O-glycan sialylated recombinant glycoproteins and cell lines for producing the same

The present invention relates to cell lines that are genetically modified to overexpress a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1), preferably human ST3Gal1, which can be used for the production of recombinant glycoproteins having highly or fully sialylated O-linked GalNAc glycans (GalNAc O-glycans), preferably core 1 GalNAc O-glycans, as well as to respective recombinant glycoproteins. Further, the present invention relates to respective methods of expressing recombinant glycoproteins, methods of increasing the degree of sialylation of recombinant glycoproteins, and methods of decreasing the micro-heterogeneity of GalNAc O-glycans. Finally, the present invention relates to respective uses of the above cell lines for the production of recombinant glycoproteins, for increasing the degree of sialylation of recombinant glycoproteins, and for decreasing the micro-heterogeneity of O-linked GalNAc glycans of recombinant glycoproteins.

Due to their clinical importance, the development of therapeutic proteins has accelerated immensely over the past years. However, the development of therapeutic proteins is often held up, in particular for complex glycosylated proteins, due to the difficulties in obtaining proteins with favorable glycosylation patterns. This often leads to suboptimal pharmacological properties such as *e.g.* reduced serum half-life or increased immunogenicity.

Another drawback is heterogeneity of the therapeutic proteins due to variations in the post-translational modification, in particular of the glycostructures. This can lead to inconsistencies in product quality between different production batches. Therefore, in order to advance the development of therapeutic proteins, it is important to achieve homogenous post-translational modification.

Glycosylation is the most common post-translational modification. Nearly all biopharmaceuticals need to be correctly glycosylated in order to display the optimal therapeutic efficacy. In general, glycosylation refers to the covalent attachment of

sugars to the protein surface, wherein the sugars are either connected to asparagine residues resulting in N-linked glycans (N-glycans) or serine or threonine residues resulting in O-linked glycans (O-glycans). The most common group of O-linked glycans are GalNAc O-glycans in which the serine or threonine is linked to N-acetyl-galactosamine (GalNAc) which in turn is linked to additional monosaccharides. As used herein, the terms "O-linked glycan" or "O-glycan" always relate to GalNAc O-glycans. As mentioned above, the glycosylation pattern can be very diverse from molecule to molecule as the attached forms can be different in monosaccharide order, branching pattern, and length (micro-heterogeneity). Additionally, not all glycosylation sites are fully occupied (macro-heterogeneity).

Glycosylation influences the solubility of proteins, their resistance to proteolysis, and their binding behavior to other proteins or to protein receptors such as *e.g.* the ASGPR (asialoglycoprotein receptor) and therefore influences the half-life of the glycoprotein in the plasma.

For small size proteins under about 50 kDa, clearance occurs mainly via renal clearance. Beside the size of a protein, also the protein surface charge has influence on the renal clearance, as the filtration of highly charged proteins in the kidney is decreased.

For large size proteins, clearance occurs mainly in the liver through specific and/or unspecific hepatic uptake. Examples for a specific uptake mediator would be the ASGPR, which binds specifically non-sialylated N-linked glycoproteins with a terminal galactose. Due to the action of this receptor, glycoproteins with terminal sialic acids covering the adjacent carbohydrate, galactose, have up to 100-fold increase in half-life as compared to their non-sialylated counterparts with terminal galactose residues on their N-linked glycans. Other receptors bind specifically to mannose, N-acetyl-glucosamine, or fucose and clear glycoproteins with these terminal sugars from the system.

From this perspective, a native N-glycosylation pattern including a high degree of terminal sialylation is crucial for therapeutic proteins since it determines the

pharmacokinetic properties of the therapeutic protein. In addition, terminal sialic acids with the proper linkage on glycoproteins reduce the immunogenicity of the glycoproteins.

The foremost strategy to obtain nearly native glycostructures and a high degree of sialylation is to produce the recombinant proteins in cell lines capable of linking mammalian-like glycostructures to proteins, *e.g.* CHO cells (Chinese hamster ovary cells). However, CHO cells lack the enzyme needed to catalyze 2,6-linkage of sialic acids, and therefore they can only catalyze the 2,3-linkage. Moreover, in addition to N-acetylneuraminic acid (sialic acid; NeuAc), they also link N-glycolylneuraminic acid (NeuGc) to glycans, a sugar not synthesized in humans and therefore immunogenic when injected into humans. Therefore, a better strategy would be to generate therapeutic proteins from cell lines derived from human cells like CAP cells (derived from human amniocytes) or HEK293 cells (derived from human embryo kidney cells). However, sialylation of therapeutic proteins secreted from mammalian cell lines during fermentation is often incomplete.

Due to the importance of sialylation for the pharmacokinetic profile of a protein, a lot of efforts were undertaken to increase the degree of sialylation. The beneficial effect of complete sialylation of N-linked glycosylation is well understood, but relatively little is known about the effect of O-glycosylation. Therefore, so far all cell-engineering efforts directed at improving the sialylation take aim on N-linked glycostructures.

O-linked GalNAc glycans always have an α-linked N-acetylgalactosamine residue linked to serine or threonine. The GalNAc can be extended with residues like galactose, GlcNAc, fucose, or sialic acid. For O-linked GalNAc glycosylation, four principal core structures can be distinguished, core 1 (GalGalNAc), core 2 (GalGlcNAcGalNAc), core 3 (GlcNAcGalNAc), and core 4 (GlcNAc2GalNAc) (Fig. 1). The terminal ends are often further modified, *e.g.* with phosphates, sulfates, carboxylic acids, or sialic acids. O-linked GalNAc glycans play a role in maintaining structures of fully folded proteins, conferring protease stability. Figure 2 shows the biosynthesis of core 1 and core 2 GalNAc O-glycans.

Although the dramatic positive impact of terminal sialic acids on N-linked glycans regarding increased half-life of therapeutic proteins is well established, the possible effect of terminal sialic acids at O-linked GalNAc glycans and the exact appearance of O-linked structures are not well understood.

Human insulin-like growth factor binding protein-6 (IGFBP6) has five O-linked glycosylation sites, and clearance from the blood is decreased for the O-glycosylated form in comparison to the deglycosylated one, indicating a general involvement of O-linked glycans while the role of sialic acids remains unclear. B cell-activating factor receptor 3 (BR3)-Fc has multiple O-linked glycosylation sites and sialylation levels vary in the manufacturing process. Separating the different forms, it could be shown that exposed galactose on the desialylated O-linked glycans of BR3-Fc was associated with rapid clearance due to uptake and degradation in the liver, in particular by non-parenchymal cell mediated clearance. Interestingly, the decrease in clearance rate that was associated with increased sialylated Gal was also observed with increased asialo GalNAc, indicating that the terminal asialo Gal might be the signal for the clearance. Adiponectin, an adipocyte-secreted, insulin-sensitizing hormone, has three putative O-linked glycosylation sites, which are not necessary for multimer formation but plasma clearance of the desialylated protein was accelerated compared with that of sialylated protein. So far, no methods have been described that will affect the sialylation or the structure of O-linked GalNAc glycans during recombinant expression of glycoproteins.

Accordingly, the technical problem underlying the present invention is to provide recombinant glycoproteins having highly or fully sialylated GalNAc O-glycans, as well as cell lines that are capable of recombinantly producing such proteins. Further, respective methods of expressing recombinant glycoproteins, increasing the degree of sialylation of recombinant glycoproteins, and decreasing the micro-heterogeneity of GalNAc O-linked glycans should be provided, as well as uses of the above cell lines for the production of recombinant glycoproteins, increasing the degree of sialylation of recombinant glycoproteins, and decreasing the micro-heterogeneity of GalNAc O-linked glycans.

The solution to the above technical problem is achieved by the embodiments characterized in the claims.

In particular, in a first aspect, the present invention relates to an animal cell line, preferably an insect, avian, or mammalian cell line, more preferably a mammalian, in particular human, cell line that is genetically modified to overexpress a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1).

The term "cell line genetically modified to overexpress a ST3Gal1" as used herein, also in the context of other transferases hereinafter, indicates that upon genetic modification, the individual cells of the cell line display a higher activity of the protein, *e.g.* the sialyltransferase, than they did before the genetic modification.

Genetic modifications that allow the overexpression of a given protein are not particularly limited and are known in the art. In a particular example, the cell line comprises an endogenous gene encoding a ST3Gal 1, such as e.g. human cell lines. In such cases, the cells can be genetically modified by inserting a promoter, enhancing element, and/or stabilizing element into the genome of the cells in a position suitable to cause overexpression of said nucleic acid. This can be done by homologous recombination using TALENS, Zn-finger proteins, CRISPR-CAS9, or other methods known in the art. Thus, in preferred embodiments, the cell line comprises an endogenous gene encoding ST3Gal1 and optionally endogenous genes encoding ST3Gal4 and/or ST6Gal1, and further has at least one genetic element, selected from the group consisting of a promoter, an enhancing element, and a stabilizing element inserted into the genome in one or more position(s) suitable to cause overexpression of ST3Gal1 and optionally ST3Gal4 and/or ST6Gal1. Suitable promoters, enhancing elements and stabilizing elements are not particularly limited and are known in the art. For example, promoters include constitutive promoters, *e.g.* a CMV, EF1alpha, SV40, RSV, UbC, CAG, BOS or PGK promoter, and inducible promoters, *e.g.* tetracycline inducible promoters or other inducible promoters known in the art. Further, enhancing elements (enhancers) include CMV enhancer, ß-globin enhancer, immunoglobulin enhancer, and PGK-enhancer. Furthermore, stabilizing elements (chromatin elements) include matrix

attachment regions (MARS), locus control regions (LCRs), and ubiquitously acting chromatin opening elements (UCOEs).

