A NOVEL STRATEGY FOR EXPRESSING RECOMBINANT HCV GLYCOPROTEINS IN CELL CULTURE: TOWARD BIOCHEMICAL, BIOPHYSICAL, AND IMMUNLOGICAL STUDIES

by

JILLIAN L. WHIDBY

A Dissertation submitted to the
Graduate School-New Brunswick
Rutgers, The State University of New Jersey
And
The Graduate School of Biomedical Sciences
University of Medicine and Dentistry of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Microbiology and Molecular Genetics
written under the direction of
Joseph Marcotrigiano, Ph.D.
and approved by

__________________________________

__________________________________

__________________________________

__________________________________

New Brunswick, New Jersey
October, 2012
Almost 4 million people are infected with hepatitis C virus (HCV) in the United States alone, with 170 million infected worldwide. It is the leading cause of liver transplantation in the US. Although infection is initially asymptomatic, HCV often leads to chronic liver disease, cirrhosis, and/or liver cancer. Many cases are treatable with combination therapy (interferon, ribavirin, and new protease inhibitors), but efficacy is dependent on the infecting strain and there is currently no vaccine. Without more effective antiviral and immunological treatments, the Centers for Disease Control and Prevention (CDC) predicts that deaths due to HCV will double or triple in the next 15 to 20 years due to prolonged disease and continued spread. The high prevalence of infection, lack of highly effective HCV-specific inhibitors, and poor response rate to the current treatment underscore the importance of developing new therapeutic strategies.

The mechanism of viral entry is an important subject of study with respect to preventing and treating infection. Two of the four HCV structural proteins, envelope 1 (E1) and
envelope 2 (E2), heterodimerize on the surface of the virion. Experimental evidence supports the roles of E1 and E2 in receptor binding, virus-cell fusion, and entry into the host cell. These factors make E1 and E2 key determinants of pathogenicity and optimal targets of vaccine design. **We hypothesize that a thorough biophysical understanding of E2, as well as improvements in the available biochemical tools for the study of HCV E2 will provide a significant advancement in the understanding of viral infection.**
ACKNOWLEDGMENTS

I would first like to express my gratitude to Dr. Joseph Marcotrigiano for being an energetic, inspiring, and dedicated advisor. His consistent encouragement has led me through many years of highs and lows. As a founding member of the Marcotrigiano lab, I am truly thankful to have had the opportunity to learn and grow as a scientist amongst a group of outstanding individuals. Fuguo Jiang, Gyehwa Shin, Samantha Yost, Ankita Basant, and Abdul Khan, and our wonderful undergraduates... you are not just colleagues, you are family for life.

Additional thanks to my committee members, Drs. Aaron Shatkin, Eddy Arnold, Victor Stollar, and Ann Stock. Your guidance and support has been unwavering. Also thanks to Dr. Martin Yarmush for his support and encouragement.

To our collaborators, Drs. Arash Grakoui, John Shires, Charles Rice, Michael Houghton, and Roberto Cattaneo (among many others), thank you for believing in our work and for helping to elevate it to new heights.

Present and former colleagues and friends at CABM, you guys ROCK. I trust that you know who you are, because I could write another thesis entirely just listing your names and how you have contributed to my great experience here. Special thanks to Drs. Peter Lobel and Gaetano Montelione and their lab members for collaborative work and exceptional advice. And to Jim Aramini for carpools and occasional crash lectures on NMR. And to Gail Ferstandig Arnold for always helping to lead me in the right direction.
I have been blessed with wonderful mentors throughout my educational experience and would like to especially thank Dr. Cristina Iftode and Dr. Andrea Bertolotti-Ciarlet. Your work ethic and guidance has been more of an inspiration to me than you will ever know.

Thank you to my closest friends, Carolyn Romanoski, Laura Contarino, and Orly Even-Dar. You have stood by me, tolerated my flakiness, and cheered the loudest.

My husband Mark, you have always been supportive and encouraging and my steady rock in my weakest moments over the last 6 years. It is truly extraordinary to share this with you. Start to finish.

Thanks to my parents, for making me believe that I could do anything I wanted to do, then not letting me quit when I tried. Without you, I wouldn’t be standing here today, and I don’t just mean that in the obvious biological sense. 😊 And to Mommom, you listen, you are patient, you never judge. That is priceless. And also thanks to my in-laws for their patience, wisdom, support, and love.

To Vincent, Mitchell, and Skyler: I don’t care how tall you are, you will always be my little brothers, the ones who taught me how to be a leader and set an example. Do what makes you happy in life and..........Don’t just leave, leave it better!
DEDICATION

This thesis is dedicated in memory of Dr. Aaron J. Shatkin. As one of my committee members, his door was always open when I needed him and his positive attitude and dedication never wavered. His impact on my graduate career will last a lifetime and his guidance throughout the course of this thesis was immeasurable.
TABLE OF CONTENTS

Abstract ............................................................................................................................. ii

Acknowledgements .......................................................................................................... iv

Dedication .......................................................................................................................... vi

Table of Contents ............................................................................................................. vii

List of Figures .................................................................................................................. xi

Introduction ......................................................................................................................... 1

1. HCV Epidemiology ......................................................................................................... 2
2. HCV Virology .................................................................................................................. 3
   a. Genome and Replication .......................................................................................... 3
   b. Envelope Glycoproteins ......................................................................................... 4
   c. Entry Receptors ....................................................................................................... 6
3. Glycoprotein Significance ............................................................................................ 6

Rationale ............................................................................................................................. 8

Materials/Methods ............................................................................................................ 9

Section I. Biophysical Analysis of HCV eE2 ..................................................................... 9

1. Production of Stable Cell Lines ................................................................................... 9
2. ELISA for eE2-Fc ........................................................................................................ 9
3. Western Blot for eE2-Fc ............................................................................................ 10
4. Expression and Purification of eE2 and eE2-C656S ..................................................... 10
5. Deglycosylation of eE2 ............................................................................................. 11
6. Mapping the N-linked Glycosylation Sites via Mass Spectrometry .............................. 11
7. Gel Filtration Analysis of eE2 and eE2-C656S .......................................................... 12
8. Free Cysteine Analysis ..................................................................................................... 13
9. Analytical Ultracentrifugation ......................................................................................... 13
10. Circular Dichroism .......................................................................................................... 14
11. eE2 ELISA Using Human sera ....................................................................................... 15
12. HCVcc Infection in the presence of purified proteins ..................................................... 16
13. Cytotoxicity .................................................................................................................. 16
14. Expression and Purification of GST and GST-CD81-LEL ............................................. 16
15. Anti-Fc CD81 Binding Assay .......................................................................................... 17

Section II. eE2 Monoclonal Antibodies for Biochemical Study ..................................... 19

1. Generation of Hybridoma Monoclonal Cultures ......................................................... 19
2. ELISA to test eE2 Monoclonal Antibodies .................................................................. 19
3. Western Blot to test eE2 Monoclonal Antibodies ....................................................... 20
4. eE2 Truncation Mutants ............................................................................................... 20
5. eE2 Truncation Mapping ELISA .................................................................................. 21
6. Purification of eE2 Monoclonal Antibodies .................................................................. 22
7. Anti-eE2 (3G4) CD81 Binding Assay ........................................................................... 23
8. Monoclonal Antibody Neutralization Assay ............................................................... 24

Section III. Examination of Disulfide Bonds in HCVcc E2 and eE2 .............................. 25

1. Site-directed Mutagenesis of HCV eE2 Cysteines and C6A Regional Mutants .... 25
2. Purification of eE2-WT, eE2- C6A, and eE2-C11A ....................................................... 26
3. Circular Dichroism ......................................................................................................... 26
4. CD81 Binding Assay ....................................................................................................... 27

Section IV. Production of Recombinant Protein Using Lentiviral Expression Vectors ...... 28

1. Lentiviral Constructs and Cloning ................................................................................ 28
2. Lentiviral Co-transfection ............................................................................................. 28
3. Lentiviral Harvest and Infection ..................................................29
4. BelloCell Operation .................................................................30

Section V. Optimization of eE2 for Crystallization Studies ..................31

1. Generation of Amino-terminal Truncations ..................................31
2. Limited Proteolysis .......................................................................31
3. Production of Endoglycosidase H ...............................................32
4. Deglycosylation with Endoglycosidase H ....................................33
5. Purification of Deglycosylated eE2 ..............................................33
6. Generation of eE2 Genotypes ......................................................34
7. Purification of Glycosylated eE2 ..................................................35

Experimental Results ......................................................................36

Section I. Biophysical Analysis of HCV eE2 ....................................36

1. Expression of J6 eE2-Fc in HEK293T Cells ..................................38
2. eE2 Glycosylation Analysis ........................................................40
3. eE2 Oligomerization and Secondary Structure ..........................44
4. eE2 is Recognized by Antibodies from HCV Infected Patients .......53
5. eE2 Blocks HCVcc Entry ............................................................55
6. Discussion ..................................................................................57

Section II. eE2 Monoclonal Antibodies for Biochemical Study ..........64

1. Production and Testing of Anti-eE2 Monoclonal Antibodies ..........66
2. Anti-eE2 Antibodies that are Cross-Reactive to H77-eE2 ..............66
3. Epitope Mapping and Purification of Anti-eE2 Antibodies ............72
4. Anti-eE2 Antibody Applications ...................................................76
5. Discussion ..................................................................................79
### Section III. Examination of Disulfide Bonds in HCVcc E2 and eE2  ........................................83

1. Disulfide Bond Mapping Using Mass Spectrometry ..................................................84
2. A Study of HCV E2 Disulfide and Regional Mutants ................................................85
3. HCV eE2-C6A and Surrounding Mutants Influence hCD81 Binding  .......................86
4. eE2-C6A is Defective in Blocking HCV Infection ...................................................89
5. Discussion ..................................................................................................................89

### Section IV. Production of Recombinant Protein Using Lentiviruses  .......................95

1. Lentivirus Expression of eE2‐PA ...........................................................................98
2. Lentivirus Expression of eE1‐PA ..........................................................................102
3. Discussion ..............................................................................................................103

### Section V. Optimization of eE2 for Crystallization Studies  ..................................106

1. Reduction of E2 Disordered Regions: HDX and Limited Proteolysis ...................107
2. Deglycosylation of eE2 Reveals Two Species .......................................................112

**Conclusion** .............................................................................................................124

**References** ...........................................................................................................128

**Abbreviations** .......................................................................................................142
LIST OF FIGURES

Figure 1. eE2 Schematic and Sequence .................................................................39

Figure 2. eE2 Western and Purification .................................................................41

Figure 3. Deglycosylation of eE2 with PNGaseF and EndoH ..................................43

Figure 4. Mapping the N-Linked Glycosylation Sites ............................................45

Figure 5. eE2 Oligomeric Species ...........................................................................46

Figure 6. Differential Labeling of Free and Disulfide-linked Cysteines ..................48

Figure 7. eE2-C656S Oligomeric Species ...............................................................50

Figure 8. Analytical Ultracentrifugation Data for eE2-C656S at pH 7 .......................52

Figure 9. CD Spectroscopy of eE2 (A) or eE2-C656A (B) at pH 7 and pH 5 ...............54

Figure 10. Functional analysis of eE2 and eE2-C656S .............................................56

Figure 11. Enzyme-linked immunoassay for CD81-LEL binding .............................58

Figure 12. anti-E2 Monoclonal Western Blots .......................................................69

Figure 13. H77 anti-E2 Monoclonal Western Blots ...............................................71

Figure 14. H77 Cross-Reactivity ELISA ..................................................................73

Figure 15. Monoclonal Antibody Truncation Map ..................................................75

Figure 16. SDS-PAGE for Purified Monoclonal Antibodies ....................................77
Figure 17. Monoclonal 3G4 CD81-Binding ELISA .......................................................... 78

Figure 18. eE2 Monoclonal Antibodies Neutralize Virus .............................................. 80

Figure 19. Mutations Affect hCD81 Binding but Secondary Structure is Conserved ....... 88

Figure 20. eE2-C6A is Impaired in Blocking Infection ................................................... 90

Figure 21. SDS-PAGE eE2-ProtA Comparison .............................................................. 100

Figure 22. pJG Vector Features and GFP Expression .................................................... 101

Figure 23. eE1-ProtA Purification ............................................................................. 108

Figure 24. HDX Results for eE2 ............................................................................... 109

Figure 25. eE2 Truncation Clone Design ..................................................................... 110

Figure 26. Gel Filtration Profiles for eE2 Truncations ................................................ 111

Figure 27. Limited Proteolysis for eE2-NTD4 ............................................................ 113

Figure 28. Purified eE2-NTD4 Deglycosylated with EndoH ........................................ 115

Figure 29. Sequence Comparison of Various HCV Genotypes .................................. 117

Figure 30. Deglycosylated eE2-NTD4 for Various Genotypes ...................................... 118

Figure 31. Gel Filtration Results for Various eE2 Genotypes ....................................... 119

Figure 32. eE2 Quasi-crystalline Material .................................................................. 120

Figure 33. Mass Spectrometry Results for Deglycosylated eE2–NTD4 ....................... 122
Figure 34. eE2 Elicits a Cross-Neutralizing Response......................................................... 126
INTRODUCTION

More than 35 years ago, an article was published reporting the high incidence of post-transfusion hepatitis in spite of rigorous blood donor screenings for hepatitis B antigen (4). At this time, researchers and clinicians were referring to this type of hepatitis as “non-A, non-B hepatitis” (NANBH). The cause of NANBH remained undefined for more than ten years, when Harvey Alter of the Department of Transfusion Medicine at the National Institutes of Health (NIH), David Bradley of the CDC, and Michael Houghton of Chiron Corporation collaboratively developed the technique that identified the virus (19). They produced a bacteriophage cDNA library using RNA from infected chimpanzee plasma. After screening about one million cDNA clones, Bradley and Houghton found a single clone that would not hybridize to either the human or chimpanzee genome. They confirmed that it was derived from an exogenous RNA molecule up to 10,000 bp in length and was associated only with infected individuals. This RNA was from the genome of a virus that they officially named “hepatitis C virus”. In the same issue of Science, published in April of 1989, Alter, Bradley, and Houghton described an assay for identifying HCV in infected patients (71). They used their recombinant cDNA clone to identify HCV specific antibodies circulating in the blood. This assay has been in use since June of 1992 for testing of patients and screening of blood donor samples.
1. HCV Epidemiology

Since its initial discovery, HCV has been found in all parts of the world, with 7 major genotypes and more than 50 subtypes isolated. Currently, 3% of the total human population is infected- three to four times more people than are infected with HIV-making HCV a serious global health problem (100, 115). In the United States, HCV infection is the leading cause of liver transplantation and resulted in more than 17,000 deaths in 2007, a number that is still on the rise (1). Although HCV infection is usually asymptomatic, more than 85% of patients will develop chronic infection (1). This often leads to chronic liver disease, cirrhosis, and sometimes hepatocellular carcinoma, though many patients are never aware are infected. In 2010, the National Academies Institute of Medicine organized a committee and published a report describing a national strategy for hepatitis awareness and prevention, citing the poor HCV awareness among healthcare providers and professionals and low prevalence of HCV testing recommendations (90). Finally in August 2012, the United States Centers for Disease Control and Prevention published a historic recommendation that all Americans born between 1945 and 1965 (“baby boomers”) get tested for HCV, as they represent 73% of HCV deaths in the United States (117). This represented true acknowledgment of the hepatitis epidemic and one of the first major steps toward control and prevention.

Despite decades of HCV research into new therapies, only about half of HCV cases in the U.S. are treatable with combination therapy (pegylated interferon-α with ribavirin),
which can be less effective for illness caused by other genotype strains throughout the world. Recently, two new protease inhibitors were approved by U.S. Food and Drug Administration (FDA) in 2011 for treatment of patients with genotype-1 infection. These compounds, called telaprevir (Vertex) and boceprevir (Merck), are highly effective but quickly promote the emergence of resistant strains. There still remains a need for a combined interferon-ribavirin therapy approach (44). While studies continue to demonstrate the possibility of an interferon-free treatment regimen, a vaccine remains elusive.

2. HCV Virology

2.a Genome and Replication

Hepatitis C virus is the only member of the Hepacivirus family within the *Flaviviridae*. Its genome consists of a 9.6 kb uncapped, positive RNA strand with a single open reading frame. Cap-independent translation via the 5' IRES produces one long polyprotein, in which the structural proteins are amino-terminal and the non-structural proteins are carboxy-terminal (81). The mature proteins, four structural and six non-structural, are cleaved co-translationally or post-translationally by cellular (signal peptidase and signal peptide peptidase) or virus-encoded (NS2-3 and NS3-4A) proteases. The structural proteins include capsid, envelope protein 1 (E1), envelope protein 2 (E2), and the p7 ion channel, while the remaining non-structural proteins function in replication of the viral genome. Upon binding to the target cell, infection proceeds by clathrin-coated
endocytosis and endosomal acidification, suggesting that fusion of the viral envelope with cellular membranes is a pH-triggered event (57, 80, 120, 133). Replication occurs cytoplasmically in association with perinuclear and endoplasmic reticulum (ER) membranes and requires the synthesis of negative strand RNA, which provides the template for the synthesis of more positive strand RNA. It is thought that when genomic RNA becomes encapsulated by core, these nucleocapsids can then bud into the ER, deriving the lipid envelope with embedded glycoproteins from the ER membrane. These particles travel through the secretory pathway and are released at the cell membrane in a non-lytic process. Evidence of viral replication also exists throughout the bloodstream, with a chronically infected patient producing up to $10^{12}$ virions per day. Recent progress in the replication of hepatitis C virus in cell culture has greatly increased our understanding of the virus life cycle (80, 133), with chimeric and adapted genomes now available for all genotypes (49, 111).

2.b. Envelope Glycoproteins

E1 and E2 are type I transmembrane proteins that form the outermost layer of the virus particle. Within the individual proteins, the amino-terminal ectodomains are soluble, while the carboxy-terminal hydrophobic domains become inserted into the envelope lipid bilayer (94). These hydrophobic domains are thought to be involved in ER retention (21, 22, 40) and in the formation of noncovalent E1/E2 heterodimers (23, 95). The ectodomains have been previously defined as the minimal constructs that result in
secretion of properly folded protein (89). Both E1 and E2 are glycosylated and contain intramolecular disulfide bonds (35, 48, 89). Folding of these proteins requires ER chaperones, particularly calnexin (35). For these reasons, overexpression of these proteins often results in misfolding and intermolecular disulfide-linked aggregates (33).

