UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Interactome screening of Human Cytomegalovirus and the Complement system & Investigating the pre-clinical immunogenicity of a novel SARS-CoV-2 Envelope (E) protein vaccine

Permalink

https://escholarship.org/uc/item/0q3809q7

Author

Lujan, Eduardo Liberato

Publication Date

2023

Peer reviewed|Thesis/dissertation

Interactome screening of Human Cytomegalovirus and the Complement system &
Investigating the pre-clinical immunogenicity of a novel SARS-CoV-2
Envelope (E) protein vaccine

By

Eduardo Liberato Lujan

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Comparative Biochemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Fenyong Liu, Chair
Professor Sarah Stanley
Professor Eva Harris

Summer 2023

ABSTRACT

Interactome screening of Human Cytomegalovirus and the Complement system &
Investigating the pre-clinical immunogenicity of a novel SARS-CoV-2
Envelope (E) protein vaccine

by

Eduardo Liberato Lujan

Doctor of Philosophy in Comparative Biochemistry

University of California, Berkeley

Professor Fenyong Liu, Chair

Human Cytomegalovirus (HCMV) is a ubiquitous virus that is a leading cause of both congenital infections in neonates and opportunistic infections in immunodeficient persons such as HIV/AIDS patients and solid-organ transplant recipients. Few human viruses compare to the complexity, broad cell tropism, and life-long persistence observed in HCMV infections. HCMV has one of the largest genomes of any known human viral pathogen with an approximate size of 230kbp of DNA that encodes roughly 175 canonical genes. The genes encoded by HCMV have been characterized by forward genetics using a single-gene deletion approach to understand the essential or dispensable nature of each HCMV gene necessary for viral replication *in vitro,* however the precise biological functions of most HCMV genes remain incompletely understood. HCMV encodes an unparalleled number of proteins that subvert both innate and adaptive immunity, yet currently the role of the human complement system in HCMV infection and immunity remain poorly understood even though this important branch of innate immunity has been well characterized in other related human herpesviruses.

Previous studies have reported that HCMV virions incubated in serum (a source of complement) are not neutralized but are nevertheless coated with activated complement proteins that normally proceed the direct lysis of microorganisms. This observation suggests that HCMV virions not only activate one or more pathways of complement, but are able to inhibit complement-mediated neutralization perhaps by a viral encoded complement regulating protein. In this dissertation, we screened 6 human complement proteins (CD55, Factor H, Mannose Binding Lectin 2, Mannan-Associated Serine Protease-1, Properdin, and C1q Binding Protein) against an HCMV genomic library in yeast containing 167 viral open reading frames. Out of 1,002 possible HCMV-Complement interactions tested, we identified 121 (8.2%) positive-protein interactions by yeast-two-hybrid among all six complement proteins tested and validated a subset of these novel HCMV-Complement interactions in human cells by co-immunoprecipitation. To date, our study is the largest, most comprehensive, and systematic investigation of interactions between HCMV the complement system, and provides a framework to further investigate HCMV-Complement interactions that may underlie undiscovered mechanisms of innate immune evasion by HCMV

and potentially inform the development of novel drugs and vaccines for the treatment and prevention of HCMV infections.

During our studies, a novel coronavirus known as Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) emerged and spread rapidly in humans, which led to a pandemic and significant loss of life globally. In response to the pandemic, SARS-CoV-2 vaccines were developed but their efficacy has been undermined by the increasing emergence of SARS-CoV-2 isolates with mutations in the Spike (S) protein which is the primary vaccine antigen in all licensed SARS-CoV-2 vaccines. These observations underscore the significant need to identify additional antigens for possible inclusion in the next-generation of SARS-CoV-2 vaccines. To address this significant unmet medical need, we investigated the vaccine potential of a highly conserved 30 amino acid transmembrane domain epitope of the SARS-CoV-2 Envelope (E) protein in mice. To increase immunogenicity, we conjugated the E-protein transmembrane domain to the immunogenic carrier protein Keyhole Limpet Hemocyanin (KLH) to generate a vaccine antigen herein referred to as KLH-E. Mice were immunized with one or two doses of KLH-E vaccine or aluminum control vaccine intramuscularly in thirty-day intervals. The KLH-E vaccine elicited serum IgG antibodies and antigen-specific T-cells to each subunit of the vaccine (KLH and E), and to our surprise induced anti-E IgG antibodies and T-cells that recognize the native E protein from the genetically divergent but related human coronavirus 229E (HCoV-229E). Furthermore, pooled sera from mice had neutralizing activity against SARS-CoV-2 pseudovirions expressing S proteins from emerging variants of concern (Beta, Delta, Omicron, XBB), and had neutralizing activity against HCoV-229E which shares only a modest 32% E protein sequence homology to the vaccine antigen. Taken together, these results suggest that the SARS-CoV-2 E protein transmembrane domain contains important epitopes that are conserved across different human coronaviruses and demonstrates that immunization with the E protein transmembrane domain from SARS-CoV-2 can induce cross-reactive immune responses that may confer protection against emerging SARS-CoV-2 variants and other genetically diverse human coronaviruses.

DEDICATION

This dissertation is dedicated to my family: Mom, Dad, Gustavo, Estrella, and Javier.

For their endless and unwavering love and support. Te amo.

TABLE OF CONTENTS

Chapter 3   **Page Number**

LIST OF FIGURES, TABLES, AND SUPPLEMENTARY MATERIALS

Chapter 3:   **Page Number**

Title

Supplemental Figures

Supplemental Tables

ABBREVIATIONS

Chapters 1, 2

AC...Assembly Complex
AD………………………………………….…Activating domain, Gal4 transcription factor
AD169..Human Cytomegalovirus, laboratory strain AD169
AIDS..Acquired Immunodeficiency Syndrome
AP………………………………………….…Alternative Pathway of Complement
BAC………………………………………..…Bacterial Artificial Chromosome
BD………………………………….....DNA Binding domain, Gal4 transcription factor
C1………………………………………….…Complement Component 1 Complex
C1q…………………………………………..Complement Component 1q Complex
C1qBP………………………………..….Complement Component 1q Binding Protein
C1s………………………………………......Complement Component 1s
C1r………………………………….....Complement Component 1r
C2……………………………………….....Complement Component 2
C2a……………………………………..Complement Component 2, a subunit
C2b……………………………………....…….Complement Component 2, b subunit
C3………………………………………….....Complement Component 3
C3a……………………………………..Complement Component 3, a subunit
C3b…………………………………………Complement Component 3, b subunit
C3bBb...C3 Convertase (Alternative pathway)
C3bBbC3b...C5 Convertase (Alternative pathway)
C4………………………………………….....Complement Component 4
C4a………………………………………Complement Component 4, a subunit
C4b………………………………………Complement Component 4, b subunit
C4b2a……………………………….......C3 Convertase (Classical and Lectin pathway)
C4b2a3b………………………………..…C5 Convertase (Classical and Lectin pathway)
C5………………………………………….....Complement Component 5
C5a………………………………………….Complement Component 5, a subunit
C5b……………………………………….Complement Component 5, b subunit
C5b-C9/C5bC6C7C8C9/MAC…………………..………....Membrane Attack Complex
C6……………………………………………Complement Component 6
C7……………………………………….....Complement Component 7
C8……………………………………….....Complement Component 8
C9……………………………………….....Complement Component 9
CCP...Complement Control Protein, domain
CD55/DAF………………………………………..CD55/Decay-Accelerating Factor
cDNA………………………………………Complementary Deoxyribose Nucleic Acid
CFH/FH……………………………………….Complement Factor H
CFP/P………………………………………Complement Factor P/ Properdin
cGAS...cytosolic GMP-AMP synthase
cGAS-STING...............................cytosolic GMP-AMP synthase Stimulator of Interferon Genes
CID………………………………………......Cytomegalovirus Inclusion Disease
Co-IP………………………………………..Co-Immunoprecipitation

HCoV-229E……………………………………………………Human Coronavirus 229E
HCoV-HKU1…………………………………………………Human Coronavirus HKU1
HCoV-NL63………………………………………………....Human Coronavirus NL63
HCoV-OC43………………………………………………...Human Coronavirus OC43
HCV...Hepatitis C Virus
HeLa...HeLa, immortalized human cell line
HIV...Human Immunodeficiency Virus
IAV...Influenza A Virus
IgA…………………………………………………………..Immunoglobulin A
IgG…………………………………………………………..Immunoglobulin G
IgG1……………………………………………Immunoglobulin G, subclass 1
IgG3……………………………………………Immunoglobulin G, subclass 3
IgM……………………………………………………Immunoglobulin M
IL-1β...Interleukin-1 beta, pro-inflammatory cytokine
IL-2...Interleukin-2, T-cell activation cytokine
IL-6.. Interleukin-6, pro-inflammatory cytokine
IL-17.. Interleukin-17, inflammatory cytokine
KLH……………………………………………………Keyhole Limpet Hemocyanin
KLH-E………Keyhole Limpet Hemocyanin-Envelope protein transmembrane domain conjugate
M…………………………………………………………………..Membrane protein
mAb...Monoclonal antibody
MERS-CoV………………………………………..Middle East Respiratory Syndrome Coronavirus
MHC..Major Histocompatibility Complex
MPERS...Membrane Proximal External Regions, glycoprotein gp41 HIV
MRC5...Medical Research Council cell strain 5, human cell line
mRNA……………………………………………………..messenger Ribonucleic Acid
N…………………………………………………………………...Nucleocapsid protein
NLRP3..NLR family pyrin domain containing 3
NMR…………………………………………………....Nuclear Magnetic Resonance, Spectroscopy
Nsp..Non-structural protein
ORF…………………………………………………………………..Open Reading Frame
PAMP...Pathogen Associated Molecular Pattern
PBS..Phosphate-Buffered Saline
PD1..Post-dose 1
PD2..Post-dose 2
PV...Poliovirus
RBD……………………………………………………Receptor Binding Domain, S-protein
RNA……………………………………………………………..…Ribonucleic Acid
RV...Rotavirus
S…………………………………………………………………...…Spike protein
SARS-CoV-1…………………………..Severe Acute Respiratory Syndrome Coronavirus Virus 1
SARS-CoV-2…………………….…Severe Acute Respiratory Syndrome Coronavirus Virus 2
STAT1..Signal Transducer and Activator of Transcription
TGF-β...Transforming Growth Factor, beta
TNF-α ....Tumor Necrosis Factor alpha, pro-inflammatory cytokine
TRIM21....................................Tripartite motif-containing protein 21, E3 ubiquitin-protein ligase

ACKNOWLEDGEMENTS:

I would first like to thank my mentor, Professor Fenyong Liu for providing me the opportunity to pursue a doctoral degree at UC Berkeley in his laboratory. Very rarely are doctoral students allowed to embark on risky research projects, but I am thankful for the scientific freedom I was provided to investigate the protein interactome between HCMV and the Complement system. Additionally, it was a great privilege to work on and contribute intellectually in a meaningful way to the coronavirus E protein vaccine and other related projects. These projects were exciting, novel, and ones which I was truly passionate about. During my time as a doctoral student in your lab I learned more than I thought possible, but the three most valuable lessons you've taught me are: (1) Never turn down an opportunity to learn, (2) Be open minded to unorthodox hypotheses that defy conventional scientific dogma, and (3) Embrace challenging experiences, because these experiences are what fortify our character.

A major milestone in a doctoral degree is the oral qualifying exam which marks the transition from doctoral student to doctoral candidate and is an experience often described as a "trial by fire". I am thankful for Drs. Lee Riley, Sarah Stanley, Gary Firestone, and Xiaohua Gong for serving on my qualifying exam committee and for their invaluable comments on my proposals, intellectually rigorous questions, and for their support.

During my doctoral degree I had the pleasure of being both a student and co-instructor in courses taught by Drs. Fenyong Liu, Eva Harris, and Sarah Stanley. I am thankful for the continued opportunities I had to learn from and work closely with such talented and inspiring professors. I am also immensely grateful to have such esteemed faculty serving on my dissertation committee. Additionally, every Spring semester I had the honor of being a co-instructor with Dr. Robert Beatty, whose profound knowledge of immunology and passion for teaching are truly remarkable.

At times, pursuing a doctoral degree felt like walking a tightrope as I tried to balance my education, lab work, teaching, and a social/personal life among many other commitments. Whenever I felt overwhelmed or paralyzed in my doctoral studies, I could always depend on late-night phone calls or visits with family to provide me with the stability and balance I needed to continue forward. Without question, my completion of a doctoral education would not have been possible without the endless and unwavering love and support from my family. During moments of weakness, I drew guidance and strength from my mother, whose love and endless sacrifices provided me with the opportunities necessary to follow my dreams. In the words of my mother "*Muchas cosas en la vida son temporales, pero la familia es para siempre*" (Many things in life are temporary, but family is forever). When I second guessed myself, my father's infinite wisdom, love, and penchant for inspirational quotes gave me the self-confidence I needed to overcome many challenges. Thank you for believing in me and for encouraging me to always strive to be my most authentic self. Thank you for always loving me separate from any conditions, accomplishments, or setbacks. I may be the first in my family to receive a doctoral degree but I hope that my success serves as an example for others in my family, or perhaps even those with similar backgrounds, to pursue higher education and follow their dreams.

Additionally, I'm eternally grateful to my academic family which includes many professors and mentors I've had along the way who all believed in me and each contributed in a unique way to the growth and development of the person I've become.

A strong support system wouldn't be complete without friends and colleagues, many of which include fellow doctoral students, lab mates, post-docs, and senior scientists in my lab and in others. I would be remiss if I did not also acknowledge my cat Rilo, who has been with me since the beginning of my undergraduate education and whose company always made my late-night computer work a little less lonely.

I am a staunch believer that people enter and exit your life for a reason, and believe it is incumbent on us to learn what we can from each of our interactions with others. Whether our interactions were brief, over the course of years, or continue to this day, I am thankful to have crossed paths with some amazing people who taught me many important life lessons. One important person I've met is Matthew Gray, who came into my life as one chapter closed and a new and exciting one began. Our time together has been short, but it's been the best time of my life. I don't know what the future holds, all I know is that I'd like to see you in it. As the Spanish saying goes "*La vida estaria vacia sin ti"*.

Finally, I'd like to acknowledge my doctoral student cohort and salute our resolve to overcome tremendous adversity. Completing a doctoral degree is a difficult feat, but never in my wildest dreams did I think I'd experience: loss, heartache, social and political unrest, wildfires, extreme heatwaves, or a global pandemic during my doctoral studies. Notwithstanding these challenges, my success and that of my peers only further underscore that we can truly achieve anything we set our minds to.

Chapter 1

Background and Introduction to Human Cytomegalovirus
and the Complement System

INTRODUCTION: HUMAN HERPESVIRUSES

The Herpesviridae Family
The Herpesviridae family is composed of large, enveloped viruses that are unique from all other viruses in various aspects including: morphology, lifecycle, cell tropism, and species specificity. Morphologically, all herpesviruses have a large, linear, double-stranded DNA genome ranging from 125-290 kbp contained within an icosahedral capsid with T=16 symmetry [1]. The capsid of herpesviruses is surrounded by tegument, a proteinaceous matrix that lies below the viral lipid envelope, the latter of which is studded with various transmembrane and membrane-associated proteins [2]. Advances in sequencing technology have led to the discovery of a plethora of novel herpesviruses over the past decade, and as of 2023 there are 113 published complete genomic sequences of known herpesviruses which have provided great insights into the evolutionary history and host specificities of the members in the Herpesviridae family. Herpesviruses infect a broad range of vertebrate animals, as well as animals of lower taxonomic groups including mollusks. Despite the identification of herpesviruses in various diverse species, one unique aspect of herpesviruses is their relatively restricted host range. Since most herpesviruses have coevolved with their host species for millions of years they have a reduced capacity to infect hosts across different species compared to other viruses [2].

Human herpesviruses
Eight herpesviruses are known to infect humans, and can be further classified based on their biological properties such as growth characteristics and cell/tissue tropism into three subfamilies: Alpha (α), Beta (β), and Gamma (ɣ) herpesviruses as listed in **Table 1-1**. α-herpesviruses include: Herpes simplex virus 1 (HSV-1/HHV-1), Herpes simplex virus 2 (HSV-2/HHV-2), and Varicella-Zoster Virus (VZV/HHV-3). HSV-1 and HSV-2 primarily cause oral and genital herpes respectively, two diseases both characterized by the appearance of fluid filled blisters on mucus membranes of the skin surrounding the mouth and genitals that occur when the viruses undergo active lytic replication [3]. During periods of latency, both HSV-1 and HSV-2 are found in the sensory ganglia and nerve cell bodies and cause no clinical manifestations [4]. Reactivation from latency can occur, and can be triggered by several environmental (UV light) [5], and biological factors (hormonal changes, infections, immunosuppression) [6, 7]. During reactivation, latent viruses become active and travel in an anterograde manner via the axon of nerves to the epithelial cells of the skin where active lytic replication once again occurs, resulting in the formation of new viral particles and herpetic sores which promote transmission of HSV-1 and HSV-2 [4]. Although both HSV-1 and HSV-2 are primarily transmitted horizontally by skin-to-skin contact, they can also be transmitted vertically before or during childbirth and cause serious complications. As with the other α-herpesviruses, VZV also has a strong tropism for epithelial cells during lytic replication and remains latent in nerve cells. Primary infection with VZV causes Varicella (Chickenpox) in children and/or Shingles-Zoster because of viral reactivation, typically observed later in adulthood or in immunosuppressed persons [8]. Features shared among all α-herpesviruses is their relatively rapid cytocidal growth cycle in epithelial cells and establishment of latent infections primarily in sensory ganglia.

Table 1-1. Human Herpesviruses				
HHV	**Name**	**Subfamily**	**Tropism**	**Associated Disease(s)**
1	Herpes Simplex Virus-1 (HSV-1)	α	Mucoepithelial and neurons	Cold sores
2	Herpes Simplex Virus-2 (HSV-2)	α	Mucoepithelial and neurons	Genital herpes
3	Varicella zoster virus (VZV)	α	Mucoepithelial and neurons	Chickenpox/Shingles
4	Epstein-Barr Virus (EBV)	ɣ	B and epithelial cells	Mononucleosis
5	Cytomegalovirus (CMV)	β	Monocytes and Lymphocytes	Cytomegalovirus
6	Roseolavirus or Herpes lymphotropic virus	β	T-cells	Roseola
7	Roseolovirus	β	T-cells	Roseola
8	Kaposi's sarcoma-associated herpesvirus (KSHV)	ɣ	Lymphocytes	Kasposi Sarcoma

β -herpesviruses that infect humans include: Human Cytomegalovirus (HCMV/HHV-5), HHV-6, and HHV-7, with the latter two often causing Roseola infantum disease (sixth disease) in young children. HCMV causes cytomegalovirus disease, most commonly in the developing fetus as congenital HCMV or in immunocompromised individuals such as HIV/AIDS patients or solid-organ transplant recipients. In contrast to α-herpesviruses, β-herpesviruses have a long and relatively slow reproductive cycle in culture, and establish latency in monocytes (HCMV) or T-cells (HHV-6 and HHV-7).

ɣ-herpesviruses include: Epstein-Barr Virus (EBV/HHV-4), which can cause mononucleosis, oropharyngeal carcinoma, and Burkitt's lymphoma. Infection with Kaposi Sarcoma associated Herpesvirus (KSHV/HHV-8) can lead to the development of Kaposi Sarcoma, and is often observed as an opportunistic infection in HIV patients with such severe immunodeficiencies that KSHV infection is considered to be an AIDS defining illness. Both EBV and KSHV establish latency in B-cells and are unique among herpesviruses in that they have the capacity to directly promote oncogenesis.

INTRODUCTION: HUMAN CYTOMEGALOVIRUS (HCMV)

Discovery. Microscopically, a hallmark of cytomegalovirus infected cells is the visual appearance of "owl's eye structures" that describe the large intranuclear inclusion bodies (cytomegaly) characteristic of HCMV infection. These inclusion bodies were first noticed in 1881 by German scientists Hugo Ribbert in samples taken from stillborn fetal tissue observed by light microscopy [9] and were not fully understood until a later report by Jesionek and Kiolemenoglou described the appearance of "protozoan-like" cells in the lungs, kidney, and liver of an 8-month luetic fetus. Later, in 1956-1957 Thomas Weller, Margaret Smith, and Wallace Rowe independently isolated and grew HCMV from humans [10-12], a feat that was only made possible due to the availability of cultured human embryonic cells in which to propagate isolated virus [13]. Since its initial discovery and propagation in laboratory culture, significant advances have been made in understanding HCMV biology and the effects of viral infection on the human body.

Virion structure. The HCMV virion is approximately 230nm in diameter and is composed of a nucleocapsid core that contains the large ~230kbp double stranded DNA genome. The icosahedral nucleocapsid of HCMV is composed of five integral proteins (UL46, UL48-UL49, UL80, UL85, UL104) and is assembled in the nucleus of infected cells after viral DNA replication. Surrounding the nucleocapsid is a disordered proteinaceous matrix called the tegument layer which is composed of approximately 32 known virion-associated tegument proteins that are acquired in both the nucleus and the cytoplasm of infected cells [14]. HCMV tegument proteins contribute to the structural integrity of the virion, but also mediate various cellular processes during viral replication and have been reported to disrupt innate and adaptive immunity [15]. Surrounding the tegument layer of HCMV virions is the viral envelope made up of a lipid membrane bilayer derived from the infected host cell that serves as the substrate for the expression of various virally encoded glycoproteins as well as some host cell proteins [16]. The surface exposed glycoproteins of HCMV engage with multiple host cell receptors to mediate viral entry and fusion, and the precise combinations of viral glycoproteins and host cell receptors varies by cell type, given the broad cell tropism of HCMV *in vivo* [17]. Details of the composition and order of specific HCMV virion components are illustrated in **Figure 1-1.**
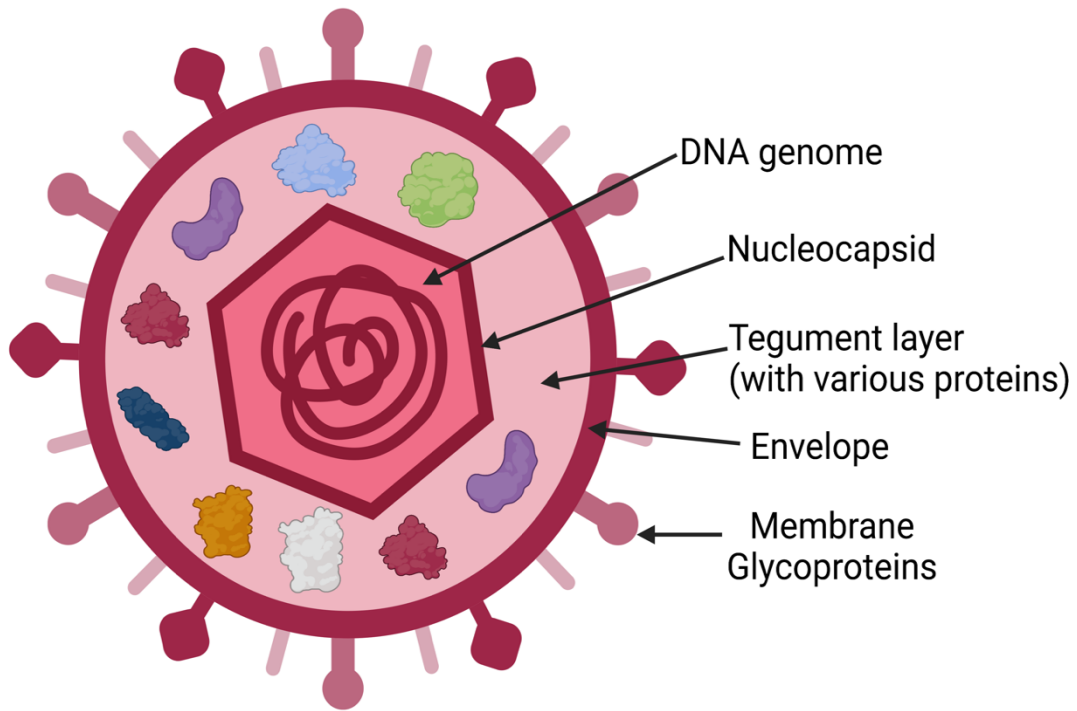
Figure 1-1. Structure of the HCMV virion.
Mature virions are coated by an envelope from which viral glycoproteins protrude and mediate HCMV entry and fusion with host cells. Below the envelope is the tegument layer, a protein rich and disordered matrix that contributes to the structural integrity of HCMV virions but also mediates various important cellular processes during viral infection. At the core of the HCMV virion is the icosahedral nucleocapsid which contains the double stranded large ~230kbp DNA genome of HCMV. Original figure designed with the assistance of Biorender software (www.Biorender.com)

Genomic architecture. The first complete HCMV genome ever sequenced belonged to strain AD169, a laboratory strain whose genome was cloned into a plasmid library to overcome technical challenges in sequencing during the early era of DNA sequencing [18]. As technology improved over the next decade, additional HCMV genomes were sequenced from bacterial artificial chromosomes [19-21], virion DNA [22], and overlapping PCR amplicons [23].

HCMV has one of the largest genomes of any known viral human pathogen with an approximate size of 230 kbp of DNA. Architecturally, the linear double stranded DNA genome consists of a unique long (UL) region and a unique short (US) region each flanked by terminal and internal repeats as illustrated in **Figure 1-2**. The terminal and internal repeats are reversely complimentary to one another. The paired sequences flanking the UL region are known as " b " or the terminal repeat long (TRL) region, and " b' " internal repeat long (IRL) region from the left to right direction respectively. The DNA sequences flanking the US region include the "c" terminal repeat short (TRS) and the " c' " region which consists of the short arm internal repeat sequence (IRS) from right to left respectively. Collectively, the configuration of the HCMV genome results in an overall sequence order of TRL-UL-IRL-IRS-US-TRS. Additionally, there are short redundant regions denoted by "a" and " a' " that occur on the outermost flanking ends of the linear DNA genome or in the short/long arm junction respectively [24, 25].
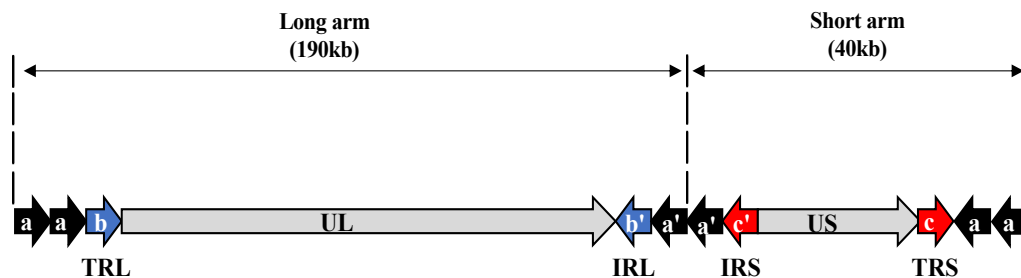
Figure 1-2. Structure of the HCMV Genome
The HCMV genome is compromised of long and short arms. Unique long (UL) region is flanked by copies of the **b** sequence: terminal repeat long (TRL) and internal repeat long (IRL) from left to right. The unique small (US) region is flanked by copies of the **c** sequence: internal repeat short (IRS) and terminal repeat short (TRS) from left to right. One or several copies of the **a** sequence are found at both ends of the long/short arm junction listed in the figure as sequence **a** (the right and left most regions of the linear sequence), or the **a'** sequence which is located at the long/short arm junction.

Gene products and kinetics. The HCMV genome encodes approximately 175 canonical genes, many of which have been characterized by forward genetics using a single-gene deletion approach to understand the essential or dispensable nature of each HCMV gene necessary for viral replication *in vitro* [19]. Subsequent studies have suggested an unprecedented level of complexity in HCMV gene expression by profiling multiple potential start sites of known open reading frames (ORF's) suggesting that HCMV may have a potential coding capacity of ~600 ORF's that also includes long non-coding RNA's and micro-RNA's with suspected regulatory functions [26, 27]. Irrespective of how HCMV ORF's are defined or delineated, HCMV has an exquisitely evolved and highly complex gene expression and regulatory network. During productive infection, HCMV gene expression occurs in a highly coordinated temporal cascade. Historically, metabolic inhibitors such as cycloheximide (a protein synthesis inhibitor), and phosphonoformate (a viral DNA replication inhibitor) were used to classify HCMV viral genes into: Immediate-Early (IE), Early (E), and Late (L) temporal expression profiles. As their name implies, IE genes of HCMV are the first to be expressed during infection and are involved in a wide range of processes including: inhibiting innate intrinsic cell immunity, regulating viral and host cell transcription, and broadly priming and optimizing the cell machinery for viral replication. HCMV-encoded E genes mostly encode viral transcription factors and polymerases involved in viral DNA replication, transcription, and protein synthesis [28]. In contrast, L genes typically encode structural proteins that are synthesized after viral DNA replication [29].

Wild-type and Laboratory strains. The redundant and complementary nature of the regions and sequences that compose the HCMV genome contribute to genetic plasticity, and as such the HCMV genome is subject to a high degree of genetic variation [30], recombination [31, 32], multiple-strain infection (superinfection) [33-35], and is especially susceptible to gene loss.

In laboratory culture, HCMV strains accumulate deletions and point mutations [36], and lose many genes after serial passage in fibroblasts. Despite the broad cell tropism of HCMV *in vivo,* not all cultured human cells are permissible to HCMV infection *in vitro*. Fibroblasts are the most commonly used cell type for propagation of HCMV because HCMV strains cultured in fibroblasts produce higher viral titers and are highly lytic, which releases virions into culture supernatant and facilitates purification. One drawback however, is that HCMV strains cultured in fibroblasts have a reduced capacity to infect non-fibroblast cell types, in particular epithelial and endothelial cells [37]. Thus, a significant challenge in studying HCMV stems from the inability to propagate clinical isolates without having them lose virulence or genetic diversity. Mutations in at least 26 genes are common during the process of adapting clinical isolates to laboratory culture in fibroblasts, and this contributes to the high genomic diversity between different HCMV strains.

For example, in the highly passaged laboratory HCMV strains AD169 and Towne, a large fragment of the UL/b' region is replaced by an inverted repeat from the left terminus of the HCMV genome. The Towne strain also has inactivating mutations in RL13 and UL130, RL13 being an immunoglobulin G (IgG) binding protein [38] and UL130 being an essential gene necessary for HCMV infection of human endothelial cells [39]. AD169 also harbors frameshift mutations in genes RL5A [40], RL13 [41], and UL131A [42] which renders these genes non-functional. Intense study of the commonly mutated genes in HCMV during adaptation to

laboratory culture in fibroblasts suggests that the reasons for the loss of genes may be due to the absence of immunological pressure *in vitro* and/or the fact that some genes lost may only be required for infection of non-fibroblast cells and are therefore dispensable.

Lifecycle of HCMV. As illustrated in **Figure 1-3,** HCMV enters human cells either through direct fusion or via the endocytic pathway. HCMV attaches and penetrates the cell in a process mediated by viral glycoproteins: gN(UL73), gH(UL75), gB(UL55), gM(UL100), gL(UL115) and specific cell receptors (the precise receptors varies by cell type). Following attachment, the viral envelope fuses with the cell membrane and releases nucleocapsids into the cytoplasm which are translocated into the nucleus by the microtubule network that docks and releases viral DNA at the nuclear pore complex. Upon release of viral DNA into the nucleus, both viral RNA transcription and DNA replication occur in a tightly regulated manner, starting with the expression of Immediate Early (IE) viral genes which optimize the cell and associated cellular machinery to produce HCMV virions. IE viral genes regulate the expression of both viral and host genes that are involved in diverse processes including but not limited to: viral replication, immunomodulation, and inhibition of apoptosis.

Shortly after, delayed early (DE) genes are expressed, and encode components of the viral DNA synthesis machinery. During this time, linear viral DNA circularizes and is copied in a rolling circle mechanism with new strands packaged into capsids that are formed in the nucleus. Thus, the major events in the nucleus include: viral IE gene expression, DNA replication, capsid assembly and loading of viral DNA. Interestingly, although many tegument proteins are expressed and localized to the nucleus during the early stages of infection, there is little evidence that tegumentation of HCMV capsids occurs in the nucleus. To exit the nucleus, nascent HCMV capsids follow a highly conserved process that is shared among other herpesviruses and involves translocation from the nucleus to the cytoplasm. In HCMV, the UL50-UL53 core genes facilitate the Nuclear Egress Complex (NEC) by interacting with viral proteins and cellular proteins including, but not limited to: C1q Binding Protein (C1qBP/p32), Lamin B receptor (LBR), Protein Kinase C (PKC). The complex network of HCMV and host proteins in the NEC mediate the dilation of the nuclear pore to allow nascent HCMV nucleocapsids to enter the cytoplasm. Nuclear egress of the newly synthesized nucleocapsids occurs in a stepwise manner with a primary envelopment step followed by de-envelopment in the cytoplasm.

The cytoplasmic stage of the HCMV viral lifecycle is complex and requires significant rearrangement of membranes and organelles within the cell. During this time, the trans-Golgi network and Endoplasmic Reticulum-Golgi Intermediate Compartment (ERGIC) undergo distinct architectural changes that contribute to the formation of a structure known as the viral Assembly Complex (AC) which is unique feature of β-herpesvirus including HCMV [43]. During the late stages of infection, capsids concentrate at the AC and acquire tegument proteins in a process known as tegumentation. Tegument proteins are some of the least understood proteins in HCMV, with many being dispensable *in vitro,* and their precise biological functions remaining elusive. Broadly speaking among herpesviruses, the tegument layer contributes to viral replication by delivering preformed proteins required for initial stages of infection, contributing to structural stability of nucleocapsids, and mediating the interactions necessary for viral envelopment. In HCMV, there are approximately 32 tegument proteins, many of which are phosphorylated and can be found in multiple subcellular compartments during infection.

Following tegumentation, HCMV viral particles undergo secondary envelopment and further maturation in the viral assembly compartment. The HCMV viral envelope is derived from a mixture of ER, ERGIC, and endosome membranes that are modified with surface exposed viral glycoproteins that function in complexes for viral entry into host cells. In the final stages of maturation after secondary envelopment and the acquisition of viral glycoproteins, the mature virion containing vesicles egress from the AC and fuse to the plasma membrane resulting in the release of HCMV viral progeny into the extracellular space.
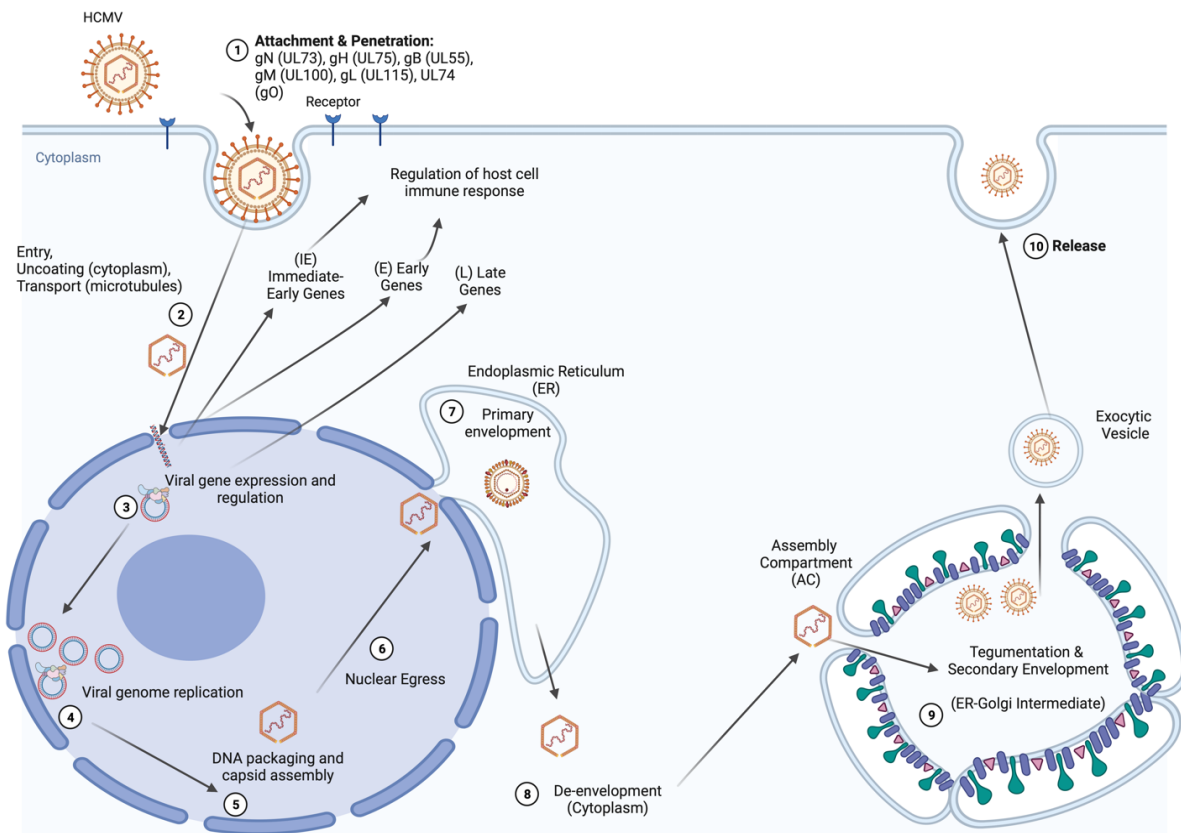
Figure 1-3. HCMV Replication cycle

The HCMV virion attaches to the cell surface via viral membrane glycoproteins in the envelope and penetrates the cell. The naked viral capsids are then translocated to the nucleus by the microtubule network in the cell, where they dock and release viral DNA at the nuclear pore complex. Upon release of DNA into the nucleus, Immediate Early (IE) genes are expressed, which regulate the expression of viral and host genes, the functions of which include but are not limited to: viral replication, immunomodulation, and inhibition of apoptosis. Following the expression of various genes, the linear viral DNA circularizes and is copied in a rolling circle mechanism, with new strands packaged into capsids that are formed in the nucleus. Nuclear egress of the newly synthesized nucleocapsids occurs in a stepwise manner with a primary envelopment step followed de-envelopment in the cytoplasm. Once in the cytoplasm the newly synthesized nucleocapsids acquire tegument proteins and bud into the Endoplasmic Reticulum-Golgi Intermediate Compartment (ERGIC) to undergo secondary envelopment and maturation. After secondary envelopment, the mature virion containing vesicles fuse with the plasma membrane resulting in the release of new viral progeny into the extracellular space. Original figure designed with the assistance of Biorender software (www.Biorender.com)

Epidemiology and Transmission. HCMV (HHV-5) is ubiquitous in human populations with a high global prevalence of infection [44]. The seroprevalence of HCMV infection ranges from 30-90% depending on age, location, gender, and socioeconomic status, with higher rates in developing countries. Transmission can occur via several routes including: saliva, sexual contact, congenitally, via breastfeeding, blood transfusion, and solid-organ transplantation.

HCMV-Associated diseases. In the immunocompetent host, primary infection with HCMV is largely asymptomatic and is characterized by a life-long latent infection [45]. However, primary infection or reactivation from latency in immunocompromised individuals can cause serious and sometimes fatal disease. The most at risk of severe HCMV disease include HIV/AIDS patients (whose infections are not controlled with antiviral medications), solid-organ transplant recipients, and the developing fetus. The clinical manifestation and broad spectrum of diseases associated with HCMV infection largely depend on which system of the body the virus infects. HCMV has a broad cell tropism *in vivo*, and this is supported by the fact that HCMV infection can cause a diverse spectrum of distinct diseases in immunocompromised persons, including but not limited to: HCMV hepatitis (which leads to fulminant liver failure), HCMV retinitis, HCMV esophagitis, HCMV colitis, and HCMV pneumonitis [46]. HCMV is also a leading viral cause of congenital birth defects including hearing loss, mental and physical retardation, and other neurological defects [47].

Antiviral agents. In 1989, Ganciclovir became the first anti-HCMV agent approved by the United States Food and Drug Administration (USFDA) for the treatment and prevention of HCMV infection and disease. Ganciclovir (1,3-dihydroxy-2-propoxymethyl guanidine) is a guanidine nucleoside analog that functions as prodrug. The inhibitory action of Ganciclovir is highly selective requiring the drug first be converted into an active triphosphate form intracellularly. The first phosphorylation of the drug is carried out by the HCMV-encoded thymidine kinase (UL97), followed by a second phosphorylation by host guanylate kinase, and a final phosphorylation by a number of host enzymes to produce the functionally active Ganciclovir triphosphate [48]. Once active, Ganciclovir triphosphate targets the HCMV DNA polymerase (UL54) where it competitively inhibits deoxyadenosine triphosphate (dATP) and replaces adenosine bases in the growing viral DNA chain. This process ultimately prevents viral DNA synthesis, as the phosphodiester bridges can no longer be synthesized due to the absence of a free 3'hydroxyl (-OH) group, which destabilizes the DNA strand. To date, Ganciclovir remains a first-line drug used to treat HCMV infections but has been undermined by two major drawbacks. First, Ganciclovir must be administered intravenously, which limits its widespread use outside of a clinical setting. Second, Ganciclovir can induce myelosuppression [49] and has been reported to have hematological toxicity, neurotoxicity, and possibly hepatotoxicity [50, 51]. As one of the most effective anti-HCMV drugs, the overuse of Ganciclovir also risks the potential to select for drug resistance, and the increasing identification of Ganciclovir-resistant HCMV isolates represents an emerging threat.

The most recent anti-HCMV drug to be licensed by the USFDA was Letermovir (3,4-dihydroquinazoline), a drug that unlike those before it targets the HCMV terminase complex [52]. HCMV genomic replication involves a rolling-circle mechanism that produces multiple genomic concatemer [53]. The HCMV core terminase complex is made up of UL51, UL56, and

UL89 [54], with all three proteins necessary to carry out the cleavage of concatemeric viral DNA into full-length genomes [55], which are then packaged into newly formed viral nucleocapsids during virion assembly [53]. Letermovir can be administered both intravenously and orally, and appears to be more well tolerated with less negative side effects compared to Ganciclovir and Valganciclovir [56]. Importantly, while Letermovir has many benefits compared to previous anti-HCMV drugs, drug resistance has been documented both *in vivo* and *in vitro* in the UL51, UL56 and UL89 viral genes targeted by the medication [57-62]. Taken together, there is an impending need for the development of novel anti-HCMV drugs to counter the emerging clinical challenge of HCMV antiviral drug resistant isolates.

Vaccine development. In the year 2000, the National Academy of Medicine listed an HCMV vaccine as the highest priority for vaccine development due to the fact that HCMV remains an important cause of global mortality and morbidity in neonates, allograft recipients, and immunocompromised persons [63]. Despite decades of research, there remains no licensed HCMV vaccine, and all vaccine candidates tested have elicited only modest protection. The development of an HCMV vaccine has been impeded by many factors including: lack of animal models, existence of many host encoded viral receptors, broad HCMV cell tropism, the absence of defined correlates of immunity, and most importantly an incomplete understanding of the full repertoire of viral immune evasion mechanisms during primary infection and the immunological factors controlling viral latency and reactivation. Taken together, the absence of any licensed HCMV vaccine represents a significant global unmet medical need.

INTRODUCTION: THE HUMAN COMPLEMENT SYSTEM

Complement. The human complement system is an important first line of defense against infection and is an evolutionarily ancient form of innate immunity with homologs identified in invertebrates [64], insects [65], and even lower organisms [66]. In humans, the complement system is composed of ~30 proteins primarily produced by the liver and secreted into the bloodstream or expressed on cell surfaces. Complement proteins typically exist in an inactive state but are rapidly activated upon infection to initiate a complex series of proteolytic and enzymatic reactions that occur in a highly coordinated stepwise manner resulting in: 1) Inflammation, 2) Pathogen membrane lysis, 3) Opsonization, and 4) Recruitment of effector immune cells to sites of infection. The complement system is composed of three distinct pathways: classical pathway (CP), lectin pathway (LP), and alternative pathway (AP). All three pathways of complement are activated by unique molecular triggers and converge at the level of complement component 3 (C3).

The Classical pathway. As depicted in **Figure 1-4**, the classical pathway of complement is triggered by IgM or other antibody-antigen complexes that are recognized by the C1 complex. The C1 complex is made up of a single molecule of C1q and a heterotetramer composed of two molecules each of C1r and C1s, forming an overall structure of C1q-C1r-C1s-C1s-C1r [67]. Within the C1 complex, C1q functions as a pattern recognition molecule that binds to the Fc domain of antibody and immune complexes. In contrast, both the C1r and C1s subcomponents of the C1 complex are serine proteases with highly restricted specificity and low enzymatic activity. Upon binding of C1q to antibody, the C1 complex is engaged resulting in the autoactivation of C1r, whose sole substrate is C1s. Activated C1r proteolytically cleaves and activates C1s, which then activates complement component 4 (C4) and complement component 2 (C2). C1s activates C4 by cleaving it into C4a and C4b. C4a functions as an anaphylatoxin and activates inflammatory immune cells, while C4b binds directly to pathogen surface and functions as an opsonin. C1s also activates C2 by cleaving it into C2a and C2b. The larger fragment C2a forms part of the complement component 3 (C3) cleaving enzyme of the classical pathway C4b2a (C3 convertase). As its name implies, the C3 convertase cleaves C3 into C3a and C3b. Like C4a, C3a acts as an anaphylatoxin and recruits inflammatory cells, while C3b binds to the C4b2a (C3 convertase) to form the C5 convertase (C4b2a3b) [68].
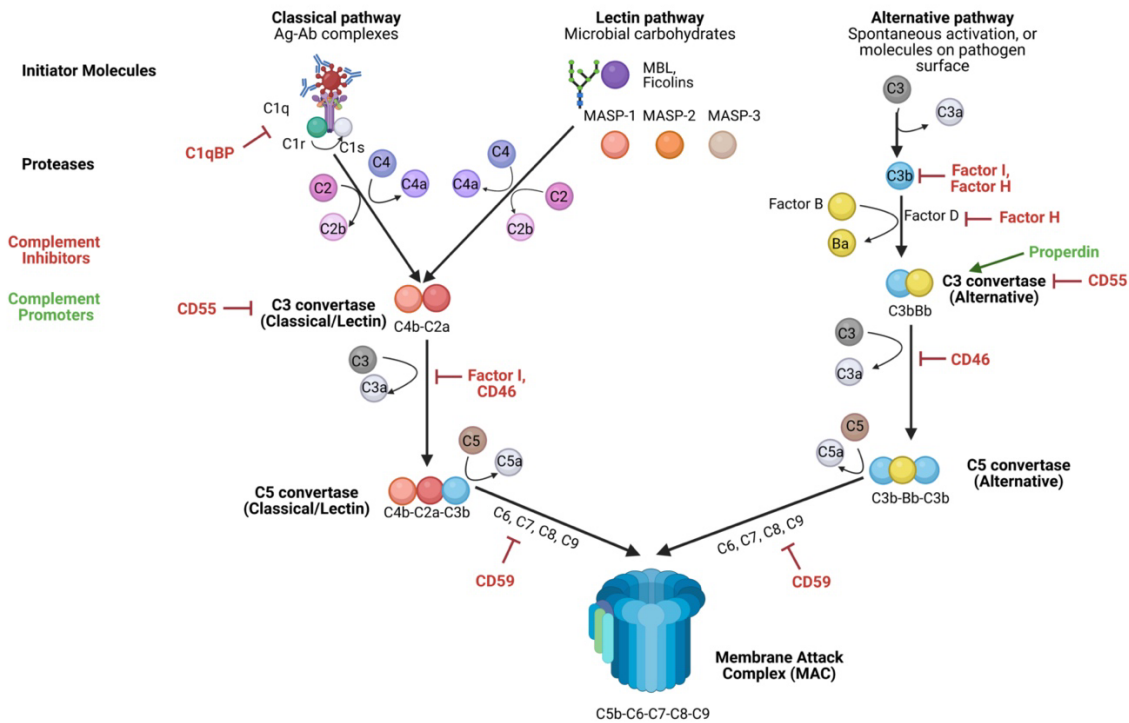
Figure 1-4. Overview of the Human Complement system. The complement system is made up of three activation pathways: the classical, the lectin, and the alternative. The classical pathway is triggered by antigen-antibody complexes recognized by C1q. C1q is found in a complex with the serine proteases C1r and C1s. The lectin pathway is triggered by carbohydrates present on the surface of microorganisms that are recognized by pattern recognition molecules (MBL2, ficolins, collectins) which are often found in complexes with serine proteases known as MASP's (MASP-1, MASP-2, MASP-3). Activation of either the classical or lectin pathway results in the cleavage of C4 and C2 which contribute to the formation of a C3 convertase (C4bC2a). Both the classical and lectin pathway share the same C3 convertase made up of C4bC2a. The alternative pathway of complement is triggered by the spontaneous hydrolysis of C3 or by PAMP's on the surface of microorganism. In the alternative pathway, activated C3 binds to Factor B (FB), which is cleaved by Factor D (FD) to form the alternative pathway C3 convertase C3bBb. Cleavage of C3 by either the classical/lectin C3 convertase (C4bC2a) or the alternative C3 convertase (C3bBb) leads to the deposition of C3b on complement activating surfaces and leads to the formation either the classical/lectin C5 convertase (C4bC2bC3b) or the alternative C5 convertase (C3bBbC3b). The C5 convertase mediates the cleavage of C5 and promotes the assembly of the membrane attack complex (MAC, C5bC7C8C9). Original figure designed with the assistance of Biorender software (www.Biorender.com).

The Lectin pathway. As illustrated in **Figure 1-4,** the lectin pathway of complement is initiated when pattern recognition molecules (Mannose binding Lectin, Collectins, Ficolins) bind to carbohydrates on the surface of microorganisms. A common feature of pattern recognition molecules in the lectin pathway is their ability to interact with and form complexes with a set of serine proteases named MASP's (MASP-1, MASP-2, MASP-3). MASP-1 and MASP-2 mediate the activation of the lectin pathway, while MASP-3 has been implicated to be involved in the activation of the alternative pathway. As with the classical pathway, initiation of the lectin pathway results in cleavage of complement component 4 (C4) and complement component 2 (C2) by MASP-1 and MASP-2, though MASP-2 is suggested to be the primary serine protease of the lectin pathway. In a process mediated by MASP's, C4 is cleaved into C4a and C4b, and C2 is cleaved into C2a and C2b, resulting in the C2a and C4b subunits coming together to produce the C3 convertase (C4bC2a) that is shared between the classical and lectin pathways. As its name implies, the C3 convertase cleaves C3 into C3a and C3b to form the C5 convertase (C4b2a3b) that is identical in both the classical and lectin pathway [68]

The Alternative pathway. In contrast to the classical and lectin pathways of complement, the alternative pathway of complement can be initiated by the spontaneous hydrolysis of C3, though it can also be initiated by various proteins, lipids, and carbohydrate structures on microbial surfaces as depicted in **Figure 1-4**. Unlike the classical and lectin pathways of complement, no pattern recognition molecules are necessary to activate the alternative pathway. Upon alternative pathway activation, Factor B (FB) forms a complex with C3b and is cleaved into Bb and Ba by the protease Factor D (FD) to produce the C3 convertase of the alternative pathway C3bBb. After formation of the alternative pathway C3 convertase, the enzymatic structure is stabilized by the addition of Properdin or complement factor P (FP/CFP), which is one of the few molecules that acts as a positive regulator of complement. Despite differences in molecular composition, the C3 convertase of the alternative pathway functions similarly to the C3 convertase of the classical/lectin pathway and cleaves C3 into C3a and C3b to produce the C5 convertase of the alternative pathway C3bBbC3b. Importantly, although the alternative pathway serves an important function on its own, it also serves as an amplification loop for the classical and lectin pathways producing significant amounts of C3b which is a critical component of subsequent steps in the complement cascade [69].

The Membrane Attack Complex (MAC). Activation of either the classical, lectin, or alternative pathways of complement results in the formation of a multimolecular complex known as the C5 convertase, which as its name implies converts complement component 5 (C5) into C5a and C5b. The classical and lectin pathways share a similar C5 convertase that is made up of C4bC2aC3b, whereas the alternative pathway C3 convertase is composed of C3bBbC3b. Irrespective of the composition of the C5 convertase, when C5 is cleaved, the C5a subunit functions as an anaphylatoxin and promotes inflammation by resident immune cells in tissues, while C5b binds to the surface of the complement activating surface and recruits one molecule each of complement component 6 (C6), complement component 7 (C7), complement component 8 (C8) and between 12-18 molecules of complement component 9 (C9) forming a molecule known as the Membrane Attack Complex (MAC). The MAC has an overall structure of C5b-C6-C7-C8-C9$^{12\text{-}18}$ [70], and the formation on the MAC produces a pore on the pathogen surface that results in direct destruction of the microorganism via lysis [71] as depicted in **Figure 1-4**.

Major proteins of the Complement system. The complement system is composed of various proteins that upon activation mediate a cascade of proteolytic reactions that result in destruction of microorganisms. Broadly speaking, complement proteins can be classified into major protein classes that include: pattern recognition molecules, activating enzymes, opsonins, inflammatory mediators, complement regulators, complement receptors, and terminal complement components that form the membrane attack complex (MAC). A summary of the proteins that make up the complement system, their functions, and their involvement in different pathways [69] can be found in **Table 1-2.**

Table 1-2. Major Proteins of the Complement System			
Class	**Name**	**Function**	**Pathway**
Pattern Recognition Molecules	C1q	Bind to Ab-Ag complexes, component of the C1 complex, initiates classical pathway activation	Classical
	MBL	Binds to foreign carbohydrates in the lectin pathway	Lectin
	Ficollins	Binds to foreign carbohydrates in the lectin pathway	Lectin
	Collectins	Binds to foreign carbohydrates in the lectin pathway	Lectin
Activating Enzymes	C1r	Serine protease that activates C1s	Classical
	C1s	Cleaves C4 and C2 in the classical pathway	Classical
	C2b	small subunit of C2 that is proteolytically cleaved by proteases in either the classical or lectin pathway	Classical, Lectin
	Factor B (FB)	Serine protease that provides catalytic activity of the alternative pathway C3 convertase and C5 convertase	Alternative
	Factor D (FD)	Serine protease that cleaves FB to produce Bb which is a component of the alternative pathway C3 convertase (C3bBb)	Alternative
	MASP-1	Serine protease that cleave C3 and C2	Lectin
	MASP-2	Serine protease that cleaves C4 and C2	Lectin
Opsonins	C4b	covalently binds to carbohydrates on the surface of the target cell. If it does not efficiently form a covalent bond with carbohydrates on the surface of the target cell then the thioester bond within the alpha chain of the C4b is irreversibly cleaved resulting in inactivation.	Classical, Lectin
	C3b	Large subunit of C3 that acts as a potent opsonin when deposited on the surface of pathogens, immune complexes, or apoptotic cells. contributes to the formation of the C3 and C5 convertase of the alternative pathway and the C5 convertase of the classical/lectin pathway.	All
Inflammatory Mediators	C3a	Small subunit of C3 that acts as a potent anaphylatoxin and is recognized by a C3a receptor on cells. Can induce vascular permeability and activate multiple inflammatory pathways.	All
	C4a	Small subunit of C4 that acts as a potent anaphylatoxin and can mediate inflammatory pathways	All
	C5a	Small subunit of C5 that acts as a potent anaphylatoxin and is recognized by a C5a receptor on cells. Can induce vascular permeability, activate multiple inflammatory pathways and trigger mast cell degranulation	All
Membrane Attack Complex (MAC)	C5b	A complex of proteins that occur in the late stage of the complement cascade which together form a pore on complement activating microbial surfaces and result in membrane lysis.	All
	C6		All
	C7		All
	C8		All
	C9		All
Complement Regulators	C1INH	Protease inhibitor that binds to and inactivates C1r and C1s proteases in the C1 complex of the classical pathway. Can block the cleavage of C4 and C2 by C1 in the classical pathway or cleavage of C4 and C2 by MASP's in the lectin pathway	Classical, Lectin
	C1qBP (gC1qBP/p32)	Binds the globular heads of C1q and inhibits C1 activation. Expressed on some cell membranes but has highest expression in the mitochondria and in the cytoplasm. May be secreted though the mechanism is unclear.	Classical
	C4bp	Inhibits the classical and lectin pathways at the level of C4 and has the ability to bind to C3b and accelerate the decay of the C3 convertase. Can also serve as a co-factor for Factor I	Classical, Lectin
	CR1 (CD35)	Acts as a receptor for C3b and C4b, mediates the cellular binding to particles and immune complexes that have activated complement	All
	MCP (CD46)	Serves as a co-factor for inactivation of C3b and C4b by Factor I and is expressed on numerous host cells to prevent self damage	All

	DAF (CD55)	CD55 interacts with C4b deposited on the surface of cells and interferes with the cleavage of C2 to C2b thus limiting classical/lectin pathway C3-convertase formation. In the context of the alternative pathway, CD55 interacts with C3b and inhibits the cleavage of Factor B into Bb by Factor D, thereby preventing the formation of the alternative pathway C3-convertase	All
	Factor H (FH)	inhibiting the assembly of the alternative pathway C3 and C5 convertases by competing with Factor B for C3b binding, facilitating the disassembly of the convertases by displacing bound Factor Bb and acting as a cofactor for Factor I mediated cleavage and inactivation of C3b	Alternative
	Factor I (FI)	Inhibits complement by cleaving and degrading C3b and C4b in the presence of co-factors such as FH, C4b-binding protein, CR1, or CD46	All
	Properdin (P)	Can bind to C3b, C3bB, and C3Bb. Promotes complement activation by stabilizing C3 convertase (C3bBb)	Alternative
	MAC-IP (CD59)	Expressed on various host cell membranes and functions to prevent the formation of the membrane attack complex (MAC) by inhibiting the polymerization of C9.	All
Complement Receptors	C1qRp	Membrane expressed receptor for C1q that participates in enhancing host cell phagocytosis	
	CR1 (CD35)	Acts as a receptor for C3b and C4b, mediates the cellular binding to particles and immune complexes that have activated complement	
	CR2 (CD21)	Binds to inactive C3b (iC3b), C3dg, or C3d. Expressed on B-cells and lowers B-cell activation threshold	
	CR3 (CD11/CD18)	Expressed on macrophages, monocytes, granulocytes and NK cells. Binds to iC3b coated particles and involved in phagocytosis. May also participate in leukocyte adhesion.	
	CR4 (CD11c/CD18)	Expressed on macrophages, monocytes, granulocytes and NK cells. Binds to iC3b, and is involved in leukocyte adhesion	

REFERENCES

1. Chiu, W. and F.J. Rixon, *High resolution structural studies of complex icosahedral viruses: a brief overview.* Virus Res, 2002. **82**(1-2): p. 9-17.
2. Davison, A.J., et al., *The order Herpesvirales.* Arch Virol, 2009. **154**(1): p. 171-7.
3. Guinan, M.E., S.M. Wolinsky, and R.C. Reichman, *Epidemiology of genital herpes simplex virus infection.* Epidemiol Rev, 1985. **7**: p. 127-46.
4. Azwa, A. and S.E. Barton, *Aspects of herpes simplex virus: a clinical review.* J Fam Plann Reprod Health Care, 2009. **35**(4): p. 237-42.
5. Taylor, J.R., et al., *Interrelationship between ultraviolet light and recurrent herpes simplex infections in man.* J Dermatol Sci, 1994. **8**(3): p. 224-32.
6. Kriesel, J.D., *Reactivation of herpes simplex virus: the role of cytokines and intracellular factors.* Curr Opin Infect Dis, 1999. **12**(3): p. 235-8.
7. Fraser, N.W. and T. Valyi-Nagy, *Viral, neuronal and immune factors which may influence herpes simplex virus (HSV) latency and reactivation.* Microb Pathog, 1993. **15**(2): p. 83-91.
8. Patil, A., M. Goldust, and U. Wollina, *Herpes zoster: A Review of Clinical Manifestations and Management.* Viruses, 2022. **14**(2).
9. Ribbert, H. *Uber protozoenartige Zellen in der Niere eines syphilitischen Neugeborenen und in der Parotis von Kindern*. Zbl All Pathol.
10. Rowe, W.P., et al., *Cytopathogenic agent resembling human salivary gland virus recovered from tissue cultures of human adenoids.* Proc Soc Exp Biol Med, 1956. **92**(2): p. 418-24.
11. Smith, M.G., *Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease.* Proc Soc Exp Biol Med, 1956. **92**(2): p. 424-30.
12. Craig, J.M., et al., *Isolation of intranuclear inclusion producing agents from infants with illnesses resembling cytomegalic inclusion disease.* Proc Soc Exp Biol Med, 1957. **94**(1): p. 4-12.
13. Weller, T.H., F.C. Robbins, and J.F. Enders, *Cultivation of poliomyelitis virus in cultures of human foreskin and embryonic tissues.* Proc Soc Exp Biol Med, 1949. **72**(1): p. 153-5.
14. Kalejta, R.F., *Tegument proteins of human cytomegalovirus.* Microbiol Mol Biol Rev, 2008. **72**(2): p. 249-65, table of contents.
15. Manandhar, T., et al., *Battle between Host Immune Cellular Responses and HCMV Immune Evasion.* Int J Mol Sci, 2019. **20**(15).
16. Coute, Y., et al., *Mass Spectrometry-Based Characterization of the Virion Proteome, Phosphoproteome, and Associated Kinase Activity of Human Cytomegalovirus.* Microorganisms, 2020. **8**(6).
17. Mahmud, J. and G.C. Chan, *Analysis of Cytomegalovirus Glycoprotein and Cellular Receptor Interactions.* Methods Mol Biol, 2021. **2244**: p. 199-211.
18. Chee, M.S., et al., *Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169.* Curr Top Microbiol Immunol, 1990. **154**: p. 125-69.
19. Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.
20. Murphy, E., et al., *Coding potential of laboratory and clinical strains of human cytomegalovirus.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14976-81.

21.   Sinzger, C., et al., *Cloning and sequencing of a highly productive, endotheliotropic virus strain derived from human cytomegalovirus TB40/E.* J Gen Virol, 2008. **89**(Pt 2): p. 359-368.

22.   Dolan, A., et al., *Genetic content of wild-type human cytomegalovirus.* J Gen Virol, 2004. **85**(Pt 5): p. 1301-1312.

23.   Cunningham, C., et al., *Sequences of complete human cytomegalovirus genomes from infected cell cultures and clinical specimens.* J Gen Virol, 2010. **91**(Pt 3): p. 605-15.

24.   Sijmons, S., M. Van Ranst, and P. Maes, *Genomic and functional characteristics of human cytomegalovirus revealed by next-generation sequencing.* Viruses, 2014. **6**(3): p. 1049-72.

25.   Marti-Carreras, J. and P. Maes, *Human cytomegalovirus genomics and transcriptomics through the lens of next-generation sequencing: revision and future challenges.* Virus Genes, 2019. **55**(2): p. 138-164.

26.   Stern-Ginossar, N., et al., *Decoding human cytomegalovirus.* Science, 2012. **338**(6110): p. 1088-93.

27.   Patro, A.R.K., *Subversion of Immune Response by Human Cytomegalovirus.* Front Immunol, 2019. **10**: p. 1155.

28.   Ye, L., et al., *Functional Profile of Human Cytomegalovirus Genes and Their Associated Diseases: A Review.* Front Microbiol, 2020. **11**: p. 2104.

29.   Gruffat, H., R. Marchione, and E. Manet, *Herpesvirus Late Gene Expression: A Viral-Specific Pre-initiation Complex Is Key.* Front Microbiol, 2016. **7**: p. 869.

30.   Rasmussen, L., A. Geissler, and M. Winters, *Inter- and intragenic variations complicate the molecular epidemiology of human cytomegalovirus.* J Infect Dis, 2003. **187**(5): p. 809-19.

31.   Mattick, C., et al., *Linkage of human cytomegalovirus glycoprotein gO variant groups identified from worldwide clinical isolates with gN genotypes, implications for disease associations and evidence for N-terminal sites of positive selection.* Virology, 2004. **318**(2): p. 582-97.

32.   Sijmons, S., et al., *High-throughput analysis of human cytomegalovirus genome diversity highlights the widespread occurrence of gene-disrupting mutations and pervasive recombination.* J Virol, 2015. **89**(15): p. 7673-7695.

33.   Meyer-Konig, U., et al., *Simultaneous infection of healthy people with multiple human cytomegalovirus strains.* Lancet, 1998. **352**(9136): p. 1280-1.

34.   Dhingra, A., et al., *Human cytomegalovirus multiple-strain infections and viral population diversity in haematopoietic stem cell transplant recipients analysed by high-throughput sequencing.* Med Microbiol Immunol, 2021. **210**(5-6): p. 291-304.

35.   Suarez, N.M., et al., *Multiple-Strain Infections of Human Cytomegalovirus With High Genomic Diversity Are Common in Breast Milk From Human Immunodeficiency Virus-Infected Women in Zambia.* J Infect Dis, 2019. **220**(5): p. 792-801.

36.   Stanton, R.J., et al., *Reconstruction of the complete human cytomegalovirus genome in a BAC reveals RL13 to be a potent inhibitor of replication.* J Clin Invest, 2010. **120**(9): p. 3191-208.

37.   Revello, M.G. and G. Gerna, *Human cytomegalovirus tropism for endothelial/epithelial cells: scientific background and clinical implications.* Rev Med Virol, 2010. **20**(3): p. 136-55.

38. Cortese, M., et al., *Recombinant human cytomegalovirus (HCMV) RL13 binds human immunoglobulin G Fc.* PLoS One, 2012. **7**(11): p. e50166.

39. Schuessler, A., et al., *Mutational mapping of UL130 of human cytomegalovirus defines peptide motifs within the C-terminal third as essential for endothelial cell infection.* J Virol, 2010. **84**(18): p. 9019-26.

40. Davison, A.J., et al., *The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome.* J Gen Virol, 2003. **84**(Pt 1): p. 17-28.

41. Yu, D., et al., *Construction of a self-excisable bacterial artificial chromosome containing the human cytomegalovirus genome and mutagenesis of the diploid TRL/IRL13 gene.* J Virol, 2002. **76**(5): p. 2316-28.

42. Akter, P., et al., *Two novel spliced genes in human cytomegalovirus.* J Gen Virol, 2003. **84**(Pt 5): p. 1117-1122.

43. Das, S., A. Vasanji, and P.E. Pellett, *Three-dimensional structure of the human cytomegalovirus cytoplasmic virion assembly complex includes a reoriented secretory apparatus.* J Virol, 2007. **81**(21): p. 11861-9.

44. Manicklal, S., et al., *The "silent" global burden of congenital cytomegalovirus.* Clin Microbiol Rev, 2013. **26**(1): p. 86-102.

45. Schottstedt, V., et al., *Human Cytomegalovirus (HCMV) - Revised.* Transfus Med Hemother, 2010. **37**(6): p. 365-375.

46. Boeckh, M. and A.P. Geballe, *Cytomegalovirus: pathogen, paradigm, and puzzle.* J Clin Invest, 2011. **121**(5): p. 1673-80.

47. Gerna, G., et al., *Congenital Human Cytomegalovirus Infection: A Narrative Review of Maternal Immune Response and Diagnosis in View of the Development of a Vaccine and Prevention of Primary and Non-Primary Infections in Pregnancy.* Microorganisms, 2021. **9**(8).

48. Matthews, T. and R. Boehme, *Antiviral activity and mechanism of action of ganciclovir.* Rev Infect Dis, 1988. **10 Suppl 3**: p. S490-4.

49. Faulds, D. and R.C. Heel, *Ganciclovir. A review of its antiviral activity, pharmacokinetic properties and therapeutic efficacy in cytomegalovirus infections.* Drugs, 1990. **39**(4): p. 597-638.

50. Perrottet, N., et al., *Valganciclovir in adult solid organ transplant recipients: pharmacokinetic and pharmacodynamic characteristics and clinical interpretation of plasma concentration measurements.* Clin Pharmacokinet, 2009. **48**(6): p. 399-418.

51. Noble, S. and D. Faulds, *Ganciclovir. An update of its use in the prevention of cytomegalovirus infection and disease in transplant recipients.* Drugs, 1998. **56**(1): p. 115-46.

52. Lischka, P., et al., *In vitro and in vivo activities of the novel anticytomegalovirus compound AIC246.* Antimicrob Agents Chemother, 2010. **54**(3): p. 1290-7.

53. Bogner, E., *Human cytomegalovirus terminase as a target for antiviral chemotherapy.* Rev Med Virol, 2002. **12**(2): p. 115-27.

54. Neuber, S., et al., *Mutual Interplay between the Human Cytomegalovirus Terminase Subunits pUL51, pUL56, and pUL89 Promotes Terminase Complex Formation.* J Virol, 2017. **91**(12).

55. Borst, E.M., et al., *The human cytomegalovirus UL51 protein is essential for viral genome cleavage-packaging and interacts with the terminase subunits pUL56 and pUL89.* J Virol, 2013. **87**(3): p. 1720-32.

