

Article

Not peer-reviewed version

Identification of Three Novel O-Linked Glycans in the Envelope Protein of Tick-Borne Encephalitis Virus

Ebba Könighofer , Ekaterina Mirgorodskaya , [Kristina Nyström](#) , [Karin Stiasny](#) , Ambjörn Kärmander , [Tomas Bergström](#) , [Rickard Nordén](#) *

Posted Date: 11 October 2024

doi: 10.20944/preprints202410.0914.v1

Keywords: E protein; N-linked glycan; O-linked glycan; tick-borne encephalitis virus



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Identification of Three Novel O-Linked Glycans in the Envelope Protein of Tick-Borne Encephalitis Virus

Ebba Könighofer ¹, Ekaterina Mirgorodskaya ², Kristina Nyström ¹, Karin Stiasny ³,
Ambjörn Kärmander ¹, Tomas Bergström ¹ and Rickard Nordén ^{1,4,*}

¹ Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Sweden

² Proteomics Core Facility, Sahlgrenska Academy, University of Gothenburg, Sweden

³ Center for Virology, Medical University of Vienna, Vienna, Austria.

⁴ Department of Clinical Microbiology, Sahlgrenska University Hospital, Gothenburg, Västra Götalandsregionen, Sweden

* Correspondence: rickard.norden@microbio.gu.se

Abstract: Tick-borne encephalitis virus is a pathogen endemic to northern Europe and Asia, transmitted through bites from infected ticks. It is a member of the Flaviviridae family, possesses a positive-sense, single-stranded RNA genome encoding a polypeptide that is processed into seven non-structural and three structural proteins, including the envelope (E) protein. The glycosylation of the E protein, a single N-linked glycan at position N154, plays a critical role in viral infectivity and pathogenesis. Here, we dissected the entire glycosylation profile of the E protein using liquid chromatography tandem mass spectrometry and identified three novel O-linked glycans which were found at relatively low frequency. One of the O-linked glycans was positioned close to the highly conserved N-linked glycan site and structural analysis indicated that it may be of relevance for viral maturation. The N154 site was found to be glycosylated with a high frequency, containing oligomannose or complex-type structures, some of which are fucosylated. An unusually high portion of oligomannose N-linked glycan structures exhibited compositions that are normally observed on proteins while they are translocated from the endoplasmic reticulum to the trans golgi network, suggesting disruption of the glycan processing pathway in the infected cells from which the E protein was obtained.

Keywords: E protein; N-linked glycan; O-linked glycan; tick-borne encephalitis virus

1. Introduction

The Tick-borne Encephalitis Virus (TBEV) is a pathogen found in large parts of northern Europe and Asia. It is transmitted via bites from infected ticks and can cause systemic and neurological infection with mortality rates ranging from >2 % to 34 % depending on subtype [1–3]. Available vaccines protect against severe disease. However, booster doses are recommended every three to five years for healthy individuals below 50 years of age, while elderly and immunosuppressed individuals require a more extended vaccine program. Rare breakthrough infections and vaccine failures occur despite adherence to the vaccination schedule [4–6].

TBEV belongs to the family *Flaviviridae*, which consists of small, enveloped viruses with positive-sense, single-stranded RNA genomes. The TBEV genome translates to a polypeptide that is further processed into seven non-structural and three structural proteins: the capsid (C) protein, the membrane (M) protein and the envelope (E) protein [7,8]. The E protein consists of three domains: domain I (ED I), which stabilizes the protein and contain the E 150-loop which is important for viral particle maturation [9,10]; domain II (ED II), which contains a conserved fusion loop essential for membrane fusion [11]; and domain III (ED III), which has an immunoglobulin-like fold and serves as a structural target for neutralizing antibodies [12]. Three linear B-cell epitopes have been identified on the E protein: one located in ED I (aa 162-179) and two in ED II (aa 51-58 and 222-239) [13].

Structural antibody epitopes, comprising discontinuous segments of the E protein or comprising residues from nearby E proteins in the quaternary structure, are also present [14–16].

Starting at N154 (TBEV numbering), ED I contains one conserved consensus site (N-X-T/S (X≠P)) for N-linked glycosylation. The N-linked glycan at N154 has been connected to important functions in viral entry, infectivity, and pathogenesis of TBEV, as well as in other flaviviruses [17–21]. Despite its importance, few studies have determined the structure of the N-linked glycan [22]. The E protein also contains 67 sites for potential O-linked glycosylation, but to our knowledge, there are no data showing O-linked glycosylation. Consequently, we mapped the glycans of the E protein from a clinical isolate of TBEV grown in human adenocarcinomic alveolar basal epithelial (A549) cells. We identified three novel O-linked glycans on the E protein and determined the composition of the N-linked glycan structures at position N154.

2. Materials and Methods

2.1. Cells and Virus

The human adenocarcinomic cell line A549 (A549-CCL-185) was obtained from ATCC and cultured in 150 cm² cell culture flasks containing 28 mL Eagle's Minimal Essential Medium (Thermo Fisher Scientific, Waltham, USA) supplemented with 8 % fetal calf serum and 1 % penicillin-streptomycin until 100 % confluence. 1 mL aliquot of virus stock prepared in A549-cells (TBEV strain F7203, a clinical strain originally isolated from the blood of a Swedish TBE patient and classified as a Western European Subtype [23], TCID₅₀ 10⁶/mL) was added to the cells, and the infection was allowed to proceed for 48-72 hours. Extracellular viral particles were harvested by removing the cell culture media from the infected cells after which the virus containing cell culture media was concentrated using density gradient centrifugation. First, cellular debris was removed by centrifugation at 5000 × g for 10 minutes at 4 °C. The supernatant was transferred to QuickSeal Ultracentrifuge Tubes (Beckman Coulter, Indianapolis, USA) and centrifuged at 100,000 × g for 1 hour at 4 °C. Proteins in the supernatant were concentrated using the Macrosep Centrifugal Device with a 10K MWCO (Pall, Port Washington, USA). The samples were loaded twice onto the centrifugal device and centrifuged at 4200 × g for 30-60 minutes at 4 °C. The membrane was washed with Tris-buffered saline (TBS), and the retentate was collected and transferred to Eppendorf tubes for storage at -20 °C prior to E protein purification using immunoaffinity columns.