As the most exposed proteins on the surface of the virus, the roles of E1 and E2 independently and as a heterodimer are extensive. The heavy glycosylation of these proteins is likely to be critical for proper folding and transport through the secretory pathway and probably provides an effective “disguise” for escape from the host immune response. Additionally, the process of viral entry is thought to involve a physical interaction between the E1/ E2 heterodimer and receptors on the cell surface. It is thought that E1 may be responsible for fusion, since a region of the protein resembles the fusion peptides of other flaviviruses (41). However, two potential fusion sequences have been proposed in E2 based on impaired fusogenicity of mutant HCV pseudoparticles (74). Based on the structures of related flavivirus glycoproteins, the HCV glycoproteins are predicted to be in the class II fusion category, containing mainly beta sheet structure (62). An HCV E2 model was published in 2010 which uses the flavivirus envelope (E) protein structure combined with an experimentally determined disulfide-bonding map to thread the E2 sequence into a predicted structure containing mostly beta sheet and random coil (70).
2.c. Entry Receptors

Many cellular receptors have been implicated in the entry of HCV. A review published in 2008 highlights evidence in support of roles for glycosaminoglycans (including heparin sulfate) as well as low-density lipoprotein receptor, an interaction attributed to HCV’s association with very low density lipoprotein (54). However, four specific receptors have been identified as playing direct roles in entry: CD81 (102), Claudin-1 (37), scavenger receptor class B type 1 (SR-BI) (110), and occludin (103). Importantly, Ploss et al. describe a thorough analysis of these receptors, implicating species tropism for CD81 and occludin (103). Their work has possibly provided the last major piece of the entry puzzle and resulted in the development of a new mouse model for HCV infection, a critical breakthrough toward the practical study of HCV pathogenesis and immunity (31).

3. Glycoprotein Significance

For over a decade, research has pointed to the roles of E1 and E2 as the primary determinants of pathogenicity. However, many researchers and pharmaceutical companies avoid the glycoproteins as targets for developing treatments or vaccines due to their high variability and mutation tolerance. Nonetheless, it has been shown that chimpanzees immunized with E1/E2 heterodimeric proteins are protected from infection with low doses of homologous hepatitis C virus (18). In 1991, researchers
identified hypervariable regions in E2, and it was later shown that deletion of hypervariable region 1 attenuated infection in chimpanzees (43, 125). Youn et al. did further work in chimpanzees and showed that an E2 antibody response correlates with lower viral titres (130). Researchers have also found that rodents injected with HCV envelope glycoproteins produce antibodies that are broadly cross-reactive in their neutralization properties (118). Importantly, Michael Houghton published a review in 2011 emphasizing the potential value of the envelope glycoproteins in generating a vaccine. He points out that the prime-boost regimen with E1 and E2 can elicit cross-reactive neutralizing antibody responses in the chimpanzee model, indicating the existence of conserved immunogenic epitopes across genotypes (56).
RATIONALE

Following the elucidation of the structure of the HIV glycoprotein (gp41) in 1998, one
group of scientists designed a small peptide which inhibits virus fusion to the cell
membrane (128). By binding to the gp41 stalk, a conformational change is inhibited and
the virus is prevented from entering the target cell. Furthermore, the annual influenza
vaccine includes three variants of the surface antigen, hemagglutinin (HA), another
highly variable and mutable glycoprotein (17). These two examples provide justification
for the exploration of HCV E2 as a potential therapeutic target. Here, we produced a
functional, native E2 protein in sufficient quantities for biophysical and biochemical
characterization and have set a strong foundation for crystallization trials. This
information provides valuable insight toward better treatments or vaccine candidates
for HCV infected patients.
MATERIALS AND METHODS

Section I. Biophysical Analysis of HCV eE2

1. Production of Stable Cell Lines

HEK293T cells were cultured in Dulbecco’s Modified Eagle Medium supplemented with 10% FBS. A 6-well plate was seeded with 0.5x10^6 cells per well and the pPro-eE2-Fc vector (from J6 strain) was transfected the following day using FuGene-HD (Roche Diagnostics, Basel, Switzerland). After three days, the cells were expanded to 150 cm^2 dishes and placed under hygromycin (Calbiochem, San Diego, CA) selection at 75 μg/mL. Individual colonies were selected, expanded, and tested for eE2-Fc expression using an anti-Fc ELISA.

2. ELISA for eE2-Fc

MaxiSorp plates (Thermo Fisher Scientific, Rochester, NY) were coated with 100 μL of eE2-Fc cell culture supernatant for 2 hours at room temperature. The wells were washed 3x with 200 μL of PBS + .05% Tween-20 (PBS-T), then blocked with 200 μL of 2% BSA in PBS-T for 1 hour at room temperature. After three more washes with PBS-T, 100 μL of goat anti-Fc antibody (Pierce, Thermo Fisher Scientific, Rochester, NY) at 1:15,000 dilution in PBS-T was incubated for 1 hour at room temperature. The ELISA was developed with TMB substrate (Thermo Scientific) and quantified using the SpectraMax 250 plate reader and SOFTMax 2.6 software.
3. Western Blot for eE2-Fc

Supernatants from mock or eE2-expressing cell lines were boiled in 2x sample buffer containing β-mercaptoethanol and run on a 10% SDS-PAGE gel with 19:1 acrylamide:bisacrylamide. The gel was transferred to a nitrocellulose membrane and blocked for 1 hour in 5 % non-fat dry milk in PBS-T. The blot was then incubated in peroxidase-conjugated goat anti-human IgG (Fcγ) (Thermo Scientific) at 1:1,000 dilution in 5 % non-fat dry milk in PBS-T for 1 hour at room temperature. After 3 ten-minute washes in PBS-T, the blot was developed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

4. Expression and Purification of eE2 and eE2-C656S

Supernatants from stable cell lines expressing either eE2-Fc or eE2-C656S-Fc were harvested, centrifuged to remove cellular debris, and filtered through a 0.22 μm membrane. The eE2-Fc (or eE2-C656S-Fc) protein was applied to protein A-conjugated resin (GE Healthcare, Piscataway, NJ) overnight with gentle rocking. The resin was washed with buffer (50 mM HEPES pH 7.5, 150 mM KCl, 5 % glycerol), and incubated with thrombin protease to release the protein from the Fc tag. After cleavage, the protein eluate was consolidated and the concentration determined by Bradford Assay, yielding 1-2 mg of eE2 per liter of supernatant.
5. Deglycosylation of eE2

eE2 was deglycosylated using either EndoH or PNGaseF according to the manufacturer’s protocol (New England Biolabs, Ipswich, MA). Briefly, 20 μg of eE2 was denatured and 10 U of EndoH or PNGaseF was added. The reaction was incubated at 37°C for 1 hour and analyzed by SDS-PAGE followed by Coomassie staining.

6. Mapping the N-linked Glycosylation Sites via Mass Spectrometry

The eE2 protein sample was denatured in 6 M urea, then reduced with 10 mM DTT for 30 min at 60°C. After denaturation, 20 mM iodoacetamide was added to alkylate sulfhydryl groups and the reaction was incubated in the dark for one hour at room temperature. Following this treatment, the sample was buffer-exchanged into 50 mM NH₄HCO₃. The sample was digested using either sequencing grade trypsin (Promega, Madison, WI) or chymotrypsin (Roche Diagnostics) according to the manufacturer’s protocol. Digested samples were dried via vacuum centrifugation and reconstituted in 50 mM NH₄HCO₃. Samples were deglycosylated with 50 U of PNGaseF (New England Biolabs, MA) by incubation for 1 hour at 37°C and the reaction was stopped with 0.1% trifluoroacetic acid (TFA).

All LC-MS/MS experiments were performed using the U3000 (Dionex, Sunnyvale, CA) in nano-LC mode on line with LTQ (Thermo Fisher Scientific). Samples were first solubilized in 0.1% TFA and loaded onto a 75 μm x 12 cm emitter column self-packed with Magic
C18AQ, 3 µm 200 Å (Michrom Bioresources Inc, Auburn, CA). The sample was eluted using a linear gradient from 98 % of 0.1 % formic acid in water to 45 % of 0.1 % formic acid in acetonitrile over 30 min. Mass spectrometry data was acquired using a data-dependent acquisition procedure with a full scan cyclic series. This was followed by zoom scans and MS/MS scans of the five most intense ions with a repeat count of two and a dynamic exclusion duration of 60 sec.

The LC-MS/MS data was searched against a human database using a local version of the Global Proteome Machine (local implementation of the Global Protome Machine (24). Carbamidomethylation of cysteine was used as the fixed modification, while oxidation of methionine and deamination of asparagine were used as potential modifications.

Manual interpretation and peak integration was performed on all peptide peaks covering potential glycosylation sites (NXT/S). (H. Zheng).

7. Gel Filtration Analysis of eE2 and eE2-C656S

Purified eE2 protein was loaded onto a Superdex200 10/300 GL size exclusion column (GE Healthcare, Piscataway, NJ) equilibrated with HEPES buffer (50 mM HEPES pH 7.5, 150 mM KCl, 5% glycerol) using the AKTA Purifier (GE).
8. Free Cysteine Analysis

To label free cysteines, the protein sample was incubated with a 20-fold molar excess of N-ethylmaleimide (NEM) and 6 M guanidine-HCl at room temperature for 1 hour in the dark. The sample was then buffer exchanged to 6 M guanidine-HCl using a spin filter and washed three times with 400 μl of 6 M guanidine-HCl to remove the NEM. Disulfide bonds were reduced by adding 10 mM DTT at 60°C for 30 min. The newly generated free sulfhydryl groups were alkylated with 20 mM iodoacetamide (IAM) at room temperature for one hour in the dark. After buffer exchange into 50 mM NH₄HCO₃, the sample was digested with trypsin protease at 37°C overnight. The sample was then deglycosylated with PNGaseF (100 U) at 37°C for 3 hours and acidified prior to LC-MS/MS analysis.

The LC-MS/MS data were searched using Sequest against the sequence of the target protein. +57 Da (alkylation by IAM) and +125 Da (alkylation by NEM) on cysteine, oxidation of methionine (+16 Da), and deamination of asparagine (+1 Da) were used as potential modifications. The identification was confirmed manually (H. Zheng).

9. Analytical Ultracentrifugation

All sedimentation experiments were performed with a Beckman Optima XL-I at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center at San Antonio. Sedimentation velocity data
were analyzed with UltraScan (26) version 9.9 (25). All measurements were made at 230 nm in intensity mode, at 20°C, and at 37,000 rpm, using standard upon 2-channel centerpieces. All samples were measured in 25 mM sodium phosphate buffer containing 50 mM KCl, adjusted to pH 7.0. Concentration dependency of the sedimentation data was assessed by sedimenting the sample at both high concentration (~0.8 optical density (OD) at 230 nm) and at low concentration (~0.25 OD at 230 nm). Hydrodynamic corrections for buffer density and viscosity were made according to methods outlined in Laue et al. as implemented in UltraScan (73). The data were analyzed by 2-dimensional spectrum analysis (2DSA) (12) using the ASTFEM-RA solution (14) with simultaneous removal of time-invariant noise. Molecular weight and shape distributions obtained in the 2DSA were further refined by Monte Carlo analysis (27). Composition comparisons were made by overlaying sedimentation coefficient distributions derived from the van Holde - Weischet analysis (28). The calculations were performed on the Lonestar cluster at the Texas Advanced Computing Center at the University of Texas at Austin, and at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio.

10. Circular Dichroism

The protein sample was buffer-exchanged into 25 mM sodium phosphate pH 5.0 or 7.0, and 50 mM KCl. The CD spectra in the wavelength range of 195-260 nm were measured at 0.5 nm intervals on an Aviv spectropolarimeter model 400 (Lakewood, NJ) at 25°C. A
quartz cell with a path length of 0.1 cm was used. The data are presented in millidegrees.

11. eE2 ELISA Using Human Sera

96-well ELISA plates (Corning, Lowell, MA) were coated with 100 µl of a 1 µg/mL solution of eE2 in NaHCO₃ overnight at 4°C. The plates were washed twice with 200 µl/well PBS-T, then blocked with a 10% solution of normal goat serum in PBS-T (Jackson ImmunoResearch, West Grove, PA) for 1 hour at 37°C. Human serum was isolated from whole blood samples (Emory University School of Medicine, PI Arash Grakoui, IRB# 1358-2004) collected in SST tubes (Becton Dickenson, Franklin Lakes, NJ) via centrifugation and frozen in aliquots at -80°C. Ten-fold serial dilutions were made for each serum sample using binding buffer composed of 0.1% normal goat serum in PBS-T. 100 µl of diluted sample was added to each well of the plates and incubated for 90 minutes at room temperature. The plates were washed eight times with PBS. 100 µl of goat anti-human IgG-Biotin conjugate (Biosource, Camarillo, CA) diluted 1:20,000 in binding buffer was added and allowed to incubate for 90 minutes at room temperature. Finally, 100 µl of streptavidin-HRP conjugate (Biosource) was added to each well at a 1:2,000 dilution and incubated for 45 min at room temperature. Using TMB substrate solution (Ebioscience, San Diego, CA), absorbance was measured using a VersaMax Microplate reader and SoftMax Pro software (Molecular Devices, Sunnyvale, CA).
12. HCVcc Infection in the presence of purified proteins

Approximately 100 TCID\textsubscript{50} of Cp7 (J6/JFH-1 recombinant) viruses (86) were incubated with two-fold dilutions of the purified eE2, eE2-C656S, GST, GST-CD81LEL or GST-mCD81 starting at 200 µg/mL. 6.0x10\textsuperscript{3} Huh-7.5 cells were seeded into a collagen-coated 96-well plate. The virus–protein mixture was incubated with the cells for three days at 37°C. Cells were stained by immunohistochemistry as previously described (86) (G. Mateu).

13. Cytotoxicity

Huh-7.5 cells were incubated with various dilutions of the purified proteins as described above. Three days later, cells were washed twice with PBS, harvested by trypsinization, and resuspended in 100 µl of PBS. Cells were stained with BD Via-Probe\textsuperscript{™} (BD Biosciences, San Jose, CA) according to the manufacturer’s instructions and counted using FACSCalibur (BD Biosciences) equipment and FlowJo (v8) analysis software (G. Mateu).

14. Expression and Purification of GST and GST-CD81-LEL

CD81-LEL was expressed with an amino-terminal GST tag and carboxy-terminal histidine tag. The GST tag alone was expressed and purified using the same method. Each plasmid was transformed into Rosetta-Gami cells (EMD, Darmstadt, Germany) and plated on agar containing ampicillin. A single colony was used to inoculate a 25 mL
starter culture of LB + ampicillin on the following day, which was shaken overnight at 37°C. On the third day, 10 mL of the starter culture was added to 1 L of LB + ampicillin. The culture was shaken at 225 rpm at 30°C until reaching an OD of ~0.5 at 600 nm. The culture was cooled for 1 hour at 4°C, then induced with 300 nM IPTG. The culture was shaken overnight at 200 rpm at 18°C. The following morning, the bacteria were pelleted at 5,000 rpm for 20 min and resuspended in 15 mL of buffer containing 50 mM HEPES pH 7.5 + 1 M KCl + 1 mM EDTA. Lysis was done via homogenization, followed by a second round of pelleting at 18,000 rpm for 40 min to clarify the lysate. A 5 mL GST column (GE, Piscataway, NJ) was equilibrated with 50 mM HEPES pH 7.5 + 150 mM KCl. After loading the bacterial lysate onto the column, the same buffer was used for thorough washing. GST-CD81 was eluted from the column using 100 mM HEPES pH 7.5 + 150 mM KCl + 15 mM reduced glutathione. Finally, GST-CD81 was dialyzed overnight into 50 mM HEPES pH 7.5 + 150 mM KCl + 20% glycerol and flash frozen in aliquots.

15. Anti-Fc CD81 Binding Assay

The protocol used for this assay was originally described by Flint et al. (42). GST-human CD81-LEL was expressed and purified as described above. 96-well plates (Nalgene Nunc, Thermo Fisher Scientific, Rochester, NY) were coated with 50 µg/mL of GST-CD81-LEL overnight at 4°C. Plates were washed 3x with PBS-T and blocked with 3% BSA in PBS-T for 1 hour at room temperature. 100 µL of supernatant from cells stably expressing eE2-Fc and eE2-C656S-Fc were added to appropriate wells and incubated overnight at
4°C. On the third day, the wells were washed 3x with PBS-T and incubated with a 1:15,000 dilution of anti-Fc HRP-conjugated antibody (Thermo Scientific) for 1 hour at room temperature. Finally, plates were washed 5x with PBS-T and patted dry. 100 µL of TMB substrate (Thermo Fisher) was added to each well and incubated for 5 mins, followed by the addition of 100 µL of 2 M sulfuric acid to stop the reaction. Absorbance readings were acquired at 450 nM using Softmax Pro software on a Spectra Max 250 (Molecular Devices, Sunnyvale, CA).
Section II. eE2 Monoclonal Antibodies for Biochemical Study

1. Generation of Monoclonal Hybridoma Cultures

BALB/c mice were immunized via intraperitoneal route with 50 µg eE2 in either Complete Freund’s Adjuvant (first immunization only), or Incomplete Freund’s Adjuvant bi-weekly for eight weeks. A final immunization of 50 µg of eE2 was given intravenously 4 days prior to collection of splenocytes. Hybridomas were generated using a cloned HAT-sensitive mouse myeloma cell line as a fusion partner. Proliferating hybridomas were screened for their ability to bind eE2 via ELISA.

2. ELISA to test eE2 Monoclonal Antibodies

96-well ELISA plates (Corning, Lowell, MA) were coated with 100 µl of a 1 µg/mL solution of purified eE2 in NaHCO₃ overnight at 4°C. The plates were washed twice with 200 µl/well PBS-T, then blocked with a 5% solution of non-fat dry milk in PBS-T for 1 hour at 37°C. 100 µl of hybridoma cell supernatant was added to each well and incubated for 90 mins at room temperature. The plates were washed four times with PBS-T. 100 µl of goat anti-mouse HRP conjugate (Southern Biotech) diluted 1:2,000 in PBS-T was added and incubated for 1 hour at room temperature. Using TMB substrate solution, absorbance was measured using a VersaMax Microplate reader and SoftMax Pro software (Molecular Devices, Sunnyvale, CA).
3. Western Blot to test eE2 Monoclonal Antibodies

Samples of either J6 eE2-Fc or H77 eE2-Fc cell culture supernatant were loaded onto a 19:1 acrylamide:bisacrylamide gel (10%) with or without the addition of β-mercaptoethanol. The SDS-PAGE gel was transferred to nitrocellulose membrane for 30 mins using a Transblot Semi-Dry apparatus (Bio-Rad). Membranes were blocked in 5% nonfat dry milk in PBST for 1 hour at room temperature. Monoclonal antibodies were diluted 1:50 into blocking buffer and incubated for 1 hour at room temperature. After 3 ten-minute washes in PBS-T, anti-mouse HRP-conjugated secondary antibody (Thermo Scientific) was added at a dilution of 1:10,000 in blocking buffer for 1 hour at room temperature. After 3 more ten-minute washes in PBS-T plus a final wash in PBS, SuperSignal West Femto (Thermo Scientific) was used to activate HRP. Optimal exposure time was approximately 5 mins.

4. eE2 Truncation Mutants

Eight different amino-terminal truncation constructs plus 8 different carboxy-terminal constructs were designed for mapping the binding epitopes of the monoclonal antibodies produced.

