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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 October 2015 (15.10.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/154783 A1

(51) International Patent Classification:
C07K 14/245 (2006.01) *A61K 39/108* (2006.01)

(21) International Application Number:
PCT/DK2015/050089

(22) International Filing Date:
10 April 2015 (10.04.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/977,838 10 April 2014 (10.04.2014) US

(71) Applicant: **SYDDANSK UNIVERSITET** [DK/DK]; Campusvej 55, DK-5230 Odense M (DK).

(72) Inventors: **BOYSEN, Anders**; Rørhatten 4, DK-5220 Odense SØ (DK). **MØLLER-JENSEN, Jakob**; Smedevænget 20, DK-5230 Odense M (DK). **PALMISANO, Giuseppe**; Via Alcide de Gasperi 8, I-70010 Turi (IT). **LARSEN, Martin Røssel**; Sanderumvej 166A, DK-5250 Odense SV (DK).

(74) Agent: **PLOUGMANN & VINGTOFT A/S**; Rued Langgaards Vej 8, 2300 Copenhagen S (DK).

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

[Continued on next page]

(54) Title: GLYCOSYLATED POLYPEPTIDES ORIGINATING FROM ENTEROTOXIGENIC ESCHERICHIA COLI (ETEC)

(57) Abstract: The present invention relates to glycosylated polypeptides that are immunogenic.These polypeptides originates from or are derived from Enterotoxigenic Escherichia coli (ETEC)and can be comprised in compositions or vaccines. The applications includes immunization, treatment, vaccination and diagnosis of ETEC.

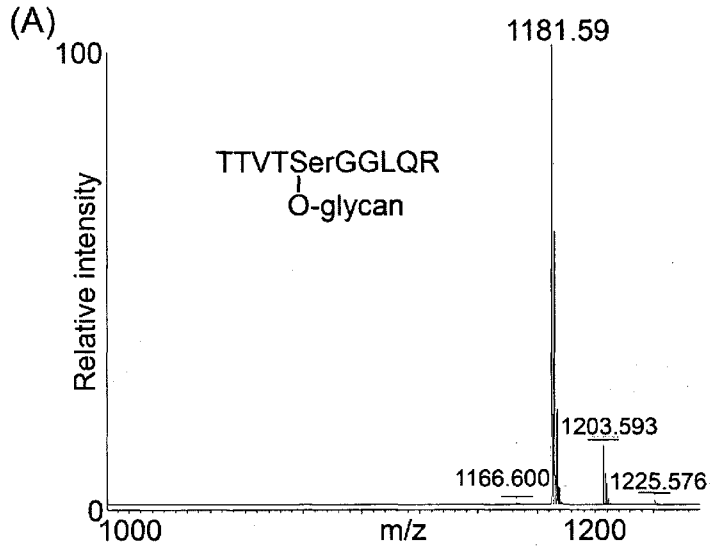

Fig. 1 (a)

WO 2015/154783 A1

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

Glycosylated polypeptides originating from enterotoxigenic Escherichia coli (ETEC)

Field of the invention

The present invention relates to glycosylated polypeptides that are immunogenic. These polypeptides originates from or are derived from enterotoxigenic *Escherichia coli* (ETEC) and can be comprised in compositions or vaccines. The applications include immunization, treatment, vaccination and diagnosis of ETEC.

Background of the invention

Enterotoxigenic *Escherichia coli* (ETEC) is the major source of *E. coli* mediated diarrhoea in humans and livestock. ETEC infections cause more than 280 million annual episodes of diarrhoea resulting in mortality numbers exceeding 300,000 deaths of children under the age of five years.

The significant negative health- and socio-economic impact of ETEC infection manifests itself mainly in the third world nations with poor sanitation and inadequate supplies of clean water. ETEC is a diverse group of pathogens defined by their ability to colonize the small intestine and secrete heat-labile and/or heat stable enterotoxins. The complex pathogenicity is further attributed to the presence of additional bacterial virulence genes on mobile genetic elements such as plasmids and chromosomal pathogenicity islands.

Much attention has been devoted to the understanding of how ETEC and other mucosa-associated pathogens interact with host tissue during infection. Recent work has revealed that bacterial protein glycosylation plays an important role in mediating adhesion, colonization and invasion of host tissue.

Up until now, the known protein glycosylation repertoire of *E. coli* was limited to just four proteins, all of which are surface-exposed adhesins with functions in bacterial pathogenesis. The prototypical ETEC strain H10407 encodes two known glycoproteins, TibA and EtpA.

While the intimate coupling between protein glycosylation and bacterial pathophysiology has become apparent, the discovery of novel glycoproteins implicated in virulence is only advancing slowly. This gap of knowledge is linked to the inherent challenges associated with glycoproteomics. The analytical tools developed for enrichment of eukaryotic O- and N-linked glycopeptides rely on a limited set of defined physiochemical properties, e.g. glycan hydrophilicity or specific lectin recognition, which are relatively rare in bacteria.

Discovery and characterization of glycoproteins is further complicated by heterogeneous glycosylation, low abundance and poor ionization of peptides modified with carbohydrates compared to the non-modified counterpart.

Mapping of O-linked glycan moieties has proven to be a particularly challenging task owing to the diverse nature of carbohydrate structures available for protein modification in bacteria. Although methods such as periodic acid/hydrazide glycan labelling and metabolic oligosaccharide engineering (MOE) have identified glycoproteins in a range of bacteria, these techniques present limitations in the form of low specificity for glycosylated proteins and dependence on sugar uptake and integration into bacterial glycoproteins, respectively.

Although they are poorly understood, bacterial glycoproteins potentially constitute an important reservoir of novel therapeutic targets, which could be used against bacterial pathogens.

Thus, there is a great need for understanding the glycosylation patterns of proteins originating from bacteria such as ETEC, and revealing the effect of the glycosylations on for example immunogenicity.

Summary of the invention

An object of the present invention is to provide glycosylated polypeptides that are immunogenic.

In one aspect of the present invention is the polypeptide NmpC (also known as ETEC_0806 or CBJ00316, SEQ ID NO: 1), TibA autotransporter (also known as CBJ01643 or ETEC_2141, SEQ ID NO: 20), EtpA (also known as CBJ04458, SEQ ID NO: 31), EatA (also known as CBJ04449, SEQ ID NO: 39), CexE (also known as CBJ04537, SEQ ID NO: 48), or Ag43 (also known as CBJ03971, SEQ ID NO: 51).

Another aspect of the present invention relates to one of these polypeptides comprising the full length sequence, a polypeptide fragment of SEQ ID NO: 1 that has at least 75 % sequence identity to the full length sequence, or a B- or T-cell epitope of the full length sequence, wherein the polypeptide is glycosylated at least in one position.

In one embodiment of the present invention is the polypeptide is glycosylated at least in two or three positions.

In a further embodiment of the present invention originates the polypeptides of the present invention from enterotoxigenic *Escherichia coli* (ETEC) or are synthetic.

Fragments of the polypeptides include the glycosylated polypeptide selected from the group of glycosylated polypeptides consisting of SEQ ID NOs: 2-19.

Yet another aspect of the present invention relates to the polypeptide, wherein the polypeptide is a fusion polypeptide.

A further aspect of the present invention relates to the polypeptides as described herein comprised in an immunogenic composition, a pharmaceutical composition, or a vaccine.

These may be formulated for intradermal, transdermal, subcutaneous, intramuscular or mucosal application.

One other aspect of the present invention relates to a nucleic acid sequence encoding a polypeptide.

Another aspect of the present invention relates to a method for immunizing a mammal, the method comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine as described herein.

A further aspect of the present invention relates to a method for treating a mammal, which is infected with ETEC comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine as described herein.

Another aspect of the present invention relates to the polypeptide, immunogenic composition, pharmaceutical composition, or vaccine as described herein for use in treating infection caused by ETEC.

Yet another aspect of the present invention relates to the polypeptide, immunogenic composition, pharmaceutical composition, or vaccine as described herein for use in the preparation of a medicament for treating infection caused by ETEC.

Another aspect of the present invention relates to the polypeptide, immunogenic composition, pharmaceutical composition, or vaccine as described herein for use in the diagnosis of an infection caused by ETEC.

Brief description of the figures

Figure 1 shows β-Elimination of glycan moiety and replacement with 2-AEP through Michael addition chemistry. (1A) MALDI MS spectrum of TTVTSGGLQR (m/z = 1181.59 Da) synthetic O-linked glycopeptide. (1B) The BEMAP reaction efficiently replaces the carbohydrate moiety with the 2-AEP molecule and produces a phosphopeptide with the mass of 1126.64 Da. Minor traces of beta-eliminated as well as intact peptide can be observed (m/z = 1001.62 Da and 1181.59, respectively). (1C) The AEP modified peptide is selectively enriched

with TiO_2 as both the glycopeptide and the beta-eliminated peptide is absent in the MALDI MS spectrum. (1D) MALDI MS peptide mass fingerprint of Tryptic digest of heptosylated protein Ag43. Ag43 can be digested into a mixture of heptosylated as well as unmodified peptides. Peptides marked with an asterisk indicate heptosylation. (1E) BEMAP converts heptosylated peptides into phosphopeptides; modified peptides are indicated. (1F) Specific TiO_2 enrichment of phosphopeptides.

Figure 2 shows gas-phase fragmentation properties of 2-AEP tagged peptide with either CID or HCD. Mapping of glycosylated residues by higher-energy collisional dissociation (HCD) fragmentation is superior to CID. Moreover, HCD fragmentation yields two characteristic ions (m/z=126.03Da and m/z=138.03Da), useful for the identification of formerly glycosylated peptides in complex MS/MS spectra.

The present invention will now be described in more detail in the following.

Detailed description of the invention

The present inventors have developed a novel mass spectrometry-based technique, termed BEMAP, that can be employed to map O-linked glycoproteins from theoretically any organism.

BEMAP combines a simple reaction scheme with a highly selective enrichment protocol to circumvent the challenges previously associated with bacterial glycoproteomics. The BEMAP reaction efficiently substitutes O-linked carbohydrate moieties with a 2-Aminoethyl phosphonic acid (AEP) group, which can be selectively isolated based on its affinity for titanium dioxide.

BEMAP has been employed to map novel protein glycosylation sites in ETEC strain H10407 and the non-pathogenic *E. coli* K-12 MG1655. Functional characterization of an H10407Δ*hldE* knockout strain revealed the importance of protein glycosylation for ETEC adhesion to human intestinal cells.

These results, together with other recent studies of bacterial glycoproteomes, highlight protein glycosylation in bacteria as an abundant, yet largely unexplored, posttranslational protein modification which is central to bacterial physiology and pathophysiology.

The ETEC glycosylated proteins (polypeptides) are important in understanding the immunogenicity of ETEC. The glycosylated polypeptides disclosed herein leads to an enhanced immunogenicity compared to the same polypeptides that are not glycosylated.

The present inventors have therefore surprisingly found that certain proteins from ETEC causes an enhanced immunogenic response due to specific glycosylations of the proteins.

Thus, an object of the present invention is to provide glycosylated polypeptides that are immunogenic.

The term glycosylation refers to O-linked glycosylation. This is the attachment of a sugar molecule to a hydroxyl oxygen of either a Serine or Threonine side chain in a protein.

The proteins include NmpC (also known as ETEC_0806 or CBJ00316, SEQ ID NO: 1), TibA autotransporter (also known as CBJ01643 or ETEC_2141, SEQ ID NO: 20), EtpA (also known as CBJ04458, SEQ ID NO: 31), EatA (also known as CBJ04449, SEQ ID NO: 39), CexE (also known as CBJ04537, SEQ ID NO: 48), and Ag43 (also known as CBJ03971, SEQ ID NO: 51).

Thus, one aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 1, a polypeptide fragment of SEQ ID NO: 1 that has at least 75 % sequence identity to SEQ ID NO: 1, or a B- or T-cell epitope of SEQ ID NO: 1, wherein the polypeptide is glycosylated at least in one position.

Another aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 20, a polypeptide fragment of SEQ ID NO: 20 that has at least 75

% sequence identity to SEQ ID NO: 20, or a B- or T-cell epitope of SEQ ID NO: 20, wherein the polypeptide is glycosylated at least in one position.