Alternatively, in cases where the cells do not comprise an endogenous gene encoding a ST3Gal1, or additionally, in cases where the cells do comprise an endogenous gene encoding a ST3Gal1, genetic modification of the cells can be achieved by introducing a nucleic acid, encoding a ST3Gal1 into the cells. Methods for introducing nucleic acids into cells are not particularly limited and are known in the art. For example, said nucleic acids could be introduced in circular or linearized form into the cells by electroporation, nucleofection, microinjection, via viral vectors, *e.g.* lentiviral vectors, reagent based methods, e.g. lipids, calcium phosphate, cationic polymers or other methods known in the art. The nucleic acids can be transiently or stably introduced into the cell by episomal systems or by stable integration of the nucleic acid into the genome. Said nucleic acids can be present in the cells in the form of one or more expression vector(s), *e.g.* pcDNA, pCEP, pLenti, pEntr, pDest, pEF, pEAK, pCMV, pStbl, or other expression vectors known in the art. Expression of the ST3Gal1 can be under the control of a constitutive promoter, *e.g.* a CMV, EF1alpha, SV40, RSV, UbC, CAG, BOS or PGK promoter, the endogenous promoter, or of an inducible promoter, *e.g.* tetracycline inducible promoter or other inducible promoters known in the art. Further, the nucleic acids encoding the ST3Gal1 can be present as one continuous nucleic acid, or can be present as separate nucleic acids, *e.g.* as separate expression vectors. Said nucleic acids can contain, in addition to the coding region and a promoter, suitable restriction sites, Kozak sequences, ribosomal binding sites, chromatin modulating elements, selection cassettes, episomal replication systems, *e.g.* Epstein-Barr Nuclear Antigen and ori P, or SV40 ori and SV40 T-large antigen, internal ribosomal entry sites (IRES), splicing signals, and polyadenylation signals known in the art. Thus, in preferred embodiments, the cell line comprises an exogenous nucleic acid encoding a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1), and optionally exogenous nucleic acids encoding a β-galactoside α-2,3-sialyltransferase 4 (ST3Gal4) and/or a β-galactoside α-2,6-sialyltransferase 1 (ST6Gal1).

Suitable genes encoding a ST3Gal1 for transfection of cell lines are not particularly limited and include any genes from any origin that encode a protein having ST3Gal1 activity, *i.e.* a protein that catalyzes the linkage of sialic acid to a Gal-beta-1,3-GalNAc structure linked to Thr or Ser. Preferably, such genes are selected from ST3Gal1 genes encoding a ST3Gal1 protein comprising in its amino acid sequence, in particular in motif 3 of the C-terminal acceptor binding site, the conserved consensus sequence (**H**/C/R)(**Y**/H/F)(**W**/Y/F)(**E**/D/H/Y), wherein the sequence HYWE (SEQ ID NO: 5) is particularly preferred. Figure 14 shows an alignment of the amino acids sequences of ST3Gal1 of various species, covering amino acids 191 to 340 of human ST3Gal1 comprising the catalytic domain, and displays said consensus sequence. More preferably the gene is selected from ST3Gal1 genes encoding a ST3Gal1 protein comprising an amino acid sequence which is at least 50%, preferably 75% and more preferably 90% % identical to amino acids 263 to 321 of the human ST3Gal1. More preferably, the gene is selected from the group consisting of ST3Gal1 genes derived from a species, selected from the group consisting of *Homo sapiens* (SEQ ID NO: 1), *Pan troglodytes* (SEQ ID NO: 6), *Macaca mulatta* (SEQ ID NO: 7), *Sus scrofa* (SEQ ID NO: 8), *Rattus norvegicus* (SEQ ID NO: 9), *Mus musculus* (SEQ ID NO: 10), *Nannospalax galili* (SEQ ID NO: 11), *Monodelphis domestica* (SEQ ID NO: 12), *Oryctolagus cuniculus* (SEQ ID NO: 13), *Cricetulus griseus* (SEQ ID NO: 14), *Canis familiaris* (SEQ ID NO: 15), *Felis catus* (SEQ ID NO: 16), *Equus caballus* (SEQ ID NO: 17), *Gallus gallus* (SEQ ID NO: 18), *Columba livia* (SEQ ID NO: 19), *Alligator sinensis* (SEQ ID NO: 20), *Latimeria chalumnae* (SEQ ID NO: 21), and *Ciona intestinalis* (SEQ ID NO: 22), wherein mammalian S3Gal1 genes, in particular the human ST3Gal1 gene, are particularly preferred.

In further preferred embodiments, the cell line is further genetically modified to overexpress a β-galactoside α-2,3-sialyltransferase 4 (ST3Gal4), preferably human ST3Gal4, and/or is further genetically modified to overexpress β-galactoside α-2,6-sialyltransferase 1 (ST6Gal1), preferably human ST6Gal1. Respective genetic modifications are preferably as defined above for ST3Gal1. Further, suitable genes encoding ST3Gal4 and suitable genes encoding ST6Gal1 are not particularly limited

and include any genes from any origin that encode a protein having the respective activity. Again, mammalian, in particular human genes are particularly preferred.

The cell lines according to the present invention can be derived from cell lines, *e.g.* mammalian cell lines, known in the art. In preferred embodiments, a cell line of the present invention can be derived from Muscovy Duck cells (AGE.CR®) African green monkey kidney epithelial cells (Vero), Madin Darby canine kidney cells (MDCK), baby hamster kidney cells (BHK), Chinese hamster ovary (CHO) cells, human hepatocarcinoma cell lines (HepG2, Huh7), human embryonic kidney 293 (HEK293) cells, human neuronal precursor cells (AGE1.HN® and NC5T11), human embryonic retinoblasts (Per.C6), myeloma cell lines (HMCLs, MM.1, U266, RPMI8226), CML tumor cell lines (NM, NM-F9,), hybrid HEK293 and lymphoma cell (HKB11), or human amniocytes (CAP; cf. EP 1 230 354 B1), wherein CHO cells, HEK293 cells and CAP cells are preferred, and CAP cells are particularly preferred.

In this context, CAP cells are permanent amniocytic cell lines comprising a nucleic acid encoding the gene products of the adenovirus, in particular adenovirus type 5 (Ad5), E1A and E1B regions. CAP cells are derived from primary human amniocytes that are transformed with a nucleic acid encoding Ad5 E1A and E1B.

Accordingly, in a preferred embodiment, the cell lines according to the present invention can be derived from human primary amniocytes comprising at least one nucleic acid encoding the gene products of the adenoviral E1 and pIX region, preferably E1 and pIX region of adenovirus type 5 (Ad5) from nt. 505 to 4079, in which E1A is under the control of the murine phosphoglycerate kinase (pgk) promoter, while E1B and pIX expression is controlled from their natural promoters. The E1B downstream intron, splice acceptor and polyA signal are replaced by corresponding motifs from SV40.

Any or all of the above preferred and/or specific embodiments described for mammalian cell lines can be combined with each other in any manner.

In a further aspect, the present invention relates to a recombinant glycoprotein having GalNAc O-glycans that are sialylated to a degree of at least 80%.

As used herein, the term "recombinant glycoprotein" indicates that the respective glycoproteins are biotechnologically produced in genetically modified organisms or cells.

The glycoproteins according to the present invention have GalNac O-glycans that are sialylated to a degree of at least 80%, preferably at least 82%, more preferably at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 95.5%, at least 96%, at least 96.5%, at least 97%, at least 97.5%, at least 98%, at least 98.5%,or at least 99%. Preferably, the glycoproteins according to the present invention have GalNac O-glycans that are sialylated to a degree of 95% or above. In this context, the term "O-glycans that are sialylated to a degree of at least X%" as used herein indicates that X% of all terminal GalNac O-glycan monosaccharide moieties in a given glycoprotein preparation are sialic acid (N-acetylneuraminic acid; NeuAc).

In preferred embodiments, at least 80%, preferably at least 82%, more preferably at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 95.5%, at least 96%, at least 96.5%, at least 97%, at least 97.5%, or at least 98% GalNac O-glycans of the glycoproteins of the present invention are core 1 GalNAc O-glycan, i.e. GalNac O-glycans having a core structure of

GP-Ser/Thr-O-GalNAc-Gal-

wherein GP is the glycoprotein, Ser/Thr-O is a serine or threonine amino acid side chain of the glycoprotein, GalNAc is N-acetylgalactosamine, and Gal is galactose (Fig. 1).

In preferred embodiments, the glycoproteins according to the present invention have core 1 GalNAc O-glycans that are disialylated to a degree of at least 25%, preferably

at least 30%, more preferably at least 35%, at least 37.5%, at least 40%, at least 42.5%, or at least 45%. In this context, the term "core 1 GalNAc O-glycans that are disialylated to a degree of at least X%" as used herein indicates that X% of all core 1 GalNAc O-glycans have two terminal sialic acid moieties. In this context, Figure 11 shows the relative amounts of the six main O-glycan species in various preparations of the glycoprotein C1 esterase inhibitor (C1 Inh).

Respective recombinant glycoproteins can be produced as described herein, *e.g.* by overexpression of a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1) together with the recombinant glycoprotein, and optionally by additional overexpression of a β-galactoside α-2,3-sialyltransferase 4 (ST3Gal4) and/or a β-galactoside α-2,6-sialyltransferase 1 (ST6Gal1). Preferably, said glycoproteins are produced in a cell line according to the present invention as described herein.

In a related embodiment, the GalNAc O-glycans of the recombinant glycoproteins of the present invention are characterized by a reduced micro-heterogeneity, *i.e.*, a reduced diversity of GalNAc O-glycan structure and composition within a given glycoprotein preparation. Respective recombinant glycoproteins can be produced as described herein, *e.g.* by overexpression of ST3Gal1 together with the recombinant glycoprotein, and optionally by additional overexpression of ST3Gal4 and/or ST6Gal1. Preferably, said glycoproteins are produced in a cell line according to the present invention as described herein.

Specific glycoproteins according to the present invention are not particularly limited, provided that said glycoproteins have GalNAc O-glycans and fulfill the respective requirements of GalNAc O-glycan sialylation. Preferably, said glycoproteins are mammalian, more preferably human glycoproteins. Glycoproteins can be selected from the group consisting of growth factors, peptide hormones, cytokines, enzymes, antibodies, antibody fragments, blood clotting factors, and protease inhibitors. Preferably, the glycoprotein is selected from the group consisting of hepatocyte growth factor (HGF), erythropoietin (EPO), Factor VIII (FVIII), Factor IX (FIX), von-Willebrand-Factor (vWF), and C1 esterase inhibitor (C1-inhibitor; C1 Inh), wherein C1 Inh is particularly preferred.

Thus, in a particularly preferred embodiment, the present invention relates to a recombinant C1 esterase inhibitor (C1-inhibitor; C1 Inh), preferably recombinant human C1 Inh, having GalNAc O-glycans that are sialylated to a degree of at least 80%. All of the above preferred and/or specific embodiments described for recombinant glycoproteins in general also apply to this specific embodiment of recombinant (human) C1 Inh.

Any or all of the above preferred and/or specific embodiments described for recombinant glycoproteins can be combined with each other in any manner.

In further aspects, the present invention relates to a method for the expression of recombinant glycoproteins having GalNAc O-glycans that are sialylated to a degree of at least 80%, to a method for increasing the degree of sialylation of GalNAc O-glycans of recombinant glycoproteins, and to a method for decreasing the micro-heterogeneity of GalNAc O-glycans of recombinant glycoproteins, all of said methods comprising the step of overexpressing with the recombinant glycoproteins a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1).

