Characterization of N-Linked Glycans of Chikungunya Virus

By
Clancy Mullan

Honors Thesis
Department of Chemistry
University of North Carolina at Chapel Hill

December 2013

Approved:

[redacted]

Abstract

Chikungunya virus is an arthropod-borne *alphavirus* that, in recent years, has presented an increased threat to human populations as the *Aedes aegypti* mosquitos which carry it have experienced resurgence in their population and by genetic adaptation to a new vector: *aedes albopictus*. Chikungunya virus contains three glycoproteins (E1, E2, E3). E2 contains two N-linked glycosylation sites (N263, N345), and E1 contains a single site (N141). These glycans may be important for modulating host interferon response and tropism, as has been shown for other alphaviruses. This thesis seeks to identify the primary oligosaccharide at each glycosylation site by endoglycosidase digestion and SDS-PAGE. Furthermore, the experiments described herein aim to determine the effect of the removal of either or both of the N-linked glycans of E2 on its ability to bind Human Prohibitin-1, a previously identified receptor.

Introduction

Chikungunya, an arthropod-borne virus of the family *Togaviridae;* genus *alphavirus*, is an enveloped, positive sense RNA virus[i]. CHIKV has become increasingly important to study as efforts to control these vectors have decreased, humans have encroached on the tropical habitats of the vectors, world travel has increased, and global warming has expanded the habitats of its vectors, *Aedes aegypti* and *Aedes albopictus*, which are native to tropical to subtropical climates and temperate climates, respectively. In the 20th century, chikungunya (CHIKV) was confined to Africa, Asia, and the Indian subcontinent; however, several large outbreaks since 1999 have occurred, including outbreaks in the Democratic Republic of the Congo (1999-2000), Gabon (2007), La Réunion (2005-2006), and South-East Asia (2005-2007), as well as the first reported cases of transmission in Europe[ii]. Currently, no specific treatment or vaccine exists for CHIKV.

The genome of CHIKV is single stranded, positive-sense RNA. It is organized into two open reading frames that encode the nonstructural and structural polyproteins of the virus. The first contains the four nonstructural proteins, which are responsible for the production of both negative and positive sense (subgenomic and full length, respectively) RNA. This is used as the template for the production of mRNA for the viral structural proteins. The structural proteins consist of capsid, two major glycoproteins, and accessory proteins. The genome possesses a 5' 7-methylguanosine cap and is polyadenylated at the 3' end[iii].

Sequence analysis has identified a total of three N-linked glycosylation sites in the E1 and E2 glycoproteins: two in E2, located at N263 and N345, and one in E1 at N141. These sites are demarcated by the sequon Asn-X-Thr or Asn-X-Ser, where X can be any amino acid except for Pro[iv]. A sugar precursor is added to the Asn residue in the lumen of the ER[v]. This precursor core is then digested by various exoglycosidases as the polypeptide is transported through the Golgi. This process leads to the production of the predominant oligosaccharide of one of three major classes of N-linked glycosylation: high-mannose, hybrid, and complex[vi]. While mammalian cells can produce each of the three types, arthropods do not produce complex N-linked glycans[vii].

N-linked glycans have been shown to have an important role in alphavirus biology. Previous work in the Heise lab has demonstrated that CHIKV lacking one or both N-linked glycans of E2 retained its ability to replicate but was severely attenuated in its ability to induce swelling and joint inflammation of infected mice[viii]. N-linked glycans have also been shown to modulate alphavirus tropism through interactions with C-type lectins[ix] and to modulate host

Type 1 interferon response[x]. Finally, it was shown that Ross River virus N-Linked glycans promote binding of mannose-binding lectin, which leads to virus-induced disease[xi].

Prohibitin-1 is a protein that has been identified as a potential receptor for CHIKV, and it has been shown to specifically bind to the E2 glycoprotein by co-immunoprecipitation[xii]. In humans, Prohibitin-1 dimerizes with Prohibitin-2. Due to its antiproliferative properties and role in apoptosis, prohibitin is very well conserved throughout eukaryotic cells[xiii,xiv] which are much like the dimerized prohibitin protein. Prohibitin has also been identified as a receptor protein for Dengue-2 in insect cells[xv]. As E2 is known to have a receptor binding role in CHIKV, it is of interest to study the effects of N-linked glycosylation of this protein on this interaction.