Yet another aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 31, a polypeptide fragment of SEQ ID NO: 31 that has at least 75 % sequence identity to SEQ ID NO: 31, or a B- or T-cell epitope of SEQ ID NO: 31, wherein the polypeptide is glycosylated at least in one position.

A further aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 39, a polypeptide fragment of SEQ ID NO: 39 that has at least 75 % sequence identity to SEQ ID NO: 39, or a B- or T-cell epitope of SEQ ID NO: 39, wherein the polypeptide is glycosylated at least in one position.

Another aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 48, a polypeptide fragment of SEQ ID NO: 48 that has at least 75 % sequence identity to SEQ ID NO: 48, or a B- or T-cell epitope of SEQ ID NO: 48, wherein the polypeptide is glycosylated at least in one position.

Yet another aspect of the present invention relates to a polypeptide comprising SEQ ID NO: 51, a polypeptide fragment of SEQ ID NO: 51 that has at least 75 % sequence identity to SEQ ID NO: 51, or a B- or T-cell epitope of SEQ ID NO: 51, wherein the polypeptide is glycosylated at least in one position.

The polypeptides of the present invention may be glycosylated at least in two positions, such as at least in three positions, at least four positions, at least five positions, at least six positions, or at least seven positions.

The polypeptides can also be glycosylated in exactly one, two three, four, five, sex or seven positions.

The polypeptides of the present invention originates from Enterotoxigenic *Escherichia coli* (ETEC) as shown in the examples of the present disclosure. The polypeptides can therefore be purified directly from ETEC but they can also be chemically synthesized.

The polypeptides can also be functional variants of the polypeptides disclosed herein. Such variance can be determined by sequence identity.

The term "sequence identity" indicates a quantitative measure of the degree of homology between two amino acid sequences of substantially equal length or between two nucleic acid sequences of substantially equal length. The two sequences to be compared must be aligned to best possible fit possible with the insertion of gaps or alternatively, truncation at the ends of the protein sequences. The sequence identity can be calculated as $\frac{(N_{ref} - N_{dif})100}{N_{ref}}$, wherein Ndif is the total number of non-identical residues in the two sequences when aligned and wherein Nref is the number of residues in one of the sequences. Hence, the DNA sequence AGTCAGTC will have a sequence identity of 75% with the sequence AATCAATC (Ndif=2 and Nref=8). A gap is counted as non-identity of the specific residue(s), i.e. the DNA sequence AGTGTC will have a sequence identity of 75% with the DNA sequence AGTCAGTC (Ndif=2 and Nref=8). Sequence identity can alternatively be calculated by the BLAST program e.g. the BLASTP program (Pearson W.R and D.J. Lipman (1988))(www.ncbi.nlm.nih.gov/cgi-bin/BLAST). In one embodiment of the invention, alignment is performed with the sequence alignment method ClustalW with default parameters as described by Thompson J., et al 1994, available at http://www2.ebi.ac.uk/clustalw/.

A preferred minimum percentage of sequence identity is at least 75 %, such as at least 80%, such as at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, and at least 99.5%.

Preferably, the numbers of substitutions, insertions, additions or deletions of one or more amino acid residues in the fusion polypeptide is limited, i.e. no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 substitutions, no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 insertions, no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 additions, and no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 deletions compared to the immunogenic polypeptide units based on polypeptides disclosed herein.

Polypeptides such as the ETEC proteins disclosed herein can contain immunogenic parts, such as B- or T-cell epitopes.

The immunogenic part of an immunogenic polypeptide is the part of the polypeptide, which elicits an immune response in an animal or a human being, and/or in a biological sample determined by any of the biological assays known to the skilled person working with immune responses.

The immunogenic part of a polypeptide may be a T-cell epitope or a B-cell epitope and can be related to one or a few relatively small parts of the polypeptide, they can be scattered throughout the polypeptide sequence or be situated in specific parts of the polypeptide.

In order to identify relevant T-cell epitopes which are recognised during an immune response, it is possible to use a "brute force" method: Since T-cell epitopes are linear, deletion mutants of the polypeptide will, if constructed systematically, reveal what regions of the polypeptide are essential in immune recognition, e.g. by subjecting these deletion mutants e.g. to assays known to the skilled person working with immune responses.

Another method utilizes overlapping oligopeptides for the detection of MHC class II epitopes, preferably synthetic, having a length of e.g. 20 amino acid residues derived from the polypeptide. These peptides can be tested in biological assays and some of these will give a positive response (and thereby be immunogenic) as evidence for the presence of a T cell epitope in the peptide.

For the detection of MHC class I epitopes it is possible to predict peptides that will bind and hereafter produce these peptides synthetically and test them in relevant biological assays. The peptides preferably having a length of e.g. 8 to 11 amino acid residues derived from the polypeptide. B-cell epitopes can be determined by analyzing the B-cell recognition to overlapping peptides covering the polypeptide of interest.

B-cell epitopes differ from T-cell epitopes in that they are conformational epitopes that require a three dimensional structure in order to raise an immune response. Not intended to be bound by theory can variants of B-cell epitopes therefore be identified through key amino acids (for example glycosylated amino acids) and a certain length of the polypeptide while remaining immunogenic.

Thus, an embodiment of the present invention therefore relates to epitopes, such as B- or T-cell epitopes on the polypeptides mentioned herein.

A common feature of the polypeptides of the present invention is their capability to induce an immunological response as illustrated in the examples. It is understood that within the scope of the present invention are variants of the polypeptides of the invention produced by substitution, insertion, addition or deletion while remaining immunogenic.

Examples of such epitopes are listed in the examples of the present disclosure. They include SEQ ID NOs: 2-19 of NmpC, SEQ ID NOs: 21-30 of TibA, SEQ ID NOs: 32-38 of EtpA, SEQ ID NOs: 40-47 of EatA, SEQ ID NOs: 49-50 of CexE, and SEQ ID NOs: 52-68 of Ag43.

Key features of these epitopes or fragments are that they comprise one or more glycosylations at central positions.

In the present context the term "substantially pure polypeptide" means a polypeptide preparation which contains at most 5% by weight of other polypeptide material with which it is associated natively or during recombinant or synthetic production (lower percentages of other polypeptide material are preferred, e.g. at most 4%, at most 3%, at most 2%, at most 1%, and at most ½%).

It is preferred that the substantially pure polypeptide is at least 96% pure, i.e. that the polypeptide constitutes at least 96% by weight of total polypeptide

material present in the preparation, and higher percentages are preferred, such as at least 97%, at least 98%, at least 99%, at least 99,25%, at least 99,5%, and at least 99,75%. It is especially preferred that the polypeptide is in "essentially pure form", i.e. that the polypeptide is essentially free of any other antigen with which it is natively associated, i.e. free of any other antigen from bacteria. This can be accomplished by preparing the polypeptide by means of recombinant methods in a host cell, or by synthesizing the polypeptide by the well-known methods of solid or liquid phase peptide synthesis, and by using appropriate purification procedures well known to the person of ordinary skill in the art.

Thus in one embodiment of the present invention are the polypeptides of the present invention substantially pure or in essentially pure form.

Fusion polypeptides

Two or more of the polypeptides disclosed herein may be fused to form fusion polypeptides.

Thus relates one embodiment of the present invention to the polypeptides as defined herein, wherein the polypeptide is comprised in a fusion polypeptide.

The polypeptides to which fusion is made may originate from ETEC or alternatively be other polypeptides that are beneficial when an enhanced immune response against ETEC is required.

Thus, another embodiment of the present invention relates to a polypeptide as defined herein, wherein the polypeptide is fused to a polypeptide originating from ETEC.

Immunogenic

An immunogenic polypeptide is defined as a polypeptide that induces an immune response. The immune response may be monitored by one of the following methods:

An *in vitro* cellular response is determined by release of a relevant cytokine such as IFN-γ, from lymphocytes withdrawn from an animal or human currently or previously infected with ETEC, or by detection of proliferation of these T cells. The induction is performed by addition of the polypeptide or the immunogenic part to a suspension comprising from $1x10^5$ cells to $3x10^5$ cells per well. The cells are isolated from either blood, the spleen, the liver or the lung and the addition of the polypeptide or the immunogenic part of the polypeptide result in a concentration of not more than 20 µg per ml suspension and the stimulation is performed from two to five days. For monitoring cell proliferation the cells are pulsed with radioactive labeled Thymidine and after 16-22 hours of incubation the proliferation is detected by liquid scintillation counting. A positive response is a response more than background plus two standard deviations. The release of IFN- γ can be determined by the ELISA method, which is well known to a person skilled in the art. A positive response is a response more than background plus two standard deviations. Other cytokines than IFN- γ could be relevant when monitoring an immunological response to the polypeptide, such as IL-12, TNF-α, IL-4, IL-5, IL-10, IL-6, TGF-β.

Another and more sensitive method for determining the presence of a cytokine (e.g. IFN-γ) is the ELISPOT method where the cells isolated from either the blood, the spleen, the liver or the lung are diluted to a concentration of preferable of 1 to 4 x 10^6 cells /ml and incubated for 18-22 hrs in the presence of the polypeptide or the immunogenic part of the polypeptide resulting in a concentration of not more than 20 µg per ml. The cell suspensions are hereafter diluted to 1 to 2 x 10^6/ ml and transferred to Maxisorp plates coated with anti–IFN-γ and incubated for preferably 4 to 16 hours. The IFN-γ producing cells are determined by the use of labelled secondary anti-IFN-antibody and a relevant substrate giving rise to spots, which can be enumerated using a dissection microscope. It is also a possibility to determine the presence of mRNA coding for the relevant cytokine by the use of the PCR technique. Usually one or more cytokines will be measured utilizing for example the PCR, ELISPOT or ELISA. It will be appreciated by a person skilled in the art that a significant increase or

decrease in the amount of any of these cytokines induced by a specific polypeptide can be used in evaluation of the immunological activity of the polypeptide.

An *in vitro* cellular response may also be determined by the use of T cell lines derived from an immune individual or an ETEC infected person where the T cell lines have been driven with either live ETEC, extracts from the bacterial cell or culture filtrate for 10 to 20 days with the addition of IL-2. The induction is performed by addition of not more than 20 µg polypeptide per ml suspension to the T cell lines containing from $1x10^5$ cells to $3x10^5$ cells per well and incubation is performed from two to six days. The induction of IFN-γ or release of another relevant cytokine is detected by ELISA. The stimulation of T cells can also be monitored by detecting cell proliferation using radioactively labeled Thymidine as described above. For both assays a positive response is a response more than background plus two standard deviations.

An *in vivo* cellular response may be determined as a positive DTH response after intradermal injection or local application patch of at most 100 µg of the polypeptide or the immunogenic part to an individual who is clinically or subclinically infected with ETEC, a positive response having a diameter of at least 5 mm 72-96 hours after the injection or application.

An *in vitro* humoral response is determined by a specific antibody response in an immune or infected individual. The presence of antibodies may be determined by an ELISA technique or a Western blot where the polypeptide or the immunogenic part is absorbed to either a nitrocellulose membrane or a polystyrene surface. The serum is preferably diluted in PBS from 1:10 to 1:100 and added to the absorbed polypeptide and the incubation being performed from 1 to 12 hours. By the use of labeled secondary antibodies the presence of specific antibodies can be determined by measuring the presence or absence of a specific label e.g. by ELISA where a positive response is a response of more than background plus two standard deviations or alternatively a visual response in a Western blot.

Another relevant parameter is measurement of the protection in animal models induced after vaccination with the polypeptide in an adjuvant or after DNA vaccination. Suitable animal models include primates, guinea pigs or mice, which are challenged with an infection of an ETEC. Readout for induced protection could be decrease of the bacterial load in target organs compared to non-vaccinated animals, prolonged survival times compared to non-vaccinated animals and diminished weight loss or pathology compared to non-vaccinated animals.

Thus are the glycosylated polypeptides described herein immunogenic when one of the above described tests is positive.

In one aspect of the present invention are the polypeptides described herein immunogenic.

In another aspect of the present invention relates to a composition comprising the polypeptides described herein. Such composition will constitute an immunogenic composition.

Antibodies

The glycosylated polypeptides disclosed herein can constitute epitopes.