56.   Piret, J. and G. Boivin, *Clinical development of letermovir and maribavir: Overview of human cytomegalovirus drug resistance.* Antiviral Res, 2019. **163**: p. 91-105.

57.   Chou, S., *Rapid In Vitro Evolution of Human Cytomegalovirus UL56 Mutations That Confer Letermovir Resistance.* Antimicrob Agents Chemother, 2015. **59**(10): p. 6588-93.

58.   Goldner, T., et al., *Geno- and phenotypic characterization of human cytomegalovirus mutants selected in vitro after letermovir (AIC246) exposure.* Antimicrob Agents Chemother, 2014. **58**(1): p. 610-3.

59.   Chou, S., *A third component of the human cytomegalovirus terminase complex is involved in letermovir resistance.* Antiviral Res, 2017. **148**: p. 1-4.

60.   Chou, S., *Comparison of Cytomegalovirus Terminase Gene Mutations Selected after Exposure to Three Distinct Inhibitor Compounds.* Antimicrob Agents Chemother, 2017. **61**(11).

61.   Lischka, P., D. Michel, and H. Zimmermann, *Characterization of Cytomegalovirus Breakthrough Events in a Phase 2 Prophylaxis Trial of Letermovir (AIC246, MK 8228).* J Infect Dis, 2016. **213**(1): p. 23-30.

62.   Douglas, C.M., et al., *Letermovir Resistance Analysis in a Clinical Trial of Cytomegalovirus Prophylaxis for Hematopoietic Stem Cell Transplant Recipients.* J Infect Dis, 2020. **221**(7): p. 1117-1126.

63.   Institute of Medicine Committee to Study Priorities for Vaccine, D., *The National Academies Collection: Reports funded by National Institutes of Health*, in *Vaccines for the 21st Century: A Tool for Decisionmaking*, K.R. Stratton, J.S. Durch, and R.S. Lawrence, Editors. 2000, National Academies Press (US)
Copyright 2000 by the National Academy of Sciences. All rights reserved.: Washington (DC).

64.   Al-Sharif, W.Z., et al., *Sea urchin coelomocytes specifically express a homologue of the complement component C3.* J Immunol, 1998. **160**(6): p. 2983-97.

65.   Levashina, E.A., et al., *Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, Anopheles gambiae.* Cell, 2001. **104**(5): p. 709-18.

66.   Kimura, A., E. Sakaguchi, and M. Nonaka, *Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, Nematostella vectensis.* Immunobiology, 2009. **214**(3): p. 165-78.

67.   Mortensen, S.A., et al., *Structure and activation of C1, the complex initiating the classical pathway of the complement cascade.* Proc Natl Acad Sci U S A, 2017. **114**(5): p. 986-991.

68.   Pangburn, M.K. and N. Rawal, *Structure and function of complement C5 convertase enzymes.* Biochem Soc Trans, 2002. **30**(Pt 6): p. 1006-10.

69.   Merle, N.S., et al., *Complement System Part I - Molecular Mechanisms of Activation and Regulation.* Front Immunol, 2015. **6**: p. 262.

70.   Podack, E.R., J. Tschoop, and H.J. Muller-Eberhard, *Molecular organization of C9 within the membrane attack complex of complement. Induction of circular C9 polymerization by the C5b-8 assembly.* J Exp Med, 1982. **156**(1): p. 268-82.

71.   Morgan, B.P., *Regulation of the complement membrane attack pathway.* Crit Rev Immunol, 1999. **19**(3): p. 173-98.

Chapter 2

**Interactome screening of human Complement proteins against the HCMV genome by
yeast-two-hybrid**

ABSTRACT

The large 230kb DNA genome of Human Cytomegalovirus (HCMV) encodes an unparalleled number of viral proteins that subvert innate and adaptive immunity. To date, little is known regarding the role of the complement system in infection and immunity to HCMV. In this study, we screened 6 human complement proteins: Decay Accelerating Factor (DAF/CD55), Factor H (FH), Mannose Binding Lectin 2 (MBL2), Mannan Associated Serine Protease 1 (MASP1), Properdin (CFP), and C1q Binding Protein (C1qBP) against an HCMV genomic library in yeast containing 167 viral open reading frames (ORF's). Out of 1,002 possible HCMV-Complement interactions tested, 121 (8.2%) total positive protein-protein interactions were identified for all six complement proteins tested. The number of positive interactions for each complement protein against the HCMV genomic library ranged from 12-31. We validated a subset of these interactions in human cells by co-immunoprecipitation and identified 6 novel interactions in HCMV-encoded proteins: UL24(MBL2), UL25(C1qBP, MASP1), UL82(MBL2, C1qBP), and UL84(MBL2). Taken together, our data suggests: 1) HCMV encodes multiple proteins that interact with one or more proteins of the human complement system, 2) HCMV-Complement interactions occur in proteins represented across all three pathways of complement, and 3) HCMV encodes proteins that interact with both positive (activating) and negative (inhibiting) complement regulating proteins. To date, our study is the largest, most comprehensive, and systematic investigation of protein-protein interactions between HCMV and the human complement system. The knowledge derived from this study can provide a framework to further investigate the functional significance of protein interactions between HCMV and the human complement system that may underlie undiscovered mechanisms of innate immune evasion by HCMV and potentially inform the development of novel drugs and vaccines for the treatment and prevention of HCMV infections.

INTRODUCTION

Viruses are obligate parasites of cells and must be equipped to confront both the extracellular and intracellular environments of cells. To survive in both microenvironments, viruses have evolved elegant strategies to evade detection and destruction by the host innate immune system. One important first line of innate immunity is the complement system. The complement system consists of ~30 proteins present in serum and on cell surfaces that collectively form three distinct yet converging pathways (classical, lectin, and alternative) [1]. The classical pathway (CP) of complement is activated by antibody-antigen complexes, the lectin pathway (LP) of complement is activated by foreign carbohydrates on the surface of microorganisms, and the alternative pathway (AP) can be activated spontaneously or by foreign molecules on the surface of microbes. All three pathways can be triggered by distinct mechanisms but ultimately converge at the level of C3, which upon activation initiates a highly-coordinated cascade of proteolytic reactions resulting in inflammation, opsonization, recruitment of immune cells to sites of infection, or direct lysis of microorganisms (via the membrane attack complex (MAC), C5b-C9).

Complement and herpesviruses. Interactions with the complement system by human herpesviruses have been documented to varying degrees among all three subfamilies: *Alphaherpesvirinae* (HSV-1, HSV-2, VZV), *Betaherspesvirinae* (HHV-6A, HCMV), and *Gammaherpesvirinae* (EBV, KSHV). The members of the *Alphaherpesvirinae* encode gC which binds to the central complement component C3 and its activation products (C3b, iC3b) [2]. gC also interferres with Factor H, Properdin, and C5, resulting in inhibition of complement [3-5]. Additionally, HSV-1 encodes glycoproteins gE and gI, which together form a complex that functions as an IgG Fc receptor (FcgR) that occludes the Fc domain of IgG and inhibits antibody-dependent complement mediated neutralization in the classical pathway [6]. For *Gammaherpesvirinae,* EBV encodes gp350 that interacts with complement receptors CR1 [7] and CR2 [8] for viral entry into B cells. In contrast, KSHV encodes KCP (ORF4), a viral complement control protein homolog of human CD55 that decays the C3 convertase and inactivates C3b and C4 to inhibit complement [9, 10]. For *Betaherspesvirinae*, HHV-6A encodes a protein complex composed of gH-gL-gQ which serves as a viral ligand for CD46 (a cell surface complement inhibitor) and this interaction is both necessary for HHV-6A infectivity and membrane fusion [11, 12]. In HCMV, no virally-encoded complement regulating proteins have been identified, yet HCMV is reported to alter the expression of several complement proteins during infection and is resistant to complement-mediated neutralization in the absence of antibody [13]. Taken together, human herpesviruses have evolved unique strategies to interact with and exploit the human complement system by various mechanisms including: inhibiting complement protein synthesis, upregulating and acquiring membrane bound host complement inhibitors, recruiting soluble complement inhibitors from serum, using cellular complement receptors to gain access to cells, or by encoding viral proteins with complement regulating functions.

HCMV virions and infected cells resist complement-mediated neutralization. The role of the complement system in immunity to HCMV remain poorly understood, even though the protective role of complement has been well characterized in other related human herpesviruses [14]. In the absence of specific antibody, HCMV virions incubated in serum (a source of complement) are not neutralized but are nevertheless coated with activated complement protein fragments (C3a, C3b, C9) which normally proceed the formation of the membrane attack complex and results in direct lysis of microbes [13]. This observation suggests that HCMV virions not only activate one or more pathways of complement but are able to inhibit complement-mediated neutralization, perhaps by a virally encoded complement regulating protein. The objective of this study is to identify novel protein interactions between HCMV-encoded open reading frames (ORF's) and a select number of human complement proteins by yeast-two-hybrid assay.

Selection of Complement proteins to investigate interactions with HCMV. In this chapter, we describe in detail the molecular structures and functions of six human complement proteins (CD55, MBL2, MASP1, C1qBP , FH, and Properdin) selected for inclusion in our study to investigate whether the HCMV genome encodes proteins that interact with components of the human complement system. The six human complement proteins were carefully selected based on a combination of the following criteria: 1) their role in one or more distinct pathways of complement, 2) regulatory functions, and 3) reported altered expression during HCMV infection *in vitro*. The molecular architecture of each complement protein selected for interactome screening is illustrated in **Figure 2-1.**
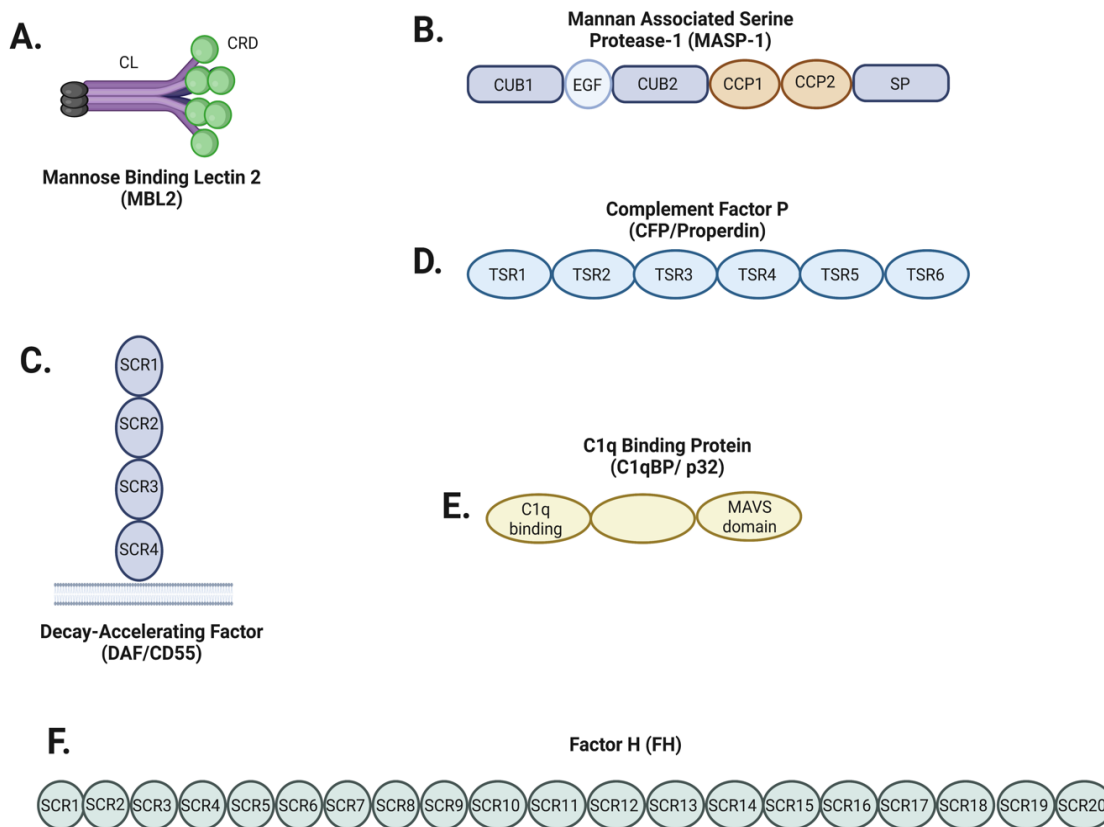
Figure 2-1. Molecular architecture of select complement proteins.
Illustrations of a subset of human complement proteins and their known protein domains. All protein structures are depicted from the N-terminus to the C-terminus in the left to right direction. (A) Mannose Binding Lectin 2 (MBL2), the primary pattern recognition molecule in the lectin pathway of complement. (B) Mannan Associated Serine Protease 1 (MASP-1), the serine protease that cooperates with MASP-2 for activation of the lectin pathway of complement. (C) Decay-Accelerating Factor (DAF/CD55), a complement regulatory protein that inhibits the formation of the C3 convertase in all three pathways of complement. (D) Complement Factor P (CFP/Properdin), a complement regulatory protein that stabilizes C3 convertases and positively regulates complement. (E) C1q Binding Protein (C1qBP/p32), a molecule that can inhibit the classical pathway by binding to the pattern recognition molecule C1q. (F) Factor H (FH), a complement regulatory protein that inhibits the alternative pathway by binding to C3b. Figured illustrations designed with Biorender (www.biorender.com).

Mannose Binding Lectin 2 (MBL2). Humans encode a single functional mannan binding lectin known as Mannose Binding Lectin 2 (MBL2). Structurally, the 248 amino acid sequence of MBL2 encodes a ~24kD polypeptide that contains four distinct regions: a cystine rich N-terminal region, a collagen-like domain, a short α-helical coiled-coil domain (neck), and a C-terminal carbohydrate recognition domain that forms the globular head of the MBL2 molecule [15] as illustrated in **Figure 2-1 (A)**. In serum, MBL2 exists in a highly ordered homotrimer structure that recognizes specific polysaccharides such as mannose or N-acetylglucosamine present on various classes of microorganisms including bacteria, viruses, protozoa, and fungi [16]. Although less common, MBL2 can also bind to phospholipids and nucleic acids [17, 18], and has been reported to contribute to the clearance of necrotic tissues [19]. In innate immunity, MBL2 is a pattern recognition molecule and potent opsonin in the lectin pathway of complement and is a major protein released by the liver into the serum as part of the "acute phase" response to infection, inflammation, or injury.

 Despite having a major role in the lectin pathway of complement, MBL2 has also been reported to neutralize pathogens in a complement-independent manner by enhancing opsonization by phagocytic cells. Although the specific phagocytic cell surface receptor for MBL2 remains unknown, it is suggested that MBL2 may bind to the protein folding chaperone calreticulin which is often found in the Endoplasmic Reticulum (ER) and contributes to antigen presentation [20]. Despite the absence of a clear molecular mechanism for the role of MBL2 in HCMV infection, genetic polymorphisms that result in low MBL2 expression and genetic deficiencies in MBL2 are correlated to an increased susceptibility to infections with various microorganisms, including HCMV [21, 22]. Taken together, MBL2 may play a protective role immunity to HCMV.

Mannan Associated Serine Protease 1 (MASP-1). The human MASP-1 and MASP-2 genes can together produce five MBL-associated serine proteases (MASP's) by alternative splicing. Of the five MASP's encoded by humans (MASP-1, MASP-2, MASP-3, Map44, Map19), MASP-1 is one of the most important serine proteases in the lectin pathway of complement and exists in concentrations much higher than that of similar proteases [23]. MASP-1 and MASP-2 cleave C3 and C4 respectively, while C2 is cleaved by both [24]. Under physiological conditions, MASP-1 is the exclusive activator of MASP-2 which allows for the formation of the C3 convertase (C4bC2a). MASP-1 can autoactivate and directly cleave C2, and is in fact the serine protease that produces an overwhelming majority of the C2a that aids in the formation of the C3 convertase (C4bC2a) [25]. In contrast, MASP-2 can only cleave C4, thus full activation of the lectin pathway of complement specifically requires both MASP-1 and MASP-2 [23, 26].

Structurally, the 699 amino acid sequence of MASP-1 encodes a ~79kD polypeptide with five distinct regions: an N-terminal CUB domain, an epidermal growth factor-like (EGF) domain, a CUB2 domain, two complement control domains (CCP1, CCP2), and a C-terminal serine protease (SP) domain [27, 28] as illustrated in **Figure 2-1 (B)**. Upon ligand binding, such as MASP-1 binding to MBL2 on complement activating surfaces, lectins undergo a conformational change that brings the serine protease domains of MASP-1 in proximity which allows for proteolytic autoactivation, and subsequent activation of other MASP's [29]. The activated MASP-1 protein structure consists of a heavy and light chain held together by a disulfide bond, with MASP-1 being unique from other lectin pathway serine proteases due to its highly promiscuous protease activity which is the result of a wide substrate binding grove that more closely resembles trypsin rather than other complement proteases [30]. Thus, compared to MASP-2 and MASP-3, MASP-1 is known to cleave the largest number of substrates [31].

Importantly, MASP-1 has also been reported to serve additional innate intracellular immune functions separate from the complement system, such as contributing to Ca^{2+} signaling, NF-kB, and p38 MAPK pathways in cells through protease-activated receptor 4 (PAR4) induced by MASP-1 proteolytic activity [32]. Additionally, MASP-1 has been reported to modulate the immune response by promoting the release of the proinflammatory molecule bradykinin [33]. In response to HCMV infection *in vitro*, MASP-1 expression is suppressed at both the transcriptional and translational level [34], an observation which may suggest that MASP-1 serves an important antiviral role during HCMV infection.

Decay Accelerating Factor (DAF/CD55). Decay Accelerating Factor (DAF/CD55) is an important regulatory glycoprotein that is highly expressed on the surface of nearly all human cells [35]. The ubiquitous expression of CD55 on human cells stems from its important function in inhibiting complement activation on the surface of host cells by recognizing C4b and C3b fragments generated during the cleavage of C4 during classical or lectin pathway activation, or C3 cleavage fragments generated via the alternative pathway [36]. CD55 can inhibit all three pathways of complement by interfering with the formation of the C3-convertases (C4bC2a or C3bBb), and thus indirectly block the formation of the membrane attack complex. In the classical and lectin pathways, CD55 interacts with C4b deposited on the surface of cells and interferes with the cleavage of C2 to C2b thus limiting classical/lectin pathway C3-convertase formation. In the context of the alternative pathway, CD55 interacts with C3b and inhibits the cleavage of Factor B into Bb by Factor D, thereby preventing the formation of the alternative pathway C3-convertase (C3bBb) [37].

Structurally, the 381 amino acid sequence of CD55 encodes a ~41kD polypeptide with five distinct regions: four short consensus repeat (SCR) domains located on the N-terminus of the protein, and a C-terminal glycosylphosphatidylinositol (GPI) anchor that attaches the proteins to the plasma membrane as illustrated in **Figure 2-1 (C)** [38]. Importantly, the ability of CD55 to regulate complement requires SCR2, SCR3, and SC4 [39].

In response to HCMV infection *in vitro*, CD55 expression is significant upregulated at both the transcriptional and translational level [34, 40]. Along with CD55, HCMV infected cells upregulate the expression of CD46 and CD59, two other host cell membrane bound complement inhibiting proteins, and all three proteins are incorporated into the viral envelope of newly formed HCMV virions [41]. While all three proteins have complement inhibiting functions, CD55 appears to be the most important in contributing to the resistance of viral particles against complement-mediated neutralization [13]. This critical mechanism is underscored by the fact that HCMV virions incubated with anti-CD55 antibodies become susceptible to neutralization by complement, though the precise viral proteins that mediate the acquisition of host complement inhibitors CD55, CD46, and CD59 into the viral envelope remain unknown. Additionally, it is unclear whether the acquisition of host CD55 during infection is a mechanism that is conserved across all HCMV isolates, since studies have reported stark differences in sensitivity to complement among diverse HCMV clinical isolates [42]. These observations suggest that HCMV may resist neutralization by complement via multiple mechanisms, that may not solely depend on the acquisition of host CD55 but could be due to additional virally encoded complement regulating proteins present in some but not all HCMV isolates.

Properdin (FP/CFP). Properdin is one of the few molecules in the complement system that functions as a positive regulator of complement by specifically augmenting the alternative pathway via stabilization of C3bBb C3-convertase complexes. Recently, properdin was also proposed to be a pattern recognition molecule with the ability to initiate complement activation [43, 44]. Properdin circulates in serum in relatively high concentrations and is unique among complement proteins in that it is primarily produced by leukocytes rather than hepatocytes [45], which are the primary cellular source of complement proteins produced by the body. Properdin has many known ligands including: C3b, C3bB, and C3bBb, as well as sufatides and various polyanionic structures [46].

Structurally, the 469 amino acid sequence of Properdin encodes a ~51kD polypeptide glycoprotein with six distinct regions beginning with a truncated N-terminal thrombospondin type repeat (TSR) that contains key conserved residues (not shown in figure), followed by six additional TSR domains (TSR 1-6) that extend to the C-terminus of the protein as illustrated in **Figure 2-1 (D)**. Under physiological conditions, properdin monomers form higher order structures that can include cyclic dimers, timers, and tetramers in approximately 1:2:1 ratio [47] with different structures being implicated in interactions with different ligands. The contributions of each TSR's in the functional activity of Properdin have been well documented. TSR4 is responsible for binding to C3b, while the contribution of both TSR4 and TSR5 are required for stabilization of the C3bBb convertase of the alternative pathway [48].

In response to HCMV infection *in vitro,* transcriptional studies have not reported the differential expression of Properdin, though this may be because the synthesis of Properdin is largely restricted to immune cells and most transcriptional studies of HCMV infection use fibroblast cell lines which are the leading cell type used to study HCMV infection.

C1q Binding Protein (C1qBP). C1qBP is known by various names including: C1qBP (C1q Binding protein), gC1qR (globular C1q receptor), Hyaluronic acid binding protein 1 (HABP1), and p32. The reason for such complex and confusing nomenclature stems from the fact that C1qBP was originally named and described based on its location and function by several independent groups. C1qBP is found in many subcellular locations including the cell membrane, mitochondria, cytoplasm, and is secreted extracellularly [49]. Given the preceding, it is unsurprising that C1qBP is a multifunctional protein that can interact with many host proteins and is involved in diverse cellular processes depending on its location including: maintaining mitochondrial homeostasis and oxidative phosphorylation [50], mediating nucleus-mitochondrial interactions [51], mitochondrial antiviral signaling (MAVS) [52], and even transcriptional regulation [53] . In the context of the complement system, C1qBP is expressed on the surface of cells and is a receptor for C1q which allows C1qBP to regulate the classical pathway of complement by inhibiting C1 activation through engagement with C1q [54].

Structurally, the 282 amino acid sequence of C1qBP encodes a ~31kD polypeptide with three distinct domains beginning with N-terminal C1q binding domain, a central disordered domain, and a C-terminal domain that is involved in MAV signaling as illustrated in **Figure 2-1 (E)**.

The role of C1qBP in complement mediated immunity to HCMV is unknown, however C1qBP is reported to play an essential intracellular role during HCMV viral replication. In response to HCMV infection *in vitro*, C1qBP expression is significant upregulated at both the transcriptional and translational level [34], though studies have not determined whether the increased expression of C1qBP also results in increased secretion of C1qBP. Extracellularly, high local concentrations of C1qBP would be expected to result in impaired classical pathway complement activation. Given this possibility, HCMV may potentially encode surface proteins that bind to secreted C1qBP and prevent neutralization of infected cells by the classical pathway of complement.

Factor H (FH). Factor H (FH) is a highly abundant serum glycoprotein that regulates the alternative pathway of complement. FH can inhibit the assembly of the alternative pathway C3 and C5 convertases (C3bBb, C3bBbC3b) by competing with Factor B for C3b binding [55], and facilitate the disassembly of the convertases by displacing bound Factor Bb (decay accelerating activity) [56], or acting as a cofactor for Factor I mediated cleavage and inactivation of C3b [57].

Structurally, the 1,231 amino acid sequence of FH encodes a ~139kD single chain polypeptide glycoprotein with twenty short consensus repeats (SCR's) that span from the N to C-terminus of the protein as illustrated in **Figure 2-1 (F).** The N-terminal SCR1-4 contribute to the decay accelerating and cofactor activities of FH [58-60], while SCR19-20 are involved in binding to host cells [61, 62]. Importantly, FH exists in multiple isoforms and can be membrane bound or secreted allowing for complement regulation either in serum or on complement activating surfaces. Thus, a unique feature of FH is its ability regulate complement activity when in the fluid phase of serum, as well as controlling complement activation on self surfaces when FH is bound to cell membranes [63]. Several FH SCR domains have been reported to participate in the binding of C3, C3 activation products, as well as various polyanions such as sialic acids, glycosaminoglycans and heparin. The FH SCR 1-6 bind C3 and C3b [64] , while SCR 19-20 bind C3 activation products such as iC3b, C3b, and C3d [64, 65].

In response to HCMV infection *in vitro*, FH expression is suppressed at both the transcriptional and translational level, though the levels of FH expression reported in transcriptome studies were low [34], likely due to the fact that fibroblasts do not produce high levels of FH [66]. Nevertheless, many viral pathogens hijack host FH to inhibit complement activation on microbial surfaces by encoding surface ligands that bind to FH [67, 68]. Thus, given that binding of FH is a common tactic by genetically diverse viruses to evade the complement system, we questioned whether HCMV might also encode a FH ligand.

RESULTS

Construction of the HCMV genomic yeast-two-hybrid library. The construction of the HCMV genomic yeast-two-hybrid library used in the present study has been previously described [69], though at the time consisted of only 59 HCMV ORF's encoding: capsid, tegument, and envelope proteins. Our laboratory cloned an additional 116 HCMV ORF's, generating a library that covers nearly all the predicted ORF's in the HCMV genome using the same methods. In total, the HCMV genomic yeast-two-hybrid library in the present study contains 175 viral ORF's. In brief, the HCMV Towne strain genomic sequence was cloned into a bacterial artificial chromosome (BAC) vector [70] and maintained as a single copy BAC-based plasmid in *E. coli*. Human cells transfected with the preceding Towne BAC vector produce infectious HCMV virions that retain wild-type growth characteristics *in vitro*, suggesting that the genomic sequence cloned into the BAC is both stable and encodes the full repertoire of viral genes responsible for productive viral replication and infection [70, 71]. A locally written Unix-based script or automation of GCG package was used to analyze the Towne-BAC sequence and determine coding sequences for viral ORF's ≥100 codons. Each codon was compared with the set of ORF's that had been predicted or reported in all of the HCMV strains for which sequences had been determined at the time of the library construction [70, 72-75]. The 175 HCMV ORF's were cloned and expressed in the yeast-two-hybrid system as previously described [69] using forward primers (from 5' to 3') homologous to the sequence immediately after the predicted translation initiation codon and 20-25 additional nucleotides of coding sequence. The reverse primers contained the reverse complement of both the predicted translation termination codon and the preceding 20-25 nucleotides at the end of the ORF. Each primer was flanked by restriction enzyme sequences to facilitate cloning of the viral ORF's into the yeast-two-hybrid vector pGBKT7. All yeast-two-hybrid constructs were verified by DNA sequence analysis and restriction enzyme profiling (data not shown). A workflow of the process for generating the HCMV genomic yeast-two-hybrid library is illustrated in **Figure 2-2.** A comprehensive list of all primers used to amplify each individual HCMV ORF for cloning into yeast-two-hybrid vectors are listed in **Supplemental Table 1**.
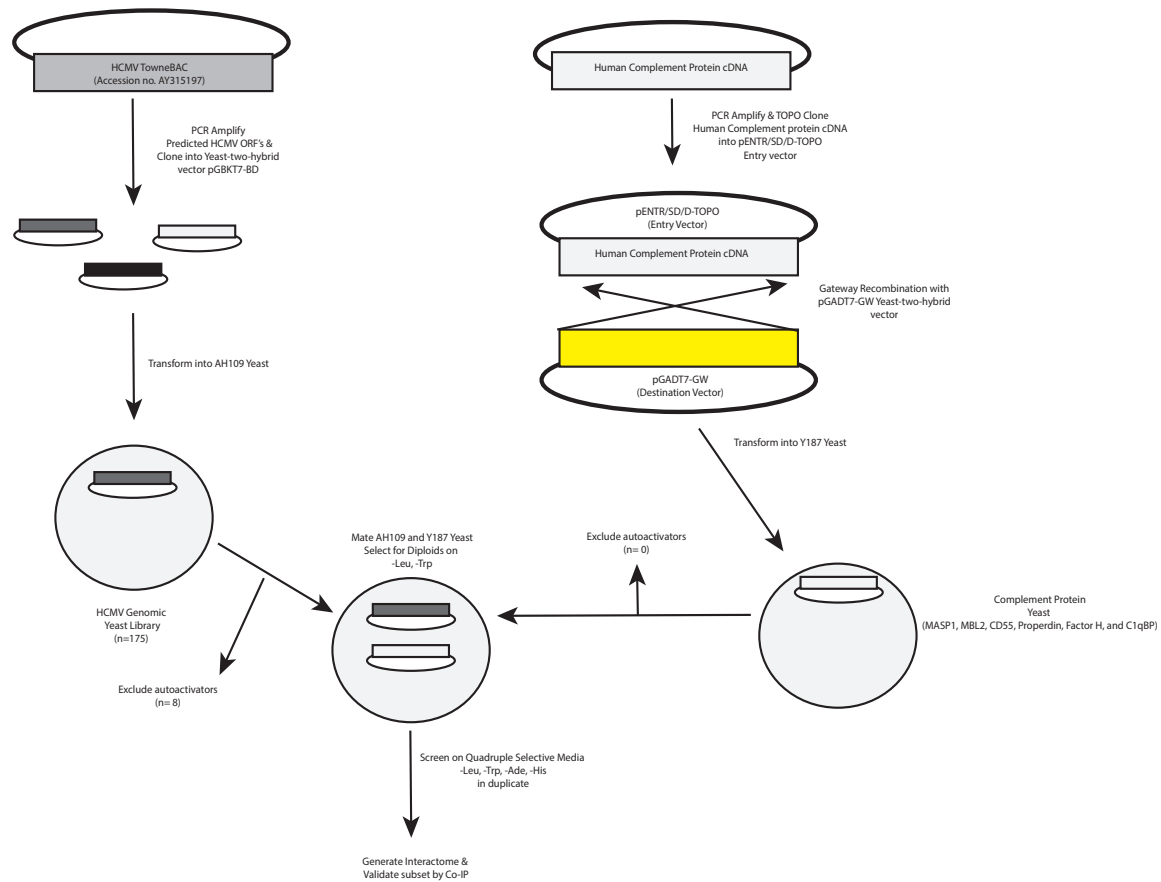
Figure 2-2. Cloning of HCMV ORF's and human complement proteins for yeast-two-hybrid analysis.

Flow chart illustrating the construction of HCMV ORF's and human complement proteins cloned into yeast expression "bait" pGBKT7-BD (HCMV) and "prey" pGADT7-AD (Complement) plasmids. pGADT7-AD plasmids encoding human complement proteins and pGBKT7-BD plasmids encoding HCMV proteins were transformed into yeast strains AH109 and Y187 respectively. All yeast transformants were tested for autoactivation, and those that were positive were subsequently excluded from further screening. A total of 1,002 unique HCMV-Complement protein interactions were tested in duplicate independent yeast matings. A subset of positive interactions were further validated by co-immunoprecipitation in human cells.

Construction of Complement yeast-two-hybrid expression vectors. A workflow of the process for generating the human complement protein yeast-two-hybrid vectors is illustrated in **Figure 2-2.** We used commercially available human cDNA's encoding Decay Accelerating Factor (DAF/CD55), Factor H (FH), Mannose Binding Lectin 2 (MBL2), Mannan Associated Serine Protease 1 (MASP1), Properdin (CFP), and Complement component 1q binding protein (C1qBP) as templates to clone and express human complement proteins in the yeast-two-hybrid system. In humans, the genes encoding complement components are complex and can be transcribed into many isoforms which can be membrane bound, secreted, or exist intracellularly. We selected the most common and canonical cDNA isoforms encoding each human complement protein and designed PCR primer pairs for each ORF as listed in **Supplemental Table 2**. In contrast to the restriction enzyme cloning used to construct the HCMV yeast-two-hybrid genomic library, we used a gateway cloning approach to clone human complement protein cDNA into yeast expression vectors due to the overlapping presence of restriction sites within the human complement protein cDNA and the multiple cloning sites of pGADT7-AD yeast expression vector (data not shown). To overcome this obstacle, we used a gateway compatible pGADT7-GW derived from the original pGADT7-AD vector [76]. The cloning strategy was as follows. The forward primer (from 5' to 3') began after the localization signal sequence encoded in each cDNA and included a four base pair (CACC) sequence necessary for directional topoisomerase (TOPO) cloning, followed by 19-31 additional nucleotides homologous to the coding sequence. The rationale for excluding the native localization signal sequence in each of the complement protein cDNA's was to avoid mislocalization of the complement proteins outside of the nucleus once expressed in yeast. The reverse primer contained the reverse complement of the translation termination codon and the preceding 18-32 nucleotides at the end of the ORF. The resultant PCR products were directionally TOPO cloned into the pENTER-SD vectors, which were later used as entry vectors to perform gateway recombination with the yeast expression vector pGADT7-GW as described in detail in the materials and methods. All human complement protein yeast-two-hybrid vectors were verified by DNA sequence analysis and restriction enzyme profiling (data not shown).

Identification of Complement-HCMV interactions by yeast-two-hybrid. To test for protein-protein interactions between human complement proteins and HCMV-encoded proteins, we performed a large-scale yeast-two-hybrid screen in 48-well format that consisted of yeast transformants each expressing a human complement protein or HCMV ORF as a hybrid protein fused with the Gal4 Activating domain (AD) or DNA binding domain (BD) respectively. Prior to performing mating experiments, all 175 HCMV pGBKT7-fusion proteins in yeast strain AH109 were tested for autoactivation. Eight AH109 strains, which contained HCMV ORF's UL26, UL48A, UL48N, UL48C, UL48.5, UL51, UL94, and US23 were determined to be autoactivators in the absence of any pGADT7-cloned ORF's and thus assigned negative for the purposes of this study, and subsequently eliminated from any further mating experiments. Similarly, all 6 human complement pGADT7-fusion proteins in yeast strains Y187 were tested for autoactivation under the same conditions, and none were found to autoactivate in the absence of any pGBKT7-cloned ORF's.

After eliminating autoactivators, the remaining 167 AH109 strains carrying HCMV pGBKT7-fusion proteins were mated with each of the 6 Y187 strains encoding single human complement proteins (CD55, FH, MBL2, MASP1, Properdin, and C1qBP) as pGADT7-fusion proteins and

plated on selective media to select for diploids. To increase fidelity, all matings were performed in duplicate and under the highest assay stringency (media lacking: -Ade, -Leu, -Trp, -His) for the selection of positive protein interactions. Aside from selecting for the presence of both plasmids pGADT7 and pGBKT7 based on leucine and tryptophan selection respectively, we also further increased the stringency of the assay by using three more reporter genes to reduce the rate of potential false positives: histidine, adenine, and Mel1, which encodes alpha-galactosidase for blue/white colony visual screening.

Figure 2-3 depicts representative protein-protein interactions between human complement proteins and HCMV ORF's identified by yeast-two-hybrid. The genes encoding human p53 (a tumor suppressor) and large T antigen (from SV40 virus) were cloned into the vectors pGBKT7 and pGADT7 respectively, and used as a positive control for protein-protein interaction since these proteins are known to interact [77]. As a negative control for protein-protein interactions, yeast transformed with pGBKT7-p53 were mated with yeast of the opposite mating type that had been transformed with pGADT7-LamC which encodes the human Lamin C protein, which is known to be a highly inert and unreactive nuclear structural protein. Using our large scale yeast-two-hybrid approach, we tested a total of 1,002 interactions between 6 human complement proteins (CD55, FH, MBL2, MASP1, Properdin, and C1qBP) and 167 HCMV-encoded ORF's.

Figure 2-4 illustrates the summary of high-confidence protein-protein interactions identified by our yeast-two-hybrid screen. For six human complement proteins tested, we identified a total of 121 positive HCMV-Complement protein interactions, representing a 12% positivity rate out of 1,002 possible unique HCMV-Complement protein interactions tested. The number of positive protein-protein interactions ranged from 12 to 31 for each of the six complement proteins tested against the HCMV genomic yeast-two-hybrid library.

Controls

(+) p53 x SV40-T   (-) p53 x Lamin C

Experimental

(+) UL7 x FH   (-) UL2 x CD55
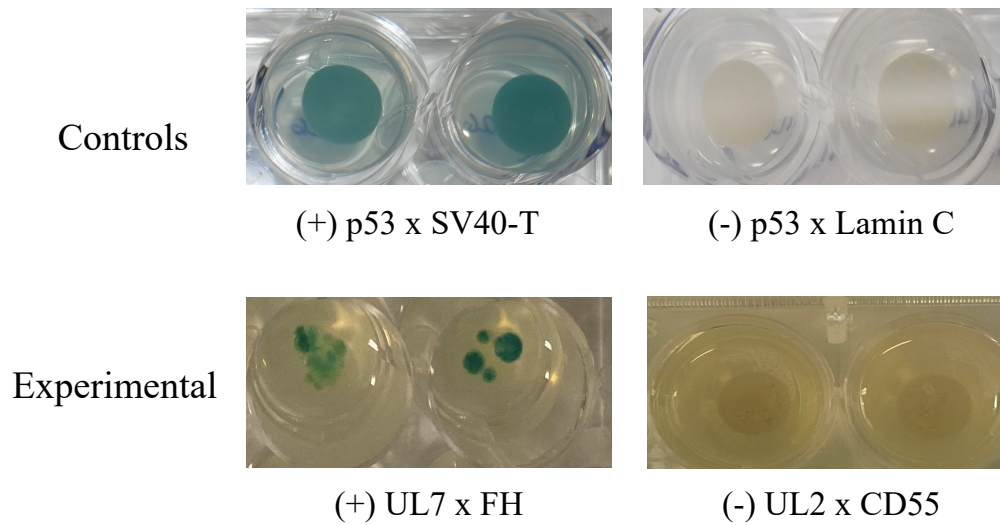
Figure 2-3. Representative protein-protein interactions between human Complement proteins and HCMV proteins identified by yeast-two-hybrid analysis.
Representative protein-protein interactions between Complement and HCMV proteins (bottom) compared to control protein-protein interactions (top). The positive control protein-protein interactions (Top left) are illustrated by blue yeast colonies in yeast expressing the human p53 protein fused to the BD domain of GAL4 and SV40 T-antigen fused to the AD domain of GAL4. The negative control protein-protein interactions (Top right) are illustrated by white colonies in the yeast expressing the human p53 fused to the BD domain of GAL4 and human Lamin C fused to the AD domain of GAL4. Positive experimental protein-protein interactions (bottom left) are illustrated between HCMV UL7 fused to the BD domain of GAL4 and FH fused to the AD domain of GAL4. Negative experimental protein-protein interactions (bottom right) are illustrated between HCMV UL2 fused to the BD domain of GAL4 and CD55 fused to the AD domain of GAL4. Each yeast-two-hybrid mating was performed in ≥ 2 independent replicates.

Summary of protein-protein interactions (PPI's) identified by yeast-two-hybrid assay
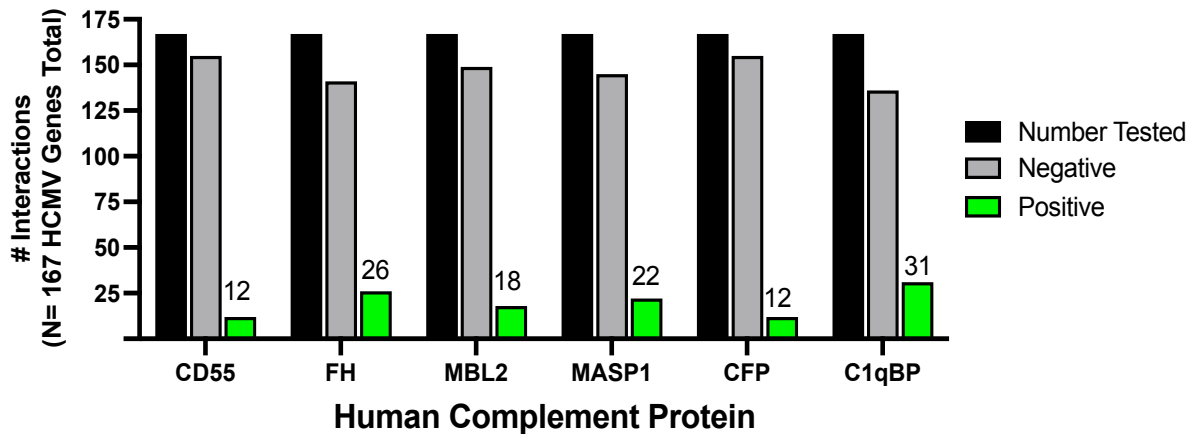
Figure 2-4. Summary of Complement-HCMV yeast-two-hybrid interactions. The vertical axis represents the number of possible HCMV genes in the yeast-two-hybrid library (excluding autoactivators). The horizontal axis lists the six human complement proteins tested for protein-protein interactions. The black bar represents the total number of proteins tested, the dark grey bar represents the number of proteins tested that resulted in no protein-protein interactions (negative), and the green bar represents the number of proteins tested that resulted in protein-protein interactions (positive). The number above the bars shows the number of HCMV-Complement interactions that were positive for each of the six human complement proteins tested (CD55, FH, MBL2, MASP-1, Properdin, C1qBP).

CD55 Interactions. For the complement inhibiting protein CD55 which regulates all three pathways of complement, we identified a total of 12 positive protein-protein interactions in our HCMV genomic yeast-two-hybrid library. As illustrated in **Figure 2-5**, of the 12 positive protein-protein interactions, 67% (n=7) occurred in proteins involved in nuclear assembly and viral replication (UL45, UL49, UL52, UL53, UL86, UL96, and US27). Of the remaining positive protein-protein interactions: 17% (n=2) were viral membrane proteins (UL6, UL132), 8% (n=1) were tegument proteins (UL24), 8% (n=1) were viral proteins involved in innate immune evasion (UL31), and 8% (n=1) were in proteins of unknown function (US26). A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with CD55 are listed in **Table 2-1.**

Figure 2-5. CD55-HCMV interactions identified by yeast-two-hybrid analysis.

#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD)	Function
			Table 2-1. CD55-HCMV Interactions identified by yeast-two-hybrid analysis
1	UL6	$D^{\Delta\dagger}$	Host membrane - single-pass membrane protein
2	UL24	$D^{\Delta\dagger}$	Virion protein, cell-tropism specific replication
3	UL31	$GD^{\Delta\dagger}$	inhibits CGAS mediated innate immunity
4	UL45	$D^{\Delta\dagger}$	Does not possess a ribonucleotide reductase activity. Betaherpesviruses probably use another strategy to expand the dNTP pool in a quiescent host cell.
5	UL49	$E^{\Delta\dagger}$	Viral DNA Replication
6	UL52	$E^{\Delta\dagger}$	Plays a role in efficient localization of neo-synthesized capsids to nuclear replication compartments, thereby controlling cleavage and packaging of virus genomic DNA.
7	UL53	$E^{\Delta\dagger}$	Plays an essential role in virion nuclear egress, the first step of virion release from infected cell. Within the host nucleus, NEC1 interacts with the newly formed capsid through the vertexes and directs it to the inner nuclear membrane by associating with NEC2. Induces the budding of the capsid at the inner nuclear membrane as well as its envelopment into the perinuclear space. There, the NEC1/NEC2 complex promotes the fusion of the enveloped capsid with the outer nuclear membrane and the subsequent release of the viral capsid into the cytoplasm where it will reach the secondary budding sites in the host Golgi or trans-Golgi network.
8	UL86	$E^{\Delta\dagger}$	Self-assembles to form an icosahedral capsid with a T=16 symmetry, about 200 nm in diameter, and consisting of 150 hexons and 12 pentons (total of 162 capsomers). Hexons form the edges and faces of the capsid and are each composed of six MCP molecules. In contrast, one penton is found at each of the 12 vertices. Eleven of the pentons are MCP pentamers, while the last vertex is occupied by the portal complex. The capsid is surrounded by a layer of proteinaceous material designated the tegument which, in turn, is enclosed in an envelope of host cell-derived lipids containing virus-encoded glycoproteins.
9	UL96	E^{Δ}, $GD^{\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription
10	UL132	GD^{Δ}, $D^{\dagger}$	Viral entry into host cell, viral life cycle
11	US26	Severe growth defect	Viral Transcription
12	US27	$GD^{\Delta\dagger}$	Plays an important role in spread of HCMV via the extracellular route. As a G-protein-coupled receptor (vGPCR), may activate signaling pathways important for virion assembly or egress processes.

Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:

Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]

† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]

Factor H Interactions. For the complement inhibiting protein FH, which negatively regulates the alternative pathway of complement, we identified a total of 26 positive protein-protein interactions. As illustrated in **Figure 2-6**, of the 26 positive protein-protein interactions, 30% (n=8) occurred in HCMV proteins of unknown function (UL12, UL13, UL30, UL61, UL63, UL67, UL81, UL131). Of the remaining positive protein-protein interactions: 27% (n=7) were viral membrane proteins (RL10, UL7, UL11, UL100, UL119, UL132, US15), 27% (n=7) were involved in nuclear assembly/viral replication (UL38, UL50, UL52, UL88, UL91, UL92, UL104), and the remaining 16% (n=4) were in tegument proteins (UL24, UL43, UL47, UL82). A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with Factor H are listed in **Table 2-2.**

Figure 2-6. FH-HCMV interactions identified by yeast-two-hybrid analysis.

Table 2-2. FH-HCMV interactions identified by yeast-two-hybrid analysis			
#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD)	Function
1	RL10	$D^{\Delta\dagger}$	Viral entry into host cell, viral life cycle
2	UL7	$D^{\Delta\dagger}$	Immunomodulation of chemo/kytokines, mediates immune CEACAM-like glycoprotein
3	UL11	GD^{Δ}, $D^{\dagger}$	membrane glycoprotein interacts with CD45 and impairs TCR signaling and T-cell proliferation
4	UL12	GD^{Δ}, $D^{\dagger}$	unknown protein
5	UL13	$D^{\Delta\dagger}$	unknown protein (possibly viral transcription)
6	UL24	$D^{\Delta\dagger}$	Tegument protein
7	UL30	$GD^{\Delta\dagger}$	unknown protein
8	UL38	$GD^{\Delta\dagger}$	Plays a role in the inhibition of host apoptosis to facilitate efficient viral replication. Promotes stabilization and inactivation of host TP53 through interaction with host MDM2 (By similarity). Induces host mTORC1 activation by antagonizing the ability of host TSC1/2 to negatively regulate mTORC1. Thus, inhibits a growth regulatory pathway to facilitate viral replication. Prevents premature cell host cell death by blocking host ubiquitin-specific protease 24/USP24-mediated autophagic ferritin degradation in lysosomes thus maintaining lysosome integrity and cellular viability
9	UL43	$D^{\Delta\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription.
10	UL47	$GD^{\Delta\dagger}$	Plays an essential role in cytoplasmic secondary envelopment during viral egress. Interacts with the capsid via the large tegument protein/LTP and participates in its transport to the host trans-Golgi network (TGN) where secondary envelopment occurs. Modulates tegumentation and capsid accumulation at the viral assembly complex.
11	UL50	$E^{\Delta\dagger}$	Plays an essential role in virion nuclear egress, the first step of virion release from infected cell. Within the host nucleus, NEC1 interacts with the newly formed capsid through the vertexes and directs it to the inner nuclear membrane by associating with NEC2. Induces the budding of the capsid at the inner nuclear membrane as well as its envelopment into the perinuclear space. There, the NEC1/NEC2 complex promotes the fusion of the enveloped capsid with the outer nuclear membrane and the subsequent release of the viral capsid into the cytoplasm where it will reach the secondary budding sites in the host Golgi or trans-Golgi network.
12	UL52	$E^{\Delta\dagger}$	Plays a role in efficient localization of neo-synthesized capsids to nuclear replication compartments, thereby controlling cleavage and packaging of virus genomic DNA.
13	UL61	$D^{\dagger}$	unknown protein
14	UL63	Unknown	unknown protein
15	UL67	D^{Δ}	unknown protein
16	UL81	Unknown	unknown protein
17	UL82	$GD^{\Delta\dagger}$	Stimulates viral immediate-early (IE) transcription. Counteracts the DAXX-mediated repression of viral transcription. Displaces a DAXX-binding protein, ATRX, from nuclear domain 10 sites (ND10) shortly after infection. Increases the basal level of SUMOylated DAXX in infected cells. May also play a role in the progression of the host cell cycle through the G1 phase (By similarity).
18	UL88	$GD^{\Delta\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocitosis, viral life cycle, viral process
19	UL91	$E^{\Delta\dagger}$	Viral transcription
20	UL92	$E^{\Delta\dagger}$	Viral transcription

21	UL100	E$^{\Delta\dagger}$	Envelope glycoprotein important for virion assembly and egress. Plays a role in the correct incorporation of gH-gL into virion membrane. Directs the glycoprotein N (gN) to the host trans-Golgi network.
22	UL104	E$^{\Delta\dagger}$	portal protein
23	UL119	D$^{\Delta\dagger}$	Serves as a receptor for the Fc part of human IgG. May thus be involved in interfering with host Ig-mediated immune responses.
24	UL131	Unknown	unknown protein
25	UL132	GD$^{\Delta}$, D†	Viral entry into host cell, viral life cycle
26	US15	D$^{\Delta\dagger}$	unknown protein
Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth _in vitro_ as reported by:			
Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]			
$\dagger$ Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]			

MBL2 Interactions. For the pattern recognition molecule MBL2 which binds foreign carbohydrates on the surface of pathogens and activates the lectin pathway, we identified a total of 18 positive protein-protein interactions. As depicted in **Figure 2-7**, of the 18 positive protein-protein interactions, 28% (n=5) occurred in HCMV membrane proteins (UL9, UL41A, UL55, UL73, US16) or proteins involved in HCMV nuclear assembly/viral replication (UL27, UL52, UL84, UL89.2, UL96) respectively. Of the remaining positive protein-protein interactions: 22% (n=4) were in HCMV proteins of unknown function (UL123.3, UL131, US25, US33), 17% (n=3) occurred in tegument proteins (UL24, UL82, US22), and 5% (n=1) occurred in proteins involved in viral transcription (UL17). A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with MBL2 are listed in **Table 2-3.**
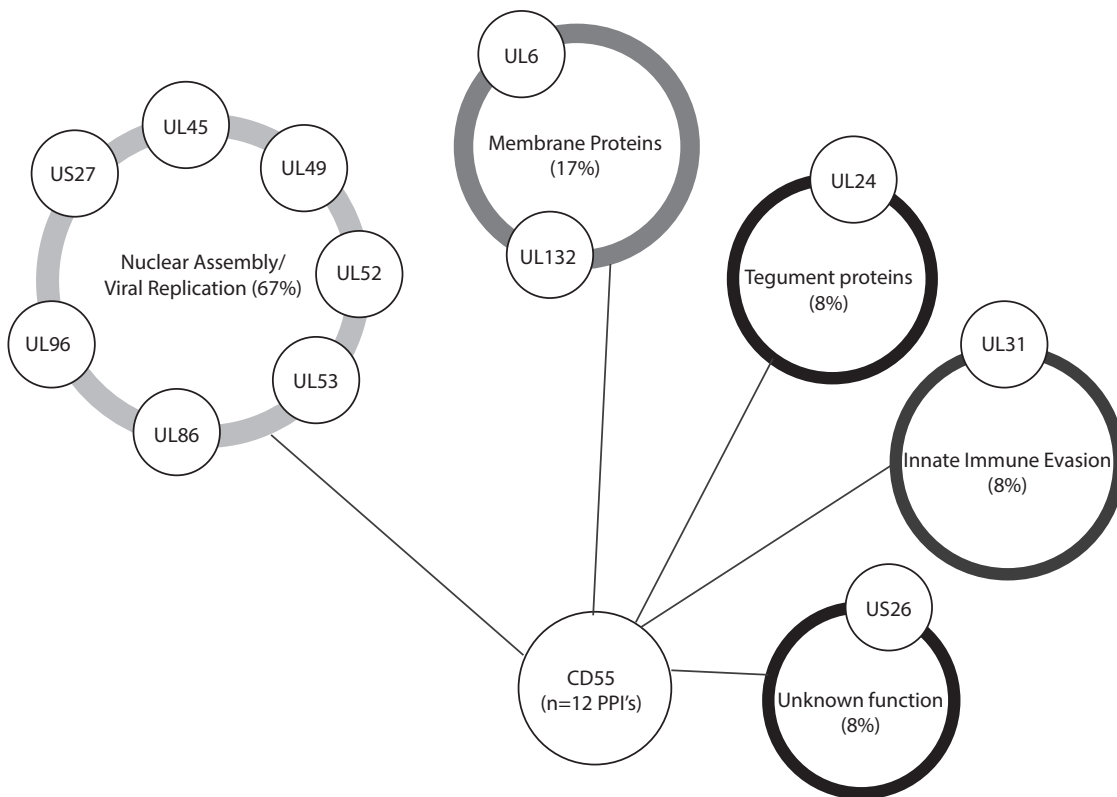

Figure 2-7. MBL2-HCMV interactions identified by yeast-two-hybrid analysis.

Table 2-3. MBL2-HCMV interactions identified by yeast-two-hybrid analysis			
#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD) Enhanced Growth (EG)	Function
1	UL9	EG$^\Delta$, D†	Putative membrane glycoprotein
2	UL17	D$^{\Delta\dagger}$	Unknown (viral transcription?)
3	UL24	D$^{\Delta\dagger}$	Virion protein, cell-tropism specific replication
4	UL27	D$^\Delta$, GD†	promotes cell cycle arrest in G0/G1
5	UL41A	D†	Virion protein
6	UL52	E$^{\Delta\dagger}$	Plays a role in efficient localization of neo-synthesized capsids to nuclear replication compartments, thereby controlling cleavage and packaging of virus genomic DNA.
7	UL55	E$^{\Delta\dagger}$	Envelope glycoprotein that forms spikes at the surface of virion envelope. Essential for the initial attachment to heparan sulfate moieties of the host cell surface proteoglycans. Involved in fusion of viral and cellular membranes leading to virus entry into the host cell. Following initial binding to its host receptors, membrane fusion is mediated by the fusion machinery composed at least of gB and the heterodimer gH/gL. May be involved in the fusion between the virion envelope and the outer nuclear membrane during virion egress.
8	UL73	E$^{\Delta\dagger}$	Envelope glycoprotein necessary for proper maturation of gM and modulation of its membrane fusion activity. Plays also a critical role in virion morphogenesis.
9	UL82	GD$^{\Delta\dagger}$	Stimulates viral immediate-early (IE) transcription. Counteracts the DAXX-mediated repression of viral transcription. Displaces a DAXX-binding protein, ATRX, from nuclear domain 10 sites (ND10) shortly after infection. Increases the basal level of SUMOylated DAXX in infected cells. May also play a role in the progression of the host cell cycle through the G1 phase (By similarity).
10	UL84	E$^{\Delta\dagger}$	Plays an essential role in viral DNA replication. May participate in the DNA replication initiation by interacting with the origin of lytic replication, oriLyt and subsequently recruiting other viral/cellular factors. Additionally, interacts with and shuttles IRS1 viral mRNA from the host nucleus to the cytoplasm (By similarity).
11	UL89.2	D$^\Delta$	unknown
12	UL96	E$^\Delta$, GD†	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription
13	UL123.3	GD$^{\Delta\dagger}$	Unknown
14	UL131	Unknown	Unknown
15	US16	D$^{\Delta\dagger}$	Viral transcription
16	US22	D$^{\Delta\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription
17	US25	D$^{\Delta\dagger}$	Unknown
18	US33	D$^{\Delta\dagger}$	Unknown
Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:			
Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]			
† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]			

MASP-1 Interactions. For the complement activating serine protease MASP-1 involved in the lectin pathway, we identified a total of 22 positive protein-protein interactions. As illustrated in **Figure 2-8**, of the 22 positive protein-protein interactions, 50% (n=11) occurred in HCMV proteins of unknown function (UL29, UL49, UL60, UL61, UL81, UL106, UL127, UL131, US25, US26, US32). Of the remaining positive protein-protein interactions: 18% (n=4) occurred in HCMV proteins involved in nuclear assembly/viral replication (UL52, UL53, UL86, UL96), 14% (n=3) were in tegument proteins (UL24, UL25, UL43), 9% (n=2) were in viral transcription, and 4.5% (n=1) were in membrane proteins (UL132), or in apoptotic inhibiting viral proteins (UL36.1) respectively. A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with MASP-1 are listed in **Table 2-4.**
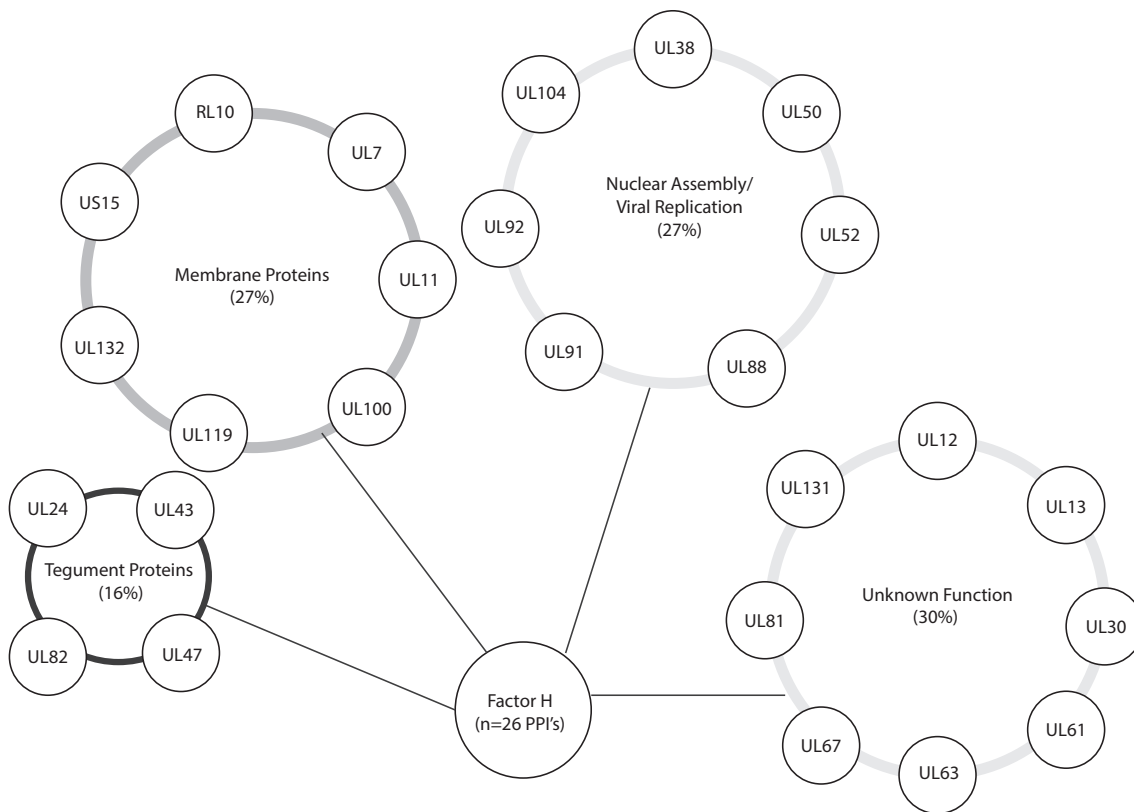

Figure 2-8. MASP1-HCMV interactions identified by yeast-two-hybrid analysis.

#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD)	Function
		Table 2-4. MASP1-HCMV interactions identified by yeast-two-hybrid analysis	
1	UL24	$D^{\Delta\dagger}$	Tegument protein
2	UL25	$D^{\Delta\dagger}$	Tegument phosphoprotein
3	UL29	$GD^{\Delta\dagger}$	Activation of immediate-early (IE) gene expression.
4	UL36.1	$D^{\Delta\dagger}$	Viral inhibitor of caspase-8-induced apoptosis
5	UL43	$D^{\Delta\dagger}$	Tegument protein
6	UL49	$E^{\Delta\dagger}$	Uncharacterized protein
7	UL52	$E^{\Delta\dagger}$	Packaging protein
8	UL53	$E^{\Delta\dagger}$	Nuclear egress protein 1
9	UL60	Unknown	Uncharacterized protein
10	UL61	$E^{\Delta\dagger}$	Uncharacterized protein
11	UL69	$GD^{\Delta\dagger}$	mRNA export factor
12	UL72	$GD^{\Delta\dagger}$	Deoxyuridine 5'-triphosphate nucleotidohydrolase
13	UL81	Unknown	Uncharacterized protein
14	UL86	$E^{\Delta\dagger}$	Major capsid protein
15	UL96	$E^{\Delta}, GD^{\dagger}$	Unknown
16	UL106	Unknown	Uncharacterized
17	UL127	D^{Δ}	Uncharacterized protein
18	UL131	Unknown	Uncharacterized protein
19	UL132	$GD^{\Delta}, D^{\dagger}$	Envelope glycoprotein
20	US25	$D^{\Delta\dagger}$	Uncharacterized protein
21	US26	$GD^{\Delta\dagger}$	Unknown
22	US32	$D^{\Delta\dagger}$	Unknown

Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:

Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]

† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]

Properdin Interactions. For the complement enhancing protein Properdin involved in the alternative pathway, we identified a total of 13 positive protein-protein interactions. As illustrated in **Figure 2-9**, of the 13 positive protein-protein interactions, 54% (n=7) occurred in HCMV proteins involved in nuclear assembly/viral replication (UL52, UL86, UL96, UL102, UL104, UL105, US27). Of the remaining positive protein-protein interactions, 31% (n=4) were viral membrane proteins (UL14, UL78, UL132, US14), and 8% (n=1) were in tegument (UL43) or proteins of unknown function (UL17) respectively. A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with Properdin are listed in **Table 2-5**.
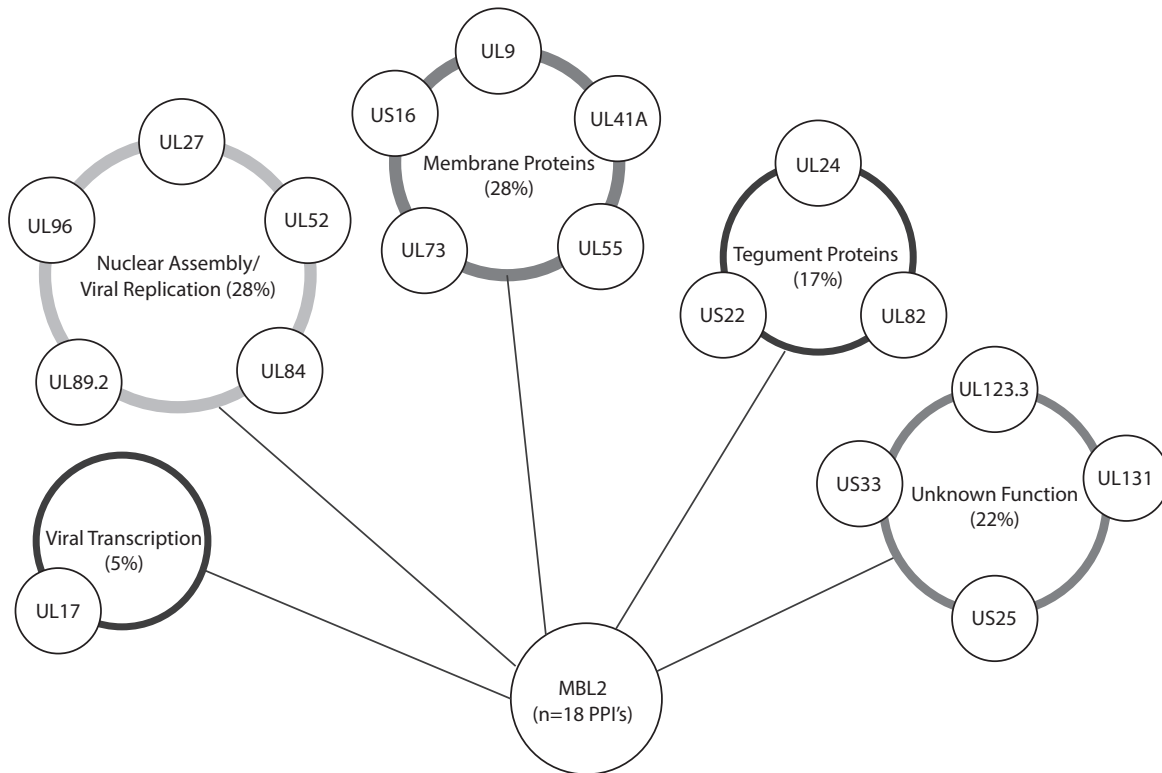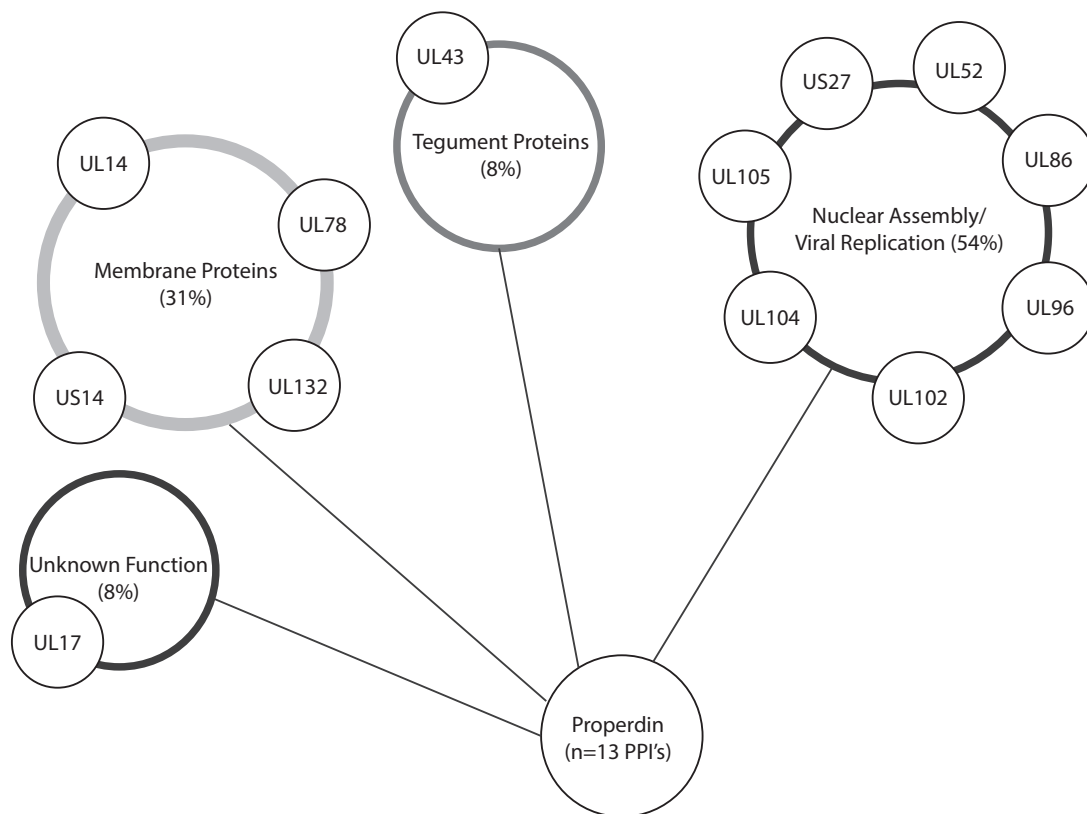

Figure 2-9. Properdin-HCMV interactions identified by yeast-two-hybrid analysis.