In particular embodiments, said methods comprise the steps of:

(a) providing a cell line according to the present invention;

(b) expressing the glycoprotein of interest in said cell line; and

(c) overexpressing a ST3Gal1 and optionally one or more, selected from the group consisting of ST3Gal4, and ST6Gal1, together with the glycoprotein of interest.

Further, the methods of the present invention can comprise the step of

(d) isolating the glycoprotein of interest.

In these aspects, all of the definitions and preferred and/or specific embodiments described for the recombinant glycoproteins of the present invention and the cell lines of the present invention apply in an analogous manner where applicable.

In particular, the recombinant glycoproteins produced in said methods, or the recombinant glycoproteins whose degree of sialylation is increased in said methods,

or the recombinant glycoproteins whose GalNAc O-glycan micro-heterogeneity is decreased in said methods, can be any glycoproteins of the present invention as defined above. Further, the cell lines provided in step (a) of the methods of the present invention can be any cell line of the present invention as defined above.

Means for the expression of proteins in the cell lines of the present invention are not particularly limited and are known in the art. In this context, the step (b) of expressing the glycoprotein of interest in said cell line encompasses the transfection of a respective coding nucleic acid into said cell line prior to the actual expression of the glycoprotein. Further, means for isolating a glycoprotein of interest from a cell culture are not particularly limited and are known in the art.

In related aspects, the present invention relates to a use of a cell line according to the present invention for the production of recombinant glycoproteins having GalNAc O-glycans that are sialylated to a degree of at least 80%, and/or for increasing the degree of sialylation of recombinant glycoproteins and/or for decreasing the micro-heterogeneity of GalNAc O-glycans of recombinant glycoproteins.

In these aspects, all of the definitions and preferred and/or specific embodiments described for the recombinant glycoproteins of the present invention and the cell lines of the present application apply in an analogous manner where applicable.

In particular, the recombinant glycoproteins produced in said use, or the recombinant glycoproteins whose degree of sialylation is increased in said use, or the recombinant glycoproteins whose GalNAc O-glycan micro-heterogeneity is decreased in said use, can be any glycoproteins of the present invention as defined above. Further, the used cell lines of the present invention can be any cell lines of the present invention as defined above.

The figures show:

Figure 1:

Four major core GalNAc O-linked glycosylation structures can be distinguished.

GalNAc O-linked glycan contains an α- linked N-acetylgalactosamine residue linked to serine or threonine. The GalNAc can be extended with residues like galactose, GlcNAc, fucose, or sialic acid. Four principal core structures can be distinguished, core 1 (GalGalNAc), core 2 (GalGlcNAcGalNAc), core 3 (GlcNAcGalNAc), and core 4 (GlcNAc2GalNAc). These core structures can further be extended and branched.

Figure 2:

Biosynthesis of core 1 and core 2 GalNAc O-glycans.

First step in the biosynthesis is the linkage of a GalNAc to a specific serine or threonine residue in the protein backbone, resulting in so called Tn-antigen. The enzyme C1GalT-1 catalyzes then the linkage of a galactose to the GalNAc, forming therefore the basis for the core 1 structure also called T-antigen. Core 1 serves as a substrate for either the C2Gnt, which catalyzes the synthesis of GlcNAc ß1-6 linked branches, resulting in core 2 structures. Core 1 structures can also additionally be sialylated by the ST3Gal1 resulting in mono-sialylated Gal1-3(NeuAc2-6)GalNAc or di-sialylated NeuAc2-3Gal1-3(NeuAc2-6)GalNAc.

Figure 3:

Serum concentration of C1 Inh after a single intravenous injection of recombinant C1 Inh (CAP C1 Inh, CAP C1 Inh ST3Gal4, CAP C1 Inh ST3Gal1/ST3GAL4 pool A and pool B) or Berinert in rats.

Pharmacokinetic studies were performed in rats injected with either Berinert (C1 Inh derived from human plasma) or recombinant C1 Inh, expressed in CAP C1 Inh, CAP C1 Inh ST3Gal4, or CAP C1 Inh ST3Gal1/ST3GAL4. Residual hC1 Inh amounts after i.v. bolus injection of a 10 mg/kg dose in Sprague Dawley female rats were determined at different time points: 5 min, 10 min, 15 min, 20 min, 30 min, 60 min, 2 h, 4 h, 6 h, and 24 h. The percentage of the residual C1 Inh was detected via ELISA. For each animal, C1 Inh concentration was normalized to the concentration at 5 min = 100%. Values were fitted and plotted versus time post injection. From

each group (n=4, except Berinert n=7) a graph for one representative animal is shown.

Figure 4:

Serum half-life of C1 Inh after a single intravenous injection of recombinant C1 Inh (CAP C1 Inh, CAP C1 Inh ST3Gal4, CAP C1 Inh ST3Gal1/ST3GAL4 pool A and pool B) or Berinert in rats.

Pharmacokinetic studies were performed in rats injected with either Berinert (C1 Inh derived from human plasma) or recombinant C1 Inh, expressed in CAP C1 Inh, CAP C1 Inh ST3Gal4, or CAP C1 Inh ST3Gal1/ST3GAL4. Residual hC1 Inh amounts after i.v. bolus injection of a 10 mg/kg dose in Sprague Dawley female rats were determined at different time points: 5 min, 10 min, 15 min, 20 min, 30 min, 60 min, 2 h, 4 h, 6 h, and 24 h. The percentage of the residual C1 Inh was detected via ELISA. For each animal, C1 Inh concentration was normalized to the concentration at 5 min = 100%. Values were fitted and plotted versus time post injection. Shown are the mean values for t1/2 (n=4, Berinert n=7), error bars=±SD.

Figure 5:

Determined AUC values for the varying recombinant C1 Inh (CAP C1 Inh, CAP C1 Inh ST3Gal4, CAP C1 Inh ST3Gal1/ST3GAL4 pool A and pool B) or Berinert.

Comparison of bioavailability between the different recombinant glyco-optimized C1 Inh samples. Shown are the mean values for AUC (n=4, Berinert n=7), error bars=±SD.

Figure 6:

N-linked glycosylation analysis of C1 Inh expressed from CAP cell lines with and without coupled expression of ST3Gal1 and/or ST3Gal4 and add-on western blot analysis.

Purified recombinant C1 Inh expressed in CAP cells migrates slower in SDS PAGE compared to Berinert. Expression of ST3Gal1 and ST3Gal4 causes a shift of the recombinant C1 Inh towards Berinert after removal of N-glycosylation by PNGase, indicating a mass reduction of the remaining O-linked-glycans. 100 ng purified recombinant C1 Inh expressed in glyco-improved C1-expressing cell lines were

digested with 500 U PNGaseF (NEB) for 1 h at 37 °C, followed by a separation on a 4 - 12 % Bis-Tris gel. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 7:

ECL lectin immunoblot of recombinant C1 Inh expressed in CAP cells in the presents or absence of sialyltransferases.

Erythrina crista galli (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore, a diminished signal in the ECL blot means an increased amount of sialylation. N-linked glycans of Berinert are almost completely sialylated, C1 Inh from CAP cells overexpressing ST3Gal4 are also nearly completely sialylated. C1 Inh from CAP cells overexpressing ST3Gal1/4 display a slightly higher amount of asialo N-glycans. In comparison, unmodified C1 Inh purified from CAP cells without overexpression of sialyltransferase provide a strong signal in the ECL blot indicating a certain amount of asialo N-glycans. As loading control same samples were treated with neuraminidase to remove the entire sialic acids content from the glycan structures. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 8:

PNA lectin immunoblot of recombinant C1 Inh expressed in CAP cells in the presents or absence of sialyltransferases.

The amount of sialylation of O-linked glycans of recombinant C1 Inh in CAP cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. C1 Inh purified from cell culture supernatants of CAP cells overexpressing ST3Gal1 displays no signal indicating complete sialylation of the O-glycans. The same is true for plasma derived C1 Inh, Berinert. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 9:

Comparison of rhC1 Inh isoform patterns by IEF analysis.

As the backbones of the different C1 Inh are identical, changes in the IEF are most likely due to changes in the sialic acid content. C1 Inh expressed in CAP with additional expression of ST3Gal4 results in a modified C1 Inh which shifts, in comparison to the unmodified C1 Inh, only slightly towards the anode, indicating a small increase in the total amount of sialic acids per molecule. In contrast, the additional expression of ST3Gal1 results in a pronounced C1 Inh shift towards the anode, indicating a significant increase in the total amount of sialic acids per molecule.

Figure 10:

MALDI-TOF mass spectrum analysis of CAP C1 Inh O-glycans of protein expressed with or without additional overexpression of ST3Gal1 and/or ST3GAL4 in comparison to Berinert.

(A) The analysis of the CAP C1 Inh sample reveals a high abundance of monosialylated O-glycans with a core 2 structure and a terminal galactose. (B) Expression of C1 Inh combined with overexpression of ST3Gal4 also results in large quantities of monosialylated O-glycans with a core 2 structure and a terminal galactose. (C) Expression of C1 Inh combined with overexpression of ST3Gal4 and ST3Gal1 leads to a shift towards core 1 O-glycan structures, which are mono- or di-sialylated, but without any terminal galactose residues. Core 2 structures are barely detectable. (D) O-glyco analysis of Berinert shows the present of only core 1 O-glycans which are mostly mono-sialylated without terminal galactose residues.

Figure 11:

Summary of the MALDI-TOF mass spectrum results of C1 Inh expressed in CAP cells coupled with or without overexpression of ST3Gal1 and/or ST3GAL4 compared to Berinert.

Displayed are the amounts of the six main glycan species in relation to the total sum of the same glycans. Glycan fragments generated by MALDI-TOF or negligible signal were not used for this analysis. In the CAP C1 Inh samples without additional overexpression of ST3Gal1 or ST3Gal4, namely CAP C1 Inh, predominantly O-

glycans with a core 2 structure are detectable (Gal1-3(Gal1-GlcNAc1-6)GalNAc-ol_Core 2) (m/z 983), or NeuAc2-3Gal1-3(Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1344), Gal1-3(NeuAc2-3Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1344) and NeuAc2-3Gal1-3(NeuAc2-3Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1706)). Core 1 structures are only rarely detectable. Additional expression of ST3Gal4 leads to a slight shift towards core 1 structures (NeuAc2-3Gal1-3GalNAc-ol_Core 1). Interestingly, the expression of C1 Inh in CAP cells coupled with the overexpression of ST3Gal1 leads to the exclusive expression of NeuAc2-3Gal1-3GalNAc-ol (Core 1) respectively Gal1-3(NeuAc2-6)GalNAc-ol (m/z 895) and NeuAc2-3Gal1-3(NeuAc2-6)GalNAc-ol (m/z 1256). "Hex" in this figure designates hexose.

