(19) DANMARK

(10) DK/EP 2756077 T3

(12) Oversættelse af europæisk patentskrift

Patent- og Varemærkestyrelsen

(51) Int.Cl.: *C 12 N 9/24 (2006.01)* *A 61 K 38/47 (2006.01)* *C 07 K 14/315 (2006.01)*
G 01 N 33/573 (2006.01) *G 01 N 33/68 (2006.01)*

(45) Oversættelsen bekendtgjort den: **2017-07-03**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2017-03-29**

(86) Europæisk ansøgning nr.: **12756523.2**

(86) Europæisk indleveringsdag: **2012-09-12**

(87) Den europæiske ansøgnings publiceringsdag: **2014-07-23**

(86) International ansøgning nr.: **EP2012067841**

(87) Internationalt publikationsnr.: **WO2013037824**

(30) Prioritet: **2011-09-13 GB 201115841**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **Genovis Ab, P.O. Box 790, 220 07 Lund, Sverige**

(72) Opfinder: **COLLIN, Mattias, Department of Clinical Sciences, Division of Infection Medicine, Biomedical Center B14, Sölvegatan 19, Lund University, 223 62 Lund, Sverige**
ALLHORN, Maria, Department of Clinical Sciences, Division of Infection Medicine, Biomedical Center B14, Sölvegatan 19, Lund University, 223 62 Lund, Sverige
SJÖGREN, Jonathan, Department of Clinical Sciences, Division of Infection Medicine, Biomedical Center B14, Sölvegatan 19, Lund University, 223 62 Lund, Sverige

(74) Fuldmægtig i Danmark: **Zacco Denmark A/S, Arne Jacobsens Allé 15, 2300 København S, Danmark**

(54) Benævnelse: **ENDOGLYCOSIDASE FRA STREPTOCOCCUS PYOGENES OG FREMGANGSMÅDER TIL ANVENDELSE HERAF**

(56) Fremdragne publikationer:
WO-A2-2008/071418
WO-A2-2009/033670
DATABASE EMBL [Online] 18 October 2008 (2008-10-18), "Streptococcus pyogenes NZ131, complete genome.", XP002686534, retrieved from EBI accession no. EM_STD:CP000829 Database accession no. CP000829 -& DATABASE UniProt [Online] 25 November 2008 (2008-11-25), "SubName: Full=Secreted Endo-beta-N-acetylglucosaminidase (EndoS);", XP002686535, retrieved from EBI accession no. UNIPROT:B5XI26 Database accession no. B5XI26 -& W. M. MCSHAN ET AL: "Genome Sequence of a Nephritogenic and Highly Transformable M49 Strain of Streptococcus pyogenes", JOURNAL OF BACTERIOLOGY, vol. 190, no. 23, 1 December 2008 (2008-12-01), pages 7773-7785, XP55043110, ISSN: 0021-9193, DOI: 10.1128/JB.00672-08
ALLHORN M ET AL: "EndoS from Streptococcus pyogenes is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity", BMC MICROBIOLOGY, BIOMED CENTRAL, LONDON, GB, vol. 8, no. 3, 1 January 2008 (2008-01-01) , pages 1-10, XP002486738, ISSN:

Fortsættes ...

1471-2180 cited in the application
COLLIN M ET AL: "EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG", EMBO JOURNAL, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 20, no. 12, 15 June 2001 (2001-06-15) , pages 3046-3055, XP002486737, ISSN: 0261-4189, DOI: 10.1093/EMBOJ/20.12.3046 cited in the application

DESCRIPTION

Field of the Invention

[0001] The present invention relates to methods of hydrolyzing the glycan of glycoproteins.

Background of the Invention

[0002] Endoglycosidase S (EndoS) is secreted by a number of serotypes of *Streptococcus pyogenes* and has a specific endoglycosidase activity on native IgG by hydrolyzing the conserved glycans attached to the asparagine 297 residue on the heavy chains of IgG, Collin and Olsen, The EMBO Journal, 2001, 20 3046-3055. WO2009/033670 discloses that EndoS can act upon IgG which is already bound to FcγRs, and describes methods of using EndoS to dissociate FcγR-IgG complexes. WO2009/033670 also discloses modified forms of EndoS which lack endoglycosidase activity, and use of such modified forms as IgG binding reagents. WO2008/071418 discloses the use of EndoS to treat or prevent diseases mediated by IgG antibodies. EndoS is the first known bacterial enzyme with a unique specificity for native IgG. In contrast, the activities of other known endoglycosidases require or are enhanced by denaturation of the glycoprotein substrate.

[0003] Antibodies such as IgG have many applications in basic research as well as in diagnostics and drug development. In some of these applications, such as immunohistochemistry, immunoassays, tumour detection, radiotherapy, crystallographic studies of antibody binding sites and immunotargeting, it is more convenient to use Fab fragments than whole IgG molecules. Some of the advantages of using Fab fragments are that they will not be affected by Fc receptors on cells or precipitate antigen, they display a reduced immunogenicity and are less susceptible to phagocytosis, and that radiolabelled Fab fragments are more rapidly cleared from tissue than whole IgG molecules. For other applications, it is desirable to use Fc fragments of IgG. In further applications, it may be desirable to use deglycosylated versions of the antibodies or other glycoproteins.

[0004] The cleavage of IgG into Fab and Fc fragments is most often carried out using proteolytic enzymes such as pepsin or papain. These enzymes often cleave other proteins, so the cleavage reaction generally has to be performed on a purified IgG fraction. Furthermore, pepsin and papain typically cleave IgG in more than one place. This means that the fragments obtained often do not correspond to whole Fab or Fc fragments, and even if cleavage does result in Fab and Fc fragments, they are typically susceptible to further cleavage into smaller fragments. The isolation of Fc fragments from Fab fragments is most often carried out using protein A or G affinity separation columns, which utilise the Fc-binding properties of the bacterial proteins A and G.

[0005] Many different glycoproteins have utility in therapeutic applications. Methods to analyse the glycosylation of such proteins have utility in the research and development of the proteins as therapeutics. It may also be desirable to provide deglycosylated versions of these proteins.

Summary of the Invention

[0006] The inventors have identified a novel endoglycosidase from serotype M49 *Streptococcus pyogenes*, referred to herein as EndoS49. EndoS49 was isolated from strain NZ131, a nephritogenic and highly transformable strain of serotype M49. NZ131 strain is a clinical isolate from a case of acute post-streptococcal glomerulonephritis in New Zealand. At a protein level, EndoS49 has less than 40% identity to EndoS, and is a 90kDa protein, compared to the 108kDa of EndoS. EndoS49 has deglycosylation activity for a broader range of proteins than EndoS.

[0007] The enzyme is a 90kDa enzyme, having a family 18 glycoside hydrolase catalytic domain. EndoS49 hydrolyzes glycan on human glycoproteins, and in particular IgG1-4, and alpha-1-microglobulin. EndoS49 can be used in the hydrolysis of glycans on human glycoproteins including IgG and alpha-1-microglobulin. EndoS49 can thus be used in glycoprofiling analysis in which the enzyme is contacted with a glycoprotein, and the products produced are separated for analysis of the glycans and the protein. EndoS49 can also be used to prepare deglycosylated proteins. The enzyme can be modified to reduce or remove endoglycosidase activity.

[0008] The present invention provides a method for complete deglycosylation of an IgG antibody comprising incubating the antibody with a polypeptide comprising

1. (a) the amino acid sequence of SEQ ID NO: 1;
2. (b) a variant thereof having at least 95% identity to the amino acid sequence of SEQ ID NO: 1 over at least 810 contiguous amino acids of SEQ ID NO: 1 and having the endoglycosidase activity of a polypeptide consisting of the amino acid sequence of SEQ ID NO: 1.

Brief Description of the Figures

[0009]

Figure 1. ClustalW alignment of EndoS49 and EndoS reveals two different proteins. EndoS49 and EndoS was aligned using ClustalW in the software MacVector. GH18 catalytic motif (D**D*D*E) is present at position 179-186 with Glu186 as the catalytic residue.

Figure 2. EndoS49 has activity on glycoproteins. A. 1 µg of EndoS49, its catalytic mutant and

the truncated versions was incubated with 3 µg human IgG in PBS overnight at 37°C and analyzed on a SDS-PAGE gel and with LCAlectin blot. B. 1 µg EndoS49 and EndoS49(E186L) was incubated with 3 µg of subclasses 1-4 of human IgG and analyzed as above. C. 1 µg EndoS49 and its mutants were incubated with 3 µg Alpha-1-microglobulin and analyzed on a 10 % SDS-PAGE gel.

Figure 3. EndoS49 binds to IgG. 4, 2 and 1 µg of EndoS49 and its mutants were immobilized on a PVDF membrane and incubated with human IgG and later with protein-G coupled to HRP.

Figure 4. The genomic context of *ndoS49* and *ndoS*. A comparison of the genes surrounding *ndoS49* and *ndoS*in GAS strains NZ131 (M49) and 5005 (M1) was carried out in MacVector.

Figure 5. Phylogenetic analysis of EndoS49 and other bacterial endoglycosidases.

Figure 6. SDS Page gel of Avastin and Erbitux after digestion with EndoS or EndoS49 followed by IdeS digestion.

Brief Description of the Sequences

[0010]

SEQ ID NO: 1 is an amino acid sequence of an EndoS49 polypeptide isolated from *S. pyogenes* M49 serotype NZ131.

SEQ ID NO: 2 is an amino acid sequence of a modified EndoS49 polypeptide (E186L) derived from the sequence of SEQ ID NO: 1.

SEQ ID NO: 3 is a nucleotide sequence encoding EndoS49 polypeptide

SEQ ID NO: 4 is an amino acid sequence of IdeS isolated from *S. pyogenes* AP1.