#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD)	Function
		Table 2-5. Properdin-HCMV interactions identified by yeast-two-hybrid analysis	
1	UL13	$D^{\Delta\dagger}$	unknown protein (possibly viral transcription)
2	UL14	GD^{Δ}, $D^{\dagger}$	membrane glycoprotein
3	UL43	$D^{\Delta\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription.
4	UL52	$E^{\Delta\dagger}$	Plays a role in efficient localization of neo-synthesized capsids to nuclear replication compartments, thereby controlling cleavage and packaging of virus genomic DNA.
5	UL78	$D^{\Delta\dagger}$	Acts as a cellular chemokine receptor and may modulate the host cell migration pattern which is the primary cellular event initiated by chemokine receptor activation.
6	UL86	$E^{\Delta\dagger}$	Self-assembles to form an icosahedral capsid with a T=16 symmetry, about 200 nm in diameter, and consisting of 150 hexons and 12 pentons (total of 162 capsomers). Hexons form the edges and faces of the capsid and are each composed of six MCP molecules. In contrast, one penton is found at each of the 12 vertices. Eleven of the pentons are MCP pentamers, while the last vertex is occupied by the portal complex. The capsid is surrounded by a layer of proteinaceous material designated the tegument which, in turn, is enclosed in an envelope of host cell-derived lipids containing virus-encoded glycoproteins.
7	UL96	E^{Δ}, $GD^{\dagger}$	Transport of viral material towards nucleus, viral entry into host cell, viral exocytosis, viral life cycle, viral process, viral transcription
8	UL102	$E^{\Delta\dagger}$	Component of the helicase/primase complex. Unwinds the DNA at the replication forks and generates single-stranded DNA for both leading and lagging strand synthesis. The primase synthesizes short RNA primers on the lagging strand that the polymerase presumably elongates using dNTPs. The primase-associated factor has no known catalytic activity in the complex and may serve to facilitate the formation of the replisome by directly interacting with the origin-binding protein and the polymerase.
9	UL104	$E^{\Delta\dagger}$	portal protein
10	UL105	$E^{\Delta\dagger}$	Component of the helicase/primase complex. Unwinds the DNA at the replication forks and generates single-stranded DNA for both leading and lagging strand synthesis. The primase synthesizes short RNA primers on the lagging strand that the polymerase elongates using dNTPs. Possesses helicase-like motifs and therefore may act as the helicase subunit of the complex.
11	UL132	GD^{Δ}, $D^{\dagger}$	Viral entry into host cell, viral life cycle
12	US14	$D^{\Delta\dagger}$	Viral Transcription
13	US27	$D^{\Delta\dagger}$	Plays an important role in spread of HCMV via the extracellular route. As a G-protein-coupled receptor (vGPCR), may activate signaling pathways important for virion assembly or egress processes.

Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:

Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]

† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]

C1qBP Interactions. For the complement inhibiting protein C1qBP involved in the classical pathway, we identified a total of 31 positive protein-protein interactions. As illustrated in **Figure 2-10**, of the 31 positive protein-protein interactions, 22% (n=7) occurred in HCMV proteins of unknown function (RL5, UL15A, UL20A, UL28, UL41, UL108, UL124) or involved in nuclear assembly/viral replication (UL52, UL53, UL56, UL84, UL85, UL86, UL102) respectively. Of the remaining positive protein-protein interactions, 19% (n=6) were tegument proteins (UL23, UL24, UL25, UL49, UL71, UL82), 16% (n=5) were proteins involved in immune evasion, 6% (n=2) were cell tropism factors (UL78, UL132) or involved in viral transcription (UL44, UL92), and the remaining 3% (n=1) were membrane proteins (UL120) or proteins involved in cell cycle regulation (UL38). A comprehensive list of the known functions and essentiality of HCMV-encoded proteins that interact with C1qBP are listed in **Table 2-6**.
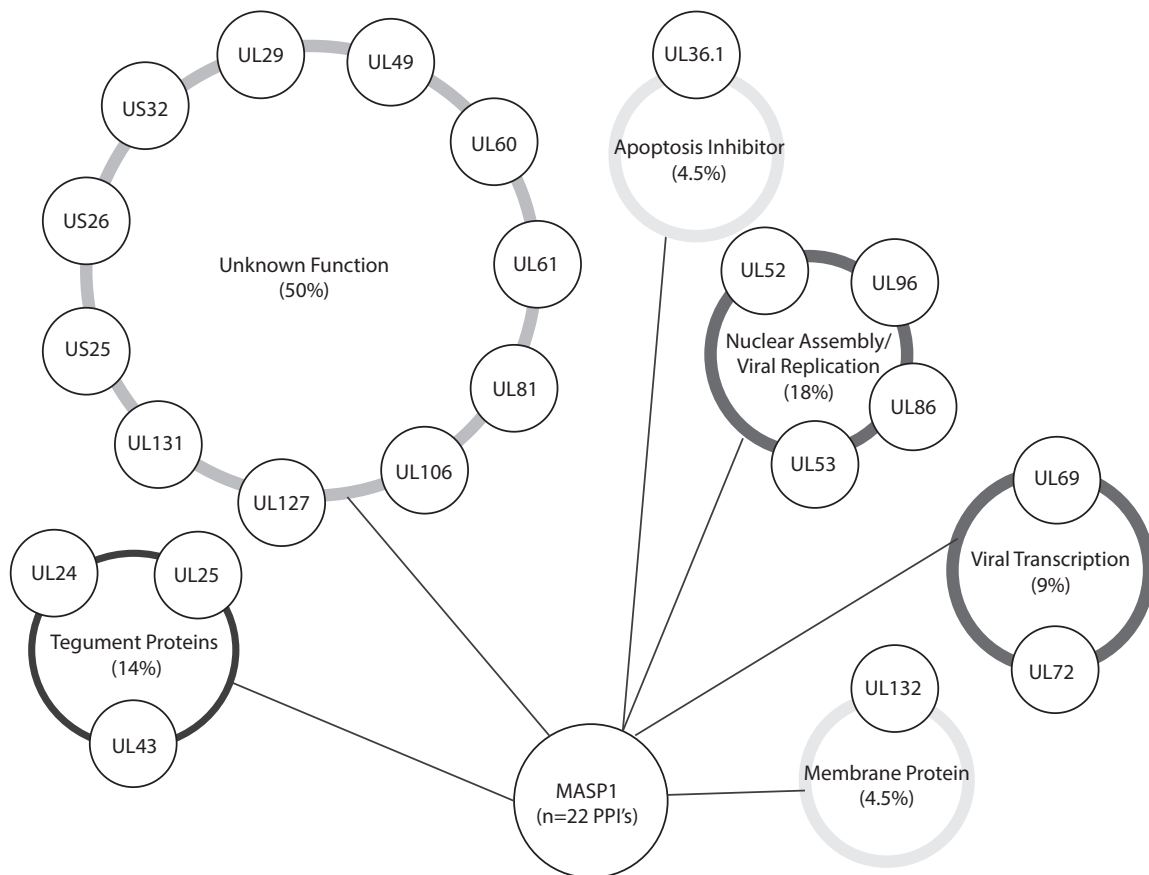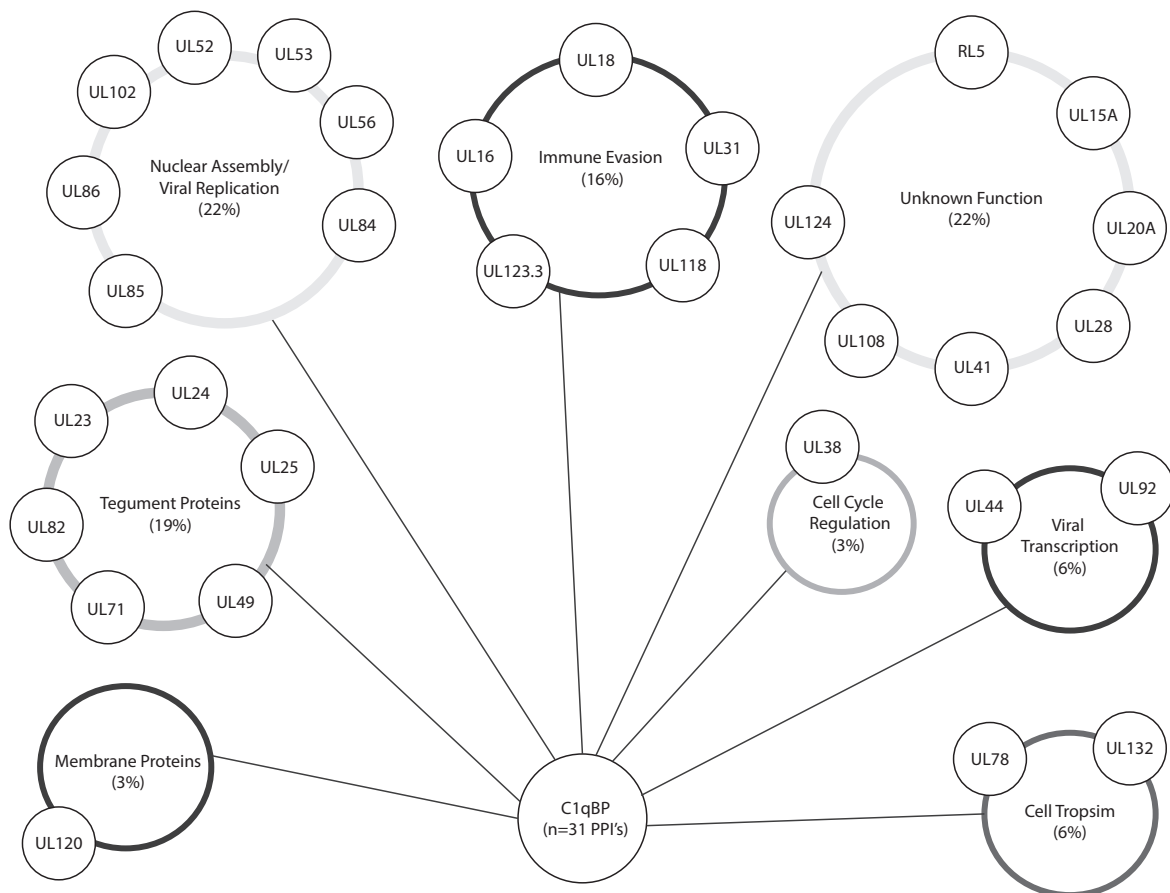

Figure 2-10. C1qBP-HCMV interactions identified by yeast-two-hybrid analysis.

#	HCMV Gene	Essential (E), Dispensable (D), Growth Defect (GD), Enhanced Growth (EG)	Function
		Table 2-6. C1qBP-HCMV interactions identified by yeast-two-hybrid analysis	
1	RL5	Unknown	Unknown
2	UL15A	$D^{\dagger}$	Unknown
3	UL16	$D^{\Delta\dagger}$	inhibits NK cell signaling
4	UL18	$D^{\Delta\dagger}$	MHC1-homolog, NK signaling/modulation
5	UL20A	Unknown	Unknown
6	UL23	$EG^{\Delta}, D^{\dagger}$	Tegument protein
7	UL24	$D^{\Delta\dagger}$	Tegument protein
8	UL25	$D^{\Delta\dagger}$	Tegument phosphoprotein
9	UL28	$GD^{\Delta\dagger}$	Cell cycle regulation, modulate apoptosis
10	UL31	$GD^{\Delta}, D^{\dagger}$	inhibits CGAS mediated innate immunity
11	UL38	$GD^{\Delta\dagger}$	Plays a role in the inhibition of host apoptosis to facilitate efficient viral replication. Promotes stabilization and inactivation of host TP53 through interaction with host MDM2 (By similarity). Induces host mTORC1 activation by antagonizing the ability of host TSC1/2 to negatively regulate mTORC1. Thus, inhibits a growth regulatory pathway to facilitate viral replication. Prevents premature cell host cell death by blocking host ubiquitin-specific protease 24/USP24-mediated autophagic ferritin degradation in lysosomes thus maintaining lysosome integrity and cellular viability
12	UL41	Unknown	Unknown
13	UL44	$E^{\Delta\dagger}$	Accessory subunit of the DNA polymerase that acts to increase the processivity of polymerization.
14	UL49	$E^{\Delta\dagger}$	Viral DNA replication
15	UL52	$E^{\Delta\dagger}$	Plays a role in efficient localization of neo-synthesized capsids to nuclear replication compartments, thereby controlling cleavage and packaging of virus genomic DNA.
16	UL53	$E^{\Delta\dagger}$	Plays an essential role in virion nuclear egress, the first step of virion release from infected cell. Within the host nucleus, NEC1 interacts with the newly formed capsid through the vertexes and directs it to the inner nuclear membrane by associating with NEC2. Induces the budding of the capsid at the inner nuclear membrane as well as its envelopment into the perinuclear space. There, the NEC1/NEC2 complex promotes the fusion of the enveloped capsid with the outer nuclear membrane and the subsequent release of the viral capsid into the cytoplasm where it will reach the secondary budding sites in the host Golgi or trans-Golgi network.
17	UL56	$E^{\Delta\dagger}$	Component of the molecular motor that translocates viral genomic DNA in empty capsid during DNA packaging. Forms a tripartite terminase complex together with TRM2 and TRM3 in the host cytoplasm. Once the complex reaches the host nucleus, it interacts with the capsid portal vertex. This portal forms a ring in which genomic DNA is translocated into the capsid. TRM1 carries an endonuclease activity that plays an important role for the cleavage of concatemeric viral DNA into unit length genomes.
18	UL71	$E^{\Delta}, GD^{\dagger}$	Plays several roles during the time course of infection, including egress of virus particles from the perinuclear space and secondary envelopment of cytoplasmic capsids that bud into specific trans-Golgi network (TGN)-derived membranes.
19	UL78	$D^{\Delta\dagger}$	Acts as a cellular chemokine receptor and may modulate the host cell migration pattern which is the primary cellular event initiated by chemokine receptor activation.

#	Gene		Function
20	UL82	GD[Δ,†]	Stimulates viral immediate-early (IE) transcription. Counteracts the DAXX-mediated repression of viral transcription. Displaces a DAXX-binding protein, ATRX, from nuclear domain 10 sites (ND10) shortly after infection. Increases the basal level of SUMOylated DAXX in infected cells. May also play a role in the progression of the host cell cycle through the G1 phase (By similarity).
21	UL84	E[Δ,†]	Plays an essential role in viral DNA replication. May participate in the DNA replication initiation by interacting with the origin of lytic replication, oriLyt and subsequently recruiting other viral/cellular factors. Additionally, interacts with and shuttles IRS1 viral mRNA from the host nucleus to the cytoplasm (By similarity).
22	UL85	E[Δ,†]	Structural component of the T=16 icosahedral capsid. The capsid is composed of pentamers and hexamers of major capsid protein/MCP, which are linked together by heterotrimers called triplexes. These triplexes are formed by a single molecule of triplex protein 1/TRX1 and two copies of triplex protein 2/TRX2. Additionally, TRX1 is required for efficient transport of TRX2 to the nucleus, which is the site of capsid assembly.
23	UL86	E[Δ,†]	Self-assembles to form an icosahedral capsid with a T=16 symmetry, about 200 nm in diameter, and consisting of 150 hexons and 12 pentons (total of 162 capsomers). Hexons form the edges and faces of the capsid and are each composed of six MCP molecules. In contrast, one penton is found at each of the 12 vertices. Eleven of the pentons are MCP pentamers, while the last vertex is occupied by the portal complex. The capsid is surrounded by a layer of proteinaceous material designated the tegument which, in turn, is enclosed in an envelope of host cell-derived lipids containing virus-encoded glycoproteins.
24	UL92	E[Δ,†]	Viral transcription
25	UL102	E[Δ,†]	Component of the helicase/primase complex. Unwinds the DNA at the replication forks and generates single-stranded DNA for both leading and lagging strand synthesis. The primase synthesizes short RNA primers on the lagging strand that the polymerase presumably elongates using dNTPs. The primase-associated factor has no known catalytic activity in the complex and may serve to facilitate the formation of the replisome by directly interacting with the origin-binding protein and the polymerase.
26	UL108	GD[Δ,†]	Latency
27	UL118	D[Δ,†]	Serves as a receptor for the Fc part of human IgG. May thus be involved in interfering with host Ig-mediated immune responses.
28	UL120	D[†]	Viral transcription
29	UL123.3	GD[Δ,†]	Unknown
30	UL124	GD[Δ], D[†]	Viral transcription
31	UL132	GD[Δ], D[†]	Viral entry into host cell, viral life cycle

Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:

Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]

† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]

Interactome map of Complement-HCMV interactions by functional grouping. We previously identified a total of 121 positive protein-protein interactions testing 6 human complement proteins against 167 proteins encoded by the HCMV genome. To understand the nature of these positive protein-protein interactions identified by yeast-two-hybrid assay within a broader context, we analyzed the interactions by concentric functional grouping as illustrated in **Figure 2-11**.
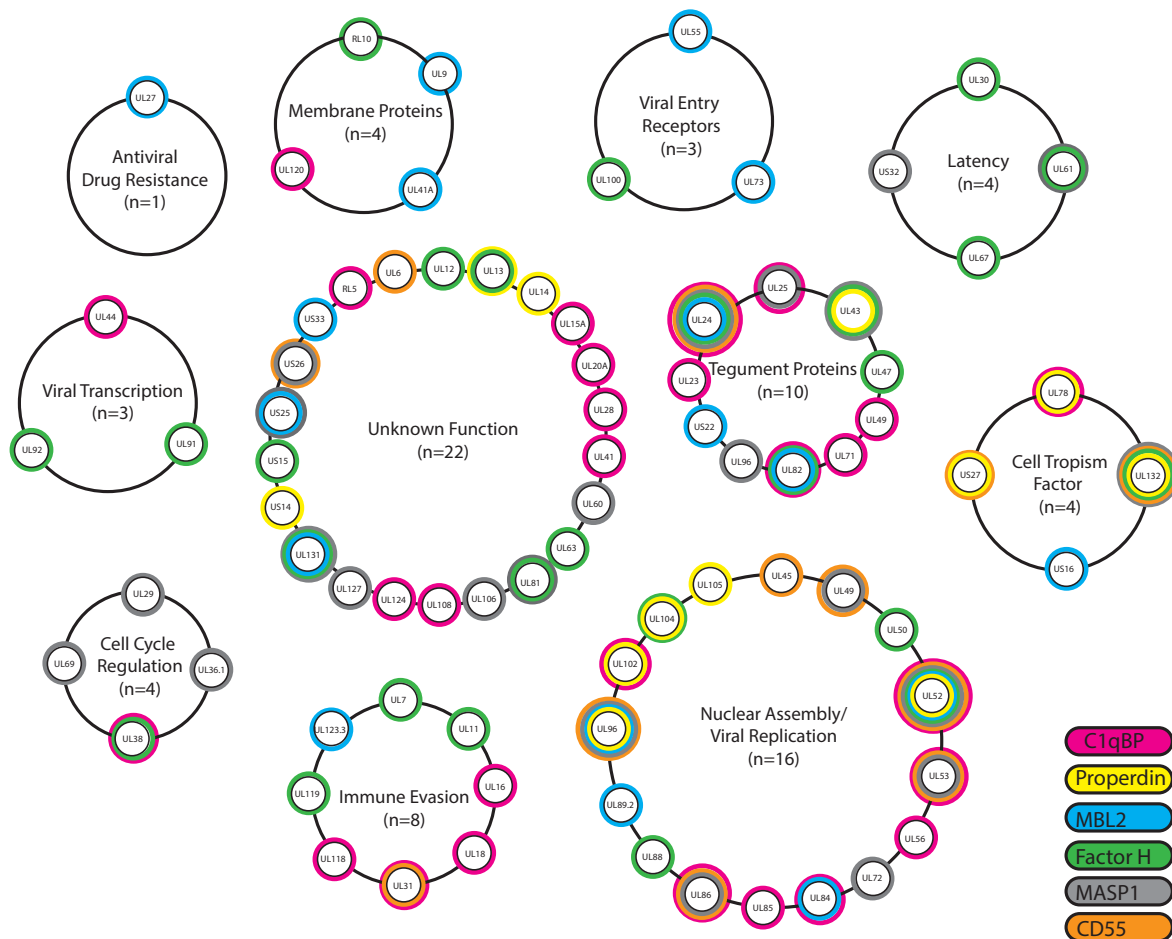
Figure 2-11. Interactome map of Complement-HCMV interactions by functional group. Interactome map of 121 Complement-HCMV interactions identified by yeast-two hybrid in complement proteins: C1qBP (magenta), Properdin (yellow), MBL2 (blue), Factor H (green), MASP-1(gray), and CD55 (orange). Interacting pairs were grouped by functional grouping according to known HCMV protein functions. HCMV proteins interacted with one or multiple complement proteins as depicted by one or more concentric rings respectively.

53

Proteins of Unknown function. The largest group of HCMV-encoded complement-interacting proteins occurred in viral proteins with no known function. This result is unsurprising, as HCMV has one of the largest genomes of any known human viral pathogen and the precise biological functions of most HCMV-encoded ORF's remain unknown. Of the 22 HCMV-encoded complement-interacting proteins, none had significant homology to other known proteins in which to infer possible function, and none were conserved across other human herpesviruses. Within this group, all HCMV proteins interacted with one or more of the six complement proteins tested, though proteins that interacted with C1qBP (RL5, UL15A, UL20A, UL28, UL41, UL108, UL124) made up the largest proportion of interactions. In contrast, HCMV proteins that interacted with CD55 (UL6) made up the smallest proportion of interactions.

Nuclear assembly/Viral replication and transcription. As a DNA virus HCMV undergoes viral replication in the nucleus, and this complex processes requires DNA replication, viral transcription, and nuclear assembly, all highly coordinated and mediated by viral encoded proteins. To our surprise, the second largest group of HCMV-encoded complement-interacting proteins occurred in proteins involved in nuclear assembly and viral replication. Of the 16 HCMV-encoded proteins 12 were conserved in other herpesviruses, and included the following HCMV genes: UL45, UL50, UL52, UL53, UL56, UL72, UL85, UL86, UL88, UL102, UL104, and UL105. Within this group of proteins, all 16 interacted with one or more of the six complement proteins tested. There was no strong association between the number of interactions identified and a specific complement protein tested. Additionally, two positive interactions we identified (UL44-C1qBP and UL84-C1qBP) have previously been reported using tandem affinity purification mass spectrometry (TAP-MS), and these interaction were nevertheless reproduced in our yeast-two-hybrid assay [79].

Tegument proteins. The tegument layer of HCMV virions is a proteinaceous matrix that lies between the genome capsid and the viral envelope. Initially, tegument proteins were thought to be nothing more than structural proteins, however an expanding body of research has demonstrated that tegument proteins are diverse and polyfunctional and have been reported to be involved in processes such as controlling viral entry, gene expression, and even immune evasion. In our study, the third largest group of HCMV-encoded complement-interacting proteins occurred in tegument proteins. Of the 10 HCMV-encoded proteins, all interacted with one or more of the six complement proteins tested, though proteins that interacted with C1qBP (UL23, UL24, UL25, UL49, UL71, UL81) made up the largest proportion of interactions. In contrast, HCMV proteins that interacted with Properdin (UL43) made up the smallest proportion of interactions.

Immune evasion. HCMV encodes an unparalleled number of viral proteins that have been reported to interfere with both innate and adaptive immunity. Unsurprisingly, the fourth largest group of HCMV-encoded complement-interacting proteins occurred in proteins involved in viral immune evasion. Of the 8 HCMV-encoded proteins known to be involved in immune evasion, none interacted with Properdin or MASP1, but they did interact with the other complement proteins tested in our yeast-two-hybrid assay. HCMV-encoded proteins that interacted with C1qBP (UL16, UL18, UL31, UL118) made up the

largest proportion of interactions, followed by those that interacted with FH (UL7, UL11, UL119). MBL2 interacted with a single HCMV gene (UL123.3). Of interest, some of the HCMV genes (UL118, UL119) that interacted with complement proteins encode Fcɣɣ receptor homologs that bind to the Fc regions of host antibody [80]. Binding of the Fc region of an antibody by a viral protein can inhibit antibody effector functions [81] such as antibody dependent complement mediated neutralization that is a hallmark of the classical pathway of the complement system.

Cell cycle regulation. HCMV virions modulate the host cell cycle during infection to create an environment and manipulate host cell machinery that is optimal for viral gene expression, DNA replication, and production of progeny virions. Of the four HCMV genes involved in cell cycle regulation interacted with complement proteins in our yeast two hybrid assay, three (UL29, UL36.1, and UL69) interacted with MASP1.

Latency. As with other members of the Herpesviridae family, HCMV has the hallmark ability to undergo viral latency. In humans, HCMV establishes latency in hematopoietic progenitor cells of the myeloid lineage and this latent state allows for life-long persistent infection. During the latent state, the HCMV viral genome is silenced with the exception of a few latency associated viral genes. We identified four latency associated HCMV genes that interacted with either FH (UL30, UL61, UL67), or MASP1 (US32).

Complement-HCMV protein interaction validation by co-immunoprecipitation. A subset of positive Complement-HCMV protein-protein interactions identified by yeast-two-hybrid were retested by co-immunoprecipitation (co-IP) assay to validate the interactions in human cells. We tested a total of 13 Complement-HCMV interactions which included: 5 interactions with C1qBP (C1qBP-UL24, C1qBP-UL25, C1qBP-UL44, C1qBP-UL82, C1qBP-UL84), 3 interactions with MBL2 (MBL2-UL24, MBL2-UL82, MBL2-UL84), 2 interactions with CD55 (CD55-UL24, CD55-UL45), 2 interactions with MASP1 (MASP1-UL25, MASP1-UL69), and 1 interaction with Properdin (CFP-UL105). Additionally, we tested UL35(Myc)-UL82(HA) and UL35(Myc)-UL83(HA), which served as positive and negative protein-protein interaction controls respectively.

In these experiments the complement proteins or HCMV proteins were expressed as a fusion protein with a carboxyl terminal c-Myc or HA epitope tag, respectively. HeLa cells were co-transfected with both complement and HCMV expression vectors as described in the materials and methods and harvested at 72 hours post-transfection. To detect expression of complement and HCMV proteins we separated raw cell lysates by SDS-PAGE and transferred them to nitrocellulose membranes before detecting proteins with anti-HA and anti-Myc monoclonal antibodies.

As illustrated in **Figure 2-12**, complement proteins: Properdin, CD55, MASP1, MBL2, and C1qBP were expressed in cell lysates as indicated by their reactivity using an anti-Myc monoclonal antibody. In lane 1, two bands of ~35-45kD were observed consistent with the predicted molecular weight of properdin (49kD). In lanes 2-3, a major ~45kD band was observed consistent with the predicted molecular weight of CD55 (43kD). Notably, even though the predicted molecular weight of CD55 is 43kD, the observed molecular weight of mature CD55 has been reported to vary between 50-100kD depending on the cell type used for expression [82] and may be due to alternative splicing or different glycosylation patterns [83, 84]. In lanes 4-5, a major band of ~70kD was observed consistent with the expected molecular weight of MASP1 (79kD). In lanes 6-8, a major band of ~40kD was observed consistent with the expected molecular weight of MBL2 (37kD). In lanes 9-13, a major band of ~45kD was observed consistent with the predicted molecular weight of C1qBP (35kD). In lanes 14-15, a major band of ~70kD was observed for control Myc-tagged UL35 consistent with the expected molecular weight of 72kD.
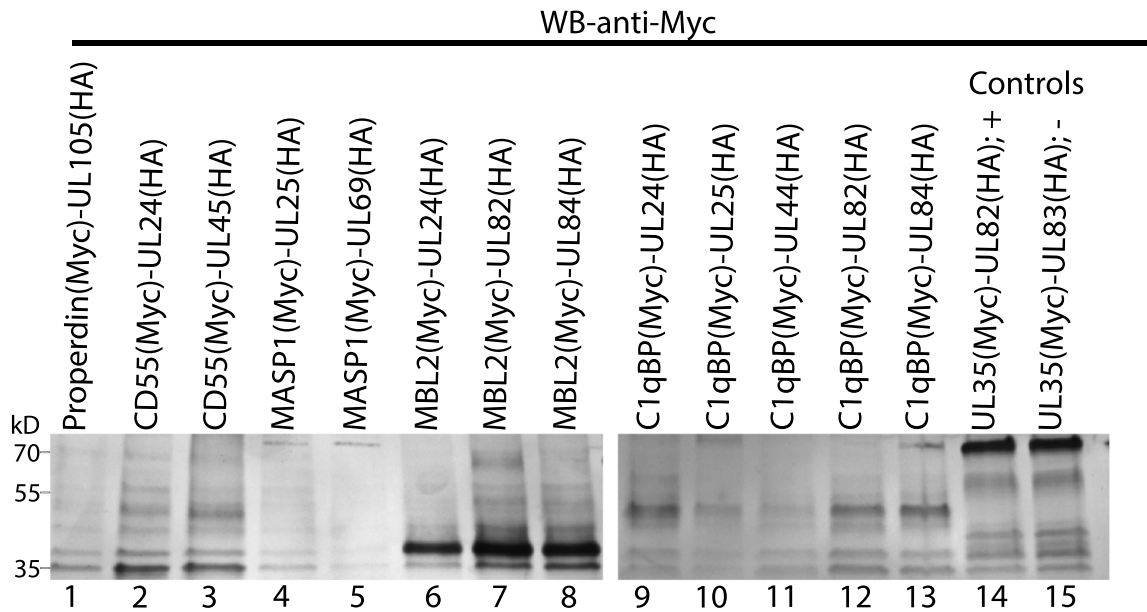
Figure 2-12. Expression of Myc-tagged Complement proteins in HeLa cell lysates. Western blot of Myc-tagged human complement proteins expressed in HeLa cell lysates detected with an anti-Myc monoclonal antibody. Properdin expression in lane 1, CD55 expression in lanes 2-3, MASP1 expression in lanes 4-5, MBL2 expression in lanes 6-8, C1qBP expression in lanes 9-13, and control UL35 expression in lanes 14-15. The predicted molecular weights of Properdin, CD55, MASP1, MBL2, C1qBP and UL35 are: 49kD, 43-100kD, 79kD, 37kD, 35kD, and 72kD, respectively.

The same cell lysates were tested for expression of HA-tagged HCMV proteins using an anti-HA monoclonal antibody as depicted in **Figure 2-13**. In lane 1 we observed a band of ~100kD consistent with the predicted molecular weight of UL105 (106kD). In lanes 2, 6, and 9 we observed major bands of ~40kD consistent with the predicted molecular weight of UL24 (40kD). In lane 3 we detected a major band of ~100kD consistent with the predicted molecular weight of UL45 (101kD). In lanes 4 and 10 we observed bands of ~75-90kD consistent with the predicted molecular weight of UL25 (73kD). In lane 5 we detected a band of ~75kD consistent with the predicted molecular weight of UL69 (83kD). In lanes 7, 12, and 14 we observed a band of ~70kD consistent with the predicted molecular weight of UL82 (73kD). In lanes 8 and 13 we detected a band of ~70kD consistent with the predicted molecular weight of UL84 (65kD). In lane 11 we observed a major band of ~55kD consistent with the predicted molecular weight of UL44 (46kD). In lane 15 we detected a major band of ~70kD consistent with the predicted molecular weight of UL83 (63kD).
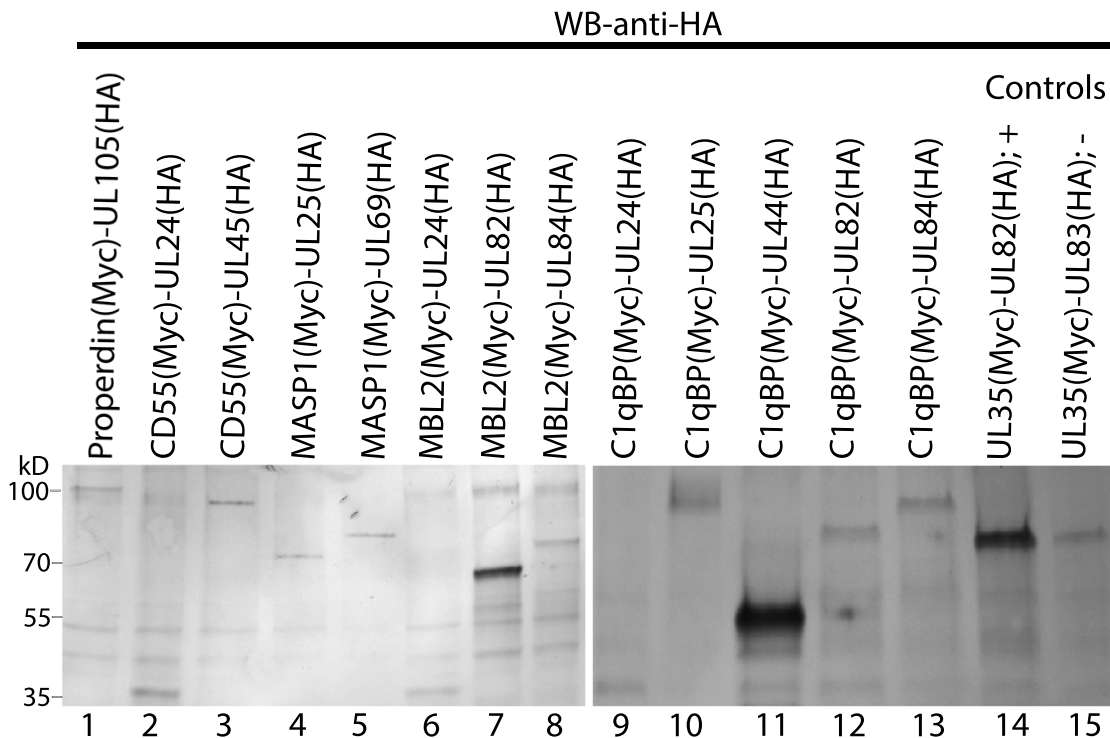
Figure 2-13. Expression of HA-tagged HCMV proteins in HeLa cell lysates.
Western blot of HA-tagged HCMV proteins expressed in HeLa cell lysates detected with an anti-HA monoclonal antibody. UL24 (lanes 2, 6, 9), UL25 (lanes 4 and 10), UL44 (lane 11), UL82 (lanes 7, 12, 14), UL84 (lanes 8, 13), UL105 (lane 1), UL45 (lane 2), UL83 (lane 15), UL69 (lane 5) with calculated molecular weights of 40kD, 73kD, 46kD, 62kD, 65kD, 106kD, 101kD, 63kD, 83kD respectively

57

After verifying expression of both Myc-tagged Complement and HA-tagged HCMV proteins in our co-transfected cell lysates, we co-immunoprecipitated the cell lysates using anti-Myc conjugated magnetic beads and tested the samples by Western blot using an anti-HA monoclonal antibody to detect bound co-immunoprecipitated protein complexes. As illustrated in **Figure 2-14**, there was no interaction identified between Properdin and UL105 as indicated by the absent protein band in lane 1. Similarly, for CD55 both UL24 and UL45 were negative for protein-protein interactions as indicated by the absence of protein bands in lanes 2 and 3 respectively. For MASP1, UL25 was found to interact as illustrated by the protein band in lane 3 of ~80kD, but UL69 did not interact as illustrated by the absent protein band in lane 4. For MBL2, UL24, UL82, and UL84 were positive for interaction as illustrated by the bands in lane 6 (~40kD), lane 7 (~70kD) and lane 8 (~80kD) which were consistent with the expected protein molecular weights of UL24, UL82, and UL84 respectively. For C1qBP, UL24 did not interact as evident by the absence of a protein band in lane 9. In contrast, C1qBP was found to interact with UL25, UL44, UL82, and UL84 as evident by the protein bands of approximately 100kD, 55kD, 70kD, and 70kD in lanes 10, 11, 12 and 13 respectively. To have confidence in the reliability of our methods to detect protein-protein interactions, we tested UL35(Myc)-UL82(HA) and UL35(Myc)-UL83(HA) as positive and negative control protein interactions respectively. UL82 was found to interact with UL35, as indicated by the ~70kD protein band in lane 14, and this positive interaction is consistent with previous studies [69] [85]. In contrast, UL83 did not interact with UL35 as evident by the absence of a protein band in lane 15, a result which is also consist with previous studies [69].
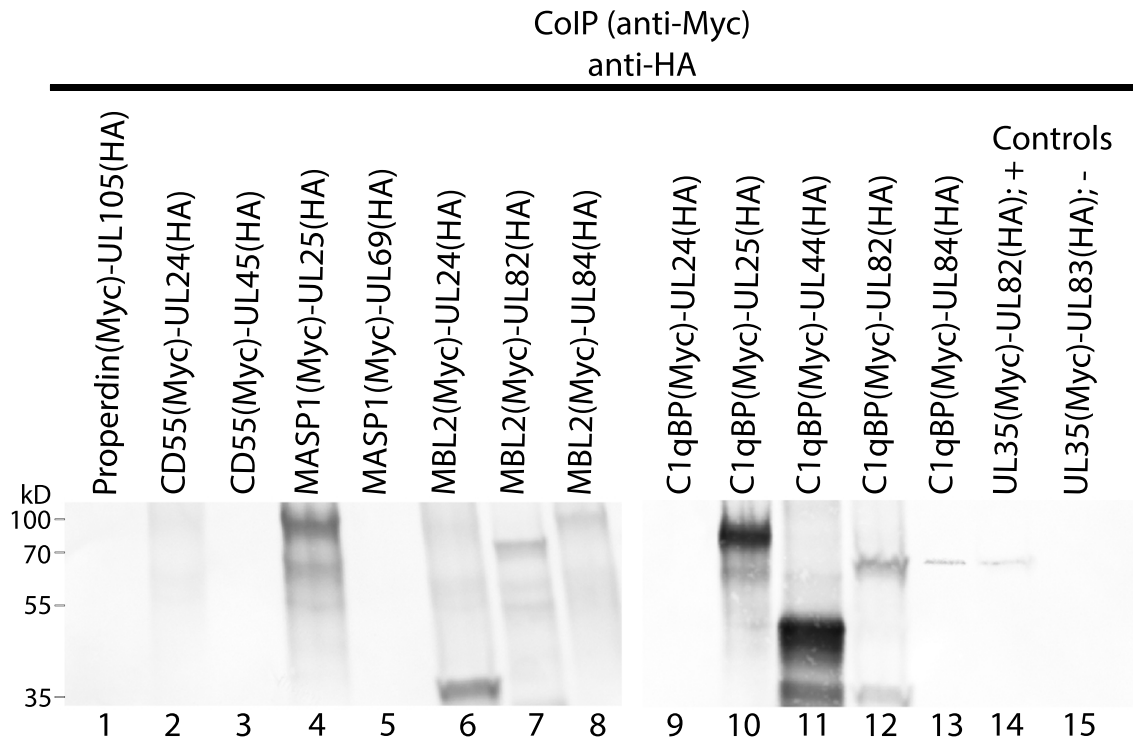
Figure 2-14. Validation of Complement-HCMV interactions by co-immunoprecipitation. Western blot of co-transfected HeLa cell lysates expressing myc-tagged complement proteins and HA-tagged HCMV proteins co-immunoprecipitated using anti-myc magnetic beads. Detection of bound HCMV proteins was detected using an anti-HA monoclonal antibody. UL24 (lane 6), UL25 (lanes 4, 10), UL44 (lane 11), UL82 (lanes 7, 12, 13), UL84 (lanes 8, 13), with calculated molecular weights of 40kD, 73kD, 46kD, 62kD, 65kD, respectively

A summary of all the Complement-HCMV protein combinations tested and the respective known functions of each of the HCMV proteins are listed in **Table 2-7**. In total we tested 13 protein-protein interactions, representing a combination of 5 different complement proteins and 9 HCMV proteins. Of the 13 protein-protein interactions that were identified by yeast-two-hybrid screening and validated by co-immunoprecipitation, 8 were positive by co-immunoprecipitation which represents a 61.5% positivity rate. Importantly, our positivity rate is within the scope, albeit slightly higher than other large scale protein-protein interactome studies of other human herpesviruses tested against the human proteome [86].

Table 2-7. Summary of HCMV-Complement interactions tested by yeast-two-hybrid and Co-IP						
Complement protein	HCMV protein	YTH	Co-IP	Status	Function	Reference for interactions
Properdin	UL105	+	-		Component of the viral helicase/primase complex involved in DNA replication	
CD55	UL24	+	-		Tegument protein, contributes to innate immune evasion by	
	UL45	+	-		Ribonucleotide-diphosphate reductase large subunit like protein. Accumulates at late stages of infection.	
MASP1	UL25	+	+	*New*	Tegument protein and major target antigen of the anti-HCMV antibody response	This study
	UL69	+	-		Involved in the regulation of host cell cycle progression, regulation of viral gene expression and nuclear export of some viral RNA's. Acts as a post-transcriptional transactivator that mediates nuclear export of unspliced RNA's.	
MBL2	UL24	+	+	*New*	Tegument protein; involved in cellular innate immune evasion by sequestering host antiviral restriction factor SAMHD1	This study
	UL82	+	+	*New*	Stimulates viral IE transcription and suppresses innate immunity by targeting the cGAS-STING pathway	This study
	UL84	+	+	*New*	Essential role in viral DNA replication initiation by interacting with the oriLyt and recruiting viral and cellular replication factors	This study
C1qBP	UL24	+	-		Tegument protein; involved in cellular innate immune evasion by sequestering host antiviral restriction factor SAMHD1	
	UL25	+	+	*New*	Tegument protein and major target antigen of the anti-HCMV antibody response.	This study
	UL44	+	+	Published	Accessory subunit of the DNA polymerase that plays an essential role in viral DNA replication and increases processivity of polymerization.	[79]
	UL82	+	+	*New*	Stimulates viral IE transcription and suppresses innate immunity by targeting the cGAS-STING pathway	This study
	UL84	+	+	Published	Essential role in viral DNA replication initiation by interacting with the oriLyt and recruiting viral and cellular replication factors	[79, 87]

DISCUSSION

Summary. The human complement system is an important first line of defense against invading microorganisms, and is an evolutionarily ancient form of innate immunity with homologs identified in invertebrates [88], insects [89], and even lower organisms [88, 90]. Several medically important obligate human viral pathogens bind to complement regulating proteins [91] [92-94], encode homologs of complement regulating proteins [95], incorporate host complement inhibitors into the viral envelope [96, 97], or use complement receptors to enter host cells [98].

The role of the complement system in immunity to HCMV has thus far been neglected and remains poorly understood. Unlike other human herpesviruses, no studies to date have reported the existence of HCMV-encoded proteins with complement regulating functions. Yet, the observation that HCMV infected cells exhibit altered expression of various complement proteins, and the resistance of HCMV virions to complement-mediated neutralization in the absence of antibody suggests that HCMV has evolved mechanisms to manipulate and subvert this important branch of innate immunity.

In the present study, we employed a largescale yeast-two-hybrid screen, testing 6 human complement proteins (CD55, MBL2, MASP1, C1qBP, FH, and CFP) against an HCMV genomic library encoding 167 viral proteins in yeast. Out of 1,002 possible HCMV-Complement interactions tested, we identified 121 (8.2%) positive protein interactions. The interactions we identified occurred in complement inhibiting molecules (CD55, FH), pattern recognition molecules (MBL2), pattern recognition receptors (C1qBP), activating enzymes (MASP1), and most surprisingly proteins that positively regulate the complement system (CFP). These results suggest that HCMV not only interacts with multiple proteins in the complement system but may potentially modulate all three pathways at different steps. In doing so, HCMV may disrupting the normal homeostasis of complement synthesis to promote viral fitness by generating a microenvironment that facilitates infection, persistence, and pathogenesis during HCMV viral replication. Currently, our study is the most comprehensive undertaking to date investigating protein interactions between components of the human complement system and HCMV-encoded proteins. Our study provides a framework to further investigate the functional significance of protein interactions between HCMV and the human complement system which may underlie novel forms of innate immune evasion by HCMV, and potentially inform the development of vaccines and therapeutics for the prevention and treatment of HCMV disease.

Yeast-two-hybrid: Strengths & Limitations. The HCMV genome encodes 175 canonical proteins, though recent studies have suggested the coding potential could be significantly greater considering the complexity of HCMV transcription and translation. With one of the largest genomes of known human viruses, HCMV has evolved eloquent mechanisms to subvert both innate and adaptive host immunity and establish active and latent infection.

The primary aim of this study was to screen proteins encoded by the HCMV genome against several complement proteins to identify interactions that could advance our understanding of how HCMV resists complement mediated neutralization. Admittedly, embarking on the present study was both ambitious and risky given significant gaps in knowledge and relatively few published studies concerning the role of the complement system in HCMV immunity. Investigating the Complement-HCMV interactome could have been achieved using various alternative methods though we opted to employ a yeast-two-hybrid approach, which is both a major strength and limitation of our study.

Similar large scale yeast-two-hybrid analyses such as the one employed in our study have been applied to rigorous interactome investigations of protein-protein interactions in many complex organisms including: *Drosophila melanogaster* [99], *Caenorhabditis elegans* [100], *Saccharomyces cerevisea* [101], *Plasmodium falciparum* [102], *Helicobacter pylori* [103], and most recently SARS-CoV-2 [104]. One important caveat of the yeast-two-hybrid system is that it is especially vulnerable to both false-positives and false-negatives, therefore the assay conditions must be carefully controlled, optimized for sensitivity and specificity, and performed in replicate.

Several factors can influence the rate of false-positives and false-negative interactions including: (1) Improper protein folding, (2) Post-translational protein modifications, (3) Protein size, (4) Conditions within the yeast nucleus where the two proteins of interest are interacting, (5) Reproducibility, and (6) Location of the activating or binding domains fused to the protein of interested in the yeast-two-hybrid system. Due to the preceding, all protein-protein interactions identified herein must be cross-validated by other protein interaction assays such as protein co-immunoprecipitation to increase the confidence of observed reactions. Notwithstanding some inherent weaknesses, the yeast-two-hybrid assay does offer many strengths compared to other protein-protein interaction assays due to its relatively low reagent cost, scalability, adjustable sensitivity, and ability to detect weaker and potentially transient interactions (for example the dissociation constant K_d for a typical yeast-two-hybrid assay is 10-100μM compared to ~10mM for GST-pulldown assays) [105].

High-throughput interactome studies. Alternative approaches to the yeast-two-hybrid method, include high-throughput quantitative mass spectrometry (MS) and tandem affinity purification (TAP-MS) mass spectrometry. Both methods have been used to develop an interactome of all cellular proteins that interact with HCMV-encoded proteins [34, 40]. Given the preceding, it is reasonable to question why previous studies failed to identify the multitude of protein interactions between HCMV and the complement system that we identified by yeast-two-hybrid. There are several potential explanations for why most of the HCMV-Complement interactions we identified may not have been previously discovered. First, large-scale interactome studies require high enrichment thresholds of interacting protein pairs which means that these studies may underestimate or miss entirely genuine interactions that may exist, but occur transiently, at different concentrations, or in low frequencies. Second, complement proteins can exist in multiple isoforms with the same protein being secreted, membrane bound, or intracellular depending on different conditions. Thus, complement proteins can exist endogenously or exogenously with respect to cell type, and may exert their function in an autocrine or paracrine fashion in the context of HCMV infection. Third, most HCMV interactome studies analyze protein interactions in infected cell lysates, and since many complement proteins are secreted these interactions may be missed entirely if they are occurring extracellularly. Lastly, human foreskin fibroblasts (HFF) cells remain the most commonly used cell type for studying HCMV *in vitro,* and while most cells are able to produce some complement components, different cell types vary greatly in both the concentration and breadth of capacity to produce some or all complement components [106-109].

Validation of Complement-HCMV protein interactions by co-immunoprecipitation
Of the 121 protein interactions we identified by yeast-two-hybrid screening, we validated a subset of complement-HCMV protein interactions by co-immunoprecipitation analysis. In total we tested 13 Complement-HCMV protein interactions representing a combination of 5 different complement proteins and 9 HCMV proteins. Of the 13 protein interactions, 8 were positive by co-immunoprecipitation and occurred in viral proteins: UL24(MBL2), UL25(C1qBP, MASP1), UL44(C1qBP), UL82(C1qBP, MBL2), and UL84(C1qBP, MBL2). Of the 8 positive Complement-HCMV protein interactions, we were initially unaware that the C1qBP-UL44 and C1qBP-UL84 interactions had been previously reported using mass spectrometry, though we were comforted by the fact that we reproduced these interactions both by yeast-two-hybrid and co-immunoprecipitation which strongly support our experimental methods.

MBL2 Interactions and Hypotheses. We identified 18 interactions between MBL2 and HCMV-encoded proteins by yeast-two-hybrid. These data are significant because no specific protein interactions between MBL2 and HCMV-encoded proteins have been previously reported in the literature, even though studies have suggested that MBL2 may serve an important role in innate immunity against HCMV. For example, certain MBL2 gene polymorphisms result in decreased levels of overall MBL2 expression and have been associated with a higher risk of invasive HCMV disease after solid organ transplantation and pediatric cytomegalovirus infection [22, 110]. Taken together, there is growing evidence that MBL2 serves an important role in protection against HCMV and specific MBL2 polymorphisms that result in decreased expression, may be a significant risk factor for HCMV infection in certain populations.

Within the classical framework of complement as an extracellular mediator of innate immunity, some interactions we identified by yeast-two-hybrid were not unexpected such as the interactions between MBL2 and UL9, UL41A, UL55(gB), and UL73(gN) since these proteins are all surface exposed in HCMV virions and have predicted N-linked glycosylation sites that are expected to be recognized by MBL2.

A previous study reported that the exogenous addition of either serum purified or recombinant MBL2 can inhibit HCMV infection of cells *in vitro,* and the inhibition could be reduced proportionally with increasing concentrations of mannan, a major ligand recognized by MBL2 [111]. Although no specific mechanism was put forward, the preceding observations suggest that MBL2 recognizes an HCMV viral protein that is surface exposed and is involved in the initial steps of viral entry. Of the five membrane proteins that interacted with MBL2, UL55(gB) and UL73(gN) are involved in viral entry and are both leading candidate antigens in HCMV vaccines currently being evaluated in human clinical trials [112]. We suspect that the MBL2 dependent mechanism of inhibiting HCMV infection reported in the previous study may be due to the binding of MBL2 to UL55 and/or UL73 and sterically hindering these molecules from engaging with host cell receptors and thereby inhibiting viral entry. Given the preceding, this may mechanistically explain the conclusions of some studies that report that low MBL2 serum levels are correlated with an increased risk of HCMV in certain populations.

In our yeast-two-hybrid screen, MBL2 also interacted with several HCMV proteins that are reported to be expressed in the nucleus, cytoplasm, and ER/Golgi of cells during viral replication. We questioned how these interactions could be possible given that MBL2 has been almost exclusively characterized at the functional level as a secreted extracellular protein. To our surprise, a lesser known alternative intracellular isoform also exists and is found in cytoplasmic granules, on the periphery of the nucleus, and is abundant in the ER [113]. In the ER, MBL2 contributes to the quality control mechanisms of proteins that travel through the ER by recognizing glycan chains attached to proteins [114]. N-linked glycosylation is one of the most common covalent protein modifications in Eukaryotic cells and this post-translational modification has significant impacts on protein folding, sorting, degradation, and ultimately secretion [115, 116]. Importantly, the post-translational modification of proteins via the attachment of glycan chains, particularly high mannose-type oligosaccharides (which are recognized by MBL2), has been shown to serve an important role in not only protein quality control in the ER, but also protein sorting into different subcellular compartments [117, 118].

64

Unlike some viruses, HCMV does not globally arrest most intracellular protein trafficking systems such as the secretory pathway, but instead adopts a more conspicuous strategy of remodeling and manipulating host protein trafficking pathways [119]. While our understanding of how HCMV manipulates host protein trafficking remains incompletely understood, HCMV infection is known to induce broad changes in protein trafficking networks and alters the structure of multiple organelles that participate in these networks [120]. For example, in HCMV viral replication, newly synthesized capsids egress from the nucleus into a perinuclear compartment known as the assembly compartment (AC) which is a pseudo structure formed by the dramatic rearrangement of multiple cellular organelles including the ER and Golgi which are critical for the production and trafficking of newly synthesized viral proteins [120, 121]. Since the ER and Golgi constitute the first steps in protein secretion, and vesicular traffic of proteins and lipids occur continuously through these organelles, we hypothesize that intracellular MBL2 may recognize heavily glycosylated or post-translationally modified HCMV proteins and modulate their localization by targeting proteins through the secretory pathway into different subcellular compartments, one of which may include the HCMV viral assembly compartment.

Within our hypothesis, we posit that there could be at least two plausible MBL2-mediated protein trafficking mechanisms that HCMV could employ to promote viral replication, for example: (1) HCMV may use MBL2 to traffic innate antiviral proteins away from cellular compartments where viral replication or assembly are occurring, or (2) HCMV may use MBL2 to traffic viral proteins to sites of viral replication and assembly.

In either scenario, HCMV could effectively misappropriate the host ER/Golgi secretory pathway in an MBL2-dependent fashion to promote viral replication. This concept is not without precedent, as MBL2-mediated viral protein trafficking between cellular compartments has been previously reported in Human Immunodeficiency Virus (HIV). HIV encodes gp120 which is a ligand for MBL2, and previous studies have shown MBL2 mediated trafficking of gp120 to different subcellular compartments including the ER/Golgi, vesicles, and in the perinuclear space [122]. Interestingly, this process occurred independent of viral replication, and did not require additional viral proteins, as it was also observed in co-transfected cells expressing gp120 and MBL2 [122]. MBL2-mediated protein trafficking of gp120 did however depend on the carbohydrate recognition domain of MBL2 (which interacts with the glycosylated regions of proteins) and cell microtubules. Nevertheless, there is indeed precedent for a novel and unconventional role of MBL2 to mediate viral protein trafficking.

Given the intracellular role of MBL2 in regulating protein trafficking, and the extracellular role of MBL2 in activating the lectin pathway of complement, high or low expression of MBL2 could have pro-viral or anti-viral effects depending on whether HCMV virions were inside or outside of the cell. For example, high levels of MBL2 in serum might prevent HCMV infection by binding to surface exposed proteins involved in viral entry (i.e. gB/UL55, gN/UL73). In contrast, if HCMV exploits the function of MBL2 to regulate protein trafficking, increased expression of intracellular MBL2 would be expected to have pro-viral effects. Taken together, future studies are needed to understand both the intracellular and extracellular role of MBL2 in HCMV infection and immunity.

MASP-1 Interactions and Hypotheses. We identified 22 protein interactions between MASP-1 and HCMV-encoded proteins involved in diverse processes such as: capsid assembly (UL86), viral packaging (UL52), nuclear egress (UL53), and tegument proteins (UL24, UL25, UL43), among others by yeast-two-hybrid. A major challenge for all viruses is coordinating with great precision the localization, processing, and assembly of viral proteins to generate nascent virions prior to the final maturation process. Such complex events generally require the assistance of central protein hubs that can interact with multiple proteins involved in different stages of the viral lifecycle and execute their assembly. Of the interactions identified by yeast-two-hybrid, we tested two by co-immunoprecipitation (MASP1-UL25, MASP1-UL69) and validated the interaction between MASP-1 and UL25.

UL25 is a 85kD phosphoprotein that is one of the most abundant outer tegument proteins in the HCMV virion [123], and is expressed exclusively in the cytoplasm during the late phase of the HCMV viral replication cycle during assembly [124]. Interestingly, UL25 is dispensable for HCMV replication *in vitro* [70], as is the case with many other HCMV tegument proteins. In a previous study by To *et al*., the UL25 tegument protein was found to interact with other tegument proteins (UL24, UL83, UL43), viral transcription factors (UL122/IE2), viral replication factors (UL44), and viral DNA packaging/cleavage factors (UL89.2) by co-immunoprecipitation. These observations strongly implicate that UL25 may serve as a "hub" protein that functions as an organizing center for connecting multiple viral proteins in the mature virion and for recruiting other virion proteins during viral assembly and maturation [69].

Although MASP-1 is most well characterized for its extracellular role as an activating serine protease enzyme in the lectin pathway of the complement system, MASP-1 has also been reported to exist intracellularly within the cytoplasm and nucleus [125]. In the context of HCMV infection, previous transcriptional studies of HCMV infected cells *in vitro* have reported that MASP-1 is significantly suppressed at both the transcriptional and translational level during the early stages of infection [34], which may suggest that MASP-1 serves an important antiviral function. While the precise role of intracellular MASP-1 in HCMV infection remains unknown, we hypothesize that MASP-1 may interact with the UL25 protein "hub" and mediate the cleavage of bound UL25 interacting partners, many of which are essential for HCMV replication. In this context, we speculate that MASP-1 may antagonize the HCMV viral replication process, which may explain why MASP-1 is suppressed at the transcriptional and translational level during HCMV infection *in vitro*.

C1qBP Interactions & Hypotheses. C1qBP (also known as gC1qR, HABP1, and p32) is an important inhibitor of the classical pathway of the complement system, though increasing studies have demonstrated that C1qBP is a multifunctional protein that interacts with many host ligands and is involved in various diverse cellular processes depending on the context of protein interaction and subcellular expression. C1qBP is expressed on the plasma membrane of cells where it regulates the classical pathway of complement, but it is also expressed in the cytoplasm, and is highly expressed in the mitochondria. Under specific conditions, such as treatment of cells with chemical reagents [126], or during infection C1qBP can also translocate to the nucleus. [52, 79, 127].

In response to HCMV infection *in vitro* C1qBP expression is significant upregulated at both the transcriptional and translational level [34], and previous studies have reported that C1qBP is one of many cellular proteins that localize to the nucleus during infection and forms complexes with multiple HCMV proteins. Despite many gaps in knowledge regarding the precise and perhaps multiple roles of C1qBP in HCMV infection, reduced expression of C1qBP during HCMV infection *in vitro* reduces the expression of essential viral proteins and significantly inhibits HCMV viral replication [128]. In contrast, over-expression of C1qBP during HCMV viral replication results in increased viral titers and greater release of viral particles [129]. These observations suggest that C1qBP serves an essential role in HCMV replication, though the precise mechanism and full repertoire of C1qBP-interacting viral proteins encoded by HCMV remain unknown.

Of all complement proteins tested in our interactome study, the highest number of interactions by far involved interactions with C1qBP. We identified 31 protein interactions between C1qBP and HCMV by yeast-two-hybrid, with interactions occurring in HCMV-encoded proteins involved in broad and diverse processes such as: immune evasion (UL16, UL18, UL31, UL118, UL123.3), nuclear assembly/viral replication (UL52, UL53, UL56, UL84, UL85, UL86, UL102), tegument proteins (UL23, UL24, UL25, UL49, UL71, UL82), and viral transcription (UL44, UL92), among others.

Given the high number of interactions we identified between C1qBP and HCMV-encoded proteins involved in transcription, capsid formation, egress, and assembly, we inferred that C1qBP may be involved in multiple steps of the HCMV viral replication cycle.

C1qBP may regulate HCMV gene expression and participate in DNA replication. During HCMV infection, viral RNA transcription and DNA replication occur concomitantly in the nucleus in a tightly regulated manner starting with the expression of immediate early (IE) genes. As their name implies, IE genes are the first to be transcribed after infection and optimize the cell for viral gene expression and replication. IE genes are transcribed from the HCMV IE loci and are reported to have multiple functions including activating expression of HCMV early (E) genes, modulating the cell cycle [130], inhibiting apoptosis [131], and suppressing innate immunity via multiple mechanisms [132]. In HCMV, many different proteins can be produced from the IE loci, however the 72kD IE1 (also known as IE1/IE72) and 86kD IE2 (also known as IE2/IE86) nuclear phosphoproteins are the most abundant and important. Unlike IE1, IE2 is essential for HCMV viral replication [133], and is extremely promiscuous in its transcription of and interactions with cellular proteins and other HCMV proteins.

A major regulator of IE2 is the polyfunctional UL84 protein, which serves as a bridge between the complex network of proteins involved in HCMV gene expression and viral DNA replication. IE2-UL84 together form a complex involved in the initiation of viral DNA synthesis from the origin of lytic replication (OriLyt) in the HCMV genome. OriLyt-dependent DNA replication requires UL84 along with six known core proteins: UL44 (DNA processivity factor), UL54 (DNA polymerase), UL70 (primase), UL105 (helicase), UL102 (primase-associated factor), and UL57 (single-stranded DNA protein) [134].

Previous studies have reported that HCMV infection results in increased expression of C1qBP and the translocation of C1qBP from the cytoplasm to the nucleus in cells [79]. Additionally, C1qBP was reported to directly interact with both UL84 and UL44 by mass-spectrometry and co-immunoprecipitation [79, 87], which suggests that C1qBP may be involved in both HCMV gene expression and viral DNA replication. C1qBP has long been suspected to participate in gene expression and RNA splicing due to the fact that it has a strong predicted transcription activation domain and has been found in complexes with host transcription factors such as Forkhead box C1 (FOXC-1) [135] as well as pre-mRNA splicing factor 2 (SF2) [136]. While the idea of a complement protein being involved in transcription is unorthodox, C1qBP is involved in the transcriptional regulation of other related human herpesviruses. For example, in Epstein-Barr virus (EBV), C1qBP participates in the transcriptional regulation of Epstein-Barr Nuclear Antigen-1 (EBNA-1) that is required for maintenance of the viral chromosome during latency in infected proliferating cells and is essential in EBV latent cell cycle DNA replication [137, 138].

As illustrated in **Supplemental Figure 2-1**, we developed a hypothetical model for the role of C1qBP in HCMV viral gene expression and DNA replication. In our model, we propose that IE2-UL84-C1qBP-UL44 together form a multiprotein complex that regulates the expression of viral genes and participates in viral DNA replication. Our model was developed based on our yeast-two-hybrid and co-immunoprecipitation experiments that demonstrate that C1qBP interacts with both UL44 and UL84, a result that is consistent with previous studies [79]. Our belief that IE2-UL84-C1qBP-UL44 form a complex that regulates the expression of viral genes and participates in viral DNA replication is supported not only by *in vitro* protein binding experiments, but also previously published fluorescent microscopy data that report all four proteins co-localizing to the nucleus of HCMV infected cells [79]. Importantly, by yeast-two-hybrid, we expanded on previous studies and identified many other C1qBP-HCMV interactions that have never

previously been reported and encompass proteins that are also involved in viral transcription (UL82, UL92, UL124) and viral DNA replication (UL49, UL52, UL56, UL102) as illustrated in **Supplemental Figure 2-1**. Taken together, our data strongly implicate a role for C1qBP in HCMV viral transcription and DNA replication, which may explain the observation that reduced expression of C1qBP during HCMV infection *in vitro* reduces the expression of essential viral proteins and significantly attenuates HCMV viral replication [128].
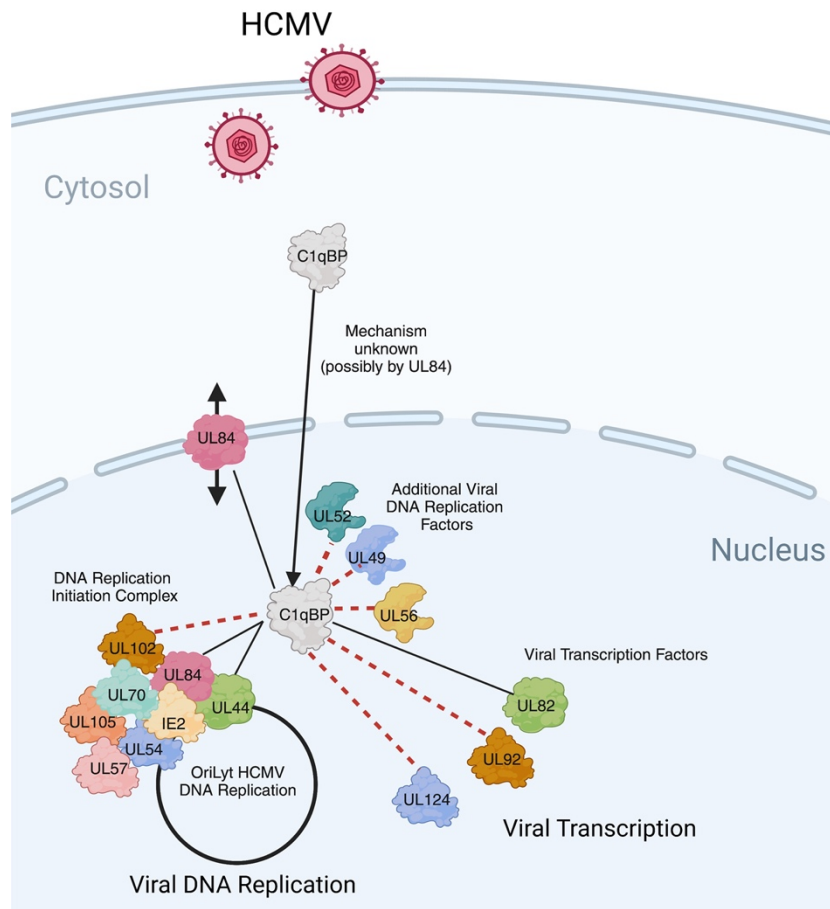
Supplemental Figure 2-1. Hypothetical model of C1qBP in HCMV gene expression and viral DNA replication.

Illustration of C1qBP interactions with HCMV proteins involved in the viral DNA replication and viral transcription. HCMV infection results in increased expression of C1qBP and the translocation of C1qBP from the cytoplasm to the nucleus in cells [79]. The precise mechanism for C1qBP translocation from the cytoplasm to the nucleus is unknown, but may be mediated by UL84 which is known to shuttle between the cytoplasm and nucleus [139]. The HCMV OriLyt-dependent DNA replication requires UL84 along with six known core proteins: UL44 (DNA processivity factor), UL54 (DNA polymerase), UL70 (primase), UL105 (helicase), UL102 (primase-associated factor), and UL57 (single-stranded DNA protein) [134]. Direct protein interactions are denoted by solid black lines and include interactions between C1qBP-UL44 and C1qBP-UL84 which interact by mass-spectrometry and co-immunoprecipitation [79], and were confirmed in our study by yeast-two-hybrid and co-immunoprecipitation. We identified the novel interaction between C1qBP-UL82 by yeast-two-hybrid, and confirmed the interaction by co-immunoprecipitation. Novel interactions identified in our study by yeast-two-hybrid are denoted by broken red lines and include HCMV proteins involved in viral transcription (UL92, UL124) and viral DNA replication (UL102). Original figure designed with the assistance of Biorender software (www.Biorender.com).

Role of C1qBP HCMV in nuclear egress. Previous studies have demonstrated that C1qBP (also known as gC1qR, HABP1, and p32) is an essential host protein involved in HCMV nuclear egress. To exit the nucleus, nascent HCMV capsids must travel through the nuclear egress complex (NEC), a multiprotein complex composed of: UL50, UL53, as well as the UL97 kinase. In addition to viral proteins, many host cellular proteins have been reported to be involved in the NEC including but not limited to: C1qBP, Lamin B receptor (LBR), Protein Kinase C (PKC) [128]. Functionally, the process of dilating the nuclear pore to allow HCMV capsids to exit the nucleus through the NEC is dependent on phosphorylation of C1qBP by the UL97 viral kinase, as well as the phosphorylation of UL50 and other proteins in the nuclear lamina [129, 140, 141].

In our study, we detected positive protein interactions in some but not all HCMV-encoded NEC components known to interact with C1qBP. For example, the interactions between C1qBP-UL50 and C1qBP-UL97 were not detected in our yeast-two-hybrid assay but have been previously reported by mass-spectrometry and co-immunoprecipitation studies [129, 142]. Although it is unclear why we did not observe a positive interaction between C1qBP-UL50, in the case of UL97 it is likely due the fact that UL97 interacts with C1qBP on its N-terminus [129] which may be inaccessible when fused to the GAL4 binding domain which is N-terminally tagged in our yeast-two-hybrid system. We did however detect a positive protein interaction between C1qBP and UL53 which is a core component of the HCMV NEC by yeast-two-hybrid assay. Different studies have come to conflicting conclusions regarding the interaction between C1qBP and UL53, with some reporting no interaction by yeast-two-hybrid [143], a weak interaction by co-immunoprecipitation [142], and a positive interaction by mass-spectrometry [128]. These discrepancies underscore a major challenge in studying the viral and host protein composition of the HCMV NEC, which likely includes many more viral and host proteins than our current knowledge. Conceivably, there are many more viral and host proteins involved in HCMV nuclear egress but that are not detected in high-throughput proteomic studies since these studies tend to bias stronger protein interactions and have certain enrichment thresholds that they use to define positive interactions. Thus, weaker albeit biologically significant host protein interactions with major and minor NEC components may be missed due to different experimental methods. Additionally, the NEC appears to be a highly dynamic structure, and its composition of accessory protein may differ at different timepoints of infection. Indeed, previous studies have reported a change in NEC composition biased towards the enrichment of cellular proteins at later timepoints in HCMV infection, compared to earlier stages of infection [128].

Thus in the current accepted model of the HCMV nuclear egress complex illustrated in **Supplemental Figure 2-2**, C1qBP acts as a major adaptor protein hub that functions in coordination with viral and host proteins to form the NEC necessary to dilate the nuclear pore large enough to allow HCMV capsids to travel from the nucleus to the cytoplasm.

By yeast-two-hybrid, we also identified many HCMV-encoded proteins that interacted with C1qBP that were involved in the steps prior to or shortly after nuclear egress. Before nuclear egress, capsids must be synthesized and loaded with HCMV viral DNA. We found that C1qBP interacts with four proteins (UL85, UL86, UL52, UL56) involved in capsid assembly. UL85 encodes a triplex capsid protein, while UL86 encodes a capsid scaffolding protein. In contrast, UL52 contributes to the efficient localization of neo-synthesized capsids to nuclear replication compartments and is required for cleavage and packaging of the HCMV genome [144], and

71

UL56 is a component of the molecular motor that translocate viral genomic DNA into empty capsids.