Figure 12:

The different core 1 and core 2 GalNAc O-glycan structures.

The MALDI-TOF-mass spectrum analysis revealed that the different samples group in two different clusters of either core 1 or core 2 O-glycan structures.

Figure 13:

ECL and PNA lectin immunoblots of recombinant HGF expressed in CAP cells in the presence or absence of sialyltransferases.

(A) *Erythrina crista galli* (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore, a diminished signal in the ECL blot means an increased amount of sialylation. Overexpression of ST3Gal4 or ST3Gal1/4 results in an increased sialylation of the N-linked glycans, whereas overexpression of ST3Gal1 has no effect compared to unmodified C1 Inh. (B) The amount of sialylation of recombinant HGF in CAP cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. Cell culture supernatant from CAP cells overexpressing ST3Gal1 results in a significant decrease of the signal indicating increased sialylation of the O-glycan. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 14:

Domain structure and sequence alignment of ST3Gal1.

A) Domain structure of ST3Gal1 Sialyltransferases. TM, transmembrane domain; Stem, stem region; L, sialyl motif L (long); sialyl motif S (short); 3, sialyl motif III; VS, sialyl motif VS (very short). The L motif is involved in binding of CMP-Sia, the motif S is involved in binding of CMP-Sia as well as acceptor, motif 3 and VS contain the catalytic consensus sequence and is involved in binding the acceptor. B) Alignment of ST3Gal1 C-terminal sequences from humans, mammals, birds, reptiles, fish, and ascidians. Amino acid residues identical in all species are shaded in light grey, amino acids identical in most species in dark gray and blocks of similar amino acids in medium gray. The consensus sequence for the catalytically active amino acids in motif 3 reads (**H**/C/R)(**Y**/H/F)(**W**/Y/F)(**E**/D/H/Y) (preferred amino acids in corresponding position in bold).

Figure 15:

ECL and PNA lectin immunoblots of recombinant C1 Inh expressed in 293F cells in the presence or absence of sialyltransferases.

(A) *Erythrina crista galli* (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Overexpression of ST3Gal4 or ST3Gal1/4 results in an increased sialylation of the N-linked glycans, whereas overexpression of ST3Gal1 alone has no effect compared to unmodified C1 Inh. Neuraminidase catalyzes the hydrolysis of N-acetyl-neuraminic acid residues from oligosaccharide, thus neuraminidase treated samples serve as positive control. (B) The amount of sialylation of recombinant C1 Inh in 293F cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Cell culture supernatant from 293F cells overexpressing ST3Gal1 results in a significant decrease of the signal indicating increased sialylation of the O-glycan. Neuraminidase treated samples serve as positive control. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 16:

PNA lectin immunoblots of recombinant C1 Inh expressed in CHO-K1 cells in the presence or absence of sialyltransferases.

The amount of sialylation of recombinant C1 Inh in CHO-K1 cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Cell culture supernatant from CHO-K1 cells overexpressing ST3Gal1 show a significant decrease of the signal indicating increased sialylation of the O-glycan compared to C1 Inh purified from CHO-K1 C1 Inh control cells. Western Blot analysis of the same protein samples served as loading control. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

Figure 17:

ECL and PNA lectin immunoblots of recombinant C1 Inh expressed in MDCK.1 cells in the presence or absence of sialyltransferases.

(A) *Erythrina crista galli* (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Overexpression of ST3Gal1 in MDCK.1 cells has no effect on the amount of sialylation of the N-linked glycans. (B) The amount of sialylation of recombinant C1 Inh upon overexpression of ST3Gal1 in MDCK.1 cells was tested by PNA lectin immunoblots. Cell culture supernatant from MDCK.1 cells overexpressing ST3Gal1 show a significant decrease of the signal indicating increased sialylation of the O-glycan. (Molecular weight marker: MagicMark™ XP Western Protein Standard).

The present invention will be further illustrated in the following examples without being limited thereto.

Examples

Experimental procedures:

Cell Culture and Fermentation

The permanent human amniocyte cell line CAP 1D5 was cultured in suspension, either in chemically defined, animal component free CAP-CDM medium (CEVEC Pharmaceuticals, Germany) supplemented with 6 mM stable glutamine (biochrom, Germany), or in serum free PEM media (Life Technologies) supplemented with 4 mM stable glutamine (biochrom, Germany).

The 293F cells from Life Technologies were cultured in suspension in Freestyle 293 expression media (Life Technologies), supplemented with 4 mM stable glutamine (biochrom, Germany).

The adherent CHO-K1 cells (ATCC, CCL-61) were cultured in F12-K media (Life Technologies), supplemented with 10 % FBS, and 2 mM stable glutamine (biochrom, Germany).

The adherent MDCK.1 cells (ATCC, CRL-2935) were cultured either in EMEM media (ATCC) or in DMEM F-12 (ATCC) supplemented with 10 % FBS and with 2 mM stable glutamine (biochrom, Germany).

CAP cells and 293F cells were cultivated at 37 °C in shaker flasks (Corning, 125 mL (25 mL wv) or 3000 mL (1000 mL wv)) at 5% CO_2, and 185 rpm. During fermentation CAP cells were fed at d3, d5, and d7 with 10 % CAP-CDM feed solution (CEVEC Pharmaceuticals, Germany) and 4 mM stable glutamine (biochrom, Germany). Adherent CHO-K1 and MDCK.1 cells were cultured at 37 °C, 5% CO_2, in 6 cm or 10 cm cell culture dishes (TPP) or 225cm^2 cell culture dishes (BD).

Cloning

For the generation of the cell lines used in the present invention, cells were sequentially nucleofected with the nucleic acid constructs encoding the glycostructure modifying enzymes ST3Gal1 and/or ST3Gal4, as well as the specific recombinant protein. Only stable cell lines were utilized. Table 1 lists all cell lines created.

Table 1: Stable cell lines used in the present invention.

Cell line	rec. protein	overexpression of the sialyltransferase(s)
CAP-C1 Inh	C1 Inh	/
CAP-C1 Inh-ST3Gal1	C1 Inh	ST3Gal1
CAP-C1 Inh-ST3Gal4	C1 Inh	ST3Gal4
CAP-C1 Inh-ST3Gal1/4	C1 Inh	ST3Gal1/ ST3Gal4
CAP-HGF	HGF	/
CAP-HGF-ST3Gal1	HGF	ST3Gal1
CAP-HGF-ST3Gal4	HGF	ST3Gal4
CAP-HGF-ST3Gal1/4	HGF	ST3Gal1/ ST3Gal4
293F-C1 Inh	C1 Inh	/
293F -C1 Inh-ST3Gal1	C1 Inh	ST3Gal1
293F -C1 Inh-ST3Gal1/4	C1 Inh	ST3Gal1/ ST3Gal4
CHO-C1 Inh	C1 Inh	/
CHO-C1 Inh-ST3Gal1	C1 Inh	ST3Gal1
MDCK.1-C1 Inh	C1 Inh	/
MDCK.1-C1 Inh-ST3Gal1	C1 Inh	ST3Gal1

For designing the ST3Gal1 cDNA, sequence information of the precursor protein and mature protein was based of the database entry UniProt Q11201 (SEQ ID NO: 1). For cloning, a ClaI restriction site and a Kozak sequence were added 5´ of the start codon of the human ST3Gal1 cDNA and an EcoRV restriction site was added 3´ of the stop codon to be inserted between the ClaI and EcoRV restriction sites in the pStbl-Puro-CMV-MCS(-) vector resulting in the expression plasmid pStbl-Puro-CMV-ST3Gal1. This vector contains a CMV promoter driving the expression of the gene of interest, followed by an SV40 intron for improved, splicing-mediated mRNA transport and a multiple cloning site for the insertion of the gene of interest. The selection marker (Puromycin) is driven by the human ubiquitin (UbC) promoter. cDNA synthesis was performed at GeneArt (Germany, Life Technologies).

For designing the ST3Gal4 cDNA, sequence information of the precursor protein and mature protein was based of the database entry UniProt Q11206 (SEQ ID NO: 2). For cloning, a ClaI restriction site and a Kozak sequence were added 5´ of the start codon of the human ST3Gal4 cDNA and a EcoRV restriction site was added 3´ of the stop codon to be inserted between the ClaI and EcoRV restriction sites in the pStbl-Puro-CMV-MCS(-) vector resulting in the expression plasmid pStbl-Puro-CMV-ST3Gal4. cDNA synthesis was performed at GeneArt (Germany, Life Technologies).

Nucleofection and Pool Generation

Nucleofection was performed using a Nucleofector (LONZA) with the appropriate Nucleofector Kit (KitV; CAP cells, 293F, and CHO or the KitT; MDCK.1 cells) according to the manufacturer's protocol. Briefly, during exponential growth phase of the culture 1×10^7 cells were harvested via centrifugation (150 g for 5 min) and resuspended in 100 µl complete nucleofector solution and mixed with a total of 5 µg plasmid. Nucleofection was performed using the X001 program (CAP and 293F cells), the U024 program (CHO-K1), or the P029 program (MDCK.1). After the pulse, cells were recovered in complete cell culture media. The cells were cultured as before.

72 to 96 h post-nucleofection cells were selected with 5 µg/ml blasticidin (therapeutic protein) and/or 200 µg/ml neomycin (pStbl-neo-CMV-ST3Gal4) and/or 2 µg/ml puromycin (pStbl-Puro-CMV-ST3Gal1), in order to generate stable pools.

Pharmacokinetic studies of recombinant C1 Inh in rats

Comparable pharmacokinetic studies were performed in rats injected with either Berinert (C1 Inh derived from human plasma) or purified recombinant C1 Inh, expressed in CAP C1 Inh, CAP C1 Inh ST3Gal4, or CAP C1 Inh ST3Gal1/ST3GAL4.

Residual hC1 Inh concentrations after i.v. bolus injection of a 10 mg/kg dose in Sprague Dawley female rats were determined at different time points: 5 min, 10 min, 15 min, 20 min, 30 min, 60 min, 2 h, 4 h, 6 h, and 24 h.

In Figure 3 the serum concentration of C1 Inh after a single intravenous injection of recombinant C1 Inh or Berinert is shown. The percentage of the residual C1 Inh was detected via ELISA. For each animal C1 Inh concentrations were normalized to the concentration at 5 min = 100%. Values were fitted and plotted versus time post injection. From each group (n=4, except Berinert n=7) a graph for one representative animal is shown. In Figure 5 the area under the curve (AUC) calculations for the different forms of C1 Inh are displayed.