Detailed Description of the Invention

General polypeptide features

[0011] The terms protein, peptide and polypeptide are used interchangeably herein. The following section relates to general features of all polypeptides disclosed herein, and in particular to variations, alterations, modifications or derivatisations of amino acid sequence which are included within the said polypeptides. It will be understood that such variations,

alterations, modifications or derivatisations of polypeptides as are described herein are subject to the requirement that the polypeptides retain any further required activity or characteristic as may be specified in subsequent sections of this disclosure.

[0012] Variants of polypeptides may be defined by particular levels of amino acid identity which are described in more detail in subsequent sections of this disclosure. Amino acid identity may be calculated using any suitable algorithm. For example the PILEUP and BLAST algorithms can be used to calculate homology or line up sequences (such as identifying equivalent or corresponding sequences (typically on their default settings), for example as described in Altschul S. F. (1993) J Mol Evol 36:290-300; Altschul, S, F et al (1990) J Mol Biol 215:403-10. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighbourhood word score threshold (Altschul *et al,* supra). These initial neighbourhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extensions for the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1992) Proc. Natl. Acad. Sci. USA 89: 10915-10919) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

[0013] The BLAST algorithm performs a statistical analysis of the similarity between two sequences; see e.g., Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873-5787. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two polynucleotide or amino acid sequences would occur by chance. For example, a sequence is considered similar to another sequence if the smallest sum probability in comparison of the first sequence to the second sequence is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001. Alternatively, the UWGCG Package provides the BESTFIT program which can be used to calculate homology (for example used on its default settings) (Devereux et al (1984) Nucleic Acids Research 12, 387-395).

[0014] It will be understood that variants of polypeptides also includes substitution variants. Substitution variants preferably involve the replacement of one or more amino acids with the same number of amino acids and making conservative amino acid substitutions. For example, an amino acid may be substituted with an alternative amino acid having similar properties, for example, another basic amino acid, another acidic amino acid, another neutral amino acid,

another charged amino acid, another hydrophilic amino acid, another hydrophobic amino acid, another polar amino acid, another aromatic amino acid or another aliphatic amino acid. Some properties of the 20 main amino acids which can be used to select suitable substituents are as follows:

Ala	aliphatic, hydrophobic, neutral	Met	hydrophobic, neutral
Cys	polar, hydrophobic, neutral	Asn	polar, hydrophilic, neutral
Asp	polar, hydrophilic, charged (-)	Pro	hydrophobic, neutral
Glu	polar, hydrophilic, charged (-)	Gln	polar, hydrophilic, neutral
Phe	aromatic, hydrophobic, neutral	Arg	polar, hydrophilic, charged (+)
Gly	aliphatic, neutral	Ser	polar, hydrophilic, neutral
His	aromatic, polar, hydrophilic, charged (+)	Thr	polar, hydrophilic, neutral
Ile	aliphatic, hydrophobic, neutral	Val	aliphatic, hydrophobic, neutral
Lys	polar, hydrophilic, charged(+)	Trp	aromatic, hydrophobic, neutral
Leu	aliphatic, hydrophobic, neutral	Tyr	aromatic, polar, hydrophobic

[0015] The polypeptides disclosed herein may be in a substantially isolated form. It will be understood that the polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. A polypeptide may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the disclosure.

[0016] The amino acid sequence of polypeptides o may be modified to include non-naturally occurring amino acids or to increase the stability of the compound. When the polypeptides are produced by synthetic means, such amino acids may be introduced during production. The polypeptides may also be modified following either synthetic or recombinant production.

[0017] Polypeptides may also be produced using D-amino acids. In such cases the amino acids will be linked in reverse sequence in the C to N orientation. This is conventional in the art for producing such polypeptides.

[0018] A number of side chain modifications are known in the art and may be made to the side chains of the polypeptides, subject to the polypeptides retaining any further required activity or characteristic as may be specified herein.

[0019] It will also be understood that polypeptides may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated, phosphorylated or comprise

modified amino acid residues. They may be modified by the addition of a signal sequence to promote insertion into the cell membrane.

[0020] Polypeptides may also be derivatised or modified to assist with their isolation or purification. For example, a polypeptide may be derivatised or modified by addition of a ligand which is capable of binding directly and specifically to a separation means. Alternatively, the polypeptide is derivatised or modified by addition of one member of a binding pair and the separation means comprises a reagent that is derivatised or modified by addition of the other member of a binding pair. Any suitable binding pair can be used. In a preferred embodiment where the polypeptide is derivatised or modified by addition of one member of a binding pair, the polypeptide is preferably histidine-tagged or biotin-tagged. Typically the amino acid coding sequence of the histidine or biotin tag is included at the gene level and the proteins are expressed recombinantly in *E. coli*. The histidine or biotin tag is typically present at one end of the polypeptide, either at the N-terminus or at the C-terminus. The histidine tag typically consists of six histidine residues, although it can be longer than this, typically up to 7, 8, 9, 10 or 20 amino acids or shorter, for example 5, 4, 3, 2 or 1 amino acids. Furthermore, the histidine tag may contain one or more amino acid substitutions, preferably conservative substitutions as defined above.

EndoS49 polypeptides having endoglycosidase activity

[0021] The EndoS49 polypeptide used in the method of the invention comprises:

1. (a) the amino acid sequence of SEQ ID NO: 1; or
2. (b) a variant thereof having at least 95% identity to the amino acid

sequence of SEQ ID NO: 1 over at least 810 contiguous amino acids of SEQ ID NO: 1 and having the endoglycosidase activity of a polypeptide consisting of the amino acid sequence of SEQ ID NO: 1.

[0022] Preferably, the polypeptide comprises, or consists of, the sequence of SEQ ID NO: 1. SEQ ID NO: 1 is the sequence of EndoS49 from *S. pyogenes*. The EndoS49 polypeptide of the invention may additionally not comprise a signal sequence.

[0023] Variants of the amino acid sequence of SEQ ID NO: 1 preferably contain residues 179 to 186 of SEQ ID NO: 1, and in particular include the motif D**D*D*E. These amino acids constitute a family 18 glycoside hydrolase catalytic domain. The glutamic acid at position 186 is essential for enzymatic activity. Most preferably, therefore, the variant of SEQ ID NO: 1 contains glutamic acid at the position equivalent to position 186 of SEQ ID NO: 1. The variant of SEQ ID NO: 1 may contain residues 179 to 186 of SEQ ID NO: 1 having one or more conservative substitutions, provided that the variant contains glutamic acid at the position equivalent to position 186 of SEQ ID NO: 1.

[0024] Typically, polypeptides which display the endoglycosidase activity of EndoS49 have at least 95%, at least 97% or at least 99% identity, with the amino acid sequence of SEQ ID NO: 1 The identity of variants of SEQ ID NO: 1 may be measured over a region of at least 810, at least 820, at least 930, at least 940 or more contiguous amino acids of the sequence shown in SEQ ID NO: 1, or more preferably over the full length of SEQ ID NO: 1.

[0025] Polypeptides for use in the present invention may be isolated from any suitable organism that expresses an EndoS49 polypeptide or a variant of an EndoS49 polypeptide. Typically, the EndoS49 polypeptide is isolated from suitable EndoS49 expressing strains of *Streptococcus,* preferably strains of *S. pyogenes,* and in particular those of serotype M49.

[0026] Isolation and purification of EndoS49 from an expressing *S. pyogenes* culture, or from cultures of other cells expressing EndoS49 is typically on the basis of endoglycosidase activity. Preferably the purification method involves an ammonium sulphate precipitation step and an ion exchange chromatography step. According to one method, the culture medium is fractionated by adding increasing amounts of ammonium sulphate. The amounts of ammonium sulphate may be 10 to 80%. Preferably the culture medium is fractionated with 50% ammonium sulphate, and the resulting supernatant is further precipitated with 70% ammonium sulphate. Pelleted polypeptides may then be subjected to ion exchange chromatography, for example by FPLC on a Mono Q column. Eluted fractions may be assayed for endoglycosidase activity and peak activity fractions may be pooled. Fractions may be analysed by SDS PAGE. Fractions may be stored at -80°C.

[0027] Polypeptides for use in the invention may also be prepared as fragments of such isolated polypeptides. Further, the EndoS49 polypeptides may also be made synthetically or by recombinant means. For example, a recombinant EndoS49 polypeptide may be produced by transfecting mammalian cells in culture with an expression vector comprising a nucleotide sequence encoding the polypeptide operably linked to suitable control sequences, culturing the cells, extracting and purifying the EndoS49 polypeptide produced by the cells.

[0028] The EndoS49 polypeptides described in this section display endoglycosidase activity. Preferably, the polypeptide hydrolyses IgG or IgG Fc fragments by hydrolysing glycan linked of a full-length IgG heavy chain polypeptide. Preferably the EndoS49 polypeptide of the invention also has endoglycosidase activity, and is capable of glycan hydrolysis of alpha-1-microglobulin.

[0029] The endoglycosidase activity may be determined by means of a suitable assay. For example, a test polypeptide may be incubated with glycoprotein such as IgG or alpha-1-microglobulin at a suitable temperature, such as 37°C. The starting materials and the reaction products may then be analysed by SDS PAGE. Typically, the molecular mass of the IgG heavy chain is reduced by approximately 3kDa to 4kDa if the test polypeptide has IgG endoglycosidase activity. Another assay for determining whether a test polypeptide has IgG endoglycosidase activity is by detection of glycosylated IgG using *Lens culinaris* agglutinin lectin (LCA), optionally using horseradish peroxidase and peroxidase substrate. Typically, the carbohydrate signal is reduced if the test polypeptide has IgG endoglycosidase activity.

Another assay for determining whether a test polypeptide has IgG endoglycosidase activity is by incubation of a test polypeptide with purified IgG Fc fragments followed by reduction of the sample with 10 mM dithiotreitol and mass spectroscopy (MALDI-TOF) analysis. Endoglycosidase activity can also be measured for EndoS49 polypeptides by using alpha-1-microglobulin in place of IgG in the assays mentioned above.

[0030] The endoglycosidase activity of the polypeptides can be further characterised by inhibition studies.

