Solution NMR Structure and Binding Studies of

Murine Hepatitis Coronavirus Envelope Protein

by

Bobby Baravati

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2020 by the
Graduate Supervisory Committee:

Petra Fromme, Co-Chair
Debra Hansen, Co-Chair
Xu Wang
Wade Van Horn

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

Coronaviruses are the causative agents of SARS, MERS and the ongoing COVID-19 pandemic. Coronavirus envelope proteins have received increasing attention as drug targets, due to their multiple functional roles during the infection cycle. The murine coronavirus mouse hepatitis virus strain A59, a hepatic and neuronal tropic coronavirus, is considered a prototype of the betacoronaviruses. The envelope protein of the mouse hepatitis virus (MHV-E) was extensively screened with various membrane mimetics by solution state nuclear magnetic resonance spectroscopy to find a suitable mimetic, which allowed for assignment of ~97% of the backbone atoms in the transmembrane region. Following resonance assignments, the binding site of the ion channel inhibitor hexamethylene amiloride (HMA) was mapped to MHV-E using chemical shift perturbations in both amide and aromatic transverse relaxation optimized spectroscopy (TROSY) spectra, which indicated the inhibitor binding site is located at the N-terminal opening of the channel, in accord with one of the proposed HMA binding sites in the envelope protein from the related SARS (severe acute respiratory syndrome) betacoronavirus. Structure calculation of residues M1-K38 of MHV-E, encompassing the transmembrane region, is currently in progress using dihedral angle restraints obtained from isotropic chemical shifts and distance restraints obtained from manually assigned NOE cross-peaks, with the ultimate aim of generating a model of the MHV-E viroporin bound to the inhibitor HMA. This work outlines the first NMR studies on MHV-E, which have provided a foundation for structure based drug design and probing interactions, and the methods can be extended, with suitable modifications, to other coronavirus envelope proteins.

DEDICATION

To my family and friends

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ix

CHAPTER 1

INTRODUCTION

Virology background

Coronaviruses and their clinical presentations

Coronaviruses (CoVs) are the largest known RNA viruses causing primarily respiratory infections in humans (Peeri et al. 2020). CoVs infect a variety of species. The known human CoVs are OC43 (Vabret et al. 2003), 229E (Hamre and Procknow 1966), HKU1 (Woo et al. 2005), NL63 (van der Hoek et al. 2004), Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) (Drosten et al. 2003), Middle East Respiratory Syndrome (MERS) (Zaki et al. 2012), and Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) (Zhou et al. 2020). SARS-CoV was responsible for the 2003 epidemic with 8096 cases and 774 deaths, and the MERS outbreak in 2006 was reported in 2,229 with 79 deaths (Schoeman and Fielding 2019). The ongoing SARS-CoV-2 pandemic has been reported in 1,226,644 cases with 66,560 deaths as of 4/5/2020 (https://coronavirus.jhu.edu/map.html). In general, CoVs target epithelial cells and consequently display pathologies of the respiratory system. The human coronaviruses 229E, NL63, OC43, and HKU1 usually cause mild to moderate upper-respiratory tract illnesses, displaying symptoms associated with the common cold. Most people have been infected with one or more of these viruses over the course of their lives. Infections can sometimes progress to lower-respiratory tract illnesses, such as pneumonia or bronchitis, in patients with weakened immune systems. SARS-CoV, SARS-CoV-2 and MERS display symptoms of greater severity with greater incidence of pneumonia and respiratory

failure (Peeri et al. 2020). In contrast, MHV causes hepatitis and neurological infections (paralysis, demyelination) with high mortality (Haring and Perlman 2001).