Carboxy-terminal truncation reverse primers:

CT1: 5’ - AAGCAGGATCCGGAGTCATTGCAGTCAGGGC - 3’
CT2: 5’ - AAGCAGGATCCTTTGAAGATGGTATAGTTAACTGTG - 3’
CT3: 5’ - AAGCAGGATCCCATGCAGTCGGGCGGCAACACAG - 3’
CT4: 5’ - AAGCAGGATCCGTCGACACAAGGTCGCGGT - 3’
CT5: 5’ - AAGCAGGATCCATCCTCATATTGCAAGGCGCCC - 3’
CT6: 5’ – AAGCAGGATCCTCAAGCCTGCAGGGTGG - 3’
CT7: 5’ - AAGCAGGATCCAGCTCTAGTACGGCAGGGTGG - 3’
CT8: 5’ - AAGCAGGATCCAACTGTGCAGGGGTAATGCCA - 3’

Amino-terminal truncation forward primers:

NT1: 5’ - AACGAGCTAGCTGGCCCCAGGCAGAAAATCCA - 3’
NT2: 5’ - AACGAGCTAGCTTTGCACACCGGCTTTATCGCG - 3’
NT3: 5’ - AACGAGCTAGCTGGCGCCTTGCAATATGAGGATA - 3’
NT4: 5’ - AACGAGCTAGCTGGAGCGCCCACTTACACGTG - 3’
NT5: 5’ - AACGAGCTAGCTTGGTTCGGCTGCACGTGGATG - 3’
NT6: 5’ - AACGAGCTAGCTCCACCCTGCCGTACTAGAGCT - 3’
NT7: 5’ - AACGAGCTAGCTGACTTCAACGCCAGCACGGAC - 3’
NT8: 5’ - AACGAGCTAGCTCCCTGCACAGTTAACTATACCAT - 3’

These truncations were cloned into the original pcDNA3.1 backbone containing the prolactin signal sequence and human Fc tag using NheI and BamHI restriction enzymes.

This construct was transfected into HEK293T adherent cells using 8 µg of DNA per 500,000 cells with 8 µL of FuGENE-HD transfection reagent (Roche). Typically, 250,000 cells were seeded per well into a 6-well plate. Three days post-transfection, supernatants were harvested for ELISA mapping experiments.

5. eE2 Truncation Mapping ELISA

50 µL of transfection supernatant was added to the wells of a 96-well Maxisorp ELISA plate (Nunc, Thermo Fisher Scientific, Rochester, NY) and incubated for 2 hours at room temperature. Wild-type J6 eE2 supernatant was used as a positive control, while non-transfected cell supernatant was used as a negative control. After the 2 hour incubation, the plate was washed 3x with 150 µL volumes of PBS-T. 100 µL of 3% BSA...
solution was added for blocking and the plate was incubated for 1 hour at room temperature. Again, the plate was washed 3x with 150 µL volumes of PBS-T. 50 µL of each hybridoma cell supernatant was added to each of 3 wells per truncation (for triplicate results) and incubated for 1 hour at room temperature. The plate was washed 3 more times with 150 µL volumes of PBS-T. A goat anti-mouse HRP-conjugated secondary (Thermo Scientific) was diluted 1:15,000 in PBST and 50 µL was added to each well for a 1 hour incubation at room temperature. Finally, the plate was washed six times with 150 µL volumes of PBS-T and patted dry. 100 µL of TMB substrate (ThermoFisher) was added to each well and incubated for 5 mins, followed by the addition of 100 µL of 2 M sulfuric acid to stop the reaction. Absorbance readings were acquired at 450 nM using Softmax Pro software on a Spectra Max 250 plate reader (Molecular Devices, Sunnyvale, CA).

6. Purification of eE2 Monoclonal Antibodies

Supernatants from eE2 hybridomas that tested positive in initial ELISA screening were purified via protein A (GE Healthcare, Piscataway, NJ) affinity chromatography. Each antibody was purified using a separate 1 mL column to avoid cross-contamination. Elution was done using 0.1 M sodium citrate pH 3 buffer directly into 1 M Tris pH 9.0 at a ratio of 5:1 v/v for immediate neutralization. Samples were dialyzed overnight into 50 mM HEPES pH 7.5 + 100 mM KCl + 10 % glycerol and frozen in aliquots.
7. Anti-eE2 (3G4) CD81 Binding Assay

The protocol used for this assay was originally described by Flint et al. (42). GST-human CD81-LEL was expressed and purified as described in Section I. 96-well plates (Nalgene Nunc, Thermo Fisher Scientific, Rochester, NY) were coated with 50 µg/mL of GST, GST-humanCD81-LEL, or GST-mouseCD81-LEL in triplicate and incubated overnight at 4°C. Plates were washed 3x with PBS-T and blocked with 3% BSA in PBST for 1 hour at room temperature. 100 µL of supernatant from cells stably expressing eE2-Fc (dilutions made with DMEM + 10% FBS + 1% Pen/Strep) or 100 µL of purified eE2 (dilutions made with HEPES buffer) were added to appropriate wells in triplicate and incubated overnight at 4°C. On the third day, the wells were washed 3x with PBS-T and incubated with a 1:1,000 dilution of 3G4 hybridoma cell culture supernatant for 1 hour at room temperature. Plates were washed 3x with PBS-T. Finally, anti-mouse HRP antibody (ThermoFisher) was added at a dilution of 1:15,000 and incubated for 1 hour at room temperature. Plates were washed 5x with PBS-T and patted dry. 100 µL of TMB substrate (ThermoFisher) was added to each well and incubated for 5 mins, followed by the addition of 100 µL of 2 M sulfuric acid to stop the reaction. Absorbance readings were acquired at 450 nM using Softmax Pro software on a Spectra Max 250 (Molecular Devices, Sunnyvale, CA).
8. Monoclonal Antibody Neutralization Assay

96-well cell culture plates were seeded with $6.0 \times 10^5$ Huh7.5 cells per well and grown overnight. The following day, purified 1F11, 2C1, 2D1, and 14F7 monoclonal antibodies were diluted as described; 50 µL of each dilution was added to the appropriate wells of the seeded culture plate. Immediately following the addition of antibody, 50 µl of \textit{R.luc} Cp7 virus (CNS2Rluc - TCID$_{50} = 10^3$) was added to each well. Four days later, the cells were lysed and the R.luciferase enzyme quantified according the instructions provided in the Renilla Luciferase Assay Kit (Promega).
Section III. Disruption of Disulfide Bonds in HCVcc E2 and eE2

1. Site-Directed Mutagenesis of eE2 Cysteines and C6A Regional Mutants

Mutations for cysteine 6 and cysteine 11 were done using the QuickChange Site-Directed Mutagenesis kit (Stratagene, La Jolla, California) in the context of the full length viral genome. The following primers and their reverse complements were used for the site directed mutagenesis:

E2_C6A: (F) 5’- CTCCCGAAGACTGCTGGCTGGCCCAGTGTACTG-3’
E2_C11A: (F) 5’- CCAGCAGGGAGCTGTGGGGCCACGGACTGTTTTAGG-3’
E2_V502A: (F) 5’- GTCTCCGGAAGACTGCGGCCCAGTGTACTGTTTACC-3’
E2_G504A: (F) 5’- CAAAGACTGCTGGCTGGCCCAGTATACGTTTACC-3’
E2_V506A: (F) 5’- CCAGAAGACTGCTGGCTGGCCCAGTGTACTGTTTACC-3’
E2_Y507A: (F) 5’- CTGTGTGTGGCCCAGTGTACTGTTTACC-3’
E2_C508A(C7A): (F) 5’- GTGTGGCCCAGTGTACTGTTTACC-3’
E2_F509A: (F) 5’- CTGTGTGTGGCCCAGTGTACTGTTTACC-3’
E2_T510A: (F) 5’- TGCTGGCCCAGTGTACTGTTTACC-3’

The sequence for expression of recombinant protein was amplified from the viral genome with the addition of Nheli forward and BamHI reverse restriction sites. The PCR product was cloned into the original pcDNA3.1 backbone containing the prolactin signal sequence and human Fc tag. A 6-well plate was seeded with 0.5x10^6 cells per well and the pPro-eE2-C6A-Fc and pPro-eE2-C11A-Fc vectors were transfected the following day using FuGene-HD (Roche Diagnostics, Basel, Switzerland). After three days, the cells were expanded to 150 cm^2 dishes and placed under hygromycin (Calbiochem, San Diego, CA) selection at 75 μg/mL. Individual colonies were selected, expanded, and tested for expression using an anti-Fc ELISA.
2. Purification of eE2-WT, eE2-C6A, and eE2-C11A

Purification of soluble eE2 mutants was performed as previously described in Materials and Methods Section I. Briefly, HEK293T stable cell lines were created to express each of the eE2 variants under the control of a CMV promoter. Each of the eE2 variants included both an amino-terminal prolactin signal sequence and a carboxyl-terminal Fc tag. A hygromycin resistance gene enabled stable clone selection. Supernatants from each of the stable cell lines were harvested, centrifuged to remove cellular debris, and filtered through a 0.22 μm membrane. The supernatants were then applied to protein A-conjugated resin (GE Healthcare, Piscataway, NJ) overnight for eE2 immobilization via Fc binding. Following extensive washing, the resin was incubated with thrombin protease for Fc tag removal. The protein eluates were then consolidated and the concentration determined by BioRad Protein Assay.

3. Circular Dichroism

Purified protein samples were desalted into 20 mM sodium phosphate pH 7.0 and 50 mM KCl. The CD spectra in the wavelength range of 195-260 nm were measured at 0.5 nm intervals on an Aviv spectropolarimeter model 400 (Lakewood, NJ) at 25°C at the Robert Wood Johnson Medical School Core Facility. A quartz cell with a path length of 0.1 cm was used. The data are presented in degree cm$^2$ dmol$^{-1}$. 
4. **CD81 Binding Assay**

The protocol used for this assay was originally described by Flint et al. (42). GST-human CD81-LEL was expressed and purified as described in Materials and Methods Section I. 96-well plates (Thermo Fisher Scientific, Rochester, NY) were coated with 50 µg/mL of GST-CD81-LEL overnight at 4°C. Plates were washed 3x with PBS-T and blocked with 3% BSA in PBST for 1 hour at room temperature. 100 µL of supernatant from cells stably expressing eE2-WT, eE2-C6A, and eE2-C11A was added to appropriate wells and incubated overnight at 4°C. On the third day, the wells were washed 3x with PBS-T and incubated with a 1:15,000 dilution of anti-Fc HRP-conjugated antibody (Thermo Scientific) for 1 hr at room temperature. Finally, plates were washed 5x with PBS-T and patted dry. 100 µL of TMB substrate (Thermo Fisher) was added to each well and incubated for 5 mins, followed by the addition of 100 µL of 2 M sulfuric acid to stop the reaction. Absorbance readings were acquired at 450 nM using Softmax Pro software on a Spectra Max 250 (Molecular Devices, Sunnyvale, CA). Following supernatant incubation, plates were washed 5x with PBST and developed with TMB substrate (Pierce, Rockford, IL).
Section IV. Production of eE2 using Lentiviral Expression Vectors

1. Lentiviral Constructs and Cloning

The pJG vector provided by John Shires (Emory University, Atlanta, GA) contains only two unique restriction sites, Pmel and PacI, and is difficult to clone into using standard restriction digestion and T4 ligation. The original prolactin signal sequence and protein-A tag were fused to the protein of interest in a pcDNA3.1 sub-cloning procedure prior to amplification for pJG lentiviral vector insertion. For genotypes and proteins other than J6 eE2, a 3X-Flag tag (Sigma Aldrich, St.Louis, MO) was fused to the carboxy terminus of the protein-A tag. The pJG vector was digested using the Pmel restriction enzyme, followed by treatment with CIP to prevent re-ligation. The eE2 or eE1 gene (including the signal sequence and tags) was amplified using Pmel complimentary forward and reverse primers, with InFusion-HD (Clontech, Mountain View, CA) pJG overhangs. InFusion-HD cloning was executed according to the manufacturer’s instructions; 3 µL of the reaction was transformed into 100 µL of HB101 cells and plated on agar containing Ampicillin.

2. Lentiviral Co-transfection

One day prior to the planned transfection, a single T-225 monolayer flask was seeded with 6.0x10^6 HEK293T cells. For transfection, 90 µg pJG-eE2, 60 µg psPAX2, 30 µg pMD2.G, and 450 µL CaCl₂ were added to a 50 mL tube and brought to a final volume of 4.5 mL with ddH₂O. 4.5 mL of 2x HEPES buffer was added at room temperature and
bubbled with a serological pipet for 10 seconds. The mixture was incubated at room temperature for 2 mins and then added directly to the culture media in the 225 cm$^2$ flask prepared the day before. The media was changed 6-8 hours later.

3. Lentiviral Harvest and Infection

Two days after the transfection, 10,000 cells were seeded into a single well of a 96-well plate in a final volume of 50 µL (2x10$^5$ cells/mL). The supernatant from the transfection, containing the recombinant lentiviruses, was harvested and centrifuged for 30 min at 4000 xg at 4°C to pellet major cellular debris. 37 mL of clarified supernatant was transferred to a Beckman Ultracentrifuge tube (Denville Scientific, Metuchen, NJ). Lentivirus was pelleted for 1.5 hours at 25,000 rpm (80,000 xg) at 4°C using maximum acceleration and deceleration in an SW28 rotor. Supernatant was decanted into a waste container filled with 1% vespheine. The tube was inverted for 5 min to dry the pellet followed by resuspension in 120 µL of DMEM + 20% FBS + 1% antibiotic/antimycotic (A/A) + 8 µg/mL of polybrene. The media was aspirated from the prepared well in the 96-well plate and 50 µL of virus suspension was added and incubated overnight. The remaining 70 µL was frozen and stored at -20°C. The following day, 50 µL of fresh DMEM + 10% FBS + 1% A/A was added to the infected well. On the third day, the media was removed and replaced with 100 µL of fresh DMEM + 10% FBS + 1% A/A. Cell expansion began as soon as confluence was reached, usually two days later.
4. **BelloCell Operation**

A single BelloCell-500 bioreactor (Cesco Bioengineering, Taiwan) was pre-equilibrated with 500 mL of DMEM + 10% FBS + 1% A/A for 30 min on the BelloStage3000 (Cesco Bioengineering, Taiwan) in a standard cell culture incubator. One bioreactor was seeded with ~5.0-8.0x10^7 HEK293T cells, the equivalent of 3-4 confluent T-175 monolayer flasks, suspended in a total volume of 50 mL and gently dispensed onto the matrix. The seeding protocol was followed in accordance with the manufacturer’s instructions, with an oscillation rate of 2.0 mm/sec upward and downward, a top hold time of 20 sec, and a down hold time of 0 sec. After 2-4 hours of seeding, the oscillation speed was reduced to 1.0 mm/s, with a top hold time of 10 sec and a bottom hold time of 1 min.

Approximately 3 days after seeding, or when the glucose reading for the media fell below 1 g/L, the media was aspirated from the bottom chamber via the central access point and refreshed with 500 mL of fresh media. After the first harvest, the media was harvested and replaced every two days.
Section V. Optimization of eE2 for Crystallization Studies

1. Generation of J6 Amino-terminal Truncations

Based on the HDX results, five different amino-terminal truncations were constructed.

In addition, the carboxy terminus was truncated to eliminate the 17\textsuperscript{th} odd cysteine.

Forward aa426: 5’ – AACGAGCTAGCTGCCCCTGAACTGCAATGACTCC – 3’
Forward aa433: 5’ – AACGAGCTAGCTTTGACACCCGGCTTTATCGCG – 3’
Forward aa449: 5’ – AACGAGCTAGCTTCGTCAGGATGTCCCGAACG – 3’
Forward aa456: 5’ – AACGAGCTAGCTATGTCCGCCTGCCGCAGTAT - 3’
Forward aa470: 5’ – AACGAGCTAGCTGGCGCCTTGCAATATGAGGATA – 3’
Reverse aa656: 5’ – GAAGTGGATCCACGATCCCCACGAGTGAAATT – 3’

The constructs were amplified by PCR, digested with NheI and BamHI restriction enzymes, and cloned into the original pcDNA3.1 backbone containing the prolactin signal sequence and protein A tag. After confirmation of positive clones by sequence analysis, the entire construct was amplified with a second round of PCR using Pmel primers designed for InFusion-HD (Clontech) cloning into the pJG lentiviral vector:

IF Prolactin Forward: 5’ - AATTTAATTAAGTTAAACATGGACAGCAAGGGTTCGTCG - 3’
IF ProtA Reverse: 5’ – TCGAGGGGATCTTAAACCTACCCGGCATCGTCTTTAAG – 3’

2. Limited Proteolysis

Purified eE2 was deglycosylated with EndoH under native conditions, then subjected to limited proteolysis with trypsin, chymotrypsin, or glu-C at both 4°C and 20°C. Each enzyme was used at a quantity of 1 µg per 100 µg of eE2 and the experiments performed in 20 mM HEPES 7.5 + 100 mM KCl. A non-digested eE2 sample (lane 0) was
saved as a positive control. At each time point (15 min, 30 min, 1 hour, 4 hours), 10 µL of the digested sample was reserved to a separate tube and flash frozen in liquid nitrogen until samples were ready to be loaded onto a gel. 10 µl of SDS-PAGE sample buffer with β-mercaptoethanol was added to each sample, then samples were boiled at 95°C for 5 minutes and loaded onto 12% polyacrylamide gels. Gels were stained by Coomassie blue.

3. Production of Endoglycosidase H

A codon-optimized endoglycosidase H (EndoH) construct was synthesized by GenScript (Piscataway, NJ). The gene was sub-cloned into pGEX-6P-1 (GE Healthcare, Piscataway, NJ) for biochemical purification via GST tag, using BamHI and XhoI restriction enzymes. GST-EndoH was transformed into UT5600 E.coli and grown on agar containing Ampicillin. A single colony was used to inoculate a 40 mL starter culture of LB + Ampicillin on the following day, which was shaken overnight at 37°C. On the third day, 10 mL of the starter culture was added to 1 L of LB + Ampicillin. The culture was shaken at 225 rpm at 30°C until reaching an OD of ~0.5 at 600 nm. The culture was cooled for 1 hour at 4°C, then induced with 300 nM IPTG. The culture was again shaken overnight at 200 rpm at 18°C. The following morning, the bacteria were pelleted at 5000 rpm for 20 min and resuspended in 15 mL of buffer containing 50 mM HEPES pH 7.5 + 1 M KCl + 1 mM EDTA. Lysis was done using homogenization, followed by a second round of pelleting at 18,000 rpm for 40 min to clarify the lysate. A 5 mL GST column was equilibrated with 50 mM HEPES pH 7.5 + 150 mM KCl. After loading the bacterial lysate
onto the column, the same buffer was used for thorough washing. GST-EndoH was eluted from the column using 100 mM TRIS pH 8.5 + 150 mM KCl + 15 mM reduced glutathione. Finally, GST-EndoH was dialyzed overnight into 50 mM HEPES pH 7.5 + 150 mM KCl + 20% glycerol and frozen in aliquots the next day.

4. Deglycosylation with Endoglycosidase H

GST-EndoH was produced as described above. eE2 protein was adjusted to compatible pH using 0.75 M sodium citrate buffer pH 5.5 to a final concentration of 100 mM. Endo H was added at a ratio of 1 mg per 20 mg of eE2 and the reaction was incubated at room temperature for 4 hours.

