ile Pro Glu Gly
420 425 430

Ser Gln Arg Val Gly Leu Val Ala Ser Gln Lys Asn Asp Leu Asp Ala
435 440 445

Val Ala Leu Met His Pro Asp Gly Ser Ala Val Val Val Val Leu Asn
450 455 460

Arg Ser Ser Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val Gly
465 470 475 480

Phe Leu Glu Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu Trp
485 490 495

His Arg Gln Asp Leu Leu Val Asp Thr Met
500 505

<210> 16
<211> 26
<212> PRT
<213> Artificial sequence

<220>
<223> Apple pectinase leader peptide

<400> 16

Met Ala Leu Lys Thr Gln Leu Leu Trp Ser Phe Val Val Val Phe Val
1 5 10 15

Val Ser Phe Ser Thr Thr Ser Cys Ser Gly
20 25

<210> 17
<211> 1305
<212> DNA
<213> Artificial sequence

<220>
<223> Alpha-gal-vac expression construct

<400> 17
gtcgacgcaa ccatggctct taagactcaa ttgctctggt ctttcgttgt tgtgttcgtg 60

gtctcctttt caacaactag ttgctccgga gaattccttg ataatggatt ggccaggaca 120

ccaactatgg gatggctgca ttgggagcgt tttatgtgta accttgactg ccaagaagag 180

cctgattctt gcatctcaga gaagctcttt atggaaatgg ctgagctgat ggtttctgaa 240

ggctggaaag acgcaggtta tgagtatttg tgcatcgacg attgctggat ggctccacag 300

agagatagtg aagggagact tcaagcagat cctcagaggt ttccacatgg tattagacag 360

ctcgccaatt atgttcactc caagggcctt aagttgggga tatacgctga tgtgggaaac 420

aaaacctgtg ctggatttcc tggttccttc ggctattacg atattgacgc acaaaccttt 480

gctgattggg gagttgacct ccttaaattc gatggctgtt actgtgattc acttgagaat 540

ctcgccgacg gttataagca catgtctctc gctttgaaca ggactggaag gagcatagtc 600

tactcctgtg aatggccact ttacatgtgg cctttccaga agccaaacta cactgagatt 660

aggcagtatt gcaaccactg gcgtaatttt gccgatatcg acgattcttg gaagtccatc 720

aagagcattc tggattggac atccttcaat caggaaagga ttgtggatgt tgctggacct 780

ggtggatgga atgatcctga catgttggtg atagggaact ttggtctctc atggaatcag 840

caagttaccc aaatggcact ctgggctatt atggctgcac cactcttcat gtctaacgac 900

ctcaggcata tctcacctca agcaaaggca ttgctgcaag acaaagacgt cattgccatt 960

aaccaggacc ctcttgggaa gcaaggttat caacttcgtc aaggcgataa ctttgaagtg 1020

tgggagagac cattgtctgg acttgcttgg gctgttgcca tgattaatag gcaagaaatc 1080

ggtggcccaa gaagttatac aatcgctgtt gcaagtcttg gaaaaggtgt cgcttgtaat 1140

cctgcctgtt tcattactca gctcctgcct gtgaaaagaa agttggggtt ctatgaatgg 1200

acttcaagac tcaggagcca tattaaccct accggaacag ttcttctcca actggagaac 1260

accatgcaaa tgtccctcaa agatcttctg gtggacacca tgtga 1305

<210> 18
<211> 430
<212> PRT
<213> Artificial sequence

<220>
<223> Alpha-gal-vac polypeptide

<400> 18

Met Ala Leu Lys Thr Gln Leu Leu Trp Ser Phe Val Val Val Phe Val
1 5 10 15

Val Ser Phe Ser Thr Thr Ser Cys Ser Gly Glu Phe Leu Asp Asn Gly
20 25 30

Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu Arg Phe Met
35 40 45

Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys Ile Ser Glu Lys
50 55 60

Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu Gly Trp Lys Asp
65 70 75 80

Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met Ala Pro Gln
85 90 95

Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg Phe Pro His
100 105 110

```
Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys Gly Leu Lys Leu
        115                 120                 125

Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala Gly Phe Pro Gly
    130                 135                 140

Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe Ala Asp Trp Gly
145                 150                 155                 160

Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp Ser Leu Glu Asn
                165                 170                 175

Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu Asn Arg Thr Gly
            180                 185                 190

Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr Met Trp Pro Phe
        195                 200                 205

Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys Asn His Trp Arg
    210                 215                 220

Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile Lys Ser Ile Leu
225                 230                 235                 240

Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp Val Ala Gly Pro
                245                 250                 255

Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly Asn Phe Gly Leu
            260                 265                 270

Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp Ala Ile Met Ala
        275                 280                 285

Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile Ser Pro Gln Ala
    290                 295                 300

Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile Asn Gln Asp Pro
305                 310                 315                 320

Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp Asn Phe Glu Val
                325                 330                 335

Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val Ala Met Ile Asn
            340                 345                 350

Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile Ala Val Ala Ser
        355                 360                 365

Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile Thr Gln Leu
    370                 375                 380

Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp Thr Ser Arg Leu
385                 390                 395                 400
```

Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu Gln Leu Glu Asn
405 410 415