[0031] The EndoS49 polypeptide is capable of hydrolyzing glycoprotein molecules present in a sample taken from a subject. Thus, where the subject is a human, the EndoS49 polypeptide is capable of hydrolyzing the glycans in glycoproteins of a subject, such as on the heavy chains of human IgG or alpha-1-microglobulin. EndoS49 is capable of hydrolyzing human IgG of all four subclasses (IgG_{1-4}). In preferred embodiments, the EndoS49 polypeptide has the ability to hydrolyze human IgG and alpha-1-microglobulin.

Methods using the endoglycosidase activity of EndoS49

[0032] As described herein, EndoS49 has endoglycosidase activity and is able to hydroylse the glycan of glycoproteins including IgG and alpha-1-microglobulin. The present invention thus provides methods for deglycosylation of glycoproteins, and in particular, hydrolysis of glycan from glycoproteins, and in particular, from IgG and alpha-1-microglobulin. Typically, such a method includes incubating a sample containing glycoprotein with EndoS49 with a glycoprotein under conditions which allow the endoglycosidase activity. Suitable conditions include use of EndoS49 at a concentration of at least 1 µg/ml, 2 µg/ml, 4 µg/ml, 6 µg/ml, 8 µg/ml, 10 µg/ml, 12 µg/ml, 15 µg/ml or 20 µg/ml, preferably at least 10 µg/ml. Suitable conditions also include incubation of the sample with EndoS49 for at least 20 minutes, 30 minutes, 40 minutes, 50 minutes, 60 minutes, 70 minutes, 80 minutes, 90 minutes or 120 minutes, preferably at least 60 minutes. Incubation preferably takes place at room temperature, more preferably at approximately 20°C, 25°C, 30°C, 35°C, 40°C or 45°C, and most preferably at approximately 37°C.

[0033] These methods may be used to provide deglycosylated glycoproteins, which may themselves be useful in research or therapy. These methods may also be used to characterise glycans on glycoproteins, for example, in glycomapping or glycoprofiling. Such glycomapping and glycoprofiling is particularly useful for antibody molecules, such as IgG molecules, for example, in the analysis of monoclonal IgG molecules. Typically, the methods involved incubating the protein with EndoS49 to hydroylase the glycans of the protein. Subsequently, the glycans and the protein or polypeptide are separated, for example, using any suitable technique such as HPLC or gel chromatography. The separated moieties can then be analysed using any suitable analytical method, such as mass spectrometry, HPLC, gel chromatography, gel electrophoresis, spectrometry, capillary electrophoresis and other standard laboratory techniques for the analysis of glycans and/or proteins.

[0034] In accordance with additional methods of the present invention, the methods may also comprise utilising additional enzymes such as IdeS so that the glycans on the Fc portion of the antibody can be analysed in more details using the methods and techniques described herein.

[0035] One example is to analyze the fucosylation of an immunoglobulin. The degree of fucosylation on the Fc glycans on an IgG molecule is important for the therapeutic potential of an IgG drug candidate. Afucosylated IgG molecules increase the ADCC (nn) effect of the therapeutic IgG molecule. Thus, in accordance with the present invention, there is provided a method for analyzing the amount of fucose in the Fc glycans of an IgG, using EndoS49.

[0036] Typically, such a method includes incubating an glycoprotein, in this case an immunoglobulin, with EndoS49 under conditions which allow the endoglycosidase activity of EndoS49. Suitable conditions include use of EndoS49 at a concentration of at least 1 µg/ml, 2 µg/ml, 4 µg/ml, 6 µg/ml, 8 µg/ml, 10 µg/ml, 12 µg/ml, 15 µg/ml or 20 µg/ml, preferably at least 10 µg/ml. Suitable conditions also include incubation of the sample with EndoS49 for at least 20 minutes, 30 minutes, 40 minutes, 50 minutes, 60 minutes, 70 minutes, 80 minutes, 90 minutes or 120 minutes, preferably at least 60 minutes. Incubation preferably takes place at room temperature, more preferably at approximately 20°C, 25°C, 30°C, 35°C, 40°C or 45°C, and most preferably at approximately 37°C. IdeS may be added after the reaction with EndoS49, or in the same reaction mixture, to induce proteolysis, dividing the immunoglobulin molecule into F(ab')2 and Fc. The two fragments are separated using a separation method before injection into a mass spectrometer. After glycan cleavage, a GlcNAc and core Fuc residue remain attached to Asn at the consensus Fc/2 glycosylation site. Since an afocusylated immunoglobulin is not core fucosylated, some Fc/2 will contain only a GlcNAc after digestion. The characteristic mass difference (-146 Da) resulting from the absence of fucose is readily apparent in the deconvoluted mass spectrum. Use of EndoS49 therefore, facilitates the direct estimation of the degree of core afucosylation of IgG.
The following Examples illustrate the invention:

Example 1

MATERIALS AND METHODS

Bacterial strains and growth

[0037] The genome of GAS strain NZ131 of serotype M49 has been sequenced and this strain was therefore selected as reference strain in this work (McShan *et al.,* 2008) (Chaussee *et al.,* 1999). GAS strain NZ131 was propagated on blood agar and*Escherichia coli* strains Top10 (Invitrogen) and BL21 pLysS (Invitrogen) were propagated on lysogeny broth (LB) agar.

For selection in *E. coli* Top10 cells, carbenicillin was used at 100 µg/mL and for *E. coli* BL21 pLysS, 100 µg/mL carbenicillin and 34 µg/mL chloramphenicol were used. Overnight cultures of *E. coli* were carried out in LB at 37°C with aeration. Genomic DNA preparation of GAS strain NZ131 was performed using Puregene DNA Purification Kit (Qiagen). Transformation was carried out using heat-shock at 42°C for 30 s. Plasmid preparations from *E. coli* were performed using Plasmid Miniprep Kit I (E.Z.N.A).All primers used in this work are listed in Table 2.

Recombinant expression of EndoS49

[0038] Recombinant expression of EndoS49 in *E. coli* was established by PCR amplification of the *ndoS49* gene from group A *Streptococcus* strain NZ131, serotype M49 with the primers ndoS49-F-BamHI, CTGTAAGGATCCAGGAGAAGACTG, and ndoS49-R-XhoI, GAAACCTCGAGTCTTTGTAATCGTAGGACTT. The *ndoS49* gene fragment was digested with restriction enzymes *BamH*I and *Xho*I and ligated into the expression vector pGEX-5X-3 (Amersham Biosciences) using DNA ligase T4 (Fermentas) creating the plasmid pGEX-ndoS49. The expression vector was transformed into *E. coli* Top10 chemically competent cells and recombinant cells were grown on 100 µg/mL carbenicillin plates and screened with PCR using primers ndoS49-F-BamHI and ndoS49-R-XhoI. Positive clones were isolated and the pGEX-*ndoS49* plasmid was purified and transformed into the *E. coli* expression strain BL21 pLysS as described above.

[0039] One recombinant clone was grown overnight at 37°C with antibiotics and diluted 1:20 in LB medium with antibiotics and grown for 3 h. The expression of the protein EndoS49 was induced with 0.1 mM IPTG for 3 h. The cells were harvested and lysed with BugBuster Protein Extraction Reagent (Novagen). Recombinant GST-EndoS49 was purified on column with Glutathione Sepharose 4B (GE Healthcare) and eluated with reduced glutathione.

Mutagenesis of EndoS49

[0040] Site-directed mutagenesis of the glutamic acid 186(Glu-186) to leucine (E186L) was carried out using QuickChange II Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. The mutagenesis primer used was CTAGATATTGATATTCTTCACGAATTTACGAAC in combination with the antisense of the sequence above and the plasmid pGEX-*ndoS49*(mutation underlined). This generated the plasmids pGEX-*ndoS49*(E186L) and, after sequencing, recombinant EndoS49(E186L) was expressed and purified as described for EndoS49. The truncated versions of EndoS49 were constructed by amplifying parts of the *ndoS49* gene from GAS NZ131 with primersndoS49(trunc1-5) containing restrictions sites BamHI and XhoI (Table 2). The fragments were digested and ligated into the pGEX vector as above, and transformed into *E. coli* Top10 and subsequently to BL21 pLysS and grown with antibiotics. The proteins were

produced as above and the proteins EndoS49(trunc1) 80 kDa, EndoS49(trunc2) 70 kDa, EndoS49(trunc3) 60 kDa, EndoS49(trunc4) 50 kDa, EndoS49(trunc5) 42 kDa were purified.

Glycoprotein glycan hydrolysis assay

[0041] 1 µg recombinant EndoS49 and its mutants were incubated with 3 µg of each glycoprotein in 20 µL PBS overnight at 37°C. Glycan hydrolysis was analyzed on a 10% SDS-PAGE gel and subsequently analyzed with LCAlectin blot as previously described (Collin and Olsén, 2001a).

Slot-blot analysis

[0042] EndoS49 and its mutants were immobilized on a methanol activated PVDF membrane at 4, 2, 1 µg in PBS per slot using Millipore slot blot equipment. The membrane was blocked with 5% skim milk (Difco) for 1 h at room temperature. Washing was consistently carried out for 3x10 minutes in PBST. The membrane was incubated with 10 µg human IgG (Sigma) in 0,5% skim milk for 1 h at 37°C and then washed. 5 µg horseradish peroxidase conjugated with protein G (Invitrogen) was added to the membrane and incubated for 1 h at 37°C. After washing the membrane was developed with Supersignal West Pico Chemiluminiscent Substrate (Thermo Scientific).

Bioinformatic analysis

[0043] The genes *ndoS49* and *ndoS* were translated into EndoS49 and EndoS and compared using the ClustalW algorithmin within the software MacVector (MacVector Inc.). The phylogenetic tree was constructed with MacVector using protein sequences from NCBI PubMed with the following accession numbers: EndoS (AF296340), EndoE (AAR20477), EndoH (NP_631673), EndoC (ADC53484), EndoF2 (P36912), EndoF3 (P36913) (Collin and Olsén, 2001b) (Collin and Fischetti, 2004) (Tarentino and Plummer, 1974) (Tarentino *et al.*, 1993).