5. Purification of Deglycosylated eE2

After deglycosylation with Endo H, eE2 was dialyzed overnight into 20 mM HEPES pH 7.5 + 50 mM KCl + 5% glycerol. The following day, eE2 was centrifuged at 18,000 rpm for 30 min to pellet precipitated material. eE2 was then loaded onto a heparin column (GE Healthcare, Piscataway, NJ) and eluted using a salt gradient from 50 mM KCl to 500 mM KCl. The highest quality protein eluted the earliest at around 100 mM KCl; aggregated material eluted above 300 mM KCl. Following elution from the heparin column, eE2 was loaded onto a Superdex 200 16/60 prep grade size exclusion column into 20 mM HEPES pH 7.5 + 100 mM KCl + 5% glycerol.
6. Generation of eE2 Genotypes

Based on the truncation results for J6, congruent truncations were generated for other HCV genotypes:

Forward H77: 5’ – AACGAGCTAGCTTTTGCCAGCTGCCGACGC – 3’
Forward J8: 5’ - AACGAGCTAGCTTTTGTCTTCTCTGCCGACG - 3’
Forward SA13: 5’ – AACGATCTAGATATGGCTAGCTGTAGGCCC – 3’
Forward S52: 5’ – AACGAGCTAGCTCTAGCAGCTGCAAGCCCAT – 3’
Forward ED43: 5’ – AACGAGCTAGCTCTCGCTGCTGCAAGAGC - 3’
Forward QC69: 5’ – AACGAGCTAGCTATGGCTTCTTGTAAACCTCTCA – 3’
Forward HK6a: 5’ – AACGAGCTAGCTATGGCTGCGTGTAAGCCCC – 3’

Reverse H77: 5’ – GATTGGATCCGCGTTCGCCCGCCTGCCA – 3’
Reverse J8: 5’ – GTTGGATCCGCGATCTCCGCGCTGAAG – 3’
Reverse SA13: 5’ – GTTGGATCCCGCTCAACACCGGTCCA – 3’
Reverse S52: 5’ – CTTAGATCTGCCTCCCCCTGGCTCC – 3’
Reverse ED43: 5’ – GTTGGATCCGACCTCCCCCTGGTCCAG – 3’
Reverse QC69: 5’ – GTTGGATCTCGCTCAACCCCGGTCCA – 3’
Reverse HK6a: 5’ – GAAGGATCCCCGCTCGCTGTGAAGC – 3’

The constructs were amplified by PCR, digested with appropriate restriction enzymes, and cloned into the original pcDNA3.1 backbone containing the prolactin signal sequence and protein A/3xFlag tag. After confirmation of positive clones by sequence analysis, each entire construct was amplified with a second round of PCR using primers designed for InFusion-HD (Clontech) cloning into the pJG lentiviral vector:

IF Prolactin Forward: 5’ - AATTTAATTAAGTTAATACATGGACAGCAAACGAGGTTGTCG - 3’
IF Flag Reverse: 5’ – TCGAGGGGATCCTTTAAACCTACTTTGTCACTGCTACTCCTT – 3’
7. Purification of Glycosylated eE2

Each batch of eE2-ProtA supernatant was centrifuged for 10 min at 5,000 rpm and filtered through a 0.22 µM membrane. Supernatant was then loaded onto an IgG FF column (GE Healthcare, Piscataway, NJ) and eluted with 0.1 M sodium citrate pH 3.0 + 20 mM KCl directly into 1 M Tris pH 9.0 for immediate neutralization. For every 1 mL of elution, 200 µL of Tris was used. A 1:50 ratio of PreScission Protease enzyme was added to the eluted protein to cleave off the protein A tag; this step required overnight dialysis into a buffer with 20 mM HEPES pH 7.5 + 250 mM KCl + 5% glycerol. The next day, the eE2 protein was desalted into 20 mM HEPES pH 7.5 + 100 mM KCl + 5% glycerol and passed over Q ion exchange. Protein A and uncleaved eE2-ProtA bound to the Q column at pH 7.5. The flow-through from the Q column was concentrated and loaded onto a Superdex200 prep grade 16/60 size exclusion column (GE Healthcare, Piscataway, NJ) using the same buffer.
EXPERIMENTAL RESULTS

Section I. Biophysical Analysis of HCV eE2

At the early stages of this thesis work, few researchers had attempted expression of the HCV envelope glycoproteins. HCV strain H77 of genotype 1a was the prototypic strain for most of the research done in the HCV field and was the first to be proven infectious in chimpanzees (69, 129). However, replication studies of HCV have shown that J6 (an isolate of the 2a strain) is strikingly conducive to \textit{in vitro} research. Using the structural proteins from this strain, a group led by Dr. Charles Rice was able to design a recombinant J6/JFH-1 system to produce infectious HCV in laboratory conditions for the first time (80). Drs. Matthew Evans and Joseph Marcotrigiano produced truncated H77 and J6 soluble E2 constructs containing only their respective amino-terminal ectodomains (amino acids 384-661 and 384-665 of the polyprotein, respectively) and cloned these into a standard pcDNA3.1 vector. Since the carboxy-terminal hydrophobic domain functions to anchor the protein within the lipid membrane, including this in the construct conflicts with attempts to make a secretable form of the protein. In addition, the ectodomain has been described as the functional unit for binding and entry, as determined by previous studies testing for CD81 and SR-BI binding (89).

Expression of E2 ectodomain (eE2) in \textit{E. coli} (82, 131), yeast (85), insect cells (107), CHO cells (11), and various other eukaryotic and viral recombinant expression systems, has consistently resulted in the formation of insoluble, misfolded and disulfide-linked,
aggregated protein. Therefore, we sought to develop a system for the expression of HCV eE2 that would yield large amounts of highly purified, active protein for functional studies. Our approach was to use cell lines that have been shown to produce functional E2, while adding an affinity tag to promote eE2 folding and facilitate purification. HEK293T cells were chosen owing to their ability to produce functional E2 in the form of replication incompetent HCV pseudoparticles (HCVpp) (7), ease of handling and robust growth rate, excellent transfectability and high capacity for recombinant protein expression.

Folding of the envelope proteins is slow and requires the ER chaperone machinery, including calnexin (20). Recombinant eE2 expression has yielded two different forms of the molecule: (i) a glycosylated protein with intramolecular disulfide bonds that is believed to be the active form and (ii) high molecular weight aggregates caused by intermolecular disulfide bonds. In fact, aggregation is so common that a non-productive folding pathway has been proposed as a physiologically relevant part of the HCV lifecycle (33). The formation of disulfide-bonded aggregates and misfolded protein has limited functional, structural, and biophysical studies on the HCV glycoproteins. Described here is a method to produce significant quantities of eE2 using human cells. The resulting protein is not aggregated, is recognized by antibodies from chronically infected HCV patient sera, and blocks HCV infection in vitro. Furthermore, we have extensively characterized the glycosylation pattern, oxidation state, and oligomeric nature of eE2 using a variety of biophysical techniques. Our purified eE2 protein
constitutes an exceptional tool for probing E2 function, and may facilitate in the design of an entry inhibitor or HCV vaccine.

1. Expression of J6 eE2-Fc in HEK293T Cells

A cytomegalovirus promoter drives expression of these constructs. Each protein construct is preceded by a prolactin signal sequence and followed by a thrombin cleavage site and human Fc domain (Figure 1A). The prolactin signal sequence promotes proper trafficking and secretion, while the Fc antibody domain facilitates secretion and constitutes a tool for biochemical assays. In addition, the Fc domain includes natural introns, which promote splicing and boost expression of the protein (Dr. Dimitar Nikolov, personal communication). The Fc domain contains glycosylation and disulfide bonds that are essential to confer binding to protein A and this ensures quality control, since E2 is also glycosylated and contains disulfide bonds. Presumably, if Fc is modified and folded properly, it is likely to correlate with proper folding and modification of the fused E2 ectodomain. For unknown reasons, the Fc region has a substantial effect in escorting E2 out of the cell. Alternative tags, including V5 and poly-histidine, were tested and resulted in nearly undetectable quantities of secreted protein. The thrombin site allows for cleavage of the E2 protein from the Fc domain, which will remain attached to Protein A sepharose following cleavage, whereas the
Figure 1. eE2 Schematic and Sequence

A) Recombinant eE2 expression construct, including all notable features required for secretion and biochemical analysis. B) Sequence of eE2 (residues 384-664 genotype J6) highlighting the conserved cysteine residues (Underlined), and the potential N-linked (Bold) and O-linked (Italics) glycosylation sites. Signal sequence residual amino acids are indicated in green, restriction enzyme residues are indicated in red, and thrombin protease residual amino acids are indicated in blue.
untagged soluble eE2 is eluted from the column. Thrombin protease is commonly used for this type of application and its activity requires a specific amino acid consensus sequence (Leu-Val-Pro-Arg X Gly-Ser). The final sequence for the secreted protein, including restriction enzyme residues and residues left behind from signal sequence cleavage and thrombin protease cleavage, is shown in Figure 1B.

An HEK293T cell line was developed that stably secretes eE2-Fc into the media. The presence of the fusion protein is detectable by Western blotting of the cell supernatants (Figure 2A) and cell lysate (data not shown). The eE2-Fc was purified using protein A resin and eE2 was subsequently separated from the immobilized Fc tag via thrombin protease cleavage, leaving the Fc tag and contaminants bound to the resin (Figure 2B). The protein that eluted from the resin by thrombin cleavage was confirmed to be eE2 by protease digestion followed by high-resolution mass spectrometry (see below) and amino-terminal sequencing (data not shown). The calculated molecular weight of the J6 eE2 protein is 33 kDa, although it migrates around 60 kDa by reducing SDS-PAGE. This molecular weight discrepancy and the diffuse nature of the band are consistent with heavily glycosylated protein.

2. eE2 Glycosylation Analysis

Glycosylation of viral envelope proteins is critical for folding, structural integrity, and immune evasion. The number and conservation of glycosylation sites varies across different HCV genotypes (48). J6-E2 contains 11 potential N-linked
**Figure 2. eE2 Western and Purification**

**A)** Supernatants from HEK293T cells transfected with GFP (mock) or eE2-FC, anti-Fc Western. **B)** Sups from cell lines expressing eE2-Fc were clarified by centrifugation (sup loaded) and applied to the resin. After incubation, the column was extensively washed to remove unbound material (flow through). E2 was eluted off the column in five fractions (elutions) following incubation with thrombin protease. Also shown is the resin before (bound resin) and after (post-cleavage resin) elution.
glycosylation sites (N-X-T/S, where X is any amino acid except proline) along with three potential O-linkage consensus sites (Figure 1B). We investigated the extent of eE2 N-linked glycosylation and the type of oligosaccharide at each site using endoglycosidases. High mannose and complex oligosaccharides can be differentiated by endoglycosidase H (Endo H) sensitivity, since Endo H will only cleave high mannose and some hybrid glycans. Peptide-N-glycosidase F (PNGase F) will remove all types of N-linked glycosylation indiscriminately. Deglycosylation of eE2 with PNGase F under denatured, reducing conditions resulted in a faster migrating band greater than the 31 kDa standard, consistent with its calculated molecular weight of 33 kDa (Figure 3). In contrast, eE2 was largely resistant to digestion with Endo H (Figure 3). This result suggests that the majority of the N-linked glycans on eE2 are of the complex form, in accordance with its mode of expression by export through the secretory pathway.

To investigate the glycosylation site usage in further detail, we employed high-resolution mass spectrometry. eE2 was digested with either trypsin or chymotrypsin and samples of the protein fragments were deglycosylated with either PNGase F or Endo H. PNGase F deaminates the asparagine residue to which the N-linked glycan is attached and converts it to aspartic acid. If the glycan is of the high mannose form, it will be sensitive to Endo H, which leaves one N-acetylglucosamine (GlcNAc) bound to the Asn. Thus, the gain of 1 Da (Asn to Asp; nitrogen to oxygen) by PNGase F or 203 Da by the GlcNAc residue left by Endo H cleavage can be resolved by mass spectrometry of the peptides. From the resulting three spectra (untreated, Endo H and PNGase F treated) we
Figure 3. Deglycosylation of eE2 with PNGase F and Endo H.

Purified eE2 was deglycosylated with PNGase F or Endo H under denaturing and reducing conditions, and then analyzed by SDS-PAGE analysis. The positions of the glycosylated eE2 (eE2/gly), deglycosylated eE2 (eE2/degly) and both enzymes (Endo H and PNGase F) are also shown.
were able to map all the glycosylation sites, estimate the approximate usage of each site, and determine whether the glycan at a particular site was complex or high mannose. We achieved almost 100% coverage of the eE2 sequence using either trypsin or chymotrypsin. Examination of peptides containing the 11 predicted N-linked glycosylation sites indicated that all sites were fully glycosylated (Figure 4, upper spectra of each panel). Only one of the 11 glycosylation sites was found to be Endo H sensitive (VGGVEHRLTAACNF, data not shown for Endo H), suggesting that the majority of the glycans are complex in nature (Figure 4, lower spectra of each panel). Peptides containing the potential O-linked glycosylation sites were resolved and shown to be unmodified (data not shown).

3. eE2 Oligomerization and Secondary Structure

Since previous reports have shown that E2 tends to aggregate, we set out to define the oligomeric state of eE2 using non-reducing SDS-PAGE, size exclusion chromatography (SEC) and analytical ultracentrifugation. SDS-PAGE analysis of eE2 under non-reducing conditions demonstrated that eE2 consisted of a mixture of two components with approximate molecular weights of ~120 kDa (dimer) and ~60 kDa (monomer) with a small amount of larger molecular mass protein (Figure 5A). To determine the oligomeric state of eE2 under native conditions, the protein was evaluated by SEC. eE2 yielded two major peaks and a slight peak found in the void volume of the column (Figure 5B). The major and minor peaks were estimated at ~123 kDa and ~75 kDa, respectively, by an inline static light scattering detector (data not shown). Both non-reducing SDS-PAGE and
Figure 4. Mapping the N-linked Glycosylation Sites.

Each panel contains the extracted ion chromatograms of proteolytic peptides of eE2 with (bottom) and without (top) PNGase F digestion. The data are plotted with the elution time on the x-axis and relative abundance on the y-axis. The peptide sequence and calculated molecular weight used for the extraction are provided at the top of each panel. The measured mass/charge ratio (m/z) and charge state for each spectrum are provided. Note that one panel has a peptide that contains two N-linked glycosylation sites. The identification of each peak was confirmed by MS/MS.
Figure 5. eE2 Oligomeric Species

A) SDS-PAGE analysis of purified eE2 in the presence and absence of β-mercaptoethanol (β-ME) and stained with Coomassie blue. B) Purified eE2 was applied to a Superdex 200 size exclusion column equilibrated with 50 mM HEPES pH 7.5, 150 mM KCl, 5% glycerol. The arrows denote the positions of the void volume, dimer, and monomer.
native size exclusion chromatography clearly demonstrated that eE2 is not aggregated. The ratio of dimer to monomer in non-reduced SDS-PAGE (Figure 5A) and gel filtration (Figure 5B) is similar, suggesting that a portion of the dimer may result from an intermolecular disulfide bond.

There are 18 conserved cysteines in full length E2, which could result in the formation of 9 disulfide bonds. The eE2 fragment contains only 17 cysteines, leaving at least one unpaired. We therefore sought to determine the oxidation state of the conserved cysteine residues in eE2. A differential cysteine labeling method followed by high-resolution mass spectrometry was employed to distinguish free cysteine residues from those involved in disulfide bonding. Briefly, eE2 was incubated with a molar excess of N-ethylmaleimide (NEM) under denaturing, non-reducing conditions to label all free cysteine residues. After disulfide bond reduction with DTT, the newly generated free cysteines were alkylated with iodoacetamide (IAM). The modified protein was digested with trypsin, deglycosylated with PNGase F, and the resulting peptides were resolved by mass spectrometry. All cysteine-containing peptides were identified and only one peptide (C\(^{656}\)NLEDRDR) was modified by NEM (Figure 6A). The expected unmodified molecular weight of this peptide is 1020.45 Da, 1077.45 Da if modified by IAM, or 1145.45 Da if modified by NEM. In Figure 4A, the spectrum shows a 573.75 m/z peak corresponding to this peptide modified by NEM and carrying a +2 charge, while no peptide appears at a position corresponding to a modification by IAM (expected ~538 m/z with a 2+ charge). All other cysteine-containing peptides were shown to have an
Figure 6. Differential Labeling of Free and Disulfide-linked Cysteines.

Free and disulfide bonded cysteines were labeled with NEM and IAM, respectively. LC-MS/MS extracted ion chromatograms for peptides containing C656 (A) and C459 (B) are shown. The extraction was done using the molecular weights of the peptides modified with IAM (upper panel) and NEM (bottom panel). C656 is free, while C459 is found in a disulfide bond. All of the other cysteine residues were labeled with IAM (data not shown), suggesting the formation of eight disulfide bonds. The data are plotted with the elution time on the x-axis and relative abundance on the y-axis. The measured mass/charge ratio (m/z) and charge state for each spectrum is provided. The identification of each peak was confirmed by MS/MS.
addition of 57 Da, indicating that they were modified only after reduction with DTT. For example, the expected molecular weight of the unmodified SAC\textsuperscript{459}RSIEAF peptide is 983.46 Da, 1040.46 Da if modified by IAM, or 1108.46 if modified by NEM. This peptide resolves at 521.32 m/z, which corresponds to the molecular weight when modified by IAM and carrying a +2 charge. It does not resolve as modified by NEM (expected ~554 m/z with a 2+ charge) (Figure 6B).

From the results of the non-reducing SDS-PAGE, cysteine labeling experiment, and SEC, we presumed that C656 may potentially be involved in the formation of an intermolecular disulfide bond. To test this hypothesis, C656 was mutated to a serine (eE2-C656S) to conserve some of the biochemical properties at this position. An HEK293T cell line that stably expresses eE2-C656S-Fc was generated and the protein was purified as before. Results from non-reducing SDS-PAGE indicated that the proportion of monomer in eE2-C656S is substantially increased relative to wild type eE2 (compare Figure 5A and Figure 7A). In order to determine the native behavior, eE2-C656S was analyzed by SEC under conditions identical to wild type eE2. The results demonstrated that the mutant is predominantly monomeric at pH7.5 with a lesser amount of dimer (Figure 7B).
Figure 7. eE2-C656S Oligomeric Species

A) SDS-PAGE analysis of purified eE2-C656S in the presence and absence of β-mercaptoethanol (β-ME) followed by staining with Coomassie blue.

B) Purified eE2-C656S was applied to a Superdex200 size exclusion column equilibrated with 50 mM HEPES pH 7.5, 150 mM KCl, 5% glycerol. The arrow denotes the position of the void, dimer, and monomer.
To confirm our SEC results, we used sedimentation velocity experiments by analytical ultracentrifugation, which can determine mass distributions and globularity from mixtures of proteins without the use of standards or interactions with a sieving matrix. Sedimentation of eE2-C656S at pH 7 indicates the presence of two dominant species, a monomeric form (60-70 kDa) and a dimeric form (80-130 kDa) (Figure 8A and 8B). There is a minor third species that has the approximate molecular weight of a trimer (150-200 kDa). Since analytical ultracentrifugation failed to detect large aggregates, we assume that trimer corresponds to the small peak that eluted in the void volume of SEC. The monomer has a frictional ratio in the range of 1.0 - 2.0 where a perfect sphere would have a frictional ratio of 1.0. It is interesting to note that the frictional ratio shows a decreasing trend with increasing molecular weight, suggesting a more globular shape for the oligomeric forms. To determine if the ratio of oligomers or protein globularity were dependent on the amount of protein measured, the analysis was performed at two different concentrations. There is no appreciable difference in the molecular weight, ratio of the different species, or shape distributions at the two concentrations used in this analysis (OD230 nm of 0.25, Figure 8A, and 0.8, Figure 8B), indicating little or no effect due to mass action at these concentrations.