CoV classification

CoVs exhibit a spherical morphology, with observed diameters of 120-160 Å (Fig. 1.1). They are classified according to genetic criteria into four types: α, β, γ and δ. The β type is further subdivided into subgroups A, B, C, and D (2012) as shown in the phylogenetic tree in Fig 2.2. MHV, which is the focus of this dissertation, belongs to the A subgroup of the βCoVs.
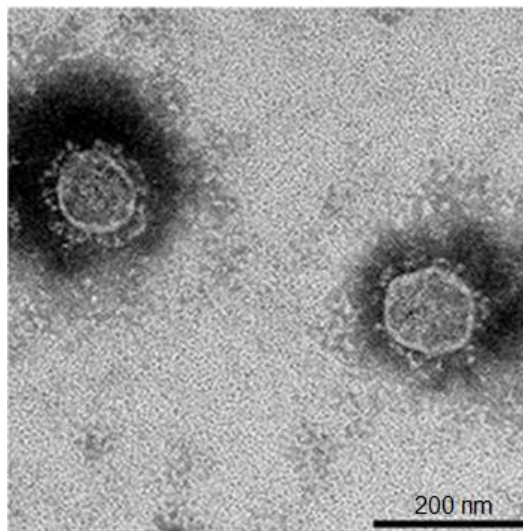
Figure 1.1. Negative-stain electron micrograph of SARS-CoV-2 virions. Reproduced from (Matsuyama et al. 2020).
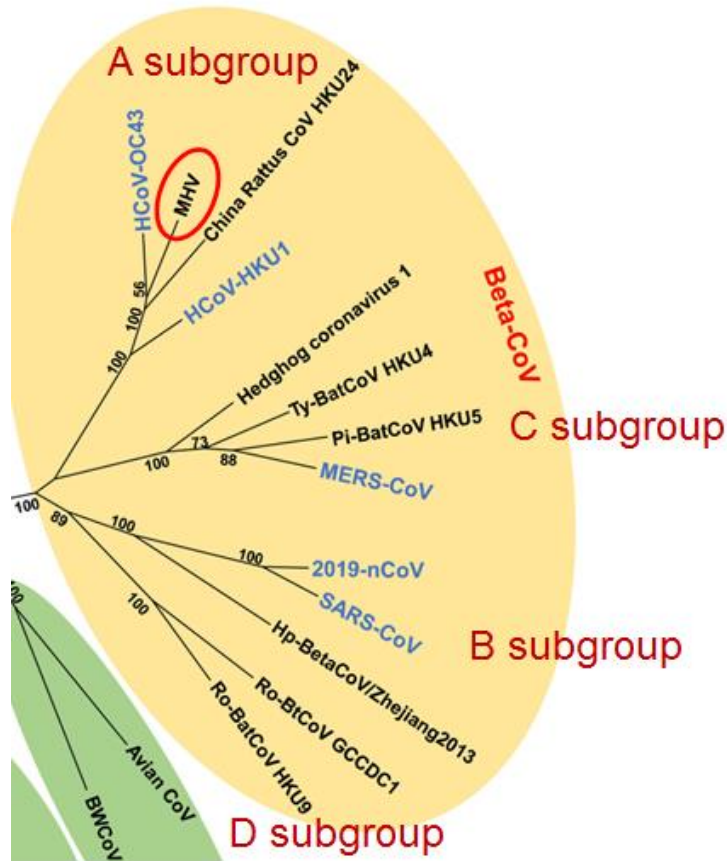
Figure 1.2. Phylogenetic tree of betacoronaviruses (βCoVs). The A, B, C and D subgroups of the βCoVs are indicated by red text. Five of the seven human CoVs belonging to the βCoVs are indicated by blue text. MHV is circled red within the A subgroup. Classification is based on the nucleotide sequences of RNA dependent RNA polymerase (RdRp). Modified from (Wang et al. 2020).

Animal reservoirs

It is thought that the highly pathogenic human CoVs, SARS-CoV, MERS, and SARS-CoV-2, originated from animal reservoirs of the virus, where the viruses evolved to cross the species barrier (Maxmen 2017; Zhou et al. 2020). In the case of SARS-CoV, it is generally thought that the virus originated from bat CoVs; The cross-species transfer of CoVs from bats most likely involves civets as an intermediate host (Li, Shi, et al. 2005). It is speculated that in the case of SARS-CoV-2 pangolins could be an intermediate due

to the high sequence similarity of the S protein receptor binding domains (RBD)

(Andersen et al. 2020). Across the entire genome however, SARS-CoV-2 is closest to the

bat CoV RaTG13 (Zhou et al. 2020).

Murine hepatitis virus (MHV)

A prototype of the β-CoVs, murine hepatitis virus (MHV) was the first CoV for which

the entire genome was sequenced (Lai and Stohlman 1981) and the first CoV for which

recombinant viruses were engineered to generate defined mutants (Fischer et al. 1997),

which have been used in animal models to understand the genetic basis of MHV

pathology (Haring and Perlman 2001). Before the SARS epidemic of 2003, MHV was

the most widely studied coronavirus. MHV is a hepatic and neuronal tropic coronavirus,

in contrast to the human CoVs, which are associated with respiratory illnesses.