An epitope, also known as antigenic determinant, is the part of an antigen that is recognized by the immune system, specifically by antibodies, B cells, or T cells.

The epitopes of protein antigens are divided into two categories, conformational epitopes and linear epitopes, based on their structure and interaction with the paratope.

A conformational epitope is composed of discontinuous sections of the antigen's amino acid sequence.

These epitopes interact with the paratope based on the 3-D surface features and shape or tertiary structure of the antigen.

By contrast, linear epitopes interact with the paratope based on their primary structure. A linear epitope is formed by a continuous sequence of amino acids from the antigen.

Thus, one aspect of the present invention relates to an antibody that binds to an epitope described herein.

Antibodies raised against the epitope may be either polyclonal or monoclonal.

The antibodies may be is suitable to generate chimeric and/ or human versions that could be appropriate for human in vivo use.

The immunoglobulin heavy chain (IgH) is the large polypeptide subunit of an antibody (immunoglobulin). A typical antibody is composed of two immunoglobulin (Ig) heavy chains and two Ig light chains.

Several different types of heavy chain exist that define the class or isotype of an antibody. These heavy chain types vary between different animals.

The immunoglobulin light chain is the small polypeptide subunit of an antibody (immunoglobulin).

There are two types of light chain in humans (as in other mammals), kappa (κ) chain, encoded by the immunoglobulin kappa locus on chromosome 2 and the lambda (λ) chain, encoded by the immunoglobulin lambda locus on chromosome 22.

Antibodies are produced by B lymphocytes, each expressing only one class of light chain.

Once set, light chain class remains fixed for the life of the B lymphocyte.

In a healthy individual, the total kappa to lambda ratio is roughly 2:1 in serum (measuring intact whole antibodies) or 1:1.5 if measuring free light chains, with a highly divergent ratio indicative of neoplasm.

The exact normal ratio of kappa to lambda ranges from 0.26 to 1.65.

Both the kappa and the lambda chains can increase proportionately, maintaining a normal ratio.

Carriers, excipients and diluents

Pharmaceutical compositions comprising the polypeptides described herein may be administered in a physiologically acceptable medium (e.g., deionized water, phosphate buffered saline (PBS), saline, aqueous ethanol or other alcohol, plasma, proteinaceous solutions, mannitol, aqueous glucose, vegetable oil, or the like).

Thus, a further embodiment of the present invention relates to a composition comprising a polypeptides as described herein that constitutes a pharmaceutical composition.

Buffers may also be included, particularly where the media are generally buffered at a pH in the range of about 5 to 10, where the buffer will generally range in concentration from about 50 to 250 mM salt, where the concentration of salt will generally range from about 5 to 500 mM, physiologically acceptable stabilizers, and the like.

The compounds may be lyophilized for convenient storage and transport.

Thus, in a further embodiment of the present invention the composition comprises one or more excipients, diluents and/or carriers.

Aqueous suspensions may contain the active materials in admixture with excipients suitable for the manufacture of aqueous suspensions.

Such excipients include suspending agents, for example sodium carboxymethylcellulose, methylcellulose, hydropropyl-methylcellulose, sodium alginate, polyvinylpyrrolidone, gum tragacanth and gum acacia; dispersing or wetting agents can be a naturally-occurring phosphatide, for example, lecithin, or condensation products of an alkylene oxide with fatty acids, for example polyoxyethylene stearate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example heptadecaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate.

Thus, an aspect of the present invention relates to a pharmaceutical composition comprising the polypeptide as described herein and at least one pharmaceutically acceptable carrier, excipient or diluent.

Vaccines, treatment and administration

The polypeptides, immunogenic compositions, and pharmaceutical composition may constitute a vaccine against ETEC.

Key features of vaccines is that they are recognized by the recipient's immune response, generate a response, and ultimately decrease the bacterial load of ETEC.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be prophylactic or therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, including, e.g., the capacity of the individual's immune system to mount an immune response, and the degree of protection desired. Suitable dosage ranges are of the order of several hundred micrograms of the fusion polypeptide of the invention per vaccination with a preferred range from about 0.1 µg to 1000 µg, such as in the range from about 1 µg to 300 µg, and especially in the range from about 10 µg to 100 µg. Suitable regimens for initial

administration and booster shots are also variable but are typified by an initial administration followed by subsequent inoculations or other administrations.

The manner of application may be varied widely. Any of the conventional methods for administration of a vaccine are applicable. These include oral, nasal or mucosal application in either a solid form containing the active ingredients (such as a pill, suppository or capsule) or in a physiologically acceptable dispersion, such as a spray, powder or liquid, or parenterally, by injection, for example, subcutaneously, intradermally or intramuscularly or transdermally applied. The dosage of the vaccine will depend on the route of administration and will vary according to the age of the person to be vaccinated and, to a lesser degree, the size of the person to be vaccinated. Currently, most vaccines are administered intramuscularly by needle injection and this is likely to continue as the standard route. However, vaccine formulations which induce mucosal immunity have been developed, typically by oral or nasal delivery. One of the most widely studies delivery systems for induction of mucosal immunity contains cholera toxin (CT) or its B subnit. This protein enhances mucosal immune responses and induces IgA production when administered in vaccine formulations. An advantage is the ease of delivery of oral or nasal vaccines. Modified toxins from other microbial species, which have reduced toxicity but retained immunostimulatory capacity, such as modified heat-labile toxin from Gram-negative bacteria or staphylococcal enterotoxins may also be used to generate a similar effect. These molecules are particularly suited to mucosal administration.

The vaccines are conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkalene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of 0.5% to 10%, preferably 1-2%. Oral formulations include such normally employed excipients as, for example, pharmatical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium

carbonate, and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and advantageously contain 10-95% of active ingredient, preferably 25-70%.

Thus, in one aspect of the present invention are the polypeptides, immunogenic compositions, and pharmaceutical compositions formulated for intradermal, transdermal, subcutaneous, intramuscular or mucosal application.

The adjuvant is preferably selected from the group consisting of dimethyloctadecylammonium bromide (DDA), dimethyloctadecenylammonium bromide (DODAC), Quil A, poly I:C, aluminium hydroxide, Freund's incomplete adjuvant, IFN-γ, IL-2, IL-12, monophosphoryl lipid A (MPL), Treholose Dimycolate (TDM), Trehalose Dibehenate and muramyl dipeptide (MDP).

The polypeptides may also be used for immunizing a mammal against ETEC or treating the mammal against ETEC.

Thus relates one aspect of the present invention to a method for immunizing a mammal, the method comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine as described herein.

Another aspect of the present invention relates to a method for treating a mammal, which is infected with ETEC comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine as described herein.

In one embodiment of the present invention is the mammal a human.

In another embodiment of the present invention is the mammal an animal selected from the group consisting of a pig, a cow, a sheep, and a horse.

A further aspect of the present invention relates to an immunogenic composition, pharmaceutical composition, or vaccine as described herein for use in treating infection caused by ETEC.

Yet another aspect of the present invention relates to an immunogenic composition, pharmaceutical, or vaccine as described herein for use in the preparation of a medicament for treating infection caused by ETEC.

Nucleic acids

By the terms "nucleic acid fragment" and "nucleic acid sequence" are understood any nucleic acid molecule including DNA, RNA, LNA (locked nucleic acids), PNA, RNA, dsRNA and RNA-DNA-hybrids. Also included are nucleic acid molecules comprising non-naturally occurring nucleosides. The term includes nucleic acid molecules of any length e.g. from 10 to 10000 nucleotides, depending on the use. When the nucleic acid molecule is for use as a pharmaceutical, e.g. in DNA therapy, or for use in a method for producing a polypeptide according to the invention, a molecule encoding at least one epitope is preferably used, having a length from about 18 to about 1000 nucleotides, the molecule being optionally inserted into a vector.

When the nucleic acid molecule is used as a probe, as a primer or in antisense therapy, a molecule having a length of 10-100 is preferably used.

According to the invention, other molecule lengths can be used, for instance a molecule having at least 12, 15, 21, 24, 27, 30, 33, 36, 39, 42, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500 or 1000 nucleotides (or nucleotide derivatives), or a molecule having at most 10000, 5000, 4000, 3000, 2000, 1000, 700, 500, 400, 300, 200, 100, 50, 40, 30 or 20 nucleotides (or nucleotide derivatives).

Thus, in another aspect of the present invention is the polypeptide as defined herein encoded nucleic acid sequence.

Diagnosis

Immunodiagnostics are well-suited for the detection of even the smallest of amounts of biochemical substances such as antibodies. Antibodies specific for a desired antigen can be conjugated with a radiolabel, fluorescent label, or color-forming enzyme and are used as a "probe" to detect it. Well known applications include pregnancy tests, immunoblotting, ELISA and immunohistochemical staining of microscope slides. The speed, accuracy and simplicity of such tests has led to the development of rapid techniques for the diagnosis of disease.

Thus, one aspect of the present invention relates to a polypeptide, immunogenic composition, or pharmaceutical composition as described herein for use in the diagnosis of an infection caused by ETEC.

The polypeptide, immunogenic composition, or pharmaceutical composition as described herein may also be used to detect the present of ETEC in a sample or used as an indication whether a sample or subject may contain ETEC.

It should be noted that embodiments and features described in the context of one of the aspects of the present invention also apply to the other aspects of the invention.

All patent and non-patent references cited in the present application, are hereby incorporated by reference in their entirety.

Examples

BEMAP method results:
BEMAP relies on β-elimination of O-linked carbohydrate modifications, Michael addition of 2-Aminoethyl phosphonic acid (AEP) and TiO_2 enrichment of phosphopeptides. Thus, BEMAP combines a firmly established in vitro chemical modification with a highly selective enrichment protocol (Thingholm et al., 2006) and the reactions take place in a single volume without the need for intermediate purification steps as described in the Experimental Procedures section.

The BEMAP method was first established using a synthetic mannosylated peptide as a model compound. As shown in Figure 1A and 1B, MALDI MS demonstrated that BEMAP efficiently replaces the carbohydrate moiety of the synthetic peptide (m/z = 1181.59 Da) with the AEP group and thus produces a phosphopeptide (m/z = 1126.64 Da).

The overall efficiency of substitution exceeds 95% (Figure 1B) without the formation of degradation products. The AEP-modified peptide was then strongly enriched using affinity chromatography with TiO_2; both the intact glycopeptide and the β-eliminated peptide (1001.62Da) were absent in the MALDI MS spectrum after enrichment (Figure 1C).

We found that BEMAP converts other glycopeptides into a phosphopeptide, independent of the identity of the linked monosaccharide (data not shown). It should be noted that the TiO_2 purification step of BEMAP also targets phosphopeptides. Therefore, as a precaution we use the enzyme Alkaline phosphatase to dephosphorylate any native phosphopeptides which otherwise may result in false positive identifications.

We analysed the gas phase-induced fragmentation properties of the converted glycopeptide. As shown in Figure 2, the exchange of a carbohydrate moiety with AEP has several advantages. The AEP addition substitutes a labile glycoside bond with a stronger covalent C-N bond which greatly improves mapping of glycosylated residues by higher-energy collisional dissociation (HCD) fragmentation. Moreover, the AEP group yielded two characteristic ions during HCD fragmentation (m/z=126.03Da and m/z=138.03Da), which are very useful for the identification of formerly glycosylated peptides in complex MS/MS spectra. It should be noted that the AEP molecule is constituted by a phosphonate functional group which is stable under CID and HCD fragmentation conditions compared to the phosphate one which is labile under these conditions. This allows unambiguous assignment of modified amino acid residues and avoids false positives in site localization assignment (data not shown).

Next, we applied BEMAP to a purified heptosylated protein: Ag43 from *E. coli* (Knudsen et al., 2008). As may be seen in Figure 1D, in-gel digestion of the glycosylated protein yielded heptosylated and unmodified peptides. Heptosylated peptides are marked by an asterisk. From the digested peptide mix, BEMAP enriched the three heptosylated peptides present in Figure 1D as well as four additional glycopeptides initially undetectable by MALDI MS (Figure 1E and 1F). It is concluded that BEMAP is a specific and sensitive method for detecting protein glycosylation.