The HCV glycoproteins are predicted to be class II fusion proteins, due to their relatedness to alphaviruses and flaviviruses. Class II fusion proteins are composed of mostly β sheet structure and do not undergo major rearrangements in secondary structure upon exposure to low pH (62). We investigated whether the changes
Figure 8. Analytical ultracentrifugation data for eE2-C656S at pH7.
Two-dimensional spectrum/Monte Carlo analysis of HCN sedimentation velocity data. Measurements of eE2 were made at low concentration (0.25 OD$_{230}$) (A) and at higher concentration (0.8 OD$_{230}$) (B). Both samples show the presence of a trimeric species. Larger species appear more globular than smaller species according to the frictional coefficient ratios. The units of the color gradient are in OD$_{230s}$. 
seen in eE2 and eE2-C656S at pH 7 and pH 5 were due to changes in secondary structure by measuring the circular dichroism (CD). The resulting CD spectra revealed that both eE2 and eE2-C656S had a minimum at about 203 nm (Figure 9A and B, respectively). Multilinear regression analysis (data not shown) suggested that the eE2 spectrum is consistent with a protein composed of predominantly β-sheet and random coil secondary structure with little to no α-helical content. The spectra at pH 7 and pH 5 are super-imposable for both constructs, indicating that changes in the oligomeric nature and globularity of eE2 by lowering the pH are not due to changes in secondary structure. From these data, we conclude that eE2 has mostly β-sheet and random coil structure and, like the flavivirus envelope protein, does not undergo major changes in secondary structure.

4. eE2 is Recognized by Antibodies From HCV Infected Patients

The presence of high levels of anti-E2 antibodies in HCV-infected human serum has been reported (6). In order to further examine if purified eE2 is conformationally similar to the E2 present on infectious HCV particles, we tested whether eE2 was recognized by antibodies from infected patient sera. An enzyme-linked immunosorbent assay (ELISA) plate was coated with eE2 and probed with serum from patients chronically infected with either genotype 1, 2 or 3. Serum from a healthy donor was tested in parallel as a negative control. Anti-human HRP was used to quantify the result. Serum from the three infected patients bound to J6 eE2 (genotype 2a) at similar titers regardless of infecting genotype, while the serum of the uninfected donor responded
Figure 9. CD Spectroscopy of eE2 (A) or eE2-C656S (B) at pH 7 and pH 5.
CD spectra are shown as millidegrees versus wavelength (nm). Error bars represent deviations from the average over a 5 second interval.
at background levels (Figure 10A). This illustrates the capacity of eE2 to be recognized by antibodies in patient sera, while also pointing out the maintenance of cross-reactive epitopes.

5. eE2 Blocks HCVcc Entry

It was shown previously that properly folded, purified E2 ectodomain from pestiviruses, bovine viral diarrhea virus and classical swine fever virus, was able to block viral infection (58, 99). In order to confirm correct folding and function of HCV eE2, we performed a similar assay using cell culture derived HCV (HCVcc). Recombinant human CD81 LEL, which has been shown to inhibit HCVcc infection, was used as a positive control for blocking infection (80). About 100 TCID₅₀ of HCVcc was incubated with serial two-fold dilutions of purified eE2, eE2-C656S, GST-human-CD81LEL (hCD81), GST-mouse-CD81LEL (mCD81) or GST. eE2, eE2-C656S, and hCD81 reduced the number of focus forming units (FFU) in a concentration-dependent manner, while mCD81 and GST proteins had no effect. This experiment yielded a 50% blocking efficiency in the range of 25-150 µg/ml for eE2, eE2-C656S, and hCD81-LEL (Figure 10B). Thus, HCVcc infection can be effectively blocked by eE2 or eE2-C656S at similar concentrations to hCD81-LEL. In order to rule out the possibility that inhibition of viral infection was due to toxicity, we quantified cell death after incubation with purified protein using fluorescence activated cell sorting analysis (FACS). Cells were incubated for 3 days with 2-fold dilutions of eE2, GST and hCD81-LEL followed by staining with Via-Probe™ to estimate viability. As shown in Figure 10C, purified eE2 and hCD81 are not toxic at 200 µg/mL
Figure 10. Functional Analysis of eE2 and eE2-C656S.

A) ELISA plates were coated with eE2 and probed with a series of ten-fold dilutions of serum from patients infected with HCV (genotypes 1, 2, or 3) and healthy donor. Antibodies in HCV-infected patient sera could detect eE2 up to 1:10,000 dilution. B) Cells were incubated with HCVcc plus GST, GST-mCD81 LEL, GST-hCD81 LEL, eE2, and eE2-C656S. Three days post-infection the cells were fixed, focus forming units were determined, and percent of inhibition was calculated. eE2, eE2-C656S, and hCD81 inhibit HCVcc infection. Error bars represent standard error of the mean for two independent experiments. Each experiment was performed in duplicate. C) Cells were incubated with eE2, GST and GST-hCD81-LEL at two concentrations (200, or 100 µg/mL). Three days later, cells were analyzed for viability using flow cytometry. The results demonstrate that eE2 is not toxic when applied to cells at the concentration that inhibits HCVcc infection.
mL (the concentration with the highest level of inhibition).

Inhibition of HCVcc entry by eE2 is thought to occur by sequestering cellular receptors. To confirm this, we analyzed the ability of eE2 to bind hCD81 in vitro by adapting an ELISA first described by Flint et al for the detection of properly folded E2 based on hCD81 binding (42). Plates were coated with GST, mCD81, and hCD81, probed with supernatants from cell lines expressing eE2-Fc or eE2-C656S-Fc, then developed with HRP-conjugated anti-Fc. The assay was executed in triplicate using undiluted cell culture supernatant and two 10-fold dilutions with media. Both eE2-Fc and eE2-C656S-Fc specifically bind hCD81 but not mCD81 or GST alone (Figure 11). Identical results were obtained using purified eE2 and eE2-C656S (Fc tag removed, data not shown). Taken together these data confirm that the eE2 protein made in the system described here maintains many of the functional characteristics of E2 found on virus.

6. Discussion

HCV E1 and E2 are primary determinants of entry and pathogenicity (34). Deletion mutagenesis has defined the ectodomain of HCV E2 to comprise amino acids 384-661 of E2 from genotype H77 (89). Functional and biophysical studies of HCV E2 have been hindered by the formation of mostly aggregated, misfolded material. Here we describe a method for production of milligram quantities of eE2, which do not contain large disulfide linked aggregates. The resulting protein can compete with HCVcc to inhibit
Figure 11. Enzyme-linked immunoassay for CD81-LEL binding.

Tissue culture supernatants of eE2-Fc fusion (no dilution, 1:10 and 1:100 dilutions) were incubated in plates coated with either GST, GST-mouse CD81-LEL, or GST-human CD81-LEL. After washing bound eE2-Fc was detected with anti-human Fc-HRP. PBS, media from wt HEK293T cells and wells without any coating were used as controls. Both eE2 and eE2-C656S binds to only human CD81.
infection, is recognized by antibodies from chronically infected patients, and can specifically bind the large extracellular loop of human CD81. These results demonstrate that the eE2 protein maintains many of the functionalities associated with E2 found on virions. Using purified eE2 we have conducted a series of biochemical and biophysical studies aimed at characterizing the glycosylation, oxidation state, and oligomeric nature of eE2.

HCV is thought to bud into the ER and egress through the cellular secretory pathway. As the virion passes through the ER and Golgi, the exposed glycans associated with E1 and E2 can be modified by processing enzymes resulting in hybrid or complex glycans. Indeed, the E1 and E2 glycans of HCVpp are mostly insensitive to EndoH digestion, suggesting that the glycans are hybrid or complex (96). One of the advantages of the expression system described here is the use of the same mammalian cell line (HEK293T) employed for HCVpp production, presumably yielding similar glycan modifications. In fact, an Endo H digest comparison of eE2 (Figure 3) and HCVpp E2 (Figure 2A from reference Op De Beck 2004) reveals remarkable similarity in the digestion pattern. All eleven of the predicted N-linked glycosylation sites in eE2 are used, which is consistent with previous data for HCVpp (47). Ten of the eleven sites have complex glycans attached, while only the most carboxy-terminal is of the high mannose form. This observation suggests that the glycan at this position is concealed from the modification enzymes in the secretory pathway, possibly due to steric affects within E2 or blocking of the site by the Fc tag attached to the carboxy terminus. As reported by Goffard et al,
mutation of the carboxy-terminal glycosylation site does not affect folding, secretion, or E1/E2 heterodimer formation, but does result in less than 50% infectivity when incorporated into HCVpp (47).

Recently, it has been reported that E2 contains O-linked glycans (38). Peptides containing the putative O-linked glycosylation sites (Figure 1A) were not modified, as determined by LC-MS/MS. However, at this time, we cannot rule out that O-linked glycosylation occurs at a low level on eE2 or on full length E2 when expressed as part of the HCV polyprotein.

Full length HCV E2 has 18 highly conserved cysteines, although the eE2 construct as defined by Michalak et. Al. and used in this study has only the first 17. Non-reducing SDS-PAGE and SEC under native conditions demonstrated that about 60-70% of wild type eE2 exists as a dimer through an intermolecular disulfide bond. Differential labeling of free and disulfide-bonded cysteines demonstrated the presence of eight disulfide bonds and one free cysteine (C656) in eE2. We reasoned that since C656 is the only free cysteine, it might be involved in the intermolecular disulfide. Mutating C656 did not affect inhibition of HCVcc entry (Figure 10B), CD81 binding (Figure 11), folding as determined by CD (Figure 9) or binding to HCV patient antibodies (data not shown). Non-reducing SDS-PAGE and SEC at pH7 of eE2-C656S showed a dramatic shift toward a mostly monomeric form. By non-reducing SDS-PAGE of eE2-C656S, there is a residual amount of higher order oligomers. However, both SEC and sedimentation velocity
measurements demonstrate that the higher order structures are pH sensitive, which would argue against the dimer being disulfide-linked.

Our results are consistent with recently published data, showing that all 18 cysteine residues of E2 are in disulfide bonds and reduction of up to half of these is compatible with HCV entry as well as antibody and CD81 binding (39). The 18th cysteine is located in the stem region found between the ectodomain and the transmembrane helix (32). The stem consists of a heptad repeat and has been proposed to be an amphipathic alpha helix with a hydrophilic and hydrophobic surface. In this model, the 18th cysteine would be facing away from the membrane. Taken together, the data suggests that the ectodomain of E2 might be tethered to the stem region by an intramolecular disulfide bond between the 17th and 18th cysteine residues.

The HCV glycoproteins, like those from alphaviruses and related flaviviruses, are predicted to be class II fusion proteins. This class of proteins is characterized as mostly beta sheet structures that do not undergo changes in secondary structure upon exposure to low pH (62). The flavivirus E protein (comprised of three domains) is responsible for receptor binding and membrane fusion. Flaviviruses have icosahedral symmetry, with E arranged as 90 dimers. Cryoelectron microscopy has demonstrated that the E protein lies flat on the surface of the viral lipid bilayer (52). Upon exposure to low pH there is a slight rotation of domain III, resulting in dissociation of the E dimers and rearrangement into trimers via a monomeric intermediate stage.
The oligomeric nature of E2 on the surface of the virus remains unknown. It is assumed that the native form of the molecule contains intramolecular disulfide bonds and heterodimerizes with E1 through their respective transmembrane regions. The ectodomain of HCV E2 with the C656S mutation appears mostly as a monomer, with a small amount of dimer and a much lower amount of trimer at pH 7. Upon exposure to low pH, there is an increase in the amount of monomer relative to dimer and trimer as determined by SEC and analytical ultracentrifugation. Sedimentation velocity measurements showed that the ratio of the three oligomers did not change at the two concentrations tested. Isolation of the monomer by SEC and reanalysis did not suggest that the various states were in equilibrium (data not shown). However, the oligomeric state of E2 may be influenced by its carboxy-terminal portion, the high concentration found on virus particles, and/or heterodimerization with E1.

CD spectroscopic analysis of HCV eE2 had a pronounced minimum at about 203 nm, consistent with a protein containing mostly β-sheet and random coil. Exposure to low pH did not cause any significant changes in the CD spectrum, which suggests the lack of major rearrangement of secondary structure. However, CD cannot rule out the possibility of structural rearrangements that preserve the overall proportion of β-sheet and random coil. In fact, changes in certain conformational epitopes on HCVpp E2 upon lowering of the pH have been reported (95).
Recombinant CD81-LEL, eE2, and eE2-C656S are capable of blocking HCVcc infection (Figure 10B). The concentration of GST-CD81-LEL needed to block HCVcc infection in our assay is comparable with previously published data from Lindenbach et al. (79). Interestingly, the amount of hCD81-LEL, eE2, or eE2-C656S required to prevent 50% inhibition of HCVcc infection was approximately the same, with a slight increase needed for eE2-C656S. Since the molecular mass of GST-CD81-LEL is about 37 kDa, the total amount of eE2 and CD81-LEL added was within two-fold. It is thought that CD81-LEL would bind to virus, while eE2 is presumed to bind to the cellular receptors. Therefore, the result that CD81 and eE2 can block infection at similar concentrations is surprising since the mechanism of blocking of infection is thought to be different. There are many factors that contribute to the similar inhibition results seen with the E2 and CD81 proteins, including the total number of CD81 and eE2 binding sites, the relative affinities of the interactions, and the kinetics of binding. At the moment, we are not certain which of these or other factors are responsible for the similar inhibition results.

The production of large quantities of functional and properly folded E2 ectodomain has allowed us to perform comprehensive biochemical and biophysical analysis of eE2, which was previously difficult due to production of mostly misfolded proteins in other expression systems.
Section II. eE2 Monoclonal Antibodies for Biochemical Study

The envelope 2 protein (E2) of hepatitis C virus (HCV) is the most highly variable protein encoded in the genome, with no more than 70-80% sequence identity across genotypes. This makes E2 a generally unattractive target for anti-HCV therapy, given the preference for drugs and vaccines that will effectively treat all or most patients. However, more than two decades of research on antibodies to E2 suggests that this is a course worth pursuing, with more and more evidence accumulating for the presence of conserved epitopes that elicit cross-reactive antibody responses.

In 1990, a group of researchers at Rochester General Hospital published their results on a purified fragment of yellow fever envelope protein (10). They found that rabbits immunized with this fragment produced antibodies that not only neutralized the homologous yellow fever virus, but also comparably cross-neutralized dengue-2 virus. In the same year, Roehrig et al. published their results on dengue-2 glycoprotein fragments tested in mice (108). Two of the peptides they tested could elicit antibodies that neutralized virus. Another account published two years later showed that monoclonal antibodies generated against hog cholera virus reacted with the homologous glycoprotein and neutralized virus (124). All of these results in HCV relatives implied potential for neutralizing antibody responses to HCV envelope protein 2, although this virus was only discovered and named 3 years prior.
By 1994, Michael Houghton’s team at Chiron Corporation had expressed the envelope 1/envelope 2 complex using recombinant vaccinia virus to infect suspension HeLa cells (recovering 1.5 mg from 120 L) (104). They used this material to immunize chimpanzees and found five out of seven animals were protected against homologous virus challenge (18). A year later, another group in Japan established the first Chinese hamster ovary (CHO) cell line expressing envelope protein 2; they were able to immunoprecipitate their E2 protein using antibodies from HCV infected patients (51). Importantly, a group at Stanford University published the first example of a particular mechanism by which the virus could be neutralized (50). Their collaborators had recently determined that CD81 was an important entry receptor for HCV (102), and here they reported a panel of antibodies that could bind to conformational epitopes of E2 across several genotypes, two of which inhibited the interaction of E2 with CD81. However, one important flaw to these experiments could not be overlooked: testing antibodies from chronically infected patients implies that these antibodies were ineffective at resolving infection and may therefore not lead to the best information about neutralization properties.

At the beginning of this thesis work, commercial or sustainable antibodies to the HCV envelope proteins were not readily available, owing to the extremely limited ability to produce E1 and E2 recombinantly. Therefore, one important aim here was to use our high-quality protein generated in HEK293T cells (as described in Section I) to generate a panel of monoclonal antibodies for various experimental purposes. Monoclonal
antibodies would a) facilitate biochemical methods that currently relied exclusively on affinity tags (including Western Blot and ELISA), b) enable a potential co-crystallization scaffold, and c) allow us to study cross-reactive and neutralizing epitopes exclusive of the confounding factors associated with HCV infection.

1. Production and Testing of Anti-eE2 Monoclonal Antibodies

In 2008, through collaboration with Arash Grakoui at Emory University, we used the J6 eE2 material described in Section I to inoculate mice. Briefly, BALB/C mice received 4 bi-weekly injections of 50 µg of purified (untagged) eE2 protein over the course of two months. Splenocytes were harvested and used to generate traditional hybridoma cell lines, which in turn were tested by Hannah Scarborough of the Grakoui laboratory. To validate the response, all successfully grown hybridomas were tested for their ability to bind eE2 purified protein using a standard ELISA technique described in Materials and Methods Section II. All supernatants testing positive for an anti-eE2 response were sent to the Marcotrigiano lab for further testing and characterization.

2. Anti-eE2 Antibodies that are Cross-Reactive to H77-eE2

Initial positive responses in anti-eE2 ELISA resulted in a panel of more than 40 monoclonal antibodies ready for comprehensive analysis by Western Blot. A list of these antibodies and a summary of the results is outlined in Table I. Supernatants from stable cell lines expressing eE2-Fc protein, using cell culture media as a negative control,
<table>
<thead>
<tr>
<th>Monoclonal</th>
<th>WB J6 R</th>
<th>WB J6 NR</th>
<th>WB H77 R</th>
<th>WB H77 NR</th>
<th>ELISA H77</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B6</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2A12</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>3E4</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3H4</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4C3</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5G1</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>5G12</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>5G9</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>6B10</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6D3</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6E8</td>
<td>-</td>
<td>+++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>6F5</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>6H12</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>8A6</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>9H9</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>12B11</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>12B8</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>12E10</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>12E6</td>
<td>+</td>
<td>+++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>12F10</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14C32</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14F7</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>14G1</td>
<td>+</td>
<td>+++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>14H6</td>
<td>+</td>
<td>++</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>15A2</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>15H12</td>
<td>-</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>15H5</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>16B10</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>17B6</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>18B11</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>18C5</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>18E4</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>18H3</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>19A10</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>19A3</td>
<td>+/-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>+</td>
</tr>
<tr>
<td>19B10</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>19B5</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>15C5</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>2D1</td>
<td>-</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2C1</td>
<td>++</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1F11</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3F5</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
<tr>
<td>3G4</td>
<td>-</td>
<td>-</td>
<td>n/a</td>
<td>n/a</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1. Biochemical summary of the eE2 monoclonal antibodies produced from mouse immunizations.
were used for these experiments. While purified protein could have been used for this experiment, the use of unpurified cell supernatant further highlights the specificity and affinity of these antibodies in an assay with relatively low sensitivity and high background potential. In strictly repeated parallel experiments with consistent exposure times, the antibodies recognized eE2-Fc with varying degrees of sensitivity as indicated by the use of one, two, or three “plus sign” symbols in Table I. Only 30% of the antibodies recognized the protein in both its reduced and non-reduced form (1B6, 3H4, 4C3, 5G9, 6D3, 6F5, 6H12, 8A6, 12E6, 14G1, 14H6, 2C1, 1F11), indicating the recognition of linear epitopes that would be undisturbed by reduction and denaturation. While 32% of the antibodies failed to recognize the protein at all by this method (5G12, 12B11, 12B8, 12E10, 14F7, 15A2, 17B6, 18C5, 18E4, 18H3, 19A10, 19B5, 3F5, 3G4), another 32 % of the antibodies recognized the protein only in its non-reduced state (2A12, 3E4, 5G1, 6B10, 6E8, 12F10, 14C32, 15H12, 15H5, 16B10, 18B11, 19B10, 15C5, 2D1). These data suggest that the protein elicits conformation-dependent antibodies to epitopes that are maintained by the intrinsic fold of the protein as well as folding maintained by disulfide-bonds. Two antibodies, 19A3 and 9H9, very weakly recognized the reduced eE2-Fc protein only, suggesting a response to an epitope that may be partially obstructed by the Fc tag on the unpurified form. The folded form of the Fc tag as presented in the non-reduced sample probably constitutes an even more significant obstruction than the reduced sample in this case. Several exemplary Western Blots, including 1F11, 14C32, and 8A6, are shown in Figure 12.
Figure 12. anti-E2 Monoclonal Western Blots.