In order to study this interaction, a technology that allows for the confirmation of suspected protein interactions using split fluorescent proteins, bimolecular fluorescence complementation (BiFC), will be employed. The BiFC constructs described here reconstitute a functional tripartite GFP molecule[xvi]. The three GFP fragments are the nine N-terminal beta strands of GFP, the 10^{th} strand, and the 11^{th} strand. The tenth and eleventh strand are added in frame to the proteins of interest in various orientations by a high-glycine linker region. If the proteins of interest interact in such a way that one of the orientations places the 10^{th} and 11^{th} strands in close proximity in the cytoplasm, then the three pieces will reconstitute the fluorescent protein.

The first goal of this thesis is to identify the predominant oligosaccharide at each of the three N-linked glycosylation sites within the glycoproteins of Chikungunya virus, which will be done by endoglycosidase digestion and Western blotting. The second is to confirm a specific E2/PHB-1 interaction using GFP-based bimolecular fluorescence complementation (BiFC) and to study the importance of the N-linked glycans of E2 in E2/PHB-1 binding.

Materials and Methods

Glycoprotein Expression Construct Design

Standard PCR based molecular cloning methodologies were used to introduce asparagine to glutamine mutations in at the correct sites (E2 N263Q, E2 N345Q, E2 N263;345Q [DM], and E1 N141Q) in full length CHIKV SL15649 (constructs provided by Charles McGee). These glycoproteins, as well as the wild type proteins, were PCR-amplified from full length clones using primer pairs that incorporated a 3' FLAG tag for the E2 constructs or a 3' V5 tag for the E1 constructs and 5' Nhe1/3' BamH1 sites. These amplicons were then cloned into pcDNA3.1 (Invitrogen) for mammalian expression or pMIB-A Δ(Honeybee melittin secretion signal) (Invitrogen) for expression in arthropod cells. Constructs were verified by PCR, restriction digest, and sequencing.

Tripartite Split-GFP Construct Design[xvii]

The BiFC constructs reconstitute a tripartite GFP molecule. The entirety of the optimized GFP sequence was produced by IDT. The three constructs are as follows. GFP1-9 encoding the first nine beta strands of GFP was PCR amplified from the IDT-GFP and cloned into pcDNA3.1. Four cassettes were designed for ease of future cloning: 5' GFP-10, 3' GFP-10, 5' GFP-11, and 3' GFP-11 with a multiple cloning site present at the junction of the linker (5' linker: GTDVGSGGGSHMGGG; 3' linker: GSGGGSGGGSTS) and the gene of interest. Each of the E2 glycan mutants described above and Prohibitin-1 were cloned into this cassette in pcDNA3.1.

Cell Culture

293T, a line of human embryonic kidney cells selected for high transfectability, cells were grown in 10 cm^2 dishes in DMEM High Glucose [Gibco] supplemented with 10% fetal bovine serum [Atlanta Biologicals], 1% Penicillin/Streptomycin, and 1% L-Glutamine at 37^oC with 5% CO_2. *Ae. albopictus* (C6/36) cells were grown in T-75 flasks in 80% L-15 Media [Cellgro], 10% Triptose phosphate buffer, 10% fetal bovine serum [Atlanta Biologicals], 100 units/mL Pencillin, 100 units/mL Streptomycin, and 1% L-Glutamine at 28^oC without CO_2.

Plasmid Transfection

Six-well plates were seeded with 293T cells 24 hours before transfection at a density of 1 x 10^6 cells per well in 2 mL maintenance per well. At the time of transfection, media was replaced with serum-free media. 30 μg plasmid was added 125μL 1M $CaCl_2$ and the volume was brought up to 500 μL with ddH_2O. 500 μL 2X BES (6.7 mL ddH_2O, 2 mL 1.4M NaCl, 1 mL 0.5M BES [Sigma], 100 μL 150 mM Na_2HPO_4, ~220 μL 1N NaOH, pH 6.95, filter sterilized) was added dropwise with pulse vortexing. The entire mix was incubated for 45 min. at room temperature then 200 μL was added dropwise to each well. 600 μL FBS was added after 90 minutes and media was changed the next day. After 48 hours, cells were lysed with 1X RIPA buffer with added protease inhibitor, centrifuged, and the supernatant collected.