PCR screening for *ndoS49*

[0044] Primers amplifying the *ndoS49* gene from GAS were designed and denoted ndoS49-F (AAAACGCGGACCACTATATGC) and ndoS49-R (AAACGTTGTCCGAGGATTTG). 42 GAS strains were propagated on blood agar and grown overnight at 37°C with 5% CO_2. Single colonies were picked and lysed in 20 µL sterile H_2O at 99°C for 10 minutes. These lysates were used as template for a stringent PCR reaction to detect *ndoS49* in the 42 GAS strains. As

a positive control, primers for the amplification of the gene *recA* were designed, recA-F (AGCCCTTGATGATGCTTTG) and recA-R (AACAATTCTGGGTGATCGG).As positive controls, both PCR reactions used genomic DNA from GAS strain NZ131 (M49) and AP1 (M1) as template.

RESULTS

ClustalW analysis reveals two different enzymes: EndoS49 and EndoS

[0045] The genes *ndoS49* and *ndoS* were *in silico* translated into proteins and compared using the ClustalWalgorithm. On the gene level the identity is 50% and 37% on the protein level. The ClustalW analysis revealed a (nearly) identical signal peptide sequence and a conserved family 18 glycoside hydrolasecatalytic domain (DGLDIDIE)(Figure 1). Experimental analysis of EndoS has shown that tryptophans are essential for the glycan-hydrolyzing activity (Allhorn *et al.,* 2008). These tryptophans are also conserved in EndoS49.

Recombinant EndoS49 show glycan hydrolyzing activity on human glycoproteins

[0046] The 90 kDa EndoS49 was successfully recombinantly expressed in *E. coli* BL21 and purified from the soluble fraction using the GST-tag. EndoS49(E186L), a catalytic mutant with the glutamic acid of the GH18 motif (E186) substituted for a leucine (L), was constructed and purified in the same way. To map the activity of the protein, 5 carboxy-terminally truncated versions of the enzymes were constructed and denoted EndoS49(trunc1) 80 kDa, EndoS49(trunc2) 70 kDa, EndoS49(trunc3) 60 kDa, EndoS49(trunc4) 50 kDa and EndoS49(trunc5) 42 kDa. This collection of enzymes was utilized to analyze the glycan hydrolyzing activity of EndoS49 on human glycoproteins. First, the enzymes were incubated with human IgG overnight and analyzed on a SDS-PAGE gel and with LCA lectin blot, detecting the mannose structures in the glycan of IgG. The gel revealed a shift of 4 kDa of the IgG heavy chain treated with EndoS49 and the LCA lectin blot confirmed this shift as a lack of the *N*-linked glycan (Figure 2A). EndoS49(E186L) showed no shift and no change in glycan composition suggesting that E186 plays a crucial role in the catalytic activity of EndoS49. Concerning the truncated enzymes, EndoS49(trunc1-4) showed activity on the glycan of IgG but EndoS49(trunc5), the smallest of the enzymes (42 kDa), showed no glycan-hydrolyzing activity (Figure 2A).

[0047] Further analysis of the IgG deglycosylation by EndoS49 was carried out by incubating IgG_{1-4} with EndoS49 and EndoS49(E186L), overnight. The IgG subclasses were analyzed as above and showed that EndoS49 has activity on all four subclasses of IgG, and in line with previous result, the catalytic mutant showed no activity (Figure 2B). Incubating the collection of enzymes with alpha-1-microglobulin, a heavy glycosylated human serum protein, and analysis

on SDS-PAGE showed glycan hydrolysis activity of EndoS49on this glycoprotein(Figure 2C). To further elucidate the specificity of EndoS49, a model substrate consisting of a *N*-acetyl-beta-D-glucosaminide coupled to the fluorescent 4-methylumbelliferylwas incubated with EndoS49 for 1, 2, 3, 4 and 16 h and the fluorescence measured. The fluorescence will increase if the sugar is cleaved but no such increase in intensity was observed (data not shown) suggesting that EndoS49 has activity only on glycoprotein substrates.

EndoS49 binding to IgG

[0048] The finding that EndoS49 has glycan hydrolyzing activity on IgG led us to believe that the enzyme binds IgG. This was evaluated with slot blot analysis where EndoS49 and its catalytic mutant and truncated versions were immobilized on a PVDF membrane and incubated with IgG and the binding detected with protein G coupled to Horseradish-peroxidase (HRP). The slot blot show an increased binding to IgG by the catalytic mutant EndoS49(E186L) (Figure 3).

The gene *ndoS49* is present in GAS serotype M49

[0049] To elucidate whether *ndoS49* is present in any other serotypes than M49 a stringent PCR was deployed to analyze the presence of the *ndoS49* gene in a selection of GAS strains. The primers ndoS49-F and ndoS49-R was used in a PCR on lysates from GAS colonies together with the positive control amplifying the *recA* gene, present in all GAS strains. The *ndoS49* gene was amplified in all selected GAS M49 serotypes and also in serotype M60, whereas no other serotype gave a PCR product (Table 1).

[0050] In the sequenced genomes of GAS strains NZ131 (M49) and MGAS5005 (M1) the genes surrounding *ndoS49* and *ndoS* were compared revealing that the genes are located in the same genomic context and that the surrounding genes are highly conserved (Figure 4). The full length EndoS49 was compared to a selection of previously described endoglycosidases, EndoS, EndoC, EndoH, EndoE, EndoF2, EndoF3) and a phylogenetic tree was reconstructed (Figure 5). This revealed that EndoS and EndoC are more closely related than EndoS and EndoS49 and that endoglycosidases from *Streptococcus* are close related compared to enzymes from other bacteria.

Table 1. The presence of *ndoS49* in a selection of GAS serotypes

Strain	**Serotype**	***ndoS49***	***recA***
5448	M1	-	+
SF370	M1	-	+
ACN1	M1	-	+
ACN2	M2	-	+
ANC3	M3	-	+

Strain	Serotype	*ndoS49*	*recA*
20224	M3	-	+
ACN4	M4	-	+
ACN5	M5	-	+
Manfredo	M5	-	+
ACN6	M6	-	+
AP6	M6	-	+
ACN9	M9	-	+
ACN11	M11	-	+
ACN12	M12	-	+
ACN18	M18	-	+
ACN19	M19	-	+
ACN22	M22	-	+
ACN24	M24	-	+
ACN28	M28	-	+
NZ131	M49	+	+
3487-05	M49	+	+
AW1	M49	+	+
AW2	M49	+	+
AW3	M49	+	+
AW4	M49	+	+
AW6	M49	+	+
AW7	M49	+	+
AW8	M49	+	+
AW9	M49	+	+
AW10	M49	+	+
AW11	M49	+	+
AW12	M49	+	+
AW13	M49	+	+
ACN49	M49	+	+
AP49	M49	+	+
CS101	M49	+	+
AP53	M53	-	+
ALAB49	M53	-	+
ACN55	M55	-	+
ACN57	M57	-	+

Strain	Serotype	*ndoS49*	*recA*
ACN60	M60	+	+
AP74	M74	-	+

Table 2. Primers used in this work

Primer name	Sequence (5'-3')
ndoS49-F-BamHI	CTGTAAGGATCCAGGAGAAGACTG
ndoS49-R-XhoI	GAAACCTCGAGTCTTTGTAATCGTAGGACTT
ndoS49(E186L)-F	CTAGATATTGATATTCTTCACGAATTTACGAAC
ndoS49(E186L)-R	GTTCGTAAATTCGTGAAGAATATCAATATCTAG
ndoS49(trunc1)-R-XhoI	ATTTCTCGAGCTGAAGACGTCCTTTAGCCACG
ndoS49(trunc2)-R-XhoI	TAAACTCGAGCCCCATCAGAAACATCTACTAAG
ndoS49(trunc3)-R-XhoI	ATTTTCTCGAGGCATTATCAACATCATAATGACC
ndoS49(trunc4)-R-XhoI	TAAACTCGAGCCAGTCATGCCTACCATAACAAGCTCAGC
ndoS49(trunc5)-R-XhoI	ATTTCTCGAGCTGTCCAACTTGTTGAATG
ndoS49-F	AAAACGCGGACCACTATATGC
ndoS49-R	AAACGTTGTCCGAGGATTTG
recA-F	AGCCCTTGATGATGCTTTG
recA-R	AACAATTCTGGGTGATCGG

The protein sequence of EndoS49

[0051] NCBIReferenceSequences

NZ131 genome: NC_011375.1

ndoS49 gene sequence: NC_011375.1

EndoS49 protein sequence: YP_002286383.1

EndoS49 Protein Sequence:

MDKHLLVKRTLGCVCAATLMGAALATHHDSLNTVKAEEKTVQTG
KTDQQVGAKLVQEIREGKRGPLYAGYFRTWHDRASTGIDCKQQHPENTMAEVPKEVDI
LFVFHDHTASDSPFWSELKDSYVHKLHQQGTALVQTIGVNELNGRTGLSKDYPDTPEG
NKALAAAIVKAFVTDRGVDGLDIDIEHEFTNKRTPEEDARALNVFKEIAQLIGKNGSD
KSKLLIMDTTLSVENNPIFKGIAEDLDYLLRQYYGSQGGEAEVDTINSDWNQYQNYID
ASQFMIGFSFFEESASKGNLWFDVNEYDPNNPEKGKDIEGTRAKKYAEWQPSTGGLKA
GIFSYAIDRDGVAHVPSTYKNRTSTNLQRHEVDNISHTDYTVSRKLKTLMTEDKRYDV

IDQKDIPDPALREQIIQQVGQYKGDLERYNKTLVLTGDKIQNLKGLEKLSKLQKLELR
QLSNVKEITPELLLPESMKKDAELVMVGMTGLEKLNLSGLNRQTLDGIDVNSITHLTSF
DISHNSLDLSEKSEDRKLLMTLMEQVSNHQKITVKNTAFENQKPKGYYPQTYDTKEGH
YDVDNAEHDILTDFVFGTVTKRNTFIGDEEAFAIYKEGAVDGRQYVSKDYTYEAFRKD
YKGYKVHLTASNLGETVTSKVTATTDETYLVDVSDGEKVVHHMKLNIGSGAIMMENLA
KGAKVIGTSGDFEQAKKIFDGEKSDRFFTWGQTNWIAFDLGEINLAKEWRLFNAETNT
EIKTDSSLNVAKGRLQILKDTTIDLEKMDIKNRKEYLSNDENWTDVAQMDDAKAIFNS
KLSNVLSRYWRFCVDGGASSYYPQYTELQILGQRLSNDVANTLKD

Example 2

[0052] Monoclonal Ig molecules can show minor variations in the Fc glycans and as a consequence the Fc glycans can appear as an inhomogeneous pool of Fc glycans. The vast majority of the glycans are identical but a minority can show variable carbohydrate structure or composition. The variety arises both from the origin of the Fc part, which can be human, humanized or from another species, and from the choice of production cell-line and cell culture conditions. In this example two well-known IgG based drugs, Avastin and Erbitux, were deglycosylated both with EndoS and with EndoS49. The samples were incubated with EndoS or EndoS49 as set out in the Table below.