Thr Met Gln Met Ser Leu Lys Asp Leu Leu Val Asp Thr Met
420 425 430

<210> 19
<211> 1308
<212> DNA
<213> Artificial sequence

<220>
<223> Alpha-gal-KDEL expression construct

<400> 19
gtcgacgcaa ccatggctct taagactcaa ttgctctggt ctttcgttgt tgtgttcgtg 60

gtctcctttt caacaactag ttgctccgga gaattccttg ataatggatt ggccaggaca 120

ccaactatgg gatggctgca ttgggagcgt tttatgtgta accttgactg ccaagaagag 180

cctgattctt gcatctcaga gaagctcttt atggaaatgg ctgagctgat ggtttctgaa 240

ggctggaaag acgcaggtta tgagtatttg tgcatcgacg attgctggat ggctccacag 300

agagatagtg aagggagact tcaagcagat cctcagaggt ttccacatgg tattagacag 360

ctcgccaatt atgttcactc caagggcctt aagttgggga tatacgctga tgtgggaaac 420

aaaacctgtg ctggatttcc tggttccttc ggctattacg atattgacgc acaaaccttt 480

gctgattggg gagttgacct ccttaaattc gatggctgtt actgtgattc acttgagaat 540

ctcgccgacg gttataagca catgtctctc gctttgaaca ggactggaag gagcatagtc 600

tactcctgtg aatggccact ttacatgtgg cctttccaga agccaaacta cactgagatt 660

aggcagtatt gcaaccactg gcgtaatttt gccgatatcg acgattcttg gaagtccatc 720

aagagcattc tggattggac atccttcaat caggaaagga ttgtggatgt tgctggacct 780

ggtggatgga atgatcctga catgttggtg atagggaact ttggtctctc atggaatcag 840

caagttaccc aaatggcact ctgggctatt atggctgcac cactcttcat gtctaacgac 900

ctcaggcata tctcacctca agcaaaggca ttgctgcaag acaaagacgt cattgccatt 960

aaccaggacc ctcttgggaa gcaaggttat caacttcgtc aaggcgataa ctttgaagtg 1020

tgggagagac cattgtctgg acttgcttgg gctgttgcca tgattaatag gcaagaaatc 1080

ggtggcccaa gaagttatac aatcgctgtt gcaagtcttg gaaaaggtgt cgcttgtaat 1140

cctgcctgtt tcattactca gctcctgcct gtgaaaagaa agttggggtt ctatgaatgg 1200

acttcaagac tcaggagcca tattaaccct accggaacag ttcttctcca actggagaac 1260

accatgcaaa tgtccctcaa agatcttagt gagaaggacg agctctga 1308

<210> 20
<211> 435
<212> PRT
<213> Artificial sequence

<220>
<223> Alpha-gal-KDEL polypeptide

<400> 20

```
Val Asp Ala Thr Met Ala Leu Lys Thr Gln Leu Leu Trp Ser Phe Val
1               5                  10                  15

Val Val Phe Val Val Ser Phe Ser Thr Thr Ser Cys Ser Gly Glu Phe
            20                  25                  30

Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp
        35                  40                  45

Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys
    50                  55                  60

Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu
65                  70                  75                  80

Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp
                85                  90                  95

Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln
            100                 105                 110

Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys
        115                 120                 125

Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala
    130                 135                 140

Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe
145                 150                 155                 160

Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp
                165                 170                 175

Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu
            180                 185                 190

Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr
        195                 200                 205

Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys
    210                 215                 220

Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile
225                 230                 235                 240

Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp
                245                 250                 255

Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly
            260                 265                 270

Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp
        275                 280                 285
```

```
Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile
    290                 295                 300

Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile
305                 310                 315                 320

Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp
                325                 330                 335

Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val
            340                 345                 350

Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile
        355                 360                 365

Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe
    370                 375                 380

Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp
385                 390                 395                 400

Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu
                405                 410                 415

Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Ser Glu Lys
            420                 425                 430

Asp Glu Leu
        435
```

<210> 21
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<220>
<221> misc_feature
<223> 5' phosphorylated

<400> 21
gatcttagtg agaaggacga gctctgataa 30

<210> 22
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Single strand DNA oligonucleotide

<220>
<221> misc_feature
<223> 5' phosphorylated

<400> 22

agctttatca gagctcgtcc ttctcactaa 30

<210> 23
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> ER retention signal peptide

<400> 23

Lys Asp Glu Leu
1

<210> 24
<211> 397
<212> PRT
<213> Artificial sequence

<220>
<223> Mature alpha-galactosidase sequence

<400> 24

Glu Phe Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu
1 5 10 15

His Trp Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp
20 25 30

Ser Cys Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val
35 40 45

Ser Glu Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp
50 55 60

Cys Trp Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp
65 70 75 80

Pro Gln Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His
85 90 95

Ser Lys Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr
100 105 110

Cys Ala Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln
115 120 125

Thr Phe Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr
130 135 140

Cys Asp Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu
145 150 155 160

Ala Leu Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro
165 170 175

Leu Tyr Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln

180 185 190

Tyr Cys Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys
195 200 205

Ser Ile Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile
210 215 220

Val Asp Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val
225 230 235 240

Ile Gly Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala
245 250 255

Leu Trp Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg
260 265 270

His Ile Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile
275 280 285

Ala Ile Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln
290 295 300

Gly Asp Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp
305 310 315 320

Ala Val Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr
325 330 335

Thr Ile Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala
340 345 350

Cys Phe Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr
355 360 365

Glu Trp Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val
370 375 380

Leu Leu Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys
385 390 395