Six-well plates were seeded with 5 x 10^6 C6/36 cells per well 24 hours before transfection in 2 mL maintenance media per well. The Reagent:DNA complex was prepared as follows: X-tremeGene HP (Roche Applied Science) warmed to room temperature and pulse vortexed, 4 μg plasmid DNA added to 200 μL empty L-15 media, 16 μL transfection reagent added to reaction, mixed, and incubated at room temperature for 15-30 minutes. The reaction mix was added

dropwise to cells. After 48 hours, cells were lysed with 1X RIPA buffer with added protease inhibitor, centrifuged, and the supernatant collected.

Endoglycosidase Digestions, SDS-PAGE and Western Blotting

Protein concentration was determined in each sample by Bradford assay. 60 μg protein was diluted to 3 μg/μL and digested with 10 U PNGase F or Endo H (New England Biolabs) overnight at 37^{o}C. 10 μL of this reaction (~30 μg protein) was analyzed by sodium dodecyl sulfate (SDS)-15% polyacrylamide gel electrophoresis (PAGE) and subjected to Western blot analysis using anti-CHIKV mouse immune ascites fluid followed by incubation with an anti-mouse horseradish-peroxidase-conjugated secondary antibody. Protein bands were visualized by ECL Plus (Amersham) immunofluorescence and developed on Amersham Hyperfilm ECL film. A titration against 18 μg E2 or E1 using anti-Chikungunya mouse serum was also performed in order to determine the efficacy of using this reagent to detect the major glycoproteins.

Coomassie Staining[xviii]

Gels were fixed in three volumes of 50% methanol, 10% acetic acid, 40% water fixing solution at room temperature for two hours on an orbital shaker. Fixing solution was removed and gel was covered in staining solution (50% methanol, 0.05% Coomassie brilliant blue R-250, 10% acetic acid, 40% water) for four hours with gentle agitation. Staining solution was removed and gel was rinsed with 50 mL fixing solution. Fixing solution was removed and gel was covered with destaining solution (5% methanol, 7% acetic acid, 88% water) for two hours at room

temperature with agitation. Destaining solution was removed, and destaining process was repeated until background was clear.

Results

For both C6/36 and 293T transfections, GFP expressing plasmids were transfected as a positive control in order to gauge transfection efficiency. Figure 1 demonstrates the efficiency of transfection in both 293T (top) and C6/36 (bottom) cells.

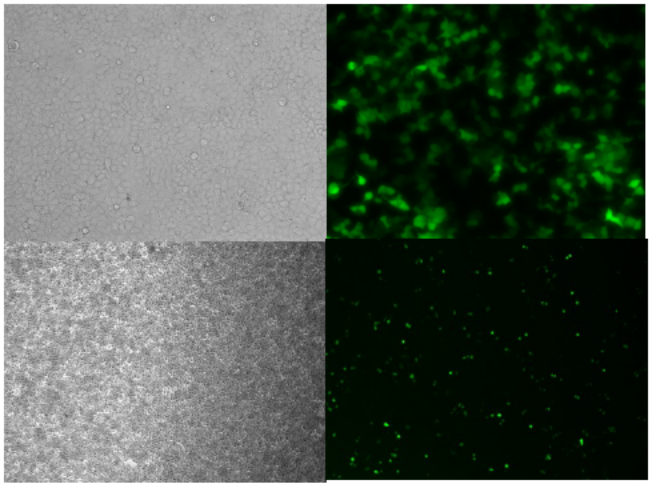


Figure 1; GFP transfection controls. 293T cells white light (top left), 293T cells expressing GFP (top right), C6/36 cells white light (bottom left), C6/36 cells expressing GFP (bottom right)