Results

The outer membrane protein fraction of H10407 was isolated and subjected to BEMAP analysis for identification of glycoproteins. This approach identified the proteins NmpC, TibA, EtpA, EatA, CexE and Ag43.

Discussion:

BEMAP relies on nucleophile tagging using 2-Aminoethyl phosphonic acid (AEP) rather than e.g. DTT. BEMAP method selectivity is achieved with the glycan-for-phosphate molecule exchange combined with a highly specific enrichment protocol for downstream sample processing (Thingholm et al., 2006). Importantly, the BEMAP chemistry can be applied in principle to any organism on a large-scale proteomics level irrespective of the chemical properties of the O-linked monosaccharide. As demonstrated in Figure 1, BEMAP replaces the carbohydrate moiety of a synthetic glycosylated peptide with a phosphotag in a chemical reaction exceeding 95% efficiency. Moreover, HCD MS/MS fragmentation of enriched BEMAP samples yields diagnostic ions instrumental for glycopeptide MS/MS spectrum identification as well as enabling unambiguous assignment of the modified amino acid residue, see Figure 2.

To identify specific pathogenic *E. coli* associated glycoproteins of potential therapeutic value we compared the outer membrane protein complement to non-pathogenic reference strain MG1655 sampled under identical conditions. By applying our BEMAP workflow we identified five ETEC vaccine candidates. Based on our analyses, we propose that novel vaccines directed against ETEC should not only be selected amongst the glycoproteins expressed by the pathogen but

can in principle also be targeting glycosylated domains of proteins which otherwise share 100% identity among *E. coli* strains.

Experimental procedures:

Lyophilized peptide sample is resuspended in 100 µl BEMAP solution consisting of 0.4 M 2-AEP (Sigma; 268674), 0.75 M NaOH (Sigma; S8045), 20 mM $Ba(OH)_2$ (Sigma; 433373) and incubate at 37°C in a heating block for 3.15 hours shaking at 1300 r.p.m. The reaction is stopped by acidification (1% TFA final concentration). Sample volume is increased to 1 ml and the peptides are purified on an Oasis HLB Plus short cartridge (Waters) as recommend by manufacturer and subsequently lyophilized. TiO_2 enrichment was performed as described by Tingholm et al., 2006.

ETEC H10407 lead molecules:

1) Putative outer membrane porin protein NmpC, ETEC_0806, CBJ00316

Primary sequence of NmpC (SEQ ID NO: 1):
MKKLTVAISAVAASVLMAMSAQAAEIYNKDSNKLDLYGKVNAKHYFSSNDADDGDTTYA
RLGFKGETQINDQLTGFGQWEYEFKGNRAESQGSSKDKTRLAFAGLKFGDYGSIDYGRN
YGVAYDIGAWTDVLPEFGGDTWTQTDVFMTQRATGVATYRNNDFFGLVDGLNFAAQYQ
GKNDRSDFDNYTEGNGDGFGFSATYEYEGFGIGATYAKSDRTDTQVNAGKVLPEVFASG
KNAEVWAAGLKYDANNIYLATTYSETQNMTVFADHFVANKAQNFEAVAQYQFDFGLRPS
VAYLQSKGKDLGVWGDQDLVKYVDVGATYYFNKNMSTFVDYKINLLDKNDFTKALGVST
DDIVAVGLVYQF

Unique to H10407 compared to other *E. coli*

Identified Tryptic glycopeptides using BEMAP combined with ESI-MS/MS (SEQ ID NOs: 2-8, respectively):

Start	end	Seq	ModAA#1	ModAA#2	ModAA#3	ModAA#4

44	60	HYFssNDADDGDttYAR		S47	S48	T56
T57						
107	117	FGDYGsIDYGR		S112		
151	158	AtGVAtYR		T152	T156	
215	226	sDRtDtQVNAGK		S215	T218	T220
227	236	VLPEVFAsGK	S234			
328	336	NMsTFVDYK	S330			
337	347	INLLDKNDFtK		T346		

Identified NmpC glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylations underlined, SEQ ID NOs: 9-19, respectively)):

NAKHYFSSNDADDG
DADDGDTTYARLGF
KFGDYGSIDYGRN
FMTQRATGVATYR
RATGVATYRNNDF
GATYAKSDRTDTQ
YAKSDRTDTQVNA
KSDRTDTQVNAGK
LPEVFASGKNAEV
YFNKNMSTFVDYK
LDKNDFTKALGVS

2) Adhesin/invasin TibA autotransporter, CBJ01643 or ETEC_2141 (SEQ ID NO: 20)

Primary sequence of TibA:

MNKVYNTVWNESTGTWVVTSELTRKGGLRPRQIKRTVLAGLIAGLLMPSMPALAAAYDN
QTIGRGETSKSMHLSAGDTAKNTTINSGGKQYVSSGGSATSTTINIGGVQHVSSGGSAT
SSTINSGGHQHVSSGGSATNTTVNNGGRQTVFSGGSAMGTIINSGGDQYVISGGSATS
ASVTSGARQFVSSGGIVKATSVNSGGRQYVRDGGSATDTVLNNTGRQFVSSGGSAAKT
TINSGGGMYLYGGSATGTSIYNGGRQYVSSGGSATNTTVYSGGRQHVYIDGNVTETTIT

SGGMLQVEAGGSASKVIQNSGGAVITNTSAAVSGTNDNGSFSIAGGSAVNMLLENGGY
LTVFDGHQASDTMVGSDGTLDVRSGGVLYGTTTLTDKGALVGDVVTNEGNLYYLNNSTA
TFTGTLTGTGTLTQEGGNTRFSGLLSQDGGIFLQSGGAMTMDALQAKANVTTQSGTTLT
LDNGTILTGNVAGDSTGAGDMAVKGASVWHLDGDSTVGALTLDNGTVDFRPSTTTRMT
PAFQAVSLALGSLSGSGTFQMNTDIASHTGDMLNVAGNASGNFVLDIKNTGLEPVSAGA
PLQVVQTGGGDAAFTLKGGKVDAGTWEYGLSKENTNWYLKADTPPPVTPPTNPDADNP
DAGNPDAGNPDAGNPDAGNPDAGKPGTGKPDAGTSSSPVRRTTKSVDAVLGMATAPAY
VFNSELDNLRFRHGDVMQNTRAPGGVWGRYTGSDNRISGGASSGYTLTQNGFETGAD
MVFDLSDSSLAVGTFFSYSDNSIKHARGGKSNVDSSGGGLYATWFDNDGYYVDGVLKY
NRFNNELRTWMSDGTAVKGDYSQNGFGGSLEAGRTFSLNENAWAQPYVRTTAFRADKK
EIRLNNGMKASIGATKSLQAEAGLKLGMTLDVAGKEVKPYLSAAVSHEFSDNNKVRINDT
YDFRNDISGTTGKYGLGVNAQLTPNAGVWAEARYENGKQTESPITGGVGFRINF

Identified Tryptic glycopeptides using BEMAP combined with ESI-MS/MS (SEQ ID NOs: 21-25, respectively):

Start	end	Seq	Mod AA #1	Mod AA #2	
70	80	SMHLsAGDTAK	S74		
81	89	NTTINsGGK	S86		
185	194	QFVssGGIVK	S188	S189	
195	203	ATSVNsGGR	S200		
223	233	QFVssGGSAAK	S226		S227

Identified TibA glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylations underlined, SEQ ID NOs: 26-30, respectively):

SKSMHL<u>S</u>AGDTAK
KNTTIN<u>S</u>GGKQYV
GARQFV<u>SS</u>GGIVKA
KATSVN<u>S</u>GGRQYV
TGRQFV<u>SS</u>GGSAAK

ETEC H10407 lead molecules:

3) Two-partner secreted adhesin EtpA, CBJ04458, ETEC_p948_0110

Primary sequence of EtpA (SEQ ID NO: 31):

MNRIYKLKFDKRRNELVVVSEITTGVGNAKATGSVEGEKSPRRGVRAMALSLLSGMMIM
AHPAMSANLPTGGQIVAGSGSIQTPSGNQMNIHQNSQNMVANWNSFDIGKGNTVQFD
QPSSSAVALNRVVGGGESQIMGNLKANGQVFLVNPNGVLFGEGASVSTSGFVASTRDIK
NDDFMNRRYTFSGGQKAGAAIVNQGELTTNAGGYIVLAADRVSNSGTIRTPGGKTVLAA
SERITLQLDNGGLMSVQVTGDVVNALVENRGLVSARDGQVYLTALGRGMLMNTVLNVS
GVVEASGMHRQDGNIVLDGGDSGVVHLSGTLQADNASGQGGKVVVQGKNILLDKGSN
ITATGGQGGGEVYVGGGWQGKDSNIRNADKVVMQGGARIDVSATQQGNGGTAVLWS
DSYTNFHGQISAKGGETGGNGGRVETSSHGNLQAFGTVSASAKKGKAGNWLLDSADIT
IVNGSNVSKTETTQSPPHTQFAPTAAGSAVSNTSINNRLNNGTSVTILTHRTRTGTAQGG
NITVNAAINKSNGSDVNLTLQAGGNITVNNSITSTEGKLNVNLSGARTSNGSITISNNAN
ITTNGGDITVGTTNTSNRVNISINNTTLNASNGNIQLTGTGTDSGILFAGNNRLTASNIAL
TGNSTSGNAINLTGTATLNATNNITLTGSSTSGNAINLKGNNTLTASNITLTGESTSGNAI
NLTDTTGTTTLNATNNITMQGTRVQIKHSNITAGNFALNATVAGSEISNTTLTATNNINLA
AKTNSASSGVYLKDARITSTNGSITANGTATANGKATHLDGNVTLNASNGRIKLTGNGH
GSASGILFAGNNRLTASNIALTGNSTSGNAINLTGTATLNATNDITLTGSSTSGNAINLTG
TATLNATNNITLTGSSTSGNAINLKGNNTLTASNITLTGESTSGNAINLTDTTGTTTLNAT
NNITMQGTRVQIKHSNITAGNFALNATVAGSEISNTTLTATNNINLAAKTNSASSGVYLK
DARITSTNGSITANGTATANGKATHLDGNVTLNASNGRIKLTGNGHGSASGILFAGNNR
LTASNIALTGNSTSGNAINLTGTATLNATNDITLTGSSTSGNAINLTGTATLNATNNITLTG
SSTSGNAINLKGNNTLTASNITLTGESTSGNAINLTDTTGTTTLNATNNITMQGTRVQIKH
SNITAGNFALNATVAGSEISNTTLTATNNINLAAKTNSASSGVYLKDARITSTNGSITANG
TATANGKATHLDGNVTLNASNGRIKLTGNGHGSASGILFAGNNRLTASNIALTGNSTSG
NAINLTGTATLNATNDITLTGSSTSGNAINLTGTATLNATNNITLTGSSTSGNAINLKGNN
TLTASNITLTGESTSGNAINLTDTTGTTTLNATNNITMQGTRVQIKHSNITAGNFALNATV
AGSEISNTTLTATNNINLAAKTNSASSGVYLKDARITSTNGSITANGTAPANDNATYLDG
NVTLNASNGSIKLTGNGNGSTSGILFAGNNTLTASNITLTGNSEVYWQ

Identified Tryptic glycopeptides using BEMAP combined with ESI-MS/MS (SEQ ID NOs: 32-34), respectively):

Seq	Mod AA #1	Mod AA #2	Mod AA #3

GNTVQFDQPsssAVALNR	S119	S120	S121	
LtGNGHGsAsGILFAGNNR	T821	S827	S829	
VsNsGTIR	S218	S220		

Identified TibA glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylations underlined, SEQ ID NOs: 35-38, respectively):