Phylogenetic analysis showed that the human coronavirus HKU1 is most closely related

to MHV (Woo et al. 2005).
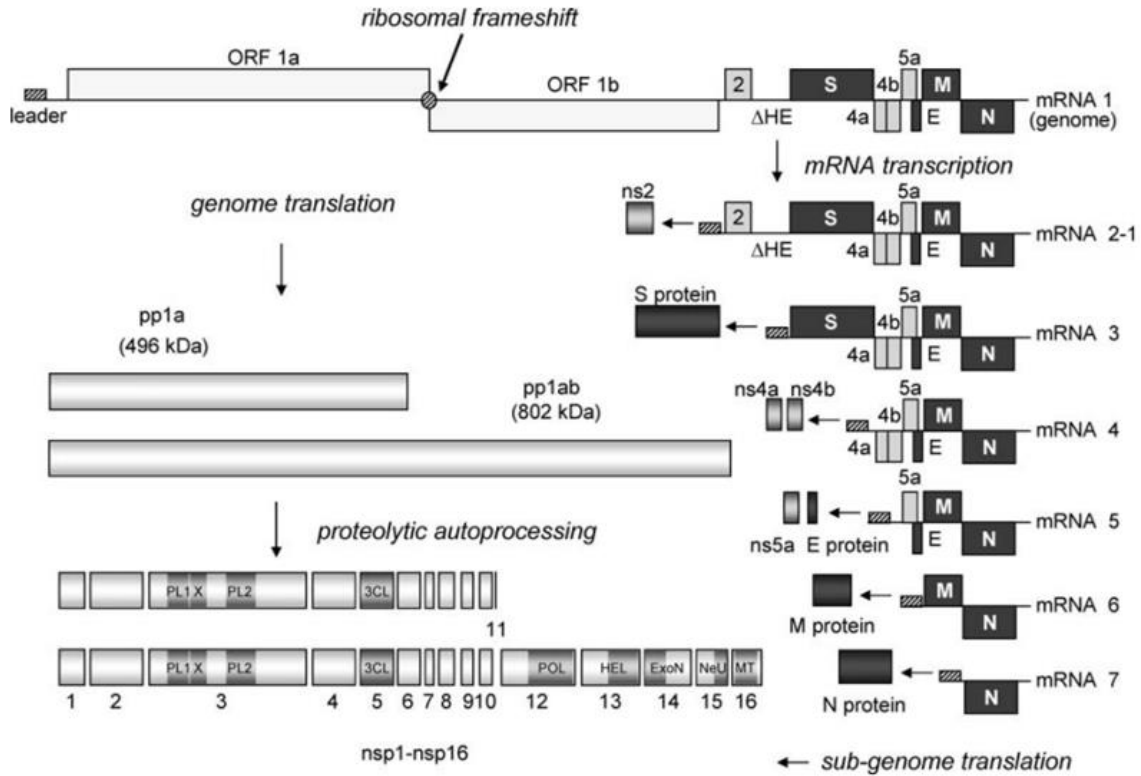
Figure 1.3. Organization of the MHV genome. In the complete MHV genome at the top of the figure, virus open reading frames (ORFs) are lightly shaded, accessory proteins are shaded, and structural proteins are heavily shaded. The genome is translated via a ribosomal frameshift (indicated by shaded circle) yielding the polyproteins pp1a and pp1b, which undergo proteolytic autoprocessing to give the non-structural proteins nsp1-nsp16. The right hand side of the figure shows transcription of a nested set of subgenomic mRNAs, which encode the indicated structural and accessory proteins. Confirmed and putative functional domains in the non-structural proteins nsp1-nsp16 are indicated on the lower right: 3CL, 3C-like cysteine proteinase; ExoN, exonuclease; HEL, superfamily 1 helicase; MT, S-adenosylmethionine-dependent 29-O-methyl transferase; NeU, endoribonuclease; PL1, papain-like protease 1; PL2, papain-like protease 2; POL, RNA-dependent RNA polymerase; X, adenosine diphosphate-ribose 19-phosphatase. Reproduced from (Sawicki et al. 2005).

The organization of the MHV genome is shown in Fig. 1.3. As is characteristic of the

order Nidovirales, gene expression occurs through a nested set of mRNAs that encode the

structural and accessory proteins (Fig. 1.3) (Perlman, Gallagher, and Snijder 2008). The

structural proteins are the nucleocapsid (N), spike (S), matrix (M), envelope (E), and

hemagglutinin-esterase (HE) (depending on the strain of MHV). In addition, there are a

number of accessory proteins, most of which are specific to the CoV species or strain, that are, in general, not required for viral replication in vitro, but play important roles in maintaining efficient replication in the host (Haring and Perlman 2001). The non-structural proteins, nsp1-nsp16, which are produced by proteolytic autoprocessing of the pp1a and pp1ab polyproteins, are primarily involved in catalytic functions and assemble to form the replication and transcription complex (RTC) which drives RNA synthesis (Snijder, Decroly, and Ziebuhr 2016).