MALDI-TOF-mass spectrum analysis

In order to desalt the samples, 100 µg protein were precipitated twice using chloroform-methanol and dried by vacuum rotation. Glycans were β-eliminated in 50 µL $NaBH_4$ (1 M in 50 mM NaOH) under argon by overnight incubation at 50 °C. After desalting with Dowex50 × 8 (H+) and the removal of borate by co-distillation of the methylester from acidified methanol, the dried residues were methylated as known in the art. MALDI-MS was performed on an UltrafleXtreme instrument (Bruker Daltonics). The methylated glycan alditols were applied to the stainless steel target by mixing 1:1 with matrix (α-cyano-4-hydroxycinnamic acid in 50 % acetonitrile/0.1 % TFA). Analyses were performed by positive ion detection in MS1 and MS2 (Post-Source-Decay-Modus). Identification of the different glycan species were based i) on the molecule ions, which give information about the monosaccharide composition and ii) the fragmentation in the MS2 analysis (B, C, Y, and Z ions).

PNGaseF digestion

PNGaseF is an amidase that cleaves between the innermost GlcNAc and asparagine residues of N-linked glycoproteins. Therefore, after treatment of C1 Inh

proteins with PNGaseF the protein backbone and the O-linked glycans remain. As the protein backbone between the plasma derived human C1 Inh and C1 Inh expressed in CAP cells is equal, the PNGaseF treatment carries indirectly information about the structure of the O-glycans.

The PNGaseF digestion was done as described in the manufacturer's instructions. Briefly, 100 ng purified protein were incubated with 500 U PNGaseF (NEB) for 1 h at 37 °C. Subsequently, samples were separated on a NuPAGE Novex 4-12 % Bis-Tris Gel under reducing conditions, according to the manufacturer's instructions. The separated proteins were transferred via an Blot Module (Invitrogen) (30 V for 60 min. at RT) onto an Amersham Hybond ECL membrane (100 V for 60 min at RT). The membrane was blocked for 1 h at RT with PBSTB (phosphate-buffered saline, pH = 7.4, supplemented with 0.1% Tween 20 and 1 % BSA). Afterwards, the membrane was incubated with the mouse monoclonal C1 Inh specific HRP-labeled antibody diluted 1:10000 in PBSTB. After washing the membrane with PBST (phosphate-buffered saline pH = 7.4 supplemented with 0.1% Tween 20), the proteins were detected using the Pierce ECL WB Substrate Kit via a chemiluminescence detector (INTAS).

Lectin Immunoblotting

Lectins are proteins that bind specific carbohydrate structures. Biotin-coupled lectins can be therefore used to analyze N-linked and O-linked glycans. *Erythrina crista galli* (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans, peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans, *Sambucus nigra* agglutinin (SNA) preferentially binds to a 2,6-linked sialic acid, whereas *Maackia amurensis* lectin (MAL) preferentially binds to 2,3-linked sialic acids.

Purified protein or cell culture supernatants with or without co-expression of ST3Gal1 and/or ST3Gal4 were separated as described above and blotted onto Hybond ECL nitrocellulose membrane (GE healthcare). The membrane was

blocked for 1 h at RT with PBSTB (phosphate-buffered saline, pH = 7.4, supplemented with 0.1% Tween 20 and 1 % BSA). Afterwards, the membrane was incubated with the lectin diluted 1:2000 (MAL 1:400) in PBSTB. After washing the membrane with PBST (phosphate-buffered saline, pH = 7.4, supplemented with 0.1% Tween 20), the membrane was stained with streptavidin-coupled horseradish peroxidase for 1 h at RT (1:1000 diluted in PBSTB). The HRP signal was amplified using anti-streptavidin IgG and anti IgG-HRP. The proteins were detected using the Pierce ECL WB Substrate Kit via a chemiluminescence detector (INTAS).

IEF analysis

Isoelectric focusing (IEF) was performed in order to analyze the isoelectric point (pI) of C1 Inh purified from CAP C1 Inh cells, or CAP C1 Inh cells transfected with ST3Gal4 and/or ST3Gal1, respectively. The degree of sialylation correlates with a given proteins acidity and therefore, with its pI. IEF analysis was done according to the manufacturers protocol (Invitrogen). Briefly, 5 µg of purified protein were loaded on pH3-7 gels and subjected to electrophoresis (1h 100 V, 1 h 200 V, 30 min 500 V). Proteins were stained with SimplyBlue SafeStain according to the manufacturers protocol (Invitrogen).

Example 1:

The protease inhibitor C1 esterase inhibitor (C1 Inh) belongs to the serpin superfamily. Its main function is the inhibition of the complement system to prevent spontaneous activation. The 500 aa protein is highly glycosylated with 7 predicted N-glycans and 8 predicted O-linked glycans.

Cells of the human amniocyte cell line CAP 1D5 previously stably transfected to express human recombinant C1 Inhibitor (rhC1 Inh, SEQ ID NO: 3) (CAP-C1 Inh) were nucleofected with a vector encoding for ST3Gal4 linearized with ScaI in order to facilitated stable integration of the construct into the genome. The vector contains

a drug expression cassette, for selection of stable integration of the linearized construct into the genome. After pool generation, the obtained stable CAP-C1 Inh-ST3Gal4 pool was subject to single cell cloning via limiting dilution. Selected clones were analyzed to proof expression of ST3Gal4. CAP single cell clones expressing ST3Gal4 were then further nucleofected with the gene encoding for ST3Gal1 gene. Cells were selected with antibiotics to obtain a pool of cells stably co-expressing rhC1 Inh, human ST3Gal4 and ST3Gal1 (CAP-C1 Inh-ST3Gal1/4).

In order to generate sufficient amounts of recombinant C1 Inh, the generated CAP-C1 Inh cell lines overexpressing human C1 Inh were cultured as described in experimental procedures. Subsequently, C1 Inh was purified from the cell culture supernatant of the following cell lines: CAP-C1 Inh, CAP-C1 Inh-ST3Gal4, and CAP-C1 Inh-ST3Gal1/4 (pool A and B), as described below.

Cell culture supernatants or purified recombinant C1 Inh and Berinert were analyzed to determine glycostructures.

Immunoblotting with a C1 Inh specific antibody (Figure 6) reveals that C1 Inh purified from plasma (Berinert) migrates faster during electrophoresis, indicating a diminished mass in comparison to recombinant C1 Inh from CAP cells. Overexpression of the sialyltransferases ST3Gal1 in combination with ST3Gal4 likewise increases migration speed. This effect is even more apparent after cleavage of N-linked glycans from the protein backbone by PNGase digestions, indicating that the faster migration is actually due to size differences in the remaining O-linked glycans.

To determine the degree of sialylation of the N-linked glycans, an ECL lectin immunoblot was performed. *Erythrina crista galli* (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore, a diminished signal in the ECL blot means an increased amount of sialylation. As shown in Figure 7, the N-linked glycans of Berinert are almost completely sialylated. N-Glycans of the C1 Inh from CAP cells overexpressing ST3Gal4 are equally sialylated, C1 Inh from CAP cells overexpressing ST3Gal1/4 display a slightly higher amount of asialo N-glycans. In

comparison, C1 Inh purified from CAP cells without overexpression of sialyltransferase give a strong signal in the ECL blot indicating a large amount of asialo N-glycans.

The degree of sialylation of O-linked glycans of recombinant C1 Inh in CAP cells upon overexpression of ST3Gal1 was initially tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a diminished signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. As depicted in Figure 8, C1 Inh purified from cell culture supernatants of CAP cells overexpressing ST3Gal1 displays no signal indicating complete sialylation of the O-glycans. The same is true for plasma derived C1 Inh, Berinert.

Results from the lectin blots could be confirmed by isoelectric focusing (IEF). As the backbones of the different C1 Inh are identical, changes in the IEF are most likely due to changes in the sialic acid content. As shown in Figure 9, by additional expression of ST3Gal4, the resulting modified C1 Inh shifts very slightly towards the anode, indicating a small increase in the total amount of sialic acids per molecule. In contrast, the additional expression of ST3Gal1 results in a pronounced shift of C1 Inh towards the anode, indicating a significant increase in the total amount of sialic acids.

In Figure 10, detailed MALDI-TOF-mass spectra of the O-linked glycans are displayed. In Figure 11, these data are summarized by outlining the peak intensity for the different O-glycan structures. Figure 12 shows an interpretation of the obtained data. In the CAP C1 Inh samples without additional overexpression of ST3Gal1 or ST3Gal4, namely CAP C1 Inh, 75 % of the detectable O-glycans are core 2 structure with only 35 % sialylated O-glycans without terminal galactose residues(Gal1-3(Gal1-GlcNAc1-6)GalNAc-ol_Core 2) (m/z 983), or NeuAc2-3Gal1-3(Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1344), Gal1-3(NeuAc2-3Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1344) and NeuAc2-3Gal1-3(NeuAc2-3Gal1-4GlcNAc1-6)GalNAc-ol (m/z 1706)). Core 1 structures are only rarely detectable. Additional expression of ST3Gal4 leads to a shift towards core 1 structures (68.7%)(NeuAc2-

3Gal1-3GalNAc-ol_Core 1). Interestingly, the expression of C1 Inh in CAP cells coupled with the overexpression of ST3Gal1 leads to the exclusive expression of core 1 O-glycans (99.1%) which are almost entirely sialylated (98.5%) without any terminal galactose residues, in particular (NeuAc2-3Gal1-3GalNAc-ol respectively Gal1-3(NeuAc2-6)GalNAc-ol (m/z 895) and NeuAc2-3Gal1-3(NeuAc2-6)GalNAc-ol (m/z 1256).

In order to determine the serum half-life of the different glyco-improved recombinant C1 Inh, CAP-C1 Inh-ST3Gal4, and CAP-C1 Inh-ST3Gal1/4, versus recombinant C1 Inh from CAP cells (wild-type), or plasma derived C1 Inh (Berinert), pharmacokinetic studies were performed.

After normalizing the concentration time curves of the different samples, it became clear that the different samples cluster into two distinct groups with almost the same shape and curve progression: on the one hand, the CAP C1 Inh without further expression of any sialyltransferases together with CAP-C1 Inh-ST3Gal4 and on the other hand CAP-C1 Inh-ST3Gal1/4 together with the plasma derived human C1 Inh, Berinert (Figure 3). The substances in one group are eliminated from the blood stream in the same manner following the same pharmacokinetics.