Immediately after nuclear egress, cytoplasmic capsids shed their primary envelope in a process known as de-envelopment and capsids concentrate near the assembly complex (AC) which is made up of ER and Golgi membranes and is located at the periphery of the nucleus. In the assembly complex, capsids acquire tegument proteins in a process known as tegumentation and undergo a secondary envelopment followed by maturation. By yeast-two-hybrid, we found that C1qBP interacts with UL71, a protein that has many roles during the time course of HCMV infection but is also highly involved in the egress of particles from the perinuclear space and secondary envelopment of cytoplasmic capsids that bud into the trans-Golgi network [145]. Taken together, these findings suggest that C1qBP may participate in the HCMV viral replication cycle shortly before nuclear egress during capsid assembly, and immediately after nuclear egress.
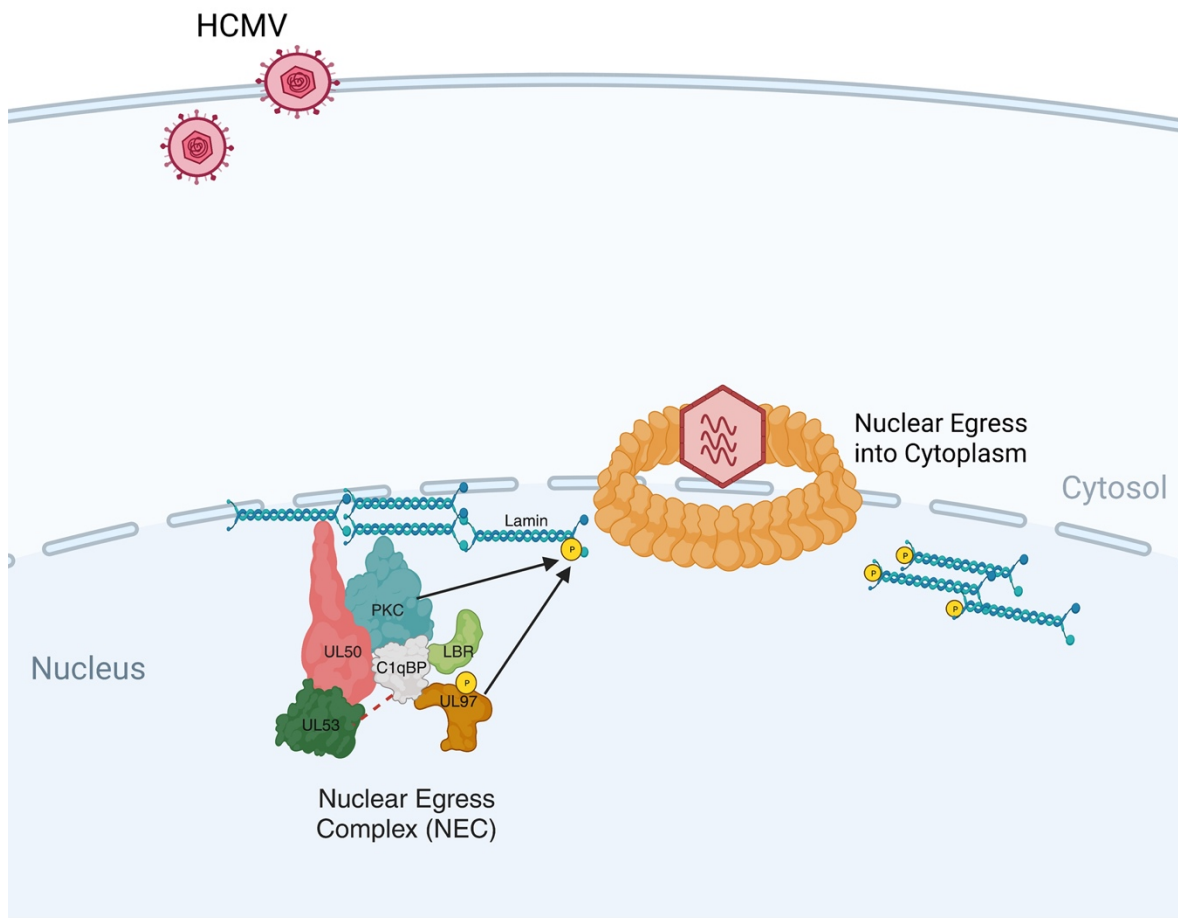
Supplemental Figure 2-2. Current model of C1qBP in the Nuclear Egress Complex (NEC). Figure depicts the current accepted model of the HCMV Nuclear Egress Complex (NEC). Nascent HCMV capsids travel through the NEC, composed of: UL50, UL53, as well as the UL97 kinase. In addition to viral proteins, the NEC also includes the following reported host proteins: C1qBP, Lamin B receptor (LBR), Protein Kinase C (PKC) [128]. Functionally, the process of dilating the nuclear pore to allow HCMV capsids to exit the nucleus through the NEC is dependent on phosphorylation of C1qBP by the UL97 viral kinase, as well as the phosphorylation of UL50 and other proteins in the nuclear lamina [129, 140, 141]. Direct protein interactions between C1qBP and major HCMV proteins that make up the NEC have been previously reported and include the following interactions: C1qBP-UL97 [129], C1qBP-UL50 [128], C1qBP-UL53 [128, 143]. We identified positive protein interactions between C1qBP and UL53 by yeast-two-hybrid. Figure designed with the assistance of Biorender software (www.Biorender.com) and includes protein interaction data from this study and other published studies.

C1qBP may cooperate with UL82 to suppress the cGAS-STING mediated cellular antiviral response at two different stages. Of the 31 proteins we identified as interacting partners with C1qBP, we validated and confirmed the novel protein interactions between C1qBP-UL25 and C1qBP-UL82. Herein, we describe the possible role of C1qBP-UL82 in suppressing the cGAS-STING mediated cellular antiviral response.

For DNA viruses such as HCMV, cytosolic DNA detected in the cytoplasm initiates the DNA sensor cyclic GMP-AMP synthase (cGAS) which in turn produces cGAMP once bound to DNA. Subsequently, cGAMP acts as a secondary messenger and binds to the endoplasmic reticulum membrane protein stimulator of IFN genes (STING). STING must then be trafficked to the Golgi to recruit host kinases and transcription factors that induce the expression of antiviral interferon genes as illustrated in **Supplemental Figure 2-3** [146-149].

HCMV is known to suppress the expression of interferons during infection, though the precise mechanisms, viral proteins involved, and contribution of host factors remain incompletely understood. Recently, the HCMV UL82 gene was reported to be involved in inhibiting the cGAS-STING pathway. During infection, UL82 is localized to the nucleus but is also abundantly found in the ER and cytoplasm [150, 151]. Specifically, UL82 inhibits the cGAS-STING pathway by interfering with the formation of STING functional translocation complexes that are necessary for STING to migrate from the ER to the Golgi, as well as to perinulear microsomes [151]. The cGAS-STING pathway has been demonstrated to be essential for mediating cellular innate immune responses to HCMV [152], and this is supported by the observation that mutant HCMV virions lacking UL82 have severe growth defects and reduced levels of viral progeny [70, 153].

In our study we found that C1qBP interacted with UL82 by yeast-two-hybrid and we validated the protein interaction by co-immunoprecipitation. These observations led us to question whether C1qBP could be involved in the cGAS-STING pathway. Under normal conditions C1qBP is predominantly localized to the mitochondria in cells, however in a recent study by Song *et al.,* the authors report that C1qBP leaks from the mitochondria into the cytoplasm during the early stages infection with herpes-simplex virus 1 (HSV-1) *in vitro*. In the cytoplasm, C1qBP was identified as a cGAS interacting protein by mass-spectrometry and co-immunoprecipitation, and the C1qBP-cGAS interaction inhibited cGAS enzymatic activity in cells and *in vitro*. Using truncation mutants of C1qBP, the authors identified the c-terminal region of C1qBP (amino acids 74-220) that is necessary and sufficient to bind to the nucleotidyltransferase (NTase) domain of cGAS and inhibit its enzymatic activity [52]. Taken together, these results implicate that cytosolic C1qBP can inhibit the cGAS-STING pathway by targeting cGAS directly.

HCMV are known to encode multiple proteins with redundant functions that can interfere with host antiviral response networks at multiple levels. Given the preceding, we hypothesize that the UL82 protein binds to C1qBP in the and forms a complex that allows for the dual and synergistic inhibition of the cGAS-STING pathway at multiple levels. In this context, an activated cGAS-STING pathway could be inhibited at the initial stage by the ability of C1qBP to inhibit the enzymatic activity of cGAS, as well as in the intermediate stage of the pathway where UL82 interferes with the formation of STING functional translocation complexes that are necessary for

STING to migrate from the ER to the Golgi. Both mechanisms would be expected to concomitantly decrease the expression of interferons and may be yet another example of HCMV regulating host signaling pathways at multiple levels.
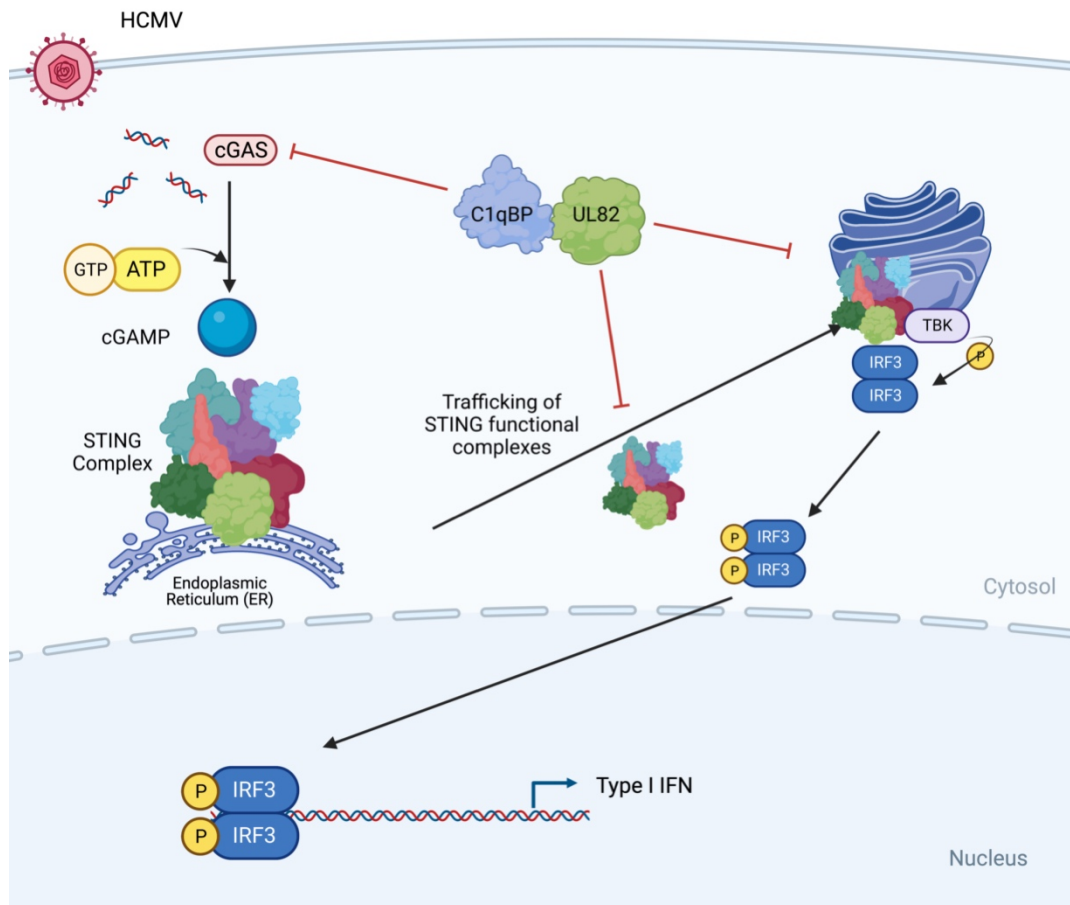
Supplemental Figure 2-3. Hypothetical model of C1qBP in suppression of cGAS-STING mediated antiviral immunity

Figure depicts a hypothetical model of the role of C1qBP in suppression of cGAS-STING mediated antiviral immunity. During HCMV infection cytosolic DNA is detected by the DNA sensor cyclic GMP-AMP synthase (cGAS) which binds to DNA and produces cGAMP that acts as a secondary messenger and binds to the endoplasmic reticulum (ER) membrane protein stimulator of IFN genes (STING). STING functional complexes are trafficked from the ER to the Golgi where it recruits kinases such as TANK-binding kinase 1 (TBK1) which phosphorylates interferon regulatory factor 3 (IRF3) which then translocates to the nucleus and induces the expression of antiviral type 1 interferon genes (type 1 IFN). Previous studies have reported that UL82 interferes with the formation of STING functional complexes that traffic from the ER to the Golgi and suppress the cGAS-STING pathway [151]. C1qBP is reported to bind to the nucleotidyltransferase (NTase) domain of cGAS and directly inhibit its enzymatic activity [52]. In our study we identified the novel interaction between C1qBP and UL82 by yeast-two-hybrid and validated the interaction by co-immunoprecipitation. We hypothesize that the C1qBP and UL82 form a complex in the cytoplasm during HCMV infection and synergistically inhibit the cGAS-STING pathway at multiple levels. Original figure designed with the assistance of Biorender software (www.Biorender.com) and includes protein interaction data from this study and other published studies.

Future directions. Complement proteins are canonically thought of as a system of innate immune proteins secreted by the liver that function extracellularly in the blood. However, despite the ancient evolutionary origins of these small proteins, our understanding of these important immune molecules is ever expanding as illustrated by many recent studies that report complement proteins having novel and unorthodox functions within cells.

In its intracellular state, the complement system (referred to as the 'complosome') has been reported to be involved in diverse processes including: sustaining T-cell homeostasis and mediating differentiation [154], regulating apoptosis [155], regulating cell metabolism [156], and has even been reported to directly neutralize intracellular pathogens. In a groundbreaking study by Tam *et al.,* the authors report that intracellular sensing of various C3b coated viral and bacterial pathogens in the cytosol triggers mitochondrial antiviral signaling (MAVS) pathways, inflammatory cytokine synthesis, and most surprisingly -destruction of pathogens by proteasomal degradation [157]. The described intracellular C3 mechanism of pathogen detection and destruction occurred independently of specific PAMP's and appears to be conserved in mammals, since C3 from various mammalian species deposited on viruses or bacteria was sufficient to activate the proceeding mechanism [157]. One caveat of this intracellular C3b sensing mechanism is that it appears to be less common for enveloped viruses than in non-enveloped viruses. Since complement components, specifically C3b, are deposited on the membrane of enveloped viruses, which fuse their membrane with host cells during attachment and penetration. Nevertheless, viruses which encode complement regulating proteins, or which co-opt host complement inhibitors to limit C3b deposition on their surface would be at a great advantage of evading this seemingly highly conserved cell autonomous immunity.

The primary objective of this study was to identify novel protein interactions between HCMV-encoded proteins and human complement proteins, with the long-term goal of characterizing the role of complement in immunity to HCMV. This study was inspired by a small number of previous reports that HCMV virions resist complement-mediated neutralization but are nevertheless coated with activated complement protein fragments that normally precede the formation of the membrane attack complex and results in direct lysis of microbes. These observations suggested that HCMV virions not only activate one or more pathways of complement but are able to inhibit complement-mediated neutralization, perhaps by a viral-encoded complement regulating protein. Though we were unable to investigate this question fully, our study is the most comprehensive undertaking to date in investigating protein interactions between HCMV and the complement system. Though the role of the complement system in immunity to HCMV remains incompletely understood, our study provides an important framework for future studies to characterize the functional significance of protein interactions between HCMV and the complement system.

MATERIALS AND METHODS

Yeast strains. The *Saccharomyces cerevisiae* yeast cells used for this study were generously provided by Dr. Sheng Luan at UC Berkeley. The yeast strains used for our yeast two-hybrid (YTH) assay were Y187 (MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met–, gal80Δ, URA3 :: GAL1UAS-GAL1TATA-lacZ) and AH109 (MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1UAS-GAL1TATA-HIS3,GAL2UAS-GAL2TATA-ADE2,URA3::MEL1UAS-MEL1TATA-lacZ).These strains of opposite mating types are one of the most common yeast mating pairs used for yeast two-hybrid screening. The resultant diploids from mating of these strains have multiple nutritional selective markers that permit optimization of the YTH assay in order to enhance or reduce the specificity and sensitivity of the protein-protein interactions.

Construction of the HCMV-Yeast-Two-Hybrid plasmid library. HCMV Towne$_{BAC}$ (accession no. AY315197) was used as a template for PCR amplification of HCMV ORFs. For genes that encode spliced transcripts, cDNAs from Towne$_{BAC}$ -infected cells were used as the templates. **Table 1** lists the primers used for cloning HCMV sequences into pGBKT7 bait and pGADT7 prey plasmids for expression in yeast (Clontech, Mountain View, CA) and in pCMV-Myc and pCMV-HA for expressing in human cells (Clontech, Mountain View, CA). Each primer sequence contained an outer and inner restriction enzyme site for cloning the HCMV sequences at the multiple cloning site (MCS) of the yeast and mammalian expression plasmids respectively. PCR amplification was performed using iProof high-fidelity DNA polymerase (Bio-Rad, Inc., Hercules, CA). The resultant constructs were confirmed by restriction digest profile and sequencing.

Construction of Human Complement protein Yeast-Two-Hybrid Plasmids. The cDNA's for human complement proteins: Decay Accelerating Factor/CD55 (CD55; SinoBiological, NCBI Ref Seq. NM_000574.3), Mannan Binding Lectin 2 (MBL2; SinoBiological, NM_000242.2), Mannan Associated Serine Protease 1 (MASP1; SinoBiological, NM_139125.2), Factor H (FH; Transomic, BC142699), Properdin (FP; Transomic, BC015756), and C1q Binding Protein (C1qBP; Transomic, BC000435) were used as a template for PCR amplification. Vectors containing the human complement protein cDNA sequence were constructed by using a pENTR/SD/D-TOPO cloning kit (Invitrogen) to generate Gateway compatible entry clones. PCR amplification was performed using iProof high-fidelty DNA polymerase (Bio-Rad, Inc., Hercules, CA) with forward primers (5' to 3') containing a CACC sequence for directional blunt end TOPO cloning. **Table 2** lists the primers used for cloning the human complement protein cDNA sequences into pENTR entry plasmids. The pENTR entry plasmids containing the human complement proteins were verified by restriction enzyme digest and DNA sequencing (data not shown) prior to being used as entry vectors in the gateway recombination reactions. Gateway recombination cloning was performed using the pENTR entry vectors containing the human complement proteins and the gateway compatible yeast-two-hybrid destination vector pGADT7-GW, mediated by the LR-Clonase enzyme (Thermofisher) according to the manufacturers instructions. The final pGADT7 vectors containing each of the six human complement proteins tested was verified for correct insert by restriction enzyme digest profiling and DNA sequencing (data not shown).

Yeast Electrotransformation. To generate a yeast cell line carrying either the HCMV or human complement protein yeast-two-hybrid plasmids (pGBKT7 or pGADT7-GW respectively), we followed the protocol as previously described by [158]. In brief, a 10mL culture of YPDA media was inoculated with a single colony of either AH109 or Y187 yeast and grown overnight at 30ºC with vigorous shaking at 200 rpm in a rotating incubator. The following day the culture was diluted to an OD_{600nm}=0.1 in 50mL volume of YPDA media, and cells were growth for 9 hours until they reached the logarithmic stage of growth OD_{600nm}=0.6. Next, cells were harvested by centrifugation at 5000 rpm for 5 minutes and the cell pellet was resuspended in 25mL of sterile deionized water to wash away residual media. Cells were once again pelleted under the same conditions and resuspended in 1M cold filter sterilized sorbitol followed by an additional round of centrifugation under the same conditions. After removing the sorbitol supernatant, the cells were treated with 2mL LiTE (1M Lithium Acetate Tris-EDTA) supplemented with 25mM DTT (Dithiothreiol) to induce competency for DNA uptake. Subsequently, cells were pelleted at 5000 rpm for 5 minutes and the supernatant discarded, followed by a wash step whereby the cell pellet was once more resuspended with 2mL of cold filter sterilized 1M sorbitol. Next, 250µl of the yeast sorbitol suspension was transferred to a sterile microcentrifuge tube and 5µl of plasmid DNA was added and gently mixed by pipetting up and down. Following pipetting, the 250µl (yeast, DNA, sorbitol) suspension was transferred to a chilled cuvette and pulse charged at 1.5 kV, 200 OHMS, 25 µF (pulse time of 5ms). Immediately after, 1mL of YPDA media was added to the cuvette and the cells were allowed to recover at 30ºC for 1 hour prior to plating the cells on selective media (SD-Leu for pGADT7 or SD-Trp for pGBKT7). Lastly, the cells on the agar plate were incubated at 30ºC for 3-5 days until single colonies appeared.

Yeast-two-hybrid (YTH) Analysis. The DNAs of the viral ORFs that were cloned into pGBKT7 and transformed into AH109 (*MAT*a) yeast, while the human complement protein cDNA cloned into pGADT7-GW were transformed into Y187 (*MAT* α), respectively (Matchmaker 3 System, Clontech). AH109 strains harboring pGBKT7 plasmids were maintained in minimal SD media with tryptophan dropout supplement (SD/-Trp), while Y187 strains harboring pGADT7-GW plasmids were maintained in minimal media with leucine dropout supplement (SD/-Leu). Prior to performing the matings, individual AH109 and Y187 strains were plated on SD/-Ade/-His/-Trp or SD/-Ade/-His/-Leu agar respectively, supplemented with 40µg/ml X-a-Gal to test for autoactivation. AH109 strains containing the sequences of eight HCMV ORF's cloned into pGBKT7 (UL26, UL48A, UL48N, UL48C, UL48.5, UL51, UL94, US23) were determined to be autoactivators in the absence of any pGADT7-cloned ORF's and subsequently eliminated from further mating experiments. None of the human complement proteins cloned into the pGADT7-GW plasmid were found to autoactivate under the same conditions in the absence of any pGBKT7-cloned ORF's. Yeast mating was carried out by inoculating fresh colonies (<2 weeks old) of both AH109 and Y187 strains into 1.5 ml microcentrifuge tubes containing 0.5 ml of YPDA media, and incubated at 30ºC with shaking at 200 rpm for 24 hours. There were a total of 1,002 possible combinations (167 AH109 x 6 Y187) and each mating combination was performed in duplicate. To select for diploids that have successfully been mated, yeast mating cultures were centrifuged at 14,000 rpm for 30 seconds, resuspended in 0.5 ml Tris-EDTA buffer, and plated in one well of a 48-well plate containing 1ml of SD/-Ade/-His/-Leu/-Trp (quadruple drop-out (QDO)) agar with 40 µg/ml of X-a-Gal. Three weeks after plating the mated diploid yeasts, the QDO plates were scored for positive protein-protein interactions. QDO/ X-a-Gal plates represent the highest selection stringency for eliminating possible false positives. AH109

strains contain three reporters-ADE2, HIS3, and MEL1 (encodes a-galactosidase) under unique GAL4 upstream activating sequences (UAS) and TATA boxes. The mating between AH109 transformed with a plasmid expressing the GAL4 binding domain (BD)-p53 fusion protein and Y187 transformed with a plasmid expressing the GAL4 activation domain (AD)-T (SV40 antigen) fusion protein was used as the positive control in each assay. The mating between AH109 transformed with a plasmid expressing the BD-Lamin C fusion protein and Y187 transformed with a plasmid expressing AD-T was used as the negative control. Growth of the diploid yeast cells representing the positive interactions was further analyzed. When scoring the duplicates, if both had no yeast growth after 3 weeks incubation at 30ºC, the fusion proteins were considered non-interacting partners. When the duplicates both had blue yeast colonies, or if only one of the wells had blue yeast colonies, that mating combination was repeated until the results from two of three independent experiments were consistent.

Cell Culture & Transfection. HeLa cells were obtained from American Type Culture Collection (ATCC) (Manassas, VA) and cultured in 12-well flat bottom tissue culture plates (Genesee Scientific) with Dulbecco modified Eagle medium (Gibco) supplemented with 5% (vol/vol) fetal bovine serum (Hyclone), 100 U of penicillin/mL, and 100µg of streptomycin/ml in an atmosphere of 5% CO_2 at 37ºC. Cells were transfected with pCMV-HA or pCMV-Myc (Clonetech, Mountain View, CA) containing either the HCMV ORF's or human complement protein sequences, using Lipofectamine 2000 (Invitrogen, Carlsbad, CA). Expression of the transgenes was verified 3 days post-transfection by Western blot analysis.

Co-immunoprecipitation analysis. Co-immunoprecipitation of HCMV-Complement protein complexes was performed using the Takara Magnetic Bead anti-Myc Immunoprecipitation Kit according to the manufacturers instructions. In brief, at 3 days post-transfection cells were removed from the 37ºC incubator and the media supernatant discarded. Each well of the 12-well flat bottom tissue culture plate was washed twice with 5mL of 1X PBS and aspirated. After washing, 1mL of Lysis buffer (150mM NaCl, 5mM EDTA [pH 8.0], 50mM Tris, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with ProteoGuard EDTA-Free Protease Inhibitor cocktail was added to each well of the plate and incubated for 30 minutes at 4ºC on a rocking platform. After the incubation, the cells were further disrupted by scraping any remaining adherent cells from the bottom of the tissue culture wells and the resultant cell lysate was collected in a 1.5mL microcentrifuge tube. The raw cell lysate was clarified by centrifugation at 12,000 rpm for 30 mins at 4ºC. To prime the anti-Myc magnetic beads for immunoprecipitation, 20µl of the anti-Myc magnetic beads were added to a fresh 1.5mL microcentrifuge tube and washed with 1mL of lysis buffer using a magnetic stand to capture the beads to the wall of the tube and discarding the residual lysis buffer. After priming, 1mL of the clarified cell lysate was added to the 1.5mL microcentrifuge tube containing magnetic beads and incubated overnight at 4ºC on a rotating shaker. The following day, the microcentrifuge tube was placed on the magnetic stand to allow the magnet to collect the immunoprecipitated and the supernatant was aspirated without disturbing the bead pellet. The magnetic beads were then washed a total of three times with 1mL of Wash buffer (150mM NaCl, 1mM EDTA, 1% Triton X-100, 10mM Tris). After washing, the remaining supernatant was discarded and the magnetic bead pellet was resuspended in 1X SDS sample buffer (0.006M Tris-HCl [pH 6.8], 10% [vol/vol] glycerol, 2% [wt/vol] SD, 5% [vol/vol] 2-mecaptoethanol, 10µg bromophenol blue per mL), vortexed and allowed to sit for 5 mins at room temperature. After, the immunoprecipitated

sample in SDS buffer was denatured at 98ºC for 5 minutes and analyzed by SDS-PAGE and Western Blot.

SDS-PAGE & Western Blot Analysis. Raw cell lysates of transfected cells were prepared by and separated on sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and Western blotting as described previously [159, 160] with a Mini-Protean II electrophoresis apparatus (Bio-Rad, Richmond, CA). Samples were suspended in SDS sample buffer (0.006M Tris-HCl [pH 6.8], 10% [vol/vol] glycerol, 2% [wt/vol] SD, 5% [vol/vol] 2-mecaptoethanol, 10μg bromophenol blue per mL) and heated to 100ºC for 10 minutes to denature before being loaded directly onto the gel. For Western blots, the gel was equilibrated with buffer (48mM Tris-HCl, 39mM glycine [pH 9.0], 20% [vol/vol] methanol) and transferred to a nitrocellulose membrane (Bio-Rad) by using a Trans-blot (Bio-Rad) semidry electrophoretic transfer cell. The nitrocellulose membranes were blocked with 5% (wt/vol) skim milk in PBS containing 0.2% (wt/vol) sodium azide. Hemmaglutinin (HA) and Myc (Myc) tagged HCMV or human complement proteins were detected using a mouse anti-HA (2-2.2.14; Invitrogen) at a 0.2μg/mL concentration or a 1:2,000 dilution of anti-Myc (Myc.A7; Invitrogen) monoclonal antibodies diluted in PBS containing 1% (wt/vol) bovine serum albumin, 1% (wt/vol) Tween-20, and 0.2% (wt/vol) sodium azide. Bound antibody was detected with 1:2,000 dilution of alkaline phosphatase (AP) conjugated polyclonal goat anti-mouse IgG (Cell signaling technologies) and developed using Sigma Fast BCIP/NBT (5-bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium) substrate (Sigma).

REFERENCES

1. Dunkelberger, J.R. and W.C. Song, *Complement and its role in innate and adaptive immune responses.* Cell Res, 2010. **20**(1): p. 34-50.
2. Fries, L.F., et al., *Glycoprotein C of herpes simplex virus 1 is an inhibitor of the complement cascade.* J Immunol, 1986. **137**(5): p. 1636-41.
3. Harris, S.L., et al., *Glycoprotein C of herpes simplex virus type 1 prevents complement-mediated cell lysis and virus neutralization.* J Infect Dis, 1990. **162**(2): p. 331-7.
4. Huemer, H.P., et al., *Herpes simplex virus glycoprotein C: molecular mimicry of complement regulatory proteins by a viral protein.* Immunology, 1993. **79**(4): p. 639-47.
5. Kostavasili, I., et al., *Mechanism of complement inactivation by glycoprotein C of herpes simplex virus.* J Immunol, 1997. **158**(4): p. 1763-71.
6. Frank, I. and H.M. Friedman, *A novel function of the herpes simplex virus type 1 Fc receptor: participation in bipolar bridging of antiviral immunoglobulin G.* J Virol, 1989. **63**(11): p. 4479-88.
7. Ogembo, J.G., et al., *Human complement receptor type 1/CD35 is an Epstein-Barr Virus receptor.* Cell Rep, 2013. **3**(2): p. 371-85.
8. Tanner, J., et al., *Epstein-Barr virus gp350/220 binding to the B lymphocyte C3d receptor mediates adsorption, capping, and endocytosis.* Cell, 1987. **50**(2): p. 203-13.
9. Spiller, O.B., et al., *Functional activity of the complement regulator encoded by Kaposi's sarcoma-associated herpesvirus.* J Biol Chem, 2003. **278**(11): p. 9283-9.
10. Spiller, O.B., et al., *Dissecting the regions of virion-associated Kaposi's sarcoma-associated herpesvirus complement control protein required for complement regulation and cell binding.* J Virol, 2006. **80**(8): p. 4068-78.
11. Santoro, F., et al., *CD46 is a cellular receptor for human herpesvirus 6.* Cell, 1999. **99**(7): p. 817-27.
12. Mori, Y., et al., *Human herpesvirus 6 variant A glycoprotein H-glycoprotein L-glycoprotein Q complex associates with human CD46.* J Virol, 2003. **77**(8): p. 4992-9.
13. Spiller, O.B., et al., *Neutralization of cytomegalovirus virions: the role of complement.* J Infect Dis, 1997. **176**(2): p. 339-47.
14. Agrawal, P., et al., *Complement Evasion Strategies of Viruses: An Overview.* Front Microbiol, 2017. **8**: p. 1117.
15. Kawasaki, T., *Structure and biology of mannan-binding protein, MBP, an important component of innate immunity.* Biochim Biophys Acta, 1999. **1473**(1): p. 186-95.
16. Kilpatrick, D.C., *Introduction to mannan-binding lectin.* Biochem Soc Trans, 2003. **31**(Pt 4): p. 745-7.
17. Kilpatrick, D.C., *Phospholipid-binding activity of human mannan-binding lectin.* Immunol Lett, 1998. **61**(2-3): p. 191-5.
18. Palaniyar, N., et al., *Nucleic acid is a novel ligand for innate, immune pattern recognition collectins surfactant proteins A and D and mannose-binding lectin.* J Biol Chem, 2004. **279**(31): p. 32728-36.
19. Poon, I.K., M.D. Hulett, and C.R. Parish, *Molecular mechanisms of late apoptotic/necrotic cell clearance.* Cell Death Differ, 2010. **17**(3): p. 381-97.
20. Pagh, R., et al., *The chaperone and potential mannan-binding lectin (MBL) co-receptor calreticulin interacts with MBL through the binding site for MBL-associated serine proteases.* FEBS J, 2008. **275**(3): p. 515-26.

21.   Ram, S., L.A. Lewis, and P.A. Rice, *Infections of people with complement deficiencies and patients who have undergone splenectomy.* Clin Microbiol Rev, 2010. **23**(4): p. 740-80.

22.   Hu, Y., et al., *Association between mannose-binding lectin gene polymorphism and pediatric cytomegalovirus infection.* Viral Immunol, 2010. **23**(4): p. 443-7.

23.   Dobo, J., et al., *The emerging roles of mannose-binding lectin-associated serine proteases (MASPs) in the lectin pathway of complement and beyond.* Immunol Rev, 2016. **274**(1): p. 98-111.

24.   Matsushita, M., et al., *Proteolytic activities of two types of mannose-binding lectin-associated serine protease.* J Immunol, 2000. **165**(5): p. 2637-42.

25.   Beltrame, M.H., et al., *MBL-associated serine proteases (MASPs) and infectious diseases.* Mol Immunol, 2015. **67**(1): p. 85-100.

26.   Thiel, S., et al., *A second serine protease associated with mannan-binding lectin that activates complement.* Nature, 1997. **386**(6624): p. 506-10.

27.   Matsushita, M., Y. Endo, and T. Fujita, *MASP1 (MBL-associated serine protease 1).* Immunobiology, 1998. **199**(2): p. 340-7.

28.   Teillet, F., et al., *Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins.* J Biol Chem, 2008. **283**(37): p. 25715-25724.

29.   Gingras, A.R., et al., *Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: implications for complement activation.* Structure, 2011. **19**(11): p. 1635-43.

30.   Presanis, J.S., et al., *Differential substrate and inhibitor profiles for human MASP-1 and MASP-2.* Mol Immunol, 2004. **40**(13): p. 921-9.

31.   Dobo, J., et al., *MASP-1, a promiscuous complement protease: structure of its catalytic region reveals the basis of its broad specificity.* J Immunol, 2009. **183**(2): p. 1207-14.

32.   Megyeri, M., et al., *Complement protease MASP-1 activates human endothelial cells: PAR4 activation is a link between complement and endothelial function.* J Immunol, 2009. **183**(5): p. 3409-16.

33.   Dobo, J., et al., *Cleavage of kininogen and subsequent bradykinin release by the complement component: mannose-binding lectin-associated serine protease (MASP)-1.* PLoS One, 2011. **6**(5): p. e20036.

34.   Nightingale, K., et al., *High-Definition Analysis of Host Protein Stability during Human Cytomegalovirus Infection Reveals Antiviral Factors and Viral Evasion Mechanisms.* Cell Host Microbe, 2018. **24**(3): p. 447-460 e11.

35.   Kim, D.D. and W.C. Song, *Membrane complement regulatory proteins.* Clin Immunol, 2006. **118**(2-3): p. 127-36.

36.   Nicholson-Weller, A. and C.E. Wang, *Structure and function of decay accelerating factor CD55.* J Lab Clin Med, 1994. **123**(4): p. 485-91.

37.   Harris, C.L., et al., *Molecular dissection of interactions between components of the alternative pathway of complement and decay accelerating factor (CD55).* J Biol Chem, 2005. **280**(4): p. 2569-78.

38.   Lukacik, P., et al., *Complement regulation at the molecular level: the structure of decay-accelerating factor.* Proc Natl Acad Sci U S A, 2004. **101**(5): p. 1279-84.

39.   Brodbeck, W.G., et al., *Structure/function studies of human decay-accelerating factor.* Immunology, 2000. **101**(1): p. 104-11.

40.   Nobre, L.V., et al., *Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions.* Elife, 2019. **8**.

41.   Spiller, O.B., et al., *Altered expression of host-encoded complement regulators on human cytomegalovirus-infected cells.* Eur J Immunol, 1996. **26**(7): p. 1532-8.

42.   Klein, M., et al., *Strain-specific neutralization of human cytomegalovirus isolates by human sera.* J Virol, 1999. **73**(2): p. 878-86.

43.   Kemper, C., J.P. Atkinson, and D.E. Hourcade, *Properdin: emerging roles of a pattern-recognition molecule.* Annu Rev Immunol, 2010. **28**: p. 131-55.

44.   Kemper, C., et al., *The complement protein properdin binds apoptotic T cells and promotes complement activation and phagocytosis.* Proc Natl Acad Sci U S A, 2008. **105**(26): p. 9023-8.

45.   Cortes, C., et al., *Local release of properdin in the cellular microenvironment: role in pattern recognition and amplification of the alternative pathway of complement.* Front Immunol, 2012. **3**: p. 412.

46.   Holt, G.D., M.K. Pangburn, and V. Ginsburg, *Properdin binds to sulfatide [Gal(3-SO4)beta 1-1 Cer] and has a sequence homology with other proteins that bind sulfated glycoconjugates.* J Biol Chem, 1990. **265**(5): p. 2852-5.

47.   Blatt, A.Z., S. Pathan, and V.P. Ferreira, *Properdin: a tightly regulated critical inflammatory modulator.* Immunol Rev, 2016. **274**(1): p. 172-190.

48.   Higgins, J.M., et al., *Characterization of mutant forms of recombinant human properdin lacking single thrombospondin type I repeats. Identification of modules important for function.* J Immunol, 1995. **155**(12): p. 5777-85.

49.   van Leeuwen, H.C. and P. O'Hare, *Retargeting of the mitochondrial protein p32/gC1Qr to a cytoplasmic compartment and the cell surface.* J Cell Sci, 2001. **114**(Pt 11): p. 2115-23.

50.   Muta, T., et al., *p32 protein, a splicing factor 2-associated protein, is localized in mitochondrial matrix and is functionally important in maintaining oxidative phosphorylation.* J Biol Chem, 1997. **272**(39): p. 24363-70.

51.   Chen, R., et al., *Identification of a novel mitochondrial interacting protein of C1QBP using subcellular fractionation coupled with CoIP-MS.* Anal Bioanal Chem, 2016. **408**(6): p. 1557-64.

52.   Song, K., et al., *Leaked Mitochondrial C1QBP Inhibits Activation of the DNA Sensor cGAS.* J Immunol, 2021. **207**(8): p. 2155-2166.

53.   Yu, L., et al., *Molecular cloning and characterization of a cellular protein that interacts with the human immunodeficiency virus type 1 Tat transactivator and encodes a strong transcriptional activation domain.* J Virol, 1995. **69**(5): p. 3007-16.

54.   Pednekar, L., et al., *Analysis of the Interaction between Globular Head Modules of Human C1q and Its Candidate Receptor gC1qR.* Front Immunol, 2016. **7**: p. 567.

55.   Weiler, J.M., et al., *Control of the amplification convertase of complement by the plasma protein beta1H.* Proc Natl Acad Sci U S A, 1976. **73**(9): p. 3268-72.

56.   Whaley, K. and S. Ruddy, *Modulation of the alternative complement pathways by beta 1 H globulin.* J Exp Med, 1976. **144**(5): p. 1147-63.

57.   Pangburn, M.K., R.D. Schreiber, and H.J. Muller-Eberhard, *Human complement C3b inactivator: isolation, characterization, and demonstration of an absolute requirement for the serum protein beta1H for cleavage of C3b and C4b in solution.* J Exp Med, 1977. **146**(1): p. 257-70.

58.   Alsenz, J., et al., *Localization of the complement-component-C3b-binding site and the cofactor activity for factor I in the 38kDa tryptic fragment of factor H.* Biochem J, 1984. **224**(2): p. 389-98.

59.   Gordon, D.L., et al., *Identification of complement regulatory domains in human factor H.* J Immunol, 1995. **155**(1): p. 348-56.

60.   Kuhn, S., C. Skerka, and P.F. Zipfel, *Mapping of the complement regulatory domains in the human factor H-like protein 1 and in factor H1.* J Immunol, 1995. **155**(12): p. 5663-70.

61.   Jokiranta, T.S., et al., *Analysis of the recognition mechanism of the alternative pathway of complement by monoclonal anti-factor H antibodies: evidence for multiple interactions between H and surface bound C3b.* FEBS Lett, 1996. **393**(2-3): p. 297-302.

62.   Oppermann, M., et al., *The C-terminus of complement regulator Factor H mediates target recognition: evidence for a compact conformation of the native protein.* Clin Exp Immunol, 2006. **144**(2): p. 342-52.

63.   Ferreira, V.P., et al., *Critical role of the C-terminal domains of factor H in regulating complement activation at cell surfaces.* J Immunol, 2006. **177**(9): p. 6308-16.

64.   Haque, A., et al., *Characterization of Binding Properties of Individual Functional Sites of Human Complement Factor H.* Front Immunol, 2020. **11**: p. 1728.

65.   Jokiranta, T.S., et al., *Each of the three binding sites on complement factor H interacts with a distinct site on C3b.* J Biol Chem, 2000. **275**(36): p. 27657-62.

66.   Katz, Y. and R.C. Strunk, *Synthesis and regulation of complement protein factor H in human skin fibroblasts.* J Immunol, 1988. **141**(2): p. 559-63.

67.   Chung, K.M., et al., *West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 19111-6.

68.   Pinter, C., et al., *Direct interaction of complement factor H with the C1 domain of HIV type 1 glycoprotein 120.* AIDS Res Hum Retroviruses, 1995. **11**(5): p. 577-88.

69.   To, A., et al., *Yeast two hybrid analyses reveal novel binary interactions between human cytomegalovirus-encoded virion proteins.* PLoS One, 2011. **6**(4): p. e17796.

70.   Dunn, W., et al., *Functional profiling of a human cytomegalovirus genome.* Proc Natl Acad Sci U S A, 2003. **100**(24): p. 14223-8.

71.   Dunn, W., et al., *Human cytomegalovirus expresses novel microRNAs during productive viral infection.* Cell Microbiol, 2005. **7**(11): p. 1684-95.

72.   Chee, M.S., et al., *Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169.* Curr Top Microbiol Immunol, 1990. **154**: p. 125-69.

73.   Davison, A.J., et al., *The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome.* J Gen Virol, 2003. **84**(Pt 1): p. 17-28.

74.   Dolan, A., et al., *Genetic content of wild-type human cytomegalovirus.* J Gen Virol, 2004. **85**(Pt 5): p. 1301-1312.

75.   Murphy, E., et al., *Coding potential of laboratory and clinical strains of human cytomegalovirus.* Proc Natl Acad Sci U S A, 2003. **100**(25): p. 14976-81.

76.   Lu, Q., et al., *Arabidopsis homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin.* Plant J, 2010. **61**(2): p. 259-70.

77.   Maier, R., et al., *Construction of a reading frame-independent yeast two-hybrid vector system for site-specific recombinational cloning and protein interaction screening.* Biotechniques, 2008. **45**(3): p. 235-44.

78. Yu, D., M.C. Silva, and T. Shenk, *Functional map of human cytomegalovirus AD169 defined by global mutational analysis.* Proc Natl Acad Sci U S A, 2003. **100**(21): p. 12396-401.

79. Du, G. and M.F. Stinski, *Interaction network of proteins associated with human cytomegalovirus IE2-p86 protein during infection: a proteomic analysis.* PLoS One, 2013. **8**(12): p. e81583.

80. Atalay, R., et al., *Identification and expression of human cytomegalovirus transcription units coding for two distinct Fcgamma receptor homologs.* J Virol, 2002. **76**(17): p. 8596-608.

81. Kolb, P., et al., *Human cytomegalovirus antagonizes activation of Fcgamma receptors by distinct and synergizing modes of IgG manipulation.* Elife, 2021. **10**.

82. Mikesch, J.H., et al., *The expression and action of decay-accelerating factor (CD55) in human malignancies and cancer therapy.* Cell Oncol, 2006. **28**(5-6): p. 223-32.

83. Kwon, Y.C., et al., *Distinct CD55 Isoform Synthesis and Inhibition of Complement-Dependent Cytolysis by Hepatitis C Virus.* J Immunol, 2016. **197**(4): p. 1127-36.

84. Caras, I.W., et al., *Cloning of decay-accelerating factor suggests novel use of splicing to generate two proteins.* Nature, 1987. **325**(6104): p. 545-9.

85. Schierling, K., et al., *Human cytomegalovirus tegument proteins ppUL82 (pp71) and ppUL35 interact and cooperatively activate the major immediate-early enhancer.* J Virol, 2004. **78**(17): p. 9512-23.

86. Uetz, P., et al., *Herpesviral protein networks and their interaction with the human proteome.* Science, 2006. **311**(5758): p. 239-42.

87. Gao, Y., K. Colletti, and G.S. Pari, *Identification of human cytomegalovirus UL84 virus- and cell-encoded binding partners by using proteomics analysis.* J Virol, 2008. **82**(1): p. 96-104.

88. Al-Sharif, W.Z., et al., *Sea urchin coelomocytes specifically express a homologue of the complement component C3.* J Immunol, 1998. **160**(6): p. 2983-97.

89. Levashina, E.A., et al., *Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, Anopheles gambiae.* Cell, 2001. **104**(5): p. 709-18.

90. Kimura, A., E. Sakaguchi, and M. Nonaka, *Multi-component complement system of Cnidaria: C3, Bf, and MASP genes expressed in the endodermal tissues of a sea anemone, Nematostella vectensis.* Immunobiology, 2009. **214**(3): p. 165-78.

91. Avirutnan, P., et al., *Complement-mediated neutralization of dengue virus requires mannose-binding lectin.* mBio, 2011. **2**(6).

92. Conde, J.N., et al., *Inhibition of the Membrane Attack Complex by Dengue Virus NS1 through Interaction with Vitronectin and Terminal Complement Proteins.* J Virol, 2016. **90**(21): p. 9570-9581.

93. Ebenbichler, C.F., et al., *Human immunodeficiency virus type 1 activates the classical pathway of complement by direct C1 binding through specific sites in the transmembrane glycoprotein gp41.* J Exp Med, 1991. **174**(6): p. 1417-24.

94. Isaacs, S.N., G.J. Kotwal, and B. Moss, *Vaccinia virus complement-control protein prevents antibody-dependent complement-enhanced neutralization of infectivity and contributes to virulence.* Proc Natl Acad Sci U S A, 1992. **89**(2): p. 628-32.

95. Kotwal, G.J. and B. Moss, *Vaccinia virus encodes a secretory polypeptide structurally related to complement control proteins.* Nature, 1988. **335**(6186): p. 176-8.

96.   Saifuddin, M., et al., *Role of virion-associated glycosylphosphatidylinositol-linked proteins CD55 and CD59 in complement resistance of cell line-derived and primary isolates of HIV-1.* J Exp Med, 1995. **182**(2): p. 501-9.

97.   Vanderplasschen, A., et al., *Extracellular enveloped vaccinia virus is resistant to complement because of incorporation of host complement control proteins into its envelope.* Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7544-9.

98.   Gaggar, A., D.M. Shayakhmetov, and A. Lieber, *CD46 is a cellular receptor for group B adenoviruses.* Nat Med, 2003. **9**(11): p. 1408-12.

99.   Giot, L., et al., *A protein interaction map of Drosophila melanogaster.* Science, 2003. **302**(5651): p. 1727-36.

100.   Li, S., et al., *A map of the interactome network of the metazoan C. elegans.* Science, 2004. **303**(5657): p. 540-3.

101.   Ito, T., et al., *A comprehensive two-hybrid analysis to explore the yeast protein interactome.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4569-74.

102.   LaCount, D.J., et al., *A protein interaction network of the malaria parasite Plasmodium falciparum.* Nature, 2005. **438**(7064): p. 103-7.

103.   Rain, J.C., et al., *The protein-protein interaction map of Helicobacter pylori.* Nature, 2001. **409**(6817): p. 211-5.

104.   Zhou, Y., et al., *A comprehensive SARS-CoV-2-human protein-protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets.* Nat Biotechnol, 2022.

105.   Mackay, J.P., et al., *Protein interactions: is seeing believing?* Trends Biochem Sci, 2007. **32**(12): p. 530-1.

106.   Lubbers, R., et al., *Production of complement components by cells of the immune system.* Clin Exp Immunol, 2017. **188**(2): p. 183-194.

107.   Morgan, B.P. and P. Gasque, *Extrahepatic complement biosynthesis: where, when and why?* Clin Exp Immunol, 1997. **107**(1): p. 1-7.

108.   Heeger, P.S. and C. Kemper, *Novel roles of complement in T effector cell regulation.* Immunobiology, 2012. **217**(2): p. 216-24.

109.   Kolev, M., G. Le Friec, and C. Kemper, *Complement--tapping into new sites and effector systems.* Nat Rev Immunol, 2014. **14**(12): p. 811-20.

110.   Cervera, C., et al., *Influence of mannose-binding lectin gene polymorphisms on the invasiveness of cytomegalovirus disease after solid organ transplantation.* Transplant Proc, 2009. **41**(6): p. 2259-61.

111.   Wu, W., et al., *Human mannose-binding lectin inhibits human cytomegalovirus infection in human embryonic pulmonary fibroblast.* APMIS, 2012. **120**(8): p. 675-82.

112.   Anderholm, K.M., C.J. Bierle, and M.R. Schleiss, *Cytomegalovirus Vaccines: Current Status and Future Prospects.* Drugs, 2016. **76**(17): p. 1625-1645.

113.   Turner, M.W., *The role of mannose-binding lectin in health and disease.* Mol Immunol, 2003. **40**(7): p. 423-9.

114.   Nonaka, M., et al., *Subcellular localization and physiological significance of intracellular mannan-binding protein.* J Biol Chem, 2007. **282**(24): p. 17908-20.

115.   Helenius, A. and M. Aebi, *Roles of N-linked glycans in the endoplasmic reticulum.* Annu Rev Biochem, 2004. **73**: p. 1019-49.

116.   Morgan, G.W., et al., *Combined biochemical and cytological analysis of membrane trafficking using lectins.* Anal Biochem, 2013. **441**(1): p. 21-31.

117.   Schrag, J.D., et al., *Lectin control of protein folding and sorting in the secretory pathway.* Trends Biochem Sci, 2003. **28**(1): p. 49-57.

118.   Sitia, R. and I. Braakman, *Quality control in the endoplasmic reticulum protein factory.* Nature, 2003. **426**(6968): p. 891-4.

119.   Cruz, L. and N.J. Buchkovich, *Rerouting the traffic from a virus perspective.* Front Biosci (Landmark Ed), 2017. **22**(11): p. 1845-1866.

120.   Alwine, J.C., *The human cytomegalovirus assembly compartment: a masterpiece of viral manipulation of cellular processes that facilitates assembly and egress.* PLoS Pathog, 2012. **8**(9): p. e1002878.

121.   Das, S. and P.E. Pellett, *Spatial relationships between markers for secretory and endosomal machinery in human cytomegalovirus-infected cells versus those in uninfected cells.* J Virol, 2011. **85**(12): p. 5864-79.

122.   Teodorof, C., et al., *Intracellular mannose binding lectin mediates subcellular trafficking of HIV-1 gp120 in neurons.* Neurobiol Dis, 2014. **69**: p. 54-64.

123.   Varnum, S.M., et al., *Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome.* J Virol, 2004. **78**(20): p. 10960-6.

124.   Battista, M.C., et al., *Expression and characterization of a novel structural protein of human cytomegalovirus, pUL25.* J Virol, 1999. **73**(5): p. 3800-9.

125.   Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.

126.   Majumdar, M., et al., *Hyaluronan binding protein 1 (HABP1)/C1QBP/p32 is an endogenous substrate for MAP kinase and is translocated to the nucleus upon mitogenic stimulation.* Biochem Biophys Res Commun, 2002. **291**(4): p. 829-37.

127.   Matthews, D.A. and W.C. Russell, *Adenovirus core protein V interacts with p32--a protein which is associated with both the mitochondria and the nucleus.* J Gen Virol, 1998. **79 (Pt 7)**: p. 1677-85.

128.   Milbradt, J., et al., *Proteomic analysis of the multimeric nuclear egress complex of human cytomegalovirus.* Mol Cell Proteomics, 2014. **13**(8): p. 2132-46.

129.   Marschall, M., et al., *Cellular p32 recruits cytomegalovirus kinase pUL97 to redistribute the nuclear lamina.* J Biol Chem, 2005. **280**(39): p. 33357-67.

130.   Song, Y.J. and M.F. Stinski, *Inhibition of cell division by the human cytomegalovirus IE86 protein: role of the p53 pathway or cyclin-dependent kinase 1/cyclin B1.* J Virol, 2005. **79**(4): p. 2597-603.

131.   Zhu, H., Y. Shen, and T. Shenk, *Human cytomegalovirus IE1 and IE2 proteins block apoptosis.* J Virol, 1995. **69**(12): p. 7960-70.

132.   Paulus, C. and M. Nevels, *The human cytomegalovirus major immediate-early proteins as antagonists of intrinsic and innate antiviral host responses.* Viruses, 2009. **1**(3): p. 760-79.

133.   Marchini, A., H. Liu, and H. Zhu, *Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes.* J Virol, 2001. **75**(4): p. 1870-8.

134.   Pari, G.S., *Nuts and bolts of human cytomegalovirus lytic DNA replication.* Curr Top Microbiol Immunol, 2008. **325**: p. 153-66.

135.   Huang, L., et al., *Human p32 is a novel FOXC1-interacting protein that regulates FOXC1 transcriptional activity in ocular cells.* Invest Ophthalmol Vis Sci, 2008. **49**(12): p. 5243-9.

136.   Petersen-Mahrt, S.K., et al., *The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation.* EMBO J, 1999. **18**(4): p. 1014-24.

137.   Van Scoy, S., et al., *Human p32: a coactivator for Epstein-Barr virus nuclear antigen-1-mediated transcriptional activation and possible role in viral latent cycle DNA replication.* Virology, 2000. **275**(1): p. 145-57.

138.   Wang, Y., et al., *P32/TAP, a cellular protein that interacts with EBNA-1 of Epstein-Barr virus.* Virology, 1997. **236**(1): p. 18-29.

139.   Gao, Y., et al., *Nucleocytoplasmic shuttling of human cytomegalovirus UL84 is essential for virus growth.* J Virol, 2010. **84**(17): p. 8484-94.

140.   Sharma, M., et al., *Human cytomegalovirus UL97 phosphorylates the viral nuclear egress complex.* J Virol, 2015. **89**(1): p. 523-34.

141.   Sharma, M., et al., *Human cytomegalovirus UL50 and UL53 recruit viral protein kinase UL97, not protein kinase C, for disruption of nuclear lamina and nuclear egress in infected cells.* J Virol, 2014. **88**(1): p. 249-62.

142.   Milbradt, J., et al., *Cytomegaloviral proteins that associate with the nuclear lamina: components of a postulated nuclear egress complex.* J Gen Virol, 2009. **90**(Pt 3): p. 579-590.

143.   Milbradt, J., S. Auerochs, and M. Marschall, *Cytomegaloviral proteins pUL50 and pUL53 are associated with the nuclear lamina and interact with cellular protein kinase C.* J Gen Virol, 2007. **88**(Pt 10): p. 2642-2650.

144.   Borst, E.M., et al., *The essential human cytomegalovirus gene UL52 is required for cleavage-packaging of the viral genome.* J Virol, 2008. **82**(5): p. 2065-78.

145.   Schauflinger, M., et al., *The tegument protein UL71 of human cytomegalovirus is involved in late envelopment and affects multivesicular bodies.* J Virol, 2011. **85**(8): p. 3821-32.

146.   Ishikawa, H. and G.N. Barber, *STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling.* Nature, 2008. **455**(7213): p. 674-8.

147.   Ishikawa, H., Z. Ma, and G.N. Barber, *STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity.* Nature, 2009. **461**(7265): p. 788-92.

148.   Zhong, B., et al., *The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation.* Immunity, 2008. **29**(4): p. 538-50.

149.   Shu, H.B. and Y.Y. Wang, *Adding to the STING.* Immunity, 2014. **41**(6): p. 871-3.

150.   Hensel, G.M., et al., *Intracellular localization and expression of the human cytomegalovirus matrix phosphoprotein pp71 (ppUL82): evidence for its translocation into the nucleus.* J Gen Virol, 1996. **77 (Pt 12)**: p. 3087-97.

151.   Fu, Y.Z., et al., *Human Cytomegalovirus Tegument Protein UL82 Inhibits STING-Mediated Signaling to Evade Antiviral Immunity.* Cell Host Microbe, 2017. **21**(2): p. 231-243.

152.   Paijo, J., et al., *cGAS Senses Human Cytomegalovirus and Induces Type I Interferon Responses in Human Monocyte-Derived Cells.* PLoS Pathog, 2016. **12**(4): p. e1005546.

153.   Bresnahan, W.A. and T.E. Shenk, *UL82 virion protein activates expression of immediate early viral genes in human cytomegalovirus-infected cells.* Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14506-11.

154.   Liszewski, M.K., et al., *Intracellular complement activation sustains T cell homeostasis and mediates effector differentiation.* Immunity, 2013. **39**(6): p. 1143-57.

155. Martin, M., et al., *Factor H uptake regulates intracellular C3 activation during apoptosis and decreases the inflammatory potential of nucleosomes.* Cell Death Differ, 2016. **23**(5): p. 903-11.

156. Schaffler, A. and C. Buechler, *CTRP family: linking immunity to metabolism.* Trends Endocrinol Metab, 2012. **23**(4): p. 194-204.

157. Tam, J.C., et al., *Intracellular sensing of complement C3 activates cell autonomous immunity.* Science, 2014. **345**(6201): p. 1256070.

158. Ito, H., et al., *Transformation of intact yeast cells treated with alkali cations.* J Bacteriol, 1983. **153**(1): p. 163-8.

159. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

160. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4350-4.

SUPPLEMENTAL TABLES

Supplemental Table 2-1. HCMV virion primer sequences		
Viral ORF	**Direction (F or R)**	**Primer Sequence (5' to 3')**
RL1	F	CCGGAATTCTTGTCGACCCCAGCCACAGCCACAAACAG
	R	CGCGGATCCGGTACCCTACCGGGGACGTTTATCCTT
RL2	F	CCGGAATTCTTGTCGACCACGTCCCCGGTAGACGGGGT
	R	CGCGGATCCGGTACCTTATAACAAAAGAAGCACAAGCTC
RL4	F	GGGAATTCCATATGTCTCGAGGTCAGCATGCACGCGTGTATGTA
	R	CGCGGATCCGCGGCCGCCTATACGGAGATCGCGGTCC
RL5	F	CCGGAATTCTTGTCGACCCATACATACACGCGTGCATGCT
	R	CGCGGATCCGGTACCCTACCATATAAAAACGCAGGGGT
RL7	F	CCGGAA TTCTTGTCGACCAAAGCAAGAGGCAGCCGAG
	R	CGCGGATCCGGTACCTCATAAGGTAACGATGCTACTT
RL9	F	GGGAATTCCATATGTTGTCGACCAAACAGCGGACAGTCCCACG
	R	CCGGAATTCGGTACCCTAGATGCCGCCAGCCACCA
RL10	F	CCGGAATTCTTGTCGACCTATCCGCGTGTAATGCACG
	R	CGCGGATCCGGTACCTCAGACGTCGTCGTCCTC
RL11	F	CCGGAATTCTTGTCGACCCAGACCTACAGCACCCCCCT
	R	CGCGGATCCGGTACCTTACTCCAAATCCCCGTCCACC
UL1	F	CCGGAATTCTTGTCGACCGGTATACAATGTAACACTAAACTA
	R	CGCGGATCCGGTACCTCACCACGGCAGCTGTCCAA
UL2	F	CCGGAATTCTTGTCGACCCCAGAAAACTCGGTCGCGAT
	R	CGCGGATCCGGTACCTTATAAAAGAGCGTCTCGAAGC
UL3	F	GGGAATTCCATATGTTGTCGACCTATTTGCAAGTCAGCAAAGAC
	R	CGCGGATCCGGTACCTCATCTTTGACTATAAGGATC
UL4	F	GGGAATTCCATATGTTGTCGACCATCGTTATGATGCTTAGAGCG
	R	CGCGGATCCGGTACCTTAGGACACGGTCAGATTGTA
UL5	F	CCGGAATTCTTGTCGACCTTTCTAGGCTACTCTGACTGT
	R	CGCGGATCCGGTACCCTACACGGTAGCGACGAGA
UL6	F	CCGGAATTCGAGATCTCTAACGGGTGGGCTGGGGTGT
	R	CGCGGATCCGGTACCCTACGTTAACAGACCACGTTCT
UL7	F	CCGGAATTCTTGTCGACCGTTTCCGGCGTGGGTTCTC
	R	CGCGGATCCGGTACCCTATAATTCTGATTTACTACACAA
UL8	F	CCGGAATTCTTGTCGACCACTAGCCCGGGACTATCTA
	R	CGCGGATCCGGTACCTCATAGTTCTGTGTCCGTCA
UL9	F	CCGGAATTCTTGTCGACCTTCCGGTACGCGTTATTACTT
	R	CGCGGATCCGGTACCTTAAGTGGCGGCGAGCATAC
UL10	F	CCGGAATTCTTGTCGACCCGGAGACTGATTAACCATATC
	R	CGCGGATCCGGTACCTCACGATAATTAGTCCGTAGT
UL11	F	CCGGAATTCTTGTCGACCCTGCTCAGGTACATTACCTTT
	R	CGCGGATCCCTCGAGTTACAGATCGGTTTGGGGGAT

UL12	F	GGGAATTCCATATGTTGTCGACCCTCTTGGTCTTTCTAGGGC
	R	CCGGAATTCGGTACCTCATCCCCCAAACCGATCTG
UL13	F	GGGAATTCCATATGTTGTCGACCCTGTGGGCCCACTGCGGTC
	R	CCGGAATTCGGTACCCTAAATCGTCATGGGCCGTCG
UL14	F	GGGAATTCCATATGTTGTCGACCTGGTCGCGTGTGGTCTTTTTAAG
	R	CCGGAATTCCTCGAGTTACTCCCGCTGCTGCTCTTC
UL15	F	GGGAATTCCATATGTCTCGAGGTCAGATAGAAATACACATCCCGT
	R	CCGGAATTCGCGGCCGCTCACAGACGCGGTACCTGT
UL15A	F	CCGGAATTCTTGTCGACCAAGCGKATGATTCGCAGTCA
	R	CGCGGATCCCTCGAGTTATTGTGCGCGATACCGCAT
UL16	F	GGGAATTCCATATGTTGTCGACCGAGCGTCGCCGAGGTACGGT
	R	CCGGAATTCCTCGAGTCAGTCCTCGGTGCGTAACC
UL17	F	GGGAATTCCATATGTTGTCGACCGATCACGCGCTCTTCACACA
	R	CCGGAATTCGGTACCTTACTCCTCCTCTTCGGGAG
UL18	F	CCGGAATTCTTGTCGACCATGACAACGTGGTGTCTGACG
	R	CGCGGA TCCCTCGAGTCA TGACGACCGGACCTTGC
UL19	F	GGGAATTCCATATGTTGTCGACCACCATCAACGCCCTGTATGAG
	R	CGCGGATCCGGTACCTCACGCGTCCCTGCGACTC
UL20	F	GGGAATTCCATATGTCTCGAGGTCTCGGGATACGGGCTATGCTG
	R	GGGAATTCCATATGGCGGCCGCTCATGCTGGCATGCAGACCAC
UL20A	F	GGGAATTCCATATGTTGTCGACCGTCTTTAACCCTTGGGATCT
	R	CCGGAATTCGGTACCTTATCGTGTTTTTGCAGCGT
UL21	F	GGGAATTCCATATGTTGTCGACCTGGCCGCTTCTGAAAAACAAC
	R	CCGGAATTCGGTACCTCATGGGATGGGGACGAGG
UL22	F	GGGAATTCCATATGTTGTCGACCGCTCGGAGGCTATGGATACTGAGCTTAGCCGTGACCTTG ACGGTGGCTTTGGCGGCACCTTCTCAGAAA TCGAAGCGCAGCGTGACGGTGGAACAACCCA
	R	CCGGAATTCGGTACCCTACTGGGTCTTTTCATTTTCT
UL23	F	CCGGAATTCTTGTCGACCAACACCGACGCTACTTCCCGT
	R	CGCGGATCCGGTACCTCACGCGTCGTCAAAAAGTTGG
UL24	F	CCGGAA TTCGAGA TCTCTGAGGAGACCCGGGCGGGA
	R	CGCGGATCCGGTACCTCAACGGTGCTGACGTCCTT
UL25	F	CCGGAA TTCAAGGCCATGGAGGCCTCGTCGCGGCGTCGCAGCT
	R	CGCGGATCCGGTACCTCAGCAACAGTATTCCCCGCT
UL26	F	CCGGAA TTCTTGTCGACCACGAGCAGGCGCGCACCC
	R	CGCGGATCCGGTACCTTACGGCAACAGCGCTGATGG
UL27	F	CCGGAATTCGAGATCTCTAACCCCGTGGATCAGCCG
	R	CGCGGATCCGGTACCTCATGTGGCGTGACCTCCGA
UL28	F	CCGGAATTCTTGTCGACCGGCGACGGTGTTTTTATAACC
	R	CGCGGATCCGGTACCTCACGACGCCCCCGTGCC
UL29	F	GGGAATTCCATATGTTGTCGACCTCCGGCCGTCGCAAGGGCT
	R	CGCGGATCCCTCGAGTCACCTACGCTTTTTGAACGG
UL30	F	CCGGAATTCTTGTCGACCACGACATCGGCGATGACGGCT
	R	CGCGGATCCGGTACCTTATAGCCATCTGCCCCCCCGT

UL31	F	GGGAATTCCATATGTTGTCGACCAAGGCGGCAAAGTCCTCCAAG
	R	CGCGGATCCGGTACCTCAGGGATGGGGAGACGTGAG
UL32	F	GGGAATTCCATATGAAGATCTCTAGTTTGCAGTTTATCGGTCTACA
	R	CGCGGATCCGCGGCCGCCTATTCCTCCGTGTTCTTAATCTT
UL33	F	CCGGAATTCTTGTCGACCACGGAGACGCTATCCGCCAT
	R	CGCGGA TCCCTCGAGTCATACCCCGCTGAGGTTATGA
UL34	F	CCGGAATTCGAGATCTCTTCTGTCATCACGGTACACAGAT
	R	CGCGGATCCGGTACCTTAAATACACAACGGGGTTATGG
UL35	F	GGGAATTCCATATGTCTCGAGGTGCTCAAGGATCGCGAGCCCCA
	R	CGCGGATCCGCGGCCGCTTAGAGATGCCGTAGATTTTCGG
UL36.1	F	CCGGAATTCTTGTCGACCGACGACCTACGGGACACGTT
	R	CGCGGATCCGGTACCTTATTGCGAACAGACGGTGCCCA
UL36.2	F	CCGGAATTCGAGATCTCTGCTTACGTCTGCTGTCAGGAA
	R	CGCGGATCCGGTACCTTAGTTGTTCATGTAGGCGTGTG
UL37.1	F	GGGAATTCCATATGTTGTCGACCTCTCCAGTCTACGTGAATCTG
	R	CCGGAATTCGGTACCTTAGTGAGACTGCTGGGGGCCGT
UL37.3	F	CCGGAATTCTTGTCGACCGTGTCTGTGTGGCTGTTGTTG
	R	CGCGGATCCGGTACCTCAGACGATCCGATGAACGTC
UL38	F	GGGAATTCCATATGTTGTCGACCACTACGACCACGCATAGCACC
	R	CCGGAATTCGGTACCCTAGACCACGACCACCATCTG
UL39	F	CCGGAATTCTTGTCGACCGGGATTTCCACCTTTTTCTTAAG
	R	CGCGGATCCCTCGAGTCACAACTCTCGGAGCGGGT
UL41	F	GGGAATTCCATATGTCTCGAGGTCTGTACCATCATTGCTGCTAC
	R	CGCGGATCCGCGGCCGCTAACGGCAAGATGACGCGAG
UL41A	F	CCGGAATTCTCTCGAGGTACACTCTTTTGCCGCACCG
	R	CGCGGATCCGCGGCCGCTTAAAAGTCTGTGTCCGACTC
UL42	F	CCGGAATTCTCTCGAGGTGCCGCGGACGCCGTCGGT
	R	CGCGGATCCGCGGCCGCTTATCCCGATGTTGACACCGTC
UL43	F	GGGAATTCCATATGTTGTCGACCGAGAAAACGCCGGCGGAGACG
	R	CCGGAATTCGGTACCTCACCTTCGAGCAAAGAGCCCCT
UL44	F	GGGAATTCCATATGTTGTCGACCGATCGCAAGACGCGCCTCTC
	R	CGCGGATCCGGTACCCTAGCCGCACTTTTGCTTCTTG
UL45	F	CCGGAA TTCTTGTCGACCAATCCGGCTGACGCGGACGA
	R	CGCGGATCCGCGGCCGCCTAAGAGGCACAGTACTTATATAC
UL46	F	GGGAATTCCATATGTCTCGAGGTGATGCACGCGCGGTGGCCAAA
	R	CGCGGATCCGCGGCCGCTCAGACGAATTCTCGAAAGTCTCC
UL47	F	CCGGAATTCTTGTCGACCATGGCGAGGCGCACGGTAGATT
	R	CGCGGATCCGGTACCTCATGGCGAGGCCGGCGGCA
UL48	F	CCGGAA TTCAAAGTCACGCAGGCCAGCTGC
	R	CCGGAATTCTTACAAAAGATAGAGAAACCGCATGT
UL48A	F	CCGGAATTCAAAGTCACGCAGGCCAGCTGC
	R	CCGGAATTCTTACAAAAGATAGAGAAACCGCATGT
UL48N	F	CCGGAATTCAAAGTCACGCAGGCCAGCTGC