Lane	Sample	Erbitux 100 µg	Avastin 100 µg	EndoS 0.1 mg/ml (µl)	EndoS49 0.1 mg/ml (µl)	ideS 1 µg
1	MW standard	-	-	-	-	-
2		-	+	5	-	+
3		-	+	50	-	+
4		-	+	-	5	+
5		-	+	-	50	+
6		-	+	-	-	+
7		+	-	5	-	+
8		+	-	50	-	+
9		+	-	-	5	+
10		+	-	-	50	+
11		+	-	-	-	+
12	MW standard	-	-	-	-	-

[0053] The results are presented in Figure 6. It was found that EndoS49 has the potential to cleave a larger variety of Fc glycans than EndoS. Even if incubation was left over night in order to minimize the effect of potential differences in the enzymatic activities the EndoS enzyme could not fully deglycosylate Erbtux Fc glycans. However EndoS49 shows a complete

deglycosylation profile and is hence the more favourable enzyme when it comes to glycan profiling of immunoglobulins.

SEQUENCE LISTING

[0054]

<110> GENOVIS AB

<120> PROTEIN AND METHOD

<130> N.114953A

<150> 1115841.7
<151> 2011-09-13

<160> 19

<170> PatentIn version 3.5

<210> 1
<211> 843
<212> PRT
<213> Streptococcus pyogenes

<400> 1

```
Met Asp Lys His Leu Leu Val Lys Arg Thr Leu Gly Cys Val Cys Ala
1               5                  10                  15

Ala Thr Leu Met Gly Ala Ala Leu Ala Thr His His Asp Ser Leu Asn
            20                  25                  30

Thr Val Lys Ala Glu Glu Lys Thr Val Gln Thr Gly Lys Thr Asp Gln
        35                  40                  45

Gln Val Gly Ala Lys Leu Val Gln Glu Ile Arg Glu Gly Lys Arg Gly
    50                  55                  60

Pro Leu Tyr Ala Gly Tyr Phe Arg Thr Trp His Asp Arg Ala Ser Thr
65                  70                  75                  80

Gly Ile Asp Gly Lys Gln Gln His Pro Glu Asn Thr Met Ala Glu Val
                85                  90                  95

Pro Lys Glu Val Asp Ile Leu Phe Val Phe His Asp His Thr Ala Ser
            100                 105                 110

Asp Ser Pro Phe Trp Ser Glu Leu Lys Asp Ser Tyr Val His Lys Leu
        115                 120                 125

His Gln Gln Gly Thr Ala Leu Val Gln Thr Ile Gly Val Asn Glu Leu
```

130 135 140

Asn Gly Arg Thr Gly Leu Ser Lys Asp Tyr Pro Asp Thr Pro Glu Gly
145 150 155 160

Asn Lys Ala Leu Ala Ala Ala Ile Val Lys Ala Phe Val Thr Asp Arg

165 170 175

Gly Val Asp Gly Leu Asp Ile Asp Ile Glu His Glu Phe Thr Asn Lys
180 185 190

Arg Thr Pro Glu Glu Asp Ala Arg Ala Leu Asn Val Phe Lys Glu Ile
195 200 205

Ala Gln Leu Ile Gly Lys Asn Gly Ser Asp Lys Ser Lys Leu Leu Ile
210 215 220

Met Asp Thr Thr Leu Ser Val Glu Asn Asn Pro Ile Phe Lys Gly Ile
225 230 235 240

Ala Glu Asp Leu Asp Tyr Leu Leu Arg Gln Tyr Tyr Gly Ser Gln Gly
245 250 255

Gly Glu Ala Glu Val Asp Thr Ile Asn Ser Asp Trp Asn Gln Tyr Gln
260 265 270

Asn Tyr Ile Asp Ala Ser Gln Phe Met Ile Gly Phe Ser Phe Phe Glu
275 280 285

Glu Ser Ala Ser Lys Gly Asn Leu Trp Phe Asp Val Asn Glu Tyr Asp
290 295 300

Pro Asn Asn Pro Glu Lys Gly Lys Asp Ile Glu Gly Thr Arg Ala Lys
305 310 315 320

Lys Tyr Ala Glu Trp Gln Pro Ser Thr Gly Gly Leu Lys Ala Gly Ile
325 330 335

Phe Ser Tyr Ala Ile Asp Arg Asp Gly Val Ala His Val Pro Ser Thr
340 345 350

Tyr Lys Asn Arg Thr Ser Thr Asn Leu Gln Arg His Glu Val Asp Asn
355 360 365

Ile Ser His Thr Asp Tyr Thr Val Ser Arg Lys Leu Lys Thr Leu Met
370 375 380

Thr Glu Asp Lys Arg Tyr Asp Val Ile Asp Gln Lys Asp Ile Pro Asp
385 390 395 400

Pro Ala Leu Arg Glu Gln Ile Ile Gln Gln Val Gly Gln Tyr Lys Gly
405 410 415

Asp Leu Glu Arg Tyr Asn Lys Thr Leu Val Leu Thr Gly Asp Lys Ile
420 425 430

```
Gln Asn Leu Lys Gly Leu Glu Lys Leu Ser Lys Leu Gln Lys Leu Glu
        435                 440                 445

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

Glu Ser Met Lys Lys Asp Ala Glu Leu Val Met Val Gly Met Thr Gly
465                 470                 475                 480

Leu Glu Lys Leu Asn Leu Ser Gly Leu Asn Arg Gln Thr Leu Asp Gly
                485                 490                 495

Ile Asp Val Asn Ser Ile Thr His Leu Thr Ser Phe Asp Ile Ser His
            500                 505                 510

Asn Ser Leu Asp Leu Ser Glu Lys Ser Glu Asp Arg Lys Leu Leu Met
        515                 520                 525

Thr Leu Met Glu Gln Val Ser Asn His Gln Lys Ile Thr Val Lys Asn
    530                 535                 540

Thr Ala Phe Glu Asn Gln Lys Pro Lys Gly Tyr Tyr Pro Gln Thr Tyr
545                 550                 555                 560

Asp Thr Lys Glu Gly His Tyr Asp Val Asp Asn Ala Glu His Asp Ile
                565                 570                 575

Leu Thr Asp Phe Val Phe Gly Thr Val Thr Lys Arg Asn Thr Phe Ile
            580                 585                 590

Gly Asp Glu Glu Ala Phe Ala Ile Tyr Lys Glu Gly Ala Val Asp Gly
        595                 600                 605

Arg Gln Tyr Val Ser Lys Asp Tyr Thr Tyr Glu Ala Phe Arg Lys Asp
    610                 615                 620

Tyr Lys Gly Tyr Lys Val His Leu Thr Ala Ser Asn Leu Gly Glu Thr
625                 630                 635                 640

Val Thr Ser Lys Val Thr Ala Thr Thr Asp Glu Thr Tyr Leu Val Asp
                645                 650                 655

Val Ser Asp Gly Glu Lys Val Val His His Met Lys Leu Asn Ile Gly
            660                 665                 670

Ser Gly Ala Ile Met Met Glu Asn Leu Ala Lys Gly Ala Lys Val Ile
        675                 680                 685

Gly Thr Ser Gly Asp Phe Glu Gln Ala Lys Lys Ile Phe Asp Gly Glu
    690                 695                 700

Lys Ser Asp Arg Phe Phe Thr Trp Gly Gln Thr Asn Trp Ile Ala Phe
705                 710                 715                 720

Asp Leu Gly Glu Ile Asn Leu Ala Lys Glu Trp Arg Leu Phe Asn Ala
                725                 730                 735
```

```
Glu Thr Asn Thr Glu Ile Lys Thr Asp Ser Ser Leu Asn Val Ala Lys
            740                 745                 750

Gly Arg Leu Gln Ile Leu Lys Asp Thr Thr Ile Asp Leu Glu Lys Met
        755                 760                 765

Asp Ile Lys Asn Arg Lys Glu Tyr Leu Ser Asn Asp Glu Asn Trp Thr
    770                 775                 780

Asp Val Ala Gln Met Asp Asp Ala Lys Ala Ile Phe Asn Ser Lys Leu
785                 790                 795                 800

Ser Asn Val Leu Ser Arg Tyr Trp Arg Phe Cys Val Asp Gly Gly Ala
                805                 810                 815

Ser Ser Tyr Tyr Pro Gln Tyr Thr Glu Leu Gln Ile Leu Gly Gln Arg
            820                 825                 830

Leu Ser Asn Asp Val Ala Asn Thr Leu Lys Asp
        835                 840
```