Intracellular viral particles and cells were lysed in 8 mL lysis buffer (Tris phosphate-buffered saline (TBS) supplemented with 2 % 3 α ,12 α -dihydroxycholelic acid sodium salt (DOC), 2 % Nonidet P-40, and 1 mM Pefabloc® SC (4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride, AEBSF)), followed by douncing in 12 mL glass douncer on ice for 1 hour before purification of the E protein using immunoaffinity columns.

2.2. Purification of E Protein Using Immunoaffinity Column

Coupling of monoclonal antibodies (mAb B2 targeting TBEV E protein DIII) [14] to CNBr-sepharose was done by swelling of 1 g sepharose (GE Healthcare, Chicago, USA) in 1mM HCl for 15 minutes during vacuum suction in a Büchner funnel, followed by rinsing with coupling buffer (0.1 M NaHCO₃, 0.5 M NaCl, pH 8). 0.5 mL sepharose was transferred to a silicon tube and 0.9 mg monoclonal anti protein E antibody was added to the tube and allowed to bind during end-to-end rotation for 2 hours at room temperature. The tube was centrifuged 5 minutes, 2000 × g at room temperature and the supernatant removed. Blocking buffer (0.2 M glycine, pH 8) was added in a volume equal to the gel and the tube again incubated 2 hours during end-to-end rotation. The centrifugation was repeated, and the supernatant removed followed by interchangeably washing and centrifugation with acetate buffer (0.1 M CH₃COONa, 0.5 M NaCl, pH 4) and coupling buffer 5 times. The final wash was performed with coupling buffer. The sepharose-coupled antibodies were stored at 4 °C in 20 % EtOH in phosphate-buffered saline (PBS) until use. The sepharose-coupled antibody column was washed with 20 mL wash buffer with detergent (0.01 M Tris-HCl, 0.5 mM NaCl, 0.1 % Nonidet P-40). The solution containing solubilized extracellular and intracellular viral particles was

applied to the column and allowed to flow through. To maximize E protein binding the flow through was allowed to pass through the column once more before washing the column with 20 mL wash buffer with detergent followed with washing with 20 mL wash buffer without detergent (0.01 M Tris-HCl, 0.5 mM NaCl). The E protein captured by the column was eluted by addition of elution buffer (0.1 M glycine-HCl, pH 2.6) in 1 mL fractions. 100 μ L neutralization buffer (1M Tris, pH 8) was added to each fraction to reestablish neutral pH in the sample.

The protein concentration in each fraction was measured using the Pierce BSA Protein Assay Kit (ThermoFischer Scientific) according to the instructions from the manufacturer and the absorbance measured at 492 nm.

2.3. NanoLC MS Analysis

Purified E protein fraction of 1 mL at an estimated concentration of 1 μ g/mL was processed using the modified filter-aided sample preparation (FASP) method [24]. In brief, the E protein fraction was reduced with 100 mM Dithiothreitol (DTT) at 60 °C for 30 min, transferred to 30 kDa MWCO Pall Nanosep centrifugation filters (Sigma-Aldrich), washed several times with 8 M urea and once with digestion buffer (DB, 50 mM TEAB, 0.5 % sodium deoxycholate (SDC)) prior to alkylation with 18 mM 2-iodoacetamide (IAM) for 30 min at room temperature (in the dark). Samples were digested with 0.1 μ g trypsin (Pierce MS grade Trypsin, Thermo Fisher Scientific, ratio 1:10) at 37 °C overnight followed by an extra 0.1 μ g trypsin addition and 2 hours incubation at 37 °C. Produced proteolytic peptides were collected by centrifugation and filters were additionally washed with 50mM TEAB to ensure that all material eluted from the filters. The collected proteolytic peptides were first purified using HiPPR™ Spin Column, followed by acidification with 10 % TFA and subsequent centrifugation to remove remaining SDC. The SDC free supernatant was desalted using Pierce Peptide Desalting Spin Columns (Thermo Scientific) prior to LC-MS/MS analysis.

The E protein proteolytic preparation was analyzed on a QExactive HF mass spectrometer interfaced with Easy-nLC1200 liquid chromatography system (ThermoFisher). Peptides were trapped on an Acclaim Pepmap 100 C18 trap column (100 μ m $\times$ 2 cm, particle size 5 μ m; ThermoFisher) and separated on an in-house packed analytical column (75 μ m $\times$ 30 cm, particle size 3 μ m, Reprosil-Pur C18; Dr. Maisch, Ammerbuch-Entringen, Germany) using a gradient from 5 % to 35 % acetonitrile in 0.2 % formic acid over 75 minutes at a flow of 300 nL/min. Each preparation was analyzed using two different MS1 scans settings, in the m/z range of 375-1500 and 600-2000, both at a resolution of 120K. MS2 analysis was performed in a data-dependent mode at a resolution of 30K, using a cycle time of 3 seconds. The most abundant precursors with charges 2–7, were selected for fragmentation using HCD at normalized collision energy settings of 28. The isolation window was set to either m/z = 1.2 for data acquired in the m/z range of 380-1500 or m/z = 3.0 for data acquired in the m/z range of 600-2000. The dynamic exclusion was set to 10 ppm for 20 seconds.

2.4. Database Search and Data Processing

The acquired data were analyzed using Proteome Discoverer version 2.4 (Thermo Fisher Scientific). Database searches were performed with either Byonic (Protein Metrics) or Sequest as search engines. The data were searched against custom database consisting of SwissProt_human database (20342 proteins) and the sequence of purified E protein. First, the data was searched with Sequest to evaluate the E protein preparation purity. After that, all glycoproteomic searches were done with Byonics (Protein Metrics) using selected N- and O-glycan database (186 and 30 glycan compositions, respectively). Precursor mass tolerance was set to 5 ppm (Sequest) or 10 ppm (Byonics) and fragment mass tolerance to 30 ppm. Tryptic peptides with up to two missed cleavages were accepted. Variable modification of methionine oxidation and fixed cysteine alkylation were allowed. Target Decoy was used for Peptide Spectrum Match (PSM) validation. Prior to the final assignment, the identified glycosylated peptides were manually evaluated based on the observed fragmentation pattern, the number of glycoforms per site, number of PSM per glycoform and the retention time windows for the different observed glycoforms (at the same site). The extracted ion chromatogram (EIC) peak intensities were used to determine the glycoform abundances, expressed as percent of

total signal for all modified and non-modified peptides sharing the same amino acid sequence. Only structures with a relative abundance of $\geq 1\%$ are listed and shown in Figure 1 and Table S1.