The infection cycle for MHV is shown in Figure 1.4. Attachment and entry of the MHV virion into the cell is mediated by interactions between the MHV S trimer and the carcinoembryonic antigen-related cell adhesion molecule 1a (CEACAM1a) host cell receptor. Following either membrane fusion or endocytosis, the genomic positive sense RNA is released and translated to give the large polyproteins pp1a and pp1ab, that undergo proteolytic autoprocessing to give the non-structural proteins nsp1-16, which form the transcription regulation complex that carries out transcription and replication of a nested set of subgenomic mRNAs. These subgenomic RNAs are translated to give the structural and accessory proteins, which generally localize to the endoplasmic reticulum (ER), ER-Golgi intermediate compartment (ERGIC) and Golgi membranes where they are involved in assembly of nascent virions. Newly formed virions are released through exocytosis which then repeat the cycle (Bergmann, Lane, and Stohlman 2006).
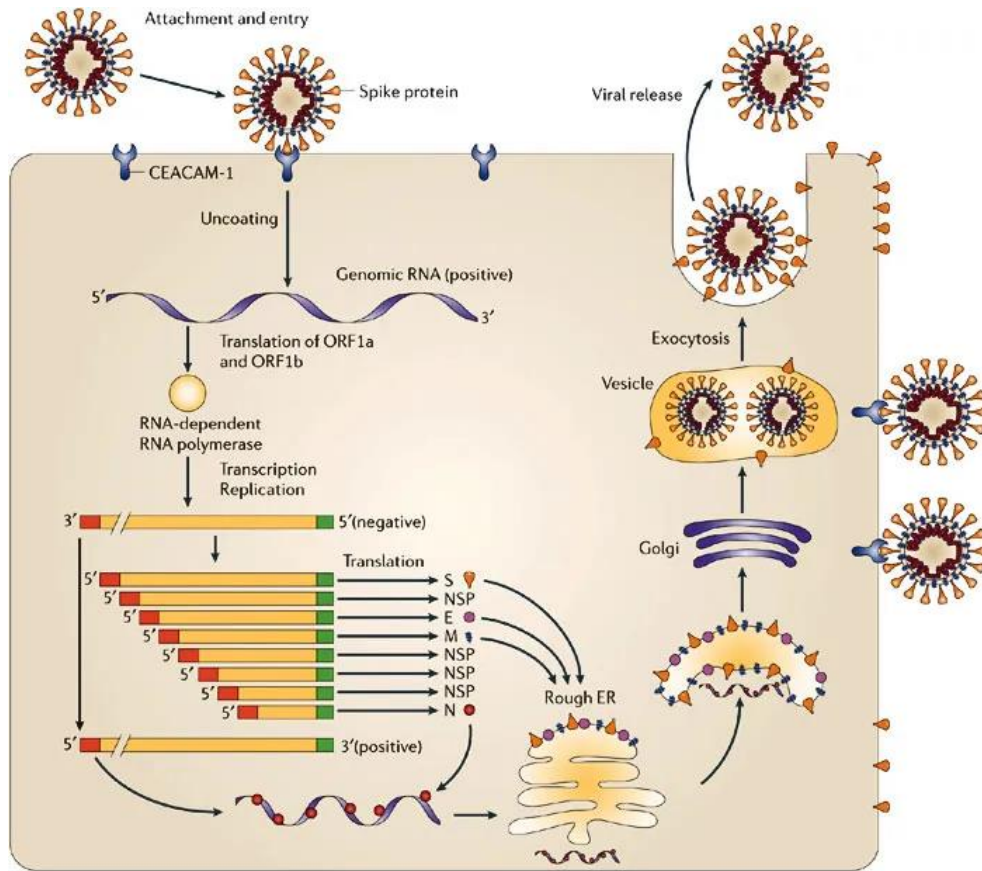
Figure 1.4. MHV infection cycle. MHV binds the host cell receptor CEACAM1a through interactions with the trimeric S glycoprotein, followed by either exocytosis or membrane fusion to release the positive sense viral RNA which is translated to give the large polyproteins pp1a and pp1ab that subsequently undergo proteolytic autoprocessing to the non-structural proteins nsp1-nsp16 which together form the RNA-dependent RNA polymerase. A full length anti-sense RNA is generated which gives rise to a nested set of subgenomic RNAs that are translated to give the structural proteins (N, S, M, E) and the accessory proteins. Assembly of nascent virions occurs in the ERGIC/Golgi membranes which are subsequently released through exocytosis to repeat the infection cycle. Figure reproduced from (Bergmann, Lane, and Stohlman 2006).