<210> 2

<211> 843

<212> PRT

<213> Streptococcus pyogenes

<400> 2

```
Met Asp Lys His Leu Leu Val Lys Arg Thr Leu Gly Cys Val Cys Ala
1               5                   10                  15

Ala Thr Leu Met Gly Ala Ala Leu Ala Thr His His Asp Ser Leu Asn
            20                  25                  30

Thr Val Lys Ala Glu Glu Lys Thr Val Gln Thr Gly Lys Thr Asp Gln
        35                  40                  45

Gln Val Gly Ala Lys Leu Val Gln Glu Ile Arg Glu Gly Lys Arg Gly
    50                  55                  60

Pro Leu Tyr Ala Gly Tyr Phe Arg Thr Trp His Asp Arg Ala Ser Thr
65                  70                  75                  80

Gly Ile Asp Gly Lys Gln Gln His Pro Glu Asn Thr Met Ala Glu Val
                85                  90                  95

Pro Lys Glu Val Asp Ile Leu Phe Val Phe His Asp His Thr Ala Ser
            100                 105                 110

Asp Ser Pro Phe Trp Ser Glu Leu Lys Asp Ser Tyr Val His Lys Leu
        115                 120                 125

His Gln Gln Gly Thr Ala Leu Val Gln Thr Ile Gly Val Asn Glu Leu
    130                 135                 140
```

```
Asn Gly Arg Thr Gly Leu Ser Lys Asp Tyr Pro Asp Thr Pro Glu Gly
145                 150                 155                 160

Asn Lys Ala Leu Ala Ala Ala Ile Val Lys Ala Phe Val Thr Asp Arg
                165                 170                 175

Gly Val Asp Gly Leu Asp Ile Asp Ile Leu His Glu Phe Thr Asn Lys
            180                 185                 190

Arg Thr Pro Glu Glu Asp Ala Arg Ala Leu Asn Val Phe Lys Glu Ile
        195                 200                 205

Ala Gln Leu Ile Gly Lys Asn Gly Ser Asp Lys Ser Lys Leu Leu Ile
    210                 215                 220

Met Asp Thr Thr Leu Ser Val Glu Asn Asn Pro Ile Phe Lys Gly Ile
225                 230                 235                 240

Ala Glu Asp Leu Asp Tyr Leu Leu Arg Gln Tyr Tyr Gly Ser Gln Gly
                245                 250                 255

Gly Glu Ala Glu Val Asp Thr Ile Asn Ser Asp Trp Asn Gln Tyr Gln
            260                 265                 270

Asn Tyr Ile Asp Ala Ser Gln Phe Met Ile Gly Phe Ser Phe Phe Glu
        275                 280                 285

Glu Ser Ala Ser Lys Gly Asn Leu Trp Phe Asp Val Asn Glu Tyr Asp
    290                 295                 300

Pro Asn Asn Pro Glu Lys Gly Lys Asp Ile Glu Gly Thr Arg Ala Lys
305                 310                 315                 320

Lys Tyr Ala Glu Trp Gln Pro Ser Thr Gly Gly Leu Lys Ala Gly Ile
                325                 330                 335

Phe Ser Tyr Ala Ile Asp Arg Asp Gly Val Ala His Val Pro Ser Thr
            340                 345                 350

Tyr Lys Asn Arg Thr Ser Thr Asn Leu Gln Arg His Glu Val Asp Asn
        355                 360                 365

Ile Ser His Thr Asp Tyr Thr Val Ser Arg Lys Leu Lys Thr Leu Met
    370                 375                 380

Thr Glu Asp Lys Arg Tyr Asp Val Ile Asp Gln Lys Asp Ile Pro Asp
385                 390                 395                 400

Pro Ala Leu Arg Glu Gln Ile Ile Gln Gln Val Gly Gln Tyr Lys Gly
                405                 410                 415

Asp Leu Glu Arg Tyr Asn Lys Thr Leu Val Leu Thr Gly Asp Lys Ile
            420                 425                 430

Gln Asn Leu Lys Gly Leu Glu Lys Leu Ser Lys Leu Gln Lys Leu Glu
        435                 440                 445
```

```
Leu Arg Gln Leu Ser Asn Val Lys Glu Ile Thr Pro Glu Leu Leu Pro
    450                 455                 460

Glu Ser Met Lys Lys Asp Ala Glu Leu Val Met Val Gly Met Thr Gly
465                 470                 475                 480

Leu Glu Lys Leu Asn Leu Ser Gly Leu Asn Arg Gln Thr Leu Asp Gly
                485                 490                 495

Ile Asp Val Asn Ser Ile Thr His Leu Thr Ser Phe Asp Ile Ser His
            500                 505                 510

Asn Ser Leu Asp Leu Ser Glu Lys Ser Glu Asp Arg Lys Leu Leu Met
        515                 520                 525

Thr Leu Met Glu Gln Val Ser Asn His Gln Lys Ile Thr Val Lys Asn
    530                 535                 540

Thr Ala Phe Glu Asn Gln Lys Pro Lys Gly Tyr Tyr Pro Gln Thr Tyr

545                 550                 555                 560

Asp Thr Lys Glu Gly His Tyr Asp Val Asp Asn Ala Glu His Asp Ile
                565                 570                 575

Leu Thr Asp Phe Val Phe Gly Thr Val Thr Lys Arg Asn Thr Phe Ile
            580                 585                 590

Gly Asp Glu Glu Ala Phe Ala Ile Tyr Lys Glu Gly Ala Val Asp Gly
        595                 600                 605

Arg Gln Tyr Val Ser Lys Asp Tyr Thr Tyr Glu Ala Phe Arg Lys Asp
    610                 615                 620

Tyr Lys Gly Tyr Lys Val His Leu Thr Ala Ser Asn Leu Gly Glu Thr
625                 630                 635                 640

Val Thr Ser Lys Val Thr Ala Thr Thr Asp Glu Thr Tyr Leu Val Asp
                645                 650                 655

Val Ser Asp Gly Glu Lys Val Val His His Met Lys Leu Asn Ile Gly
            660                 665                 670

Ser Gly Ala Ile Met Met Glu Asn Leu Ala Lys Gly Ala Lys Val Ile
        675                 680                 685

Gly Thr Ser Gly Asp Phe Glu Gln Ala Lys Lys Ile Phe Asp Gly Glu
    690                 695                 700

Lys Ser Asp Arg Phe Phe Thr Trp Gly Gln Thr Asn Trp Ile Ala Phe
705                 710                 715                 720

Asp Leu Gly Glu Ile Asn Leu Ala Lys Glu Trp Arg Leu Phe Asn Ala
                725                 730                 735

Glu Thr Asn Thr Glu Ile Lys Thr Asp Ser Ser Leu Asn Val Ala Lys
```