	R	CCGGAATTCTTACAGAAGTTGCGGACCCAAAATTTG
UL48C	F	CCGGAATTCATCCATGAGACGCAGCAGGCC
	R	CCGGAATTCTTACAAAAGATAGAGAAACCGCATGT
UL48.5	F	CCGGAATTCTTGTCGACCTCTAACACCGCGCCGGGACC
	R	CGCGGATCCGGTACCTCAGCGCCGGGTGCGCGAC
UL49	F	CCGGAATTCTCTCGAGGTGCCAATCGCCGTCTCCGACA
	R	CGCGGATCCGCGGCCGCTTAGACATGGGGCAGGCCGTG
UL50	F	CCGGAATTCTCTCGAGGTGAGATGAACAAGGTTCTCCATCA
	R	CGCGGATCCGCGGCCGCTCAGTCGCGGTGTGCGGAGC
UL51	F	CCGGAATTCTTGTCGACCTCCTGGGCTAAGCAGCGGGT
	R	CGCGGATCCGGTACCTTATTTACCCGGCGCCGGCT
UL52	F	GGGAATTCCATATGAAGATCTCTAATCCGAGTACCCACGTGAG
	R	CGCGGATCCGGTACCCTAGACATACTTGTCTATCACGTA
UL53	F	CCGGAATTCTTGTCGACCTCTAGCGTGAGCGGCGTGC
	R	CGCGGATCCGGTACCTCAAGGCGCACGAATGCTGTT
UL54	F	GGGAATTCCATATGTTTTTCAACCCGTATCTGAGCGG
	R	CCGGAATTCTCAACAGCATTCGTGCGCCTTG
UL55	F	CGCGGATCCGGGCCATGGAGGCCGAATCCAGGATCTGGTGCCTG
	R	CGCGGATCCGGTACCTCAGACGTTCTCTTCTTCGTCG
UL56	F	GGGAATTCCATATGTTGTCGACCGAGATGAATTTGTTACAGAAACTA
	R	CCGGAATTCGCGGCCGCTTAACGCAGACTACCAGGCAC
UL57	F	GGGAATTCCATATGGCGGCCGCTAGCCACGAGGAACTAACCGC
	R	CGCGGATCCGCGGCCGCTTACAACCGGCTGCGTTTGGC
UL58	F	CCGGAATTCTTGTCGACCCGGTTTACAAGCAGAAACACC
	R	CGCGGATCCGGTACCTTACTTTTTACAAGACGAAAAATGT
UL59	F	CCGGAATTCGAGATCTCTGCGCTCGTTGACCAGGAAAG
	R	CGCGGATCCGGTACCTTACCAAAAAACACATTTTTCGTCT
UL60	F	CCGGAATTCTTGTCGACCCCCGGCACGGGGCTCGCG
	R	CGCGGATCCGGTACCCTACGTATACCGGATGCTAGGC
UL61	F	CCGGAATTCGAGATCTCTTTGTGGAGGGGCGTACTGAC
	R	CGCGGATCCGGTACCTCAGGAGCTCCGGGAGCG
UL62	F	CCGGAATTCGAGATCTCTGCGCCGCTCGGCCAGACT
	R	CGCGGATCCGGTACCTCAATTGCCGTCAGTACGCCC
UL63	F	CCGGAATTCTTGTCGACCATTTCTCATCAGTATCATGAAAG
	R	CGCGGATCCGGTACCTTATACAACAGCGACAGTTCG
UL64	F	CCGGAATTCTTGTCGACCGTGAGAGGTCGCCACCGAG
	R	CGCGGATCCGGTACCCTAACCTTTAAAACCGTATCCTTC
UL66	F	CCGGAATTCGAGATCTCTTCATGTTACTACGTGTGTGTT
	R	CGCGGATCCGGTACCTCATGAGCACCACCTGTCG
UL67	F	CCGGAATTCTTGTCGACCGTGAGGAGTTTGGAGGAAATA
	R	CGCGGATCCGGTACCTTATATTGTGTGTATATAGATAGG
UL68	F	CCGGAATTCTTGTCGACCATGATAATGATGTACGCGTGC
	R	CGCGGATCCGGTACCCTACCACCCCGATAAAAACC

UL69	F	CCGGAATTCTCTCGAGGTGAGCTGCACTCACGCGGCCGT
	R	CGCGGATCCGCGGCCGCTTAGTCATCCATATCATCGCTGTAAC
UL70	F	GGGAATTCCATATGTTGTCGACCGATGCGCTGCGATTTCTTGCA
	R	CGCGGATCCGGTACCTCAGACGGCGGTCGCCGG
UL71	F	CCGGAATTCTTGTCGACCAGCAGAATCATACTCTGTTGCGA
	R	CGCGGATCCGCGGCCGCTCAGGGATGACGGGGAGGTTT
UL72	F	CCGGAATTCTCTCGAGGTCTTACGATGTTTACCGATCGAAT
	R	CGCGGATCCGCGGCCGCTCAGGGATGACGGGGAGGTTT
UL73	F	CCGGAATTCTCTCGAGGTGAGTGGAACACACTAGTACTAG
	R	CGCGGATCCGCGGCCGCTCAATAGCCTTTGGTGGTGGT
UL74	F	CGCGGATCCTGTCGACCGGGAGAAAGGGAGAGATGAGA
	R	CGCGGATCCGGTACCTTACTGCAACCACCACCAAAGG
UL75	F	CCGGAATTCTTGTCGACCCGACCAGGCCTCCCCTCCTA
	R	CCGGAATTCGGTACCTCAGCATGTCTTGAGCATGCGGT
UL76	F	CCGGAATTCTTGTCGACCCGTCCGGGTGTGGGGAC
	R	CGCGGATCCGGTACCCTATAAAGACCGTGTGGGACG
UL77	F	GGGAATTCCATATGAAGATCTCTAGTCTGTTGCACACCTTTTGGC
	R	CCGGAATTCGGTACCTTACAACACCGCCACGCTCGGA
UL78	F	CCGGAATTCGAGATCTCTTCCTCTTCTGCGGAGAAGACT
	R	CGCGGATCCGGTACCTCATAATGCCGTTACCGTCGC
UL79	F	CCGGAATTCTCTCGAGGTATGGCCCGCGACGAAGAGAA
	R	CGCGGATCCGCGGCCGCTCACACGTCGTTAGCCAGCGT
UL80	F	GGGAATTCCATATGAAGATCTCTACGATGGACGAGCAGCAGC
	R	CCGGAATTCGGTACCTTACTCGAGCTTATTGAGCGCA
UL80A	F	GGGAATTCCATATGTTGTCGACCTCGCACCCTCTGAGTGCTGC
	R	CCGGAATTCGGTACCTTACTCGAGCTTATTGAGCGCA
UL81	F	CCGGAATTCGAGATCTCTGCGTGCAGCACGCCGCAC
	R	CGCGGATCCGGTACCCTAATAAAGCACACCGCGGTT
UL82	F	CCGGAATTCTTGGCCATGGAGGCCTCTCAGGCATCGTCCTCGCC
	R	CGCGGATCCCTCGAGCTAGATGCGGGGTCGACTGC
UL83	F	CCGGAATTCTTGTCGACCGAGTCGCGCGGTCGCCGTTGT
	R	CGCGGATCCGGTACCTCAACCTCGGTGCTTTTTGGGCGT
UL84	F	GGGAATTCCATATGAAGATCTCTCCACGCGCCGACCCCAACC
	R	CCGGAATTCGGTACCTTAGAGATCGCCGCAGACCATG
UL85	F	CCGGAATTCTTGTCGACCGCGGCCATGGAGGCCAACAT
	R	CGCGGATCCGGTACCTCAGCCTTTAAATATGCAGGTC
UL86	F	GGGAATTCCATATGAAGATCTCTGAGAACTGGTCGGCGCTCGAG
	R	CCGGAATTCGGTACCTCACGAGTTAAATAACATGGATTGC
UL87	F	GGGAATTCCATATGTCTCGAGGTGCTGGTGCTGCGCCGCGC
	R	CCGGAATTCCTCGAGTCATCGTGATGCAAACCGCGCT
UL88	F	GGGAATTCCATATGTTGTCGACCATGGAAGCCGCGGCCGCTG
	R	CCGGAATTCCTCGAGCTAGGCACGCAGCAGAGCCA
UL89.1	F	CCGGAATTCTTGTCGACCTTGCGCGGAGACTCGGCC

	R	CGCGGATCCGGTACCTTAGTTGGTGTTGTAGCAACTGG
UL89.2	F	CCGGAATTCTCTCGAGGTAGCATCCGAGGACAAAACTTC
	R	CGCGGATCCGCGGCCGCCTAGCTGACCCTGAAACGGA
UL90	F	CCGGAATTCTTGTCGACCGCGGCTGTGGCGGGTCG
	R	CGCGGATCCGGTACCTTACGCACGCTCGGCTGGAA
UL91	F	CCGGAATTCGAGATCTCTAACTCGTTGCTGGCGGAACT
	R	CGCEGATCCGGTACCTCATGTCACAGGCGCCCGA
UL92	F	CCGGAATTCTTGTCGACCTGCGACGTCTCGGGCGCCT
	R	CGCGGATCCGGTACCTCAAACGCCAGATCCGAATACA
UL93	F	GGGAATTCCATATGTTGTCGACCGAAACGCACCTGTATTCGGATC
	R	CGCGGATCCGGTACCCTAAAGATCGTCGAACGGCAAG
UL94	F	CCGGAATTCTCTCGAGGTGCTTGGCGCAGCGGGCTTTG
	R	CGCGGATCCGCGGCCGCTTAGTGCACTAGGTTCTTAAGCAC
UL95	F	GGGAATTCCATATGGGTACCATGGCGGCGGCGGTGGTGC
	R	GGGAATTCCATATGGGTACCTCATAGATTCAACGTGATGAGACC
UL96	F	CCGGAATTCTCTCGAGGTACGTCGGTCAACAAACAGCTC
	R	CGCGGATCCGCGGCCGCTTAGACGGCGTCGTCGACCT
UL97	F	GGGAATTCCATATGTTGTCGACCTCCTCCGCACTTCGGTCTCG
	R	CCGGAATTCGGTACCTTACTCGGGGAACAGTTGGC
UL98	F	CCGGAATTCTTGTCGACCTGGGGCGTCTCGAGTTTGGA
	R	CGCGGATCCGCGGCCGCTCAGGGGCTCACCGGGCGT
UL99	F	CCGGAATTCTCTCGAGGTGGTGCCGAACTCTGCAAACGA
	R	CGCGGATCCGCGGCCGCTTAAAAGGGCAAGGAGGCGGC
UL100	F	CCGGAATTCTCTCGAGGTGCCCCCTCGCACGTGGATAA
	R	CGCGGATCCGCGGCCGCTCAAGCGTCCTCGAAGTCTTCA
UL101	F	CCGGAATTCTTGTCGACCCGGCGGCCCGCTCGCTC
	R	CGCGGATCCGGTACCTCACCGCGCACCACGAACG
UL102	F	CCGGAATTCGCGGCCGCAACCGCTCAGCCGCCGCTGC
	R	CGCGGATCCGCGGCCGCTTAAGCGTTGAGCCGGAAAAACC
UL103	F	CCGGAATTCTTGTCGACCGAGGCCCTGATGATCCGCG
	R	CGCGGATCCGGTACCTCACTCTTCCTCTCCCCGTTC
UL104	F	CCGGAATTCGCGGCCGCAGAGCGAAACCACTGGAACGAA
	R	CGCGGATCCGCGGCCGCCTAGTGAAATCCGTATGGACCT
UL105	F	GGGAATTCCATATGAAGATCTCTTCGATGACGGCCTCGTCATC
	R	CCGGAATTCAGATCTTCAAAAAATAAGCGTGGTGCGTT
UL106	F	CCGGAATTCTCTCGAGGTATGACCGACCGCACGGAG
	R	CGCGGATCCGCGGCCGCCTACGGGTGGTACCTGCAA
UL107	F	CCGGAATTCTTGTCGACCAAGCTTCCTTACAGCATAACT
	R	CGCGGATCCGGTACCTTATAACATGTATTTTGAAAAATTG
UL108	F	GGGAATTCCATATGTTGTCGACCGGTCATCGTCGTCCACGAC
	R	CGCGGATCCGGTACCCTACTTTAGAGATAATAAGACACC
UL109	F	CCGGAATTCTTGTCGACCATGATACACGACTACCACTG
	R	CGCGGATCCGGTACCTTACGAGCAAGAGTTCATCA

UL110	F	CCGGAATTCTTGTCGACCCCTTTGAAAACCAAGCCTATTG
	R	CGCGGATCCGGTACCTTAACTCGTGAGTGTGATGGTT
UL111	F	CCGGAATTCTTGTCGACCTCCTCTGGACGACACTGC
	R	CGCGGATCCGGTACCCTATTCTAACCGCGGAAGAA
UL111A	F	CCGGAATTCTTGTCGACCCTGTCGGTGATGGTCTCTTC
	R	CGCGGATCCGGTACCTCACCCGACACGCGGAAAAG
UL112.1	F	GGGAATTCCATATGTTGTCGACCGATCTGCCTACTACCGTCGT
	R	CCGGAATTCGGTACCTTAGTCGTTCTCGGAGGAGGGA
UL112.2	F	GGGAATTCCATATGAAGATCTCTGGCTCTCCTCCCCTCCGG
	R	CGCGGATCCGGTACCTTAATCGTCGAAAAACGCCGC
UL114	F	GGGAATTCCATATGTTGTCGACCGCCCTCAAGCAGTGGATGCT
	R	CCGGAATTCGGTACCTCACCCACAGAGTCGCCAGT
UL115	F	GGGAATTCCATATGTTGTCGACCTGCCGCCGCCCGGATTGCG
	R	CGCGGATCCGGTACCTTAGCGAGCATCCACTGCTTGA
UL116	F	GGGAATTCCATATGTCTCGAGGTTATCAACGTCAGCTTGTTGAC
	R	CGCGGATCCGCGGCCGCTCAAGTCTGCGGCACGATG
UL117	F	GGGAATTCCATATGTTGTCGACCGTCATGTTCTCCCAGGACCAC
	R	GGGAATTCCATATGGGTACCTCATGAGGTGGGCAGGCGAA
UL118	F	CCGGAATTCTTGTCGACCATTTTATCCACAGGAAAACTAC
	R	CGCGGATCCGGTACCCTACCACTGCTTGAAGTAGG
UL119	F	CCGGAATTCTTGTCGACCTGTCCCGTACTGGCGATCG
	R	CGCGGATCCGGTACCTTACTTACCTTTATCACCGGG
UL120	F	CCGGAATTCTTGTCGACCTACCGTGCGGGGGTGACG
	R	CGCGGATCCGGTACCTCACCTCTTATGACGCTTAAAC
UL121	F	CCGGAATTCTCTCGAGGTTGGGGGTGCGGGTGGTCC
	R	CGCGGATCCGCGGCCGCCTATGATTTATAGCTTTTGGACAC
UL122	F	CCGGAATTCTTGTCGACCCTGCCCCTCATCAAACAGGAA
	R	CGCGGATCCGGTACCTTACTGAGACTTGTTCCTCAG
UL123.3	F	CCGGAATTCTTGTCGACCCCCGAGACACCCGTGACCA
	R	CGCGGATCCGGTACCTTAGAGTTCTGCCAGGACATCTTT
UL123.4	F	CCGGAATTCTTGTCGACCGTCAAACAGATTAAGGTTCGA
	R	CGCGGATCCGGTACCTTACTGGTCAGCCTTGCTTC
UL124	F	CCGGAATTCTTGTCGACCGAAAGGAACAGTCTGTTAGTCT
	R	CGCGGATCCGGTACCTTAAACATAGCGTGGGATCTC
UL125	F	CCGGAATTCTTGTCGACCTCCATAACATGGCTCTTTGC
	R	CGCGGATCCGGTACCTTAGGAGCAAGGAGCTGC
UL126	F	CCGGAATTCTTGTCGACCCGGTTTGACTCACGGGGAT
	R	CGCGGATCCGGTACCTCACCTATAGCATAAGGAAGC
UL127	F	CCGGAATTCTTGTCGACCTGCCAGCTTGATGTCGCCTC
	R	CGCGGATCCGGTACCCTAAAATGGGCTATATGCTGCA
UL129	F	GGGAATTCCATATGTTGTCGACCTTACACCTTCTGCACCCATC
	R	CGCGGATCCGGTACCCTACTCTGGCATTCCCGAG
UL130	F	CCGGAATTCTCTCGAGGTCTGCGGCTTCTGCTTCGTCA

	R	CGCGGATCCGCGGCCGCTCATGACGCGCGGTTTTCAAA
UL131	F	CCGGAATTCTTGTCGACCTGTATGATGTCTCATAATAAAG
	R	CGCGGATCCGGTACCTCAAAGTTGTCCAAGCCGT
UL132	F	CCGGAATTCTTGTCGACCCCAGCACCGCGGGGTCTCCTT
	R	CGCGGATCCGGTACCCTAGTCGTACTCGGGATCTCTGA
US12	F	GGGAATTCCATATGTTGTCGACCGTACAGATCCAGTTTCACCAAG
	R	CGCGGATCCGGTACCCTATAGGGCGAATTGGAGCT
US13	F	CCGGAATTCTTGTCGACCGACCCGCCGCTGCCGTCG
	R	CGCGGATCCGGTACCCTACGAGCCACCGCCACCTT
US14	F	CCGGAATTCTTGTCGACCGAGACAGTTTCCACGCAGC
	R	CGCGGATCCGGTACCTCAGGCAGCCTTGCTCTGG
US15	F	CCGGAATTCGAGATCTCTGCGACCGGTTCTCGCTCAG
	R	CGCGGATCCGGTACCCTACAGCTTGTCAGAGGAAAAG
US16	F	CCGGAATTCGAGATCTCTGGTCTGCGCTTTCCCACC
	R	CGCGGATCCGGTACCCTAGGGCGAGAGGGTGGA
US17	F	CCGGAATTCTTGTCGACCTCTCCGAACTCAGAGGCCAC
	R	CGCGGATCCGGTACCTTACGCCATGGTTCGCGTGA
US18	F	CCGGAATTCTTGTCGACCGGCGACACCGCCTCGGTTT
	R	CGCGGATCCGGTACCTTATAACAAGCTGAGGAGACTC
US19	F	CCGGAATTCTTGTCGACCCTTCATGTCGTCCCGCTAGAA
	R	CGCGGATCCGGTACCTCATGGGCTCCACAACCAGA
US20	F	CCGGAATTCTTGTCGACCCCTCGGTGGTGTGCGATAC
	R	CGCGGATCCGGTACCTTAGGACTTCCCCGTCGTACT
US21	F	CCGGAATTCTTGTCGACCTCGCTCCGCGGTCAGGTCC
	R	CGCGGATCCGGTACCTTAGGAGACGAACTGGGGCGA
US22	F	CCGGAATTCTTGTCGACCTCCCTACTCACCAAAGCCGC
	R	CGCGGATCCCTCGAGTTAGGGACCCGGGTCTGGTC
US23	F	GGGAATTCCATATGTCTCGAGGTTGGCGTACACGTTGGGAAGACG
	R	CGCGGATCCGCGGCCGCCTACACAAAGTGCTCCCGAAAATC
US24	F	GGGAATTCCATATGTCTCGAGGTATGGATCCGGCTGCGGGTTCT
	R	CCGGAATTCGCGGCCGCTCAAATCTGGATGTACTCGCGCA
US25	F	CCGGAATTCTTGTCGACCGCGACACTCTCAGCCGGACA
	R	CGCGGATCCGGTACCTCATCGGCACGTGCTCTTTTTTG
US26	F	GGGAATTCCATATGGGTACCCGCCAGTCCTATCGCTACG
	R	CCGGAATTCGGTACCTTACATCCAAAAGGGCAAGTAAG
US27	F	CCGGAATTCTTGTCGACCACTACATCTACTACAACTACCAC
	R	CGCGGATCCGGTACCTTACAACAGAAATTCCTCCTCC
US28	F	CCGGA TTCTTGTCGACCACACCGACGACGACGACC
	R	CGCGGATCCCTCGAGTTACGGTATAATTTGTGAGACG
US29	F	CCGGAATTCTCTCGAGGTCGGTGTTTCCGATGGTGGCT
	R	CGCGGATCCGCGGCCGCTTACTCGGAGGCGTCAACAAC
US30	F	GGGAATTCCATATGTCTCGAGGTGGAAATCCCCGGAGCCCGCT
	R	GGGAATTCCATATGGCGGCCGCTTATACAGGCGGAAACACAGAC

98

US31	F	CCGGAATTCGAGATCTCTAGAGACGGGAGGCTGATCG
	R	CGCGGATCCGGTACCTTAGTGTTGTTCACCCCGTGT
US32	F	CCGGAATTCGAGATCTCTGCTATGTACACATCCGAATCC
	R	CGCGGATCCGGTACCTTACACGTAAGTTTCCCGGTG
US33	F	CCGGAATTCTTGTCGACCGTTCATCGCCGTCGCATCG
	R	CGCGGATCCGGTACCTCAGAGAATCGCATAAGAACG
US34	F	CCGGAATTCTCTCGAGGTAACTTAGAGCAACTCATCAAC
	R	CGCGGATCCGCGGCCGCTCAGCATACCCCGGGCTC
US34A	F	CCGGAATTCTTGTCGACCCTGAAATTCTTCTTAAAATTAC
	R	CGCGGATCCGGTACCTCACGGAAAAGACTGCCGA
US35	F	CCGGAATTCTTGTCGACCGAACATCGACACACGACTG
	R	CGCGGATCCGGTACCTCATTTCTGGATGTATTTGTTAC

List of primer sequences used for cloning in yeast expression vectors. The 5' end of each primer contains a restriction site for cloning into the pGADT7 or pGBKT7 restriction site. The homology sequence to the ORF contains at least 18 nucleotides.

Supplemental Table 2-2 Complement primer sequences			
Human Complement Protein cDNA	**NCBI Ref Seq**	**Direction (F or R)**	**Primer Sequence**
CD55	NM_000574.3	F	CACCGACTGTGGCCTTCCCCCAGATGTACC (26 bp homology)
		R	CTAAGTCAGCAAGCCCATGGTTACTAGCGTCC (32 bp)
Factor H	BC142699	F	CACCGAAGATTGCAATGAACTTCCTC (22 bp homology)
		R	CTATCTTTTTGCACAAGTTGGATACTCCAG (30 bp)
MBL2	NM_000242.2	F	CACCGAAACTGTGACCTGTGAGGATGCCC (25 bp homology)
		R	TCAGATAGGGAACTCACAGACGG (23 bp)
MASP1	NM_139125.2	F	CACCCACACCGTGGAGCTAAACAATATGTTTGGCC (31 bp homology)
		R	TCACCGTTCCACCTGGGG (18 bp)
Properdin	BC015756	F	CACCGACCCCGTGCTCTGCTTCA (19 bp homology)
		R	AGAGTTCCTCTTCCTCAGGGTCTTTGCAAGC (31 bp)
C1qBP	BC000435	F	CACCATGCTGCCTCTGCTGC (16bp homology)
		R	CTGGCTCTTGACAAAACTCTTGAGGTC (27bp)
List of primer sequences used for cloning human complement protein cDNA into yeast expression vectors. The 5' end of each forward primer contains the CACC sequence to allow directional blunt end ligation of PCR products into pENTR/D-TOPO vectors (Invitrogen). pENTER/D-TOPO vectors were verified by restriction digest and sequencing prior to serving as a donor for Gateway recombination with pGADT7-GW (Addgene #61702). The final pGADT7 yeast expression plasmid encoding each human complement protein was verified by restriction digest and sequencing.			

Supplemental Table 2-3. HCMV-Complement mating matrix							
HCMV		**Human Complement Protein**					
Essential (E), Dispensable (D), Growth Defect (GD), Enhanced Growth (EG)	**Gene (BD)**	**Decay-Accelerating Factor (DAF/CD55)**	**Mannan Binding Lectin 2 (MBL2)**	**Properdin (CFP)**	**Factor H (FH)**	**Mannan Associated Serine Protease 1 (MASP1)**	**C1q Binding Protein (C1qBP)**
D^{Δ†}	RL1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	RL2	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	RL4	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	RL5	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	RL7	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	RL9	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	RL10	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
D^{Δ†}	RL11	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D[†]	UL1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD^Δ, D[†]	UL2	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL3	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL4	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL5	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL6	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL7	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL8	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
EG^Δ, D[†]	UL9	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ†}	UL10	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD^Δ, D[†]	UL11	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)

GD$^\Delta$, D†	UL12	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	UL13	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Negative (-, -)	Negative (-, -)
GD$^\Delta$, D†	UL14	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL15	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D†	UL15.A	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
D$^{\Delta\dagger}$	UL16	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
D$^{\Delta\dagger}$	UL17	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	UL18	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
D$^{\Delta\dagger}$	UL19	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^\Delta$, D†	UL20	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL20.A	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
GD$^\Delta$	UL21	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL22	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
EG$^\Delta$, D†	UL23	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
D$^{\Delta\dagger}$	UL24	Positive (+, +)	Positive (+, +)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Positive (+, +)
D$^{\Delta\dagger}$	UL25	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)
GD$^{\Delta\dagger}$	UL26	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
D$^\Delta$, GD†	UL27	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^{\Delta\dagger}$	UL28	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
GD$^{\Delta\dagger}$	UL29	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
GD$^{\Delta\dagger}$	UL30	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
GD$^\Delta$, D†	UL31	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
GD$^{\Delta\dagger}$	UL32	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	UL33	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

$E^{\Delta\dagger}$	UL34	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta}, D^{\dagger}$	UL35	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL36.1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$D^{\Delta\dagger}$	UL36.2	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
E^{Δ}	UL37.1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL37.3	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL38	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)
D^{Δ}	UL39	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL41	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$D^{\dagger}$	UL41.A	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL42	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL43	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Positive (+, +)	Negative (-, -)
$E^{\Delta\dagger}$	UL44	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$D^{\Delta\dagger}$	UL45	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL46	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL47	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$E^{\Delta}, GD^{\dagger}$	UL48	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL48.A	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
Unknown	UL48-N	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
Unknown	UL48-C	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
$E^{\Delta\dagger}$	UL48.5	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
$E^{\Delta\dagger}$	UL49	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)
$E^{\Delta\dagger}$	UL50	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL51	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator

$E^{\Delta\dagger}$	UL52	Positive (+, +)	Positive (+, +)	Positive (+, +)	Positive (+, +)	Positive (+, +)	Positive (+, +)
$E^{\Delta\dagger}$	UL53	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)
$E^{\Delta\dagger}$	UL54	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL55	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL56	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$E^{\Delta\dagger}$	UL57	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL58	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL59	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL60	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$E^{\dagger}$	UL61	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Negative (-, -)
D^{Δ}	UL62	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL63	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL64	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL66	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL67	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
Unknown	UL68	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL69	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$E^{\Delta\dagger}$	UL70	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta}, GD^{\dagger}$	UL71	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$GD^{\Delta\dagger}$	UL72	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$E^{\Delta\dagger}$	UL73	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL74	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL75	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta}, GD^{\dagger}$	UL76	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

$E^{\Delta\dagger}$	UL77	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL78	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$E^{\Delta\dagger}$	UL79	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL80	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL80-A	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL81	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Negative (-, -)
$GD^{\Delta\dagger}$	UL82	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)
$D^{\Delta\dagger}$	UL83	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL84	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$E^{\Delta\dagger}$	UL85	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$E^{\Delta\dagger}$	UL86	Positive (+, +)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)	Positive (+, +)
$E^{\Delta\dagger}$	UL87	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta}, D^{\dagger}$	UL88	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL89.1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL89.2	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
E^{Δ}	UL90	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL91	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL92	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)
$E^{\Delta\dagger}$	UL93	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta}, GD^{\dagger}$	UL94	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
$E^{\Delta\dagger}$	UL95	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta}, GD^{\dagger}$	UL96	Positive (+, +)	Positive (+, +)	Positive (+, +)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$GD^{\Delta\dagger}$	UL97	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL98	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

105

$E^{\Delta\dagger}$	UL99	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL100	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
Unknown	UL101	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL102	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$GD^{\Delta\dagger}$	UL103	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL104	Negative (-, -)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL105	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL106	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
Unknown	UL107	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD^{Δ}	UL108	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL109	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
D^{Δ}	UL110	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL111	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D^{Δ}	UL111-A	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL112.1	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL112.2	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL114	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$E^{\Delta\dagger}$	UL115	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	UL116	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$GD^{\Delta\dagger}$	UL117	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\dagger}$	UL118	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$D^{\Delta\dagger}$	UL119	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
$D^{\dagger}$	UL120	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
$D^{\Delta\dagger}$	UL121	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

E$^{\Delta\dagger}$	UL122	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^{\Delta\dagger}$	UL123.3	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
GD$^{\Delta\dagger}$	UL123.4	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^{\Delta}$, D†	UL124	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)
Unknown	UL125	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL126	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta}$	UL127	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
GD$^{\Delta}$	UL129	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	UL130	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	UL131	Negative (-, -)	Positive (+, +)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Negative (-, -)
GD$^{\Delta}$, D†	UL132	Positive (+, +)	Negative (-, -)	Positive (+, +)	Positive (+, +)	Positive (+, +)	Positive (+, +)
D$^{\Delta\dagger}$	US12	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^{\Delta}$, D†	US13	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US14	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US15	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US16	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US17	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US18	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US19	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US20	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US21	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
D$^{\Delta\dagger}$	US22	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
GD$^{\Delta\dagger}$	US23	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator	Auto activator
D$^{\Delta}$, GD†	US24	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

$D^{\Delta\dagger}$	US25	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$GD^{\Delta\dagger}$	US26	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
$D^{\Delta\dagger}$	US27	Positive (+, +)	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	US28	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	US29	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$EG^{\Delta}, D^{\dagger}$	US30	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	US31	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	US32	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Positive (+, +)	Negative (-, -)
D^{Δ}	US33	Negative (-, -)	Positive (+, +)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\Delta\dagger}$	US34	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
$D^{\dagger}$	US34-A	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)
Unknown	US35	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)	Negative (-, -)

Essentiality determined by gene deletion analysis and measuring effect on mutant HCMV growth *in vitro* as reported by:

Δ Functional profiling of the human cytomegalovirus genome (Towne) [70]

† Functional map of human cytomegalovirus AD169 defined by global mutational analysis [78]

Legend:

Positive (+, +): indicates positive protein interaction as determined by ≥2 independent yeast-two-hybrid mating experiments

Negative (-,-): indicates negative protein interactions as determined by ≥2 independent yeast-two-hybrid mating experiments

Autoactivator: indicates a non-specific autoactivating false-positive as determined by mating an HCMV protein expression vector with a cognate "empty" vector and observing blue colonies. Yeast carrying autoactivating plasmids were considered negative and not included in the final data analysis.

CHAPTER 3:

Investigating the humoral and cellular immune responses in mice to a highly conserved coronavirus Envelope (E) protein epitope

ABSTRACT

To date, all currently licensed SARS-CoV-2 vaccines target the Spike (S) protein as a primary vaccine antigen; however, there is an increasing emergence of SARS-CoV-2 isolates with spike protein mutations that confer resistance to vaccines, monoclonal antibody therapeutics, and natural immunity. Taken together, there is a significant unmet need to investigate the potential of additional antigens for possible inclusion in the next generation of SARS-CoV-2 vaccines. In the present study, we investigated the vaccine potential of a highly conserved 30 amino acid transmembrane domain in the SARS-CoV-2 Envelope (E) protein in mice. To increase immunogenicity, we conjugated the SARS-CoV-2 E peptide fragment to Keyhole Limpet Hemocyanin (KLH) and immunized mice with one or two doses of vaccine, administered intramuscularly in thirty-day intervals. The KLH-E conjugate vaccine elicited serum IgG antibodies and antigen-specific T-cells in mice against both the SARS-CoV-2 E peptide fragment, as well as the KLH carrier. The KLH-E vaccine elicited modest antibody titers to the SARS-CoV-2 E peptide fragment at one month post-immunization after a single dose but these titers decreased by 1.9-fold one month later (GMT= 87 at 30 days vs. GMT=46 at 60 days post-dose 1; P=0.1325). A second dose of the KLH-E vaccine increased antibody titers to the SARS-CoV-2 E peptide fragment but did not result in a significant increase in antibody compared to the same time point after a single dose (GMT=87 at 30 days post-dose 1 vs. GMT=91 at 30 days post-dose 2; P=0.05201). In contrast, the KLH-E vaccine elicited robust antibodies to the KLH carrier after a single dose, which increased 6.5-fold after a second dose (GMT=226 at 30 days post-dose 1 vs. GMT=1481 at 30 days post-dose 2; **P=0.0047). The KLH-E vaccine also elicited antigen-specific T-cells to both the KLH carrier and the SARS-CoV-2 E peptide fragment, though most of the cellular immunity elicited by the vaccine were directed at the KLH carrier. To our surprise, the anti-E IgG antibodies and T-cells of KLH-E mice recognize the native E protein from the genetically divergent but related human coronavirus HCoV-229E. Additionally, pooled sera from mice had neutralization activity against SARS-CoV-2 pseudovirions expressing S proteins from emerging variants of concern (Beta, Delta, Omicron, XBB), as well as neutralizing activity against HCoV-229E which shares only a modest 32% E protein sequence homology to the vaccine antigen. Taken together, these results implicate that immunization with the E protein transmembrane domain from SARS-CoV-2 can induce cross-reactive immune responses that may confer protection against emerging SARS-CoV-2 variants and other genetically diverse human coronaviruses.

INTRODUCTION

Coronaviruses. Coronaviruses are a highly diverse family of positive-sense single-stranded enveloped RNA viruses that infect many mammalian and avian species. The genus *Coronaviridae* was officially recognized in 1975 by David Tyrrell and colleagues after working with a number of human and animal viruses that had a distinct yet similar morphology by electron microscopy, with a characteristic crown-like appearance due to projections (spike glycoproteins) on the surface of the virions [1-3]. The family *Coronaviridae* is composed of four subfamilies: *alphacoronaviruses*, *betacoronaviruses*, *gammacoronaviruses*, and *deltacoronaviruses* [4]. Members in the *alphacoronavirus* and *betacoronavirus* subfamilies exclusively infect mammals, while members in the *gammacoronavirus*, and *deltacoronavirus* subfamilies have a much broader tropism that also includes avian species [5].

Seven coronaviruses are known to infect humans and include members from either the *alphacoronavirus* subfamily (HCoV-229E, HCoV-OC43, HCoV-NL63, HCoV-HKU1) or *betacoronavirus* subfamily (SARS-CoV-1, MERS-CoV, and SARS-CoV-2). The four coronaviruses that infect humans in the *alphacoronavirus* subfamily regularly circulate and are endemic in human populations globally, generally causing mild seasonal respiratory infections like that of the "common cold" [6]. Despite their endemicity, all four *alphacoronaviruses* that infect humans are believed to have zoonotic origins based on their phylogenetic relationship to other coronaviruses found in animals. Progenitors of HCoV-NL63 and HCoV-229E have recently been identified in African bats which are their presumed natural reservoir [7, 8], with camelids being a suspected intermediate host for HCoV-229E prior to infecting humans [9]. Coronaviruses HCoV-OC43 and HCoV-HKU1 are believed to have originated in rodents prior to infecting humans [10, 11].

Similarly, the three coronaviruses that infect humans in the *betacoronavirus* subfamily are all examples of recent emerging zoonotic viruses that have caused multi-country outbreaks (SARS-CoV-1 and MERS-CoV), or in the case of SARS-CoV-2 a global pandemic. Human coronaviruses in the *betacoronavirus* subfamily generally cause more severe disease than members in the *alphacoronavirus* subfamily and are all believed to have zoonotic origins based on their phylogenetic relationship to other coronaviruses found in wild animals. In particular, bat species appear to be important natural animal reservoir harboring SARS-like coronaviruses [12].

SARS-CoV-1. The first *betacoronavirus* to infect humans was SARS-CoV-1 during November 2003 in Guangdong province, China. The early cases of SARS-CoV-1 in Guangdong reportedly occurred in restaurant workers handling wild animals as exotic food [13]. Later, studies would identify SARS-like viruses in masked palm civets (*Paguma larvata)* at a live animal market in southern China that supplied animals for consumption in restaurants in Guangdong province [14]. The SARS-like virus isolated from masked palm civets had a 99.8% genomic sequence homology to the SARS-CoV-1 isolates in humans, and convalescent serum from infected humans was capable of restricting the growth of the SARS-like viruses (isolated from wild animals) in culture [14], suggesting that the two viruses were highly similar. Despite the high degree of genetic similarity, masked palm civets are believed to be the intermediate hosts, rather than the natural reservoir for SARS-CoV-1 due to three important lines of reasoning. First, progenitor

viruses have been identified in horseshoe bats (genus *Rhinolophus)* with greater sequence homology to SARS-like viruses isolated in civets compared to SARS-CoV-1. Second, progenitor coronaviruses in horseshoe bats show significantly more genetic diversity compared to those isolated from civets or SARS-CoV-1 in humans, which would be consistent with the evolutionary arms race between a virus and its natural reservoir [12]. Third, molecular studies of the Angiotensin Converting Enzyme 2 (ACE2) receptor- Spike (S) protein interaction indicate that progenitor coronaviruses in bats are unable to efficiently infect human cells and thus rapid viral evolution in an intermediate host may be necessary to adapt the virus for human infection [15]. Being highly transmissible among humans, SARS-CoV-1 rapidly spread to Hong Kong and other provinces across China and then to 28 countries [13, 16]. By July 2003, SARS-CoV-1 caused 8,096 confirmed cases of infection in 29 countries, with a case fatality rate of approximately 9.6% (http://www.who.int/csr/sars/country/table2004_04_21/en/). Interestingly, after 2004 there was no further transmission of SARS-CoV-1 and the virus disappeared as mysteriously as it first appeared.

MERS-CoV. On June 13, 2012 in Jeddah, Saudi Arabia, a novel coronavirus was isolated from a man with pneumonia, and would later be named MERS-CoV [17]. Like its predecessor SARS-CoV-1, MERS-CoV represented the second Coronavirus from the *betacoronavirus* subfamily to infect humans and caused a similar clinical manifestation. Unlike SARS-CoV-1 which was much more transmissible, MERS-CoV infections occurred in independent clusters and was mostly limited to countries in the Middle East, particularly in the Kingdom of Saudi Arabia. Later, multi-country outbreaks of MERS-CoV were reported in Africa, Europe, and in the United States, but were all epidemiologically linked by recent travel to the Middle East. The second largest outbreak of MERS-CoV occurred in the Republic of Korea in 2015, and was caused by a single person who travelled to the Middle East, resulting in 185 confirmed cases of MERS-CoV infection and 36 deaths [18]. Since its discovery in 2012, across the globe 27 countries have reported cases of MERS-CoV causing 2,519 confirmed cases of infection with a case fatality of 34.3% (https://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html), a case-fatality rate over three times higher than that of SARS-CoV-1. The progenitor of MERS-CoV is believed to have originated in bats [19-21], before infecting dromedary camels [22-24] and entering humans. Unlike SARS-CoV-1 which has not been detected in humans since 2004, a small number of isolated primary cases of MERS-CoV have occurred even as recently as April 2022, and all but one case reported recent contact with camels (Saudi Arabia= 3 cases, Qatar=2 cases, Oman=1 case) (https://www.ecdc.europa.eu/en/middle-east-respiratory-syndrome-coronavirus-mers-cov-situation-update).

SARS-CoV-2 and the COVID-19 pandemic. During December 2019, a widespread outbreak of pneumonia of unknown cause was reported in Wuhan city, China. During this time, the etiological agent, precise modes of transmission, and the communicability between humans was incompletely understood, though many people infected were epidemiologically linked to a large seafood and wet animal market in Wuhan city in the Hubei province [25]. Due to the epidemiological link with the seafood and wet animal market, it was highly speculated that the outbreak of pneumonia was due to a zoonotic pathogen, as this had been the setting previously for SARS-CoV-1, the virus originating in animal reservoirs (civet cats) which caused the SARS outbreak in 2003 in Guangdong, China [26-28]. The viral agent responsible for the pneumonia outbreak in Wuhan, China, was eventually identified as SARS-CoV-2, a novel Coronavirus from

the *betacoronavirus* subfamily that had never been previously identified in humans. As the number of SARS-CoV-2 infections increased and crossed continents, nations struggled to respond on how best to limit transmission -with some imposing draconian lockdown measures [29], while others took the opposite approach and hoped (but ultimately failed) to achieve a meaningful threshold of population level immunity [30]. On March 11, 2020 after more than 118,000 cases in 114 countries and 4,291 deaths, the World Health Organization (WHO) officially declared a global COVID-19 pandemic [31]. As of the date of publication of this dissertation, there have been 768,983,095 confirmed infections and 6,953,743 deaths globally due to SARS-CoV-2 according to the WHO. Importantly, these significant figures are likely an underestimate of the true number of cases and deaths due to SARS-CoV-2.

BIOLOGY OF SARS-COV-2

The SARS-CoV-2 genome. Members of the family *Coronaviridae* have large positive-sense single-stranded RNA genomes ranging from 25,000-32,000 nucleotides, representing one of the largest genomes of all RNA virus families [32]. The genome of SARS-CoV-2 is approximately 29,903 nucleotides [33] and encodes 29 proteins [34]. Upon infection the positive-sense genome directly translates two polyproteins from the ORF1a and ORF1b regions of the genome, which are then cleaved by two viral proteases (Papain-like protease PLpro or nsp3, and 3-chymotrypsin-like protease 3CLpro or nsp5). The organization of the SARS-CoV-2 genome is illustrated in **Figure 3-1** compared to the genomes of the other seven coronaviruses that also infect humans, with significant overlap in both protein sequence homology and functions across different coronavirus lineages.
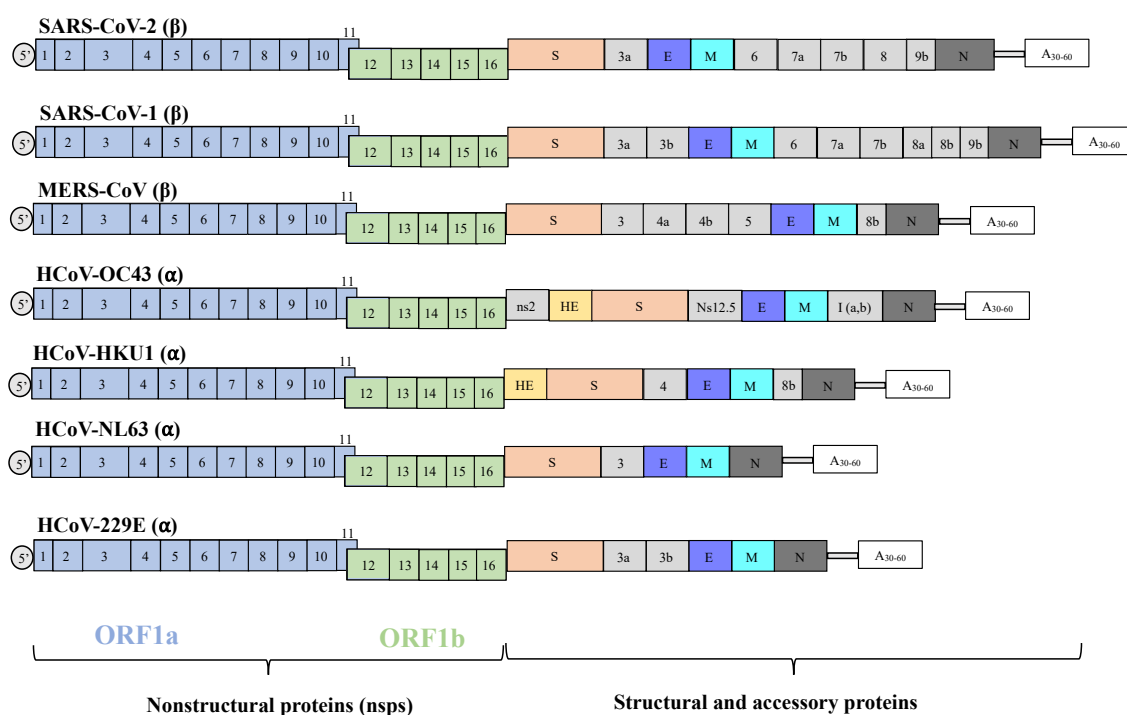

Figure 3-1. Genome structure of human coronaviruses

Genome structure and coding potentials of human coronaviruses. The viral genome is a single-stranded, positive-sense RNA with a cap (grey circle) at the 5′-end and a poly-A tail (A30-60) at the 3′ end. The genome encodes 16 non-structural proteins (ORF1a → nsp1-11 and ORF1b → nsp12-16) from the left three-fourth of the genome, and 4–5 structural proteins (S, spike; E, envelope; M, membrane; N, nucleocapsid; HE, hemagglutinin esterase) and various number of accessory proteins (numbered boxes) from the right one-fourth of the genome.

Nonstructural proteins. SARS-CoV-2 produces 16 nonstructural proteins (Nsp1-16) which vary in size from 13 amino acid residues (Nsp11) to 1299 amino acid residues (Nsp3) [35] and perform various important functions during infection. Nsp1 inhibits translation of host proteins [36], Nsp2 interacts with host proteins (prohibitins) involved in mitochondrial function and intracellular signaling [37, 38], Nsp3 is a papain-like protease [21] along with Nsp5 which is a 3C-like protease [39], with both involved in viral and host peptide processing. Nsp4 and Nsp6 form double membrane vesicles associated with replication/transcription complexes [40], Nsp7 and Nsp8 function as accessory factors of the viral encoded RNA-dependent RNA polymerase Nsp12. Together, Nsp9, Nsp12, and Nsp13 form the coronavirus viral replication transcription complexes [41]. Nsp10 functions as a cofactor of Nsp14, a protein involved in viral RNA replication and suppression of innate immunity [42]. Nsp16 also contributes to the suppression of innate intracellular antiviral immunity by specifically interfering with cytosolic RNA sensing [43]. The precise biological function of Nsp11 is unknown, while Nsp15 is a uridine-specific endoribonuclease and interferon agonist [44].

Structural proteins. SARS-CoV-2 encodes 4 major structural proteins: the Spike (S) protein, Nucleocapsid (N) protein, Membrane (M) protein, and the Envelope (E) protein, all of which are required to produce an infectious viral particle. The S protein mediates attachment of the virus to the host cell surface receptor ACE2 resulting in fusion and viral entry. The M protein is the most abundant protein, which defines the shape of the viral envelope and serves an essential role in virus assembly and morphogenesis [45]. The N protein is a multifunctional protein that mediates several important processes, including serving as an RNA-binding substrate critical for viral genome packing [46]. Lastly, the E protein is the smallest structural protein on the SARS-CoV-2 virion and appears to be polyfunctional and critical for the SARS-CoV-2 viral lifecycle.

SARS-CoV-2 Viral replication. As illustrated in **Figure 3-2**, SARS-CoV-2 viral replication begins with attachment of mature virions to host cells mediated by the Spike (S) glycoprotein and the human angiotensin-converting enzyme 2 (ACE2) receptor. Upon receptor binding, the S protein undergoes a conformational change and can activate fusion or endocytosis of the virus with the cell membrane after being cleaved by cell surface proteases. Following fusion, genomic positive single strand RNA is released inside the cytoplasm of the cell and replicated to form two RNA strands, one which serves as a template for replication of viral genomic RNA, and a second strand which is translated into the four major structural proteins (N, S, M, E). The N protein is translated in the cytoplasm by ribosomes, while the S, M, and E proteins are translated in the lumen of the Endoplasmic Reticulum (ER). Of note, during viral replication the E protein is highly expressed in infected cells, yet only a small amount of the total E protein produced is found in assembled virions [47]. Following synthesis of all four structural proteins, the N protein assembles with the positive-strand genomic RNA and assembly of virions occurs at the transitional zone between the ER and Golgi-apparatus. Subsequent viral maturation occurs as the virions are assembled and bud off from the Golgi membrane as vesicles, which are translocated to the host cell membrane where they fuse and are released via exocytosis [4, 48] .
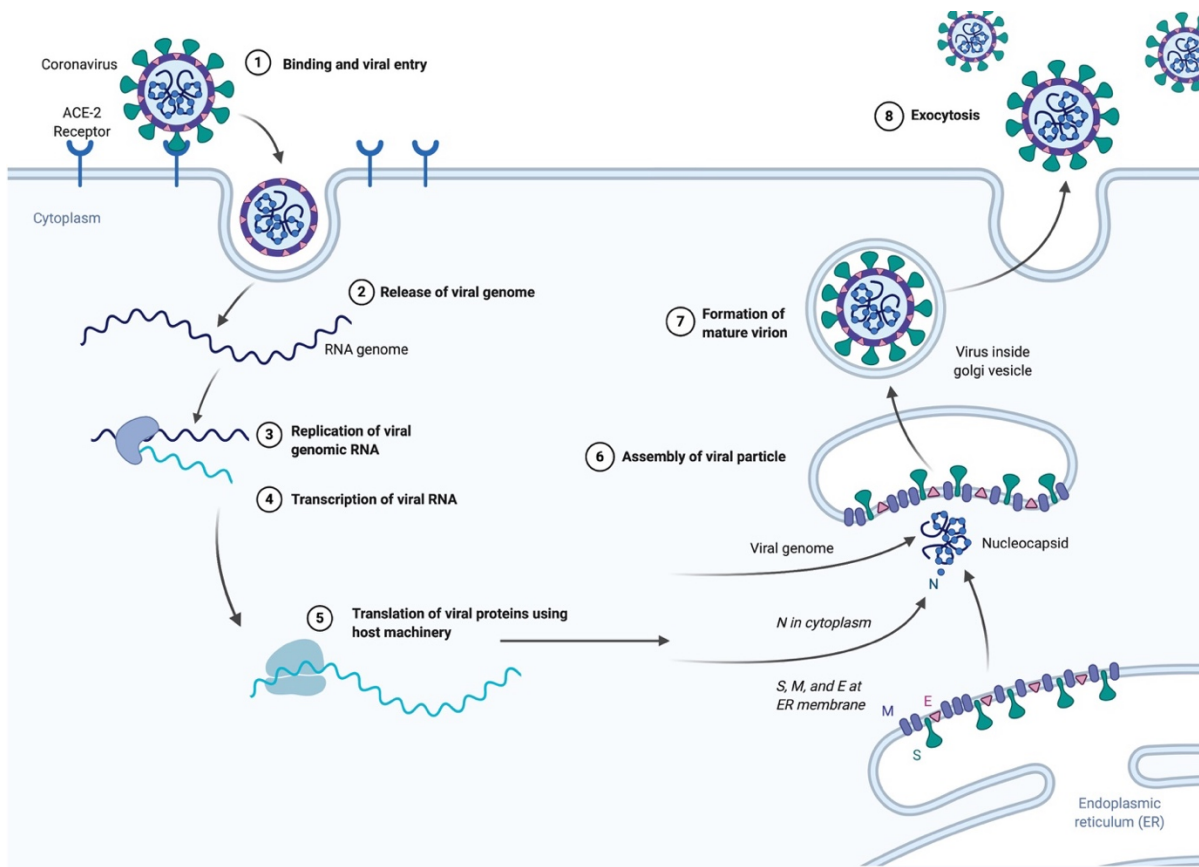
Figure 3-2. SARS-CoV-2 viral replication. The life cycle of SARS-CoV-2 [48]. Upon binding to the host membrane receptor ACE2 by the viral Spike (S) protein, the SARS-CoV-2 virion enters the host cell and releases its positive sense genome. The positive strand RNA translates pp1a and pp1ab, which are cleaved into multiple non-structural proteins (Nsps) including an RNA-dependent RNA polymerase (Nsp12). The RNA-dependent polymerase transcribes a negative-strand genomic RNA which is used as a template to generate more positive-strand genomic RNA (viral genomic replication) in addition to other protein coding RNA's. The protein coding RNA's are translated into the four major structural proteins (N, S, M, E) which assemble with the positive strand genomic RNA to form a mature virion in the at the transitional zone between the ER and Golgi-apparatus. Mature virions then exit the cell via exocytosis. Figure designed with the assistance of Biorender software (www.biorender.com).

SARS-CoV-2 ENVELOPE (E) PROTEIN

Envelope protein molecular architecture. Of the four SARS-CoV-2 structural proteins the E protein is the smallest, most sparse, and least characterized of all structural proteins in the mature virion. The molecular architecture of the 75 amino acid SARS-CoV-2 E protein is composed of three distinct protein domains: a short hydrophilic amino acid N-terminus consisting of 7-12 amino acids, followed by a large hydrophobic transmembrane domain of 25 amino acids, followed by a long hydrophilic carboxyl C-terminus which compromises the majority of the protein [49] . In SARS-CoV-2, the E protein is polyfunctional and is associated with assembly, budding, and envelope formation in the viral replication cycle based on its protein-protein interactions with other viral structural and accessory proteins [50], and high (95%) sequence homology to the SARS-CoV-1 E protein which has been more extensively studied [47].

Coronavirus E protein is a viroporin. The coronavirus E protein contains structural features common to the viroporin class of proteins. Viroporins are a broad class of virally encoded small (60-120 amino acid) proteins with a distinct hydrophobic transmembrane domain, which intimately interacts with membrane surfaces to produce a pore with selective ion and small molecule specificity. Viroporins have been identified in many medically relevant and highly pathogenic human viruses including: Human Immunodeficiency Virus (HIV) [51], Hepatitis C Virus (HCV) [52], Rotavirus (RV) [53], Poliovirus (PV) [54], and Influenza A Virus (IAV) [55]. Via the function of viroporins, viruses can modulate ionic gradients, pH, membrane vascularization, and cell permeability. Perturbations in the homeostasis of these systems can affect diverse cellular processes including protein trafficking, signal transduction, apoptosis, or contribute synergistically to produce an optimal cellular environment for viral replication.

The viroporin activity of the envelope protein has been documented in human Coronaviruses including: SARS-CoV-1 [56-59], MERS-CoV [60], SARS-CoV-2 [61], HCoV-229E [62], and HCoV-OC43 [63]. Although the E protein is encoded by all human coronaviruses, some encode accessory proteins with predicted viroporin activity based on their structures SARS-CoV-1 (ORF3a), HCoV-229E (4a), and HCoV-OC43 (ns12.9), SARS-CoV-2 (ORF3a) [64]. This redundancy in encoding multiple viroporin-like proteins may partially explain why the E protein is dispensable in some but essential in other human coronaviruses (SARS-CoV-2). Nevertheless, the E protein viroporin activity appears to serve an important role in coronavirus pathogenesis.

The transmembrane domain is essential for E protein activity. In all coronaviruses, the transmembrane domain of the E protein is specifically required to form higher order structures and functional viroporin activity via homotypic-oligomeric interactions [65]. The SARS-CoV-2 E protein differs in only three amino acid substitutions (T55S, V56F, E69R) and one deletion (G70-del) compared to the E protein sequence of SARS-CoV-1, but both share complete sequence identity in the transmembrane region. The E protein transmembrane domains for both SARS-CoV-1 and SARS-CoV-2 are reported to oligomerize *in vitro* and form a pentameric structure with a central pore that serve as a cationic selective channel [56, 62, 66-68]. The selective cationic permeability of the E protein however can be influenced by specific ions (K^+, Na^+, Ca^{2+}, and Mg^{2+}) [56, 69], membrane composition [62], and pH conditions [70].

E protein viroporin activity is related to virulence. In SARS-CoV-1 the introduction of mutations (N15A, V25F) in the transmembrane domain of the E protein that ablate viroporin activity reduces viral fitness *in vitro* [71] and strongly attenuates pathogenesis *in vivo* in animal models [72, 73]. Although uncommon, SARS-CoV-1 isolates from humans with mutations in the E protein transmembrane domain have been identified, and are also highly attenuated and result in reduced viral titer *in vivo* [74]. The importance of the E protein viroporin activity in coronavirus pathogenesis is underscored by the observation that introduction of mutations or deletions in the SARS-CoV-1 transmembrane domain are restored by additional compensatory mutations after several cell passages *in vitro*, implicating a strong evolutionary selective pressure to maintain viroporin function [73].

Membrane gradients within cells and subcellular compartments require specific ion compositions and concentrations to maintain normal cellular homeostasis. When these systems are perturbed by viroporins in the context of viral infection, cellular innate immune effectors are engaged. Disruptions in calcium homeostasis are commonly observed after viroporin expression during viral infections, causing Ca^{2+} to leak from intracellular stores such as mitochondria, ER, and Golgi complexes into the cytoplasm [64]. Ca^{2+} can act as a cytosolic signal and regulate multiple cellular processes and even induce apoptosis at elevated levels [75]. For both SARS-CoV-1 and SARS-CoV-2, the E protein transmembrane domain forms a calcium Ca^{2+} selective ion channel that mediates the activation of the NLRP3 inflammasome and triggers the overproduction of proinflammatory cytokine IL-1ß [64, 73, 76, 77]. In turn, IL-1ß can activate the release of other proinflammatory cytokines such as TNF-α and IL-6 which are primary mediators of the systemic acute phase response [78]. Elevated inflammatory cytokines IL-1ß, TNF-α, and IL-6 are a hallmark of acute respiratory distress syndrome (ARDS) [79] that cause pulmonary and multiorgan damage, and are and a major cause of SARS-CoV-2 related death [80, 81].

Even in the absence of infection, ectopic expression of the SARS-CoV-2 E protein can induce cell death in many cultured cell lines (HeLa, CaCO-2, Vero E6, 16HBE) via an atypical pyroptosis-like mechanism that remains poorly understood. Similar cell death is also observed when the SARS-CoV-1 E protein (95% identical; 100% transmembrane domain identical) was expressed in Vero E6 cells [63]. Interestingly, the ectopic expression of the E protein from HCoV-OC43, a related human coronavirus that causes mild seasonal respiratory infections, also caused cell death in Vero E6 cells but required much greater expression [63].

E protein vaccine potential. To date, all licensed SARS-CoV-2 vaccines include the S protein as the primary vaccine antigen. Antibodies directed at the S protein, particularly the receptor binding domain (RBD), can inhibit the engagement of the S protein SARS-CoV-2 viral receptor with the cognate ACE2 host cell receptor and thereby prevent infection. Unfortunately, the SARS-CoV-2 S protein is extremely mutable, and the emergence of several SARS-CoV-2 variants have already reduced or eliminated the efficacy of monoclonal antibody therapeutics, currently licensed vaccines, and natural immunity. Breakthrough infections in vaccinated persons remain increasingly common, though generally result in more mild disease. While the mRNA vaccine platform offer the promise of rapid vaccine reformulations to keep up with emerging variants, production costs and the long clinical trial timelines necessary for the authorization of new vaccine formulations have undermined their utility. The need for booster doses, recommended dose intervals, mixing-and-matching of different vaccines, among other factors have contributed to poor vaccine uptake by much of the population.

In contrast to other SARS-CoV-2 antigens, largescale genomic studies of publicly available viral genomes have demonstrated that the E protein remains highly conserved and therefore represents an attractive vaccine antigen with the potential to confer broad protection against SARS-CoV-2 variants [82]. Furthermore, since the E protein is moderately conserved across various human coronaviruses, the induction of humoral and cellular immune responses to this conserved antigen has the potential to theoretically elicit pan-coronavirus immunity.

Objective. The objective of the current study is to (1) test an unusual and novel vaccine concept with a highly hydrophobic transmembrane peptide as the immunogen, and (2) develop a new vaccine against SARS-CoV-2, the causative agent of COVID-19. The antigen used in our study is a 30-amino acid long peptide corresponding to the transmembrane region of the E protein from SARS-CoV-2. The E protein is essential for SARS-CoV-2 replication and the transmembrane domain is moderately conserved among human coronaviruses. We will test the immunogenicity of the 30 amino acid domain in eliciting effective immune responses in mice against SARS-CoV-2 pseudovirions and other related human coronaviruses to determine if the host immune system can recognize and mount an effective immune response to a completely occluded transmembrane peptide.

RESULTS

Bioinformatic analysis. Using a bioinformatic approach, we first aligned the amino acid sequences of the full-length E proteins from all seven coronaviruses known to infect humans. The E protein is composed of an external region accessible at the surface of the virion, a transmembrane region, and an intravirion region. As illustrated in **Figure 3-3 (bottom)**, the N-terminal region of the E protein is predicted to be accessible at the surface of the virion and shares some sequence homology across all seven human coronaviruses with notable differences in HCoV-OC43 which has an insertion at position 3 (M; Methionine) and 4 (N; Asparagine) of the primary amino acid sequence that is missing from other similar coronaviruses, even within the same *alphacoronavirus* subfamily. Additionally, HCoV-HKU1 and HCoV-OC43 both have a conserved tryptophan (W) residue at amino acid position 13 that is absent in the related *alphacoronaviruses* HCoV-229E and HCoV-NL63. With respect to the C-terminal intravirion region of the E protein, we also observed many gaps within the aligned primary amino acid sequences with limited punctuated amino acid positions with some sequence conservation or similar amino acid properties. In contrast, we observed the highest degree of sequence conservation and amino acid homology in the transmembrane region of the E protein of all seven coronaviruses, which was surprising given the phylogenetic relationship of the E protein amino acid sequences which clustered according to *alphacoronavirus* and *betacoronavirus* subfamilies (**Figure 3-3, top**).

```
229E       MF--LKLVDDHA-LVVNVLLWCVVLIVILLVCITIIKLIKLCFTCHMFCNRTVYGPIKNVYHIYQSYM-----HIDPFPKRVIDF
NL63       MF--LRLIDDNG-IVLNSILWLLVMIFFFVLAMTFIKLIQLCFTCHYFFSRTLYQPVYKIFLAYQDYM-----QIAPVPAEVLNV
HKU1       MV--DVFFTDTAWYVGQIFFLVLSCVIFLIFVVALLATIKLCIQICGFCNIFIISPSAYVYNRGRQLYKSYSEHVIPSTLD-DLI
OC43       MFMADAYLADTVWYVGQIIFIVAICLLVTIVVVAFLATFKLCIQLCGMCNTLVLSPSIYVFNRGRQFY-EFYNDVKPPVLDVDDV
MERS-CoV   ML--PFVQERIGLFIVNFFIFTVVCAITLLVCMAFLTATRLCVQCMTGFNTLLVQPALYLYNTGRSVYVKFQDSKPPLPPD-EWV
SARS-CoV-1 MY--SFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPTVYVYSRVKNL------NSSEGVPD-LLV
SARS-CoV-2 MY--SFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYVYSRVKNL-------NSSRVPD-LLV
           *                :  :  .:          .  :. ::::    .**        . :  *    ::    ..                  .
```

```
Virion Surface
Transmembrane
Intravirion
Legend: (-)gaps, . or :(some conservation, indicate different properties of AA's at these locations)
```

Figure 3-3. Phylogenetic relationship and amino acid sequence alignment of human coronavirus Envelope (E) proteins. (Top) Neighbor-joining phylogram of envelope protein amino acid sequences from seven coronaviruses known to infect humans: SARS-CoV-1 (NC_004718), MERS-CoV (YP_009047209), SARS-CoV-2 (YP_009724392), 229E-CoV (NP_0735554), NL63-CoV (YP_003769), OC43-CoV (YP_009555243), HKU1-CoV(YP_173240) computed with MEGA 11 [83]. The test of phylogeny was performed using the Bootstrap method with 100 bootstrap replications. The values at each node of the phylogenetic tree represent the bootstrap confidence value. The scale bar indicates 20% amino acid sequence divergence. (Bottom) Multiple sequence alignment of the amino acid sequences of the seven coronaviruses known to infect humans. The regions of the proteins are labelled corresponding to their relative predicted location: virion surface (gray), transmembrane (yellow), intravirion (white).

Vaccine design and immunization. We observed the highest degree of sequence conservation and homology concentrated in the transmembrane domain of the coronavirus E protein as illustrated in **Figure 3-3**. The observation that the transmembrane domain of the E protein was moderately conserved across genetically diverse human coronaviruses prompted us to design a prototype vaccine to investigate whether it was possible to elicit immunity to this conserved region. Previous immunoinformatic studies predicted the presence of multiple B and T-cell epitopes, and possible MHC binding sites on the SARS-CoV-2 E protein despite its relatively small size at only 75 amino acids [84]. However, we were skeptical whether a synthetic peptide alone corresponding to the hydrophobic 30 amino acid (TLIVNSVLLFLAFVVFLLVTLAILTALRLC) SARS-CoV-2 transmembrane domain would be capable of inducing an immune response, even though the amino sequence was hypothetically long enough to be recognized by MHC class I and MHC class 2 molecules which generally require a minimum of 8-10 and 13-25 amino acids for immune presentation respectively.

To enhance immunogenicity, we had the SARS-CoV-2 E protein transmembrane domain sequence synthesized and conjugated to the immunogenic carrier protein Keyhole Limpet Hemocyanin (KLH), producing an antigen we herein refer to as KLH-E. The KLH carrier is a very large, copper-containing protein derived from the hemolymph of the mollusk *Megathura crenulata* and is highly immunogenic to the mammalian immune system. KLH is also a T-cell dependent antigen that is non-toxic and elicits robust humoral and cellular immune responses when administered via a number of different routes (intramuscular, intradermal, subcutaneous, etc.) [85]. To investigate the immunogenicity of our vaccine, we vaccinated groups of inbred mice with one or two doses (50µg per dose) of KLH-E or control vaccine (aluminum) intramuscularly at 30 day intervals as described in the materials and methods. After the end of the vaccination study, mice were terminally bled and their serum and spleens collected and analyzed to assess the immunogenicity of the KHL-E vaccine.

Pilot study. We carried out a pilot experiment to assess and characterize the humoral responses to the antigens used for the immunization. Specifically, sera was pooled from each group of mice and was tested for reactivity against the KLH-E conjugate and the 30 amino acid SARS-CoV-2 E protein transmembrane domain. Three serum pools were from mice immunized with the KLH-E vaccine collected at 30 days post-dose 1 (n=8 mice), 60 days post-dose 1 (n=8 mice), or 30 days post-dose 2 (n= 8 mice). Three serum pools were from mice immunized with the control aluminum vaccine collected at 30 days post-dose 1 (n= 5 mice), 60 days post-dose 1 (n=4 mice), or 30 days post-dose 2 (n=5 mice).

As illustrated in **Figure 3-4**, by ELISA the titers of the IgG antibodies specific to the KLH-E antigen were reduced proportionally with dilution folds of the sera obtained from each group of mice receiving either one or two doses of KLH-E vaccine. As expected, serum pools from mice immunized with two doses of vaccine yielded higher IgG KLH-E antibody titers when compared to pools from mice immunized with a single dose of vaccine.
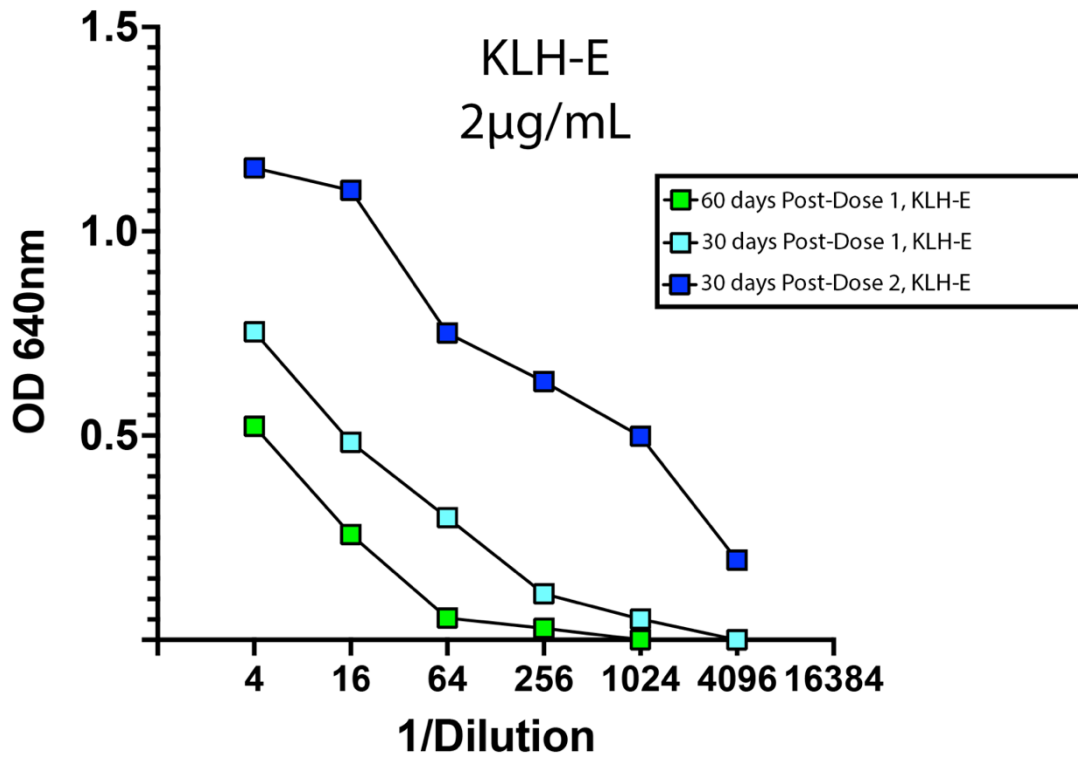
Figure 3-4. Antibody response of pooled mouse sera measured against KLH-E. Pooled sera from groups of mice immunized with the KLH-E conjugate (squares) were combined in equal parts and measured for IgG antibodies to the KLH-E vaccine antigen. Mice were given one or two doses of vaccine intramuscularly in 30 day intervals. Each sample was tested in replicate in two independent assays.