Organization of the CoV virion and overview of the CoV structural proteins

The organization of the coronavirus virion is shown in Figure 1.5. The nucleocapsid (N)

protein is an RNA-binding phosphoprotein involved in genome encapsidation and RNA

synthesis and translation. The ribonucleoprotein complex (RNP) is surrounded by a lipid

envelope containing the S, M, E structural proteins (and HE, depending on strain). Lipids

7

found in the envelope originate at the site of virion budding at the endoplasmic reticulum-
Golgi intermediate compartment (ERGIC) and Golgi membranes (Lopez et al. 2006).
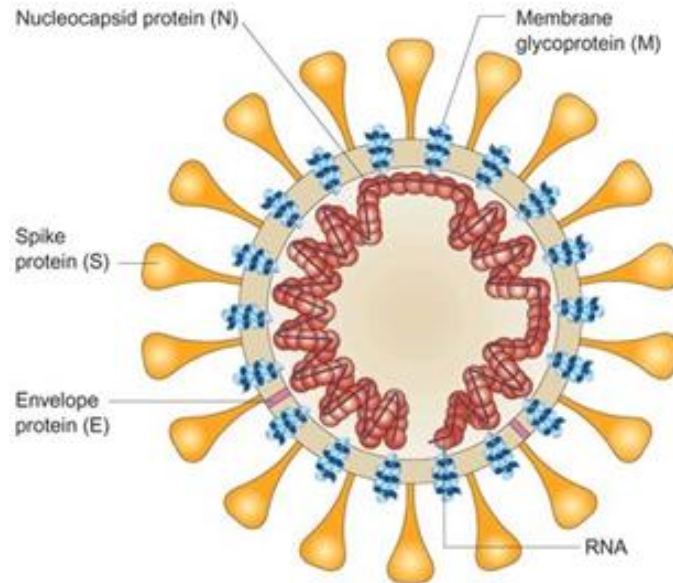
Figure 1.5. Structural organization of the coronavirus virion. The nucleocapsid (N)
protein binds viral RNA to form a helical capsid, called the ribonucleoprotein complex
(RNP), which is enclosed by a lipid envelope containing the spike (S), membrane (M)
and envelope (E) structural proteins. Figure reproduced from (Peiris, Guan, and Yuen
2004).

The spike (S) glycoprotein, which is observed as ~20 nm projections from the lipid

envelope in Fig. 1.1, is a homotrimeric type I membrane protein that mediates viral entry

via the host cell receptor. The MHV S protein targets the CEACAM1a receptor (Peng et

al. 2011), and the 4 Å resolution cryo-EM structure of the MHV S protein (Fig. 1.6)

indicated that the distal loops of domain A interact with the CEACAM1a receptor (Walls

et al. 2016). The S protein of SARS-CoV (Li, Li, et al. 2005) and SARS-CoV-2 target the

angiotensin-converting enzyme 2 (ACE2) receptor to gain entry into the cell (Hoffmann

et al. 2020); however, there are large structural differences in their receptor binding

domains (RBDs) (Fig. 1.7). The SARS CoV-2 S has ~10- to 20-fold higher ACE-2

binding affinity than SARS CoV S, which is thought to account for the high tramsmission

rate of SARS CoV-2 (Wrapp et al. 2020). The S protein is the main inducer of

neutralizing antibodies which target the highly variable residues in the ectodomain N-

terminal domain. Depending on the MHV strain, there may be an additional protein,

hemagglutinin esterase (HE) in the virion envelope, which projects ~5 nm from the