While overexpression of ST3Gal4 alone had no beneficial effect on serum half-life of the recombinantly expressed C1 Inh, the additional co-expression of ST3Gal1 enhanced serum half-life about 6-fold (Figure 4). AUC and therefore the bioavailability increase around 6 fold compared to the unmodified C1 Inh and is equivalent to the AUC values measured for Berinert (Figure 5). Briefly, these results point out that for glycoproteins exhibiting N-glycans as well as O-glycans, broad sialylation of the O-glycans is as important as it is for the N-glycans.

Example 2:

Hepatocyte growth factor is a potent mitogen for mature parenchymal hepatocyte cells, seems to be a hepatotrophic factor, and acts as a growth factor for a broad

spectrum of tissues and cell types. The 728 aa sized protein contains 4 predicted N-glycans and 1 predicted O-linked glycan.

CAP 1D5 cells stably expressing human recombinant hepatocyte growth factor (SEQ ID NO: 4) were nucleofected either with a vector encoding for ST3Gal4, ST3Gal1, or ST3Gal4 and ST3Gal1 linearized with ScaI in order to facilitate stable integration of the construct into the genome. The vectors contain a drug expression cassette, which facilitates selection for cells with a stable integration of the linearized construct into the genome. After pool generation the obtained stable CAP cell pools, CAP-HGF, CAP-HGF-ST3Gal1, CAP-HGF-ST3Gal4, and CAP-HGF-ST3Gal1/4 were cultured as described in experimental procedures. Cell culture supernatants containing recombinant human HGF were examined by ECL and PNA lectin blots to determine the existing N- and O-linked glycostructures.

Erythrina crista galli (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore a diminished signal in the ECL blot means an increased amount of sialylation. As shown in Figure 13, overexpression of ST3Gal4 or ST3Gal1/4 results in an increased sialylation of the N-linked glycans, whereas overexpression of ST3Gal1 has no effect.

The amount of sialylation of recombinant HGF in CAP cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. As depicted in Figure 13, overexpression of ST3Gal1 alone or in combination with ST3Gal4 resulted in a significant increase in sialylation of the O-glycan of recombinant HGF.

Example 3:

The following experiment was executed in order to investigate if the observed increased sialylation of the O-glycans of glycoproteins upon overexpression of the

sialyltransferase ST3Gal1 is common feature shared by diverse cell lines utilized for the manufacture of recombinant proteins or viruses for pharmaceutical production and/or biomedical research.

293F cells (Life Technologies, R-970-07), immortalized human embryonic kidney cells, were nucleofected either with a vector encoding for ST3Gal1, or ST3Gal1 and ST3Gal4 linearized with ScaI in order to facilitate stable integration of the construct into the genome. The vectors contain a drug expression cassette, which facilitates selection for cells with a stable integration of the linearized construct into the genome. After pool generation the obtained stable 293F cell pools, 293F-ST3Gal1, and 293F-ST3Gal1/4 and wildtype 293F cells were further nucleofected with the gene encoding for hC1 Inh. Cells were selected with antibiotics to obtain pools of cells stably expressing i) rhC1 Inh, ii) rhC1 Inh and human ST3Gal1, iii) rhC1 Inh, human ST3Gal1, and human ST3Gal4. C1 Inh was purified from C1 Inh containing cell culture supernatant as described in the method section and examined by ECL and PNA lectin blots to determine the existing N- and O-linked glycostructures.

Erythrina crista galli (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore a diminished signal in the ECL blot means an increased amount of sialylation. As shown in Figure 15, overexpression of ST3Gal1/4 results in an increased sialylation of the N-linked glycans, whereas overexpression of ST3Gal1 alone has no effect.

The amount of sialylation of recombinant C1 Inh in 293F cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. As depicted in Figure 15, overexpression of ST3Gal1 alone or in combination with ST3Gal4 resulted in a significant increase in sialylation of the O-glycan of recombinant human C1Inh.

Example 4:

CHO-K1 (ATCC, CCL-61) cell line was derived as a subclone from the parental CHO cell line initiated from a biopsy of an ovary of an adult Chinese hamster.

In order to investigate if the observed increased sialylation of O-linked glycans upon overexpression of human sialyltranserase ST3Gal1 also takes place in non-human mammalian cell lines, human C1 Inh was expressed in these cells in the presence or absence of the sialyltransferase ST3Gal1.

CHO-K1 cells were nucleofected with a vector encoding for ST3Gal1 linearized with ScaI in order to facilitate stable integration of the construct into the genome. The vector contains a drug expression cassette, which facilitates selection for cells with a stable integration of the linearized construct into the genome. After pool generation the obtained stable CHO-K1 cell pool, CHO-ST3Gal1, and wildtype CHO-K1 cells were further nucleofected with the gene encoding for hC1 Inh. Cells were selected with antibiotics to obtain pools of cells stably expressing i) rhC1 Inh, ii) rhC1 Inh and human ST3Gal1.

Cells were expanded as described in the method section. For the production of cell culture supernatant containing human C1 Inh cells were seeded in 10 cm cell culture dishes, 3 d post seeding cells were extensively washed with 1xPBS in order to remove fetal bovine serum, subsequently adding fresh serum free media. Four days later cell culture supernatant were harvest, cells and cell debris were removed by centrifugation and filtration via a 0.22µm filter. C1 Inh was purified as described in the method section.

The amount of sialylation of O-glycans of recombinant C1 Inh in CHO-K1 cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. As shown in Figure 16,

overexpression of ST3Gal1 alone resulted in a significant increase in sialylation of the O-glycan of recombinant human C1Inh.

Example 5:

Canine MDCK.1 cells (ATCC, CRL-2935) were stably transfected. The resulting stable MDCK.1 cell pool, MDCK.1 -ST3Gal1, and wildtype MDCK.1 cells were further stably transfected to obtain pools of cells stably expressing *i*) rhC1 Inh, *ii*) rhC1 Inh and human ST3Gal1.

MDCK.1 cells were expanded as described in the method section. For the production of cell culture supernatant containing human C1 Inh cells were seeded onto 225 cm^2 cell culture dishes, 2 d post seeding cells were extensively washed with 1xPBS in order to remove fetal bovine serum, subsequently adding fresh serum free media. Five days later cell culture supernatant were harvest, cells and cell debris were removed by centrifugation and filtration via a 0.22µm filter. C1 Inh was purified as described in the method section and examined by ECL and PNA lectin blots to determine the existing N- and O-linked glycostructures.

Erythrina crista galli (ECL) lectin detects ß1-4 linked terminal galactose on N-linked glycans. Therefore a diminished signal in the ECL blot means an increased amount of sialylation. As shown in Figure 17, overexpression of ST3Gal1 does not result in an increased sialylation of the N-linked glycans.

The amount of sialylation of O-glycans of recombinant C1 Inh in MDCK.1 cells upon overexpression of ST3Gal1 was tested by PNA lectin immunoblots. Peanut agglutinin (PNA) detects ß1-3 linked terminal galactose on O-linked glycans. Therefore, a decreased signal in the PNA lectin blot implies an increased level of sialylation of the galactose residues of O-glycans. As shown in Figure 17, overexpression of ST3Gal1 alone resulted in a significant increase in sialylation of the O-glycans of recombinant human C1Inh.

Discussion:

In the experiments presented above, an ST3Gal1 catalyzing the transfer of sialic acid from CMP-sialic acid to galactose-containing substrates was overexpressed in mammalian cells, either alone or in combination with an ST3Gal4.

Overexpression of ST3Gal1 resulted in nearly complete sialylation of co-expressed O-glycosylated recombinant proteins. Additionally, the heterogeneity of the GalNAc O-glycans was significantly reduced.

Surprisingly, the overexpression of ST3Gal4 alone had no effect on the pharmacokinetic profile of the examined glycoprotein, whereas additional overexpression of ST3Gal1 and the resulting increase in the degree of sialylation of the GalNAc O-glycans increased the serum half-life about 6-fold.

Modifying the glycan structures of therapeutic proteins in order to improve their pharmacokinetic profile is a very potent tool. In the case of the present invention, in particular two common weaknesses of therapeutic proteins are addressed, first the limited serum half-life and second the heterogeneity of glyco-structures. This can be achieved by a forced expression of ST3Gal1 enzyme resulting in secreted glycoproteins showing nearly complete sialylation of GalNAc O-glycans.

This invention is not limited to one particular cell line from one host but is rather applicable to a broad range of animal cell lines. Additionally, it is not restricted to one particular group of glycoproteins but is applicable to a wide range of glycoproteins containing at least one GalNAc O-linked glycan, *e.g.* growth factors, peptide hormones, cytokines, enzymes, antibodies, antibody fragments, blood clotting factors, or protease inhibitors.

The present invention relates to the following amino acid and nucleotide sequences.

SEQ ID NO: 1

Human ST3Gal1

```
MVTLRKRTLK VLTFLVLFIF LTSFFLNYSH TMVATTWFPK QMVLELSENL KRLIKHRPCT  60
CTHCIGQRKL SAWFDERFNQ TMQPLLTAQN ALLEDDTYRW WLRLQREKKP NNLNDTIKEL 120
FRVVPGNVDP MLEKRSVGCR RCAVVGNSGN LRESSYGPEI DSHDFVLRMN KAPTAGFEAD 180
VGTKTTHHLV YPESFRELGD NVSMILVPFK TIDLEWVVSA ITTGTISHTY IPVPAKIRVK 240
QDKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSMHV CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDADF ESNVTATLAS INKIRIFKGR                       340
```

SEQ ID NO: 2

Human ST3Gal4

```
MVSKSRWKLL AMLALVLVVM VWYSISREDR YIELFYFPIP EKKEPCLQGE AESKASKLFG  60
NYSRDQPIFL RLEDYFWVKT PSAYELPYGT KGSEDLLLRV LAITSSSIPK NIQSLRCRRC 120
VVVGNGHRLR NSSLGDAINK YDVVIRLNNA PVAGYEGDVG SKTTMRLFYP ESAHFDPKVE 180
NNPDTLLVLV AFKAMDFHWI ETILSDKKRV RKGFWKQPPL IWDVNPKQIR ILNPFFMEIA 240
ADKLLSLPMQ QPRKIKQKPT TGLLAITLAL HLCDLVHIAG FGYPDAYNKK QTIHYYEQIT 300
LKSMAGSGHN VSQEALAIKR MLEMGAIKNL TSF                              333
```