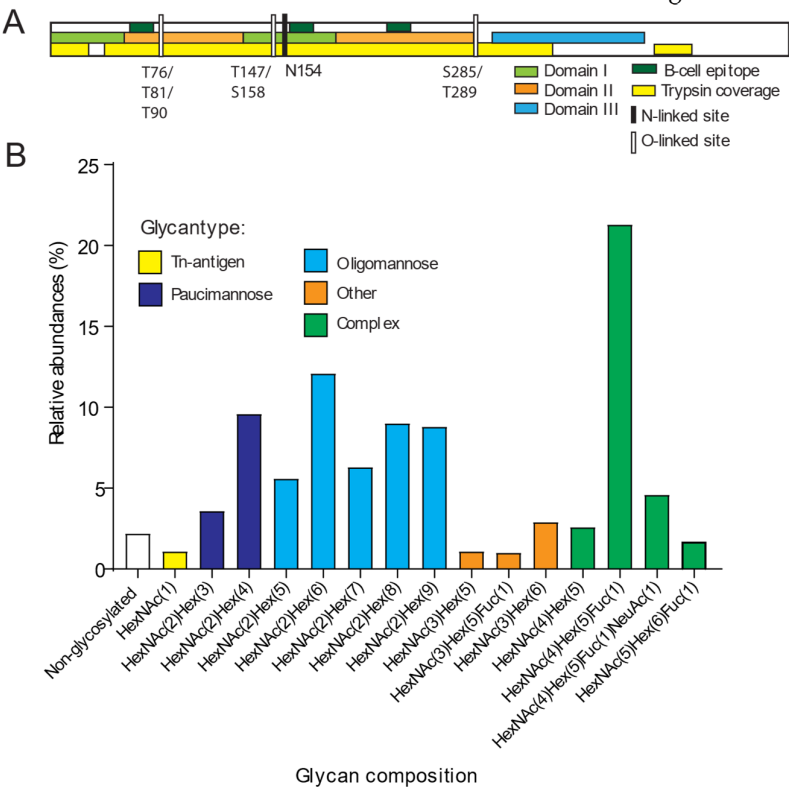


Figure 1. The positions and structures of the O- and N-linked glycans within the E protein of TBEV. **(A)** Schematic drawing of the E protein, showing the B cell-epitopes identified by Kuivanen S et al [13], domains D I-III, the trypsin coverage used in our glycoproteomic analysis, and the identified sites for N- and O-linked glycosylation. **(B)** Glycoform distribution of identified structures detected on the peptide covering the N-linked glycosite. Structures included in the category “other” includes both N-linked glycans of hybrid type, and a combination of N-linked and O-linked glycoforms present on the same peptide. Structures with a relative abundance of $\leq 1\%$ are excluded.

2.5. In Silico Glycosylation

Crystal structures of the E-protein (PDB ID 7QRE and 1SVB) was downloaded from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB). Heteroatoms and the precursor membrane protein fragment of 7QRE was removed using ChimeraX [25]. The attachment of glycans, as well as the calculation of the solvent accessible surface area were done by uploading the structures to the glycoprotein builder from GLYCAM Web [26]. The condensed glycan notations of the attached glycans are as follows:

O-linked glycans (residues T76, T81, T147, S158 and S285): DGalpNAc1-OH

N-linked glycans (residues N154): DManpa1-6[DManpa1-3]DManpa1-6[DManpa1-2DManpa 1-3]DManpb1-4DGlcpNAcb1-4DGlcpNAcb1-OH

3. Results and Discussion

The clinical TBEV isolate, F7203, was grown in the A549 cell line. Extra- and intracellular viral particles were harvested and lysed before purification of the E protein on an immunosorbent column coated with a monoclonal antibody targeting an epitope within ED III. Proteolytic cleavage of the isolated E proteins with trypsin resulted in a peptide coverage of 70 %, which were analyzed for glycan modifications using liquid chromatography tandem-mass spectrometry (LC-MS/MS) (Figure 1A). The peptide coverage included the linear B-cell epitopes [13], the consensus site for N-linked

glycosylation N154-T156, and 54 out of 67 possible sites (80 %), i.e. Serine (S) or Threonine (T), for O-linked glycosylation.

The N-linked consensus site N154-T156 was highly occupied, with 90 % of the peptides carrying N-linked glycan modifications at N154. Of these peptides, 42 % carried an oligomannose type glycan (Figure 1B), with $\text{Man}_6\text{GlcNac}_2$ being the most abundant (present on 12 % of analyzed peptides). High degrees of oligomannose structures have been observed on, among others, the human immunodeficiency virus (HIV) envelope protein gp120, which might indicate a low degree of glycan processing [27]. In gp120, the densely located N-linked glycans sterically hinder enzymatic elongation, however the oligomannose structures of the single N-linked glycan of the E proteins might instead be important for correct protein folding during maturation [28] or for interaction with receptor proteins, regulating infectivity and pathogenesis [17–21]. Interestingly, a large fraction of the oligomannose structures contained -8 or -9 mannose residues. These types of structures are the major glycoforms prior to the exit from the endoplasmic reticulum (ER) during glycan synthesis and are subsequently trimmed down in the trans golgi network (TGN) before elongation and generation of complex type N-linked glycan structures [29]. During the sample preparation, we harvested both intra- and extracellular viral particles prior to the isolation of the E protein. Thus, it is possible that a significant part of the isolated E protein was captured before it was fully processed by the glycosylation machinery. We did not calculate the ratio between intra- and extracellular viral particles before isolating the E protein and it is not possible to determine if there is a selection bias regarding the glycosylation status of the E protein in the immunosorbent column. In addition, flavivirus particles show a large heterogeneity upon release from the infected cell, both with respect to the state of maturation as well as internal protein architecture [30]. This suggests that the protein maturation process, within the ER and the TGN, is subjected to considerable disruptions which could also impact the glycosylation.