To assess whether the KLH-E vaccine elicited anti-E IgG antibodies, we repeated the experiments using the same serum pools tested against the 30-amino acid transmembrane domain of the E protein as the coating antigen in the ELISA. We observed similar results, with the titers of the IgG antibodies specific to the E-peptide transmembrane domain reduced proportionally with dilution folds of the sera obtained from each group of mice receiving either one or two doses of vaccine. Similarly, the anti-E IgG antibody titers in pooled sera were higher in mice that received two doses of vaccine compared to a single dose at one month post-immunization (titer of 64 vs. 33, respectively) as illustrated in **Figure 3-5**.
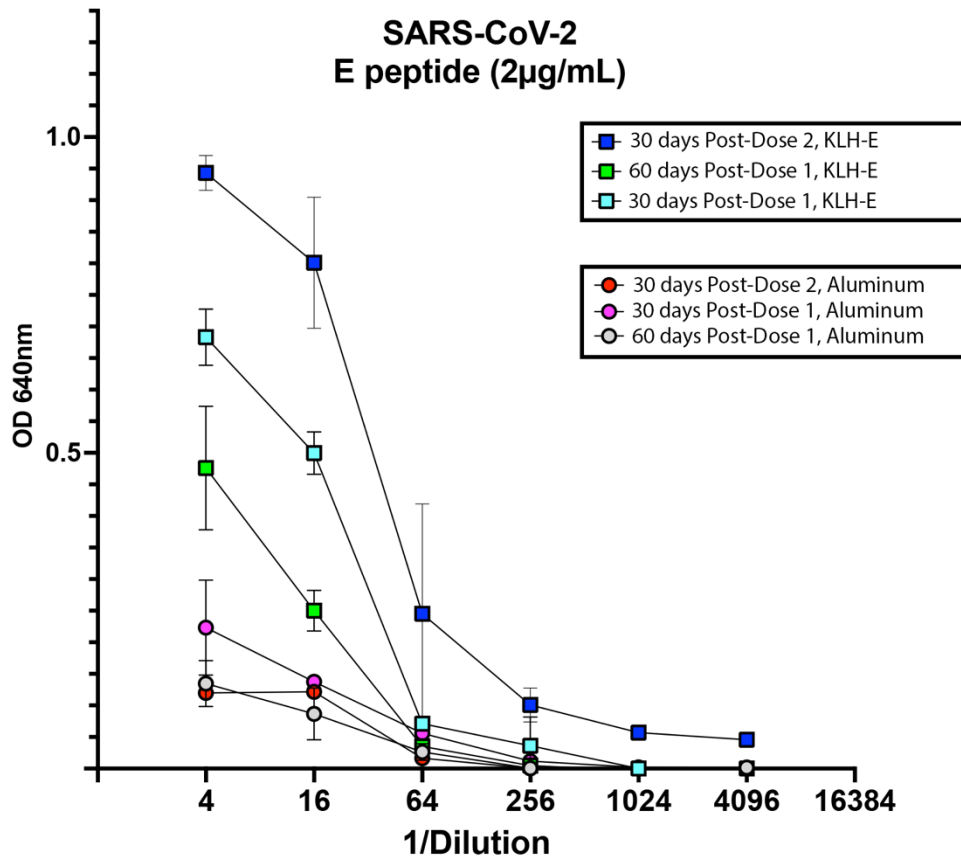
Figure 3-5. Antibody response of pooled mouse sera measured against E-peptide.
Pooled serum from groups of mice immunized with the KLH-E (squares) or aluminum control vaccine (circles) were combined in equal parts and measured for IgG antibodies to the SARS-CoV-2 E transmembrane domain peptide. Mice were given one or two doses of vaccine intramuscularly in 30 day intervals. Error bars at each dilution point represent the +/- standard deviation of each sample tested in replicate in two independent assays.

The results of this pilot experiment suggest that mice immunized with the KLH-E vaccine elicited normal antibody responses with IgG titers corresponding to the dose of immunization. Additionally, we observed that mice immunized with the KLH-E vaccine generated IgG antibody responses to the vaccine antigen KLH-E as well as generating specific anti-E IgG antibody responses. The humoral responses we observed in our ELISA experiments were also specific to vaccination, as none of the pooled sera from mice immunized with the aluminum control vaccine had KLH-E or E-peptide specific IgG antibodies above background in our ELISA assay. While our preliminary study provided important insight into the immunogenicity of our KLH-E vaccine, it should be noted that some serum pools included more mice than others. Although serum pools are generally helpful in assessing trends in antibody responses, individual animal titers are always preferred since different volumes of serum used to prepare the pools as well as high/low responding mice can skew and potentially confound interpretation of antibody responses. To investigate further the humoral responses elicited by KLH-E vaccination in mice, we quantified the IgG antibody responses of individual mice in each vaccine group against: KLH-E (the vaccine antigen), KLH (the immunogenic carrier), and the SARS-CoV-2 E-peptide transmembrane domain by ELISA.

Antibody response to KLH-E vaccine. As illustrated in **Figure 3-6**, against the KLH-E antigen, mice immunized with a single dose the KLH-E vaccine had an IgG antibody geometric mean titer (GMT) of 220 (range: 159-358) at 30 days post-immunization. A small, but significant decrease in KLH-E IgG titer was observed in mice immunized with a single dose of the KLH-E vaccine at 60 days post-immunization (GMT: 220 vs. 132, 30 days vs. 60 days post-dose 1; *P=0.0194). The 0.6-fold reduction in geometric mean titers at 30 days compared to 60 days post-immunization was not surprising, as it is widely understood that antibody levels induced by vaccination ultimately decline over time irrespective of the vaccine, especially after a single dose. Despite the waning antibody titers, a second dose of KLH-E vaccine elicited a robust 7.2-fold significant increase in geometric mean titer in mice at 30 days post-immunization (GMT: 220 vs. 1588, 1 dose vs. 2 doses at 30 days; **P=0.0060). In contrast to the robust IgG antibody responses observed in KLH-E vaccinated mice, mice immunized with one or two doses of the control aluminum vaccine had no detectable IgG antibodies to KLH-E as expected. Since the IgG antibody responses measured against the KLH-E vaccine antigen represent a mixture of antibodies to KLH as well as to the SARS-CoV-2 E peptide transmembrane domain, we proceeded to measure the IgG antibody responses of individual mice to each individual subunit of the vaccine to delineate the relative contribution of each in the overall humoral response elicited by the KLH-E vaccine.

Figure 3-7 depicts the KLH carrier IgG antibody responses of individual mice induced by the KLH-E vaccine. The geometric mean KLH IgG titer of mice immunized with a single dose of KLH-E was 226 at 30 days post-immunization. We observed a similar small, but significant decrease in the KLH specific IgG titer of mice immunized with a single dose of KLH-E vaccine at 60 days post immunization (GMT: 226 vs. 137, 30 days vs. 60 days post-dose 1; *P=0.0227). A second dose of KLH-E vaccine elicited a 7.2-fold increase in KLH specific geometric mean IgG antibody titers at one month post-immunization (GMT: 226 vs. 1481, 1 dose vs. 2 doses at 30 days; **P=0.0047). No KLH IgG responses were detected in mice immunized with the control aluminum vaccine. The highly similar trends in the geometric means and individual IgG antibody titers to KLH-E and KLH in vaccinated mice, led us to infer that most of the antibodies

induced by vaccination were directed at the immunogenic carrier KLH. These results are unsurprising, considering that the SARS-CoV-2 E peptide is only 30 amino acids in length with an approximate molecular weight of 3.29 kD compared to the 3,414 amino acid KLH carrier with a molecular weight of 370 kD. Since the SARS-CoV-2 E peptide is the primary antigenic target of our immunogenicity study, we quantified the anti-E specific IgG antibody responses of individual mice immunized with the KLH-E vaccine.
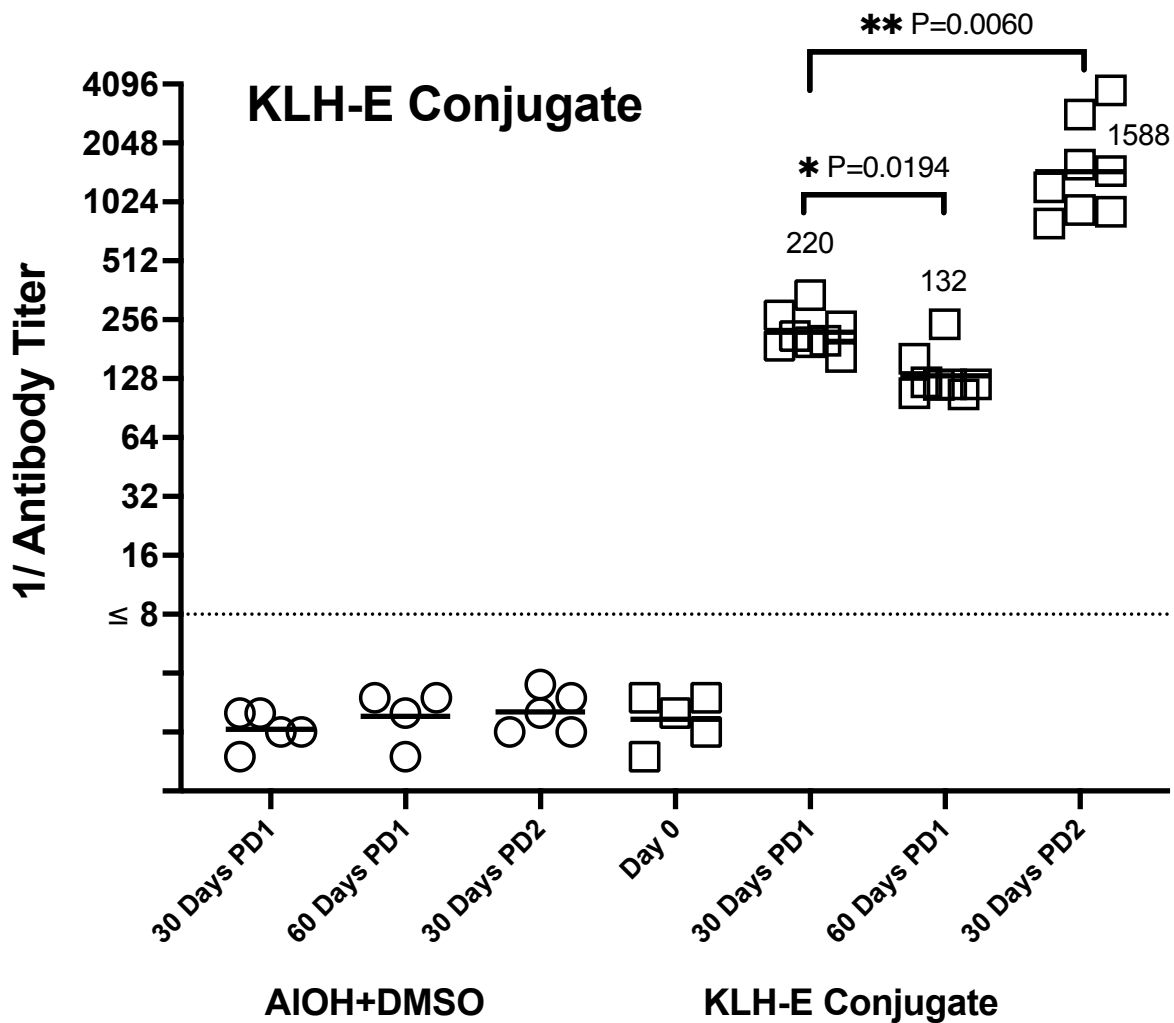
Figure 3-6. IgG antibody titers of individual measured against KLH-E.
Serum from individual mice immunized with the KLH-E conjugate (squares) or aluminum control vaccine (circles) measured for IgG antibodies against KLH-E. The responses of mice illustrated in the figure represent antibodies to both the KLH carrier, as well as the E-peptide fragment. Mice were given one or two doses of vaccine intramuscularly in 30 day intervals. Each symbol represents the reciprocal antibody titer of individual animals in each vaccine group, with each data point representing the average antibody titer enumerated in duplicate, in two independent assays. Mice with no detectable IgG to the KLH-E conjugate were assigned a titer of one-half the lower limit of the assay (i.e. a titer of <1:8 was assigned as 4). Numbers above each group represent the geometric mean titer (GMT) IgG antibody titer for all animals in the respective vaccine group. Differences between vaccine groups were assessed using a two-tailed students *t* test of the log-transformed reciprocal titers. Differences with a probability of <0.5 (two-tailed) were considered significant.
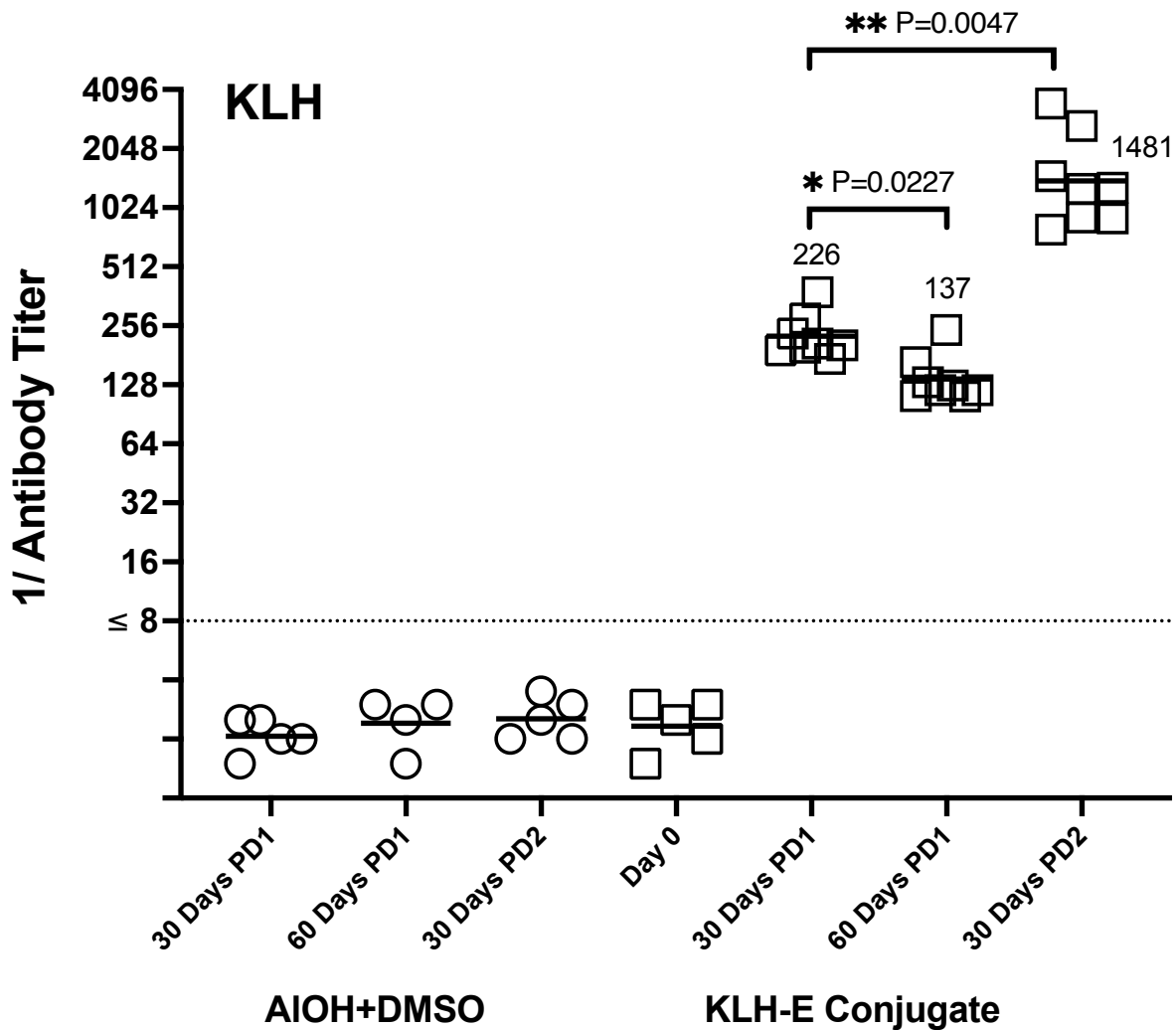
Figure 3-7. IgG antibody titers of individual mice measured against KLH.
Serum from individual mice immunized with the KLH-E conjugate (squares) or aluminum
control vaccine (circles) measured for IgG antibodies against the vaccine carrier KLH. The
responses of mice illustrated in the figure represent antibodies only to the KLH carrier. Mice
were given one or two doses of vaccine intramuscularly in 30 day intervals. Each symbol
represents the reciprocal antibody titer of individual animals in each vaccine group, with each
data point representing the average antibody titer enumerated in duplicate, in two independent
assays. Mice with no detectable IgG to KLH were assigned a titer of one-half the lower limit of
the assay (i.e. a titer of <1:8 was assigned as 4). Numbers above each group represent the
geometric mean titer (GMT) IgG antibody titer for all animals in the respective vaccine group.
Differences between vaccine groups were assessed using a two-tailed students *t* test of the log-
transformed reciprocal titers. Differences with a probability of <0.5 (two-tailed) were considered
significant.

Since the SARS-CoV-2 E peptide transmembrane domain is the primary antigenic target of our immunogenicity study, we quantified the anti-E transmembrane specific IgG antibody responses of individual mice immunized with the KLH-E vaccine as illustrated in **Figure 3-8**. The geometric mean anti-E IgG titer of mice immunized with a single dose of KLH-E was 87 at 30 days post-immunization. As with the antibodies directed at KLH, we observed a similar decline in the anti-E specific IgG antibody titers of mice immunized with a single dose of KLH-E vaccine at 60 days post-immunization, though this decrease was not significant (GMT: 87 vs. 46, 30 days vs. 60 days post-dose 1; $P=0.1325$, ns). A second dose of KLH-E vaccine elicited an increase in anti-E specific antibody, but there was no significant difference in the geometric mean anti-E antibody titers at 30 days post-immunization when comparing mice that received one or two doses of vaccine (GMT: 87 vs. 91, 1 vs. 2 doses at 30 days; $P=0.5201$, ns).

These results are surprising and interesting for two reasons. First, these results suggest that antibodies can indeed be induced against the E peptide, which represents the complete transmembrane domain sequence of the E protein which is extremely hydrophobic and small (30-amino acids). Second, there was no significant difference in anti-E specific IgG titers in mice immunized with one or two doses of KLH-E vaccine at 30 days post-immunization. This is in stark contrast to the "booster" response observed in antibodies directed at the KLH carrier after a second dose of KLH-E vaccine. Taken together, these results suggest that one dose of KLH-E vaccine may already induce the maximal level of antibody responses, and that immunization with additional doses may restore the waning effects but do not increase the overall level of anti-E antibody titers. Whether this observation is due to the nature of the E peptide transmembrane domain being extremely hydrophobic, or simply a consequence of a skewed antibody repertoire towards KLH due its greater size and immunogenicity may warrant further investigation. A comprehensive and detailed overview of the ELISA serum IgG antibody curves, dilutions tested, and pattern of reactivity for each animal tested against the KLH-E, KLH, and E-peptide antigens can be found in **Supplemental Figures 3-1, 3-2, 3-3, 3-4, 3-5, 3-6, 3-7** and are summarized in **Table 3-1**.
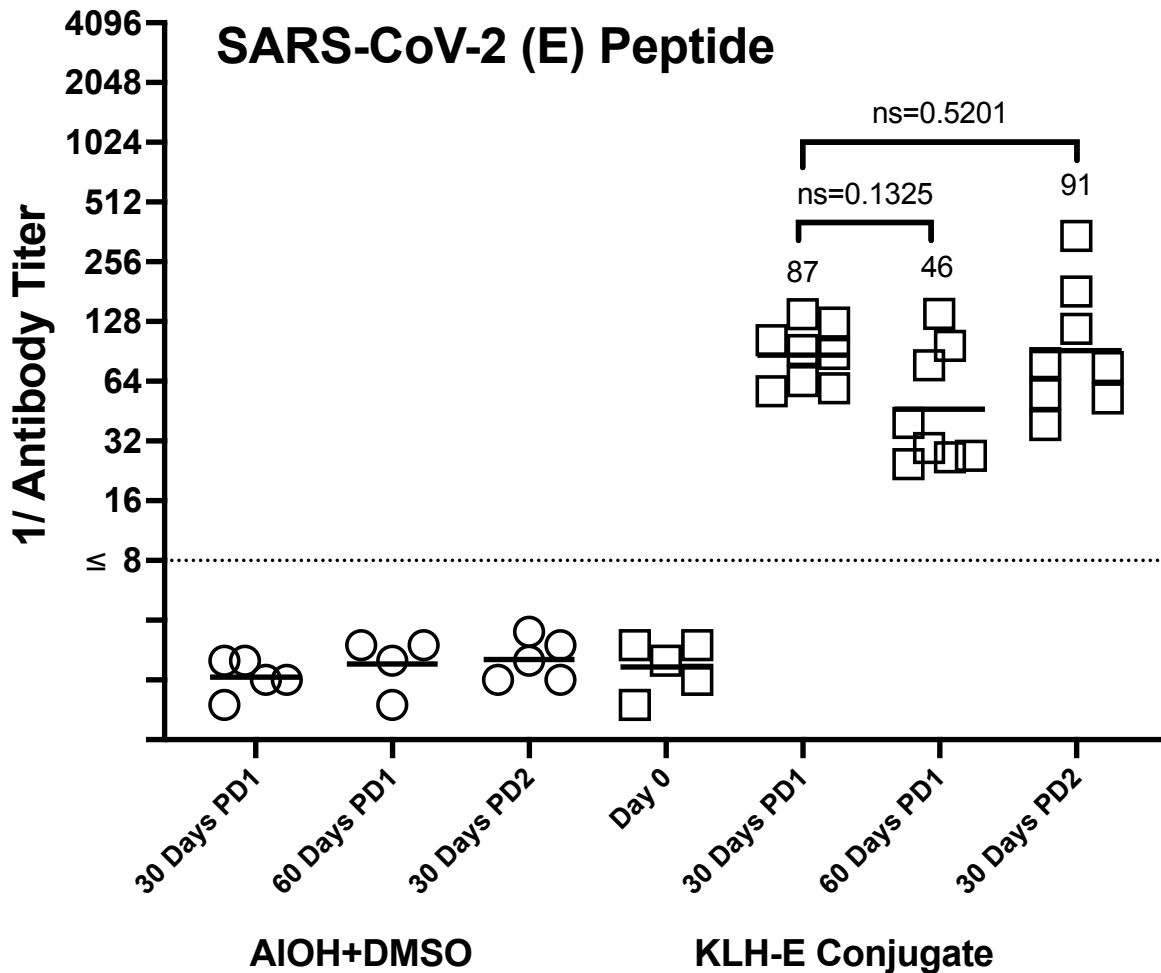
Figure 3-8. IgG antibody titers of individual mice measured against E-peptide.
Serum from individual mice immunized with the KLH-E conjugate (squares) or aluminum control vaccine (circles) measured for IgG antibodies against the SARS-CoV-2 E-peptide. The responses of mice illustrated in the figure represent antibodies only to the E-peptide transmembrane domain. Mice were given one or two doses of vaccine intramuscularly in 30 day intervals. Each symbol represents the reciprocal antibody titer of individual animals in each vaccine group, with each data point representing the average antibody titer enumerated in duplicate, in two independent assays. Mice with no detectable IgG to the E-peptide were assigned a titer of one-half the lower limit of the assay (i.e. a titer of <1:8 was assigned as 4). Numbers above each group represent the geometric mean titer (GMT) IgG antibody titer for all animals in the respective vaccine group. Differences between vaccine groups were assessed using a two-tailed students *t* test of the log-transformed reciprocal titers. Differences with a probability of <0.5 (two-tailed) were considered significant.

130

Table 3-1: Summary of Mouse Immunization Groups and IgG Antibody Titers										
Vaccine Group	Vaccine	No. Doses	Blood collection post immunization	Mouse No.	KLH-E		KLH		E peptide	
1 (n=5)	KLH-E	1	Day 0 *Immunized with KLH-E and sacrificed	1	< 8	< 8	< 8	< 8	< 8	< 8
				2	< 8	< 8	< 8	< 8	< 8	< 8
				3	< 8	< 8	< 8	< 8	< 8	< 8
				4	< 8	< 8	< 8	< 8	< 8	< 8
				5	< 8	< 8	< 8	< 8	< 8	< 8
2 (n=5)	AlOH+ DMSO	1	30 days	1	< 8	< 8	< 8	< 8	< 8	< 8
				2	< 8	< 8	< 8	< 8	< 8	< 8
				3	< 8	< 8	< 8	< 8	< 8	< 8
				4	< 8	< 8	< 8	< 8	< 8	< 8
				5	< 8	< 8	< 8	< 8	< 8	< 8
3 (n=4)	AlOH+ DMSO	1	60 days	1	< 8	< 8	< 8	< 8	< 8	< 8
				2	< 8	< 8	< 8	< 8	< 8	< 8
				3	< 8	< 8	< 8	< 8	< 8	< 8
				4	< 8	< 8	< 8	< 8	< 8	< 8
4 (n=5)	AlOH+ DMSO	2	30 days	1	< 8	< 8	< 8	< 8	< 8	< 8
				2	< 8	< 8	< 8	< 8	< 8	< 8
				3	< 8	< 8	< 8	< 8	< 8	< 8
				4	< 8	< 8	< 8	< 8	< 8	< 8
				5	< 8	< 8	< 8	< 8	< 8	< 8
5 (n=8)	KLH-E	1	30 days	1	224	250	240	227	86	90
				2	197	202	185	199	58	60
				3	218	205	200	196	54	60
				4	268	270	282	275	133	120
				5	193	200	203	212	67	61
				6	186	191	200	207	106	100
				7	159	171	168	179	94	90
				8	358	325	370	388	148	131
6 (n=8)	KLH-E	1	60 days	1	173	151	164	170	98	95
				2	223	263	237	250	130	150
				3	118	121	113	127	26	28
				4	126	113	119	121	40	39
				5	109	136	124	140	25	24
				6	110	103	118	105	26	27
				7	125	113	120	134	76	78
				8	99	117	107	117	29	30
7 (n=8)	KLH-E	2	30 days	1	1378	1562	1221	1371	113	122
				2	2721	2999	2555	2780	173	190
				3	3721	3883	3318	3621	334	357
				4	1562	1613	1420	1542	70	88
				5	915	944	900	922	51	59
				6	932	891	899	903	36	41
				7	771	816	801	782	49	56
				8	1210	1224	1298	1252	72	79

T-cell response to KLH-E vaccine. Effective clearance of viral infections, such with SARS-CoV-2, require both antibody and cellular adaptive immunity. CD4+ helper T-cells are master regulators of adaptive immunity that secrete cytokines necessary for B-cell antibody production and CD8+ cytotoxic T-cell effector functions. We previously demonstrated that a novel KLH-E vaccine elicited antibody responses in mice to each subunit of the KLH-E vaccine and sought to investigate whether the vaccine also elicited antigen-specific cellular immunity. To address this question, we homogenized spleen tissue from KLH-E and aluminum vaccinated mice to isolate spleenocytes, a term that describes a variety of cell populations such as B-cells, T-cells, Macrophages and Dendritic cells purified from splenic tissue. We developed a unique ELISPOT-based assay for quantifying the antigen-specific T-cell responses of mice immunized with a novel KLH-E vaccine. In the assay, mouse spleenocytes are incubated in the presence or absence of purified antigen using a nitrocellulose membrane coated with a mouse monoclonal IL-2 antibody. The mouse spleen is rich in T-cells and antigen presenting cells (B-cells, Dendritic cells, Macrophages) that are capable of processing exogenous antigen and presenting it to T-cells for recognition. A major hallmark of T-cell activation is the production and secretion of IL-2 cytokine, which is then captured by the monoclonal antibodies conjugated to the nitrocellulose membrane. IL-2 foci are then developed and enumerated, producing an assay output as illustrated in **Supplemental Figure 3-8**.

The cryopreservation of spleenocytes allows them to be conveniently used in immunological assays, but the process also has the potential to reduce overall cell viability. Since our ability to measure functional and specific T-cell activation depends on antigenic-specificity, cell viability, and functional activity, prior to quantifying the antigen-specific responses of KLH-E vaccinated mice we performed two pilot experiments to ensure that the T-cells in our splenic samples were both viable and retained functional activity. First, to measure viability we enumerated freshly thawed splenic samples using a hemocytometer and trypan blue staining as described in the materials and methods. Of the total number of cells counted in each splenic sample, we observed ≥90% cell viability by trypan blue staining (data not shown). The trypan blue dye method has long been used to assess cell viability since live cells possess intact membranes that exclude the dye, whereas dead cells do not. Second, to ensure that the T-cells in our splenic samples retained the ability become activated and produce IL-2, we stimulated spleenocytes with Concanavalin A (ConA), a lectin that binds to and cross-links components of the T-cell receptor to induce activation resulting in IL-2 production in an antigen-independent manner. **Supplemental Figure 3-9** illustrates that mouse T-cells retain functional activity after cryopreservation, irrespective of vaccine dose or formulation, as determined by the presence of IL-2 secreting foci. Stimulation with ConA resulted in similar proportions of activated T-cells in both KLH-E and control Aluminum vaccinated mice as expected. These findings suggest that T-cells from the splenic samples collected from immunized mice are fully functional and can be activated, further supporting that these cells collected from mice using our experimental procedures fully retain their activity.

132

After establishing the functionality of our T-cell assay, we investigated the antigen-specific T-cell response to each vaccine subunit induced by KLH-E immunization in representative mice from each vaccine group. As depicted in **Figure 3-9**, a single dose of KLH-E vaccine elicited similar proportions of T-cells at 30 and 60 days post-immunization (44 vs. 37 IL-2 secreting foci; 30 days vs. 60 days) when measured against the KLH-E conjugate that was used as the vaccine antigen. There was a slight decrease in KLH-E specific T-cells at 60 days vs. 30 days post-immunization with a single dose but this decrease was not statistically significant (P=0.808). A second dose of KLH-E vaccine induced a significant increase in KLH-E specific T-cells at 30 days post-immunization, compared to a single dose at either time point (104 vs. 44 [**P=0.018] or 104 vs. 37 [***P=0.009] IL-2 secreting foci; 2 doses vs. 1 dose at 30 or 60 days post-immunization, respectively).
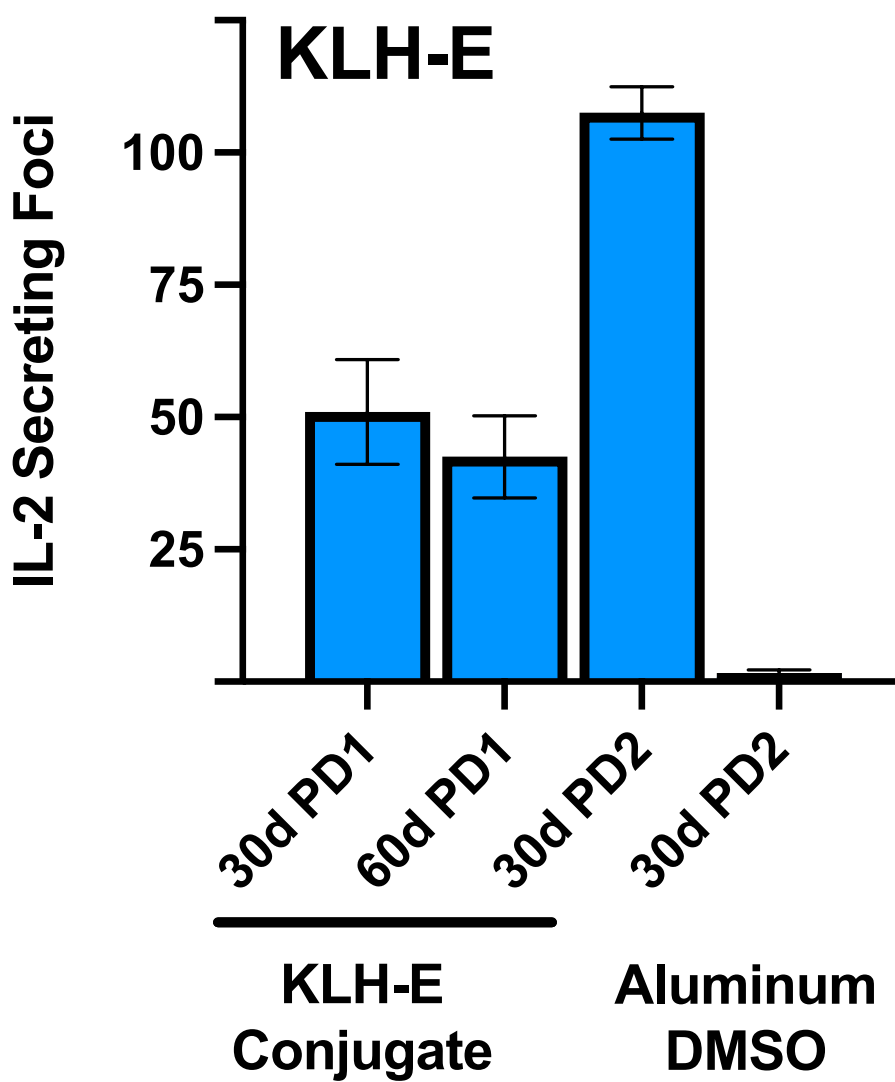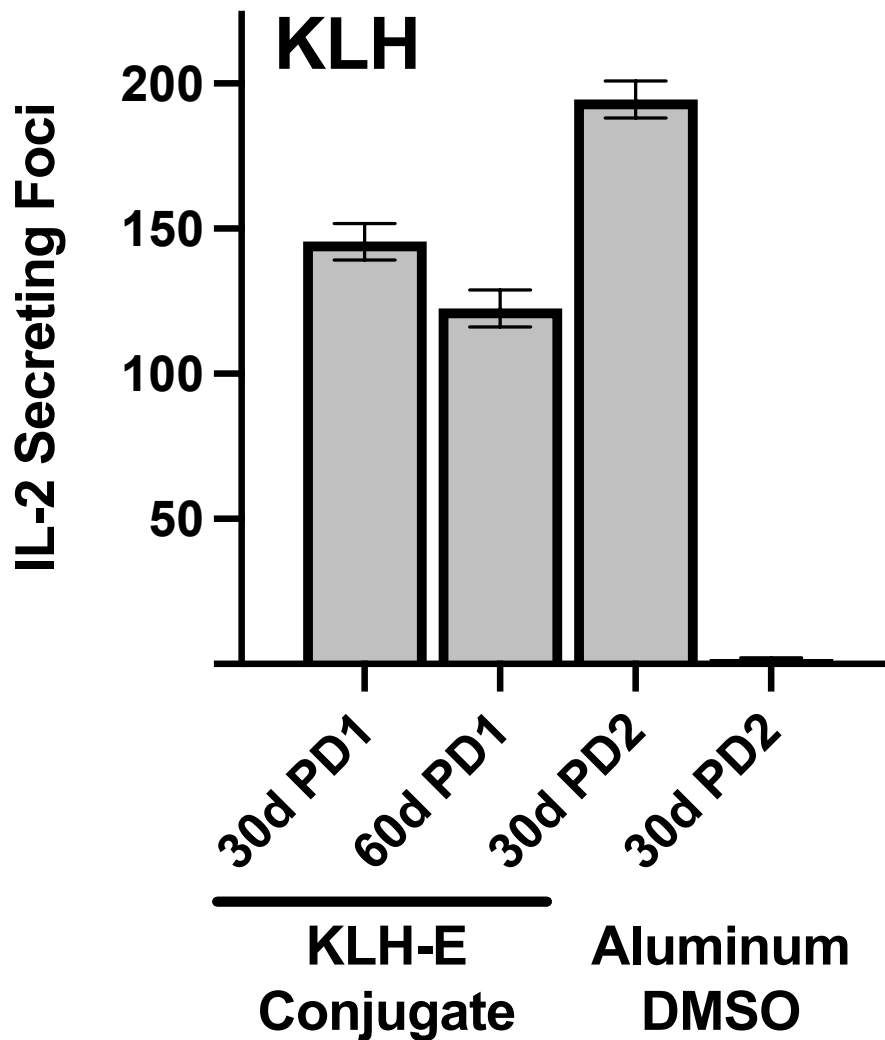
Figure 3-9. T-cell responses of mice measured against KLH-E.
The horizontal axis lists individual representative animals from each vaccine group at select time points, while the verticle corresponds to the enumerated IL-2 secreting foci in the antigen-specific T-cell activation assay. A total of 750,000 spleenocytes from representative mice in each vaccine group and timepoint were stimulated with 2µg of KLH-E and tested for T-cell activation. Data represent the mean of two independent replicates and error bars represent the +/- standard deviation.

A single dose of KLH-E vaccine elicited similar proportions of T-cells at 30 and 60 days post-immunization (141 vs. 127 IL-2 secreting foci; 30 days vs. 60 days) when measured against the KLH carrier protein as illustrated in **Figure 3-10**. There was a slight decrease in KLH-specific T-cells at 60 days vs. 30 days post-immunization with a single dose, but this decrease did not reach statistical significance (P=0.068). A second dose of KLH-E vaccine induced a significant increase in KLH-specific T-cells at 30 days post-immunization, and this was greater than a single dose at either time point (199 vs. 141 [**P=0.016] or 199 vs. 127 [***P=0.008] IL-2 secreting foci; 2 doses vs. 1 dose at 30 or 60 days post-immunization, respectively).

Figure 3-10. T-cell responses of mice measured against KLH.
The horizontal axis lists individual representative animals from each vaccine group at select time points, while the vertical axis corresponds to the enumerated IL-2 secreting foci in the antigen-specific T-cell activation assay. A total of 750,000 spleenocytes from representative mice in each vaccine group and timepoint were stimulated with 100μg of KLH and tested for T-cell activation. Data represent the mean of two independent replicates and error bars represent the +/- standard deviation.

Since the KLH and KLH-E antigens differ only by the presence of the 30-amino acid SARS-CoV-2 E protein transmembrane domain, one could question why it was that there were more IL-2 secreting foci in mouse spleenocytes stimulated with KLH vs. KLH-E even at the same spleenocyte cell concentration tested in the assay. The reason for this, can be explained by two factors. First, it is possible that the differences we observed are due to variation in cellular immune responses of individual mice. Second, although the same concentration of spleenocytes were used from different animals, we stimulated the cells with starkly different concentrations of exogenous antigen (KLH=100µg vs. KLH-E=2µg). Thus, while there is a 50X order of magnitude difference in the concentration of antigen used for stimulation, the conclusion from both experiments are equivalent and consistent with the body of literature regarding the immunogenicity of the KLH antigen. Specifically, that KLH is highly immunogenic and elicits a robust T-cell response.

As illustrated in **Figure 3-11**, a single dose of KLH-E vaccine elicited similar proportions of T-cells at 30 days and 60 days post-immunization (12 vs. 5 IL-2 secreting foci; 30 days vs. 60 days) when measured against the 30-amino acid E-peptide corresponding to the SARS-CoV-2 E transmembrane domain. There was a significant decline in the E-peptide specific T-cells at 60 days vs. 30 days post-immunization with a single dose (**P=0.038). A second dose of KLH-E vaccine induced a three to six-fold increase in E-peptide specific T-cells compared to a single dose at either timepoint (29 vs. 12, or 29 vs. 5, IL-2 secreting foci; 2 doses vs. a single dose at 30 or 60 days post-immunization). The induction of elevated anti-E specific T-cells after a second dose of vaccine was statistically significant when compared to a single dose at 30 days post-immunization (**P=0.033) or 60 days post-immunization (**P=0.018). Taken together, our studies demonstrate that KLH-E vaccination induces both antibodies and antigen-specific T-cells to the KLH-E antigen used for vaccination, as well as to each of the subunits of the KLH-E vaccine. A comprehensive summary of IL-2 secreting foci enumerated in our T-cell assays including replicate values and different concentrations of spleenocytes tested are listed in **Table 3-2.**
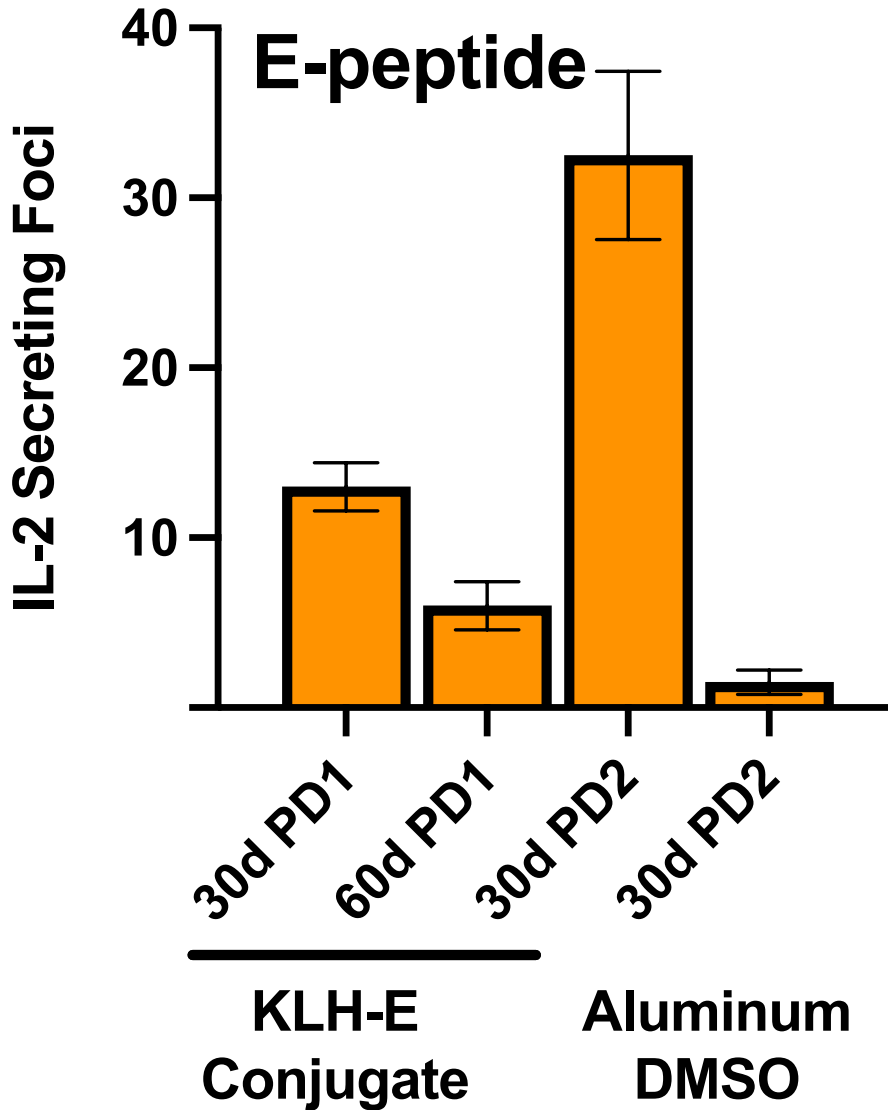
Figure 3-11. T-cell responses of KLH-E vaccinated mice measured against E-peptide. The horizontal axis lists individual representative animals from each vaccine group at select time points, while the vertical axis corresponds to the enumerated IL-2 secreting foci in the antigen-specific T-cell activation assay. A total of 750,000 spleenocytes from representative mice in each vaccine group and timepoint were stimulated with 1µg of E-peptide (TLIVNSVLLFLAFVVFLLVTLAILTALRLC) and tested for T-cell activation. Data represent the mean of two independent replicates and error bars represent the +/- standard deviation.

Mouse No.	Vaccine	Doses	Spleenocyte collection post-immunization	Spleenocytes Tested	KLH-E (2μg)		KLH (100μg)		E peptide (1μg)		ConA	
#5	KLH-E	1	30 days	2,500	2	1	2	2				
				25,000			11	18				
				250,000	9	13	141	150	1	1		
				750,000	44	58			2	3	199	199
				2,250,000					12	14		
#7	KLH-E	1	60 days	2,500	1	1	0	0				
				25,000			12	20	0	0		
				250,000	10	15	127	118	1	1	185	170
				750,000	37	48			5	7		
#4	KLH-E	2	30 days	2,500	7	10	15	20				
				25,000			75	90				
				250,000	15	17			1	3		
				750,000	104	111	199	199	5	3	180	188
				2,250,000					29	36		
#5	Aluminum	2	30 days	2,500	0	0						
				25,000	0	0	0	0				
				250,000	0	1	0	1	0	0		
				750,000			1	1	0	0	151	144
				2,250,000					1	0		

Table 3-2. Summary of Mouse IL-2 secreting foci following antigen stimulation

Cross-reactivity of Antibodies and T-cells induced by KLH-E vaccination. Since the E protein transmembrane domain is highly conserved across emerging SARS-CoV-2 variants of concern, and moderately conserved across different human coronaviruses, we investigated the cross-reactivity of antibodies and T-cells induced by KLH-E vaccination. First, we measured by ELISA the ability of anti-E antibodies induced by KLH-E vaccination to recognize SARS-CoV-2 E protein expressed on the surface of a hybrid alphavirus SARS-CoV-2 pseudovirions (HA-CoV-2). The HA-CoV-2 pseudovirions all expressed the four SARS-CoV-2 structural proteins (S, M, N, and E) and each of the four pseudovirions tested, expressed a different S protein sequence from a SARS-CoV-2 variant of concern (Beta, Delta, Omicron, XBB). As illustrated in **Figure 3-12**, at 30 days post-immunization pooled sera from mice vaccinated with two doses of KLH-E, but not the control aluminum vaccine, contained anti-E antibodies that recognize the E protein on the surface of HA-CoV-2 pseudovirions. As expected, pooled sera from KLH-E vaccinated mice reacted similarly by ELISA, irrespective of the S protein sequence expressed by the HA-CoV-2 pseudovirions. As an additional assay control, we also measured the ELISA reactivity of control rabbit monoclonal antibody (mAb 27VB1) that is broadly neutralizing and recognizes the receptor binding domain (RBD) of the SARS-CoV-2 S protein. The mAb bound to all four SARS-CoV-2 pseudovirions tested to varying degrees (data not shown) which is consistent with both *in vitro* and *in vivo* studies that report SARS-CoV-2 variants of concern (particularly omicron and XBB) have divergent S protein sequences that result in impaired antibody binding [86].

The same serum pools were also tested by ELISA against HCoV-229E infected human MRC5 whole cell lysates to determine whether anti-E antibodies induced by KLH-E vaccination could recognize native full length E protein from a genetically divergent human coronavirus. As depicted in **Figure 3-13,** pooled serum from mice immunized with the control aluminum vaccine had no reactivity to both uninfected and HCoV-229E infected MRC5 cells irrespective of the number of doses or timepoint as expected. Likewise, as illustrated in **Figure 3-14,** pooled serum from mice immunized with KLH-E had no reactivity to uninfected MRC5 cell lysates irrespective of the number of vaccine doses or timepoints. To our surprise, we observed that serum pools from mice immunized with the KLH-E vaccine contained antibodies that were reactive to HCoV-229E infected MRC5 cell lysates with serum pools from mice immunized with two doses of KLH-E vaccine yielded higher antibody reactivity compared to pools from mice immunized with a single dose of vaccine. Moreover, we observed a decline in antibody reactivity in pooled sera from mice immunized with a single dose of KLH-E vaccine at 60 days post-immunization compared to 30 days post-immunization. These findings, using the same mouse serum pools, share similar trends with the previously observed antibody reactivity measured against the 30-amino acid E-peptide corresponding to the SARS-CoV-2 E protein transmembrane domain. One important distinction however, was that the same dilution of pooled KLH-E mouse sera produced approximately half the amount of signal when tested against the HCoV-229E infected cell lysate compared to the 30-amino E-peptide, which may be due to the low sequence homology (32%) between the E protein transmembrane domains of SARS-CoV-2 and HCoV-229E.

While using HCoV-229E infected cell lysates as a source of native E protein has significant benefits, one challenge was ensuring that our observations were indeed specific for anti-E antibodies. We concluded that the antibody reactivity of pooled KLH-E vaccinated mouse sera measured against HCoV-229E infected human MRC5 cell lysates was due to the presence of specific anti-E antibodies based on the observation that the same serum pool did not react to control uninfected MRC5 cell lysates. These observations suggests that antibodies induced by KLH-E vaccination recognize an antigen present in HCoV-229E infected, but not uninfected cells.

Another important consideration when using HCoV-229E infected cell lysates as a source of native E protein is ensuring that sufficient test lysate is present on the ELISA plate while also keeping the overall concentration of coating antigen at a minimum to reduce non-specific background signal. Conceivably, differences in the concentration of uninfected vs. HCoV-229E infected cell lysates could result in differences in antibody reactivity that could confound interpretation. To eliminate this possibility, we controlled for differences in coating antigen concentration in each of our ELISA plates by first measuring the total protein concentration of each cell lysate preparations as described in the materials and methods. As an additional control, to show that the uninfected and HCoV-229E infected MRC5 cell lysate preparations were equivalent in total protein concentration and overall protein composition (excluding the E protein), we measured the level of an unrelated antigen that should normally be present in both cell lysate preparations irrespective of HCoV-229E infection. **Supplemental Figure 3-10**, depicts the levels of β-actin in each MRC5 cell lysate preparation for uninfected and HCoV-229E infected cells. As expected, we observed no significant difference in the levels of β-actin in each of the MRC5 cell lysate preparations, with both curves nearly overlapping. This observation further supports that the reactivity we observed in pooled KLH-E immune mouse sera measured against HCoV-229E infected cell lysates was specific for anti-E antibodies, and that both plates were coated with equivalent total protein concentrations.
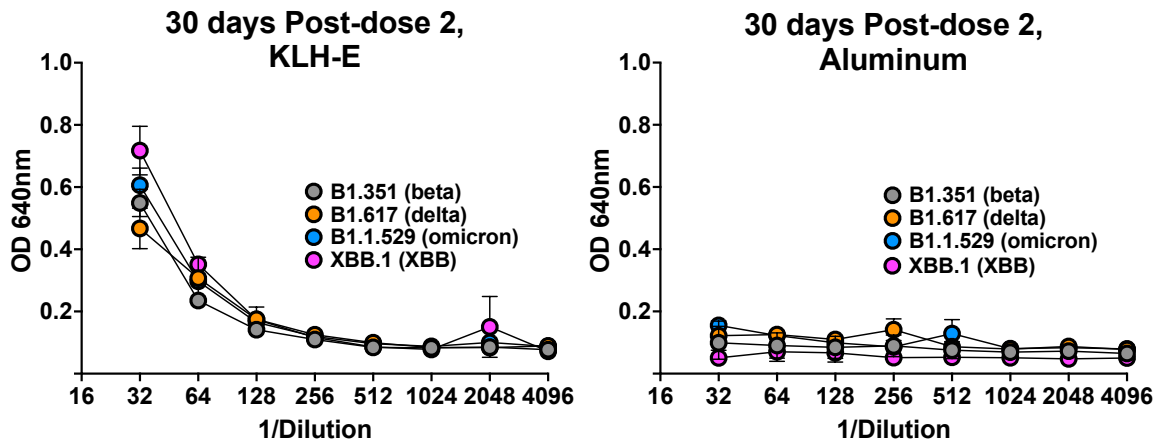
Figure 3-12. Antibody response of pooled mouse sera measured against HA-CoV-2 pseudovirions.

Figure depicts the anti-E IgG of pooled sera from mice immunized intramuscularly with two doses of KLH-E or aluminum control vaccine at 30 days post-immunization and measured for anti-E IgG antibody binding against hybrid alphavirus SARS-CoV-2 pseudovirions (HA-CoV-2) by ELISA. The HA-CoV-2 pseudovirions all expressed the four SARS-CoV-2 structural proteins (Spike (S), Matrix (M), Nucleocapsid (N), Envelope (E)), but each of the four HA-CoV-2 pseudovirions tested expressed a S protein sequence corresponding to a SARS-CoV-2 variant of concern (Beta, Delta, Omicron, XBB).

MRC5 Uninfected

Legend:
- 30 days Post Dose 1
- 60 days Post Dose 1
- 30 days Post Dose 2

Aluminum

OD640nm vs 1/Serum Dilution

MRC5 Uninfected

Legend:
- 30 days Post Dose 1
- 60 days Post Dose 1
- 30 days Post Dose 2

MRC5 HCoV-229E Infected

Aluminum
- 30 days Post Dose 1
- 60 days Post Dose 1
- 30 days Post Dose 2

MRC5 HCoV-229E Infected

Aluminum
- 30 days Post Dose 1
- 60 days Post Dose 1
- 30 days Post Dose 2
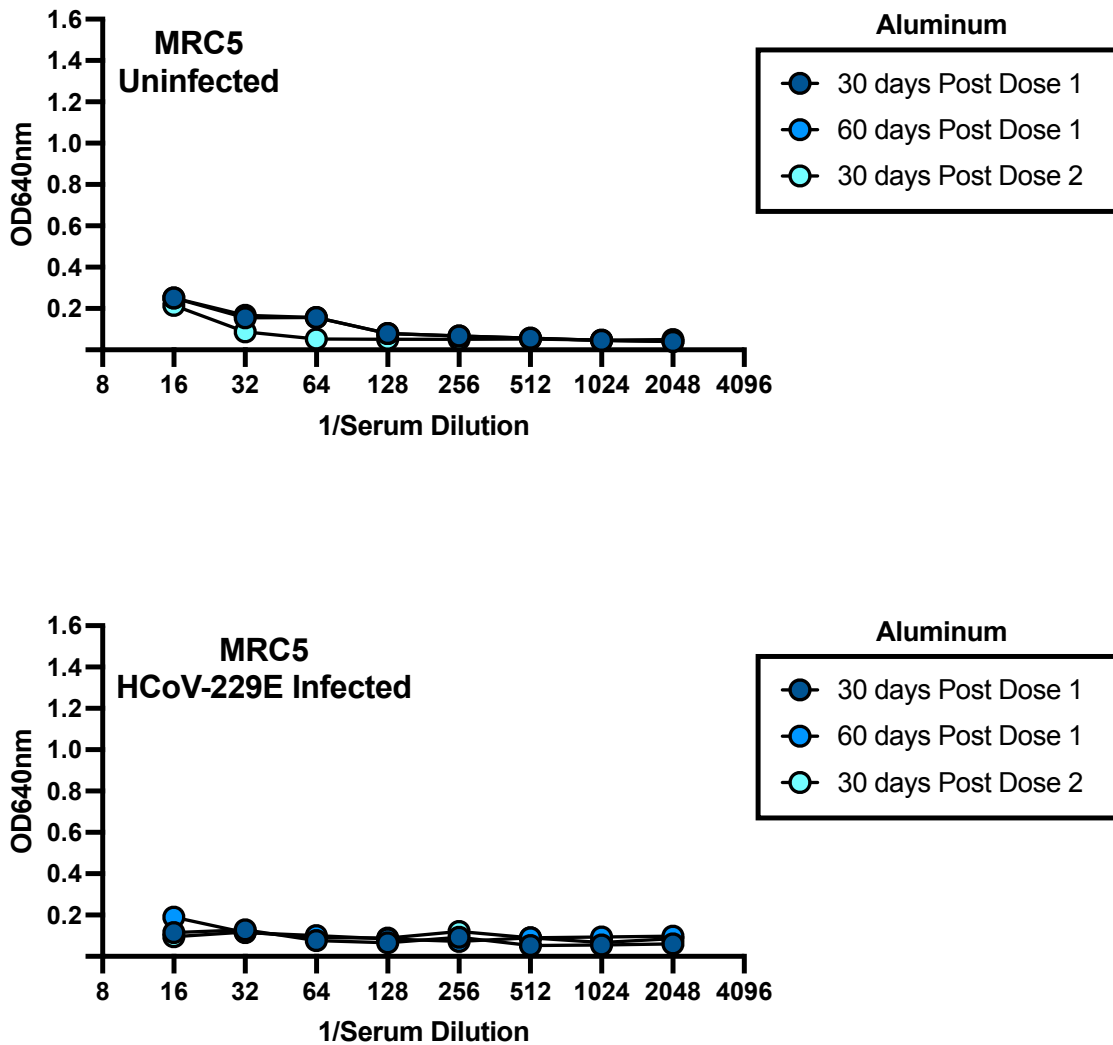
Figure 3.13. Pooled mouse serum of mice immunized with aluminum control vaccine measured against uninfected or HCoV-229E infected MRC5 cell lysates. Pooled serum from groups of mice immunized with the aluminum control vaccine were combined in equal parts using serum from each vaccine group and measured for IgG antibodies to the native E protein expressed in uninfected (top) or HCoV-229E infected MRC5 cell lysates (bottom). The serum pools were from mice immunized with the control aluminum vaccine collected at 30 days post-dose 1 (n= 5 mice), 60 days post-dose 1 (n=4 mice), or 30 days post-dose 2 (n=5 mice).

143
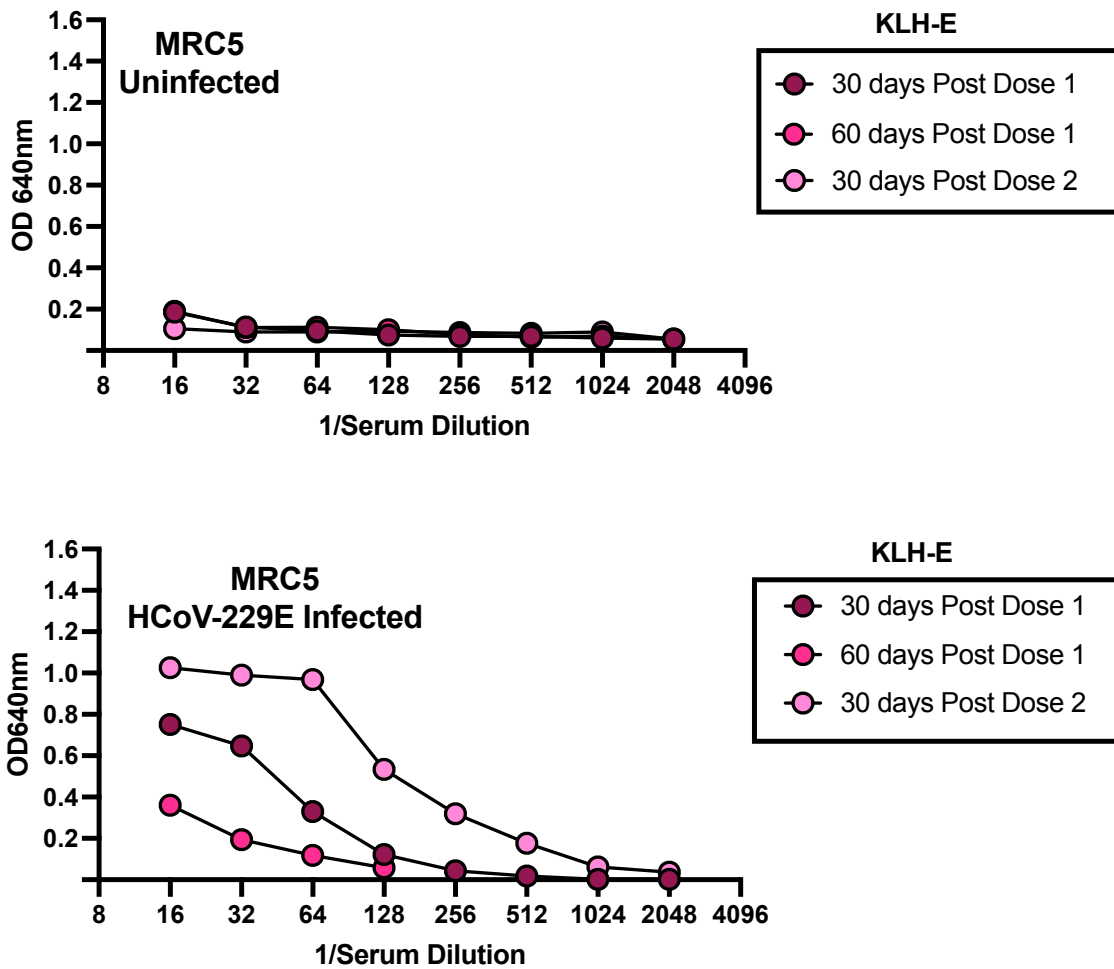
Figure 34. Pooled mouse serum of mice immunized with KLH-E vaccine measured against uninfected or HCoV-229E infected MRC5 cell lysates. Pooled serum from groups of mice immunized with the KLH-E vaccine were combined in equal parts using serum from each vaccine group and measured for IgG antibodies to the native E protein expressed in uninfected (top) or HCoV-229E infected MRC5 cell lysates (bottom). The serum pools were from mice immunized with the control KLH-E vaccine collected at 30 days post-dose 1 (n= 8 mice), 60 days post-dose 1 (n=8 mice), or 30 days post-dose 2 (n=8 mice).

The results of these experiments are significant and unexpected for two reasons. First, the surface accessibility of the SARS-CoV-2 native E protein and its domain architecture in the SARS-CoV-2 virion remain incompletely understood in the absence of a high-resolution crystal structure. Though the SARS-CoV-2 E protein transmembrane domain has been modeled based on nuclear magnetic resonance (NRM) microscopy data of the protein in lipid bilayers [70], the current model lacks the N and C termini of the full length protein. Since the amino acid charges on both the N and C termini of the SARS-CoV-2 E protein, as well as the composition of the lipid membrane that surrounds the protein, can greatly influence the accessibility of the transmembrane domain, the topology and surface accessibility of epitopes on the E protein in SARS-CoV-2, and coronaviruses broadly, remain highly debated in the literature and largely inconclusive [87]. Nevertheless, our ELISA experiments demonstrate that anti-E transmembrane antibodies induced by KLH-E vaccination recognize an epitope that is accessible even when the native full length E protein is expressed in different modalities (pseudovirions vs. infected cell lysates) and independent of any specific SARS-CoV-2 E protein conformation.

Second and most surprising, is that anti-E antibodies induced by KLH-E vaccination appear to be cross-reactive and can recognize divergent E protein sequences from other human coronaviruses, presumably epitopes located within the transmembrane domain. In fact, among all seven coronaviruses known to infect humans the sequence homology of the full-length E proteins relative to the SARS-CoV-2 E protein ranges from: 94.74% (SARS-CoV-1) to 18.46% (HCoV-NL63), as described in detail in **Table 3-3**. Notwithstanding, restricting the sequence analysis to only the E transmembrane domain sequences of SARS-CoV-2 and HCoV-229E there is a mere 32% sequence homology. Yet, despite this low sequence homology, anti-E antibodies induced by KLH-E vaccination recognize HCoV-229E native E protein in infected cell lysates, which suggests that the SARS-CoV-2 E protein transmembrane domain may contain important epitopes that are shared across genetically divergent human coronaviruses. The preceding observations led us to inquire whether the KLH-E vaccine also induced cross-reactive T-cells.

To address this question, we measured antigen-specific T-cell responses as before but used HCoV-229E infected and uninfected human MRC5 cell lysates for antigen stimulation. **Figure 3-15** depicts the antigen-specific T-cell responses of mice immunized with two doses of KLH-E or aluminum control vaccine at 30 days post-immunization stimulated with uninfected and HCoV-229E infected human MRC5 cell lysates. As expected, there was no significant T-cell response in cells stimulated with the uninfected human MRC5 cell lysates irrespective of vaccination (1 vs. 1 IL-2 secreting foci, P=1). In contrast, when the same cells were treated with HCoV-229E infected cell lysates, we observed a 14-fold increase in IL-2 secreting foci in mice receiving two doses of KLH-E vaccine compared to mice immunized with the same number of doses of aluminum vaccine (14 vs. 1 IL-2 secreting foci; **P=0.036).

To confirm that the differences we observed were due to antigenic-specificity of the T-cells in mice for the E protein in HCoV-229E infected cell lysates and not due to defects in the ability of T-cells to become activated, we tested the same cells using the control Concanavalin A (ConA), which strongly induces T-cell activation and IL-2 secretion in an antigen-independent manner. Treatment of mouse spleenocytes with ConA resulted in similar proportions of IL-2 secreting foci irrespective of the vaccine administered (205 vs. 206 IL-2 secreting foci; KLH vs. Aluminum vaccinated). A comprehensive list of the IL-2 secreting foci enumerated and number of spleenocytes tested can be found in **Table 3-4**.

This experiment led to two important insights. First, it suggests that native E protein from HCoV-229E infected cell lysates can be processed by antigen presenting cells and recognized by T-cells in mouse splenic samples used in our T-cell activation assay. Second, it implicates that during antigen processing, a region of the HCoV-229E E protein (presumably the transmembrane region) is presented and can be recognized by KLH-E vaccine induced T-cells. Taken together, these results suggest that KLH-E vaccination not only induces cross-reactive antibodies, but also induces cross-reactive T-cells that recognize a conserved epitope in the E protein even when the overall sequence homology is low.

Table 3-3. Envelope protein amino acid sequence homology of human coronaviruses				
Subfamily	Human coronavirus	E protein (Uniprot code)	Full Length (amino acids)	Homology % to SARS-CoV-2
betacoronavirus	SARS-CoV-2	P0DTC4	75	100%
	SARS-CoV-1	P59637	76	94.74%
	MERS-CoV	K9N5R3	82	36.00%
alphacoronavirus	HCoV-HKU1	Q0ZJ83	82	31.58%
	HCoV-OC43	Q4VID3	84	31.15%
	HCoV-229E	P19741	77	27.14%
	HCoV-NL63	H9EJA2	77	18.46%

Table 3-3. Envelope protein amino acid sequence homology of human coronaviruses
The total amino acid length of published human coronavirus envelope protein sequences for: SARS-CoV-2, SARS-CoV-1, MERS-CoV, HCoV-HKU1, HCoV-OC43, HCoV-229E, HCoV-NL63 were obtained from Uniprot and aligned using Clustal Omega (EMBL-EBI) for sequence homology analysis using standard default parameters. The sequence homology % of each human coronavirus envelope protein is listed in comparison to the SARS-CoV-2 envelope protein.
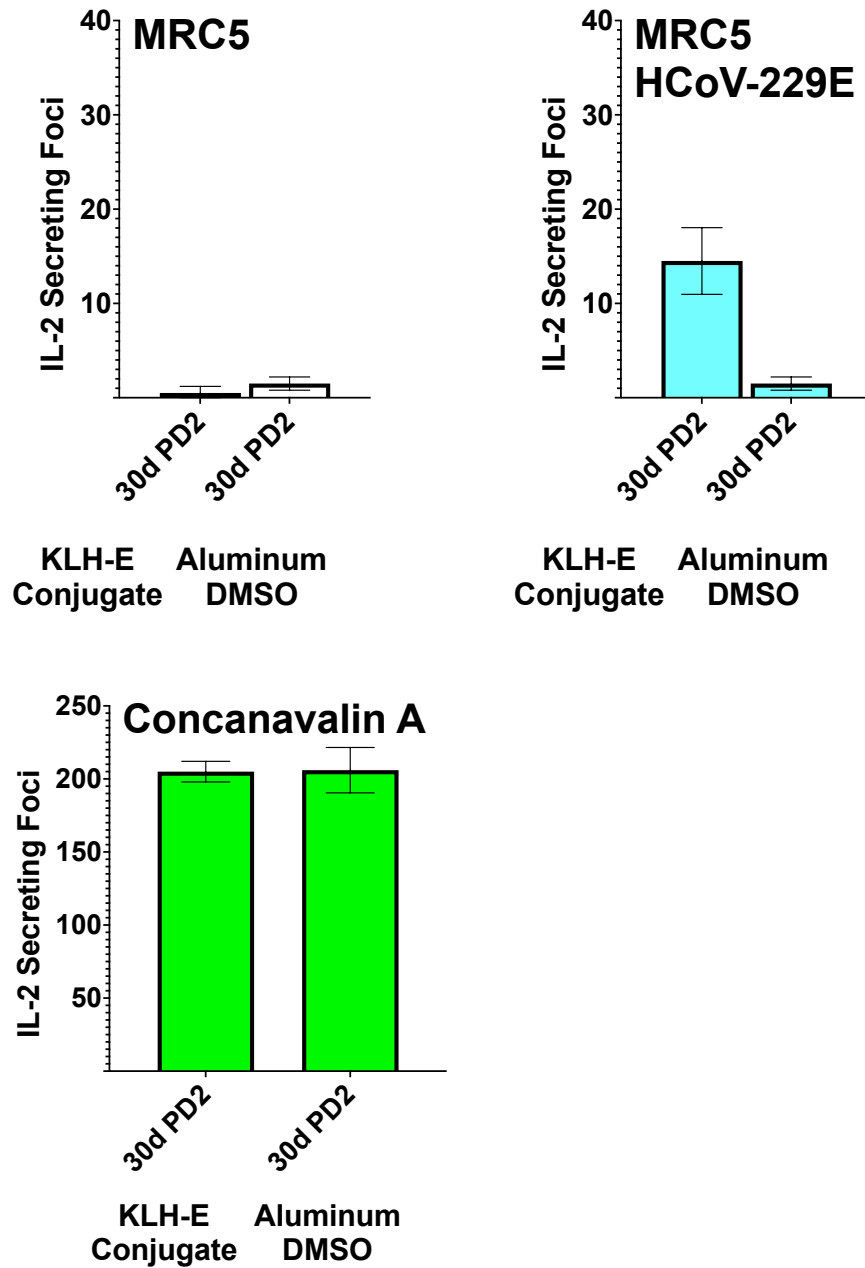
Figure 3-15. T-cell responses of KLH-E vaccinated mice measured against uninfected or HCoV-229E infected MRC5 cell lysates.