VQFDQPSSSAVALNR
NGRIKLTGNGHGS
TGNGHGSASGILF
NGHGSASGILFAG

4) Secreted autotransporter protein EatA, CBJ04449, ETEC_p948_0020

Primary sequence of EatA (SEQ ID NO: 39):
MNKVFSLKYSFLAKGFIAVSELARRVSVKGKLKSASSIIISPITIAIVSYAPPSLAATVNADI
SYQTFRDFAENKGAFIVGASNINIYDKNGVLVGVLDKAPMPDFSSATMNTGTLPPGDHTL
YSPQYVVTAKHVNGSDIMSFGHIQNNYTVVGENNHNSLDIKIRRLNKIVTEVAPAEISSV
GAVNGAYQEGGRFKAFYRLGGGLQYIKDKNGNLTPVYTNGGFLTGGTISALSSYNNGQM
ITAPTGDIFNPANGPLANYLNKGDSGSPLFAYDSLDKKWVLVGVLSSGSEHGNNWVVTT
QDFLHQQPKHDFDKTISYDSEKGSLQWRYNKNSGVGTLSQESVVWDMHGKKGGDLNA
GKNLQFTGNNGEIILHDSIDQGAGYLQFFDNYTVTSLTDQTWTGGGIITEKGVNVLWQV
NGVNDDNLHKVGEGTLTVNGKGVNNGGLKVGDGTVILNQRPDDNGHKQAFSSINISSG
RATVILSDANQVNPDKISWGYRGGTLDLNGNNVNFTRLQAADYGAIVSNNNKNKSELTL
KLQTLNENDISVDVKTYEVFGGHGSPGDLYYVPASNTYFILKSKAYGPFFSDLDNTNVWQ
NVGHDRDKAIQIVKQQKIGESSQPYMFHGQLNGYMDVNIHPLSGKDVLTLDGSVNLPEG
VITKKSGTLIFQGHPVIHAGMTTSAGQSDWENRQFTMDKLRLDAATFHLSRNAHMQGDI
SAANGSTVILGSSRVFTDKNDGTGNAVSSVEGSSIATTAGDQSYYSGNVLLENHSSLEV
RENFTGGIEAYDSSVSVTSQNAIFDHVGSFVNSSLLLEKGAKLTAQSGIFTNNTMKIKEN
ASLTLTGIPSVGKPGYYSPVTSTTEGIHLGERASLSVKNMGYLSSNITAENSAAIINLGDS
NATIGKTDSPLFSTLMRGYNAVLQGNIMGPQSSVNMNNALWHSDRNSELKELKANDSQI
ELGVRGHFAKLRVKELIASNSVFLVHANNSQADQLNVTDKLQGSNNTILVDFFNKAANG
TNVTLITAPKGSDENTFKAGTQQIGFSNITPEIRTENTDTATQWVLTGYQSVADARASKIA
TDFMDSGYKSFLTEVNNLNKRMGDLRDSQGDAGGWARIMNGTGSGESGYRDNYTHVQ

IGADRKHELNGIDLFTGALLTYTDNNASSQAFSGKTKSLGGGVYASGLFESGAYFDLIGK
YLHHDNRYTLNFASLGERSYTSHSLYAGAEIGYRYHMSENTWVEPQMELVYGSVSGKSF
NWKDQGMQLSMKDKDYHPLIGRTGVDVGRAFSGDTWKVTVRAGLGYQFDLLANGETV
LQDASGKKHFKGEKDSRMLMNVGTNVEVKDNMRFGLELEKSAFGRYNIDNSINANFRYY
F

(SEQ ID NOs: 40-43, respectively)	Mod AA #1	Mod AA #2	Mod AA #3
GGtLDLNGNNVNFTR	T500		
LQAADYGAIVsNNNK	S523		
VGEGTLtVNGK	T434		
AANGtNVTLITAPK	T1011		

Identified EatA glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylations underlined, SEQ ID NOs: 44-47, respectively):

WGYRGGTLDLNGN
DYGAIVSNNNKNK
VGEGTLTVNGKGV
NKAANGTNVTLIT

5) CexE, CBJ04537

Primary sequence of CexE (SEQ ID NO: 48):
MKKYILGVILAMGSLSAIAGGGNSERPPSVAAGECVTFNSKLGEIGGYSWKYSNDACNE
TVAKGYAIGVAMHRTVNYEGGYSIQSSGIVKPGSDFIMKGGKTYKGHKKVSAGGDTPY
WYK

Seq(SEQ ID NO: 49)	Mod AA #1
YsNDACNETVAK	S53

Identified CexE glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylation underlined, SEQ ID NO: 50):

GYSWKYSNDACNE

6) Putative antigen Ag43, CBJ03971, ETEC_4462

Primary sequence of Ag43 (SEQ ID NO: 51):

MKGESVYARDVSLCPITLFCTCRYPLLWVPAFLFTLCKEKLMKRHLNTSYRLVWNHITGTL
VVASELARSRGKRTGVAVALSLATATSVPALAADSVVQAGETVSGGTLENHDNQIVFGTT
NGITISTGLEYGPDNEANTGGQWVQDGGTASNTTISSGGLQFVGAGGKATDTIINEGGG
QSLKGLALNTTLNGGEQWMHEGAIATGTVINDKGWQVVKPGAVATDTVVNTGAEGGPD
AENADTGQFVRGDAVRTTINKNGRQIVVATGVANTTVVYAGGDQTVHGYALDTTLNGG
NQYVHNGGTASDTVVNSDGWQIIKEGGLADFTTVNQKGKLQVNAGGTATNVTLKQGG
ALVTSTAATVLGSNRLGNFTVENGKADGVVLESGGRLDVLEGHSAWKTLVDDGGILAVS
AGGKATDVTMTSGGALIADSGATVEGTNASGKFSIDGISGQASGLLLENGGSFTVNAGG
QAGNTTVGHRGTLTLAAGGSLSGRTQLSKGASMVLNGDVVSTGDIVNAGEIHFDNQTT
QDAVLSRAVAKSNSPVTFHKLTTTNLTGQGGTINMRVSLDGSNASDQLVINGGQATGKT
WLAFTNVGNSNLGVATSGQGIRVVDAQNGATTEEGAFALSRPLQAGAFNYTLNRDSDE
DWYLRSENAYRAEVPLYTSMLTQAMDYDRILAGSRSHQTGVNGENNSVRLSIQGGHLG
HDNNGGIARGATPESSGSYGFVRLEGDLLRTEVAGMSVTAGVYGAAGHSSVDVKDDDG
SRAGTVRDDAGSLGGYLNLVHTSSGLWADIVAQGTRHSMKASTDNNDFRARGWGWLG
SLETGLPFSITDNLMLEPQLQYTWQGLSLDDGKDNAGYVKFGHGSAQHVRAGFRLGSH
NDMTFGEGTSSRAPLRDSAKHSMRELPVNWWVQPSVIRTFSSRGDMSMGTAAAGSNM
TFSPSRNGTSLDLQAGLEARVRENITLGVQAGYAHSVIGSSAEGYNGQATLNVTF

Start	End	Sequence ID Nos 52-57, respectively	Mod AA#		
228	240	NGGVAGNTTVNQK	T235		
803	811	AstDNNDFR	S804	T805	
379	378	LGNFtVEnGK	T373		
682	695	sHQtGVnGENNsVR	S682	T685	S693
566	587	VsLDGsNAsDQLVInGGQAtGK		S567	S571 S574
	T585				
611	642	VVDAQNGAttEEGAFALSRPLQAGAFnYTLNR		T619	T620

Identified Ag43 glycopeptides listed as probable epitopes presented by Antigen Presenting Cells (glycosylations underlined, SEQ ID NOs: 58-68, respectively):

GGLADFTTVNQKG
RHSMKASTDNNDFR
NRLGNFTVENGKA
ILAGSRSHQTGVN
GSRSHQTGVNGEN
VNGENNSVRLSIQ
TINMRVSLDGSNA
RVSLDGSNASDQL
LDGSNASDQLVIN
INGGQATGKTWLA
DAQNGATTEEGAFA

Example 2 – Immunogenicity of ETEC glycosylated proteins

An immunogenic polypeptide is defined as a polypeptide that induces an immune response.

The immune response may be monitored by one of the following methods: An *in vitro* cellular response is determined by release of a relevant cytokine such as IFN-γ, from lymphocytes withdrawn from an animal or human currently or previously infected with ETEC, or by detection of proliferation of these T cells. The induction is performed by addition of the polypeptide or the immunogenic part to a suspension comprising from $1x10^5$ cells to $3x10^5$ cells per well. The cells are isolated from either blood, the spleen, the liver or the lung and the addition of the polypeptide or the immunogenic part of the polypeptide result in a concentration of not more than 20 µg per ml suspension and the stimulation is performed from two to five days. For monitoring cell proliferation the cells are pulsed with radioactive labeled Thymidine and after 16-22 hours of incubation the proliferation is detected by liquid scintillation counting. A positive response is a response more than background plus two standard deviations. The release of IFN- γ can be determined by the ELISA method, which is well known to a person skilled in the art. A positive response is a response more than background plus two standard deviations. Other cytokines than IFN- γ could be relevant when monitoring an immunological response to the polypeptide, such as IL-12, TNF-α, IL-4, IL-5, IL-10, IL-6, TGF-β.

Another and more sensitive method for determining the presence of a cytokine (e.g. IFN-γ) is the ELISPOT method where the cells isolated from either the blood, the spleen, the liver or the lung are diluted to a concentration of preferable of 1 to 4 x 10^6 cells /ml and incubated for 18-22 hrs in the presence of the polypeptide or the immunogenic part of the polypeptide resulting in a concentration of not more than 20 µg per ml.

The cell suspensions are hereafter diluted to 1 to 2 x 10^6/ ml and transferred to Maxisorp plates coated with anti–IFN-γ and incubated for preferably 4 to 16 hours. The IFN-γ producing cells are determined by the use of labelled secondary anti-IFN-antibody and a relevant substrate giving rise to spots, which can be enumerated using a dissection microscope. It is also a possibility to determine the presence of mRNA coding for the relevant cytokine by the use of the PCR technique. Usually one or more cytokines will be measured utilizing for example the PCR, ELISPOT or ELISA. It will be appreciated by a person skilled in the art that a significant increase or decrease in the amount of any of these cytokines induced by a specific polypeptide can be used in evaluation of the immunological activity of the polypeptide.

An *in vitro* cellular response may also be determined by the use of T cell lines derived from an immune individual or an ETEC infected person where the T cell lines have been driven with either live ETEC, extracts from the bacterial cell or culture filtrate for 10 to 20 days with the addition of IL-2. The induction is performed by addition of not more than 20 µg polypeptide per ml suspension to the T cell lines containing from 1x10^5 cells to 3x10^5 cells per well and incubation is performed from two to six days. The induction of IFN-γ or release of another relevant cytokine is detected by ELISA. The stimulation of T cells can also be monitored by detecting cell proliferation using radioactively labeled Thymidine as described above. For both assays a positive response is a response more than background plus two standard deviations.

An *in vivo* cellular response may be determined as a positive DTH response after intradermal injection or local application patch of at most 100 µg of the polypeptide or the immunogenic part to an individual who is clinically or

subclinically infected with ETEC, a positive response having a diameter of at least 5 mm 72-96 hours after the injection or application.

An *in vitro* humoral response is determined by a specific antibody response in an immune or infected individual. The presence of antibodies may be determined by an ELISA technique or a Western blot where the polypeptide or the immunogenic part is absorbed to either a nitrocellulose membrane or a polystyrene surface. The serum is preferably diluted in PBS from 1:10 to 1:100 and added to the absorbed polypeptide and the incubation being performed from 1 to 12 hours. By the use of labeled secondary antibodies the presence of specific antibodies can be determined by measuring the presence or absence of a specific label e.g. by ELISA where a positive response is a response of more than background plus two standard deviations or alternatively a visual response in a Western blot.

Another relevant parameter is measurement of the protection in animal models induced after vaccination with the polypeptide in an adjuvant or after DNA vaccination. Suitable animal models include primates, guinea pigs or mice, which are challenged with an infection of an ETEC. Readout for induced protection could be decrease of the bacterial load in target organs compared to non-vaccinated animals, prolonged survival times compared to non-vaccinated animals and diminished weight loss or pathology compared to non-vaccinated animals.

The glycosylated polypeptides described herein are immunogenic when one of the above described tests is positive.

Example 3 - Schematic overview of assays and experiments used to characterize glycosylated as well as non-glycosylated NmpC or TibA protein properties. An "X" indicate lead molecule specific assay.