Complex type N-linked glycan structures were found on 30 % of the peptides, mainly of the composition $\text{HexNac}_4\text{Hex}_5\text{Fuc}_1$ (present on 21 % of the analyzed peptides) (Figure 1B). A fucose group was identified on 28 % of the complex type glycans, and manual inspection of the MS spectra suggest core fucosylation of the innermost GlcNAc. The addition of a core fucose is synthesized by the enzyme fucosyltransferase 8 (FUT 8) which has been shown to be upregulated in hepatitis B virus transfected hepatoma cells [31,32], and over expression of FUT8 was correlated with increased binding of Hepatitis B virus (HBV)-like particles to the cells [32]. Since viral infection can alter gene expression in the host cell to create glycan compositions in favor for viral spread and replication, the high degree of core fucosylation might be advantageous for the infectivity of TBEV [33–35].

Only 5 % of the N-linked glycans carried one or more sialic acid residues, indicating limited access by sialyltransferases. Sialic acid plays an important role in the recognition of pathogens by macrophages via the sialic acid-binding immunoglobulin-like lectin sialoadhesin (Siglec-1), found on macrophage subsets [29]. Siglec-1 mainly recognizes α 2-3 linked sialic acids, and binding leads to increased uptake of viral particles and enhanced infection, as shown for HIV [29,36]. It has been shown that TBEV infects macrophages and a possible connection to Siglec-1 warrants further investigation [37].

13 % of the peptides covering the N-linked consensus site carried N-linked glycans of paucimannose type ($\text{Man}_{3-4}\text{GlcNac}_2(\text{Fuc}_{1-2})$). Paucimannose decorates the E protein when the virus is grown in tick cells [22] and is commonly present in invertebrates. Vertebrates, however, typically extend the glycan precursor into complex-type N-linked glycans [38]. Previous studies have also identified paucimannosidic structures in tumor cells [39,40], which could explain our findings. Almost 5 % of the N-linked glycans were assigned as “other” in Figure 1B. These structures can either comprise hybrid-type structures, or a combination of N-linked structures and O-linked structures present on the same peptide. PNGaseF-treatment prior to reanalysis would remove the N-linked glycans, enabling identification and analysis of O-linked glycoforms. However, the limited quantity of purified E protein at our disposal prevented further differentiation of these glycoforms.

Three O-linked glycans were found on the E protein (Figure 1A). First, manual inspection of the fragment ion spectra suggests a single HexNAc monosaccharide present in the region covering amino

acids C74-R94, containing three potential sites for O-linked glycosylation (T76, T81 and T90) (Figure S1). Our methodological approach did not permit quantification of glycan structures that were present on less than 1 % of the identified peptides. Therefore, we can only conclude that this O-linked glycan is of low abundance in our E protein preparation. However, despite this site being mainly non-glycosylated our data implies that the glycan is present under certain conditions, but the significance of it, and under what conditions it appears remain unknown. Secondly, a confirmed O-linked glycan was identified on peptides covering amino acids V143-R160, containing two potential sites for O-linked glycosylation (T147 and S158). These sites are close to the N-linked glycan at N154. Of the total number of analyzed peptides covering T147 and S158, 1.1 % carried a single N-acetylgalactosamine (GalNAc). The presence of an O-linked glycoform was confirmed by a high-intensity oxonium ion at m/z 126 during manual inspection of the ion spectra (Figure S2). As the O-linked site on peptides covering amino acids V143-R160 is close to the N-linked site N154, it is possible that this glycan may impact viral entry, infectivity, or pathogenesis either alone or by interaction with the N-linked glycan. This O-linked glycan is also near one of the three B-cell epitopes (covering amino acids 163-180) [13]. Our group has previously shown that antibody reactivity to a peptide decorated with a single GalNAc varies dependent on the site of the modification and that specific sugar residues can both enhance and reduce antibody reactivity [41,42]. The third identified O-linked glycan has two possible sites, S285 or T289. A single GalNAc, as determined by manual inspection of the ion spectra, was detected on peptides covering the amino acids (S285-K296) (Figure S3). Comparable to the O-linked glycan on sites T76, T81 or T90, also this glycan is low-abundant (present on less than 1 % of the analyzed peptides).

Next, we compared the amino acid sequences of 762 samples (European, Siberian and Far Eastern TBEV subtypes) deposited at GenBank, of which, 396 were full length protein E classified as European strain (Table S2 and S3). As expected, the N-linked consensus site N154 was conserved in all samples, exhibiting a 100 % identity among all 762 analyzed sequences. There was also a high degree of conservation among the amino acids which we identified as potential carriers of O-linked glycan structures. T76, T90 and T289 showed 100 % identity in all sequences, and T147 was in four cases replaced with S147, thereby still permitting O-linked glycosylation. T81, S158 and S285 showed 94.5 %, 98.8 % and 99.5 % identity among all strains respectively. Among the European strains, all potential glycosites showed 100 % identity, except for T81 (96.0 %) and S158 (98.7 %). In summary, it appears that most of the potential O-linked glycan sites are conserved among all analyzed protein E sequences. For each O-linked glycan identified, there is at least one potential site with 100 % identity.

To further investigate the potential impact of the O-linked glycans we performed structural analysis of the E protein (PDB ID: 7QRE [10]) to the Glycoprotein Builder from GLYCAM Web [26], after removing the precursor membrane protein fragment and non-standard residues using ChimeraX [25]. We tested all potential sites for O-linked glycosylation as well as the single site for N-linked glycosylation, by adding single GalNAc residues or the most frequently observed N-linked oligomannose structure ($\text{Man}_6\text{GlcNAc}_2$) to N154, with the E 150-loop in its open configuration (Figure 2). Sites T90 and T289 could not harbor GalNAc residues in the model due to a low calculated solvent accessible surface area. Interestingly, T147 and S158 constitute parts of the E 150-loop [10] and both sites are accessible for addition of GalNAc in this model.