Western blots for three different monoclonal sera are shown here to illustrate variable reactivity. Media is loaded as a negative control.
Considering the excellent results in Western Blot experiments against homologous J6 eE2-Fc protein, we next determined the response to genotype 1a protein produced by the same stable cell line production method. H77 eE2-Fc supernatant was loaded into an SDS-PAGE gel under both reducing and non-reducing conditions, using J6 eE2-Fc supernatant as a positive control. Only the ~65% (28 total) of antibodies that tested positive against any form of J6 eE2-Fc were tested for reactivity to H77 eE2-Fc. Of these 28 antibodies, 12 did not react at all to H77 eE2-Fc (1B6, 3E4, 3H4, 4C3, 6D3, 9H9, 12F10, 14C32, 16B10, 2D1, 2C1, 1F11). Nine antibodies only reacted to the non-reduced form of the protein and with much weaker affinity compared to J6 protein, indicating not only the importance of maintaining disulfide-linked epitopes for reactivity in this assay, but also the conserved nature of these epitopes (2A12, 5G1, 5G9, 6B10, 6E8, 12E6, 14G1, 15H5, 18B11, 19B10, 15C5). Of the remaining antibodies, four bound reasonably well to both reduced and non-reduced protein (6F5, 6H12, 8A6, 15H12). Western Blots similar to those shown in Figure 12 for J6 eE2-Fc protein are also shown for H77 eE2-Fc (Figure 13) for comparison.

Although only about 30% of the monoclonal antibodies were cross-reactive to genotype 1a protein, it was likely that cross-reactivity would improve in an experiment that preserved the conformation of the protein. Therefore we decided to go back and test these antibodies for reactivity in the original ELISA assay performed by Hannah Scarborough in the Grakoui laboratory. Using the same experimental technique described in Section II.1 for the characterization of the antibody panel, Hannah tested
Figure 13. H77 anti-E2 Monoclonal Western Blots.

Western blots for three different monoclonal sera are shown here to illustrate H77 cross-reactivity variability. J6 eE2 cell supernatant is loaded as a positive control.
purified H77 and J6 proteins in parallel. Impressively, 27 out of 43 antibodies recognized H77 eE2 protein with signals comparable to J6 eE2 (Figure 14).

3. Epitope Mapping and Purification of Anti-eE2 Antibodies

As described in the introduction for this section, antibodies to eE2 would prove most useful for therapeutic purposes if they could be generated against highly conserved and neutralizing epitopes. Although these antibodies were raised against eE2 protein outside the context of viral infection, we wanted to determine which epitopes were the most immunogenic. One traditional method for mapping epitopes is to use alanine scanning mutagenesis, which is more practical for studying smaller regions. Instead, we chose to produce truncated versions of the antigenic protein for our mapping studies. We generated a panel of eight amino-terminal (NT1...NT8) and eight complementary carboxy-terminal (CT1...CT8) truncation clones using the same cloning approach described in Section I (prolactin signal sequence, Fc tag). These clones were designed based on the approximate borders of hypervariable and highly conserved regions as described by Falkowska et al., with some slight compromises made to adjust for the positions of glycosylation sites (38). The truncation clones were transiently transfected in HEK293T cells and the supernatants tested to confirm secreted protein by anti-Fc Western Blot (data not shown). All truncation supernatants were used to coat an ELISA plate, then each antibody was tested for binding to each truncated protein in triplicate using full length J6 eE2-Fc as a positive control.
Figure 14. H77 Cross-Reactivity ELISA.

All monoclonal antibodies that tested positive for J6 E2 binding (red bars) by ELISA were tested for H77 E2 binding (blue bars) in parallel.
Based on the results of this ELISA, most of the antibodies could be mapped to one of seven defined regions of the protein, designated A through G (Figure 15). Antibodies that bound to all eight C-terminal truncation clones but to none of the N-terminal truncation clones would be defined as binding to hypervariable region 1 at the extreme N-terminus, Region A (2C1, 2D1, 1F11). Antibodies that bound CT1-CT7 and only NT1 and NT2 were categorized to Region B (3H4, 4C3, 6F5, 12B8, 14C32, 18E4). Antibodies that bound CT1-CT6 and NT1-NT3 were categorized to Region C (3G4, 8A6, 19E9, 18H3, 12E10, 14F1, 14H6). Although this region is defined broadly to include hypervariable region 2 and highly conserved region 2, it is possible that these antibodies map to the hypervariable region since this is expected to be more exposed. However, several antibodies mapping to this region are cross-reactive to H77 protein (Table I. and Figure 14). Antibodies that bound to CT1-CT5 and NT1-NT4 were categorized to Region D (5G9, 15C9, 19B5). Only two antibodies bound to CT1-CT4 and NT1-NT7, Region E (1B6, 12F10), which is interesting since this region is predicted to be hypervariable and likely to be solvent exposed. More expectedly, only two antibodies bound to the highly conserved Region F (CT1-CT2, NT1-NT7), 6D3 and 18C5, although neither of these is cross-reactive to H77 protein. Finally, antibodies that bound only to the N-terminal truncation clones were designated to Region G (12E8, 6E8, 14F7, 5G12, 19B10, 15H12, 16B10, 15H5, 6H12, 9H9, 14G1, 2A12).

Based on the results of the epitope mapping studies, we chose sample antibodies reactive to each region to attempt purifications by protein A affinity. This would help to
Figure 15. Monoclonal Antibody Epitope-Binding Map.

In order to establish approximate binding epitopes for our collection of monoclonal antibodies, complimentary amino- and carboxy-terminal eE2 fragments were expressed transiently and used an ELISA to test for monoclonal antibody recognition. This figure illustrates the defined binding regions (A-G) and lists the antibodies by region.
concentrate the samples and prepare them for more rigorous downstream biochemical applications. Hybridomas were expanded to 1 L spinner cultures by the Grakoui laboratory and sent to the Marcotrigiano laboratory for purification. Successfully purified antibodies included 1F11 and 2C1 (Region A), 6F5 (Region B), 14F1 (Region C), 19B5 (Region D), 1B6 (Region E), 6D3 (Region F), 16B10 and 2A12 (Region G), and 6B10 (not successfully mapped). Eluted fractions were loaded onto SDS-PAGE gels and stained with Coomassie; a sample elution for each antibody (under reducing conditions) is shown in Figure 16, with heavy chain and light chain indicated to the right. Note that the hybridoma cultures were early heterogeneous populations, therefore the amount of antibody produced varied amongst the clones.

4. Anti-eE2 Antibody Applications

In order to validate the usefulness of the monoclonal antibodies generated against eE2, the antibodies were tested in various biochemical experiments. As described in Section I, eE2 protein binds well to human CD81 large extracellular loop in an anti-Fc based binding assay. In order to test for CD81 binding of untagged protein, a few of the monoclonal antibodies were tested for their ability to function in this assay. Anti-eE2 antibody 3G4 gave the best results, with strong titrating signals for both purified eE2 and eE2-Fc in cell supernatants bound to human CD81, and minimal background binding to the negative controls (Figure 17).
Figure 16. SDS-PAGE for Purified Monoclonal Antibodies.

Antibodies were purified via Protein A column. Sample elutions are shown here under reducing conditions, with heavy chain (HC) and light chain (LC) indicated to the right.
Figure 17. Monoclonal 3G4 CD81-Binding ELISA.

3G4 was used to detect the interaction between CD81-LEL and eE2-Fc (diluted in DMEM) or purified eE2 (diluted in HEPES buffer). As expected, eE2 bound human CD81 only and the antibody showed minimal background in anti-mouse HRP quantification.
Next, we tested the activity of some of the purified antibodies in a neutralization assay against HCVcc expressing *Renilla luciferase*. Four purified monoclonal antibodies were tested (2C1, 2A12, 1F11, 14F7) in concentrations ranging from 0.02 µg/mL to 20 µg/mL. Both 1F11 and 2C1 (mapping to the hypervariable N-terminus) effectively neutralized virus at the lowest concentrations tested (Figure 18). 2A12 and 14F7 did not demonstrate a significant neutralization response.

5. Discussion

In 2006, the results of a Phase II clinical trial reported that an anti-E2 human monoclonal antibody under study for the treatment of HCV-infected liver transplant patients failed to significantly reduce HCV viremia below the limit of detection (112). Although their results were disappointing, the antibody was effective to some degree at higher doses and was well tolerated by the patients, suggesting the potential for this form of therapy in combination with other treatment regimens. Over the course of our antibody studies, a significant amount of work has been published supporting the role of E2 in eliciting neutralizing antibodies, with a great deal of emphasis toward a greater understanding of conserved and cross-reactive epitopes. In 2007, Michael Houghton again published on the subject. His group of researchers immunized guinea pigs with recombinant envelope proteins 1 and 2 and found that these immune sera could neutralize HCVcc and HCVpp, identifying the N-terminal domain as the most immunogenic part of E2 (118). In three different studies published in 2008, a specific E2 epitope was identified which gave rise to antibodies that could neutralize multiple genotypes of HCVcc
Figure 18. eE2 Monoclonal Antibodies Neutralize Virus.

Purified antibodies 2C1, 2A12, 1F11, 14F7 were tested for their ability to neutralize virus in concentrations ranging from 0.02 µg/mL to 20 µg/mL. 1F11 (red box) and 2C1 (green inverted triangle) neutralized virus at the lowest concentrations tested while 2A12 (yellow triangle) and 14F7 (blue diamond) did not demonstrate a significant neutralization response.
and HCVpp which overlaps with the CD81 binding domain (67, 75, 98). It was from one of these studies that the popular monoclonal antibody CBH-5 was identified (98).

Interestingly, another group published their results on an epitope spanning residues 434-446 (H77 sequence), which elicits antibodies that can interfere with the activity of neutralizing antibodies at a conserved epitope spanning 412-426 (132). These results likely explain why potent neutralizing antibodies are present in chronically infected patients, but still failed to address the issues of mutability.

Finally, in 2009, a report was published acknowledging the escape mutant conundrum of neutralizing human antibodies to E2. A group led by Steven K.H. Foung used HCVpp to examine a panel of neutralizing human monoclonal antibodies and identified compensatory mutations outside of the CD81-binding domain which allowed neutralization escape at a cost to viral fitness (66). This prompted a search for alternative epitopes that could elicit neutralizing antibodies while avoiding escape variants. In 2012, Keck et al. published their results, identifying nine antibodies that are cross-neutralizing and map to a previously unidentified epitope on genotype 2a E2 (68). This epitope is highly conserved across genotypes and antibodies to this region did not promote escape mutations in an HCVcc repetitive neutralization study. These results further support the potential for E2 in developing new treatments to HCV.

Aside from the potential clinical use of monoclonal antibodies against E2 and the valuable information that can be gleaned from studying neutralizing epitopes, the Marcotrigiano lab is specifically interested in using antibody fragments (Fab) as a
potential co-crystallization scaffold. As early as 1990, HIV capsid p24 was co-crystallized with an Fab for diffraction to 2.7 Å resolution. Here, the authors purified anti-p24 mAb25.4 from ascites fluid and digested the Fab fragments from the Fc using papain cleavage. The structure was ultimately solved in 1996 (92). In 1991, Jacobo-Molina et al. produced the first crystals of HIV-1 reverse transcriptase (RT), which diffracted to 3.5 Å resolution in complex with an Fab and double-stranded DNA (59). Two years later, the structure was solved to 3.0 Å resolution (60), then resolved to higher resolution in 1998 using alternative DNA template-primers (29). To further emphasize the usefulness of this technique, Peter Kwong published the structure of HIV gp120 in complex with a fragment of the CD4 binding domain and neutralizing Fab 17b (72).

In order to produce and purify significant quantities of Fab fragments for E2 crystallography studies, the Marcotrigiano lab acquired a specific collection of hybridoma cell lines from the Grakoui lab. We are now able to improve the clonal selection of these hybridomas by diluting cultures to individual colonies and testing these colonies for their E2 binding strength by ELISA. Depending on which E2 construct is tested in a co-crystallization experiment, Fab candidates may include 1F11, 8A6, 3G4, 14F1, 5G9, 1B6, 6D3, 18H3, 15C9, 12F10, and 18C5.
Section III. Examination of Disulfide Bonds in HCVcc E2 and eE2

In 1990, a group at Genentech led by Timothy Gregory performed a comprehensive disulfide-mapping analysis of HIV gp120 envelope protein, which contains 19 cysteine residues engaged in 9 disulfide bonds (78). Their experimental method relied on enzymatic digestion and deglycosylation of the protein as produced in CHO cells, followed by reversed-phased HPLC to successfully identify peptides engaged in disulfide linkages and map the disulfide pattern. Using their studies as a model, we intended to produce similar results with HCV eE2 produced in HEK293T cells, which has 18 absolutely conserved cysteine residues expected to be engaged in disulfide bonds. Of these eighteen, seventeen are present in the ectodomain of E2 (eE2), and one is present in the proposed stem region (32).

Disulfide bonds have been implicated in the processes of folding and structural integrity of the mature protein, helping to prevent degradation (9) and immune system recognition (63) while also assisting in the folding process (8). For viruses specifically, disulfide linkages in enveloped proteins can constitute a trigger mechanism, the rearrangement of which can catalyze the structural reorganization necessary to enable viral entry (2, 5, 13, 46, 61). For the hepatitis C virus, intermolecular disulfide bonds also help to stabilize heterodimer formation between envelope protein 1 and 2 in viral particles (121). Importantly, manipulation of the oxidation state often leads to misfolded and aggregated proteins that do not leave the endoplasmic reticulum (83).

At the beginning of this thesis work, the significance of these cysteine residues to HCV
viral replication, virion production, and infectivity had not been studied. Here, several experimental techniques used to address the roles of disulfide bonding in HCV eE2 are described.

1. Disulfide Bond Mapping Using Mass Spectrometry

With the successful results of eE2 glycosylation characterization described in Section I, an attempt to map the disulfide linkages using a similar experimental approach was executed. Samples of eE2 were digested with either trypsin or chymotrypsin and subjected to LC-MS for identification under reducing and non-reducing conditions. A manual search of the data identified several peaks corresponding to the expected m/z of disulfide linked peptides. However, depending on the experimental conditions for the sample, a few cysteine residues showed linkage to multiple other cysteines. Therefore, the method was considered to be flawed. In 2010, an alternative LC-MS approach was published, describing the map of 17 disulfide linkages for tissue plasminogen activator (127). This method relied on electron transfer dissociation (ETD) of fragments digested with combinations of trypsin, Lys-C, and Glu-C and successfully identified glycosylated peptides as well. This method has potential for the mapping of HCV E2 disulfides as well, depending on the availability of the right equipment and expertise.
2. A Study of HCV E2 Disulfide and Regional Mutants

Shortly after our attempts to map disulfide linkages as described above, another group of researchers used a similar method and published the HCV E2 disulfide map, utilizing this map and flavivirus homology as a scaffold upon which to thread a structural model (70). In addition, researchers demonstrated that reduction of up to half of the 9 disulfide linkages in E2 does not significantly affect receptor binding (39). Using these results as a foundation for further study, a mutagenesis approach was utilized in order to extrapolate more relevant information about the relationship between disulfide linkages and viral infectivity. In collaboration with Arash Grakoui at Emory University, we compared the results of cysteine-alanine mutations in HCVcc and in soluble eE2 mutants.

After sequence analysis confirmed the conservation of all 18 cysteine residues in the E2 ectodomain, individual mutations to alanine were produced in the context of J6/JFH-1 HCVcc (labeled C1A through C18A according to their relative position from the amino-terminus). A careful analysis of protein production and initial intracellular RNA replication indicated that the mutant viruses were comparable to wild type at this level although replication diminished overtime indicating reduced viral spread (data produced by G. Mateu, not shown). In fact, none of these mutants produced infectious virus particles as confirmed by measuring the extracellular release of core protein (G. Mateu, not shown), indicating the importance of disulfides in folding and assembly of virus particles. Interestingly, mutants C1A and C11A produced equivalent levels of core
following transfection, but only C11A had demonstrated any initial level of infectivity. However, C6A produced approximately double the amount of extracellular core than either C1A or C11A and a 10-fold higher level than the negative control (G. Mateu, not shown). This prompted further investigation of the C6A mutant and the mechanism by which this mutation prevents entry.

To zero in on the role of C6A in preventing viral infection, further alanine mutations were introduced in the surrounding region, including valine 502, glycine 504, valine 506, tyrosine 507, phenylalanine 509, and threonine 510. Mutants V502A, G504A, V506A, and Y507A retained very low infectivity, while T510A showed 10-fold less infectivity than wild type. Mutants C505A (C6A), C508A (C7A), and F509A were noninfectious (G. Mateu, data not shown). None of these mutants demonstrated a phenotype similar to the original C6A mutant virus, leading us to hypothesize that the production of non-infectious viral particles by the C6A mutant may be due to disulfide bond disruption. This disruption may have been directly leading to impaired receptor binding.

3. HCV eE2-C6A and Surrounding Mutants Influence hCD81 Binding

To further investigate the role of C6 in HCV infection, we sought to determine the effect of this mutation and the nearby mutations described above on CD81 binding, using purified recombinant E2 and large extracellular loop (LEL) of human CD81. A variety of previous research has demonstrated that CD81 is necessary for HCVcc infectivity (80, 122, 133). The soluble form of eE2 used here was described in Section I and has been
used to characterize E2-CD81 interactions. Briefly, eE2 mimics native E2 in human CD81 (hCD81) binding, blockage of HCVcc infection, and recognition by antibodies from patients chronically infected with HCV (126) (70). Since substitutions in the highly conserved region of amino acids 480 to 510 were shown to impair infectivity, the hCD81 binding ability of each of these mutant recombinant eE2 proteins was analyzed in vitro as previously described (126). As shown in Figure 19A, T510A recombinant protein binds to hCD81 almost as well as wild type eE2, while mutants V502A, G504A, V506A, Y507A, C7A, and F509A bind at only 30% to 60% the capacity of wild type eE2. Interestingly, the recombinant mutant protein C6A showed the weakest interaction with CD81, with less than 20% binding capacity compared to wild type eE2.