740 745 750

Gly Arg Leu Gln Ile Leu Lys Asp Thr Thr Ile Asp Leu Glu Lys Met
755 760 765

Asp Ile Lys Asn Arg Lys Glu Tyr Leu Ser Asn Asp Glu Asn Trp Thr
770 775 780

Asp Val Ala Gln Met Asp Asp Ala Lys Ala Ile Phe Asn Ser Lys Leu
785 790 795 800

Ser Asn Val Leu Ser Arg Tyr Trp Arg Phe Cys Val Asp Gly Gly Ala
805 810 815

Ser Ser Tyr Tyr Pro Gln Tyr Thr Glu Leu Gln Ile Leu Gly Gln Arg
820 825 830

Leu Ser Asn Asp Val Ala Asn Thr Leu Lys Asp
835 840

<210> 3

<211> 2532

<212> DNA

<213> Streptococcus pyogenes

<400> 3

atggataaac atttgttggt aaaaagaaca ctagggtgtg tttgtgctgc aacgttgatg 60

ggagctgcct tagcgaccca ccatgattca ctcaatactg taaaagcgga ggagaagact 120

gttcaaacag gaaagacaga tcagcaggtt ggtgctaaat tggtacagga aatccgtgaa 180

ggaaaacgcg gaccactata tgctggttat tttaggacat ggcatgatcg tgcttcaaca 240

ggaatagatg gtaaacagca acatccagaa aatactatgg ctgaggtccc aaaagaagtt 300

gatatcttat ttgtttttca tgaccataca gcttcagata gtccattttg gtctgaatta 360

aaggacagtt atgtccataa attacatcaa cagggaacgg cacttgttca gacaattggt 420

gttaacgaat taaatggacg tacaggttta tctaaagatt atcctgatac tcctgagggg 480

aacaaagctt tagcagcagc cattgtcaag gcatttgtaa ctgatcgtgg tgtcgatgga 540

ctagatattg atattgagca cgaatttacg aacaaaagaa cacctgaaga agatgctcgt 600

gctctaaatg tttttaaaga gattgcgcag ttaataggta aaaatggtag tgataaatct 660

aaattgctca tcatggacac taccctaagt gttgaaaata atccaatatt taaagggata 720

gcggaagatc ttgattatct tcttagacaa tattatggtt cacaaggtgg agaagctgaa 780

gtggatacta taaactctga ttggaaccaa tatcagaatt atattgatgc tagccagttc 840

atgattggat tctccttttt tgaagaatct gcgtccaaag ggaatttatg gtttgatgtt 900

aacgaatacg accctaacaa tcctgaaaaa gggaaagata ttgaaggaac acgtgctaaa 960

aaatatgcag agtggcaacc tagtacaggt ggtttaaaag caggtatatt ctcttatgct 1020

attgatcgtg atggagtggc tcatgttcct tcaacatata aaaataggac tagtacaaat 1080

ttacaacggc atgaagtcga taatatctca catactgact acaccgtatc tcgaaaatta 1140

aaaacattga tgaccgaaga caaacgctat gatgtcattg atcaaaaaga cattcctgac 1200

ccagcattaa gagaacaaat cattcaacaa gttggacagt ataaaggcga tttggaacgt 1260

tataacaaga cattggtgct tacaggagat aagattcaaa atcttaaagg actagaaaaa 1320

ttaagcaagt tacaaaaatt agagttgcgc cagctatcta acgttaaaga aattactcca 1380

gaacttttgc cggaaagcat gaaaaaagat gctgagcttg ttatggtagg catgactggt 1440

ttagaaaaac taaaccttag tggtctaaat cgtcaaactt tagacggtat agacgtgaat 1500

agtattacgc atttgacatc atttgatatt tcacataata gtttggactt gtcggaaaag 1560

agtgaagacc gtaaactatt aatgactttg atggagcagg tttcaaatca tcaaaaaata 1620

acggtgaaaa atacggcttt tgaaaatcaa aaaccgaaag gttattatcc tcagacgtat 1680

gataccaaag aaggtcatta tgatgttgat aatgcagaac atgatatttt aactgatttt 1740

gtttttggaa ctgttactaa acgtaatacc tttattggag acgaagaagc atttgctatc 1800

tataaagaag gagctgtcga tggtcgacaa tatgtgtcta aagactatac ttatgaagct 1860

tttcgtaaag actataaagg ttacaaggtt catttaactg cttctaacct aggagaaaca 1920

gttacttcta aggtaactgc tactactgat gaaacttact tagtagatgt ttctgatggg 1980

gaaaaagttg ttcaccacat gaaactcaat ataggatctg gtgccatcat gatggaaaat 2040

ctggcaaaag gggctaaagt gattggtaca tctggggact ttgagcaagc aaagaagatt 2100

ttcgatggtg aaaagtcaga tagattcttc acttggggac aaactaactg gatagctttt 2160

gatctaggag aaattaatct tgcgaaggaa tggcgtttat ttaatgcaga gacaaatact 2220

gaaataaaga cagatagtag cttaaacgtg gctaaaggac gtcttcagat tttaaaagat 2280

acaactattg atttagaaaa aatggacata aaaaatcgta aagagtatct gtcgaatgat 2340

gaaaattgga ctgatgttgc tcagatggat gatgcaaaag cgatatttaa tagtaaatta 2400

tccaatgttt tatctcggta ttggcggttt tgtgtagatg gtggagctag ctcttattac 2460

cctcaatata ccgaacttca aatcctcgga caacgtttat caaatgatgt cgctaatacg 2520

ctgaaggatt ga 2532

<210> 4

<211> 310

<212> PRT

<213> Streptococcus pyogenes

<400> 4

Asp Ser Phe Ser Ala Asn Gln Glu Ile Arg Tyr Ser Glu Val Thr Pro
1 5 10 15

Tyr His Val Thr Ser Val Trp Thr Lys Gly Val Thr Pro Pro Ala Asn
20 25 30

Phe Thr Gln Gly Glu Asp Val Phe His Ala Pro Tyr Val Ala Asn Gln
35 40 45

Gly Trp Tyr Asp Ile Thr Lys Thr Phe Asn Gly Lys Asp Asp Leu Leu
50 55 60

Cys Gly Ala Ala Thr Ala Gly Asn Met Leu His Trp Trp Phe Asp Gln

65 70 75 80

Asn Lys Asp Gln Ile Lys Arg Tyr Leu Glu Glu His Pro Glu Lys Gln
85 90 95

Lys Ile Asn Phe Asn Gly Glu Gln Met Phe Asp Val Lys Glu Ala Ile
100 105 110

Asp Thr Lys Asn His Gln Leu Asp Ser Lys Leu Phe Glu Tyr Phe Lys
115 120 125

Glu Lys Ala Phe Pro Tyr Leu Ser Thr Lys His Leu Gly Val Phe Pro
130 135 140

Asp His Val Ile Asp Met Phe Ile Asn Gly Tyr Arg Leu Ser Leu Thr
145 150 155 160

Asn His Gly Pro Thr Pro Val Lys Glu Gly Ser Lys Asp Pro Arg Gly
165 170 175

Gly Ile Phe Asp Ala Val Phe Thr Arg Gly Asp Gln Ser Lys Leu Leu
180 185 190

Thr Ser Arg His Asp Phe Lys Glu Lys Asn Leu Lys Glu Ile Ser Asp
195 200 205

Leu Ile Lys Lys Glu Leu Thr Glu Gly Lys Ala Leu Gly Leu Ser His
210 215 220

Thr Tyr Ala Asn Val Arg Ile Asn His Val Ile Asn Leu Trp Gly Ala
225 230 235 240

Asp Phe Asp Ser Asn Gly Asn Leu Lys Ala Ile Tyr Val Thr Asp Ser
245 250 255

Asp Ser Asn Ala Ser Ile Gly Met Lys Lys Tyr Phe Val Gly Val Asn
260 265 270

Ser Ala Gly Lys Val Ala Ile Ser Ala Lys Glu Ile Lys Glu Asp Asn
275 280 285

Ile Gly Ala Gln Val Leu Gly Leu Phe Thr Leu Ser Thr Gly Gln Asp
290 295 300

Ser Trp Asn Gln Thr Asn

305 310

<210> 5

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Cleavage product

<400> 5

```
Gly Pro Ser Val Phe Leu Phe Pro
1               5
```

<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 6
ctgtaaggat ccaggagaag actg 24

<210> 7
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 7
gaaacctcga gtctttgtaa tcgtaggact t 31

<210> 8
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 8
ctagatattg atattcttca cgaatttacg aac 33

<210> 9
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 9

gttcgtaaat tcgtgaagaa tatcaatatc tag 33

<210> 10
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 10
atttctcgag ctgaagacgt cctttagcca cg 32

<210> 11
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 11
taaactcgag ccccatcaga aacatctact aag 33

<210> 12
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 12
attttctcga ggcattatca acatcataat gacc 34

<210> 13
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 13
taaactcgag ccagtcatgc ctaccataac aagctcagc 39

<210> 14
<211> 29

<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 14
atttctcgag ctgtccaact tgttgaatg 29

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 15
aaaacgcgga ccactatatg c 21

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 16
aaacgttgtc cgaggatttg 20

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 17
agcccttgat gatgctttg 19

<210> 18
<211> 19
<212> DNA
<213> Artificial Sequence

<220>

<223> Primer

<400> 18
aacaattctg ggtgatcgg 19

<210> 19
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> GH18 catalytic motif present from residues 179 to 186 of SEQ ID NO: 1

<220>
<221> MISC_FEATURE
<222> (2)..(3)
<223> Xaa refers to an unconserved amino acid

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa refers to an unconserved amino acid

<220>
<221> MISC_FEATURE
<222> (7)..(7)
<223> Xaa refers to an unconserved amino acid

<400> 19

```
Asp Xaa Xaa Asp Xaa Asp Xaa Glu
1               5
```

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO2009033670A [0002] [0002]
- WO2008071418A [0002]

Non-patent literature cited in the description

- **COLLINOLSEN**The EMBO Journal, 2001, vol. 20, 3046-3055 [0002]
- **ALTSCHUL S. F.**J Mol Evol, 1993, vol. 36, 290-300 [0012]
- **ALTSCHUL, S, F et al.**J Mol Biol, 1990, vol. 215, 403-10 [0012]
- **HENIKOFFHENIKOFF**Proc. Natl. Acad. Sci. USA, 1992, vol. 89, 10915-10919 [0012]
- **KARLINALTSCHUL**Proc. Natl. Acad. Sci. USA, 1993, vol. 90, 5873-5787 [0013]
- **DEVEREUX et al.**Nucleic Acids Research, 1984, vol. 12, 387-395 [0013]

PATENTKRAV

1. Fremgangsmåde til fuld deglycosylering af et IgG-antistof omfattende inkubation af antistoffer med et polypeptid, der omfatter
(a) aminosyresekvensen ifølge SEQ ID NO: 1;
(b) en variant deraf med mindst 95 % identitet med aminosyresekvens ifølge SEQ ID NO: 1 over mindst 810 sammenhængende aminosyrer ifølge SEQ ID NO: 1 og med endoglycosidaseakvititeten for et polypeptid, der består af aminosyresekvensen ifølge SEQ ID NO: 1.

2. Fremgangsmåde ifølge krav 1, der endvidere omfatter vurdering af glycosyleringsprofilen for antistoffer ved analyse af de produkter, der fremstilles ved inkubationen.

3. Fremgangsmåde ifølge krav 2, der omfatter:
(a) etablering af kontakt mellem antistoffer og polypeptidet for at hydrolysere glycan fra antistoffet;
(b) separatopn af glycanet fra det deglycosylerede protein;
(c) analyse af det således fremstillede glycan og/eller deglycosylerede protein.

4. Fremgangsmåde ifølge et hvilket som helst af kravene 1 til 3, hvor antistoffet omfatter et monoklonalt IgG-antistof.

DRAWINGS

Figure 1.