SEQ ID NO: 3

Human C1 Inh

```
MASRLTLLTL LLLLLAGDRA SSNPNATSSS SQDPESLQDR GEGKVATTVI SKMLFVEPIL  60
EVSSLPTTNS TTNSATKITA NTTDEPTTQP TTEPTTQPTI QPTQPTTQLP TDSPTQPTTG 120
SFCPGPVTLC SDLESHSTEA VLGDALVDFS LKLYHAFSAM KKVETNMAFS PFSIASLLTQ 180
VLLGAGENTK TNLESILSYP KDFTCVHQAL KGFTTKGVTS VSQIFHSPDL AIRDTFVNAS 240
RTLYSSSPRV LSNNSDANLE LINTWVAKNT NNKISRLLDS LPSDTRLVLL NAIYLSAKWK 300
TTFDPKKTRM EPFHFKNSVI KVPMMNSKKY PVAHFIDQTL KAKVGQLQLS HNLSLVILVP 360
QNLKHRLEDM EQALSPSVFK AIMEKLEMSK FQPTLLTLPR IKVTTSQDML SIMEKLEFFD 420
FSYDLNLCGL TEDPDLQVSA MQHQTVLELT ETGVEAAAAS AISVARTLLV FEVQQPFLFV 480
LWDQQHKFPV FMGRVYDPRA                                            500
```

SEQ ID NO: 4

Human HGF

MWVTKLLPAL LQHVLLHLL LLPIAIPYAE GQRKRRNTIH EFKKSAKTTL IKIDPALKIK 60
TKKVNTADQC ANRCTRNKGL PFTCKAFVFD KARKQCLWFP FNSMSSGVKK EFGHEFDLYE 120
NKDYIRNCII GKGRSYKGTV SITKSGIKCQ PWSSMIPHEH SFLPSSYRGK DLQENYCRNP 180
RGEEGGPWCF TSNPEVRYEV CDIPQCSEVE CMTCNGESYR GLMDHTESGK ICQRWDHQTP 240
HRHKFLPERY PDKGFDDNYC RNPDGQPRPW CYTLDPHTRW EYCAIKTCAD NTMNDTDVPL 300
ETTECIQGQG EGYRGTVNTI WNGIPCQRWD SQYPHEHDMT PENFKCKDLR ENYCRNPDGS 360
ESPWCFTTDP NIRVGYCSQI PNCDMSHGQD CYRGNGKNYM GNLSQTRSGL TCSMWDKNME 420
DLHRHIFWEP DASKLNENYC RNPDDDAHGP WCYTGNPLIP WDYCPISRCE GDTTPTIVNL 480
DHPVISCAKT KQLRVVNGIP TRTNIGWMVS LRYRNKHICG GSLIKESWVL TARQCFPSRD 540
LKDYEAWLGI HDVHGRGDEK CKQVLNVSQL VYGPEGSDLV LMKLARPAVL DDFVSTIDLP 600
NYGCTIPEKT SCSVYGWGYT GLINYDGLLR VAHLYIMGNE KCSQHHRGKV TLNESEICAG 660
AEKIGSGPCE GDYGGPLVCE QHKMRMVLGV IVPGRGCAIP NRPGIFVRVA YYAKWIHKII 720
LTYKVPQS 728

SEQ ID NO: 5

ST3Gal1 motif 3 consensus sequence

HYWE 4

SEQ ID NO: 6

Pan troglodytes ST3Gal1

MVTLRKRTLK VLTFLVLFIF LTSFFLNYSH TMVATTWFPK QMVLELSENL KRLIKHRPCT 60
CTHCIGQRKL SAWFDERFNQ TVQPLLTAQN ALLEDDTYRW WLRLQREKKP NNLNDTIKEL 120
FRVVPGNVDP MLEKRSVGCR RCAVVGNSGN LRESSYGPEI DSHDFVLRMN KAPTAGFEAD 180
VGTKTTHHLV YPESFRELGD NVSMILVPFK TIDLEWVVSA ITTGTISHTY VPVPAKIRVK 240
QDKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSMHV CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDADF ESNVTATLAS INKIRIFKGR 340

SEQ ID NO: 7

Macaca mulatta ST3Gal1

```
MVTLRKRTLK VLTFLVLFIF LTSFFLNYSH TMVTTTWFPK QMVLELSENL KRLIKHRPCT  60
CTHCIGQRKL SVWFDERFNQ TVQPLLTAQN ALLEDDTYRW WLRLQREKKP NNLNDTIKEL 120
FRVVPGNVDP MLEKRSVGCR RCAVVGNSGN LRESSYGPEI DRHDFVLRMN KAPTAGFEAD 180
VGTKTTHHLV YPESFRELGD NVSMILVPFK TIDLEWVVSA TTTGTISHTY VPVPAKIRVK 240
QDKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSMHV CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDADF ESNVTATLAS INKIRIFKGR                       340
```

SEQ ID NO: 8

Sus scrofa ST3Gal1

```
MAPMRKKSTL KLLTLLVLFI FLTSFFLNYS HTVVTTAWFP KQMVIELSEN FKKLMKYPYR  60
PCTCTRCIEE QRVSAWFDER FNRSMQPLLT AKNAHLEEDT YKWWLRLQRE KQPNNLNDTI 120
RELFQVVPGN VDPLLEKRLV SCRRCAVVGN SGNLKESYYG PQIDSHDFVL RMNKAPTEGF 180
EADVGSKTTH HFVYPESFRE LAQEVSMILV PFKTTDLEWV ISATTTGRIS HTYVPVPAKI 240
KVKKEKILIY HPAFIKYVFD RWLQGHGRYP STGILSVIFS LHICDEVDLY GFGADSKGNW 300
HHYWENNPSA GAFRKTGVHD GDFESNVTTI LASINKIRIF KGR                   343
```

SEQ ID NO: 9

Rattus norvegicus ST3Gal1

```
MVNMRKRTLK YLTFFLLFIF LTSFVLNYSN SGVPSAWFPK QMVLEFSENF RKFIKSQPCT  60
CRHCISQGKV SYWFDQRFNK TMQPLLTAHN ALMEEDTYRW WLRLQRERKP NNLSDTVKEL 120
FRLVPGNVDP MLNKRLVGCR RCAVVGNSGN LKDSSYGPEI DSHDFVLRMN RAPTVGFEAD 180
VGSRTTHHLV YPESFRELGE NVNMVLVPFK ITDLQWVISA TTTGTITHTY VPVPPKIKVK 240
QEKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSIHI CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDGDF EYNVTTTLAA INKIRIFKGR                       340
```

SEQ ID NO: 10

Mus musculus ST3Gal1

```
MRRKTLKYLT FFLLFIFLTS FVLNYSNTGV PSAWFPKQML LELSENFRRF IKSQPCTCRH  60
CISQDKVSYW FDQRFNKTMQ PLLTVHNALM EEDTYRWWLR LQRERKPNNL SDTVKELFRL 120
VPGNVDPMLN KRLVGCRRCA VVGNSGNLKD SSYGPEIDSH DFVLRMNKAP TVGFEADVGS 180
RTTHHLVYPE SFRELGENVN MVLVPFKTTD LQWVISATTT GTITHTYVPV PPKIKVKQEK 240
ILIYHPAFIK YVFDNWLQGH GRYPSTGILS IIFSIHICDE VDLYGFGADS KGNWHHYWEN 300
NPSAGAFRKT GVHDGDFEYN ITTTLAAINK IRIFKGR                          337
```

SEQ ID NO: 11

Nannospalax galili ST3Gal1

```
MVNLRKKIVK WLTFLLLFVF LTSCFLNYSN SGVPITWFPK QMVLELSENF QKLIKQRPCT  60
CTHCISQSKV SSWFDQRFNQ TMQPLLTASN AMMEEDTYQW WLRLQRERKP NNLSDIVKEL 120
FSLVPGNVDP VLDKRSVGCR RCAVVGNSGN LRASSYGSDI DSHDFVLRMN RAPTVGFEAD 180
VGSRTTHHLV YPESFRELGE NVNMVLVPFK TTDLQWVISA TTTGTITHTY VPVPPKIKVK 240
QEKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSMHV CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDGDF ESNVTTTLAS INKIRIFKGR                       340
```

SEQ ID NO: 12

Monodelphis domestica ST3Gal1

```
MAAIKKKRLK VFTFVLLLVS LTSFFLNYAH TTATYTWFPK QMVMHFSEHF KRFMKYPQRP  60
CSCSQCISET GFAPWFDERF NHTMQPLLNR QNAFLENDTY TWWMKLQRER TPKRLNETFM 120
DLFSIIPGDV DPLLQKGPLI CRRCAVVGNS GNLKESHYGK DIDSHDFVLR MNRAPTAGFE 180
VDVGRKTTHH LVYPESFREL AGNVSMILVP FKTMDLQWLI SALTKGTINF TYVPVPRKIH 240
VNREKILIYH PAFIKYVFDS WLQAHGRYPS TGILSVILSL HICDKVDLYG FGADSKGNWH 300
HYWENNPSAG AFRKTGVHDG DFESNVTSTL ASLNKIRIFK GR                    342
```

SEQ ID NO: 13

Oryctolagus cuniculus ST3Gal1

```
MVTPRKRTLK ALAFLMLFIF LTSFLLNYSH TMVATTWFPK QMVLEFSENL RKLIKTRPCT  60
CAHCVGQRKL SAWFDERFNQ TMQPLLTAHN ALMEEDTYRW WLKLQREKKP NNLNDTIKEL 120
FSVVPGDVDP VLEKRSVGCR RCAVVGNSGN LRESSYGPDI DSHDFVLRMN KAPTVGFEGD 180
VGSKTTHHLV YPESFRELGE NVSMVLVPFK TIDLQWVVSA TTTGTISHTY VPVPAKIKVK 240
QDKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSMHI CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDADF ESNVTATLAA INKIRIFKGR                       340
```

SEQ ID NO: 14

Cricetulus griseus ST3Gal1

```
MMTTQKKVLK VLTFLVLLIF LTSFVLNFAH TTVPAAWFPK QMVLELSQNL RKLIKPPPCT  60
CTHCISQRKV SAWFDKRFNQ TVQPLLTAHN AVLEEDTYQW WLRLQREKKP SNLSDTIREL 120
FSVVPGNVDP VLEKKSGSCR RCAVVGNSGN LRESSYGPEI DSHDFVLRMN RAPTVGFEAD 180
VGSKTTHHLV YPESFRELGE DVSMILVPFK TIDLQWVVSA TTTGTISHTY VPVPKKIKVK 240
QDKILIYHPA FIKYVFDNWL QGHGRYPSTG ILSVIFSLHV CDEVDLYGFG ADSKGNWHHY 300
WENNPSAGAF RKTGVHDGDF ESNVTATLAA INKIRIFTGR                       340
```