Once transfection efficiency was established, Western blot analysis was used to test whether plasmid based expression of E2 or E1 was successful and detectable. Using mouse anti-Chikungunya ascites fluid, Western blotting failed to identify 293T expressed E2 or E1, as seen in Figure 2. The 293T expressed GFP control is quite pronounced, suggesting that either the proteins failed to express correctly, were unstable following expression, or the ascites fluid does not contain antibodies that bind to the chikungunya glycoproteins. Serum from 25-day-old CHIKV

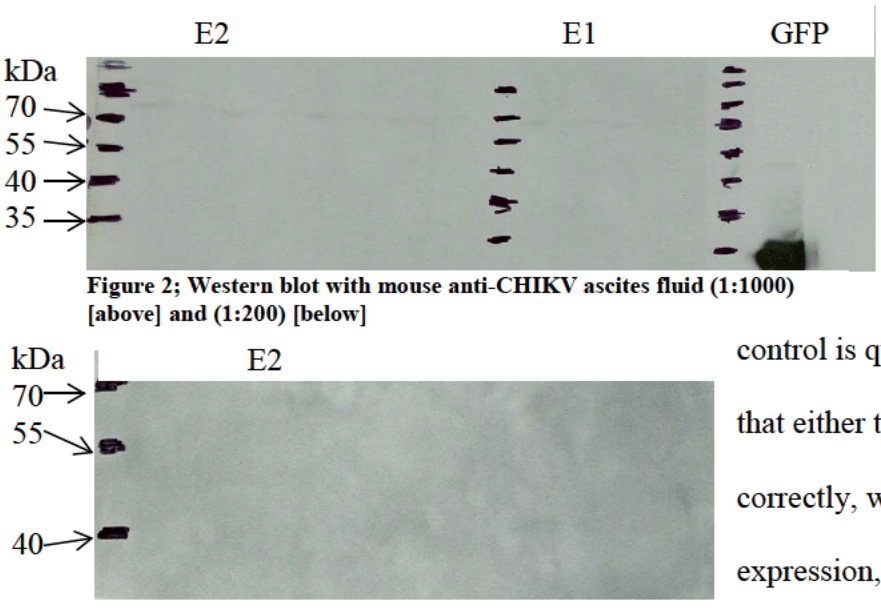


Figure 2; Western blot with mouse anti-CHIKV ascites fluid (1:1000) [above] and (1:200) [below]

infected mice was also tested as a primary antibody in a titration of serum concentrations from 1:1000 to 1:100 against a constant 10 µg total protein, which is displayed as Figure 3. In order to determine whether the protein expression or the antibody being used was problematic, the polyacrylamide gels were stained with Coomassie blue (Figure 4) for whole protein visualization.