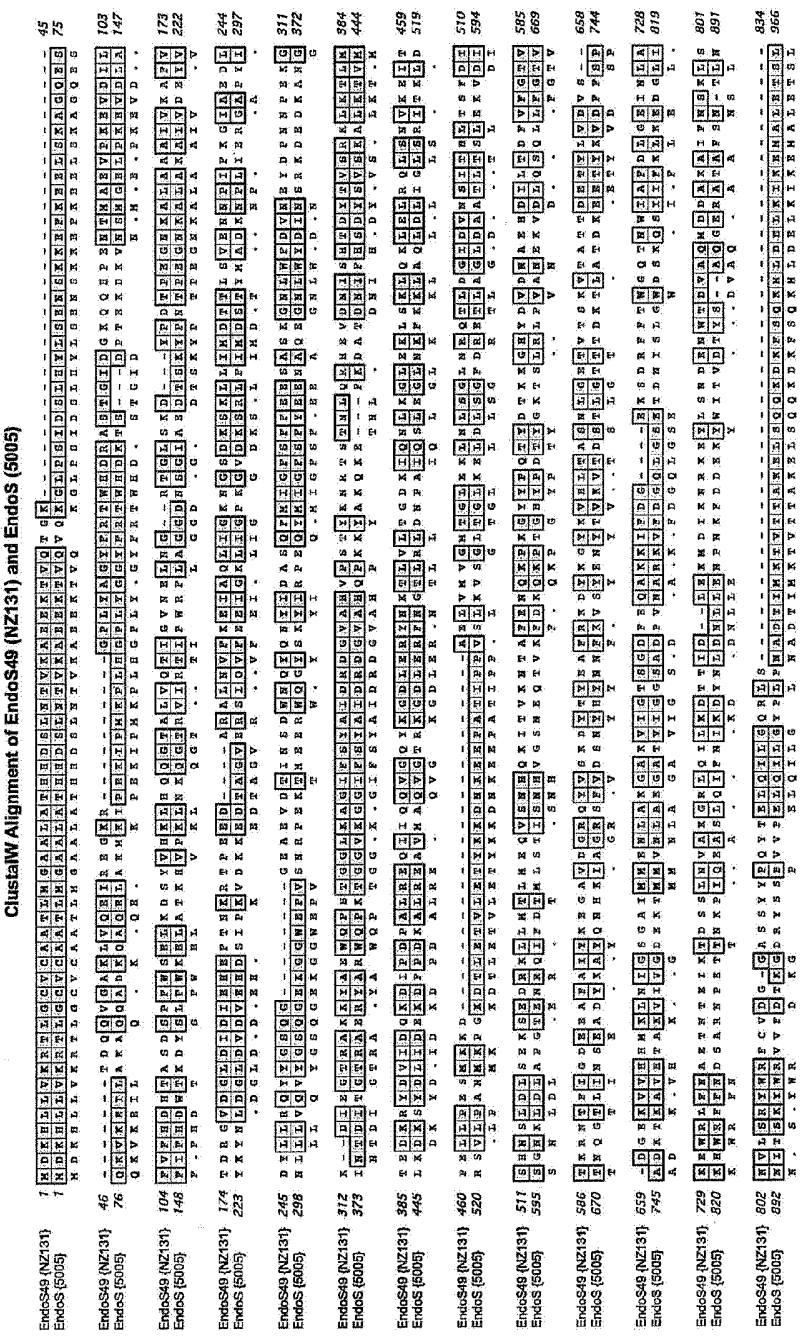

Figure 2.

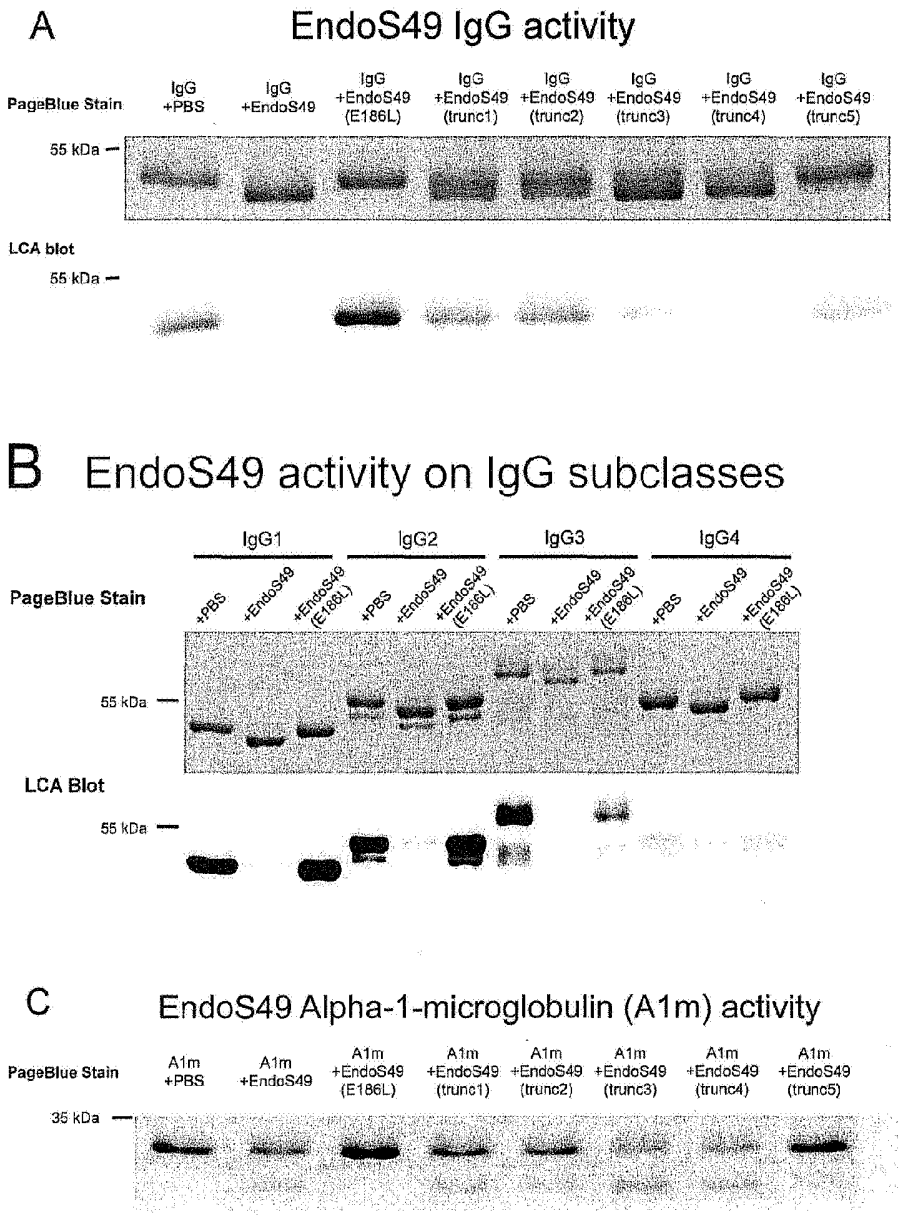
A
EndoS49 IgG activity
PageBlue Stain
IgG +PBS
IgG +EndoS49
IgG +EndoS49 (E186L)
IgG +EndoS49 (trunc1)
IgG +EndoS49 (trunc2)
IgG +EndoS49 (trunc3)
IgG +EndoS49 (trunc4)
IgG +EndoS49 (trunc5)
55 kDa
LCA blot
55 kDa
B
EndoS49 activity on IgG subclasses
IgG1
IgG2
IgG3
IgG4
PageBlue Stain
+PBS
+EndoS49
+EndoS49 (E186L)
55 kDa
LCA Blot
55 kDa
C
EndoS49 Alpha-1-microglobulin (A1m) activity
PageBlue Stain
A1m +PBS
A1m +EndoS49
A1m +EndoS49 (E186L)
A1m +EndoS49 (trunc1)
A1m +EndoS49 (trunc2)
A1m +EndoS49 (trunc3)
A1m +EndoS49 (trunc4)
A1m +EndoS49 (trunc5)
35 kDa

Figure 3.

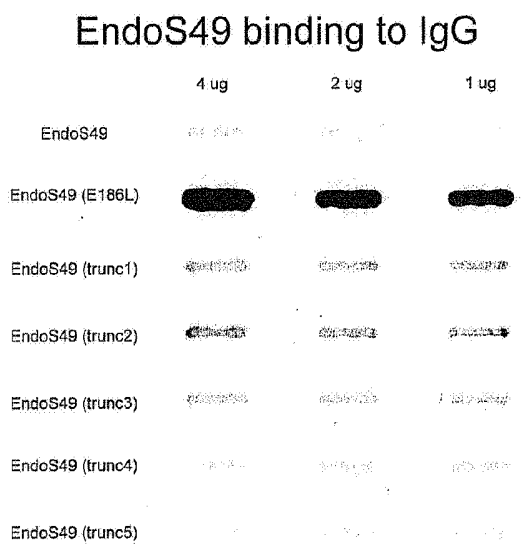

Figure 4.

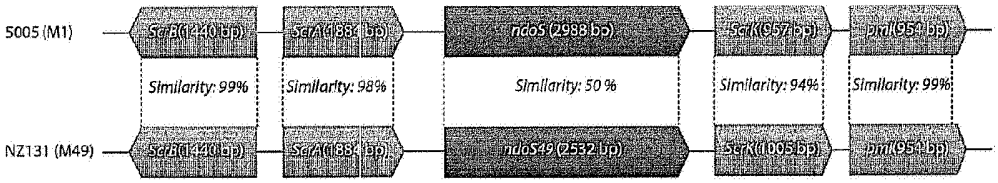

Figure 5. Phylogenetic tree of endoglycosidases

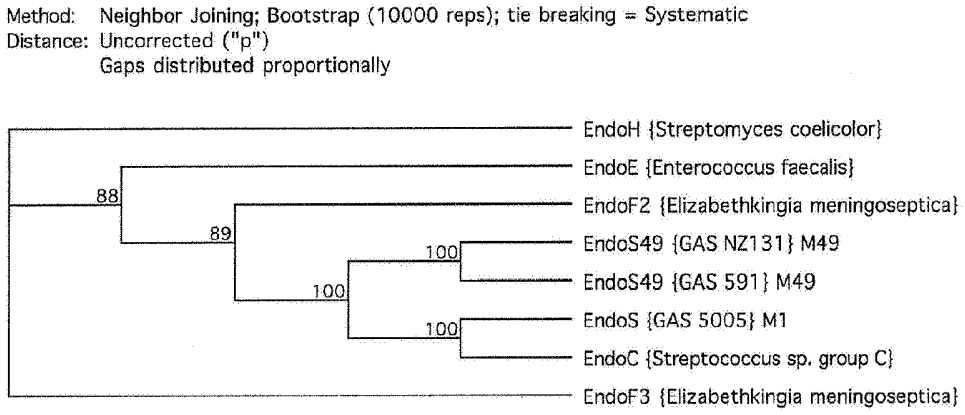

Figure 6.