To determine whether the lack of CD81 binding was due to changes in secondary structure, circular dichroism (CD) analysis was performed on eE2-wt and eE2-C6A. The results showed that the CD spectra of C6A was identical to that generated by wild type eE2, indicating that substitution of C6 by alanine did not generate gross changes in the secondary structure of eE2 (Figure 19B). These results suggest that this region likely harbors a CD81-binding domain, and provide compelling evidence that the C6A mutant was unable to establish productive infection in cell culture due to a defect in CD81 co-receptor binding. Further experiments performed by Elizabeth Elrod (Grakoui lab) have confirmed this prediction; purified mutant virus was used in an adapted CD81 binding assay and was indeed impaired in hCD81 binding compared to wild type (data not shown).
Figure 19. Mutations Affect hCD81 Binding but Secondary Structure is Conserved.

A) ELISA for CD81-binding demonstrates greatly reduced binding efficiency for eE2-C6A compared to wild type eE2 and surrounding mutants. B) Circular dichroism analysis shows that the secondary structure for eE2-C6A is identical to wild type eE2.
1. **eE2-C6A is Defective in Blocking HCV Infection**

As with the blocking experiments described in Section I, we wanted to determine whether eE2-C6A mimicked the function of wild type eE2 *in vitro* in terms of blocking HCVcc infection (70, 126). Huh-7.5 cells were incubated for three days with 100 TCID$_{50}$ of Cp7 virus mixed with 50 µg/ml of purified eE2, eE2-C17S, eE2-C6A, hCD81, or GST. eE2-C17S was included as a control since we previously showed that this recombinant protein was recognized by antibodies from patients infected with HCV, blocked HCVcc infectivity, and bound hCD81 (126). As shown in Figure 20, control proteins eE2, eE2-C17S, and hCD81 blocked 60-70% virus infectivity while C6A blocked only 25%. These results provide further support for the role of C6 in binding to CD81.

2. **Discussion**

This section emphasizes the importance of disulfide bonding in maintaining the structural integrity of eE2, demonstrating that mutations in highly conserved cysteine residues cannot be tolerated. Although the disulfide linkages were difficult to map by mass spectrometry, our results with HCVcc mutants and recombinant mutant eE2 protein were more informative.
Figure 20. eE2-C6A is Impaired in Blocking Infection.

Purified eE2, eE2-C17S, and hCD81 blocked 60-70% virus infectivity as expected, while C6A blocked only 25%. GST alone was tested in parallel as a negative control.
Previous research has shown that point mutations in the E2 protein can alter the incorporation of E1 and E2 in HCVpp even when these proteins seem to run similarly to the WT protein in Western blots (74). Therefore, we expected that mutagenesis of cysteine residues would have a serious detrimental effect on E2 folding and incorporation into HCVcc particles. Indeed, none of the cysteine mutant genotypes produced infectious viral particles, nor was there any increase in viral particle production over time. This reduction in particle replication with serial passages was attributed to lack of viral spread, suggesting that these mutants either failed to revert or failed to generate compensating mutations to produce infectious virus. These results are further supported by the observation that disulfide bonds are necessary for proper folding of E2 in the ER, helping to stabilize the basic structure of the protein (70). This is in contrast to many other viruses, where cysteine residues instead function in viral fusion by mediating changes in the secondary structure upon exposure to low pH (5) (116) (97) (2) (109) (88) (46) (13) (61).

Interestingly, cysteine 6 exhibited the most unique phenotype by resulting in a measurable amount of non-infectious particle release. Krey et al. produced a model for E2 in which C6 is disulfide-linked to C7, wherein the surrounding peptide region plays a possible role in virus fusion. However, a role in CD81 binding has not been described for this region of the protein. We therefore took careful measure to determine the role of C6 in direct binding to CD81 by introducing mutations in the conserved proximity of C6 and C7. The results indicated that the disruption of this disulfide bond determined the
behavior of the C6A mutant, since none of the surrounding mutations in the putative fusion peptide abolished infectivity without also affecting particle release.

Once the physiological relevance of the mutations was established in HCVcc, we addressed the mechanism of C6A dysfunction and found that changes associated with this phenotype directly impaired CD81 binding. As stated above, this region has not been previously shown to play a role in CD81 binding. Therefore, we recapitulated the C6A and surrounding mutations in recombinant protein. Here we showed that C6A, V502A, G504A, V506A, and Y507A mutants were impaired in their ability to bind CD81 (Figure 19A), implicating this domain in CD81 binding. Although C6A may be directly involved in CD81 binding, it is also possible that this mutation caused local minor changes undetectable by CD analysis, or that the resulting free cysteine (C7) yielded an E2 protein that is no longer able to bind CD81.

The next step was to investigate the ability of the eE2-C6A protein to block infection of HCVcc in Huh7.5 cells. Indeed, eE2-C6A was dramatically impaired in its ability to block infection when compared to wild type and surrounding mutant recombinant proteins. Again, this suggests the possibility that CD81 binding was prevented. The results were confirmed by HCVcc CD81 binding experiments, in which C6A mutant virus was impaired in CD81 binding.

In a similar approach to understanding E2 disulfide linkages, McCaffrey et al. introduced alanine point mutants at all 18 conserved cysteine residues of genotype 1a H77 E2 in
both HCVpp and recombinant E2 ectodomain. They found that 12 out of 16 cysteines in the E2 ectodomain were critical for CD81 binding. However, they could not draw conclusions about the roles of all disulfides in assembly and egress, which precludes their role in CD81 binding, because only a subset of mutants were ultimately tested in HCVcc. Nevertheless, in agreement with our results, they found that the C6A mutant was unable to bind CD81 in an HCVpp immunoprecipitation with CD81-LEL.

One of the most perplexing questions that arises as a result of this study is why proposed mutations of complementary cysteines (based on Felix Rey’s model) do not necessarily produce the same viral phenotype. Initially, we were inclined to question our data and re-think the interpretation of our results. However, McCaffrey reached similar conclusions with HCV pseudoparticles, indicating that the disulfide map may be correct and the experiment mapping methods insufficient to propose Rey’s model. Alternatively, it is possible that one cysteine in the pair has a greater effect on the immediate local environment than the other. As suggested by McCaffrey, this may be tested by substituting with a more conservative mutation than alanine (serine, for example) (87). Finally, an article published in 2011 provides evidence for free thiols in a pre-attachment E2, while an earlier study is in complete contradiction (39, 45). It is still possible that there is a difference in disulfide linkage in pre- and post-attachment forms of E2.
In summary, these results have shown that a cysteine 6 mutant genotype is capable of generating non-infectious virus particles. This mutant did not revert, was capable of assembling competent particles, but was incapable of infecting naïve cells. If these particles can elicit neutralizing antibodies in an immune-competent mouse model, this work could present the foundation for an attenuated HCV vaccine.
Section IV. Production of Recombinant Protein Using Lentiviruses

Production of recombinant proteins relies traditionally on expression systems developed in bacteria or lower eukaryotes (yeast, baculovirus, etc.) owing to their rapid growth and inexpensive media. However, as discussed above, HCV E2 (and many important human proteins) has posttranslational modifications (e.g. glycosylation and disulfide bonding) that make production in non-mammalian cells extremely challenging, since these expression systems fail to accurately replicate these modifications. Based on published results for the expression of E2 using various techniques and expression systems, mammalian cells were the best choice in spite of the drawbacks.

Unfortunately, research in mammalian protein expression has progressed slowly over the last decade, owing to the high cost of media, serum, and disposable plastics, technically skilled personnel and facility requirements. However, the utilization of the Fc tag in combination with a high capacity adherent bioreactor had made our system choice more practical. From 2007 to late 2010, the work presented for this thesis was based on the stable selection of HEK293T cells producing the protein of interest (as described in Section I). In 2010, we determined that the Fc tag was responsible for undesirable multimerization, which prevented more complex biophysical studies. We switched to the Protein A tag to solve this problem, preserving the high affinity and highly specific purification step essential to isolating eE2 from cell culture supernatants. This resulted in a major decrease in protein yield, compelling us to explore alternative methods to amplify expression.
Early on, recombinant protein expression in mammalian cells was performed by transient DNA transfection of suspension cultures using the CaCl₂ method (65, 113). While this type of method has remained the most common for large-scale protein production in mammalian cells, there are significant drawbacks. For example, protein expression levels slowly decrease since the DNA plasmids are not integrated into the host genome and are therefore lost during several rounds of cell division. If more protein is needed, the entire process must be repeated costing time and money. Consequently, the expression is limited to a few weeks per round of transfection.

Production of the stable expression cell line previously used for eE2-Fc utilized a DNA plasmid with a hygromycin selection marker. This helped to circumvent the transient expression pitfalls, but required more than a month to establish each mutant or variant.

Attempting to improve expression yields and feasibility of mammalian expression, Aricescu et al. published a description of HEK293 adherent cells transiently transfected using polyethylenimine (3). While they reported the recovery of 1-40 mg/L of media (4 roller bottles) amongst 24 protein targets, this method required excessive plastics/disposable usage, which is both labor-intensive and costly. A review published in the same year described the large-scale transfection of various suspension cell lines, outlining optimal transfection reagents, expression vectors, and media (101).

Volumetric limitations and yield compromises were discussed but no significant improvement in the technology was reported. In 2009, Lee et al. described a similar approach, where HEK293T cells were transfected with milligram quantities of DNA and
grown in Corning CellStacks (77). Still, large amounts of DNA and cost-prohibitive transfection reagents and plastics made this approach unrealistic for our applications.

To address the ongoing challenges associated with recombinant E2 expression in mammalian cells, we integrated several different elements of cell culture technology for a synergistic result. The BelloCell-500 bioreactor (Cesco Bioengineering, Taiwan) was chosen for its ideal investment/output ratio. Its innovative design addresses several adherent cell culture challenges (123). Relative to the systems described above, the BelloCell a) consumes significantly less disposable products and therefore produces less waste, resulting in a substantial cost savings of almost 50 % over roller bottles b) utilizes a three dimensional attachment matrix, allowing for better use of space and a higher degree of cell density and c) provides a more ideal growth environment in terms of temperature distribution and gas exchange. The BelloCell operates on a programmable stage called the BelloStage 3000, which holds 4 bioreactors and fits inside a standard cell culture incubator. In our hands, a single bioreactor can be maintained for 3 months with continuous media harvesting and replacement. Successful protein expression was achieved for both HEK293T and HeLa cells, resulting in as much as a 10-fold increase in production over conventional roller bottles. However, the use of this bioreactor alone was no longer sufficient for expression of E2 with the new Protein A tag (eE2-ProtA), so we explored ways to amplify expression within the BelloCell system.
1. **Lentivirus Expression of eE2-ProteinA**

The most important component to overcome the Protein A barrier and improve the efficiency of our adherent cell method was the introduction of second generation (3-vector components) lentiviral expression vectors. Lentiviral vectors have emerged over the last decade as powerful, reliable, and safe tools for stable gene transfer in a wide variety of mammalian cells (93). In addition, lentiviral expression systems provide a cost- and labor-effective method of generating stably expressing cell lines within a short period of time -- on the order of days to a week. Lentiviruses are a genus of the *Retroviridae* family that can efficiently and conveniently infect a wide variety of mammalian cells, including non-dividing cells. Upon infection, the viral RNA genome is reverse-transcribed into DNA, which is transported to the nucleus and integrated into the genome by a viral-encoded integrase enzyme. The use of separate plasmids (replication vector, glycoprotein vector, and packaging vector) prevents recombination of genetic material to produce an intact viral genome, resulting in a pseudotype lentivirus capable of only a single round of infection. The retroviral envelope glycoproteins have been replaced with the glycoprotein protein of vesicular stomatitis virus (VSV-G) to allow for greater host cell tropism and added stability. The entire process of stable cell line production using lentiviral vectors can be performed in a week, circumventing the time consuming procedure of drug selection.

More important than time efficiency, lentivirus expression also results in a substantial increase in protein yield due to the high copy number of expression cassettes.
introduced per cell. The number of copies of the expression cassette introduced into a cell correlates to the number of virus particles entering the cell, called the “multiplicity of infection” (MOI). Our protocol, adapted from John Shires at Emory University, is designed such that each cell is infected at an extremely high MOI (100-1,000). This eliminates the need to culture cells in hygromycin selective media to identify clones that contain the expression cassette. Figure 21 illustrates an approximate 100-fold increase in expression for eE2-ProtA in lentiviral infected cells (right) versus the original eE2-ProtA stable clone (left).

Originally described by Trono et al., two components of the lentiviral system described here are currently available through AddGene (93). pMD2.G encodes VSV glycoproteins for incorporation into the viral envelope and versatility of cell tropism, and psPAX2 contains the standard HIV Gag/Pol cassette. The third plasmid, pJG, was re-engineered to further improve viral titers. pJG uses a CMV promoter to drive expression, with additional enhancement provided by the Rev Response Element and a woodchuck hepatitis promoter response element (Figure 22) (30). In both its current and original design, this vector contains a green fluorescent protein (GFP) gene immediately following the cloning site, separated by an internal ribosome entry site (IRES). This provides an early visible marker for infection efficiency, which correlates to the expression of the protein of interest (84).
Figure 21. SDS-PAGE eE2-ProtA Comparison.
eE2-ProtA was purified via IgG and eluted with low pH. Elutions are shown for BelloCell supernatant using standard transfection and stable clone selection (left) versus lentiviral infection (right).
Figure 22. pJG Vector Features and GFP Expression.

A) The pJG vector includes a CMV promoter and enhancement elements including the Rev Response Element (RRE) and woodchuck hepatitis promoter response element (WPRE). B) GFP is separated from the cloned gene by an IRES to provide a visual correlation of expression.
Finally, to address the potential problem of glycosylation heterogeneity produced in HEK293T cells, the cell culture media was supplemented with mannosidase inhibitors, producing more uniform glycosylation. We tested several inhibitors over the course of this work, including NB-DNJ and swainsonine, but decided that the best option in terms of cost, availability, and glycan homogeneity was the inhibitor kifunensine. First described in 1990, kifunensine is used to prevent complex and hybrid glycosylation from progressing throughout the secretory pathway (36). The effectiveness and expression results with the use of kifunensine have been thoroughly described (15). This compound works at 1 mg/L and conveniently doubles the protein yield, although the mechanism for this enhancement is not clear. The combined expression results of the BelloCell bioreactor, lentiviral infection, and kifunensine addition increased the yield of eE2 100-fold over plasmid/stable cell culturing.

2. Lentivirus Expression of eE1-ProteinA

The extraordinary success of eE2-ProtA expression using the lentivirus method prompted us to revisit a 5-year persistent struggle: the expression of envelope protein 1 (E1). The prevailing opinion in the literature, from the 1990’s to the present, is that E1 cannot be expressed in the absence of E2 and in fact probably depends on co-expression of E2 for proper folding (33). One group of researchers described their recent experiences studying neutralizing antibodies to the E1E2 complex as follows: “...isolated folded E1 is unavailable, and our attempts to express soluble E1 in mammalian
and *Drosophila* systems resulted only in misfolded aggregates,” (75). Undeterred by this general dogma, we have explored many different E1 expression constructs and expression methods since 2007 and were successful only in expressing E1 transiently or from cell lysates and only at levels detectable by Western blot.

J6 eE1 was cloned into the pJG lentiviral vector as described in the Materials and Methods. A 3x Flag tag was fused to the carboxy-terminus of the protein A tag, since there are no readily available antibodies for direct immunodetection of E1. The first attempt to produce the clone resulted in yields of approximately 10 mg/L (Figure 23). Shortly after my original E1 expression, postdoctoral fellow Abdul Khan (in our laboratory) was able to reproduce the expression results with H77 E1 and Con1 (genotype 1b consensus) E1. So far, the Con1 E1 protein has the least aggregation of all three genotypes tested and biophysical experiments are underway.

**3. Discussion**

The patent for this experimental method was submitted in June 2012. Although the system is should be useful for numerous experimental approaches, the concentration of protein achievable using the lentivirus system in conjunction with the BelloCell bioreactor should facilitate crystallography for both E1 and E2. Although both proteins can be subjected to deglycosylation, the glycans probably contribute to their
Figure 23. eE1-ProtA Purification

Lentivirus expressed eE1-ProtA was purified from BelloCell supernatants using IgG and the eluted fractions were loaded onto SDS-PAGE and stained with Coomassie.
unusually high solubility and both proteins can reach concentrations nearly impractical for crystallization trials. Considering the expected local concentration of the envelope glycoproteins on the surface of a viral particle, this is not surprising. This emphasizes the importance of the expression capacity of this system.

Furthermore, the ability to express significant quantities of both E2 and E1 opens new possibilities for the study of their interactions with each other and with cellular receptors. Abdul Khan has successfully expressed another challenging target using this system, HCV entry factor scavenger receptor class B type I (SR-BI). Since the Marcotrigiano lab also has the capacity to express Claudin-1 and CD81-LEL using bacterial expression, this collection of entry factors and envelope proteins should enable experimental approaches beyond what has been reported for other laboratories. Binding assays using ELISA and isothermal calorimetry are underway.
Section V. Optimization of eE2 for Crystallization Studies

The crystallization of glycoproteins is a challenging task for many reasons, as has been discussed in previous sections of this thesis. There are several examples of glycoprotein structures in the literature which we have referred to for guidance, though each one uses quite different methods owing to the empirical nature of crystallography studies. One of the earlier examples, mentioned previously in this thesis, is HIV gp120. This protein was produced in drosophila S2 cells, which are limited in the complexity of glycosylation produced; the protein was ultimately deglycosylated and its structure solved in complex with a CD4 (HIV receptor) fragment and Fab (72). Then, in 2004, a flavivirus group published the pre-fusion structure of the dengue virus envelope glycoprotein produced in the same S2 cells, but without considering glycosylation as a problem (91). Another group of researchers published the gp120 structure of SIV produced in Hi-5 (cabbage looper) insect cells, again without deglycosylating the protein (16). The deletion of three flexible variable loops enabled sufficient stability to crystallize the protein without additional ligands, suggesting that the contribution of the variable loops to the flexibility of the protein was a more significant problem than the contribution of the glycans. Indeed, they acquired resolution for all 13 glycans. One year later, the structure of the Epstein-Barr virus envelope protein, gp350, was published to 3.5 Å resolution. Using SF9 insect cells, this group also crystallized glycosylated protein and acquired resolution for 14 out of 15 glycans, indicating intermolecular interactions for many (119). Shortly after, Jeroen Mesters published an
article on the crystallization of glycoproteins, emphasizing their ability to participate in
crystal contacts regardless of heterogeneity (64). Furthermore, a search of the Protein
Data Bank returned hundreds of results for carbohydrate-modified structures (a simple
search for “glycosylated structure” returned 272 entries). Finally, the structure of Ebola
virus glycoprotein was published in 2008 solved in complex with an antibody; this
protein was produced transiently in HEK293T cells followed by deglycosylation (76).