Experiment type	Candidate molecule	
	NmpC	TibA
Mouse challenge	X	X
Serum and mucosal antibody responses	X	X
Antibody mediated inhibition of ETEC binding to Caco-2	X	X
Antibody mediated inhibition of ETEC binding to Caco-2; cAMP release measurement	X	X
Surface-plasmon-resonance-based (SPR) to determine k_{off} and k_{on} constants	X	X
MIC assay to test antibiotics susceptibility	X	

NmpC and TibA have been tested for a wide variety of features using several different assays. An overview can be seen in the table above.

Example 4 - Vaccination with glycosylated NmpC or TibA affords better protection against intestinal colonization of ETEC in mice compared to the non-glycosylated protein versions.
Assay type: Mouse challenge studies

Materials and methods:
Five groups of CD-1 mice were immunized with either adjuvant only (control), or appropriate amount of adjuvant + e.g. 25 µg of glycosylated NmpC, or adjuvant + e.g. 25 µg of non-glycosylated NmpC, or adjuvant + e.g. 25 µg of glycosylated TibA, or adjuvant + e.g. 25 µg of non-glycosylated TibA, on days 0, 14, 28. On day 40, mice were treated with streptomycin [e.g. 5 g per liter] in drinking water for 24 hours, followed by drinking water alone for 18 hours. After administration of famotidine to reduce gastric acidity, mice were challenged with 10^6 cfu of a chloramphenicol-resistant ETEC strain by oral gavage. Fecal samples (6 pellets/mouse) were collected on day 42 before oral gavage, re-suspended in buffer (10mM Tris, 100mM NaCl, 0.05% Tween 20, 5mM Sodium Azide, pH 7.4) overnight at 4°C, centrifuged to pellet insoluble material, and recover supernatant for fecal antibody testing (below). Twenty-four hours after infection, mice were sacrificed, sera were collected, and dilutions of saponin small-intestinal lysates were plated onto Luria agar plates containing chloramphenicol (40 µg/ml).

Experimental outcome: As determined by CFU counting, fecal samples from mice immunized with glycosylated antigen NmpC and TibA contained fewer ETEC compared to fecal samples from mice immunized with non-glycosylated antigen versions.

Example 5 - Immunization with glycosylated antigen NmpC or TibA generates robust serum and mucosal antibody responses

Assay type: Kinetic enzyme-linked immunosorbent assay probing relative levels of IgA, IgM and IgG

Materials and methods:
Murine immune responses to adjuvant, glycosylated and non-glycosylated versions of NmpC and TibA were determined using kinetic ELISA. Briefly, ELISA wells were incubated at 4°C overnight with proteins at a final concentration of 4 µg/ml in 0.1 M $NaHCO_3$ buffer (pH 8.6), washed the following day with Tris-buffered saline containing 0.005% Tween 20 (TBS-T), and blocked with 1% bovine serum albumin (BSA) in TBS-T for 1 h at 37°C prior to the addition of the samples. Sera were diluted 1:500 in TBS-T with 1% BSA, and 100 µl was added to each ELISA well, followed by incubation at 37°C for 1 h. After three washes with TBS-T, horseradish peroxidase-conjugated secondary antibody (either goat anti-mouse IgA, IgM, or IgG) was added at a final dilution of 1:5,000, followed by incubation for an additional hour before washing and development with TMB (3,3′,5,5′-tetramethylbenzidine)-peroxidase substrate (KPL). Kinetic ELISA data are expressed as Vmax in milliunits/min.

Experimental outcome: Immunization with glycosylated antigen NmpC or TibA generates robust IgA, IgG and IgM antibody responses as compared to non-glycosylated versions.

Reference for experiment: EatA, an immunogenic protective antigen of enterotoxigenic Escherichia coli, degrades intestinal mucin.

Example 6 - Monoclonal antibodies raised against glycosylated NmpC and TibA inhibits ETEC binding to intestinal epithelial cells to a higher extent compared antibodies raised against non-glycosylated NmpC and TibA protein version.

Assay type: Adhesion Assay

Materials and methods:
In vitro, Caco-2 epithelial cell monolayers were infected with ETEC H10407 at multiplicities of infection of approximately 100 (bacteria/cell) approximately 18 to 22 days after seeding into wells. Cultures of bacteria were grown overnight in Luria broth from frozen glycerol stocks, diluted 1:100, and grown for 1 h. One microliter of bacterial culture is added to confluent Caco-2 monolayers seeded into 96-well plates preincubated with or without antibodies. Two hours after inoculation, the monolayers were washed 3 times with tissue culture medium after which bacteria were isolated, serial diluted and plated to count CFU the following day.

Experimental outcome: Monoclonal antibodies raised against glycosylated NmpC and TibA inhibits ETEC binding to intestinal epithelial cells to a higher extent compared monoclonal antibodies raised against non-glycosylated NmpC and TibA protein version.

Example 7 - Monoclonal antibodies raised against glycosylated NmpC and TibA inhibits ETEC binding to intestinal epithelial cells to a higher extent compared to monoclonal antibodies raised against non-glycosylated NmpC and TibA protein version.

Assay type: Adhesion Assay coupled to cAMP enzyme immunoassay

Materials and methods:
In vitro, Caco-2 epithelial cell monolayers were infected with ETEC H10407 at multiplicities of infection of approximately 100 (bacteria/cell). Cultures of bacteria were grown overnight in Luria broth from frozen glycerol stocks, diluted 1:100, and grown for 1 h. One microliter of bacterial culture is added to

confluent Caco-2 monolayers seeded into 96-well plates preincubated with or without antibodies. Two hours after inoculation, the monolayers were washed 3 times with tissue culture medium, and the medium was replaced with 100 µl of fresh medium/well and returned to the incubator (37°C, 5% CO2) for 2.5 h. Subsequently, cyclic AMP (cAMP) enzyme immunoassay (EIA) (Arbor Assays, Ann Arbor, MI) was used to examine the efficiency of toxin delivery.

Experimental outcome: Addition of antibodies raised against glycosylated NmpC and TibA results in lower levels of released cAMP into the growth medium compared to monoclonal antibodies raised against non-glycosylated NmpC and TibA protein version.

Example 8 - Monoclonal antibodies raised against NmpC and TibA glycoepitopes results in lower equilibrium dissociation constants ($K_d = k_{off} / k_{on}$) compared to monoclonal antibodies raised against non-glycosylated NmpC and TibA epitopes.

Assay type: Surface-plasmon-resonance-based (SPR) sensor (IBIS MX96)

Materials and methods:
NmpC and TibA were digested ON at 25C using an appropriate protease. Subsequently, peptides covering whole protein length or sequences known to be glycosylated were isolated using HPLC. Standard protocols were employed to create an array containing peptides (+/- glycosylated variants) as well as controls; human IgG, human serum albumin (HSA) and empty spots. After interaction with the antibodies the array was regenerated (cleaned) using an acid step to wash and elute the antibody and to prepare the sensor for the subsequent Ab sample. In the continuous flow microspotter (CFM), various peptides were arrayed onto this layer where each feature contained a small peptide bound at saturating concentrations. The CFM-arrayed sensor was then inserted in the IBIS MX96 and a series of antigen binding antibodies were then sequentially injected and passed over the complete sensor surface.

Experimental outcome: Glycosylation changes the binding and affinity properties of the antibody dramatically as shown for monoclonal antibodies targeting the same linear sequence where one is modified and the other is not.

Example 9 - Antibiotics susceptibility (Cephalothin and Polymyxin B) can be reduced with antibodies (polyclonal and monoclonal antibodies) targeting glycosylated NmpC epitopes

Assay type: MIC assay to test antibiotics susceptibility.

Materials and methods:
The MIC for a panel of antibiotics (e.g. Cephalothin and Polymyxin B) was determined by serial dilution in Mueller Hinton broth as recommended by the Clinical and Laboratory Standards Institute.

Experimental outcome: An nmpC isogenic knock-out strain is less susceptible to antibiotics compared to wild type H10407. By including antibodies which targets NmpC epitopes in the MIC assay, the uptake of antibiotics is reduced resulting in antibiotics susceptibility levels comparable to the nmpC isogenic knock-out strain.

Claims

1. A polypeptide comprising:

a) SEQ ID NO: 20,

b) a polypeptide fragment of SEQ ID NO: 20 that has at least 75 % sequence identity to SEQ ID NO: 20, or

c) a B- or T-cell epitope of SEQ ID NO: 20,

wherein the polypeptide is glycosylated at least in one position.

2. The polypeptide as defined in any of claim 1, wherein the polypeptide is glycosylated at least in two positions.

3. The polypeptide as defined in any of claims 1-2, wherein the polypeptide is glycosylated at least in three positions.

4. The polypeptide as defined in any of claims 1-3, wherein the polypeptide originates from Enterotoxigenic Escherichia coli (ETEC).

5. The polypeptide as defined in any of claims 1-4, wherein the glycosylated polypeptide is selected from the group of glycosylated polypeptides consisting of SEQ ID NOs: 21-30.

6. The polypeptide as defined in any of claims 1-5, wherein the polypeptide is a fusion polypeptide.

7. The polypeptide as defined in any of claims 1-6, wherein the polypeptide is fused to a polypeptide originating from ETEC.

8. An immunogenic composition comprising the polypeptide as defined in any of claims 1-7.

9. A pharmaceutical composition comprising the polypeptide as defined in any of claims 1-7 and at least one pharmaceutically acceptable carrier, excipient or diluent.

10. The immunogenic composition as defined in claim 8 or the pharmaceutical composition of claim 9, which is a vaccine against ETEC.

11. An immunogenic composition, a pharmaceutical composition, or a vaccine as defined in any of claims 8-11, which is formulated for intradermal, transdermal, subcutaneous, intramuscular or mucosal application.

12. A nucleic acid sequence encoding a polypeptide as defined in any of claims 1-7.

13. A method for immunizing a mammal, the method comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine according to any of the preceding claims.

14. A method for treating a mammal, which is infected with ETEC comprising administering to the mammal the immunogenic composition, pharmaceutical composition or vaccine according to any of the preceding claims.

15. The method as defined in any of claims or polypeptide for use as defined in any of claims 13-14, wherein the mammal is a human.

16. The polypeptide as defined in any of claims 1-7, immunogenic composition as defined in claim 8, pharmaceutical composition as defined in claim 9, or vaccine as defined in claim 10 for use in treating infection caused by ETEC.

17. The polypeptide as defined in any of claims 1-7, immunogenic composition as defined in claim 8, pharmaceutical composition as defined in claim 9, or vaccine as defined in claim 10 for use in the preparation of a medicament for treating infection caused by ETEC.

18. The polypeptide as defined in any of claims 1-7, immunogenic composition as defined in claim 8, pharmaceutical composition as defined in claim 9, or vaccine as defined in claim 10 for use in the diagnosis of an infection caused by ETEC.