surface of the lipid envelope and is also involved in host cell receptor interactions.
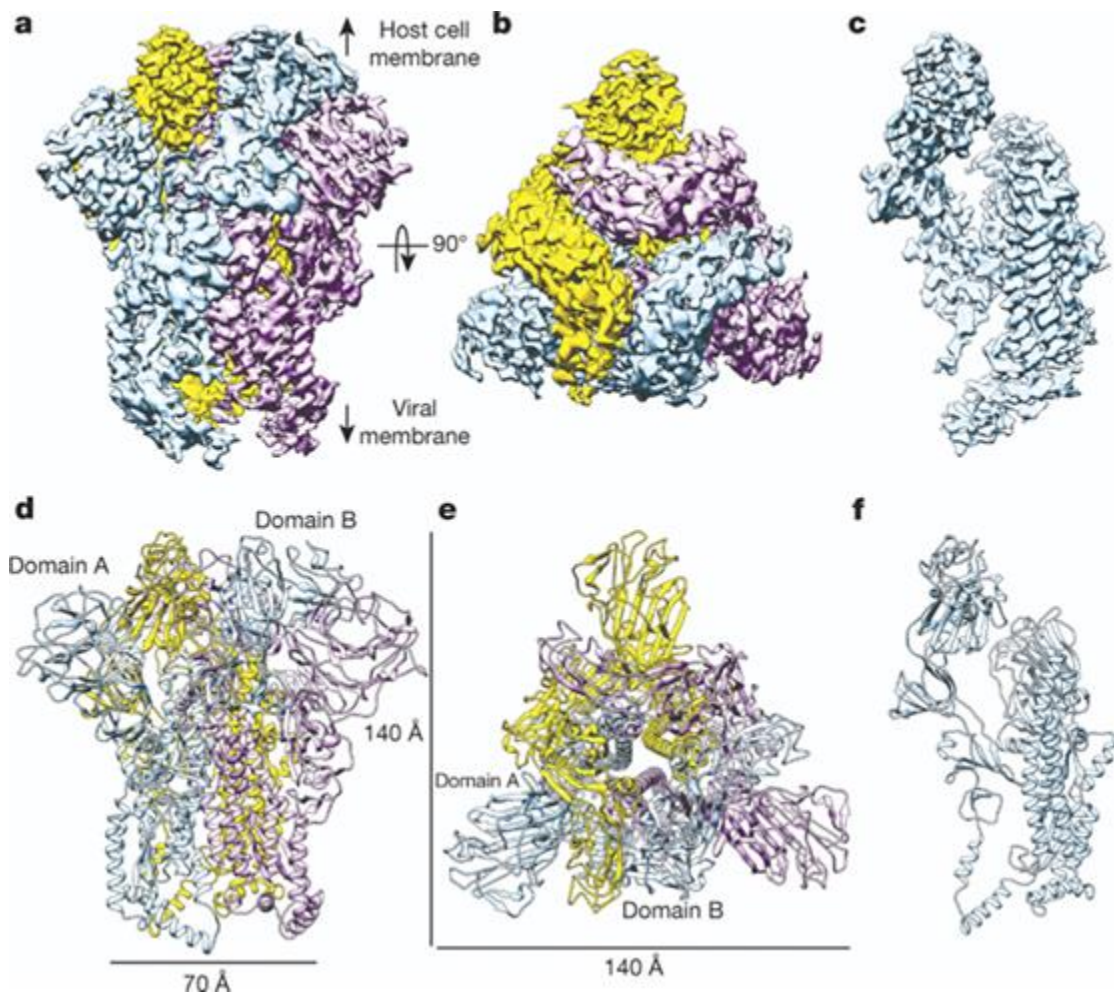
Figure 1.6. Cryo-EM structure of MHV-S. **a-c.** 4.0 Å cryo-EM map of the MHV S trimer colored by individual monomeric subunits. **(a)** Side view. **(b)** Top view looking towards the viral envelope and **(c)** side view of MHV S monomer. **d-f,** ribbon diagrams of **a-c.** Reproduced from (Walls et al. 2016).
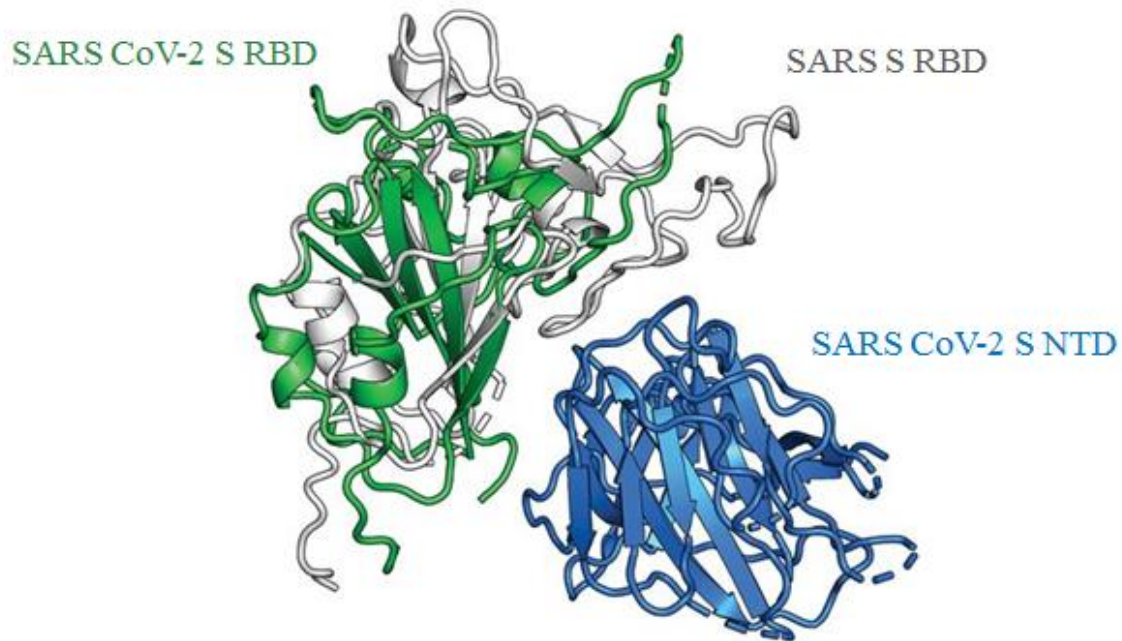
Figure 1.7. Overlay of SARS and SARS CoV-2 S protein ACE2 receptor binding domains. The receptor binding domains (RBD) of SARS-CoV and SARS-CoV-2 are aligned based on the position of the adjacent N-terminal domain (NTD), which shows large structural differences in the RBDs. Reproduced from (Wrapp et al. 2020).