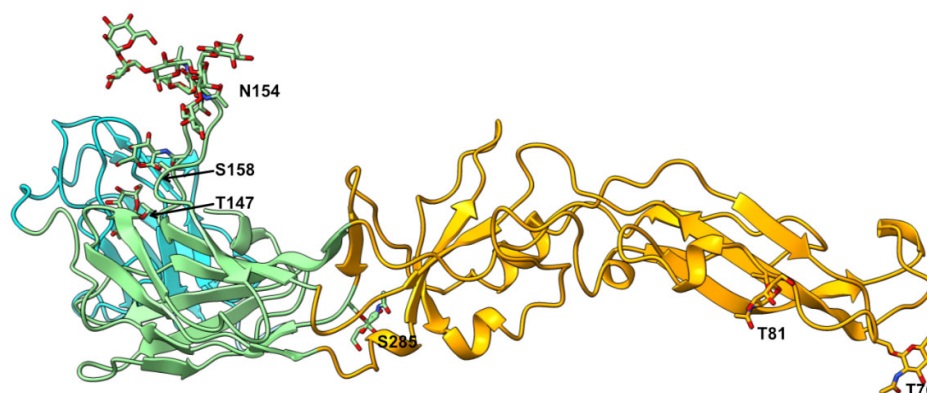


Figure 2. Structural overview of the N- and O-linked glycosylated E protein, with the E 150-loop in open configuration. A representation of the most abundant oligomannose ($\text{Man}_6\text{GlcNac}_2$) is attached to N154. Single GalNAc monosaccharides are attached to T76, T81, T147, S158 and S285. Domain I: blue, Domain II: green, Domain III: orange.

When analyzing site occupancy of O-linked glycans within the E protein of PDB ID 1SVB [16,30] where the E 150-loop is in its closed conformation, the T147 site could not harbor the GalNAc residue. A closer inspection of the E 150-loop indicates that it would sterically hinder O-linked glycosylation at site T147 in its closed conformation (Figure 3 and Figure S4).

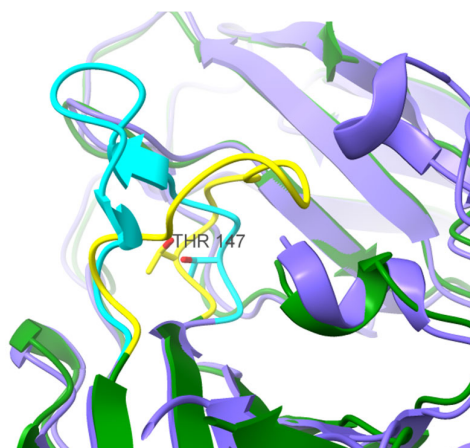


Figure 3. Superimposed non-glycosylated protein structures of PDB ID 1SVB and 7QRE, the proteins are colored green and purple respectively and residues 146-160, constituting the E150 loop, are colored yellow and blue. The E150 loop of 7QRE is kept open at acidic pH by interactions with the precursor to M protein (prM, not shown), while the yellow loop of 1SVB is in a closed conformation. The location of T147 varies between the two models and could not be glycosylated in silico for 1SVB.

Despite only 1.1 % of the peptides carrying an O-linked glycan close to position N154 (T147 and S158), and an even lower number of peptides carrying O-linked glycans at the other positions within the E protein, this type of post translational modification may have implications for viral pathogenesis. The high number of viral particles in the blood during TBEV viremia means that a significant number of viral particles may harbor O-linked glycans [43]. Interestingly, the sequence comparison showed that T147 was highly conserved, but the threonine was exchanged for a serine in four cases, possibly indicating selective pressure to maintain the ability for O-linked glycosylation. Moreover, the prediction tool indicated that T147 is only accessible for the addition of an GalNAc residue when the fusion loop is in the open conformation. We consider three possible scenarios that are compatible with O-linked glycosylation of site T147 or S158; (i) Addition of an GalNAc residue at

T147 occur early in the TGN which may help stabilize the fusion loop in an open configuration. In this scenario the GalNAc needs to be cleaved before the fusion loop is closed in a pH-dependent manner [10]. (ii) Addition of an GalNAc residue occurs at site S158 which may or may not have an impact on the E 150-loop. (iii) We harvested the E protein from infected cells, and it cannot be ruled out that a proportion of the peptides come from misfolded proteins that are destined for degradation. Thus, the glycosylation that we observe at site T147 or S158 may be an artefact. At this stage, we cannot speculate as to which of these scenarios is most likely, but the N-linked glycan at position N154 at the distal tip of the E150-loop is important for viral infectivity and may affect cell tropism, including crossing the blood-brain-barrier [17,21]. The mechanism has however not been fully uncovered, and the impact of the composition of the N-linked glycan and the presence of an adjacent O-linked glycan warrants further investigation.

Moreover, the identified low-abundance O-linked glycans at sites T76, T81 or T90 and at sites S285 or T289 are within or close to ED II. This domain contains the highly conserved fusion loop at aa 98-110 [44], which mediates membrane fusion and dimer formation by interactions with precursor to M protein (prM), and is the main target for cross-reacting antibodies [45–47]. Post translational modification by glycosylation of B cell epitopes can potentially alter antibody recognition. Since the glycans on viral proteins are synthesized by the host cell glycosylation machinery, they could be considered “self” by the immune system, and therefore do not induce an immune response. Instead, these large self-like glycans might prevent binding of neutralizing antibodies due to physical hindrance, thereby limiting the protective effect of the antibody epitope [48]. However, small glycans, such as a single GalNAc, can enhance antibody recognition of a peptide. For example, 70 % of the tested sera from patients infected with herpes simplex virus type 2 (HSV-2) have IgG directed toward a heptamer within glycoprotein G, carrying a single GalNAc modification at T504. This serum reactivity is lost when the glycan moiety is absent [49]. Additionally, depending on the site of the glycan modification, individual sera can confer diverse responses towards the same glycosylated epitope [41]. Altogether, this implies that glycosylation can influence antigen processing and presentation by antigen-presenting cells (APCs), resulting in a shifted immunodominance where the antibody repertoire is directed towards certain epitopes to a higher degree than others [50]. Also, many of the vaccines used against TBEV are based on inactivated virus produced in Chick Embryo Fibroblasts (CEF), and the viral particles hence likely contain glycan signatures of the CEF cells. A recombinant vaccine, with a glycosylation profile significantly different from the actual glycosylation present on the viral particle might confer a lower protection than a vaccine with a matching glycosylation profile [51]. Thus, the glycosylation profile should be considered when new vaccine candidates are developed.