Discussion

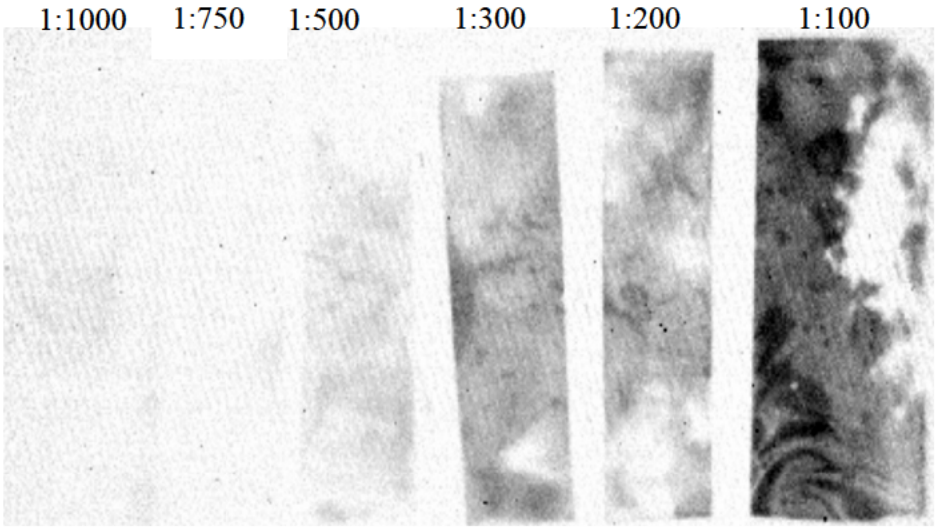
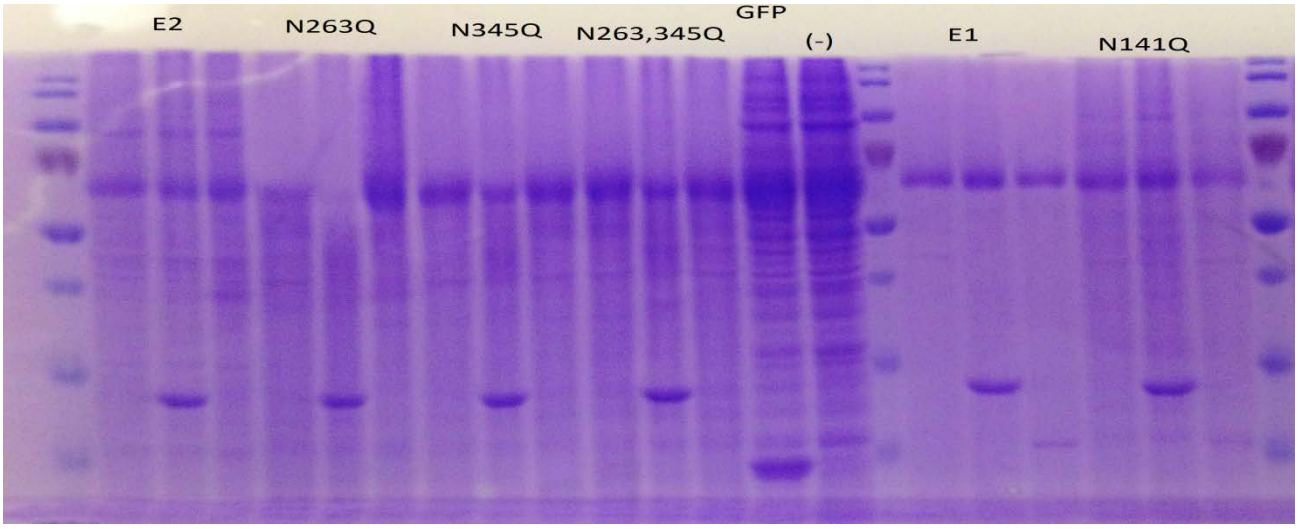


Figure 1; Titration of anti-CHIKV serum from 25-day-old mice against 10µg E2 and E1 with concentrations listed above respected lanes

Given all of the methods of visualization, it appears that the CHIKV glycoproteins were not successfully expressed in mammalian cells, which is most clearly seen in the Coomassie stained gel (Figure 4). Furthermore it appears to be approximately the same size as a major cellular product, which further obfuscates the results. I postulate that this is due to the differences between plasmid-based expression of the glycoproteins and native expression, which may result in incorrect protein conformation or protein degradation. In CHIKV, the glycoproteins are expressed as a single polyprotein (E3, E2, 6K, E1) that undergoes posttranslational processing into the mature glycoproteins. Since E2 is typically expressed as pE2 (a precursor) that includes the chaperone protein E3, it is likely that E2 is unable to fold correctly when expressed alone. Furthermore, it is possible that the glycoproteins fail to locate to the membrane when expressed without the entire casette, which would likely lead to protein degradation.

In order to evaluate this, several constructs will be produced that should allow this to be explored. First, wild-type E2 and E1 will be fused to GFP in sequence and transfected into 293T cells. If GFP, and, by assumption, the glycoproteins, are successfully produced, the transfected cells will be detectable by fluorescent microscopy.



Second, the entire casette containing E3 to E1 will be placed in an expression construct.

Figure 2; Coomassie stain of SDS-15%-PAGE gel containing lysates of 293T cells transfected with expression plasmids of each protein listed above each lane. For each CHIKV protein, three lanes were run: undigested, PNGase F digested, and EndoH digested (left to right). The 36 kDa band in each PNGase F lane is the endoglycosidase.

This will be transfected as above and tested by SDS-PAGE, Western blotting, and Coomassie staining to determine whether wild-type E2 and E1 are successfully being expressed in this way. If so, then the predominant oligosaccharide will be characterized as described above. Several monoclonal antibodies against CHIKV E2 and E1 will also be tested against E2 and E1 produced in this way.