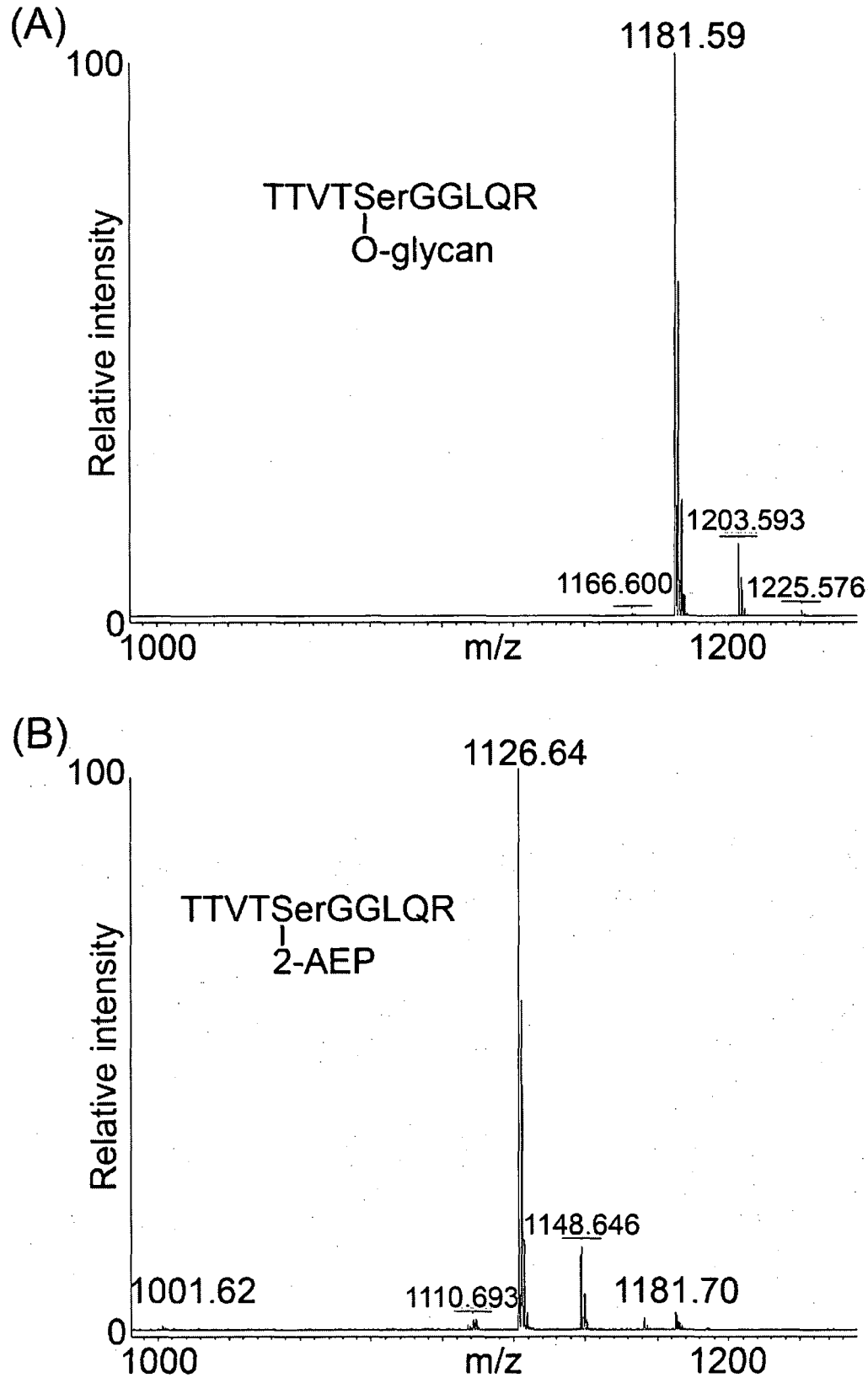
(A)
1181.59
100
TTVTSerGGLQR
O-glycan
Relative intensity
1203.593
1166.600
1225.576
0
1000
m/z
1200
(B)
1126.64
100
TTVTSerGGLQR
2-AEP
Relative intensity
1148.646
1001.62
1110.693
1181.70
0
1000
m/z
1200

Fig. 1 (a) (b)

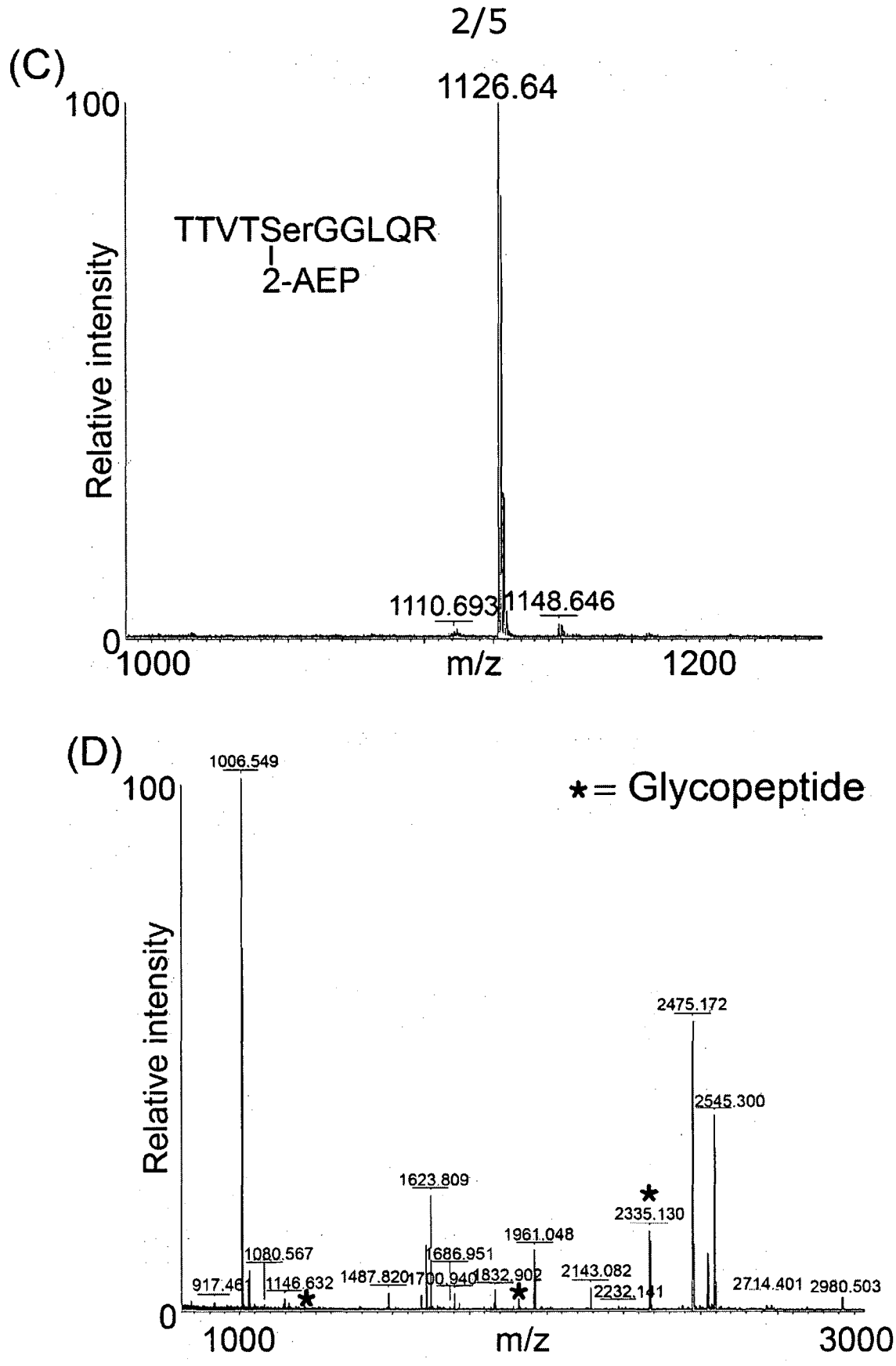
(C)
100
Relative intensity
0
1126.64
TTVTSerGGLQR
2-AEP
1110.693
1148.646
1000
m/z
1200
(D)
1006.549
100
Relative intensity
0
*= Glycopeptide
2475.172
2545.300
1623.809
2335.130
1961.048
1080.567
686.951
917.46
1146.632
1487.820
1700.940
1832.902
2143.082
2232.141
2714.401
2980.503
1000
m/z
3000

Fig. 1 (c)(d)

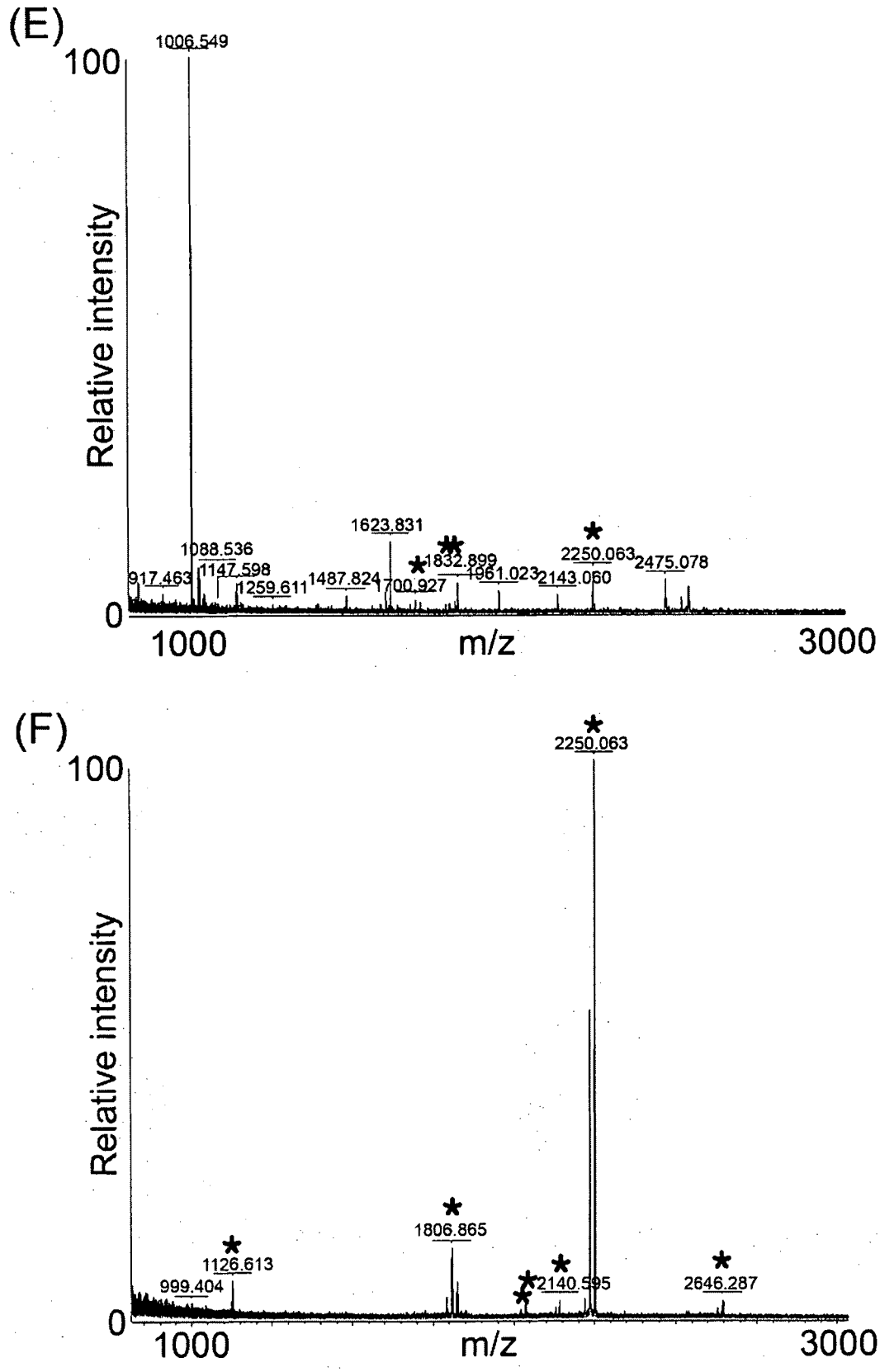
(E)
100
Relative intensity
0
1000
m/z
3000
1006.549
1623.831
1088.536
1147.598
1832.899
2250.063
2475.078
917.463
1259.611
1487.824
700.927
1961.023
2143.060
(F)
100
Relative intensity
0
1000
m/z
3000
2250.063
1806.865
1126.613
999.404
2140.595
2646.287

Fig. 1 (e)(f)