4. Conclusions

We have performed a detailed characterization of the glycan profile of the E protein and found three novel O-linked glycans in addition to the previously described N-linked glycan. A comparison of E protein sequences from European, Siberian and Far Eastern TBEV subtypes indicates a high degree of conservation for most of the potential sites for O-linked glycosylation as well as for the N-linked glycan site. Despite the fact the number of peptides which carry O-linked glycans are of low abundance it is intriguing that two positions (T147 and S158) are close to the E 150-loop which is shown to be essential for maturation of the TBEV particle.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: LC-MS/MS spectra showing the fragmentation patterns of the naked pep-tide C74-R94 and the corresponding peptide carrying a single HexNAc modification; Figure S2: LC-MS/MS spectra showing the fragmentation patterns of the naked pep-tide V143-R160 and the corresponding peptide carrying a single HexNAc modification; Figure S3: LC-MS/MS spectra showing the fragmentation patterns of the naked peptide V285-R295 and the corresponding peptide carrying a single HexNAc modification; Figure S4: Location of T154 in the E protein structure visualized as solvent accesibel surface; Table S1: Observed glycopeptides of site N154 of TBEV strain F7203 grown in the human adenocarcinomic cell line A549; Table S2: Amino acid similarity at potential glycosylation sites within 762 protein E se-quences of European, Siberian and

far eastern TBEV subtypes; Table S3: Amino acid similarity at potential glycosylation sites within 396 protein E sequences of European TBEV strains.

Author Contributions: Conceptualization, E.K., T.B. and R.N.; Methodology, E.K., E.M., K.N. and A.K.; Data curation, E.M.; Validation, E.K., E.M., K.N. and A.K.; Formal analysis, E.K., E.M., K.N. and A.K.; Investigation, E.K., E.M., K.N. and A.K.; Resources, E.M., K.S., T.B. and R.N.; Writing – Original draft Preparation, E.K. and R.N.; Writing – Review & Editing, E.K., E.M., K.N., K.S., A.K., T.B. and R.N.; Visualization, E.K., E.M., K.N. and A.K.; Supervision, E.M. and R.N.; Project administration, R.N.; Funding acquisition, R.N.

Funding: The study was funded by the Sahlgrenska University Hospital. Project grant FoU laboratoriemedicin and project grant C4A.

Data Availability Statement: The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Conflicts of interest: The authors declare no conflict of interest.

References

1. Ruzek, D.; Avšič Županc, T.; Borde, J.; Chrdle, A.; Eyer, L.; Karganova, G.; Kholodilov, I.; Knap, N.; Kozlovskaya, L.; Matveev, A.; Miller, A.D.; Osolodkin, D.I.; Överby, A.K.; Tikunova, N.; Tkachev, S.; Zajkowska, J., Tick-borne encephalitis in Europe and Russia: Review of pathogenesis, clinical features, therapy, and vaccines. *Antiviral Res* **2019**, *164*, 23-51.
2. Haglund, M.; Günther, G., Tick-borne encephalitis--pathogenesis, clinical course and long-term follow-up. *Vaccine* **2003**, *21 Suppl 1*, S11-S18.
3. Anastasina, M.; Füzik, T.; Domanska, A.; Pulkkinen, L.I.A.; Šmerdová, L.; Formanová, P.P.; Straková, P.; Nováček, J.; Růžek, D.; Plevka, P.; Butcher, S.J., The structure of immature tick-borne encephalitis virus supports the collapse model of flavivirus maturation. *Sci Adv* **2024**, *10* (27), ead1888.
4. Kubinski, M.; Beicht, J.; Gerlach, T.; Volz, A.; Sutter, G.; Rimmelzwaan, G.F., Tick-Borne Encephalitis Virus: A Quest for Better Vaccines against a Virus on the Rise. *Vaccines (Basel)* **2020**, *8* (3).
5. Hansson, K.E.; Rosdahl, A.; Insulander, M.; Vene, S.; Lindquist, L.; Gredmark-Russ, S.; Askling, H.H., Tick-borne Encephalitis Vaccine Failures: A 10-year Retrospective Study Supporting the Rationale for Adding an Extra Priming Dose in Individuals Starting at Age 50 Years. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2020**, *70* (2), 245-251.
6. Lotrič-Furlan, S.; Bogovič, P.; Avšič-Županc, T.; Jelovšek, M.; Lusa, L.; Strle, F., Tick-borne encephalitis in patients vaccinated against this disease. *J Intern Med* **2017**, *282* (2), 142-155.
7. Chambers, T.J.; Hahn, C.S.; Galler, R.; Rice, C.M., FLAVIVIRUS GENOME ORGANIZATION, EXPRESSION, AND REPLICATION. *Annual review of microbiology* **1990**, *44* (1), 649-688.
8. Pulkkinen, L.I.A.; Barrass, S.V.; Domanska, A.; Överby, A.K.; Anastasina, M.; Butcher, S.J., Molecular Organisation of Tick-Borne Encephalitis Virus. *Viruses* **2022**, *14* (4).
9. Chao, L.H.; Klein, D.E.; Schmidt, A.G.; Peña, J.M.; Harrison, S.C., Sequential conformational rearrangements in flavivirus membrane fusion. *eLife* **2014**, *3*, e04389.
10. Vaney, M.C.; Dellarole, M.; Duquerroy, S.; Medits, I.; Tsouchnikas, G.; Rouvinski, A.; England, P.; Stiasny, K.; Heinz, F.X.; Rey, F.A., Evolution and activation mechanism of the flavivirus class II membrane-fusion machinery. *Nature communications* **2022**, *13* (1), 3718.
11. Allison, S.L.; Schlich, J.; Stiasny, K.; Mandl, C.W.; Heinz, F.X., Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. *Journal of virology* **2001**, *75* (9), 4268-4275.
12. Zhang, X.; Jia, R.; Shen, H.; Wang, M.; Yin, Z.; Cheng, A., Structures and Functions of the Envelope Glycoprotein in Flavivirus Infections. *Viruses* **2017**, *9* (11).
13. Kuivanen, S.; Hepojoki, J.; Vene, S.; Vaheri, A.; Vapalahti, O., Identification of linear human B-cell epitopes of tick-borne encephalitis virus. *Virol J* **2014**, *11*, 115.
14. Kiermayr, S.; Stiasny, K.; Heinz, F.X., Impact of quaternary organization on the antigenic structure of the tick-borne encephalitis virus envelope glycoprotein E. *Journal of virology* **2009**, *83* (17), 8482-91.
15. Jarmer, J.; Zlatkovic, J.; Tsouchnikas, G.; Vratskikh, O.; Strauß, J.; Aberle, J.H.; Chmelik, V.; Kundi, M.; Stiasny, K.; Heinz, F.X., Variation of the specificity of the human antibody responses after tick-borne encephalitis virus infection and vaccination. *Journal of virology* **2014**, *88* (23), 13845-57.
16. Rey, F.A.; Heinz, F.X.; Mandl, C.; Kunz, C.; Harrison, S.C., The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **1995**, *375* (6529), 291-8.
17. Pletnev, A.G.; Bray, M.; Lai, C.J., Chimeric tick-borne encephalitis and dengue type 4 viruses: effects of mutations on neurovirulence in mice. *Journal of virology* **1993**, *67* (8), 4956-4963.
18. Martina, B.E.E.; Koraka, P.; van den Doel, P.; Rimmelzwaan, G.F.; Haagmans, B.L.; Osterhaus, A.D.M.E., DC-SIGN enhances infection of cells with glycosylated West Nile virus in vitro and virus replication in human dendritic cells induces production of IFN- α and TNF- α . *Virus research* **2008**, *135* (1), 64-71.