The membrane (M) glycoprotein is a type III integral membrane protein having a predicted triple spanning $N^{exo} C^{endo}$ topology that forms an extended lattice with the inner leaflet of the viral membrane, which is thought to account for the thickness of the CoV envelope. The M protein interacts with the small envelope (E) protein in the formation of the viral envelope and virus like particles (VLPs). The interaction of M and E is thought to occur at their C-terminal domains (Ruch and Machamer 2012; Corse and Machamer 2003).

Coronavirus envelope (CoV E) proteins

Common structural features of CoV E proteins

The E protein is perhaps the most multi-faceted of the three structural proteins that are embedded in the viral envelope in terms of the varied functional roles that have been proposed (Schoeman and Fielding 2019). Despite large variations in sequence identity between coronavirus E proteins (Fig. 1.8), there are conserved structural features: the size and location of the transmembrane domains, the conserved proline at the center of a proposed β-turn (Li, Surya, et al. 2014), a cysteine-rich membrane proximal region subject to palmitoylation (Lopez et al. 2008), alignment of polar residues within the lumen of the channel (Ye and Hogue 2007) and a putative C-terminal PDZ binding motif (Teoh et al. 2010).
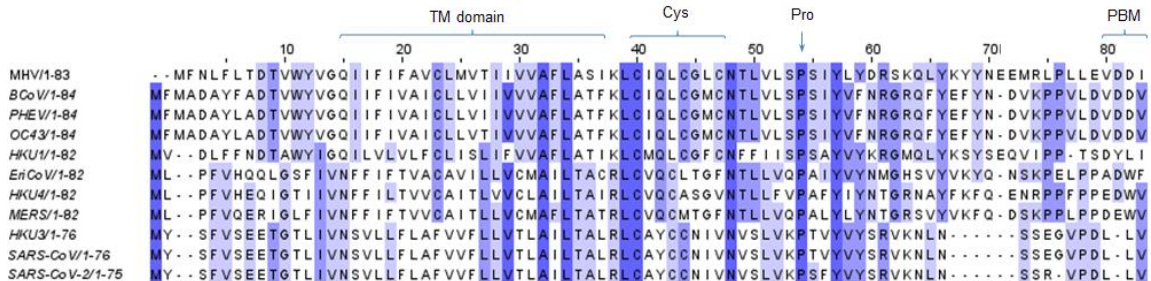
Fig. 1.8. Multiple sequence alignment of CoV E proteins. Residues are numbered in relation to MHV and shaded by % sequence identity. The transmembrane domain is denoted by TM domain. Cys denotes the cysteine-rich membrane proximal region, Pro denotes the conserved proline of the proposed β-turn, and PBM denotes the PDZ binding motif. Multiple sequence alignment (MSA) was done using T-Coffee (Notredame, Higgins, and Heringa 2000) and viewed with JalView (Clamp et al. 2004).