Third, the Chikungunya 181/25 vaccine strain will be produced from molecular clones containing the various combinations of glycosylation site ablations by first transcribing the viral RNA *in vitro* followed by electroporating the RNA into BHK cells. This analysis will then be repeated with BHK and C6/36 produced virus.

Finally, this casette (E3 to E1) will be placed in a lentiviral vector for expression. Similarly to the above expression constructs, the expression of wild-type E2 and E1 will first be analyzed by SDS-PAGE, Western blotting, and Coomassie staining, followed by identification of the predominant oligosaccharide if expression if successful.

In conclusion, it appears that the CHIKV E2 and E1 glycoproteins were not successfully expressed by these constructs, and new approaches will be necessary to express the proteins in order to determine the predominant oligosaccharides present at each N-Linked glycosylation site. Once the proteins are successfully produced, the BiFC constructs will be made and experiments regarding the E2/Human Prohibitin-1 interaction and the role of glycans in this interaction can begin.

---

[i] Fields, Bernard N, David M. Knipe, Peter M. Howley, and Diane E. Griffin. *Fields Virology*. Philadelphia: Lippincott Williams & Wilkins, 2001. Print.

[ii] "Chikungunya." . World Health Organization, n.d. Web. 28 Nov 2012.a <http://www.who.int/mediacentre/factsheets/fs327/en/>.

[iii] Khan, A. H., K. Morita, M. C. Parquet, F. Hasebe, E. G. M. Mathenge, and A. Igarashi. 2002. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. J. Gen. Virol. 83:3075-3084.

[iv] Brooks, Susan, M Dwek, and Udo Schumacher. *Functional and Molecular Glycobiology*. 2002.

[v] Burda P, Aebi M (January 1999). "The dolichol pathway of N-linked glycosylation". *Biochim. Biophys. Acta* 1426 (2): 239–57

[vi] Drickamer K, Taylor ME (2006). *Introduction to Glycobiology* (2nd ed.). Oxford University Press, USA.

[vii] Butters, T. D., R. C. Hughes, and P. Vischer. 1981. Steps in the biosynthesis of mosquito cell membrane glycoproteins and the effects of tunicamycin. Biochim. Biophys. Acta 640672-686

[viii] Unpublished. McGee, Charles, Bronwyn Gunn, Mark Heise.

[ix] Klimstra, WB, KD Ryman, et al. "DC-SIGN and L-SIGN can act as attachment receptors for alphaviruses and distinguish between mosquito cell- and mammalian cell-derived viruses." *Journal of Virology*. 77.22 (2003): n. page. Print.

[x] Shabman, Reed, Mark Heise, et al. “Differential induction of Type I interferon responses in myeloid dendritic cells by mosquito and mammalian-cell-derived alphaviruses.” *Journal of Virology*. 2007 Jan;81(1):237-47. Epub 2006 Nov 1.

[xi] Unpublished. Gunn, Bronwyn, Mark Heise.

[xii] Wintachai, P., N. Wikan, et al. "Identification of Prohibitin as a Chikungunya Virus Receptor Protein." *Journal of Medical Virology*. 84 (2012): 1757-1770. Print.

[xiii] Zhang L, *et al*. Asian Pac J Cancer Prev, 2012. PMID 23098474.

[xiv] Najm MZ, *et al*. Asian Pac J Cancer Prev, 2012. PMID 23244120.

[xv] Kuadkitkan A, Wikan N, Fongsaran C, Smith DR. 2010. Identification and characterization of prohibitin as a receptor protein mediating DENV-2 entry into insect cells. Virology 406:149–161.

[xvi] Cabantous, S., G.S. Waldo, et al. "A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association." *Scientific Reports*. (2013): n. page. Print.

[xvii] Cabantous, S., G.S. Waldo, et al. "A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association." *Scientific Reports*. (2013): n. page. Print.

[xviii] Sasse, Joachim, and Sean Gallagher. "Staining Proteins in Gels." *Current Protocols in Molecular Biology*. (2001): n. page. Web. 3 Dec. 2013.