The horizontal axis lists individual representative animals from mice immunized intramuscularly with two doses of the KLH-E or aluminum vaccine at 30 days post-immunization. The vertical axis corresponds to the enumerated IL-2 secreting foci in the antigen-specific T-cell activation assay. A total of 750,000 splenocytes from representative mice in each group were stimulated with 200µg of uninfected or HCoV-229E infected human MRC5 cell lysates. As a control, the same concentration of cells were treated with Concanavalin A (ConA).

Mouse No.	Vaccine	Doses	Spleenocyte collection post-immunization	Spleenocytes Tested	Uninfected (200µg)		HCoV-229E infected (200µg)		ConA		Cell only	
#3	Aluminum	2	30 days	250,000	0	0	0	0	74	70	0	0
				750,000	0	1	1	2	200	210	0	0
#8	KLH-E	2	30 days	250,000	0	0	4	7	56	80	0	0
				750,000	1	2	17	12	217	195	0	0

Table 3-4. Summary of mouse IL-2 secreting foci following stimulation with uninfected or HCoV-229E infected cell lysates

Neutralization activity of Antibodies induced by KLH-E vaccination. To date, the most widely used immunological metric to infer SARS-CoV-2 vaccine efficacy is the induction of neutralizing antibodies. To investigate the functional activity of anti-E transmembrane antibodies induced by KLH-E vaccination, we measured the neutralization activity of pooled sera from mice receiving two doses of KLH-E or aluminum control vaccine at 30 days post-immunization against the four HA-CoV-2 pseudovirions used in previous ELISA experiments. Importantly, the HA-CoV-2 system used in this study has been validated as a robust platform for the rapid quantification of neutralizing antibodies against SARS-CoV-2 and its emerging variants, and correlates quite accurately (R^2 =0.87) with neutralizing titers measured against the actual SARS-CoV-2 virus [88].

As illustrated in **Figure 3-16**, pooled KLH-E immune sera had similar neutralization activity (% infection, range: 0 - 4%) against all four HA-CoV-2 pseudovirions expressing different spike protein sequences from SARS-CoV-2 variants of concern (Beta, Delta, Omicron, XBB). As expected, pooled serum from control aluminum vaccinated mice had negligible or no neutralization activity (% infection, range: 97-100%) suggesting that the neutralization of HA-CoV-2 pseudovirions was mediated by specific antibody and not due to a global serum sensitivity. As a control, we tested a broadly neutralizing rabbit monoclonal antibody (mAb 27VB1) directed at the receptor binding domain (RBD) of the SARS-CoV-2 S protein and observed disparate levels of neutralization activity (% infection, range: 0-65%). Specifically, HA-CoV-2 Omicron and XBB pseudovirions had greater resistance to neutralization, consistent with both *in vivo* and *in vitro* reports that these highly mutated S protein sequences confer resistance to vaccine induced and natural immunity mediated by anti-S antibody against SARS-CoV-2 variants of concern [89-91].
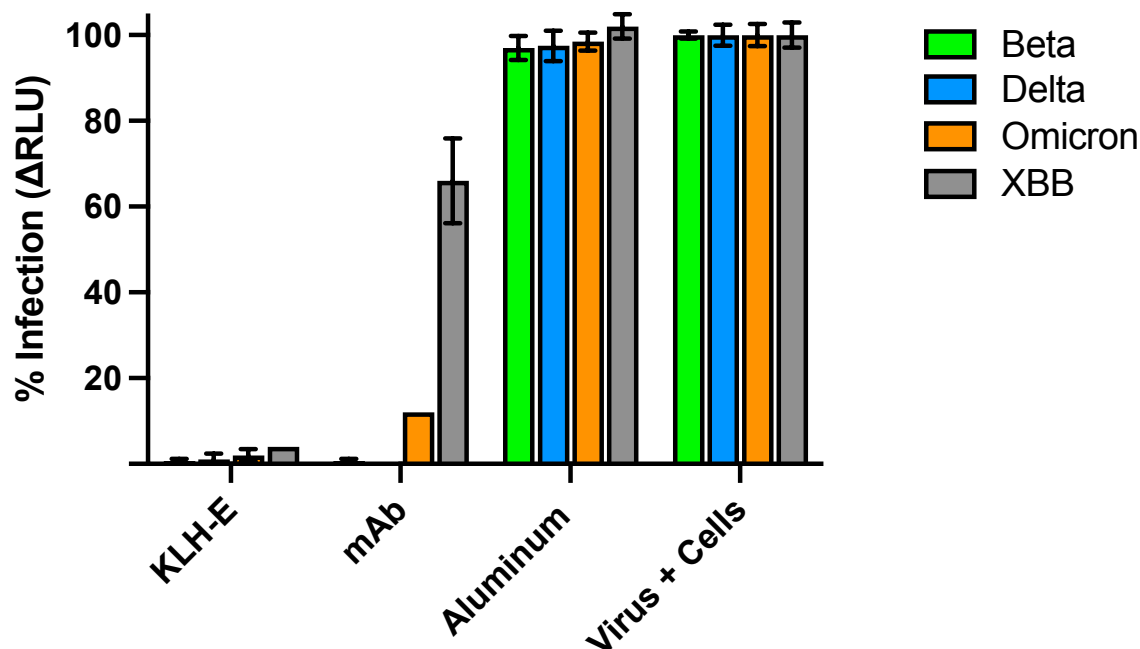
Figure 3-16. Neutralization of HA-CoV-2 Pseudovirions.
The horizontal axis lists the pooled serum, monoclonal antibody, or virus and cell only compared to the % infection on the vertical axis. The figure depicts the quantification of neutralizing antibodies in pooled sera at 30 days post-immunization from mice vaccinated with two doses of KLH-E or aluminum control vaccine. Each bar represents an individual HA-CoV-2 pseudovirion expressing a S protein sequence from a SARS-CoV-2 variant of concern (Beta, Delta, Omicron, XBB). The % infection was defined as the change (Δ) in relative luciferase units of pseudovirions incubated in test serum or control monoclonal antibody (mAb) compared to the luciferase signal of pseudovirions and cells only. The background luciferase signal of uninfected cells was also measured but was negligible (data not shown). The control antibody (27VB1) used in the assay is a broadly neutralizing rabbit anti-RBD S antibody (Virongy Biosciences). Each pooled serum was tested at a 1:2 dilution, and the control mAb was tested at a concentration of 5µg. All samples were tested in independent replicates and error bars represent the +/- standard deviation.

To further elucidate the breadth of immunity conferred by KLH-E vaccination, we measured the neutralization activity of pooled sera from mice receiving one or two doses of KLH-E or control aluminum vaccine against HCoV-229E, a genetically diverse human coronavirus with a highly divergent E protein sequence to that of our vaccine antigen (32% sequence homology). As illustrated in **Figure 3-17**, the data of which was derived from **Supplemental Figure 3-11**, there was no neutralization activity in pooled sera from mice immunized with one or two doses of aluminum vaccine as expected. In contrast, there was a significant reduction in viral infection when HCoV-229E virions were incubated with pooled sera from mice immunized with one or two doses of KLH-E at 30 days post-immunization. The magnitude of neutralization against HCoV-229E virions was proportional to the number of doses of KLH-E vaccine in pooled serum, with two doses of KLH-E yielding a significant reduction in viral infection compared to a single dose of aluminum vaccine at either time point (100% vs. 20% infection; 30 days post dose 1 aluminum vs. KLH-E; ***P=0.00002) (100% vs. 10% infection; 30 days post dose 2 aluminum vs. KLH-E; ***P=0.000016).

Collectively, these sets of experiments demonstrate that anti-E antibodies directed at the SARS-CoV-2 E protein transmembrane domain have functional neutralizing activity *in vitro,* and serve as proof-of-principle that such antibodies can potentially confer broad immune responses that can protect against emerging SARS-CoV-2 variants. This is exemplified by the near equal neutralization of KLH-E sera against all four HA-CoV-2 pseudovirions tested (Beta, Delta, Omicron, XBB), which is in stark contrast to the neutralization mediated by a control neutralizing anti-S monoclonal antibody. Additionally, the observation that anti-E antibodies induced by KLH-E vaccination have neutralizing activity against HCoV-229E further underscores the possibility of developing vaccines with broad immunity against various human coronaviruses.
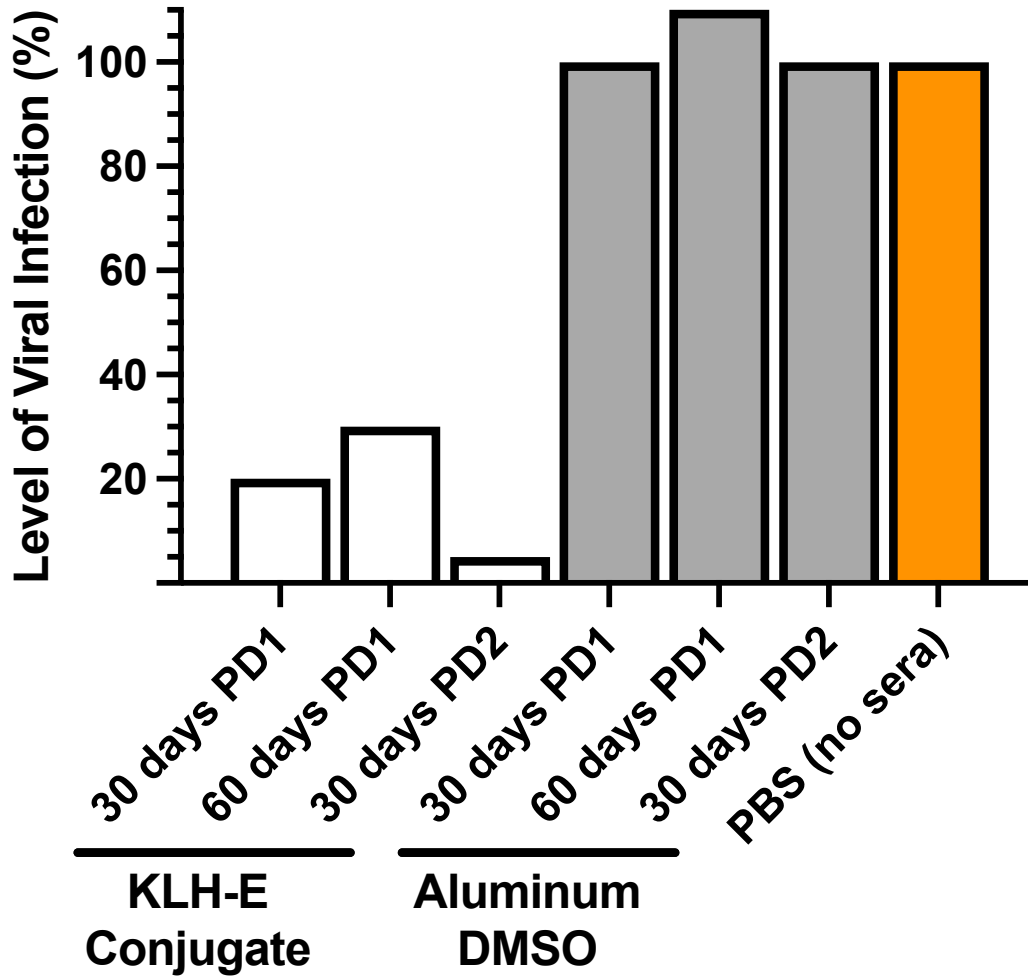
HCoV-229E

Figure 3-17. Neutralization of HCoV-229E. Quantification of neutralizing antibodies induced by one or two doses of KLH-E or control aluminum vaccine measured against human Coronavirus 229E (HCoV-229E). The horizontal axis lists the pooled serum from mice in each vaccine group against the vertical axis which lists the level of viral infection measured (%). Each pooled serum was tested at a 1:2 dilution by combining diluted serum and virus as described in the materials and methods. By 21 days post-infection, if most cells in a tissue culture well were lysed with observable cytopathic effect (CPE) by microscopy, the serum was defined as lacking neutralization activity. In contrast, if cells appeared normal in comparison to uninfected cells, the serum was defined as having neutralizing activity.

Cytokine profiling HCoV-229E infection in mouse lungs.

We investigated the immunogenicity of a novel KLH-E vaccine in mice and demonstrated that the vaccine induced protection *in vitro* via the induction of neutralizing antibodies against Ha-CoV-2 pseudovirions and HCoV-229E. The observation that antibodies induced by KLH-E vaccination had broad neutralizing activity against a genetically distant human coronavirus prompted us to investigate the feasibility of using HCoV-229E as a model coronavirus for future *in vivo* studies. The purpose of the following experiment was not to determine whether inbred BALB/c mice were capable of productive HCoV-229E infection, but rather to assess whether HCoV-229E was sufficiently virulent to evoke a similar inflammatory cytokine profile that is often observed in humans infected with SARS-CoV-2 and other coronaviruses broadly.

To address whether HCoV-229E was sufficiently virulent to evoke an inflammatory cytokine profile in infected BALB/c mice lungs, we intranasally infected mice with a high inoculum (1×10^5 PFU) of HCoV-229E and 24 hours later administered a mock treatment with PBS as described in the materials and methods. As a control, uninfected mice were also treated with a mock treatment of PBS in the same dosing interval. As illustrated in **Figure 3-18,** at 7 days post-treatment, we observed elevated levels of various cytokines and chemokines in both uninfected and HCoV-229E infected mouse lungs with the expression profiles of all mice in each experimental group generally in agreement and homogenous (see error bars; mean +/- standard deviation). To determine differences in the lung cytokine profiles of mice in each experimental group we performed a multiple t-test as described in the materials and methods. As depicted in **Figure 3-19**, there was a statistically significant decrease in CXCL1 (a neutrophil chemokine) in HCoV-229E infected mice compared to uninfected controls (**P=0.0098). In contrast, there was a significant increase in IL-2 (*P=0.0383) and IL-17 (*P=0.0441).
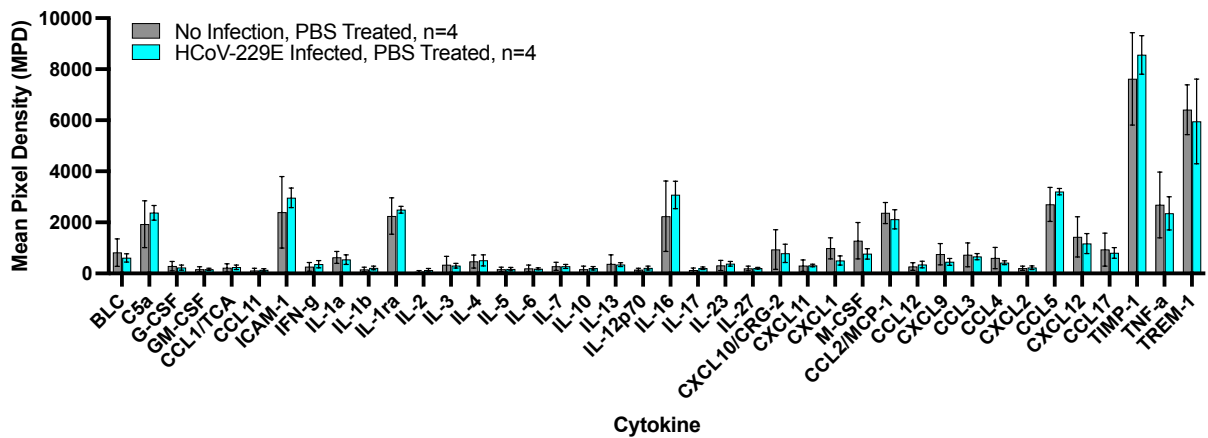
Figure 3-18. Effect of HCoV-229E infection on BALB/c mouse lung cytokines.
The horizontal axis lists the cytokines measured in independent replicates in homogenized mouse lungs tested in the absence of infection (n=4, gray bars) or infected with HCoV-229E (n=4, blue bars). The vertical axis is the calculated mean pixel density (MPD) of the cytokine array panel measured using ImageJ software as described in the materials in methods. Each bar represents the mean of four animals from each group and the error bars represent the +/- standard deviation.
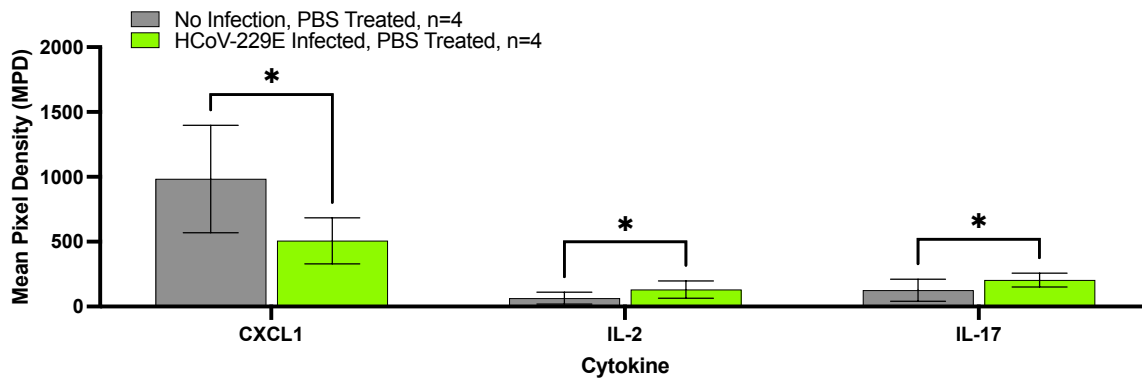
Figure 3-19. Statistically significant changes in cytokines in response to HCoV-229E infection The horizontal axis lists the cytokines for which there was a statistically significant change in expression of cytokine in homogenized mouse lungs comparing uninfected (gray bars) vs. HCoV-229E infected (green bars) mice. The vertical axis is the calculated mean pixel density (MPD) of the cytokine array panel measured using ImageJ software as described in the materials and methods. Each bar represents the mean of four animals from each group and the error bars represent the +/- standard deviation. Expression of CXCL1 decreased significantly (P=0.009) in HCoV-229E infected mice, whereas there was a modest but statistically significant increase in expression of IL-2 (P=0.038) and IL-17 (P=0.044) compared to uninfected control mice.

The lower levels of CXCL1 in uninfected compared to HCoV-229E infected mice was surprising, as the inverse would be expected given the function of CXCL1 as a neutrophil chemokine. CXCL1 is an interferon gamma induced protein and therefore we expected it to be more abundantly expressed in the context of viral infection [92]. Conceivably, HCoV-229E viral mechanisms that reduce the host interferon responses would be expected to benefit viral survival and infection. Our observation of reduced levels of CXCL1 in mice following infection with HCoV-229E are however in agreement with previous transcriptional studies in human cells infected with HCoV-229E which report that many interferon-stimulated genes and other transcripts of early antiviral response are repressed shortly after infection [93]. Surprisingly, compared to other highly pathogenic coronaviruses, previous studies report that HCoV-229E induces a more potent INF-I response than SARS-CoV-1 [93] . Furthermore, although expression of the human receptor aminopeptidase N (APN) is sufficient for HCoV-229E infection in cultured cells *in-vitro,* transgenic animal models require both the expression of human APN <u>and</u> genetic deletion of signal transducer and activator of transcription 1 (STAT1) [94], suggesting that interferons and STAT1 pathways play an essential role in controlling HCoV-229E infection [95] [96].

The significantly elevated levels of IL-2 and IL-17 cytokines in HCoV-229E infected mice compared to uninfected mice are consist with activation of the adaptive immune system and inflammation in the lung respectively. In humans infected with either HCoV-229E or SARS-CoV-2, IL-2 and IL-17 are both upregulated, and in the case of IL-17 contribute to the acute respiratory syndrome (ARD) severe pathology of SARS-CoV-2 infections [97]. Based on data generated in this pilot experiment on the similar cytokine profiles of mice infected with HCoV-229E compared to SARS-CoV-2 presented in the literature, we concluded that HCoV-229E may have some potential to serve as a model coronavirus for pre-clinical vaccine studies. Future *in vivo* studies are needed to investigate whether KLH-E vaccination protects against SARS-CoV-2 and other related human coronaviruses such as HCoV-229E.

DISCUSSION

Summary. In the present study, we are the first to investigate the immunogenicity of the SARS-CoV-2 E protein transmembrane domain using the immunogenic carrier KLH to enhance the immune response to the transmembrane domain hapten. We characterized in detail the humoral and cellular immune responses of mice induced by a novel KLH-E vaccine formation and presented *in vitro* evidence that antibodies directed at the SARS-CoV-2 E protein transmembrane domain can confer broad immunity against a genetically divergent coronavirus. These results are significant because they demonstrate for the first time to our knowledge, cross-protection against coronaviruses in different subfamilies mediated by antibodies targeting the envelope protein. These results suggest that the SARS-CoV-2 E protein transmembrane domain contains epitopes that can induce broadly cross-reactive immune responses.

Natural infection informed coronavirus vaccine development. Prior to the COVID-19 pandemic there was a body of scientific research on related coronaviruses that greatly informed and accelerated the development of SARS-CoV-2 vaccines. In response to the SARS-CoV-1 epidemic in 2002, scientist profiled the immune responses of convalescent persons and discovered that anti-S antibodies are immunodominant in SARS patients and had neutralizing activity [98, 99], an observation in agreement with subsequent S-based vaccines against SARS-CoV-1 and MERS-CoV in preclinical models [100, 101]. These findings set the precedent for current S protein vaccines, with the S protein being the primary antigen is all currently licensed SARS-CoV-2 vaccines. However, the increasing emergence of SARS-CoV-2 variants with S protein mutations that confer resistance to vaccines and natural immunity represent a growing concern.

In contrast to S-based vaccines, other coronavirus structural proteins have been tested as possible SARS-CoV-2 vaccine candidates with limited success. The N (Nucleocapsid) protein is abundant and expressed during the early stages of infection and is moderately conserved across coronaviruses. However, the N protein is not displayed on the surface of coronaviruses and does not induce neutralizing antibodies [102]. Although, one study found CD4+ and CD8+ T-cell dependent partial immunity against SARS-CoV-1 and MERS-CoV when N was expressed using a viral vector vaccine [103]. Similarly, the M (Membrane) protein is the most abundant structural protein on the surface of SARS-CoV-2 virion and although it is reported to induce high antibody titers in animals, anti-M antibodies appear to lack neutralization activity [104].

As with the related SARS-CoV-1, during natural infection with SARS-CoV-2, IgM, IgG, and IgA antibodies are induced to varying degrees against all four structural proteins: N, S, M, and E. Data from several large serological studies report that anti-E antibodies are induced in significantly lower orders of magnitude compared to other coronavirus structural proteins [105]. Broadly speaking, protein antigens occurring in high abundance, repeating units (multimers), or which contain pathogen associated molecular patterns (PAMP's) induce greater responses by the adaptive immune system. Conceivably, the weak induction of anti-E antibodies during SARS-CoV-2 infection may be a result of the E protein being both the smallest and least abundant protein on the surface of the SARS-CoV-2 virion.

Previous E vaccine studies. Relatively few studies have investigated the immunogenicity of the coronavirus E protein. Although it is moderately conserved among coronaviruses, the E protein is small at only 75-85 amino acids which means there are fewer B and T-cell epitopes for the immune system to target. Additionally, the E protein is the least abundant structural protein on coronavirus virions. Previous studies have investigated the immunogenicity of the coronavirus E protein alone or in combination with other structural proteins and have reported mixed results. In a study by Buchholz *et al*, the authors expressed the SARS-CoV-1 structural proteins S, M, E, and N alone or in combination using a recombinant parainfluenza virus vector vaccine administered as a single intranasal dose in a hamster model susceptible to SARS-CoV-1 infection. A single dose of vaccine expressing only the S protein induced a high titer of neutralizing serum antibodies, but a vaccine co-expressing S+M+E did not augment the neutralizing antibody response. Conversely, in the absence of S, co-expression of M+E or M+N did not induce serum neutralizing antibodies. Interestingly, hamsters immunized with a S+M+E vaccine had slightly greater protection than hamsters immunized with S vaccine alone when challenged with SARS-CoV-1 *in vivo*. In contrast, a M+E+N vaccine conferred no protection in hamsters challenged with SARS-CoV-1 [106]. These results suggest that *in vitro* neutralization assays may underestimate the protection of antibodies induced to non-S structural proteins, and while antibodies to non-S structural proteins may lack neutralization activity *in vitro*, they may cooperate synergistically with anti-S antibodies to confer protection *in vivo*.

Later, when it was found that the coronavirus E protein is a virulence factor, subsequent studies targeted the E protein not as a vaccine candidate but rather as a target for the development of live-attenuated SARS-CoV-1 vaccines. Deletion of the E protein in SARS-CoV-1, results in highly attenuated SARS-CoV-1ΔE virions that retain immunogenicity and confer protection against lethal challenge with wildtype SARS-CoV-1 in animal models [107]. It is interesting to note that although the SARS-CoV-1 and SARS-CoV-2 E proteins share 95% sequence homology, for unknown reasons the E protein is dispensable in SARS-CoV-1 but essential in SARS-CoV-2. For this reason, targeted disruption of the E protein used to generate experimental live-attenuated SARS-CoV-1 vaccines cannot be used to generate live-attenuated self-replicating SARS-CoV-2 vaccines.

In a recent study by Chen *et al*., mice immunized intramuscularly with three doses of a DNA vaccine expressing either the SARS-CoV-2 E, M, or E+M fusion proteins, failed to elicit antibodies but did generate robust antigen-specific T-cells [108]. Using synthetic 15-mer overlapping peptides that covered the entire sequence of the SARS-CoV-2 E protein, the authors report that the immunodominant T-cell epitopes were located at the c-terminus of the protein. Interestingly, the third largest population of T-cells recognized a peptide sequence in the transmembrane domain of the E protein (E07: LAILTALRLCAYCCN). In this context, our data agree with the induction of robust cellular immunity induced by our KLH-E vaccine which shared common epitopes (Epitope E07 in Chen *et al*: <u>LAILTALRL</u>CAYCCN vs. TLIVNSVLLFLAFVVFLLVT<u>LAILTALRL</u>C in the KLH-E vaccine). In contrast, in a previous but similar study by Jin *et al.*, mice immunized intramuscularly with five doses of a DNA vaccine each expressing the SARS-CoV-1 E, M, or N protein induced robust antibodies and antigen-specific T-cells to all three structural proteins [109]. These discordant result are two examples among many additional studies that report mixed results using DNA vaccines to investigate the immunogenicity of non-S coronavirus structural proteins [109-112]

Nevertheless, the discordant results of the limited studies investigating the immunogenicity of the coronavirus E protein have many possible explanations. First, it is possible that the coronavirus E protein may not maintain its native structure or be expressed in sufficient quantities to be immunogenic when expressed on the surface of viral vectors. Second, given the small size of the E protein, even if expressed in its native state the E protein may be overshadowed by larger and more immunogenic endogenous proteins present in the viral vector used for vaccination. Third, the immune responses to DNA vaccines may vary greatly depending on factors such as number of vaccine doses, route of administration, and inclusion of specific adjuvants. Notwithstanding, our data serve as proof-of-principle that the SARS-CoV-2 E protein transmembrane domain can indeed be immunogenic at least when conjugated to KLH, and we provide direct experimental evidence that the coronavirus E protein transmembrane domain contains important protective epitopes that were previously predicted to exist in published bioinformatic studies [113].

Safety concerns. A theoretical safety concern of a SARS-CoV-2 based E protein vaccine is the possible induction of a maladaptive host inflammatory response and immunopathology. The SARS-CoV-2 E protein was recently reported to be a Toll-like receptor 2 (TLR) ligand [114] [115], and the protein alone in the absence of infection is sufficient to induce inflammation and cytokine dysregulation when administered to mice [114, 116]. These hyperactive inflammatory cytokine responses may be related to the viroporin activity of the full-length SARS-CoV-2 E protein. In SARS-CoV-1 which shares a 100% E protein transmembrane domain sequence homology to SARS-CoV-2, the E protein viroporin activity interferes with cellular calcium Ca^{2+} homeostasis which serves as a trigger for the NOD-like family pyrin domain containing 3 (NLRP3) inflammasome leading to IL-1ß overproduction and subsequent inflammation and immunopathology *in vitro* and *in vivo* [73, 76, 77]. A recombinant transmembrane protein epitope vaccine conjugated to an immunogenic carrier, such as our KLH-E vaccine, may mitigate some of the potential safety risks associated with E protein-based vaccines. Additionally, the incorporation of discrete mutations (N15A, V25F) that ablate viroporin activity [77] of the coronavirus E protein transmembrane domain could further reduce potential safety concerns if such mutations did not affect protective epitopes. Notably, the SARS-CoV-2 E protein is not the only structural protein and vaccine candidate for which safety concerns exist. Recently, studies have reported that the SARS-CoV-2 S protein, the primary antigen in all currently licensed vaccines, can trigger endothelial and epithelial barrier dysfunction *in vitro* and vascular leak *in vivo* independent of viral infection and the ACE2 receptor. The mechanisms underlying endothelial and epithelial barrier dysfunction and vascular leak in response to the S protein appear to be mediated by glycosaminoglycans, integrins, and the TGF-ß signaling axis [117]. Despite the possible safety concerns regarding both the S and E proteins, they should not be disqualified as important vaccine antigens. For example, toxoid-based vaccines stimulate immunity to microbial toxins that would normally be harmful to the host, but the toxoids are genetically or chemically inactivated to make them safe and tolerable as vaccine antigens Given this precedent, a similar approach may be helpful in reducing possible safety concerns with SARS-CoV-2 S and E proteins in future vaccines.

Cross-reactivity of anti-E antibodies and T-cells induced by KLH-E vaccination. The ability of both humoral and cellular arms of immunity to exert their effector functions depends on precise recognition between a specific epitope and antibody/t-cell receptor. We demonstrated that a novel KLH-E vaccine could induce specific anti-E antibodies and T-cells as measured using a 30-amino acid synthetic peptide corresponding to the SARS-CoV-2 E protein transmembrane domain. Given that the epitope of interest was conjugated to an immunogenic carrier we had initial concerns about whether antibodies and T-cells induced to the SARS-CoV-2 E transmembrane domain could recognize native full-length E protein. Despite the availability of commercially produced purified recombinant SARS-CoV-2 E protein, we questioned whether such protein produced in bacteria or other host organism would retain its native structure given the fact that the E protein is an integral single pass protein that is predicted to require a lipid membrane to maintain its structure. Furthermore, most recombinant SARS-CoV-2 E proteins are modified at the sequence level to facilitate periplasmic expression in *E.coli* or fused to tags such as maltose binding protein to improve solubility and purification. All the preceding factors could reasonably influence the overall E protein structure and possible conformational epitopes. To our surprise, anti-E antibodies and T-cells induced by KLH-E vaccination recognize native full-length E protein in cell lysates infected with the genetically divergent but related human coronavirus HCoV-229E *in vitro* by ELISA and by T-cell activation assay. This result was surprising given that the E protein transmembrane domains between SARS-CoV-2 and HCoV-229E have extremely low sequence homology at a mere 32%. These results are significant because they suggest that the SARS-CoV-2 E protein transmembrane domain contains epitopes that can induce broad immune responses that cross-react with genetically divergent human coronaviruses.

In vitro neutralization activity. We demonstrated that the KLH-E vaccine not only could elicit anti-E transmembrane antibodies, but we measured the functional activity of the antibodies *in vitro* by neutralization assay and found that pooled sera from KLH-E vaccinated mice had neutralizing activity against HCoV-229E and HA-CoV-2 pseudovirions. These results are important because they suggest that the SARS-CoV-2 E protein transmembrane domain appears to be accessible to antibody binding even when the E protein is expressed in different modalities. Furthermore, our experiments demonstrate that KLH-E vaccination induces both specific and functional anti-E antibodies that have neutralization activity against HCoV-229E and HA-CoV-2 pseudovirions.

HCoV-229E as a model coronavirus. Given that the primary antigenic target of our vaccine is the SARS-CoV-2 E protein transmembrane domain, it is reasonable to question why we used HCoV-229E as a model virus rather than SARS-CoV-2. There are several important justifications and benefits of using HCoV-229E as a model to investigate the potential of HCoV-229E as a model coronavirus for future *in vivo* vaccine studies. First, HCoV-229E can be handled under lower biosafety conditions and unlike SARS-CoV-2, HCoV-229E generally only causes mild respiratory infections in immunocompetent individuals. Second, since HCoV-229E is a member of the alphacoronavirus subfamily, it allows us to develop a framework to understand the limitations and breadth of immunity potentially induced by novel vaccines against coronaviruses in different subfamilies. Lastly, if the 30 amino acid SARS-CoV-2 E protein transmembrane domain sequence in the KLH-E vaccine induced cross-reactive antibodies that recognize other

160

divergent E protein sequences, this information could potentially inform the development of epitope-based pan-coronavirus vaccines.

Mouse models of coronavirus infection. Since the emergence of SARS-CoV-2, several animals have been evaluated for *in vivo* studies including: mice, hamsters, ferrets, and non-human primates as models for infection [118]. While each model system has benefits and limitations, together the use of animal models has led to a greater understanding of the host response, immunopathogenesis, and surrogates for the development of vaccines and therapeutics. Several inbred mouse species have been evaluated for susceptibility to SARS-CoV-2 infection as well as for other human coronaviruses. In the absence of the human ACE2 receptor, most mice do not exhibit severe clinical signs when infected with SARS-CoV-1 or SARS-CoV-2 with the exception of senescent or immunocompromised mice [119, 120]. However, even transgenic mice expressing the human ACE2 receptor do not fully recapitulate the scope and severity of SARS-CoV-2 clinical manifestations in humans [121, 122]. Thus, while different transgenic mouse models are susceptible to SARS-CoV-2 infection they vary in disease severity, which may be due to differences in tissue distribution and expression level of the human ACE2 transgene.

In our *in vivo* studies, we used inbred BALB/c mice in a pilot experiment to understand whether HCoV-229E maintains its virulence and can induce inflammatory responses in the lung typically observed in SARS-CoV-2 and other human coronaviruses. We found that despite HCoV-229E having lower pathogenicity compared to other human coronaviruses, intranasal infection in BALB/c mice resulted in elevated inflammatory cytokines. HCoV-229E uses the human aminopeptidase N (hAPN/CD13) receptor for viral entry which is not present in BALB/c mice, therefore infection with HCoV-229E is not expected to result in high levels of productive viral infection. Though clearly, the introduction of HCoV-229E into the lungs of mice was sufficient to evoke some inflammatory responses that are like those observed in humans. A transgenic mouse model has been developed for HCoV-229E that expresses the human APN receptor [123]. Interestingly, while expression of APN is sufficient for HCoV-229E infection in cultured cells, it is necessary but not sufficient for HCoV-229E infection *in vivo.* Productive infection of HCoV-229E in mice requires both the expression of the human APN receptor, as well as genetic ablation of endogenous mouse signal transducer and activator of transcription 1 (STAT1) which controls the regulation of many important immune genes. Thus, while a transgenic mouse model exists for HCoV-229E, it comes with significant drawbacks that immunologically do not fully recapitulate the infections observed in humans.

Limitations. A limitation of the present study is the absence of complementary vaccine immunogenicity data for the unconjugated SARS-CoV-2 E protein transmembrane domain. Our decision to conjugate the SARS-CoV-2 E protein transmembrane domain to the KLH carrier was justified based on concerns that the hapten alone would be insufficiently large to be immunogenic. Although conjugation with KLH likely enhanced the immune responses to the SARS-CoV-2 E protein transmembrane domain it may have also contributed to some suboptimal results in our study. First, whether due to size or differences in relative immunogenicity, an overwhelming majority of antibodies and T-cells induced by the KLH-E vaccine were directed at the KLH carrier. Second, despite the induction of anti-E IgG after a single dose of KLH-E vaccine, a second dose of vaccine did not elicit an overall increase in anti-E IgG. In contrast, the KLH-E vaccine induced robust anti-KLH IgG after a single dose, and significantly increased

following subsequent immunization. These results suggest a skewed antibody repertoire directed at the KLH carrier rather than the SARS-CoV-2 E protein transmembrane domain. Conceivably, primary immunization with KLH-E and boosting with a different immunogenic carrier E conjugate, could potentially overcome the skewed antibody repertoire. At 30 days post-immunization with a single dose of KLH-E vaccine we observed a decline in anti-E IgG antibodies and T-cells, and although these decreases did not reach statistical significance, they suggest that immunity may wane rapidly over time which is a common feature of many recombinant protein subunit vaccines. Long-term studies are necessary to understand the kinetics of humoral and cellular immunity induced by SARS-CoV-2 E protein transmembrane domain vaccines. The most significant limitation of our study is the absence of a clear molecular mechanism to explain how anti-E antibodies neutralize HCoV-229E virions *in vitro,* and the absence of complementary *in-vivo* data of the protection induced by KHL-E vaccination. These experiments are ongoing, but unfortunately due to limitations of time our *in-vivo* KLH-E vaccination and challenge experiments are not included in the present dissertation.

Hypothetical mechanism of anti-E antibody mediated neutralization. Neutralization of viruses by antibody generally occurs via three distinct mechanisms.

The first mechanism of antibody mediated viral neutralization is straightforward and involves binding of antibody to a surface exposed viral entry protein and sterically hindering the engagement of the entry protein with a host cell receptor. Antibodies directed at the S protein of SARS-CoV-2 are believed to function in this manner due to several lines of evidence. First, the natural humoral response in SARS-CoV-2 infection is strongly biased towards the S protein, and antibodies specifically directed at the receptor binding domain (RBD) of the S protein have the greatest neutralization potency. As SARS-CoV-2 evolves, there is significance evidence that the RBD domain of the S protein is under strong selective pressure and is highly mutable, which may explain why natural infection and vaccination do not prevent re-infection by SARS-CoV-2 variants with highly mutated RBD domains of the S protein.

A second mechanism of viral neutralization depends on antibody but is mediated by the complement system. In this mechanism, antibody-antigen complexes are recognized by innate immune molecules known as complement. Specifically, antibody-antigen complexes are recognized by the globular heads of the C1q molecule which serves as a substrate for subsequent complement effector molecule activation in the classical pathway and can directly lead to membrane lysis and thus neutralization of the viral particle. Classical pathway complement mediated neutralization has been reported *in vitro* in SARS-CoV-2 for anti-S antibodies and is more frequent in individuals who elicited high levels of certain IgG subclasses (IgG1 and IgG3) [124].

 A third mechanism of antibody mediated viral neutralization occurs via phagocytosis. In this scenario, "non-neutralizing" antibodies bind to one or more surface exposed viral protein that may or may not be involved in cell entry. Viruses coated in "non-neutralizing" antibodies are highly attractive since the Fc region of antibodies act as a potent opsonin and are recognized by Fc receptors on the surface of phagocytic cells. Via this mechanism, viral particles coated in "non-neutralizing" antibodies can be transported into the cytoplasm and be inactivated by lysosomal degradation. Alternatively, viral particles that escape the phagolysosome can be

detected by cytoplasmic Fc receptor sensors such as TRIM21 which mediates proteosome-dependent destruction of virus particles [125]. The ability of "non-neutralizing" antibodies to contribute to opsonization mediated neutralization has been described previously in SARS-CoV-2 using "non-neutralizing" anti-S monoclonal antibodies [126]. Interestingly, in the previously cited study by *Bahnan et al.,* the authors report that "non-neutralizing" anti-S monoclonal antibodies can contribute to opsonization mediated neutralization both *in vitro* (using THP-1 monocytic cells) as well as *in vivo* in ACE2 transgenic mice infected with SARS-CoV-2.

In our study we demonstrated that vaccination with KLH-E can induce specific and functional anti-E antibodies directed at the SARS-CoV-2 E protein transmembrane domain. We measured the functional activity of the anti-E antibodies against HCoV-229E and HA-CoV-2 pseudovirions by neutralization assay but did not dissect in detail the molecular mechanisms responsible for the neutralization we observed. Based on the three mechanisms of antibody dependent neutralization, it is plausible that anti-E antibodies could mediate protection by one or more mechanism and may depend on the *in vitro* assay conditions. In our neutralization assay using the HA-CoV-2 pseudovirion system, it is unlikely that the anti-E neutralization we observed was due to inhibiting receptor mediated viral entry. In the study that first developed and validated the HA-CoV-2 pseudovirion system, the authors demonstrated that expression of the SARS-CoV-2 S protein was both necessary and sufficient to infect the target HEK293T(ACE2/TMRPSS2) cells used in their system, and infection did not require the expression of other SARS-CoV-2 structural proteins (N, M, E). The alphavirus the authors used as a pseudovirion backbone could not infect HEK293T(ACE2/TMRPSS2) cells without expressing the S protein [88]. Taken together, the anti-E antibody neutralization we observed was likely not due to sterically hindering engagement of the S protein on the HA-CoV-2 pseudovirions and the ACE2 receptor on the HEK293T (ACE2/TMRPSS2) cells.

We also observed anti-E antibody neutralization of HCoV-229E in human foreskin fibroblast (HFF) cells. Unlike SARS-CoV-2, the S protein of HCoV-229E uses the human amino peptidase N (APN/CD13) receptor for viral entry, which is abundantly and constitutively expressed in human fibroblasts such as our HFF cells [127]. Among human coronaviruses the E protein has not been reported to be involved in viral entry and is the sparsest structural protein on the surface of viral particles, thus intuitively we would not expect anti-E antibodies to sterically hinder engagement of HCoV-229E spike to the APN entry receptor.

Although we did not measure whether anti-E antibody neutralization of HA-CoV-2 pseudovirions or HCoV-229E was complement-dependent, this could potentially be ruled out by depleting the test serum of endogenous complement by heat inactivation and measuring neutralization activity. For neutralization studies, we used small freshly thawed aliquots of test serum which contained endogenous and presumably active complement. Nevertheless, both the HA-CoV-2 pseudovirions and HCoV-229E did not appear to be sensitive to complement alone in the absence of anti-E antibody, since virions from both viruses were highly infectious and serum from aluminum vaccinated mice had negligible neutralization activity.

Hypothetically, anti-E antibodies could function in an opsonophagocytic dependent neutralization manner mediated by specific Fc receptors expressed on cells used for *in vitro* neutralization assays. There are many Fc receptors that exist in humans, and each varies with

163

respect antibody isotype and subtype they recognize. Fc receptors are generally expressed on immune cells (B-cells, Mast cells, Monocytes, Macrophages, NK cells, Eosinophils); however, FcRn is one example of a Fc receptor that is expressed on immune cells as well as endothelial and epithelial cells [128]. Epithelial and Endothelial cells are two of the most widely used cell types used to assess the neutralization activity of antibodies against viruses *in vitro,* including SARS-CoV-2. Importantly, FcRn binds to IgG and has the greatest ligand binding affinity in cellular endosomes, as would be the case if anti-E antibodies neutralized coronaviruses in an opsonophagocytic manner. Taken together, the concept that "non-neutralizing" antibodies could indeed neutralize SARS-CoV-2 via unorthodox mechanisms such as opsonization is both surprising and provocative. In this context, the true neutralization capacity of both structural and non-structural antibodies directed at SARS-CoV-2 proteins which were previously deemed "non-neutralizing" may perhaps warrant re-examination.

Antigenic stability of the E protein. Unlike the highly mutable S protein, largescale genomic studies of publicly available viral genomes have demonstrated that the SARS-CoV-2 E protein is highly conserved. In a study investigating the mutational landscape of the SARS-CoV-2 E protein in global isolates from the start of the pandemic till August 2020, Rahman *et al.* analyzed 81,818 SARS-CoV-2 genomes and reported that 98.8% of the E protein in globally circulating SARS-CoV-2 strains is highly conserved [82]. These initial observations are further supported by one of the largest studies to date by Abavisani *et al.* which analyzed the amino acid sequences of all SARS-CoV-2 structural proteins globally from the declaration of the COVID-19 pandemic to January 2022, where the authors report that 96.40% (n=6,524,654 genomes analyzed) of the E protein amino acid sequences exhibited no mutation from the original Wuhan-2019 SARS-CoV-2 isolate [129]. Notably, the latter study also reported that in the extremely small number of SARS-CoV-2 isolates with E protein mutations, most had ≤4 non-synonymous mutations total, with mutations occurring almost exclusively at the N and C terminus of the E protein [129]. Taken together, these studies implicate that mutations in the SARS-CoV-2 E protein accumulate more slowly than in other structural proteins (S, M, and N), and suggest an evolutionary pressure to maintain E protein transmembrane domain amino acid residues. For these reasons, as well as the data presented in our study, the SARS-CoV-2 E protein transmembrane domain represents an attractive vaccine antigen that warrants serious consideration as it has the potential to confer broad protection against emerging SARS-CoV-2 variants and potentially other genetically diverse human coronaviruses.

Broad Implications. Traditional vaccine approaches target highly abundant and surface exposed epitopes, therefore the concept that antibodies can be induced to a transmembrane domain of a protein and have neutralizing activity is both provocative and unexpected. Conventionally, it would be assumed that even if antibodies could be induced to a transmembrane domain, they would be unable to exert their effector functions if the epitope was inaccessible on the surface of the microorganism. To date, only the transmembrane domain of the SARS-CoV-2 E protein (PDB: 7K3G) has been determined by nuclear magnetic resonance (NMR) spectroscopy data [70], while the N and C termini structures of the protein remain in incompletely understood. Overall, the SARS-CoV-2 E protein appears to be conformationally flexible, with some residues possibly accessible under different temperature, membrane, and pH conditions. Taken together, we cannot fully exclude the possibility that the transmembrane domain of the SARS-CoV-2 E protein may be exposed or more accessible than our current assumption based on the primary

amino acid hydrophobic residues and current molecular models. Nevertheless, there is some precedent for unorthodox and experimental vaccines designed to induce immunity to occluded regions of proteins such as the hemagglutinin stalk-based vaccines in Influenza virus [130-132] and membrane proximal external regions (MPERS) of gp41 in Human Immunodeficiency Virus (HIV) [133, 134]. One important distinction between the preceding vaccines is that while both aim to induce immunity to occluded regions of proteins, neither qualifies as a true transmembrane domain vaccine. The present study serves as proof-of-principle that both humoral and cellular immunity can indeed be induced to a transmembrane region of a protein using the SARS-CoV-2 E protein as a model. Furthermore, our study demonstrates that antibodies directed at the SARS-CoV-2 E protein transmembrane domain can have functional activity, a result which challenges and expands our traditional understanding of basic vaccinology. In this context, the vaccine potential of many highly conserved transmembrane proteins in various human microbial pathogens may warrant reconsideration.

Future Directions. Future studies are needed to determine whether antibodies directed at the transmembrane domain of the SARS-CoV-2 E protein are broadly and equally neutralizing against SARS-CoV-2 and other human coronaviruses. Most of our current understanding of cross-protection against genetically diverse coronaviruses in different subfamilies relies on the contributions of anti-S antibodies, which are believed to offer some limited cross protection since 3 of 7 (SARS-CoV-1, SARS-CoV-2, HCoV-NL63) human Coronaviruses encode S proteins that use human ACE2 as a viral entry receptor. Since the transmembrane domain of the E protein is moderately conserved among human coronaviruses, it represents an attractive vaccine antigen with the potential to induce pan-coronavirus immunity. Despite overall sequence conservation at the transmembrane level, it is unclear why the E protein is essential in some coronaviruses, yet dispensable in others. One possible explanation is that some human coronaviruses encode accessory proteins with redundant functions to that of the E protein. Nevertheless, the inclusion of the SARS-CoV-2 E protein transmembrane domain in future generation coronavirus vaccines may broaden protection not only against emerging SARS-CoV-2 variants, but also endemic human coronaviruses (HCoV-229E, HCoV-NL64, HCoV-OC43, HCoV-HKU1) that cause seasonal respiratory infections [135].

MATERIALS AND METHODS

SARS-CoV-2 E peptide. The 30 amino acid sequence (H-TLIVNSVLLFLAFVVFLLVTLAILTALRLC-OH) corresponding to the full length transmembrane domain of the SARS-CoV-2 Envelope protein was synthetically produced (Pepscan, Netherlands) as either a purified peptide (77.1% Purity, UPLC/UV_{215}) or conjugated to the Keyhole Limpet Hemocyanin carrier protein (77.1% Purity, UPLC/UV_{215}). The extreme hydrophobicity of the transmembrane domain sequence proved challenging for synthesis and is reflected in the purity of each custom peptide.

Mouse immunization. The protocol was approved by the Institutional Animal Care and Use Committees (ACUC) at the University of California, Berkeley. At approximately 6 weeks of age, groups of male inbred BALB/c mice were purchased from Jackson Laboratories and housed at the University of California, Berkeley animal facility for 1 week prior to beginning the immunogenicity study. At the time of the first dose of vaccine, the ages of the BALB/c mice were 8 weeks, and mice were assigned to different vaccine groups. The mice were immunized with 50 µg of KLH-E per dose intramuscularly with 0.5mg Aluminum hydroxide salt in DMSO.

Viral strains and cell lines. Human coronavirus 229E (HCoV-229E) was supplied by American Type Culture Collection (ATCC). Human primary foreskin Fibroblast (HFF) (CC-2509) were obtained from Clonetics (San Diego). Cells were maintained in Dulbecco's modified eagles medium (DMEM) (Gibco) supplemented with 10% heat inactivated fetal bovine serum (Bio Whittaker, Walkersville, Maryland) and incubated at 37ºC and 5% CO_2. Cells were passaged twice a week using trypsin (0.25%)-EDTA and were not used beyond passage 12. Viral stocks were prepared by infecting cells at a multiplicity of infection of 0.01 for 4 to 7 days until significant cytopathic effect (CPE) was observed. Infected cells were subjected to 3 freeze/thaw cycles, and infected-cell lysate was stored at -80ºC.

Recombinant protein ELISA. Serum IgG antibody titers to Keyhole Limpet Hemocyanin (KLH) (EDM Millipore, CAT# 374817-50MG), KLH-E (Pepscan, Netherlands), and the purified 30 amino acid (TLIVNSVLLFLAFVVFLLVTLAILTALRLC) E-peptide conjugate (Pepscan, Netherlands). 96-well Medisorp ELISA plates (Thermo Fisher) were coated with 2µg of antigen (KLH, KLH-E, or E peptide) in ELISA coating buffer (50mM carbonate-bicarbonate buffer, 1.59g Na2CO3 + 2.93g NaHCO3 in 1L diH20, pH 9.4) and sensitized overnight at 4ºC. Next, plates were extensively washed by adding 200µl of ELISA wash buffer (0.05M Tris, 0.138M NaCl, 0.0027 M KCl, 0.05% Tween-20 pH 8.0 at 25ºC) to each well and aspirated, with each wash repeated in triplicate. After washing, the wells of the ELISA plate were blocked with 200µl per well of 5% non-fat powdered milk in PBS at room temperature for two hours. After blocking, the plate was washed as described previously and 100µl of serum diluted in ELISA dilution buffer (5% Bovine Serum Albumin, BSA; VWR) was added to the wells of the plate and incubated for one hour at room temperature on an orbital shaker (Roto mix, Thermolyne). After the primary antibody incubation, the plate was once again washed as previously described and 100µl of goat anti-mouse IgG (H+L) (Cell signaling technologies, Massachusetts, USA) at a 1:2,000 dilution and incubated for one hour at room temperature on an orbital shaker. Next, the plate was once again washed as previously described, and 100µl of chemiluminescent TMB

substrate (BioLegend, California, USA) was added to each well and optical density measured at OD=640nm in a plate reader (Spectramax ® M2, Molecular Devices).

Infected cell lysate ELISA. Serum anti-E IgG antibody reactivity was measured against HCoV-229E infected human MRC5 cell lysates or uninfected control cell lysates.
Cell lysates were prepared by growing MRC5 cells in Eagles Minimum Essential Medium (EMEM) (Gibco) supplemented with 10% heat inactivated fetal bovine serum (Bio Whittaker, Walkersville, Maryland) and incubated at 37ºC and 5% CO_2. Once the cells reached 80% confluency, they were infected with HCoV-229E at an MOI= 1, or mock infected with 1X PBS. After 7 days, cytopathic effect (CPE) was observed and both infected and mock infected MRC5 cells were subjected to 3 freeze/thaw cycles, quantified by measuring absorbance at 280nm in spectrophotometer, followed by storing at -80ºC. 96-well Medisorp ELISA plates (Thermo Fisher) were coated with 50µg/mL of cell lysate diluted in coating buffer (50mM carbonate-bicarbinate buffer, 1.59g Na2CO3 + 2.93g NaHCO3 in 1L diH20, pH 9.4) and sensitized overnight at 4ºC. Next, plates were extensively washed by adding 200µl of ELISA wash buffer (0.05M Tris, 0.138M NaCl, 0.0027 M KCl, 0.05% Tween-20 pH 8.0 at 25ºC) to each well and aspirated, with each wash repeated in triplicate. After washing, the wells of the ELISA plate were blocked with 200µl per well of 5% non-fat powdered milk in PBS at room temperature for two hours. After blocking, the plate was washed as described previously and 100µl of serum diluted in ELISA dilution buffer (5% Bovine Serum Albumin, BSA; VWR) was added to the wells of the plate and incubated for one hour at room temperature on an orbital shaker (Roto mix, Thermolyne). After the primary antibody incubation, the plate was once again washed as previously described, and 100µl of goat anti-mouse IgG (H+L) AP secondary antibody (Cell signaling technologies, Massachusetts, USA) was added at a 1:2,000 dilution and incubated for one hour at room temperature on an orbital shaker. Next, the plate was once again washed as previously described, and 100µl of chemiluminescent TMB substrate (BioLegend, California, USA) was added to each well and optical density measured at OD=640nm in a plate reader (Spectramax ® M2, Molecular Devices).

Pseudovirion ELISA. Serum anti-E IgG antibody reactivity was measured against four commercially available hybrid alphavirus pseudovirions which expressed all four SARS-CoV-2 structural proteins (Spike (S), Membrane (M), Nucleocapsid (N), Envelope (E)) (Virongy Biosciences; Manassas, VA, USA). The four pseudovirions tested differed only in the Spike protein sequences they expressed and corresponded to sequences from SARS-CoV-2 variants of concern: Beta (B1.351), Delta (B1.617), Omicron (B1.1.529), XBB (XBB.1). The HA-CoV-2 pseudovirions have been previously described in detail and validated for rapid quantification of neutralizing antibodies. To measure anti-E IgG antibody reactivity, 96-well Medisorp ELISA plates (Thermo Fisher) were coated with 2000 PFU/mL of HA-CoV-2 Pseudovirions (Virongy Biosciences, Manassas, VA, USA) diluted in coating buffer (50mM carbonate-bicarbinate buffer, 1.59g Na2CO3 + 2.93g NaHCO3 in 1L diH20, pH 9.4) and sensitized overnight at 4ºC. Next, plates were extensively washed by adding 200µl of ELISA wash buffer (0.05M Tris, 0.138M NaCl, 0.0027 M KCl, 0.05% Tween-20 pH 8.0 at 25ºC) to each well and aspirated, with each wash repeated in triplicate. After washing, the wells of the ELISA plate were blocked with 200µl per well of 5% non-fat powdered milk in PBS at room temperature for two hours. After blocking, the plate was washed as described previously and 100µl of serum diluted in ELISA dilution buffer (5% Bovine Serum Albumin, BSA; VWR) was added to the wells of the plate and

incubated for two hours at room temperature on an orbital shaker (Roto mix, Thermolyne). After the primary antibody incubation, the plate was once again washed as previously described, and 100µl of goat anti-mouse IgG (H+L) AP secondary antibody (Cell signaling technologies, Massachusetts, USA) was added at a 1:2,000 dilution and incubated for one hour at room temperature on an orbital shaker. Next, the plate was once again washed as previously described, and 100µl of chemiluminescent TMB substrate (BioLegend, California, USA) was added to each well and optical density measured at OD=640nm in a plate reader (Spectramax ® M2, Molecular Devices).

Isolation of mouse spleenocytes. Individual mouse spleenocytes were isolated from freshly sacrificed animals and placed in a small sterile petridish with 5mL of Hanks Balanced Salt Solution (HBSS) (Sigma). In a sterile biosafety hood, the spleen was carefully minced into small pieces (~0.2cm) with a sterile razor blade before transferring the material to a 70µm cell strainer over a 50mL falcon conical tube. Using the plunger end of a sterile syringe, the material was homogenized through the cell strainer and 10mL of sterile 1X PBS (Gibco) was used to wash the cell strainer. After collecting the homogenized spleen tissue, the suspension was centrifuged for 5 minutes (400-600 x g) at 4ºC. After centrifugation, the supernatant was discarded and the cell pellet was resuspended in 5mL of sterile 1X Red Blood Cell (RBC) Lysis Buffer [155 mM NH_4Cl, 12 mM $NaHCO_3$, 0.1 mM EDTA) and incubated for 5 minutes on ice. After incubation, the cell suspension was centrifuged for 5 minutes (400-600 x g) at 4ºC and the supernatant discarded. The cell pellet was then washed with 10mL of cold sterile 1X PBS and centrifuged again as before. After the final centrifugation, the supernatant was discarded and the spleenocyte cell pellet was resuspended in 2mL of freezing buffer (90% heat inactivated Fetal Bovine Serum (ΔFBS, Bio Whittaker; Walkersville, Maryland) and 10% DMSO (ATCC)). The 2mL volume was separated into two 1mL aliquots, flash frozen, and stored at -80ºC.

Antigen-specific T-cell assay. Interleukin 2 (IL-2) is an important cytokine with various immunological functions. IL-2 is highly expressed by T-cells following activation and can therefore serve as an indirect surrogate of T-cell activation. To measure the activation of spleenocytes from vaccinated mice, we adapted a commercially available IL-2 ELISPOT[PLUS] assay (Mabtech, catalog #3441-4APW-2). In brief, a 96-well plate precoated with a mouse anti-IL2 monoclonal antibody was washed four times with sterile 1X PBS (200µl/well). After washing, the wells of the plate were conditioned with 200µl of RPMI media (Gibco; Montana, USA) supplemented with 10% heat inactivated Fetal Bovine Serum (ΔFBS, Bio Whittaker; Walkersville, Maryland) and incubated for 30 minutes at room temperature. A hemocytometer (Reichert, USA) and 4% Trypan Blue stain (Sigma) was used to quantify the number of freshly thawed spleenocytes and asses overall viability of cells to be added to each well. After the total number of spleenocytes in each sample was enumerated and diluted in RMPI media, the medium from the 96-well plate was removed and fixed concentrations of spleenocytes (2,500-750,000 cells per well) from immunized mice were added in each well with a fixed concentration of antigen (KLH=100µg, KLH-E=2µg, SARS-CoV-2 E peptide=1µg) or infected cell lysate (uninfected MRC5 cell lysate=200µg or HCoV-229E infected MRC5 cell lysate=200µg) for stimulation. A total volume of 200µl of the spleenocyte/antigen/RPMI mixture was added to each well of the covered 96-well plate and stored in a 37ºC humidified incubator with 5% CO_2 and incubated for 48 hours. As additional assay controls, some wells contained only spleenocytes (no exogenous antigen), or spleenocytes treated with a 1:500 dilution of 500x Concanavalin A

(Thermofisher). To detect IL-2 secreting foci after the 48 hour incubation, the plate was removed and washed five times with PBS (200µl/well). The detection antibody (5HA-biotin) was diluted to 1µg/mL in PBS containing 0.5% FBS and 100µl was added to each well and incubated for two hours at room temperature. After the primary antibody incubation, the plate was washed as previously described, and 100µl/well of the streptavidin-ALP secondary antibody (1:1,000 dilution in PBS containing 0.5% FBS) was added and incubated for one hour at room temperature followed by another wash as previously described. Lastly, 100µl of BCIP/NTB-plus substrate was added to each well of the 96-well plate and incubated until distinct foci were visible. After color development, the plate was washed extensively with 200µl/well of distilled water and allowed to dry. To enumerate foci, the spots in each well were counted manually using a dissection microscope. All T-cell activation experiments were performed in duplicate at each concentration and for each antigen tested. All enumeration of foci were confirmed by two independent observers for accuracy.

Serum neutralization of HA-CoV-2 pseudovirions. HEK293T (ACE2/TMRPSS2) cells were obtained from Virongy Biosciences (Manassas, VA, USA) and maintained in DMEM+10%ΔFBS (Bio Whittaker, Walkersville, Maryland). Cells were plated at a density of $2.5x10^4$ cells/well in 100µl volume of DMEM+10%ΔFBS in a sterile 96-well tissue culture treated plate and allowed to adhere for 4 hours prior to infection. In a separate sterile 96-well plate, a 75µl reaction mixture was prepared of diluted test sera (15µl of diluted serum + 25µl pseudovirion + 35µl media), broadly neutralizing rabbit anti-RBD Spike antibody control 27VB1 (Virongy Biosciences) (1µl mAb + 25µl pseudovirion + 49µl media), or virus only (25µl pseudovirion + 50µl media) and incubated for 1 hour at 37ºC. After incubation, the media from the adherent cells was removed and replaced with the neutralization mixture and allowed to infect for 24 hours in a 37ºC incubator. After infection, the liquid in each well was removed with a multichannel pipette and discarded, followed by carefully washing each well with 200µl of sterile 1X PBS to wash away excess assay components. After washing, 20µl per well of 1X Promega Lysis buffer was added to each well and allowed to lyse for 5 minutes at room temperature. After lysis, 100µl of Luciferase assay reagent was added to each well of the 96-well plate and read using a plate reader (Spectramax ® M2, Molecular Devices). Neutralization activity was calculated by comparing the reduction in relative luciferase units (ΔRLU) in wells containing test serum or neutralizing monoclonal antibody to wells containing pseudovirion and cells in the absence of antibody.

Serum neutralization of human coronavirus 229E (HCoV-229E). In a 24-well tissue culture plate (Genesee Scientific), $5x10^4$ HFF cells were plated in 1mL of DMEM media (Gibco) supplemented with 10% FBS (Bio Whittaker, Walkersville, Maryland) and incubated at 37ºC for 3 days or until the HFF cells in each well reached 90% confluency. In a sterile PCR tube, 20µl of 229E viral stock ($1x10^5$ PFU/mL), or 2,000 PFU of virus were incubated with 20µl of test serum or media only as a negative control. The mixture of virus and serum was mixed and centrifuged for 30 seconds and incubated at 37ºC for 1 hour in a PCR thermal cycler (Peltier Thermal Cycler PTC-200, MJ Research). After incubation, the 40µl serum and virus mixture was added to each well of the 24-well tissue culture plate containing 960µl of fresh DMEM media supplemented with 10% FBS and incubated at 37ºC in a CO_2 incubator overnight. After 24 hours of infection, the media in each well was replaced with 1mL of fresh DMEM media and cultured at 37ºC in a CO_2 incubator for 20 days. By 21 days post infection, if most cells in a well were lysed with observable cytopathic effect (CPE) by microscopy, the serum was defined as lacking

169

neutralization activity. In contrast, if cells appeared normal in comparison to uninfected cells, the serum was defined as having neutralizing activity.

Cytokine profiling of HCoV-229E infection in mouse lungs. Groups of inbred BALB/c mice were infected intranasally with 1×10^5 PFU of HCoV-229E, or uninfected, and 24 hours later given an intranasal "mock treatment" of 50µl of PBS. At 7 days post treatment, mice were sacrificed, and lungs were isolated and homogenized using a mortar and pestle and immediately frozen at -80ºC. Tissue lysates were analyzed using a mouse cytokine antibody array panel according to the manufacturer's instructions (R&D Systems, CAT#ARY006) for the parallel determination of relative levels of selected mouse cytokines and chemokines. Tissue lysates were incubated with the cytokine array membrane for 24 hours at 4ºC and developed according to the manufacturer's instructions. Images of the cytokine array panels for each mouse was collected using a Bio-Rad ChemiDoc touch imaging system under standard settings according to the manufacturer's instructions. To quantify signal intensity of each cytokine and chemokine, raw unprocessed images were analyzed using the ImageJ analysis program as previously described [136]. Each pair of duplicate cytokine and chemokine measurement was individually analyzed by measuring the Mean Pixel Intensity (MPI) and selecting the total area under the curve for each histogram. The replicate data for each cytokine measured for each animal was plotted and tested for statistical significance using Graphpad Prism software. For statistical analysis of differentially expressed cytokines we used an unpaired multiple t-test analysis on each replicate cytokine datapoint for each animal in the two groups. For our analysis we assumed Gaussian distribution and that both samples in each row represented populations with the same standard deviation. Differences with a probability of <0.05 (two-tailed) were considered significant.

Statistical methods (ELISA). The proportions of mice responding to the vaccine with serum neutralizing titers ≥1:8 in post-immunization sera were compared by Fishers exact test. For calculation of geometric mean titers, titers below the lower limit of detection were assigned a value that was half of the lower limit (i.e., a titer of <1:8 was assigned a titer of 1:4), and the reciprocal titers were $\log_{10}$ transformed. To determine whether the geometric mean serum antibody titers between the two independent groups of mice were different, we used a two-tailed Students *t* test and the log-transformed reciprocal titers. Differences with a probability of <0.05 (two-tailed) were considered significant.

REFERENCES

1. McIntosh, K., W.B. Becker, and R.M. Chanock, *Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease.* Proc Natl Acad Sci U S A, 1967. **58**(6): p. 2268-73.

2. Witte, K.H., M. Tajima, and B.C. Easterday, *Morphologic characteristics and nucleic acid type of transmissible gastroenteritis virus of pigs.* Arch Gesamte Virusforsch, 1968. **23**(1): p. 53-70.

3. Tyrrell, D.A., et al., *Coronaviridae.* Intervirology, 1975. **5**(1-2): p. 76-82.

4. V'Kovski, P., et al., *Coronavirus biology and replication: implications for SARS-CoV-2.* Nat Rev Microbiol, 2021. **19**(3): p. 155-170.

5. Hidalgo, P., M. Valdes, and R.A. Gonzalez, *Molecular biology of coronaviruses: an overview of virus-host interactions and pathogenesis.* Bol Med Hosp Infant Mex, 2021. **78**(1): p. 41-58.

6. Corman, V.M., et al., *Hosts and Sources of Endemic Human Coronaviruses.* Adv Virus Res, 2018. **100**: p. 163-188.

7. Huynh, J., et al., *Evidence supporting a zoonotic origin of human coronavirus strain NL63.* J Virol, 2012. **86**(23): p. 12816-25.

8. Tao, Y., et al., *Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History.* J Virol, 2017. **91**(5).

9. Corman, V.M., et al., *Link of a ubiquitous human coronavirus to dromedary camels.* Proc Natl Acad Sci U S A, 2016. **113**(35): p. 9864-9.

10. Su, S., et al., *Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses.* Trends Microbiol, 2016. **24**(6): p. 490-502.

11. Forni, D., et al., *Molecular Evolution of Human Coronavirus Genomes.* Trends Microbiol, 2017. **25**(1): p. 35-48.

12. Wang, L.F., et al., *Review of bats and SARS.* Emerg Infect Dis, 2006. **12**(12): p. 1834-40.

13. Zhong, N.S., et al., *Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003.* Lancet, 2003. **362**(9393): p. 1353-8.

14. Guan, Y., et al., *Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China.* Science, 2003. **302**(5643): p. 276-8.

15. Li, W., et al., *Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions.* J Virol, 2006. **80**(9): p. 4211-9.

16. Chinese, S.M.E.C., *Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China.* Science, 2004. **303**(5664): p. 1666-9.

17. Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.* N Engl J Med, 2012. **367**(19): p. 1814-20.

18. Korean Society of Infectious, D., C. Korean Society for Healthcare-associated Infection, and Prevention, *An Unexpected Outbreak of Middle East Respiratory Syndrome Coronavirus Infection in the Republic of Korea, 2015.* Infect Chemother, 2015. **47**(2): p. 120-2.