These reports on the structures of glycoproteins suggested many possible method for
purification and crystallization of the HCV envelope 2 protein. With an ideal system in
place to express large quantities of E2 protein in a manageable time period, it became
feasible to explore clone/construct optimization of the protein for crystallography
studies. Although most of our efforts had previously focused on a full-length
ectodomain construct (residues 384-664 of the polyprotein), one of our first priorities
was to examine disorder.

1. Reduction of E2 Disordered Regions: HDX and Limited Proteolysis

In 2009, the Montelione group here at Rutgers published a highly sensitive method for
examining protein disorder by mass spectrometry (114). Their technique exploits the
rapid rate of exchange between hydrogen and deuterium ions (HDX) within relatively
disordered regions of proteins, combined with the high sensitivity of liquid
chromatography-mass spectrometry to identify exchanges within seconds. We provided
a sample of our J6 E2 purified protein to the mass spectrometry facility at CABM,
including buffer prepared with D$_2$O. Haiyan Zheng performed the experiment according to their protocol, although the digested protein fragments had to be searched manually due to the presence of glycosylated residues. Although peptide coverage was ultimately still incomplete, we were able to draw some critical conclusions about the flexibility of the protein. The amino-terminal portion of E2 showed a very high rate of exchange at the shortest time period tested all the way through methionine residue 449 (residue 57, Figure 24). Additional disorder, though less extensive, was predicted through to tyrosine residue 460 (residue 75, Figure 24). The Precission Protease cleavage site at the carboxy-terminus showed a surprising low rate of exchange, but the carboxy-terminus of the protein was highly disordered from glutamic acid residue 659 (residue 260) to serine 664 (residue 271).

We decided to generate a variety of amino-terminal truncation clones to minimize disorder without interrupting disulfide linkages, referred to in Figure 25 as NTD1 through NTD5. We also concluded it would be prudent to truncate the carboxy-terminus to arginine residue 656, eliminating the last cysteine in the sequence which we suspected to be unpaired. Each construct was cloned into the lentiviral backbone and expanded to the BelloCell stage for purification and analysis by gel filtration. Only two of the clones resulted in the purification of monomeric protein as shown in Figure 26. NTD1 was the longest construct and did not exclude any cysteine residues compared to the original construct. NTD4 excluded two cysteines, yet preserved the monomeric species. Interestingly, attempts to exclude a single cysteine (NTD2 and NTD3) or
Figure 24. HDX Results for eE2.

The eE2 sample was diluted with buffer containing D$_2$O and quenched at 10 sec, 100 sec, and 1000 sec intervals. The rate of exchange is indicated by the color scale, where red represents a high rate of exchange and therefore an exposed/flexible portion of the protein and blue represents the lowest rate of exchange and therefore an unexposed/stable region. Areas in white were not covered in a manual search of the data, possibly due to disulfide bond interference.
Based on the HDX results, five new amino-terminal truncation clones were designed based on the potential interruption of disulfide bonds. These were designated NTD1-NTD5 and were intended to explore the most significant deletion possible while maintaining monomeric protein.

**Figure 25. eE2 Truncation Clone Design.**

NTD1: aa433 (smallest deletion)
NTD2: aa470 (-1 Cysteine)
NTD3: aa426 (-1 Cysteine)
NTD4: aa449 (-2 Cysteines)
NTD5: aa456 (-3 Cysteines)
Figure 26. Gel Filtration Results for N-Terminal Deletion Constructs.
Each of the five new amino-terminal truncation clones was scaled-up to the BelloCell level and purified for gel filtration analysis on a Superdex 200 column. As shown, NTD1, which deletes no cysteine residues, yields monomeric protein. NTD4, which excludes two cysteine residues, also results in monomeric protein.
exclude 3 cysteines (NTD5) resulted in dimers or aggregated species. Since NTD4 was the shortest construct resulting in a monomeric species, all work from that point forward was performed with this deletion construct.

Before moving forward with crystallization attempts, we confirmed the integrity of the shortened eE2 construct by limited proteolysis. The protein was deglycosylated with Endo H and subjected to protease digestion with trypsin, chymotrypsin, and glu-C at both 4°C and 20°C and for time points ranging from 15 mins to 4 hours. Results indicated an extremely stable core fragment, with only minor proteolysis occurring predominantly in the 20°C sample and at the 4 hour time point (Figure 27). Chymotrypsin showed the most significant amount of proteolysis, with some activity at both temperatures. We concluded from this experiment and from analytical ultracentrifugation data acquired by Paul Leonard of Ann Stock’s lab (data not shown) that it was reasonable to move forward with this construct.

2. Deglycosylation of eE2 Reveals Two Species

Initial trials with J6 eE2 indicated that the protein may still be refractory to crystallization. In spite of the mixed results described by various groups in the introduction of this section, we acknowledged that our issue could be due in part to the heterogeneity of the glycans, since they account for more than half of the molecular weight of the protein. Even with the addition of the kifunensine inhibitor, it
Figure 27. Limited Proteolysis for eE2-NTD4.
eE2-NTD4 was deglycosylated with Endo H and subjected to protease digestion with trypsin, chymotrypsin, and glu-C at both 4°C (left) and 20°C (right) at the indicated time points. Results indicated an extremely stable core fragment, with minor proteolysis occurring predominantly in the 20°C sample at the 4 hour time point, with the exception of chymotrypsin.
remained likely that we were generating a range of high mannose glycans and thought it may be best to trim them all back. For this reason, we adapted a large-scale Endo H deglycosylation protocol and began producing Endo H in our laboratory for greater cost-effectiveness (production of Endo H is described in the Section V. Materials and Methods). The standard protocol for Endo H digestion requires pH 5.5 sodium citrate and heating the sample to 37°C for 1 hour. We believed that heating the sample to this temperature could compromise structural integrity and adapted our deglycosylation procedure to 4 hours at room temperature. As indicated by a Coomassie stain of deglycosylated protein, the experiment never resulted in a single sharp band (Figure 28).

According to work published in 2002 by Hogg et al., a protein that exists as a mixture (part glycosylated, part non-glycosylated), can still be suitable for crystallization studies (55). In their experiments with pokeweed antiviral protein, they found that the crystallization conditions selected for the glycosylated version of the protein in the mixture. The glycan itself played an integral role in the homodimeric contacts. Therefore, we attempted crystallization trials with the mixture, while working toward production of a single species. In our first attempt, we expressed the same construct in HeLa cells to see if the discrepancy was cell-line related. However, the deglycosylated protein looked identical to protein produced in the HEK293T cells (data not shown). Our next thought was to express the identical construct for alternative HCV genotypes, expecting that the degree of heterogeneity of sequence and glycosylation
Figure 28. Purified eE2-NTD4 Deglycosylated with EndoH.

The pH of the purified eE2 solution was adjusted to 5.5 using sodium citrate buffer, then the protein was deglycosylated with a 1:20 mg/mg quantity of EndoH for four hours at room temperature. This resulted in the presence of two predominant species as indicated by Coomassie stained SDS-PAGE.
sites may offer a natural solution to the problem. A sequence alignment is provided in Figure 29A, in addition to a chart summarizing the relative homology amongst the genotypes expressed (Figure 29B). Fortunately, we were able to construct homologous clones (as indicated by the position of the arrows in Figure 29A) for all genotypes, although significant expression was only achieved for J8, QC69, HK6a, and H77. As indicated in Figure 30, all genotypes tested produced the same mixture of species as J6. Although this was discouraging, we learned that genotype H77 expressed at higher quantities than any other genotype, including J6. Figure 31 shows the expression levels of J6, H77, HK6a, and SA13 purified under comparable conditions and loaded onto gel filtration. This led us to pursue the H77 clone for biophysical studies from this point forward.

Initial crystallography attempts focused on deglycosylated protein as described above. Although some material precipitated during the deglycosylation process, the supernatant was easily clarified by high-speed centrifugation. Conveniently, deglycosylated eE2 binds well to heparin, one of HCV’s proposed entry factors (54). This step allowed the separation of aggregated material and resulted in a highly concentrated elution of approximately 5 mL of monomer, an ideal volume for loading onto the gel filtration column. High-throughput crystallization screens were set up for various Hampton and Qiagen kits at both 4°C and 20°C using the GryphonHT robot in Dr. Eddy Arnold’s laboratory. As shown in Figure 32, the closest we came to
Figure 29. Sequence Comparison of Various HCV Genotypes.

A) A sequence alignment for all 7 HCV genotypes, with NTD4 construct borders indicated by arrows. B) This chart summarizes the % Identity and % Similarity of all genotypes with respect to J6 (2a).
Figure 30. Deglycosylated eE2-NTD4 for Various Genotypes.
As indicated by Coomassie stained SDS-PAGE, when protein of various genotypes was deglycosylated with Endo H, the result was still a mixture of two species.
Figure 31. Gel Filtration Results for Various eE2 Genotypes.

As indicated by Superdex 200 gel filtration results, H77 (genotype 1a) in orange expressed most strongly, although all resulted in monomeric species.
Figure 32. eE2 Quasi-crystalline Material

High throughput screening in 96-well nano-drop plates gave promising results in 0.1 M BIS-TRIS propane pH 7 1.5 M Li₂SO₄. Results could not be repeated manually.
crystal-like formations was in 0.1 M BIS-TRIS propane pH 7 with 1.5 M Li₂SO₄ at a protein concentration of 8 mg/mL. After struggling to repeat this condition in manual screens, we followed the advice of several other crystallographers and began to set up crystallization screens with fully glycosylated protein.

First, the purification protocol for eE2 had to be adjusted since the heparin binding site is obstructed by the glycans (data not shown). Fortunately, the cleaved protein A tag and uncleaved eE2-ProtA both bind to an anion exchange column at pH 7.5. Although this type of subtractive chromatography is sub-optimal, a final gel filtration step resulted in highly pure eE2. However, it was still necessary to consider that the heterogeneity of the high-mannose glycans produced with kifunensine may inhibit crystal formation. We therefore tested the expression of the protein in the N-acetylgalactosaminyltransferase-deficient (GnTI-) HEK293 cell line that was used for the structural determination of rhodopsin in 2002 (105). A comparison of mass spectrometry results for deglycosylated eE2 protein produced in HEK293T with kifunensine versus protein produced in GnTI- cells showed a calculable difference in the number of mannose residues applied to the single remaining glycan (Figure 33). The expected molecular weight of the protein without glycans is 23390 Da. For kifunensine treated protein (left panel), an addition of 1 GlcNac residue on each of 6 sites (+203 Da), plus a single Man₉GlcNac₂ (+1864 Da), minus 2 Da per disulfide bond equals 25460 Da, corresponding well to the experimentally determined value of 26462 m/z. For GnTI- produced protein (right panel), an addition of an addition of 6 GlcNac residues (+203 Da), plus a single
**Figure 33. Mass Spectrometry Results for Deglycosylated eE2 –NTD4.**

eE2 was deglycosylated and purified, then subjected to MALDI-TOF in order to compare the m/z difference between protein produced in HEK293T + kifunensine (left) or in HEK293 GnTI- cells (right). As expected, the m/z corresponded to the predicted molecular weight plus 1 Man$_3$GlcNac$_2$ (left) and 1 Man$_3$GlcNac$_2$ (right).
Man$_3$GlcNac$_2$ (+ 1216 Da), minus 2 Da per disulfide bond equals 25812 Da, corresponding well to the experimental value of 25803 m/z. With fewer mannose additions, the relative homogeneity of the GnTI- produced protein was considered superior. Crystallization trials continued with this form of the eE2 protein. The lentiviral-infected GnTI- cell line expressing H77 eE2-NTD4-ProtA secreted protein at quantities between 35 and 40 mg/L at peak production. Although 50% of the protein mass at this stage is accounted for by the ProtA tag alone, the importance of this secretion level is emphasized by the fact that the glycosylated form of the purified protein concentrates effortlessly beyond 150 mg/mL. This concentration of protein must be handled manually due to viscosity. Fortunately, high-throughput screens at lower concentrations indicated that the eE2 protein showed activity in PEG. All manual screens set up in 48-well or 24-well trays using hanging drop vapor diffusion were designed to screen around various forms of PEG ranging from PEG400 to PEG6000 and mixtures of different PEGs. The most favorable conditions appeared to be around 9% PEG, with optimal activity in PEG3350, followed by a 50% mixture of PEG400 and PEG3350. Phase separation was common in most PEG conditions at 13% or higher. None of the conditions tested resulted in crystal formation. Further experiments will be performed, testing co-crystallization conditions with Fab fragments as described in Section II.
CONCLUSION

We recently discovered a company in Massachusetts, Immuno Diagnostics, Inc, selling purified recombinant H77 E2 produced in eukaryotic cells for $3,950.00 per 10 µg. A second company, Austral Biologicals, provides product information for HCV E2 produced in CHO cells, priced at $395.00 for 5 µg, although this has been “temporarily unavailable” since our earliest awareness of the product in 2008. The value of this material for many types of studies and the difficulty in producing it are apparent considering this information as well as results described in recent literature. Although other groups have repeated the expression of recombinant E2 in much lower quantities, we are now in a position to do experiments that other researchers cannot. While crystallographic studies are ongoing, the usefulness of the recombinant protein we are able to produce has enabled us to establish collaborative studies with a number of outstanding researchers across the globe.

Since 2007, we have collaborated with some of the leading scientists in the HCV field. Some of our earliest work with Dr. Charles Rice at Rockefeller University helped to validate the antibodies we produced against our recombinant protein (data not shown). More recently, we collaborated with Dr. Rice and Dr. Roberto Cattaneo at the Mayo Clinic in an experiment to test purified eE2 in a vaccine boost strategy (106). In this study, HCV envelope proteins from genotype 1a (H77) were engineered into a recombinant measles virus, which was used to immunize mice. When immunized mice
received a vaccine boost with recombinant H77 (1a) eE2 protein, this elicited an antibody response that was cross-neutralizing to HCVcc of genotypes 2a (J6) and 1b (Con1) (Figure 34). The recombinant protein boost resulted in a more robust response than the measles virus boost, indicating that recombinant protein may constitute a more useful vaccine strategy than previously assumed due to its hypervariable nature.

An immunization strategy where recombinant E2 is both the prime and the boost has yet to be studied.

Since 2011, we have also been working on a collaborative project with Dr. Michael Houghton at the University of Alberta. Dr. Houghton has advocated the potential of HCV envelope glycoproteins in vaccine design since 2000 (53) (56). The Houghton group has solicited our help in a vaccine strategy which involves the immunization of goats with recombinant E1 and E2. Preliminary results have been promising, with a cross-reactive antibody response and cross-neutralization response to genotypes 2a and 3a reported (unpublished data). By immunizing with various genotypes and truncated constructs, we hope to gain a better understanding of immunogenic versus conserved/neutralizing domains and determine if one particular genotype is more ideal for eliciting cross-neutralizing responses. These studies should lay important groundwork for testing HCV envelope protein as a legitimate vaccine candidate.
Figure 34. eE2 Elicits a Cross-Neutralizing Response.

Green dashed lines represent the anti-CD81 positive control while blue dashed lines represent anti-H77E2 (homologous protein) neutralization. As indicated by the red bars, an eE2 vaccine boost elicited antibodies that neutralized H77 HCvcc, J6 HCvcc, and Con1 HCvcc. MV-E1/Ft-E2/Ft refers to a construct in which the E1 and E2 glycoproteins are fused to the transmembrane domains and cytoplasmic tails of measles virus F protein.
We have established many other collaborative efforts as a result of our ability to express HCV envelope proteins as well as other interesting biological targets. These collaborations will help to keep the Marcotrigiano laboratory at the forefront of major achievements in HCV research. Furthermore, to the best of my knowledge, we continue to be a leading group in attempting crystallography studies with HCV envelope protein E2 and the only group with the capacity to express multiple co-crystallization factors including antibodies, E1, CD81-LEL, and SR-BI. Our work sets the stage for a better understanding of HCV infection, and how to treat this disease and prepare anti-HCV vaccines.
REFERENCES


25. **Demeler, B.** 2009, posting date. A comprehensive data analysis software package for analytical ultracentrifugation experiments. [Online.]


40. **Flint, M., and J. A. McKeating.** 1999. The C-terminal region of the hepatitis C virus E1 glycoprotein confers localization within the endoplasmic reticulum. The Journal of general virology **80 (Pt 8)**:1943-1947.


43. **Forns, X., R. Thimme, S. Govindarajan, S. U. Emerson, R. H. Purcell, F. V. Chisari, and J. Bukh.** 2000. Hepatitis C virus lacking the hypervariable region 1 of the second envelope protein is infectious and causes acute resolving or persistent infection in chimpanzees. Proc Natl Acad Sci U S A **97**:13318-13323.


ABBREVIATIONS

A/A: antibiotic/antimycotic

BSA: bovine serum albumin

βME: β-mercaptoethanol

CD: circular dichroism

CD81-LEL: CD81 large extracellular loop

cDNA: complementary DNA

CHO: Chinese hamster ovary (cell line)

CIP: calf intestinal phosphatase

CMV: cytomegalovirus

ddH$_2$O: “double distilled” filtered water

DMEM: Dulbecco’s Modified Eagle Medium

DNA: deoxyribonucleic acid

E1: hepatitis C virus envelope protein 1

E2: hepatitis C virus envelope protein 2

eE2: hepatitis C virus envelope protein 2 soluble ectodomain

ELISA: enzyme-linked immunosorbent assay

ER: endoplasmic reticulum

Fab: light chain antibody fragment

FACS: fluorescence activated cell sorting
FBS: Fetal bovine serum

Fc: antibody heavy chain fragment

FDA: Food and Drug Administration

GFP: green fluorescent protein

GlcNac: N-acetyl glucosamine

gp41: HIV glycoprotein, 41 kDa

gp120: HIV glycoprotein, 120 kDa

gp350: Epstein-Barr virus glycoprotein, 350 kDa

GST: glutathione-s-transferase

HEK: Human embryonic kidney (cell line)

HeLa: cervical cancer, Henrietta Lacks (cell line)

HCV: hepatitis C virus

HCVcc: cell culture derived hepatitis C virus

HCVpp: hepatitis C virus pseudoparticles (replication incompetent)

HDX: hydrogen/deuterium exchange

HIV: human immunodeficiency virus

HPLC: high performance liquid chromatography

HRP: horseradish peroxidase

IAM: iodoacetamide

IgG: immunoglobulin, isotype G
IRES: internal ribosome entry site

LB: luria broth

LC-MS: liquid chromatography-mass spectrometry

MOI: multiplicity of infection

NEM: N-ethylmaleamide

NIH: National Institutes of Health

OD: optical density

PBS: phosphate-buffered saline

PBS-T: phosphate-buffered saline + 0.05 % Tween-20

PCR: polymerase chain reaction

PEG: polyethylene glycol

RNA: ribonucleic acid

SDS-PAGE: sodium dodecyl sulfate- polyacrylamide gel electrophoresis

SEC: size exclusion chromatography

SIV: simian immunodeficiency virus

SR-BI: scavenger receptor class B type I

TCID$_{50}$: tissue culture infectious dose, 50 %

TFA: trifluoroacetic acid

VSV: vesicular stomatitis virus