CID

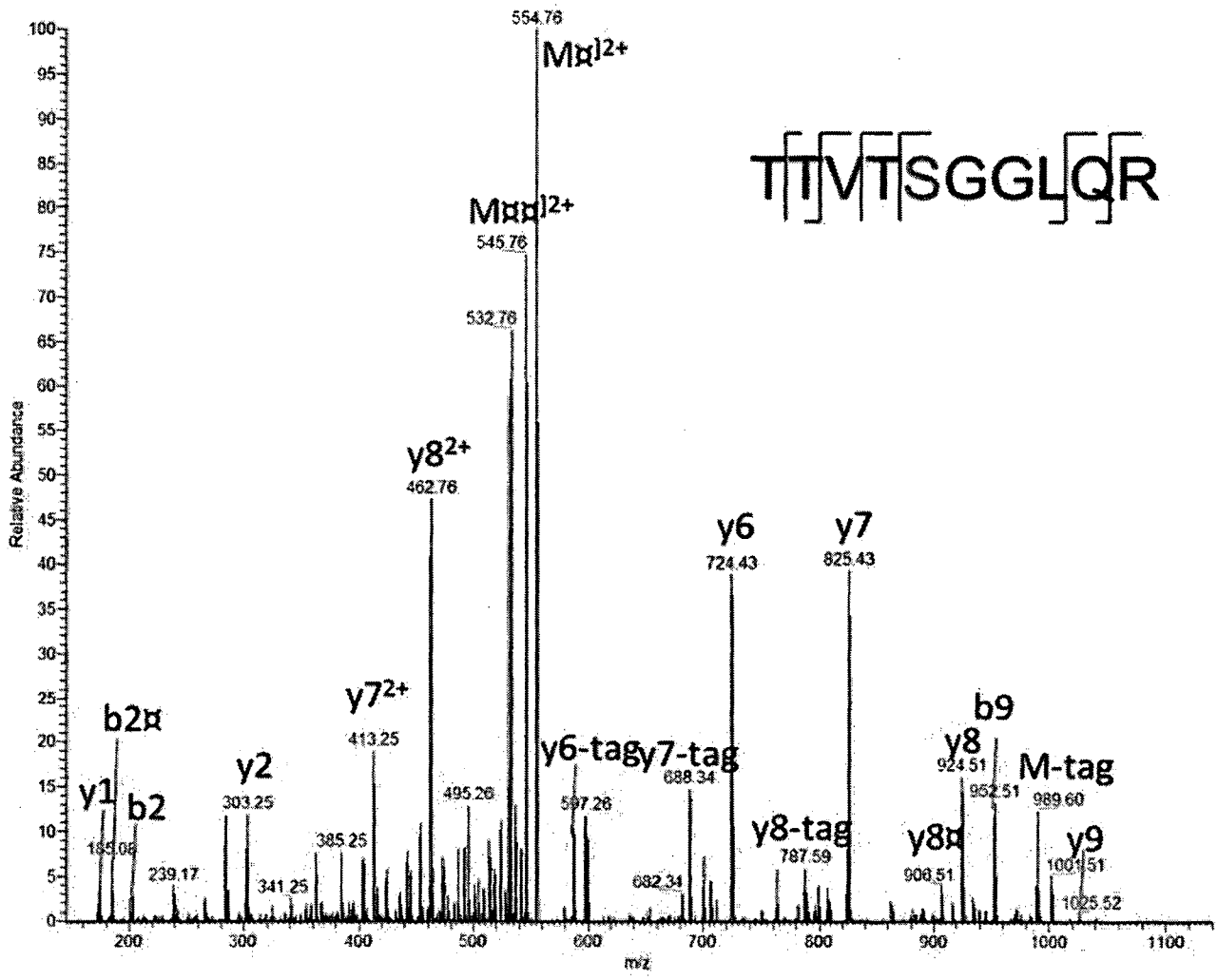
TTVTSGGLQR
Relative Abundance
m/z
554.76
M¤2+
M¤¤2+
545.76
532.76
y8 2+
462.76
y6
724.43
y7
825.43
y7 2+
413.25
b2¤
y2
303.25
y1
b2
y6-tag
y7-tag
688.34
b9
y8
924.51
M-tag
989.60
y8-tag
787.59
y8¤
906.51
y9
1001.51
1025.52
682.34
597.26
495.26
385.25
341.25
239.17

Fig. 2 (CID)

HCD

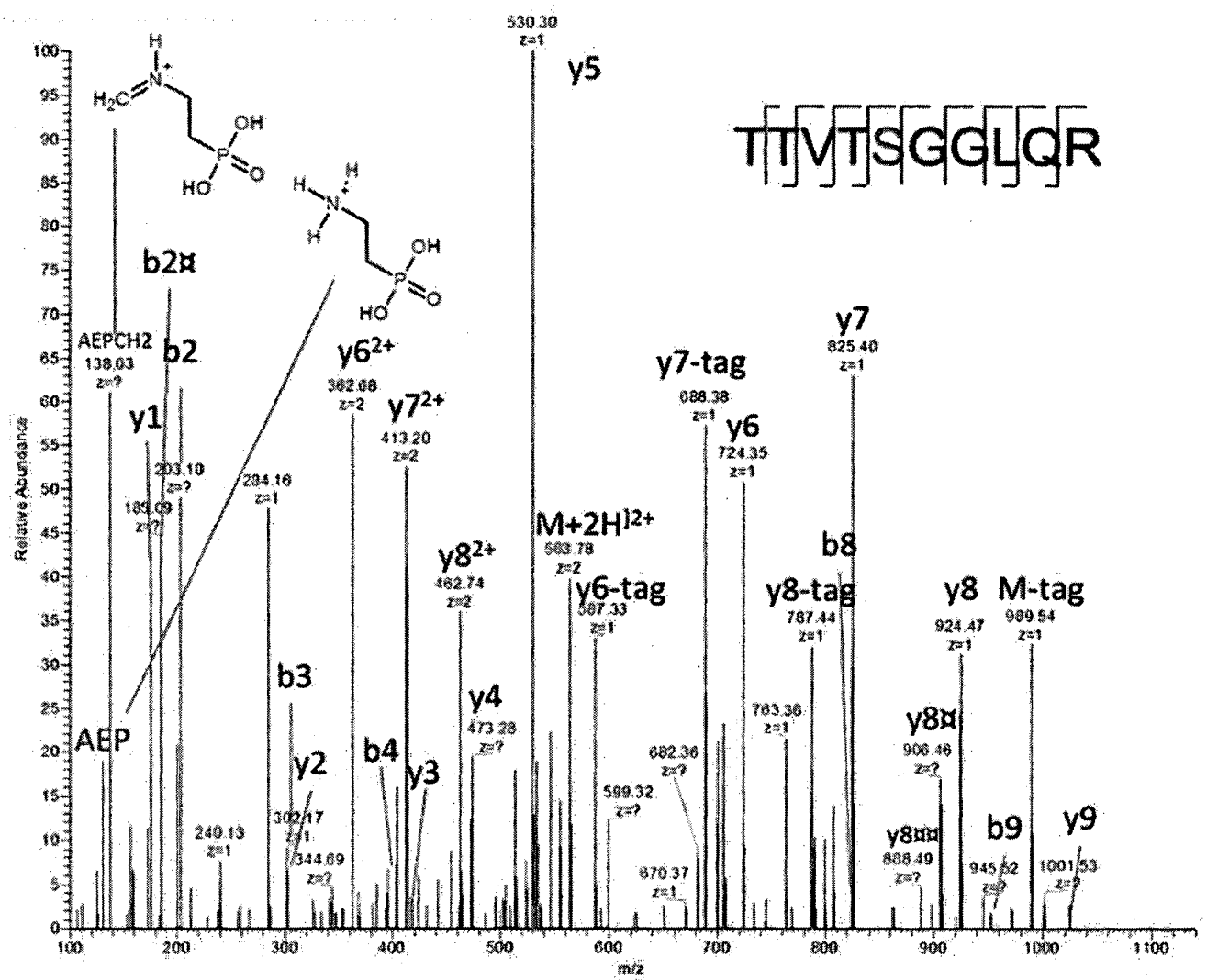
TTVTSGGLQR
y5
530.30
z=1
b2¤
AEPCH2
138.03
z=?
b2
y1
y6²⁺
382.68
z=2
y7²⁺
413.20
z=2
y7-tag
088.38
z=1
y7
825.40
z=1
y6
724.35
z=1
M+2H]2+
583.78
z=2
y8²⁺
462.74
z=2
y6-tag
587.33
z=1
b8
y8-tag
787.44
z=1
y8
924.47
z=1
M-tag
989.54
z=1
b3
y4
473.28
z=?
AEP
y2
b4
y3
y8¤
906.46
z=?
y8¤¤
888.49
z=?
b9
y9
1001.53
z=?
Relative Abundance
m/z

Fig. 2 (HCD)

INTERNATIONAL SEARCH REPORT

International application No
PCT/DK2015/050089

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/245 A61K39/108
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, Sequence Search, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt [Online] 1 November 1999 (1999-11-01), "RecName: Full=Adhesin/invasin TibA autotransporter; Contains: RecName: Full=Adhesin/invasin TibA; AltName: Full=Glycoprotein TibA; Contains: RecName: Full=Adhesin/invasin TibA translocator; Flags: Precursor;", XP002740838, retrieved from EBI accession no. UNIPROT:Q9XD84 Database accession no. Q9XD84 The query sequence SEQ ID NO:20 has 100.00 % identity (100.00 % similarity) over 989 positions in a common overlap (range (q:s): 1-989:1-989) with subject UNIPROT:Q9XD84 (length: 989).; the whole document -/--	1-4

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 June 2015	07/07/2015

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	van Heusden, Miranda

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	-& Christoph Lindenthal ET AL: "Identification of a Glycoprotein Produced by Enterotoxigenic Escherichia coli", Infection and Immunity, vol. 67, no. 8 1 August 1999 (1999-08-01), 1 August 1999 (1999-08-01), pages 4084-4091, XP055195000, UNITED STATES Retrieved from the Internet: URL:http://iai.asm.org/content/67/8/4084.abstract page 4085, column 2, paragraph 5 - page 4086, column 1, paragraph 2; figure 1; table 1 page 4090, column 2, paragraph 1 page 4089, column 1, paragraph 2	1-4,8,9, 12
X	C. LINDENTHAL ET AL: "Enterotoxigenic Escherichia coli TibA Glycoprotein Adheres to Human Intestine Epithelial Cells", INFECTION AND IMMUNITY, vol. 69, no. 1, 1 January 2001 (2001-01-01), pages 52-57, XP055195452, ISSN: 0019-9567, DOI: 10.1128/IAI.69.1.52-57.2001 page 54, column 1, paragraph 2 - column 2, paragraph 2 page 53, column 2, paragraph 3-4; figure 1c	1-4,6-9, 11-13
X	K. ROY ET AL: "Enterotoxigenic Escherichia coli Elicits Immune Responses to Multiple Surface Proteins", INFECTION AND IMMUNITY, vol. 78, no. 7, 10 May 2010 (2010-05-10), pages 3027-3035, XP055195400, ISSN: 0019-9567, DOI: 10.1128/IAI.00264-10 page 3029, column 2, paragraph 6; table 2	18
A	HARRIS J A ET AL: "Directed evaluation of enterotoxigenic escherichia coli autotransporter proteins as putative vaccine candidates", PLOS NEGLECTED TROPICAL DISEASES, PUBLIC LIBRARY OF SCIENCE, US, vol. 5, no. 12, 6 December 2011 (2011-12-06), pages e1428-1, XP008151760, ISSN: 1935-2727, DOI: 10.1371/JOURNAL.PNTD.0001428 page 7; figure 4e page 8, column 1, paragraph 3	1-18

-/--

2

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ROY K ET AL: "Vaccination with EtpA glycoprotein or flagellin protects against colonization with enterotoxigenic Escherichia coli in a murine model", VACCINE, ELSEVIER LTD, GB, vol. 27, no. 34, 23 July 2009 (2009-07-23) , pages 4601-4608, XP026266826, ISSN: 0264-410X, DOI: 10.1016/J.VACCINE.2009.05.076 [retrieved on 2009-06-11] the whole document -----	1-18
A	RAJ K. UPRETI ET AL: "Bacterial glycoproteins: Functions, biosynthesis and applications", PROTEOMICS, vol. 3, no. 4, 1 April 2003 (2003-04-01), pages 363-379, XP055007473, ISSN: 1615-9853, DOI: 10.1002/pmic.200390052 page 375, column 2, paragraph 2 page 376, column 2, paragraph 2 -----	1-18
X,P	LU QIUHE ET AL: "An Iron-Containing Dodecameric Heptosyltransferase Family Modifies Bacterial Autotransporters in Pathogenesis", CELL HOST & MICROBE, ELSEVIER, NL, vol. 16, no. 3, 11 September 2014 (2014-09-11), pages 351-363, XP029054253, ISSN: 1931-3128, DOI: 10.1016/J.CHOM.2014.08.008 the whole document -----	1-4,8,9, 12

2