19. Goto, A.; Yoshii, K.; Obara, M.; Ueki, T.; Mizutani, T.; Kariwa, H.; Takashima, I., Role of the N-linked glycans of the prM and E envelope proteins in tick-borne encephalitis virus particle secretion. *Vaccine* **2005**, *23* (23), 3043-3052.
 20. Sayce, A.C.; Alonzi, D.S.; Killingbeck, S.S.; Tyrrell, B.E.; Hill, M.L.; Caputo, A.T.; Iwaki, R.; Kinami, K.; Ide, D.; Kiappes, J.L.; Beatty, P.R.; Kato, A.; Harris, E.; Dwek, R.A.; Miller, J.L.; Zitzmann, N., Iminosugars Inhibit Dengue Virus Production via Inhibition of ER Alpha-Glucosidases—Not Glycolipid Processing Enzymes. *PLOS Neglected Tropical Diseases* **2016**, *10* (3), e0004524.
 21. Beasley, D.W.C.; Whiteman, M.C.; Zhang, S.; Huang, C.Y.H.; Schneider, B.S.; Smith, D.R.; Gromowski, G.D.; Higgs, S.; Kinney, R.M.; Barrett, A.D.T., Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. *Journal of virology* **2005**, *79* (13), 8339-8347.
 22. Lattová, E.; Straková, P.; Pokorná-Formanová, P.; Grubhoffer, L.; Bell-Sakyi, L.; Zdráhal, Z.; Palus, M.; Ruzek, D., Comprehensive N-glycosylation mapping of envelope glycoprotein from tick-borne encephalitis virus grown in human and tick cells. *Scientific reports* **2020**, *10* (1), 13204.
 23. Haglund, M.; Vene, S.; Forsgren, M.; Günther, G.; Johansson, B.; Niedrig, M.; Plyusnin, A.; Lindquist, L.; Lundkvist, A., Characterisation of Human Tick-Borne Encephalitis Virus From Sweden. *Journal of Medical Virology* **2003**, *71* (4), 610-621.
 24. Wiśniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M., Universal sample preparation method for proteome analysis. *Nat Methods* **2009**, *6* (5), 359-62.
 25. Meng, E.C.; Goddard, T.D.; Pettersen, E.F.; Couch, G.S.; Pearson, Z.J.; Morris, J.H.; Ferrin, T.E., UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **2023**, *32* (11), e4792.
 26. Group, W. GLYCAM Web. <http://glycam.org>.
 27. Struwe, W.B.; Chertova, E.; Allen, J.D.; Seabright, G.E.; Watanabe, Y.; Harvey, D.J.; Medina-Ramirez, M.; Roser, J.D.; Smith, R.; Westcott, D.; Keele, B.F.; Bess, J.W., Jr.; Sanders, R.W.; Lifson, J.D.; Moore, J.P.; Crispin, M., Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen. *Cell Rep* **2018**, *24* (8), 1958-1966.e5.
 28. Jitsuhara, Y.; Toyoda, T.; Itai, T.; Yamaguchi, H., Chaperone-Like Functions of High-Mannose Type and Complex-Type N-Glycans and Their Molecular Basis1. *The Journal of Biochemistry* **2002**, *132* (5), 803-811.
 29. In *Essentials of Glycobiology*, Varki, A.; Cummings, R.D.; Esko, J.D.; Stanley, P.; Hart, G.W.; Aebi, M.; Darvill, A.G.; Kinoshita, T.; Packer, N.H.; Prestegard, J.H.; Schnaar, R.L.; Seeberger, P.H., Eds. Cold Spring Harbor Laboratory Press
- Copyright 2015-2017 by The Consortium of Glycobiology Editors, La Jolla, California. All rights reserved.: Cold Spring Harbor (NY), 2015.
30. Rey, F.A.; Stiasny, K.; Heinz, F.X., Flavivirus structural heterogeneity: implications for cell entry. *Current opinion in virology* **2017**, *24*, 132-139.
 31. Schneider, M.; Al-Shareffi, E.; Haltiwanger, R.S., Biological functions of fucose in mammals. *Glycobiology* **2017**, *27* (7), 601-618.
 32. Takamatsu, S.; Shimomura, M.; Kamada, Y.; Maeda, H.; Sobajima, T.; Hikita, H.; Iijima, M.; Okamoto, Y.; Misaki, R.; Fujiyama, K.; Nagamori, S.; Kanai, Y.; Takehara, T.; Ueda, K.; Kuroda, S. i.; Miyoshi, E., Core-fucosylation plays a pivotal role in hepatitis B pseudo virus infection: a possible implication for HBV glycotherapy. *Glycobiology* **2016**, *26* (11), 1180-1189.
 33. Norden, R.; Samuelsson, E.; Nystrom, K., NFkappaB-mediated activation of the cellular FUT3, 5 and 6 gene cluster by herpes simplex virus type 1. *Glycobiology* **2017**, *27* (11), 999-1005.
 34. Nyström, K.; Nordén, R.; Muylaert, I.; Elias, P.; Larson, G.; Olofsson, S., Induction of sialyl-Lex expression by herpes simplex virus type 1 is dependent on viral immediate early RNA-activated transcription of host fucosyltransferase genes. *Glycobiology* **2009**, *19* (8), 847-59.
 35. Nyström, K.; Grahn, A.; Lindh, M.; Brytting, M.; Mandel, U.; Larson, G.; Olofsson, S., Virus-induced transcriptional activation of host FUT genes associated with neo-expression of Ley in cytomegalovirus-infected and sialyl-Lex in varicella-zoster virus-infected diploid human cells. *Glycobiology* **2007**, *17* (4), 355-66.
 36. Zou, Z.; Chastain, A.; Moir, S.; Ford, J.; Trandem, K.; Martinelli, E.; Cicala, C.; Crocker, P.; Arthos, J.; Sun, P.D., Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids. *PloS one* **2011**, *6* (9), e24559.
 37. Ahantarig, A.; Růzek, D.; Vancová, M.; Janowitz, A.; St'astná, H.; Tesarová, M.; Grubhoffer, L., Tick-borne encephalitis virus infection of cultured mouse macrophages. *Intervirology* **2009**, *52* (5), 283-90.
 38. Kost, T.A.; Condreay, J.P.; Jarvis, D.L., Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nature biotechnology* **2005**, *23* (5), 567-75.
 39. Becker, Y.; Förster, S.; Gielen, G.H.; Loke, I.; Thaysen-Andersen, M.; Laurini, C.; Wehrand, K.; Pietsch, T.; Diestel, S., Paucimannosidic glycoepitopes inhibit tumorigenic processes in glioblastoma multiforme. *Oncotarget* **2019**, *10* (43), 4449-4465.