19. Corman, V.M., et al., *Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat.* J Virol, 2014. **88**(19): p. 11297-303.

20.   Ithete, N.L., et al., *Close relative of human Middle East respiratory syndrome coronavirus in bat, South Africa.* Emerg Infect Dis, 2013. **19**(10): p. 1697-9.

21.   Anthony, S.J., et al., *Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus.* mBio, 2017. **8**(2).

22.   Memish, Z.A., et al., *Human infection with MERS coronavirus after exposure to infected camels, Saudi Arabia, 2013.* Emerg Infect Dis, 2014. **20**(6): p. 1012-5.

23.   Chu, D.K., et al., *MERS coronaviruses in dromedary camels, Egypt.* Emerg Infect Dis, 2014. **20**(6): p. 1049-53.

24.   Lau, S.K., et al., *Polyphyletic origin of MERS coronaviruses and isolation of a novel clade A strain from dromedary camels in the United Arab Emirates.* Emerg Microbes Infect, 2016. **5**(12): p. e128.

25.   Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019.* N Engl J Med, 2020. **382**(8): p. 727-733.

26.   Wang, M., et al., *SARS-CoV infection in a restaurant from palm civet.* Emerg Infect Dis, 2005. **11**(12): p. 1860-5.

27.   Xu, R.H., et al., *Epidemiologic clues to SARS origin in China.* Emerg Infect Dis, 2004. **10**(6): p. 1030-7.

28.   Xu, H.F., et al., *[An epidemiologic investigation on infection with severe acute respiratory syndrome coronavirus in wild animals traders in Guangzhou].* Zhonghua Yu Fang Yi Xue Za Zhi, 2004. **38**(2): p. 81-3.

29.   Wang, H., et al., *Lockdown measures during the COVID-19 pandemic strongly impacted the circulation of respiratory pathogens in Southern China.* Sci Rep, 2022. **12**(1): p. 16926.

30.   Orlowski, E.J.W. and D.J.A. Goldsmith, *Four months into the COVID-19 pandemic, Sweden's prized herd immunity is nowhere in sight.* J R Soc Med, 2020. **113**(8): p. 292-298.

31.   Cucinotta, D. and M. Vanelli, *WHO Declares COVID-19 a Pandemic.* Acta Biomed, 2020. **91**(1): p. 157-160.

32.   Zmasek, C.M., et al., *Genomic evolution of the Coronaviridae family.* Virology, 2022. **570**: p. 123-133.

33.   Wu, F., et al., *Author Correction: A new coronavirus associated with human respiratory disease in China.* Nature, 2020. **580**(7803): p. E7.

34.   Kim, D., et al., *The Architecture of SARS-CoV-2 Transcriptome.* Cell, 2020. **181**(4): p. 914-921 e10.

35.   Yoshimoto, F.K., *The Proteins of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS CoV-2 or n-COV19), the Cause of COVID-19.* Protein J, 2020. **39**(3): p. 198-216.

36.   Schubert, K., et al., *Author Correction: SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation.* Nat Struct Mol Biol, 2020. **27**(11): p. 1094.

37.   Suryawanshi, R.K., et al., *Dysregulation of Cell Signaling by SARS-CoV-2.* Trends Microbiol, 2021. **29**(3): p. 224-237.

38.   Cornillez-Ty, C.T., et al., *Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling.* J Virol, 2009. **83**(19): p. 10314-8.

39.   Scott, B.M., et al., *Predicted coronavirus Nsp5 protease cleavage sites in the human proteome.* BMC Genom Data, 2022. **23**(1): p. 25.

40.   Angelini, M.M., et al., *Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles.* mBio, 2013. **4**(4).

41.   Kirchdoerfer, R.N. and A.B. Ward, *Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors.* Nat Commun, 2019. **10**(1): p. 2342.

42.   Zaffagni, M., et al., *SARS-CoV-2 Nsp14 mediates the effects of viral infection on the host cell transcriptome.* Elife, 2022. **11**.

43.   Russ, A., et al., *Nsp16 shields SARS-CoV-2 from efficient MDA5 sensing and IFIT1-mediated restriction.* EMBO Rep, 2022. **23**(12): p. e55648.

44.   Frazier, M.N., et al., *Characterization of SARS2 Nsp15 nuclease activity reveals it's mad about U.* Nucleic Acids Res, 2021. **49**(17): p. 10136-10149.

45.   Zhang, Z., et al., *Structure of SARS-CoV-2 membrane protein essential for virus assembly.* Nat Commun, 2022. **13**(1): p. 4399.

46.   Wu, W., et al., *The SARS-CoV-2 nucleocapsid protein: its role in the viral life cycle, structure and functions, and use as a potential target in the development of vaccines and diagnostics.* Virol J, 2023. **20**(1): p. 6.

47.   Nieto-Torres, J.L., et al., *Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein.* Virology, 2011. **415**(2): p. 69-82.

48.   Masters, P.S., *The molecular biology of coronaviruses.* Adv Virus Res, 2006. **66**: p. 193-292.

49.   Schoeman, D. and B.C. Fielding, *Coronavirus envelope protein: current knowledge.* Virol J, 2019. **16**(1): p. 69.

50.   Boson, B., et al., *The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles.* J Biol Chem, 2021. **296**: p. 100111.

51.   Ewart, G.D., et al., *The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels.* J Virol, 1996. **70**(10): p. 7108-15.

52.   Griffin, S.D., et al., *The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine.* FEBS Lett, 2003. **535**(1-3): p. 34-8.

53.   Hyser, J.M., et al., *Rotavirus disrupts calcium homeostasis by NSP4 viroporin activity.* mBio, 2010. **1**(5).

54.   Madan, V., N. Redondo, and L. Carrasco, *Cell permeabilization by poliovirus 2B viroporin triggers bystander permeabilization in neighbouring cells through a mechanism involving gap junctions.* Cell Microbiol, 2010. **12**(8): p. 1144-57.

55.   Pinto, L.H., L.J. Holsinger, and R.A. Lamb, *Influenza virus M2 protein has ion channel activity.* Cell, 1992. **69**(3): p. 517-28.

56.   Wilson, L., et al., *SARS coronavirus E protein forms cation-selective ion channels.* Virology, 2004. **330**(1): p. 322-31.

57.   Liao, Y., J.P. Tam, and D.X. Liu, *Viroporin activity of SARS-CoV E protein.* Adv Exp Med Biol, 2006. **581**: p. 199-202.

58.   Castano-Rodriguez, C., et al., *Role of Severe Acute Respiratory Syndrome Coronavirus Viroporins E, 3a, and 8a in Replication and Pathogenesis.* mBio, 2018. **9**(3).

59.   Verdia-Baguena, C., et al., *Transport mechanisms of SARS-CoV-E viroporin in calcium solutions: Lipid-dependent Anomalous Mole Fraction Effect and regulation of pore conductance.* Biochim Biophys Acta Biomembr, 2021. **1863**(6): p. 183590.

60.   Surya, W., et al., *MERS coronavirus envelope protein has a single transmembrane domain that forms pentameric ion channels.* Virus Res, 2015. **201**: p. 61-6.

61.   Cao, Y., et al., *Characterization of the SARS-CoV-2 E Protein: Sequence, Structure, Viroporin, and Inhibitors.* Protein Sci, 2021. **30**(6): p. 1114-1130.

62.   Verdia-Baguena, C., et al., *Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids.* Virology, 2012. **432**(2): p. 485-94.

63.   Xia, B., et al., *SARS-CoV-2 envelope protein causes acute respiratory distress syndrome (ARDS)-like pathological damages and constitutes an antiviral target.* Cell Res, 2021. **31**(8): p. 847-860.

64.   Breitinger, U., et al., *Viroporins: Structure, function, and their role in the life cycle of SARS-CoV-2.* Int J Biochem Cell Biol, 2022. **145**: p. 106185.

65.   Torres, J., et al., *The transmembrane oligomers of coronavirus protein E.* Biophys J, 2005. **88**(2): p. 1283-90.

66.   Torres, J., et al., *Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein.* Protein Sci, 2007. **16**(9): p. 2065-71.

67.   Pervushin, K., et al., *Structure and inhibition of the SARS coronavirus envelope protein ion channel.* PLoS Pathog, 2009. **5**(7): p. e1000511.

68.   Cabrera-Garcia, D., et al., *The envelope protein of SARS-CoV-2 increases intra-Golgi pH and forms a cation channel that is regulated by pH.* J Physiol, 2021. **599**(11): p. 2851-2868.

69.   Sultan, F., K. Ahuja, and R.K. Motiani, *Potential of targeting host cell calcium dynamics to curtail SARS-CoV-2 infection and COVID-19 pathogenesis.* Cell Calcium, 2022. **106**: p. 102637.

70.   Mandala, V.S., et al., *Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers.* Nat Struct Mol Biol, 2020. **27**(12): p. 1202-1208.

71.   DeDiego, M.L., et al., *A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo.* J Virol, 2007. **81**(4): p. 1701-13.

72.   Lamirande, E.W., et al., *A live attenuated severe acute respiratory syndrome coronavirus is immunogenic and efficacious in golden Syrian hamsters.* J Virol, 2008. **82**(15): p. 7721-4.

73.   Nieto-Torres, J.L., et al., *Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis.* PLoS Pathog, 2014. **10**(5): p. e1004077.

74.   Alam, I., et al., *Functional Pangenome Analysis Shows Key Features of E Protein Are Preserved in SARS and SARS-CoV-2.* Front Cell Infect Microbiol, 2020. **10**: p. 405.

75.   Madan, V., A. Castello, and L. Carrasco, *Viroporins from RNA viruses induce caspase-dependent apoptosis.* Cell Microbiol, 2008. **10**(2): p. 437-51.

76.   Nieto-Torres, J.L., et al., *Relevance of Viroporin Ion Channel Activity on Viral Replication and Pathogenesis.* Viruses, 2015. **7**(7): p. 3552-73.

77.   Nieto-Torres, J.L., et al., *Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome.* Virology, 2015. **485**: p. 330-9.

78.   Martinon, F., A. Mayor, and J. Tschopp, *The inflammasomes: guardians of the body.* Annu Rev Immunol, 2009. **27**: p. 229-65.

79.   Meduri, G.U., et al., *Persistent elevation of inflammatory cytokines predicts a poor outcome in ARDS. Plasma IL-1 beta and IL-6 levels are consistent and efficient predictors of outcome over time.* Chest, 1995. **107**(4): p. 1062-73.

80.   Guan, W.J., et al., *Clinical Characteristics of Coronavirus Disease 2019 in China.* N Engl J Med, 2020. **382**(18): p. 1708-1720.

81.   Narasaraju, T., *Histopathologic Changes and SARS-CoV-2 Immunostaining in the Lung of a Patient With COVID-19.* Ann Intern Med, 2020. **173**(4): p. 323-324.

82.   Rahman, M.S., et al., *Mutational insights into the envelope protein of SARS-CoV-2.* Gene Rep, 2021. **22**: p. 100997.

83.   Tamura, K., G. Stecher, and S. Kumar, *MEGA11: Molecular Evolutionary Genetics Analysis Version 11.* Mol Biol Evol, 2021. **38**(7): p. 3022-3027.

84.   Tilocca, B., et al., *Immunoinformatic analysis of the SARS-CoV-2 envelope protein as a strategy to assess cross-protection against COVID-19.* Microbes Infect, 2020. **22**(4-5): p. 182-187.

85.   Swaminathan, A., et al., *Keyhole limpet haemocyanin - a model antigen for human immunotoxicological studies.* Br J Clin Pharmacol, 2014. **78**(5): p. 1135-42.

86.   Wang, Q., et al., *Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants.* Cell, 2023. **186**(2): p. 279-286 e8.

87.   Mukherjee, S., D. Bhattacharyya, and A. Bhunia, *Host-membrane interacting interface of the SARS coronavirus envelope protein: Immense functional potential of C-terminal domain.* Biophys Chem, 2020. **266**: p. 106452.

88.   Hetrick, B., et al., *Development of a hybrid alphavirus-SARS-CoV-2 pseudovirion for rapid quantification of neutralization antibodies and antiviral drugs.* Cell Rep Methods, 2022. **2**(3): p. 100181.

89.   Patel, A., et al., *Molecular basis of SARS-CoV-2 Omicron variant evasion from shared neutralizing antibody response.* Structure, 2023.

90.   He, Q., et al., *An updated atlas of antibody evasion by SARS-CoV-2 Omicron sub-variants including BQ.1.1 and XBB.* Cell Rep Med, 2023. **4**(4): p. 100991.

91.   Chakraborty, C., et al., *The SARS-CoV-2 Omicron recombinant subvariants XBB, XBB.1, and XBB.1.5 are expanding rapidly with unique mutations, antibody evasion, and immune escape properties - an alarming global threat of a surge in COVID-19 cases again?* Int J Surg, 2023. **109**(4): p. 1041-1043.

92.   Rauch, I., M. Muller, and T. Decker, *The regulation of inflammation by interferons and their STATs.* JAKSTAT, 2013. **2**(1): p. e23820.

93.   Tang, B.S., et al., *Comparative host gene transcription by microarray analysis early after infection of the Huh7 cell line by severe acute respiratory syndrome coronavirus and human coronavirus 229E.* J Virol, 2005. **79**(10): p. 6180-93.

94.   Lassnig, C., et al., *Development of a transgenic mouse model susceptible to human coronavirus 229E.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8275-80.

95.   Zhao, X., et al., *Interferon Control of Human Coronavirus Infection and Viral Evasion: Mechanistic Insights and Implications for Antiviral Drug and Vaccine Development.* J Mol Biol, 2022. **434**(6): p. 167438.

96.   Duncan, J.K.S., et al., *Interferon regulatory factor 3 mediates effective antiviral responses to human coronavirus 229E and OC43 infection.* Front Immunol, 2023. **14**: p. 930086.

97.   Mendoza, V.M.M., *Interleukin-17: A potential therapeutic target in COVID-19.* J Infect, 2020. **81**(2): p. e136-e138.

98.   Cao, Z., et al., *Potent and persistent antibody responses against the receptor-binding domain of SARS-CoV spike protein in recovered patients.* Virol J, 2010. **7**: p. 299.

99.   Zhong, X., et al., *B-cell responses in patients who have recovered from severe acute respiratory syndrome target a dominant site in the S2 domain of the surface spike glycoprotein.* J Virol, 2005. **79**(6): p. 3401-8.

100.   Li, J., et al., *Immunogenicity and protection efficacy of monomeric and trimeric recombinant SARS coronavirus spike protein subunit vaccine candidates.* Viral Immunol, 2013. **26**(2): p. 126-32.

101.   Tai, W., et al., *Recombinant Receptor-Binding Domains of Multiple Middle East Respiratory Syndrome Coronaviruses (MERS-CoVs) Induce Cross-Neutralizing Antibodies against Divergent Human and Camel MERS-CoVs and Antibody Escape Mutants.* J Virol, 2017. **91**(1).

102.   Wang, N., et al., *Subunit Vaccines Against Emerging Pathogenic Human Coronaviruses.* Front Microbiol, 2020. **11**: p. 298.

103.   Zhao, J., et al., *Airway Memory CD4(+) T Cells Mediate Protective Immunity against Emerging Respiratory Coronaviruses.* Immunity, 2016. **44**(6): p. 1379-91.

104.   He, Y., et al., *Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus.* J Clin Microbiol, 2005. **43**(8): p. 3718-26.

105.   Szymczak, A., et al., *Antibodies specific to SARS-CoV-2 proteins N, S and E in COVID-19 patients in the normal population and in historical samples.* J Gen Virol, 2021. **102**(11).

106.   Buchholz, U.J., et al., *Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9804-9.

107.   Regla-Nava, J.A., et al., *Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates.* J Virol, 2015. **89**(7): p. 3870-87.

108.   Chen, J., et al., *DNA Vaccines Expressing the Envelope and Membrane Proteins Provide Partial Protection Against SARS-CoV-2 in Mice.* Front Immunol, 2022. **13**: p. 827605.

109.   Jin, H., et al., *Induction of Th1 type response by DNA vaccinations with N, M, and E genes against SARS-CoV in mice.* Biochem Biophys Res Commun, 2005. **328**(4): p. 979-86.

110.   Zhu, M.S., et al., *Induction of SARS-nucleoprotein-specific immune response by use of DNA vaccine.* Immunol Lett, 2004. **92**(3): p. 237-43.

111.   Dutta, N.K., et al., *Search for potential target site of nucleocapsid gene for the design of an epitope-based SARS DNA vaccine.* Immunol Lett, 2008. **118**(1): p. 65-71.

112.   Shi, S.Q., et al., *The expression of membrane protein augments the specific responses induced by SARS-CoV nucleocapsid DNA immunization.* Mol Immunol, 2006. **43**(11): p. 1791-8.

113.   Bhattacharya, S., A. Banerjee, and S. Ray, *Development of new vaccine target against SARS-CoV2 using envelope (E) protein: An evolutionary, molecular modeling and docking based study.* Int J Biol Macromol, 2021. **172**: p. 74-81.

114.   Zheng, M., et al., *TLR2 senses the SARS-CoV-2 envelope protein to produce inflammatory cytokines.* Nat Immunol, 2021. **22**(7): p. 829-838.

115.   Planes, R., et al., *SARS-CoV-2 Envelope (E) Protein Binds and Activates TLR2 Pathway: A Novel Molecular Target for COVID-19 Interventions.* Viruses, 2022. **14**(5).

116.   Su, W., et al., *SARS-CoV-2 envelope protein triggers depression-like behaviors and dysosmia via TLR2-mediated neuroinflammation in mice.* J Neuroinflammation, 2023. **20**(1): p. 110.

117.   Biering, S.B., et al., *SARS-CoV-2 Spike triggers barrier dysfunction and vascular leak via integrins and TGF-beta signaling.* Nat Commun, 2022. **13**(1): p. 7630.

118.   Bi, Z., et al., *Animal models for SARS-CoV-2 infection and pathology.* MedComm (2020), 2021. **2**(4): p. 548-568.

119.   Roberts, A., et al., *Aged BALB/c mice as a model for increased severity of severe acute respiratory syndrome in elderly humans.* J Virol, 2005. **79**(9): p. 5833-8.

120.   Subbarao, K. and A. Roberts, *Is there an ideal animal model for SARS?* Trends Microbiol, 2006. **14**(7): p. 299-303.

121.   Golden, J.W., et al., *Human angiotensin-converting enzyme 2 transgenic mice infected with SARS-CoV-2 develop severe and fatal respiratory disease.* JCI Insight, 2020. **5**(19).

122.   Rathnasinghe, R., et al., *Comparison of transgenic and adenovirus hACE2 mouse models for SARS-CoV-2 infection.* Emerg Microbes Infect, 2020. **9**(1): p. 2433-2445.

123.   Baric, R.S. and A.C. Sims, *Humanized mice develop coronavirus respiratory disease.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8073-4.

124.   Jarlhelt, I., et al., *SARS-CoV-2 Antibodies Mediate Complement and Cellular Driven Inflammation.* Front Immunol, 2021. **12**: p. 767981.

125.   Kiss, L. and L.C. James, *The molecular mechanisms that drive intracellular neutralization by the antibody-receptor and RING E3 ligase TRIM21.* Semin Cell Dev Biol, 2022. **126**: p. 99-107.

126.   Bahnan, W., et al., *Spike-Dependent Opsonization Indicates Both Dose-Dependent Inhibition of Phagocytosis and That Non-Neutralizing Antibodies Can Confer Protection to SARS-CoV-2.* Front Immunol, 2021. **12**: p. 808932.

127.   Piela-Smith, T.H. and J.H. Korn, *Aminopeptidase N: a constitutive cell-surface protein on human dermal fibroblasts.* Cell Immunol, 1995. **162**(1): p. 42-8.

128.   Baker, K., et al., *Immune and non-immune functions of the (not so) neonatal Fc receptor, FcRn.* Semin Immunopathol, 2009. **31**(2): p. 223-36.

129.   Abavisani, M., et al., *Mutations in SARS-CoV-2 structural proteins: a global analysis.* Virol J, 2022. **19**(1): p. 220.

130.   Steel, J., et al., *Influenza virus vaccine based on the conserved hemagglutinin stalk domain.* mBio, 2010. **1**(1).

131.   Krammer, F., et al., *Chimeric hemagglutinin influenza virus vaccine constructs elicit broadly protective stalk-specific antibodies.* J Virol, 2013. **87**(12): p. 6542-50.

132.   Khanna, M., et al., *Protective immunity based on the conserved hemagglutinin stalk domain and its prospects for universal influenza vaccine development.* Biomed Res Int, 2014. **2014**: p. 546274.

133.   Melnychuk, L., et al., *Development of a DNA vaccine expressing a secreted HIV-1 gp41 ectodomain that includes the membrane-proximal external region.* Vaccine, 2017. **35**(20): p. 2736-2744.

134.   Cerutti, N., et al., *Antigp41 membrane proximal external region antibodies and the art of using the membrane for neutralization.* Curr Opin HIV AIDS, 2017. **12**(3): p. 250-256.

135.   Wen, C., et al., *Clinical Study of Human Coronavirus NL63, OC43, 229E, HKU1 Infentions in Hospitalized Children from 2015 to 2020.* Infect Drug Resist, 2022. **15**: p. 1093-1101.

136.   Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
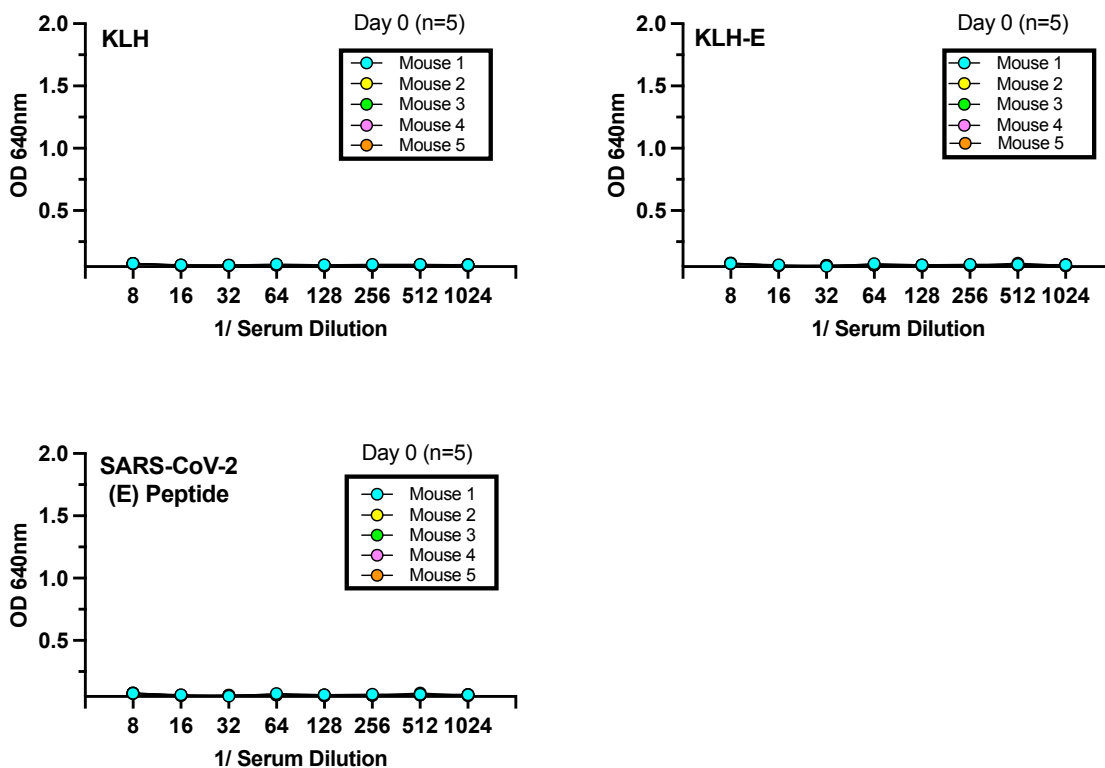
Supplemental Figure 3-1. KLH-E vaccinated mice (day 0) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of serum from individual mice in the Day 0 vaccine group were measured for antibody binding to the antigens used for immunization in the study. Each line represents an individual animal (n=5 total) from the Day 0 KLH-E vaccination group. Each serum sample was tested in replicate in two independent assays.
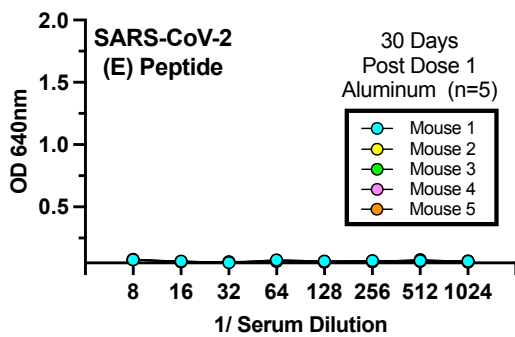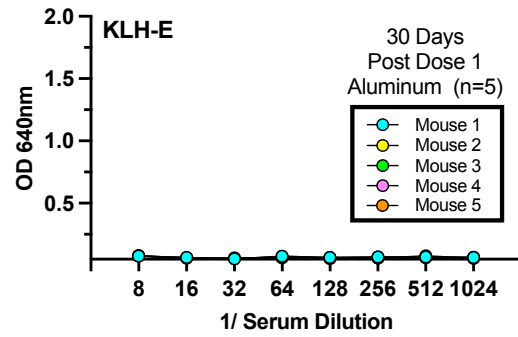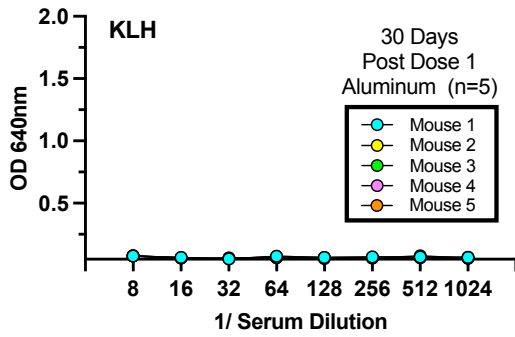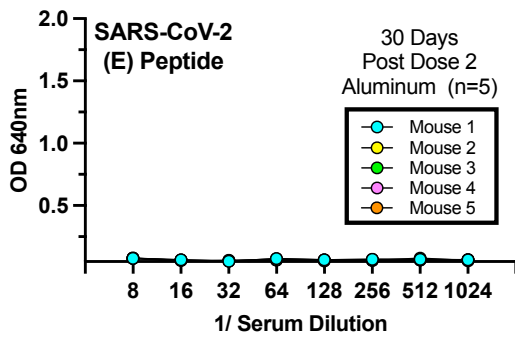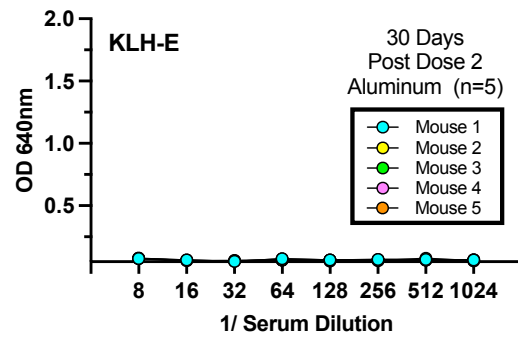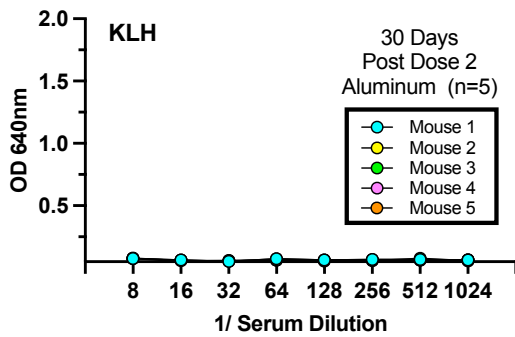
Supplemental Figure 3-2. Aluminum vaccinated mice (30 days post-dose 1) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of serum from individual mice in the 30 days post-dose 1 Aluminum vaccine group were measured for antibody binding to the antigens used for immunization in the study. Each line represents an individual animal (n=5 total) from the 30 days post-dose 1 Aluminum vaccination group. Each serum sample was tested in replicate in two independent assays.
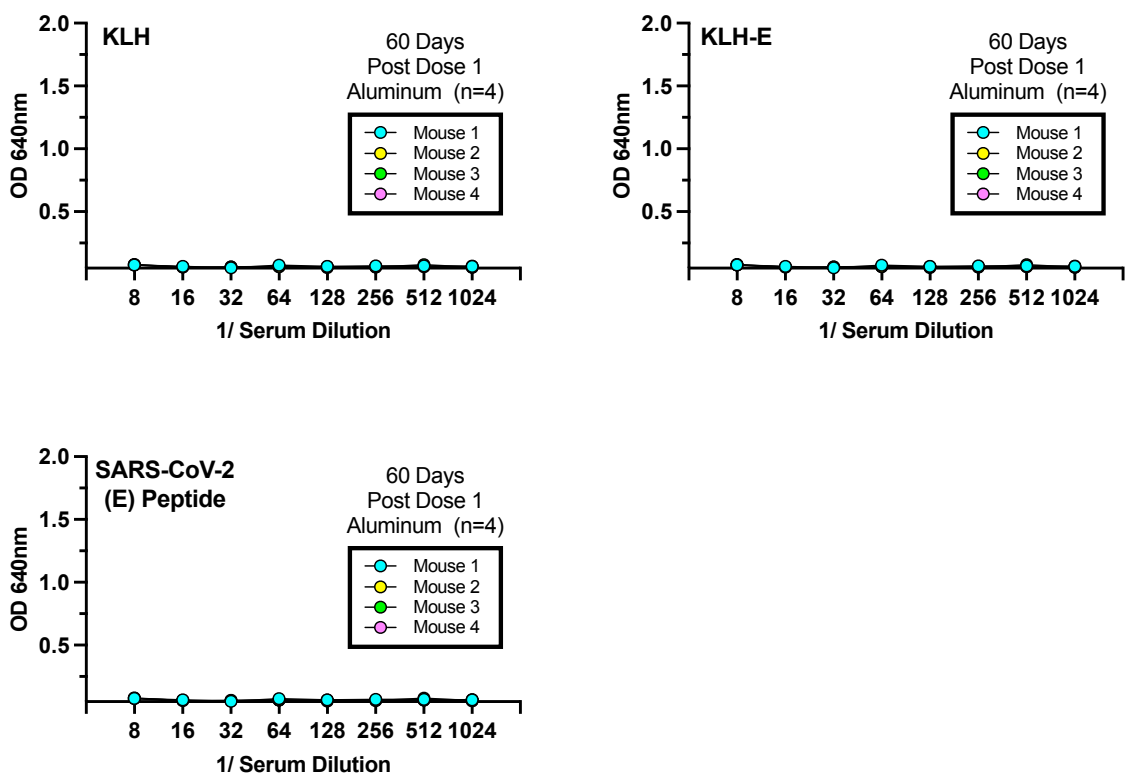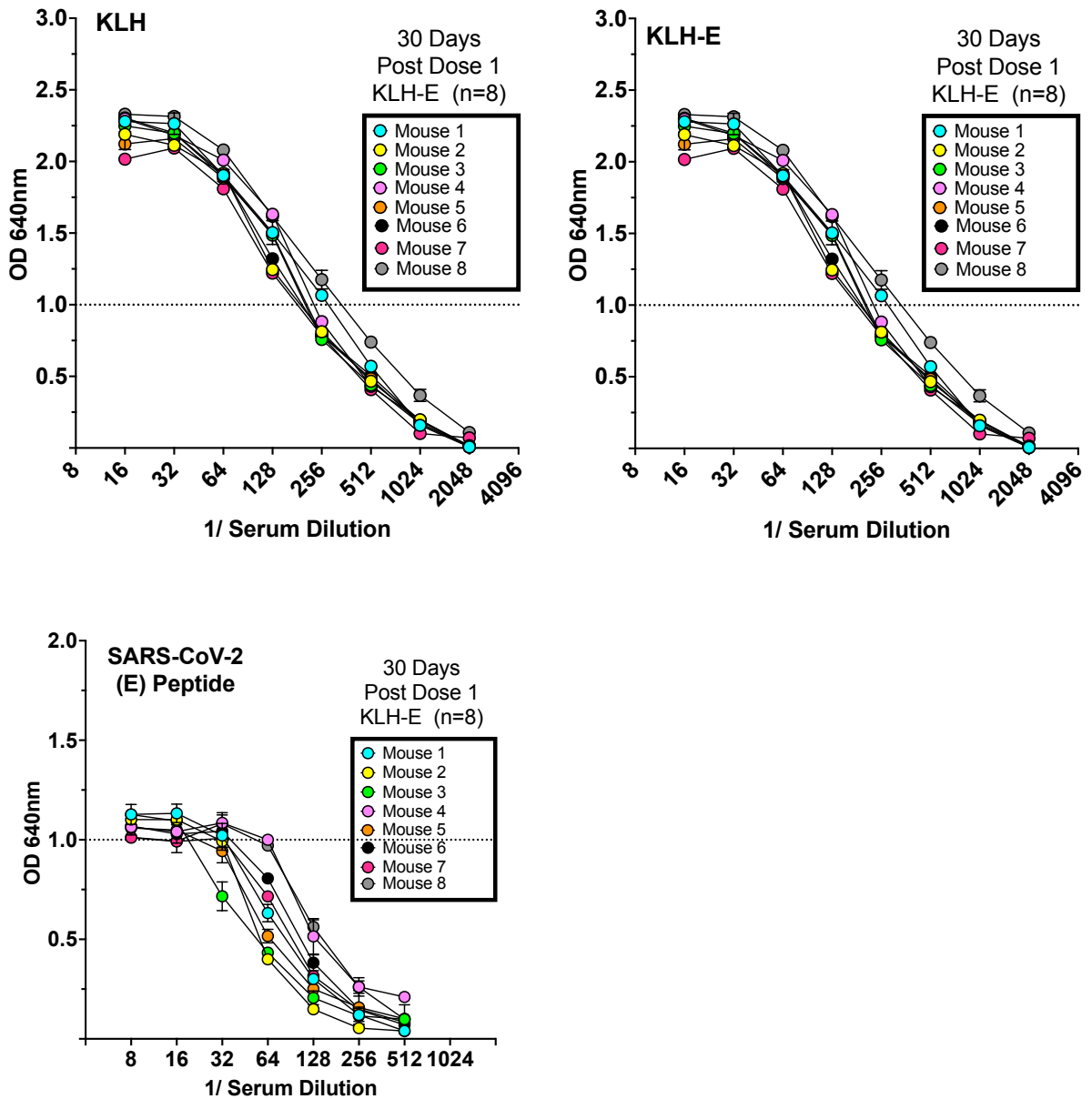
Supplemental Figure 3-3. Aluminum vaccinated mice (60 days post-dose 1) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of
serum from individual mice in the 60 days post-dose 1 Aluminum vaccine group were measured
for antibody binding to the antigens used for immunization in the study. Each line represents an
individual animal (n=4 total) from the 60 days post-dose 1 Aluminum vaccination group. Each
serum sample was tested in replicate in two independent assays. Error bars represent the +/-
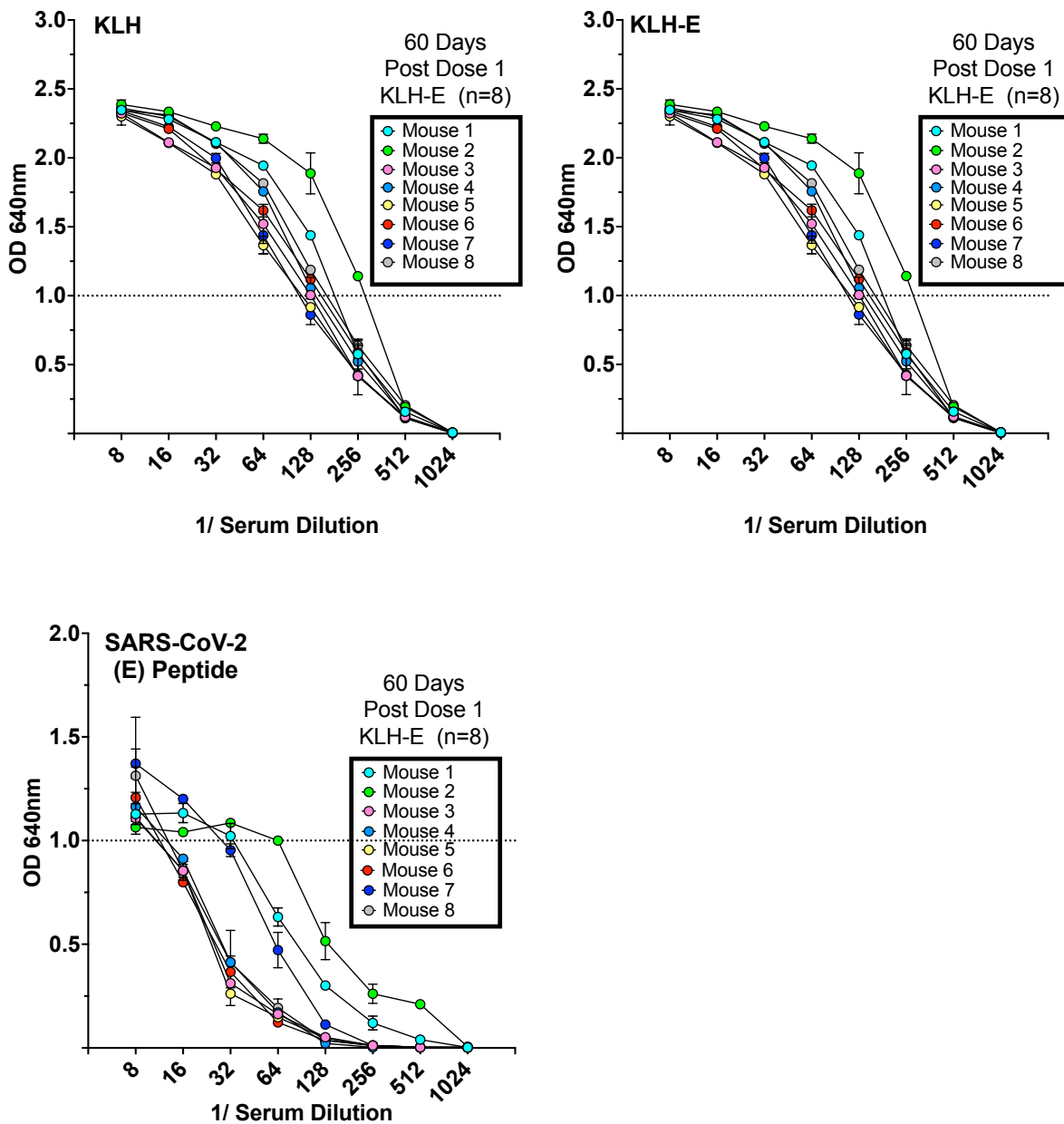standard deviation.

181

Supplemental Figure 3-4. Aluminum vaccinated mice (30 days post-dose 2) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of serum from individual mice in the 30 days post-dose 2 Aluminum vaccine group were measured for antibody binding to the antigens used for immunization in the study. Each line represents an individual animal (n=5 total) from the 30 days post-dose 2 Aluminum vaccination group. Each serum sample was tested in replicate in two independent assays. Error bars represent the +/- standard deviation.
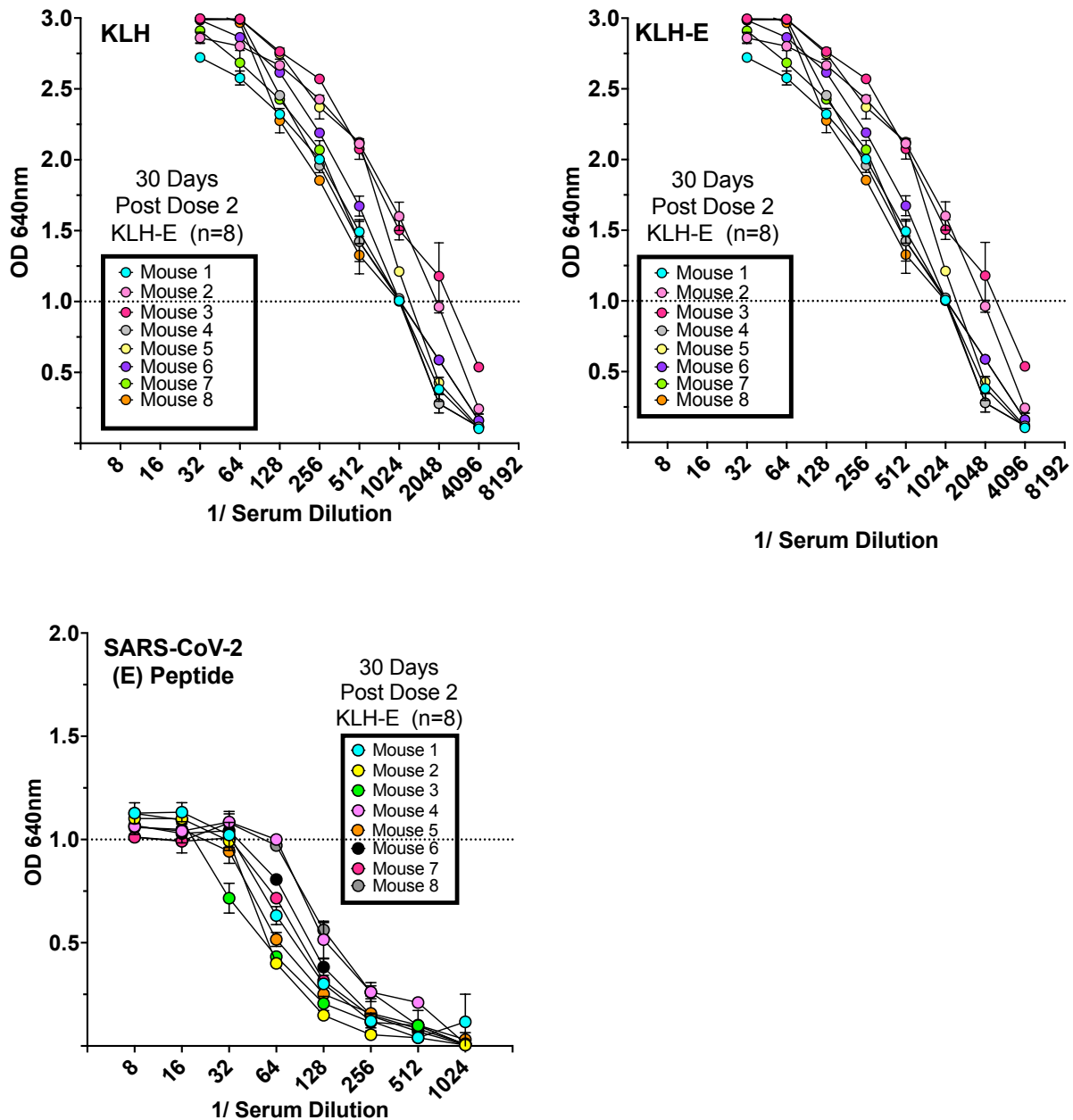
Supplemental Figure 3-5. KLH-E vaccinated mice (30 days post-dose 1) ELISA.
ELISA plates were coated with 2μg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of serum from individual mice in the 30 days post-dose 1 KLH-E vaccine group were measured for antibody binding to the antigens used for immunization in the study. Each line represents an individual animal (n=8 total) from the 30 days post-dose 1 KLH-E vaccination group. Each serum sample was tested in replicate in two independent assays. Error bars represent the +/- standard deviation.

Supplemental Figure 3-6. KLH-E vaccinated mice (60 days post-dose 1) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of serum from individual mice in the 60 days post-dose 1 KLH-E vaccine group were measured for antibody binding to the antigens used for immunization in the study. Each line represents an individual animal (n=8 total) from the 60 days post-dose 1 KLH-E vaccination group. Each serum sample was tested in replicate in two independent assays. Error bars represent the +/- standard deviation.
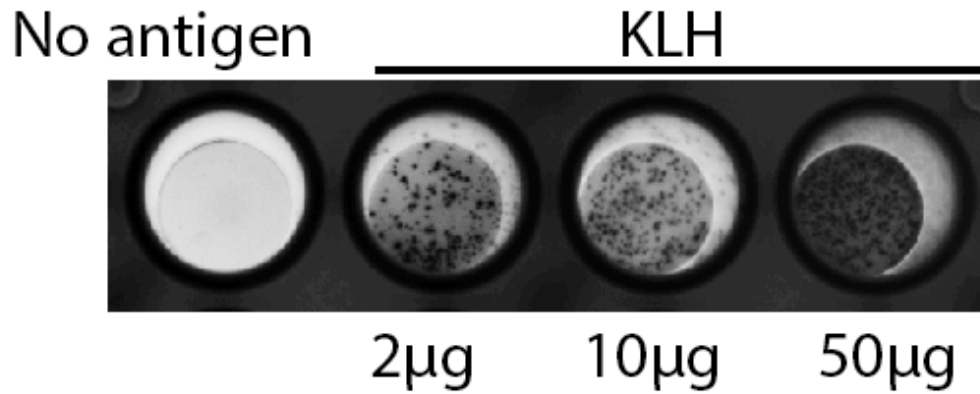
Supplemental Figure 3-7. KLH-E Vaccinated mice (30 days Post-dose 2) ELISA.
ELISA plates were coated with 2µg of KLH, KLH-E, or SARS-CoV-2 E peptide and dilutions of
serum from individual mice in the 30 days post-dose 2 KLH-E vaccine group were measured for
antibody binding to the antigens used for immunization in the study. Each line represents an
individual animal (n=8 total) from the 30 days post-dose 2 KLH-E vaccination group. Each
serum sample was tested in replicate in two independent assays. Error bars represent the +/-
standard deviation.

Supplemental Figure 3-8. Representative image of ELISPOT output data. Figure illustrates representative ELISPOT output data. Each well contains 500,000 spleenocytes from a single mouse immunized with two doses of KLH-E vaccine administered intramuscularly in 30 day intervals. KLH-specific T-cells were measured in spleenic samples by stimulating cells with different test concentrations (2μg, 10μg, 50μg) of purified KLH. The dots in each of the respective well represents an IL-2 secreting foci of activated T-cells.
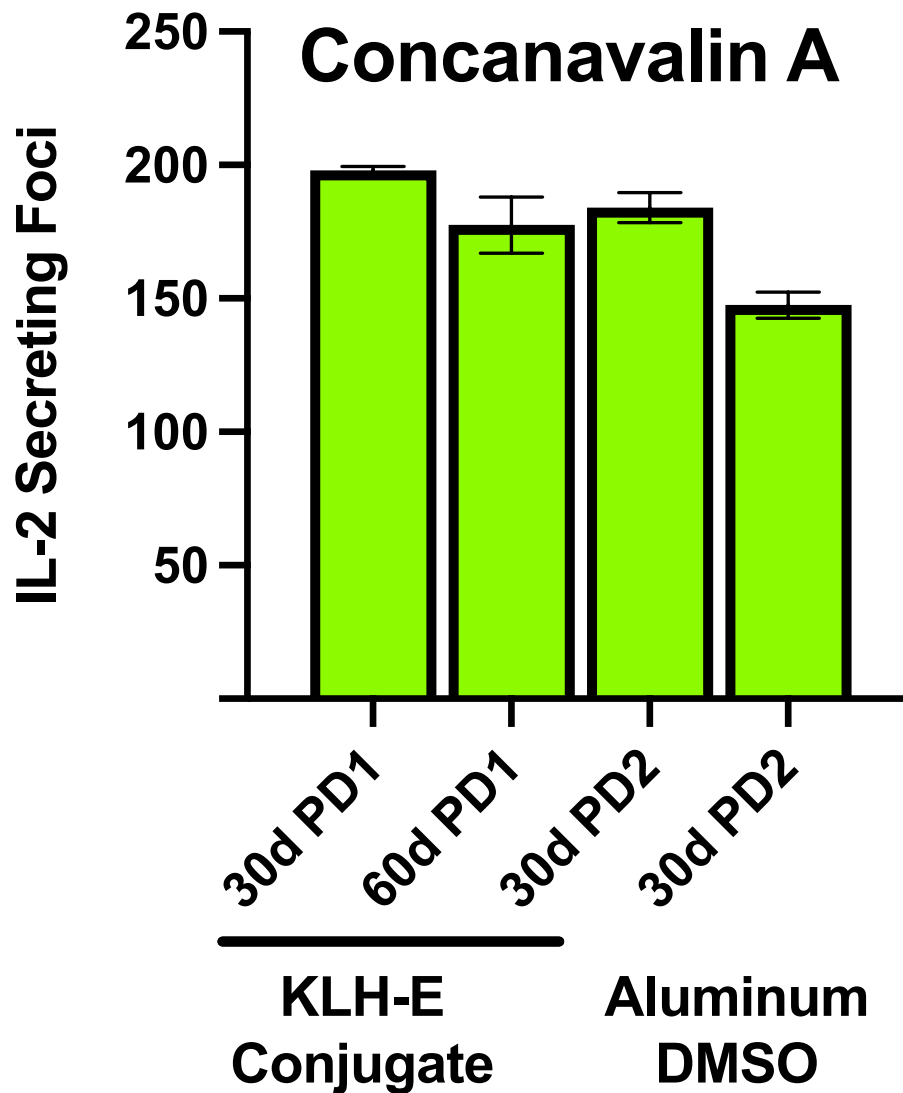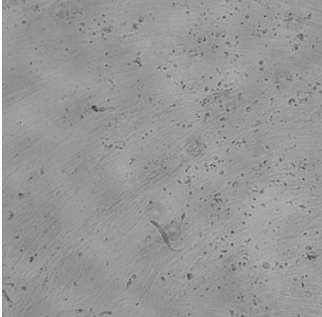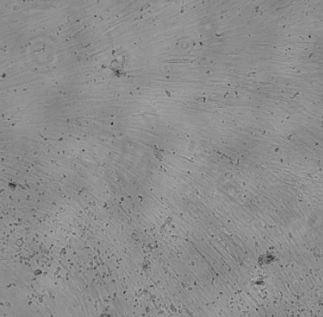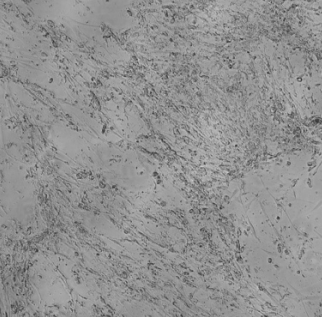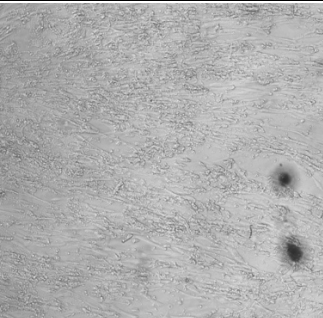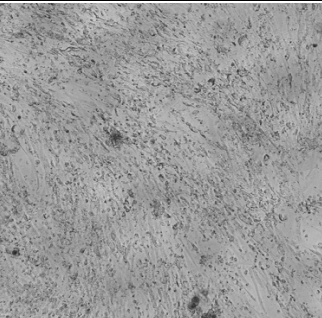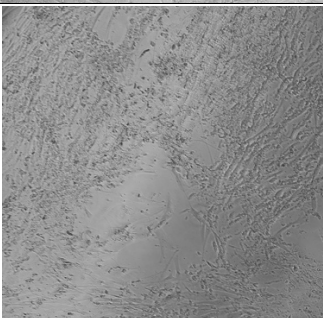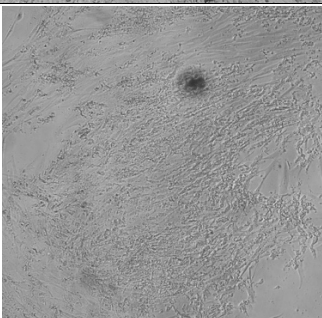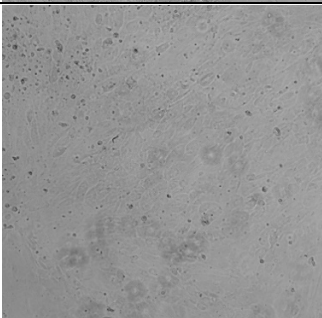
Supplemental Figure 3-9. Mouse T-cells retain functional activity after cryopreservation.
The horizontal axis lists individual representative animals from each vaccine group at select time points, while the vertical corresponds to the enumerated IL-2 secreting foci in the antigen-specific T-cell activation assay. A total of 750,000 spleenocytes were tested from cryopreserved samples and carefully thawed in a 37ºC water bath before enumerating and assessing viability of cells as described in the materials and methods. Cells were stimulated with Concanavalin A in the absence of exogenous specific antigen and measured for T-cell activation. Data represent the mean of two independent replicates and error bars represent the +/- Standard Deviation (SD).
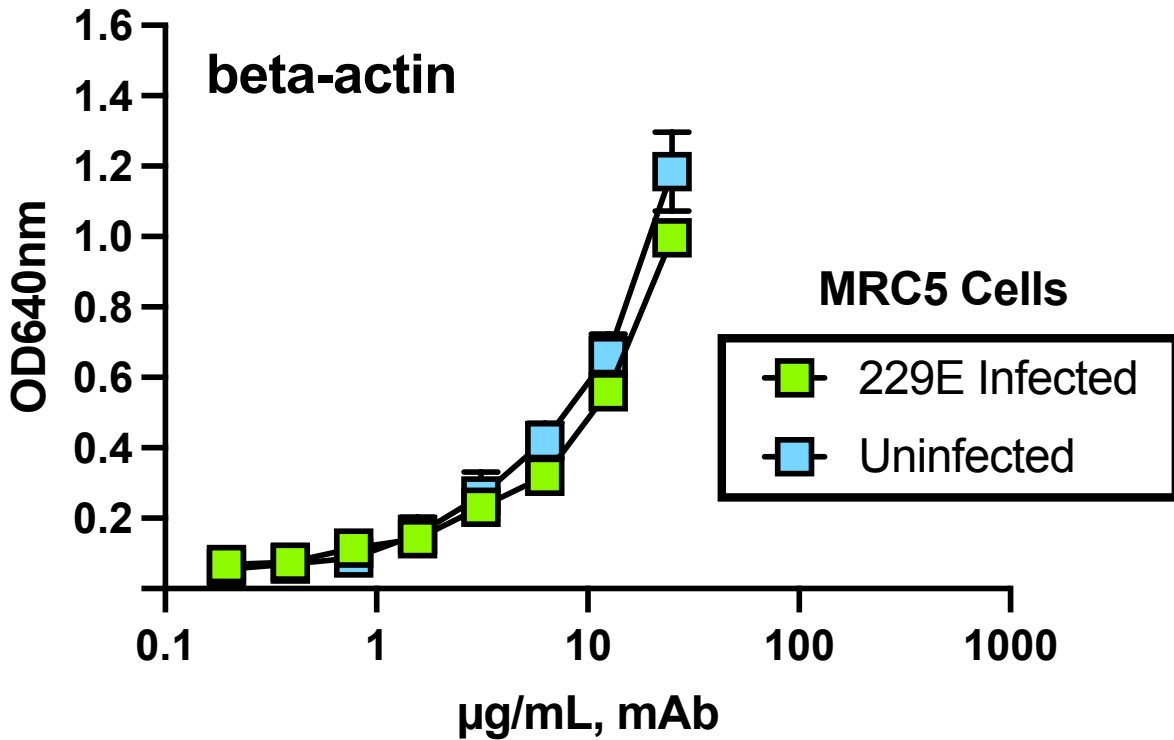
Supplemental Figure 3-10. Comparison of uninfected and HCoV-229E-infected MRC5 cell lysates using the control antigen β-actin.
ELISA plates were coated with 50μg/mL of uninfected or HCoV-229E infected MRC5 cell lysate that were subsequently used in antibody and t-cell assays. To demonstrate that the two MRC5 cell lysate preparations were functionally equivalent with respect to protein concentration and concentrations of an unrelated but shared antigen, each preparation was measured for binding with a mouse β-actin using a monoclonal antibody (Invitrogen; AC-15, CAT#AM4302). The horizontal axis lists the monoclonal antibody (mAb) concentration plotted against the vertical axis which lists the optical density measured by ELISA. Each cell lysate preparation was tested in replicate in two independent assays. Error bars represent the +/- standard deviation.

	Replica 1	Replica 2
Row 1 30 days post-dose 2 KLH-E		
Row 2 30 days post-dose 1 KLH-E		
Row 3 30 days post-dose 2 Aluminum		
Row 4 Virus Only No serum		
Row 5 Cell Only No Virus		

189

Supplemental Figure 3-11. Neutralization activity of pooled KLH-E immunized mouse sera against HCoV-229E. 5×10^4 HFF cells were plated with 1 ml of DMEM with 10% FBS in each well of a 24-well tissue culture plate until 90% confluent. To assess neutralization activity, 20µl of a 1×10^5 PFU/ml stock of HCoV-229E (~2,000 PFU) were incubated with 20µl of freshly thawed pooled mouse sera and incubated at 37ºC for 1 hour. After incubation, the 40µl virus-serum suspension was added to wells in the 24-well tissue culture plate and cultured overnight at 37ºC as described in the methods. The following day, media was replaced with 1mL of fresh DMEM complete medium and continued to culture at 37ºC for 20 days. By day 21 post-infection, cells were visualized to qualify any observed cytopathic effects (CPE).

SUPPLEMENTAL TABLES

Supplemental Table 3-1. Mean Pixel Density (MPD) of mouse lung cytokines (no infection)							
Cytokine/ Chemokine	Mouse 1 (C0241)		Mouse2 (C0242)		Mouse 3 (C0253)		Mouse 4 (C0254)	
BLC	468.778	404.485	452.314	428.314	1754.021	1551.021	797.728	687.728
C5a	2740.849	2542.142	1739.142	1661.728	659.142	605.728	2758.849	2726.849
G-CSF	85.364	117.485	116.899	140.192	452.192	468.899	440.314	460.435
GM-CSF	58.071	65.778	74.778	98.899	195.192	276.899	235.314	285.607
CCL1/TCA	112.778	91.778	118.485	82.071	533.192	351.485	278.899	192.899
CCL11	37.071	49.778	48.778	56.192	120.192	135.778	161.192	305.314
ICAM-1	2979.556	3124.263	2776.435	2614.142	202.071	218.778	3627.556	3640.263
IFN-g	162.071	351.192	206.192	243.899	70.657	55.95	511.899	477.899
IL-1a	485.899	492.899	712.899	923.728	372.778	337.778	902.728	804.728
IL-1b	94.778	126.485	48.364	75.778	247.778	134.657	275.607	242.899
IL-1ra	2316.142	2409.435	2305.728	2369.142	1127.314	1320.142	3191.849	2954.849
IL-2	43.364	52.778	14.828	34.364	112.657	154.071	63.485	43.364
IL-3	84.071	127.485	143.485	89.485	1012.314	684.899	305.607	254.192
IL-4	162.485	180.899	595.021	361.192	951.607	541.192	418.314	550.021
IL-5	58.071	63.778	104.314	114.899	199.485	209.192	241.607	293.314
IL-6	61.778	67.192	84.778	114.192	225.899	386.314	377.314	253.607
IL-7	111.192	149.607	163.899	183.607	557.485	257.364	396.314	417.314
IL-10	33.95	67.485	108.899	87.021	141.485	353.192	346.728	166.192
IL-13	113.485	139.899	99.485	154.485	730.728	1093.435	340.899	298.192
IL-12p70	262.314	175.112	116.607	96.607	95.657	169.071	151.485	130.192
IL-16	2705.971	2892.263	2381.678	2380.849	139.657	149.657	3641.263	3620.556
IL-17	58.071	108.192	18.243	124.192	54.95	263.657	153.485	222.485
IL-23	88.071	125.485	200.485	207.485	539.485	651.899	307.899	354.899
IL-27	100.778	87.071	175.192	196.607	198.364	389.192	239.192	185.192
CXCL10/ CRG-2	464.192	511.899	165.071	121.778	2091.728	2015.435	1000.021	1148.314
CXCL11	105.071	131.778	286.899	41.243	761.607	380.485	337.607	377.021
CXCL1	438.192	268.485	1053.728	940.021	1358.314	1305.607	1256.435	1246.728
M-CSF	636.899	629.607	693.314	557.314	2079.728	2111.142	1911.435	1630.314
CCL2/MCP-1	1718.314	1875.314	2161.728	2438.142	2873.556	2582.556	2592.728	2692.435
CCL12	147.485	152.485	231.899	243.899	360.778	521.485	413.485	101.657
CXCL9	494.314	431.899	481.314	448.021	1260.607	1394.728	1071.971	452.899
CCL3	454.607	478.899	608.314	524.314	1339.607	1590.021	433.899	399.192
CCL4	517.435	407.314	369.607	132.071	1291.142	1217.021	439.314	443.607
CXCL2	120.899	177.607	111.899	121.485	308.192	320.485	252.192	206.192
CCL5	2609.728	2355.021	3277.849	3097.142	1753.142	1831.556	3408.849	3306.435
CXCL12	894.192	1027.607	983.607	1097.021	919.314	1118.314	2733.849	2664.728
CCL17	242.971	187.728	751.577	672.577	2023.698	1618.991	737.749	1233.698
TIMP-1	7331.062	7359.062	9587.083	8729.012	5193.77	4881.477	8307.77	9576.426
TNF-a	1115.456	1036.627	2557.234	2426.941	4764.77	4069.527	2874.527	2627.698
TREM-1	5732.234	5679.527	7692.305	7951.719	5420.426	5654.891	6681.941	6479.527

Supplemental Table 3-2. Mean Pixel Density (MPD) of mouse lung cytokines (HCoV-229E infected)								
Cytokine/ Chemokine	Mouse 1 (C0086)		Mouse 2 (C0087)		Mouse 3 (C0255)		Mouse 4 (C0256)	
BLC	751.556	667.728	523.607	446.899	712.021	864.314	434.314	475.314
C5a	2702.142	2787.556	2124.728	2220.728	2434.849	2572.849	2067.142	2095.728
G-CSF	408.849	276.314	105.778	96.778	237.192	227.192	255.607	220.192
GM-CSF	222.314	195.849	89.485	99.899	187.899	145.192	172.899	175.728
CCL1/TCA	250.314	237.142	129.485	121.192	330.607	367.728	248.485	254.192
CCL11	146.899	91.192	102.192	71.778	223.899	201.899	132.071	131.192
ICAM-1	3366.971	3029.142	2509.728	2818.435	3425.849	3363.849	2612.142	2570.142
IFN-g	414.021	175.192	234.778	281.899	523.192	582.899	312.778	374.485
IL-1a	467.435	471.021	330.485	284.485	620.192	684.192	642.314	835.314
IL-1b	254.314	211.314	149.192	190.607	279.192	328.899	199.485	117.071
IL-1ra	2521.849	2332.435	2401.142	2697.849	2622.142	2551.142	2328.435	2494.435
IL-2	201.728	228.314	67.485	132.192	149.899	159.899	55.364	51.364
IL-3	464.142	369.021	261.021	224.192	388.607	292.899	190.192	243.607
IL-4	482.849	481.435	268.314	164.485	559.728	675.314	811.021	651.021
IL-5	257.607	284.899	128.485	79.778	164.778	152.485	142.192	177.192
IL-6	233.899	209.899	158.021	215.607	159.485	128.071	186.192	174.778
IL-7	301.021	293.021	203.314	129.071	394.899	298.192	282.485	311.192
IL-10	177.899	337.021	162.485	116.192	193.607	219.192	179.778	239.485
IL-13	331.728	303.021	245.314	252.607	413.899	340.899	450.314	409.607
IL-12p70	210.071	344.607	118.364	161.607	265.607	279.021	147.071	197.778
IL-16	2753.435	2680.435	3474.142	2902.435	3792.556	3811.556	2744.849	2441.849
IL-17	214.607	254.607	180.192	252.192	196.192	266.899	113.778	154.071
IL-23	395.021	470.142	312.728	241.485	316.607	339.607	408.192	529.192
IL-27	235.899	242.021	161.192	220.021	210.899	215.899	176.899	172.778
CXCL10/ CRG-2	640.314	683.314	562.899	549.899	1435.021	1232.607	412.899	813.607
CXCL11	321.607	406.021	252.899	258.607	266.192	310.485	328.607	345.899
CXCL1	535.607	736.728	286.485	229.899	467.778	530.485	688.314	580.899
M-CSF	1171.142	841.021	496.899	590.021	740.192	662.485	840.314	807.607
CCL2/MCP-1	2312.263	2281.849	1685.728	1585.021	2605.728	2548.021	1947.607	2000.314
CCL12	448.021	522.728	198.485	250.485	257.071	211.778	468.899	435.192
CXCL9	378.728	418.728	267.899	303.314	607.899	588.192	569.899	505.192
CCL3	883.142	584.435	557.607	594.021	764.899	738.899	546.192	617.607
CCL4	552.435	422.435	374.314	339.314	425.485	434.192	503.899	332.192
CXCL2	345.142	300.021	271.899	227.192	163.485	177.778	146.071	194.485
CCL5	3346.971	3051.142	3085.142	3214.142	3169.142	3095.435	3333.142	3346.142
CXCL12	1056.263	883.142	704.607	756.314	1353.899	1231.192	1777.021	1614.314
CCL17	887.941	813.87	521.971	801.92	655.506	1075.163	581.092	1068.991
TIMP-1	9182.255	8979.719	7582.062	7469.941	8964.184	9574.184	8301.891	8419.477
TNF-a	1842.991	2292.527	1836.527	1595.113	2308.577	2322.991	3349.941	3285.991
TREM-1	5824.598	5461.305	3751.113	3883.941	6737.648	5837.527	8427.477	7707.062

	Uninfected (Mean; n=4 mice)	HCoV-229E infected (Mean; n=4 mice)	Difference	SE of difference	T ratio	P-value
Supplemental Table 3-3. Statistical Analysis of mouse lung cytokine expression levels						
Cytokine/ Chemokine						
BLC	818.0	609.5	208.6	197.8	1.055	0.309504
C5a	1929	2376	-446.4	338.3	1.319	0.208199
G-CSF	285.2	228.5	56.73	73.43	0.7727	0.452562
GM-CSF	161.3	161.2	0.1591	38.21	0.004165	0.996736
CCL1/TCA	220.2	242.4	-22.19	63.81	0.3478	0.733160
CCL11	114.3	137.6	-23.35	36.95	0.6320	0.537558
ICAM-1	2398	2962	-564.1	512.8	1.100	0.289862
IFN-g	260.0	362.4	-102.4	78.72	1.301	0.214188
IL-1a	629.2	541.9	87.25	106.3	0.8205	0.425690
IL-1b	155.8	216.3	-60.47	39.36	1.536	0.146806
IL-1ra	2249	2494	-244.4	256.4	0.9532	0.356665
IL-2	64.86	130.8	-65.92	28.82	2.287	0.038264
IL-3	337.7	304.2	33.48	123.5	0.2712	0.790184
IL-4	470.1	511.8	-41.68	117.4	0.3551	0.727779
IL-5	160.6	173.4	-12.84	38.96	0.3297	0.746513
IL-6	196.4	183.2	13.14	49.21	0.2670	0.793354
IL-7	279.6	276.6	2.948	62.94	0.04684	0.963300
IL-10	163.1	203.2	-40.09	49.08	0.8169	0.427680
IL-13	371.3	343.4	27.90	129.4	0.2156	0.832438
IL-12p70	149.6	215.5	-65.89	33.14	1.988	0.066718
IL-16	2239	3075	-836.2	523.5	1.597	0.132541
IL-17	125.4	204.1	-78.66	35.56	2.212	0.044106
IL-23	309.5	376.6	-67.16	77.70	0.8644	0.401951
IL-27	196.4	204.5	-8.003	34.62	0.2311	0.820549
CXCL10/ CRG-2	939.8	791.3	148.5	302.0	0.4917	0.630570
CXCL11	302.7	311.3	-8.576	82.27	0.1042	0.918462
CXCL1	983.4	507.0	476.4	159.4	2.989	0.009766
M-CSF	1281	768.7	512.5	262.1	1.956	0.070784
CCL2/MCP-1	2367	2121	246.0	196.9	1.250	0.231931
CCL12	271.6	349.1	-77.44	69.79	1.110	0.285904
CXCL9	754.5	455.0	299.5	153.5	1.951	0.071414
CCL3	728.6	660.9	67.76	169.4	0.4000	0.695179
CCL4	602.2	423.0	179.2	150.3	1.192	0.253192
CXCL2	202.4	228.3	-25.89	38.96	0.6646	0.517096
CCL5	2705	3205	-500.2	240.8	2.077	0.056680
CXCL12	1430	1172	257.7	311.3	0.8278	0.421650
CCL17	933.6	800.8	132.8	239.9	0.5536	0.588576
TIMP-1	7621	8559	-938.5	693.4	1.354	0.197342
TNF-a	2684	2354	329.8	509.2	0.6477	0.527674
TREM-1	6412	5954	457.7	679.5	0.6736	0.511529