SEQ ID NO: 15

Canis familiaris ST3Gal1

```
MVTMRKRTLK VLTLLVLFIF LTSFFLNYSH TMVTTTWFPK QMVVELSENF KKFMKYTHRP  60
CTCARCIGQQ RVSAWFDERF NRSMQPLLTA QNALLEEDTY SWWLRLQREK QPNNLNDTIR 120
ELFQVVPGNV DPLLEKRSVG CRRCAVVGNS GNLRESWYGP QIDSHDFVLR MNKAPTAGFE 180
MDVGSKTTHH LVYPESFREL AENVSMVLVP FKTTDLEWVV SATTTGTISH TYVPVPAKIK 240
VKKDKILIYH PAFIKYVFDS WLQGHGRYPS TGILSVIFSL HICDEVDLYG FGADSKGNWH 300
HYWENNPSAG AFRKTGVHDG DFESNVTATL ASINKIRIFK GR                    342
```

SEQ ID NO: 16

Felis catus ST3Gal1

```
MVTVRKRTLK VLTLLVLFIF LTSFFLNYSH TMVATTWFPK QMVVELSENF KKFMKYAHRP  60
CTCARCIGQQ RVSPWFDERF NRSMQPLLTA QNALLEEDTY SWWLRLQREK QPNNLNDTIK 120
ELFQVVPGNV DPLLEKKSGG CRRCAVVGNS GNLRESWYGP QIDGHDFVLR MNKAPTAGFE 180
ADVGSKTTHH LVYPESFREL GENVSMVLVP FKTTDLEWVV SATTTGTISH TYVPVPAKIK 240
VKKNKILIYH PAFIKYVFDN WLQGHGRYPS TGILSVIFSL HICDEVDLYG FGADSKGNWH 300
HYWENNPSAG AFRKTGVHDG DFESNVTATL ASINKIRIFK GR                    342
```

SEQ ID NO: 17

Equus caballus ST3Gal1

```
MATHRRRILK VLTLLILFIF LTSFFLNYSH TVVTTAWFPK QMVLELSENF KKLVQYSHRP  60
CSCARCIGQQ KVSSWFDERF NRSMQPLLTV QNAFLEEDAY NWWLRLQREK EPSNLNDTIK 120
ELFRVVPGNV DPLLGKRSVG CRRCAVVGNS GNLKESSYGP QIDSHDFVLR MNKAPTAGFE 180
AYVGSKTTHH LVYPESFREL GENVSMVLVP FKTTDLEWVV SATTTGTISH TYVPVPAKIK 240
VKQDKILIYH PAFIKYVFDN WLQGHGRYPS TGILSVIFSL HICDEVDLYG FGADSRGNWH 300
HYWENNPSAG AFRKTGVHDG DFESNVTATL ASIDKIRIFK GR                    342
```

SEQ ID NO: 18

Gallus ST3Gal1

```
MVTVRKRNVK VFTFAFVLIT VTSFLLNYKH QVTMTTWDPK HIISQFSEQV RKLIKFPRRP  60
CSCSTCISEL GHSLWFDQRF NSTMQPFLTS QNALIPEDSY RWWLKLQGEK SPKNINDTLK 120
ELFGIIPGDR DPLQERGTFS CRRCAVVGNS GNLRQSQYGQ DIDSHDFVLR MNRAPTIGYE 180
SDVGSKTTHH FVYPESYKEL AENVSMIVIP FKTLDLRWIV TALTTGTINF TYVPVPRKIK 240
VRKEKVLIYN PSFIKYVYEN WLQNHGRYPS TGLLSVIFAL HVCDEVNVYG FGADSKGHWH 300
HYWENNASAG AFRQTGVHDG DFEFNVTLTL ASIEKIKFFK GR                    342
```

SEQ ID NO: 19

Columba livia ST3Gal1

```
MVVVRRRNVK VFTFAFLLIT VTSFLLNYTH QVTTTTWDPR HLVMQFSEQV QKLFKYPRRP  60
CSCRSCISEL GHSLWFDQRF NSTMQPFLTS QNALIPEDSY RWWLKLQGEK TPKNINATLK 120
ELFEFIPGDG DPLQERGTST CRRCAVVGNS GNLLQSQYGQ DIDSHDFVLR MNRAPTTGYE 180
SDVGSKTTHH FVYPESYKEL AENVSMILIP FKTLDLRWIV TALTTGTINF TYVPVPRKIK 240
VKKEKILIYN PTFMKYVYEN WLQHHGRYPS TGLLSLIFAL HVCDEVNVYG FGADSRGHWH 300
HYWENNGSAG AFRKTGVHDG DFEFNVTLTL ASIEKINFFK GR                    342
```

SEQ ID NO: 20

Alligator sinensis ST3Gal1

```
MRRRHLKMFS FLFVFIAAMS FFLNYNHYEA MVTWAPQQIV MQFSEQFKKL MKHPRRPCSC  60
KACVSELGLS LWFDERFNQT MQPLLTTQNA LISQDSYRWW LKLQGEKNPK NINETIKELF 120
ETISGDGSQL QERSSSMCRR CAVVGNSGNL RQSHYGQDID SHDFVLRMNR APTVGFESDV 180
GSKTTHHFVY PESFKELPEN VSMIVIPFKT LDLRWIVSAL TTGTINHTYV PVPRKIKVKK 240
EKILVYHPDF LKYVFDHWLQ RHGRYPSTGI LSVVFALHVC DEVNLYGFGA NSKGHWHHYW 300
ENNPSAGAFR QTGVHDGDFE SNITSTLAAV NKIHLFKGR                       339
```

SEQ ID NO: 21

Latimeria chalumnae ST3Gal1

```
MARHNHRIMW LLTIILLLCV YMVIYDMGED KQKLIKIPSI RRLSGRTIVL DKKLCSCEKC  60
VSEKEESAWF DERFDPNFQP ILMTEVQDIP SHALQWWLSL QAGNKNYNLS ESIAKLFTVV 120
PRTNHSGIRD PAHCRKCAVV GNSGNLKGSN HGKEIDAHHF VIRMNRARTA GFEPDVGIKT 180
THHLMYPESS QDLQPGVHLV LLPFKIMDFE WIRSALTTGE ITRTYFRVQQ FIKADKDKVL 240
IINPTFFKYV CDHWTEHHGR YPSTGMTALV FALHICDEVS VFGYGADSNG NWHHYWENNR 300
NGGAFRRTGV HSGDFESQII KKLADEGKII FYK                             333
```

SEQ ID NO: 22

Ciona intestinalis ST3Gal1

```
MLINFKLSRV IAMLLVVAIF LTYSWLLLWS TKTALQTNRK NKAGQDEVPV INVIKEDSYV  60
QQKTQNLNKG KRFDLGRVNH SHPREEIQQN NKCGHQLDAS QTRWFRARFN PEIEPVWTQS 120
ALEIDYLVYD WWLSLQSSEA ENLDKTFEAL YKEGVPRKDP FARLTHDREA GCRSCAVVGN 180
SGNILNSNYG NVIDGHDFVI RMNKGPTYNY ENDVGSKTTH RFMYPTTAAS SLPQGVSLVL 240
VPFQPLDIKW LLSALTTGEI TRTYQPLVRR VTCDKSKITI ISPTFIRYVH DRWTQHHGRY 300
PSTGLLALIY ALHECDEVDV YGFGANRAGN WHHYWEDLPP HVAGAFRKTG VHDSAQENEI 360
IDQLHIHGLL RVHRSEQSS                                              379
```

Claims

1. A cell line that is genetically modified to overexpress a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1).

2. The cell line according to claim 1, which is further genetically modified to overexpress a β-galactoside α-2,3-sialyltransferase 4 (ST3Gal4) and/or to overexpress a β-galactoside α-2,6-sialyltransferase 1 (ST6Gal1).

3. The cell line according to claim 1 or claim 2, wherein the cell line comprises an endogenous gene encoding ST3Gal1 and optionally endogenous genes encoding ST3Gal4 and/or ST6Gal1, and further has at least one genetic element, selected from the group consisting of a promoter, an enhancing element, and a stabilizing element inserted into the genome in one or more position(s) suitable to cause overexpression of ST3Gal1 and optionally ST3Gal4 and/or ST6Gal1.

4. The cell line according to any one of claims 1 to 3, wherein the cell line comprises an exogenous nucleic acid encoding a β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1), and optionally exogenous nucleic acids encoding a β-galactoside α-2,3-sialyltransferase 4 (ST3Gal4) and/or a β-galactoside α-2,6-sialyltransferase 1 (ST6Gal1).

5. The cell line according to any one of claims 1 to 4, wherein the cell line is a mammalian cell line, preferably a human cell line.

6. The cell line according to any one of claims 1 to 4, wherein the cell line is derived from a cell line, selected from the group consisting of AGE.CR® cells, Vero cells, MDCK cells, BHK cells, CHO cells, HEK293 cells, HepG2 cells, Huh7 cells, AGE1.HN® cells, NC5T11 cells, Per.C6 cells, HMCLs cells, MM.1

cells, U266 cells, RPMI18226 cells, HKB11 cells, NM cells, NM-F9 cells, and CAP cells.

7. The cell line according to any one of claims 1 to 4, wherein the cell line is derived from human primary amniocytes comprising at least one nucleic acid encoding the gene products of the adenoviral E1 and pIX regions.

8. A recombinant glycoprotein having GalNAc O-glycans that are sialylated to a degree of at least 80%.

9. The recombinant glycoprotein according to claim 8 having GalNAc O-glycans that are sialylated to a degree of at least 95%.

10. The recombinant glycoprotein according to claim 8 or claim 9, wherein at least 80% of the GalNAc O-glycans are core 1 GalNAc O-glycans.

11. The recombinant glycoprotein according to claim 10, wherein the core 1 GalNAc O-glycans are disialylated to a degree of at least 25%.

12. The recombinant glycoprotein according to any one of claims 8 to 11, wherein the glycoprotein is selected from the group consisting of growth factors, peptide hormones, cytokines, enzymes, antibodies, antibody fragments, blood clotting factors, and protease inhibitors.

13. The recombinant glycoprotein according to any one of claims 8 to 12, wherein the glycoprotein is selected from the group consisting of hepatocyte growth factor (HGF), erythropoietin (EPO), Factor VIII (FVIII), Factor IX (FIX), von-Willebrand-Factor (vWF), and C1 esterase inhibitor (C1-inhibitor; C1 Inh).

14. The recombinant glycoprotein according to any one of claims 8 to 13, wherein the glycoprotein is a human glycoprotein.

15. A method for the expression of recombinant glycoproteins having GalNAc O-glycans that are sialylated to a degree of at least 80%, comprising the step of overexpressing with the recombinant glycoproteins β-galactoside α-2,3-sialyltransferase 1 (ST3Gal1).