40. Pont, L.; Kuzyk, V.; Benavente, F.; Sanz-Nebot, V.; Mayboroda, O.A.; Wuhler, M.; Lageveen-Kammeijer, G.S.M., Site-Specific N-Linked Glycosylation Analysis of Human Carcinoembryonic Antigen by Sheathless Capillary Electrophoresis-Tandem Mass Spectrometry. *J Proteome Res* **2021**, *20* (3), 1666-1675.
41. Nordén, R.; Nilsson, J.; Samuelsson, E.; Risinger, C.; Sihlbom, C.; Blixt, O.; Larson, G.; Olofsson, S.; Bergstrom, T., Recombinant Glycoprotein E of Varicella Zoster Virus Contains Glycan-Peptide Motifs That Modulate B Cell Epitopes into Discrete Immunological Signatures. *International journal of molecular sciences* **2019**, *20* (4).
42. Samuelsson, E.; Mirgorodskaya, E.; Nyström, K.; Bäckström, M.; Liljeqvist, J.; Nordén, R., Sialic Acid and Fucose Residues on the SARS-CoV-2 Receptor-Binding Domain Modulate IgG Antibody Reactivity. *ACS Infect Dis* **2022**, *8* (9), 1883-1893.
43. Saksida, A.; Jakopin, N.; Jelovšek, M.; Knap, N.; Fajs, L.; Lusa, L.; Lotrič-Furlan, S.; Bogovič, P.; Arnež, M.; Strle, F.; Avšič-Županc, T., Virus RNA Load in Patients with Tick-Borne Encephalitis, Slovenia. *Emerg Infect Dis* **2018**, *24* (7), 1315-1323.
44. Chakraborty, S., Computational analysis of perturbations in the post-fusion Dengue virus envelope protein highlights known epitopes and conserved residues in the Zika virus. *F1000Res* **2016**, *5*, 1150-1150.
45. Roby, J.A.; Setoh, Y.X.; Hall, R.A.; Khromykh, A.A., Post-translational regulation and modifications of flavivirus structural proteins. *The Journal of general virology* **2015**, *96* (Pt 7), 1551-69.
46. Chiou, S.S.; Fan, Y.C.; Crill, W.D.; Chang, R.Y.; Chang, G.J., Mutation analysis of the cross-reactive epitopes of Japanese encephalitis virus envelope glycoprotein. *The Journal of general virology* **2012**, *93* (Pt 6), 1185-1192.
47. Rey, F.A.; Stiasny, K.; Vaney, M.C.; Dellarole, M.; Heinz, F.X., The bright and the dark side of human antibody responses to flaviviruses: lessons for vaccine design. *EMBO Rep* **2018**, *19* (2), 206-224.
48. Olofsson, S.; Blixt, O.; Bergstrom, T.; Frank, M.; Wandall, H.H., Viral O-GalNAc peptide epitopes: a novel potential target in viral envelope glycoproteins. In *Reviews in medical virology*, 2015/11/03 ed.; 2016; Vol. 26, pp 34-48.
49. Clo, E.; Kracun, S.K.; Nudelman, A.S.; Jensen, K.J.; Liljeqvist, J.A.; Olofsson, S.; Bergstrom, T.; Blixt, O., Characterization of the viral O-glycopeptidome: a novel tool of relevance for vaccine design and serodiagnosis. *Journal of virology* **2012**, *86* (11), 6268-78.
50. Angeletti, D.; Gibbs, J.S.; Angel, M.; Kosik, I.; Hickman, H.D.; Frank, G.M.; Das, S.R.; Wheatley, A.K.; Prabhakaran, M.; Leggat, D.J.; McDermott, A.B.; Yewdell, J.W., Defining B cell immunodominance to viruses. *Nature immunology* **2017**, *18* (4), 456-463.
51. Avci, F.Y.; Li, X.; Tsuji, M.; Kasper, D.L., A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design. *Nature medicine* **2011**, *17* (12), 1602-9.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.