Proposed functions of CoV E proteins: E induces curvature in the viral envelope

In the context of the viral envelope, the E protein is found in low copy numbers and is involved in interactions with the M protein at their large C-terminal extra-membrane domains (Corse and Machamer 2003), where the E protein is thought to induce curvature in the viral envelope (Fischer et al. 1998). Mutations at the C-terminal extramembrane domain of MHV-E gave aberrations in morphology where some exhibited an ovoid rather than spherical morphology (Fischer et al. 1998). Interaction between M and E is sufficient for the formation of virus-like particles (VLPs) (Vennema et al. 1996).

Proposed functions of CoV E proteins: CoV E proteins are critical in the assembly of virions

Recombinant CoVs lacking E show attenuated replication. For murine hepatitis virus lacking E protein (MHV ΔE) replication is 1000 times lower than the wild type MHV (Fig 1.9A) (Wilson, Gage, and Ewart 2006b), whereas for SARS ΔE, virus replication is only 20- to 200-fold lower than the wild type (Fig. 1.9B) (DeDiego et al. 2007). The differences in the level of attenuation seem to suggest varying requirements of the E protein in viral replication.
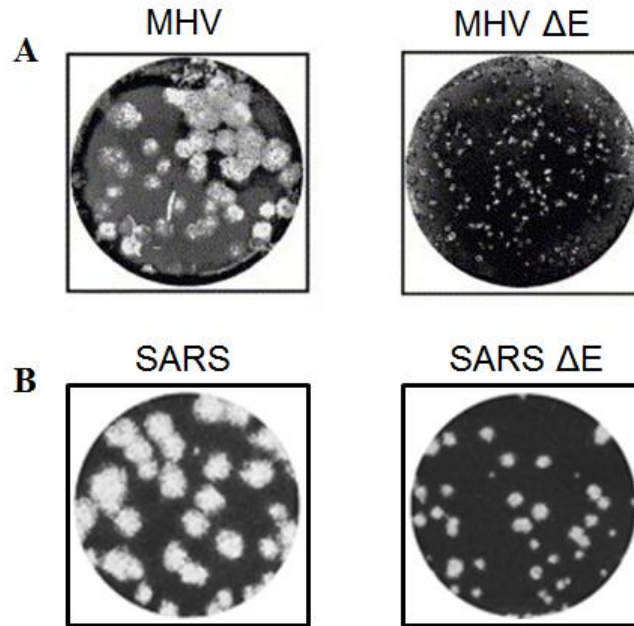
Figure 1.9. CoV E proteins are critical for viral replication. **(A)** MHV ΔE replication is 1000 times lower than the wild type MHV. Reproduced from (Wilson, Gage, and Ewart 2006b). **(B)** SARS ΔE replication is only 20- to 200-fold lower than that of the wild type virus. Reproduced from (DeDiego et al. 2007).

CoV E viroporins

Several of the CoV E proteins (Wilson et al. 2004; Wilson, Gage, and Ewart 2006b) assemble into pentameric channels, called viroporins, which are thought to be implicated in viral budding and scission (Nieva, Madan, and Carrasco 2012; OuYang and Chou 2014). Black lipid membrane (BLM) electrophysiology measurements and corresponding current-voltage (IV) plots for MHV E, IBV E, HCoV-229 E and SARS-CoV E viroporins are shown in Figure 1.10. The relative ion permeabilities (Table 1.1) indicate that MHV E forms a cation-selective channel that is highly selective for Na^+, with the viroporin being ~69 times more permeable to Na^+ than K^+. The discrimination between mono-valent cations is lower for the SARS-CoV E viroporin. HCoV-229 E is slightly more selective for K^+ than Na^+. The nonlinearity in the current-voltage plot for IBV E indicates

13

rectification of the viroporin. The ion conductance and selectivity is influenced in large part by the lipid composition and charge of the membrane (Verdia-Baguena et al. 2012). This study additionally examined the effect of single-point mutations on the ion channel activity of SARS-CoV E, which found that N15A and V25F mutants abolished channel activity, whereas the T11A and T35A mutants exhibited higher channel currents than that of the wild type (Fig. 1.11). CoV E proteins are also thought to indirectly modify host cell ion channels (Torres et al. 2015).
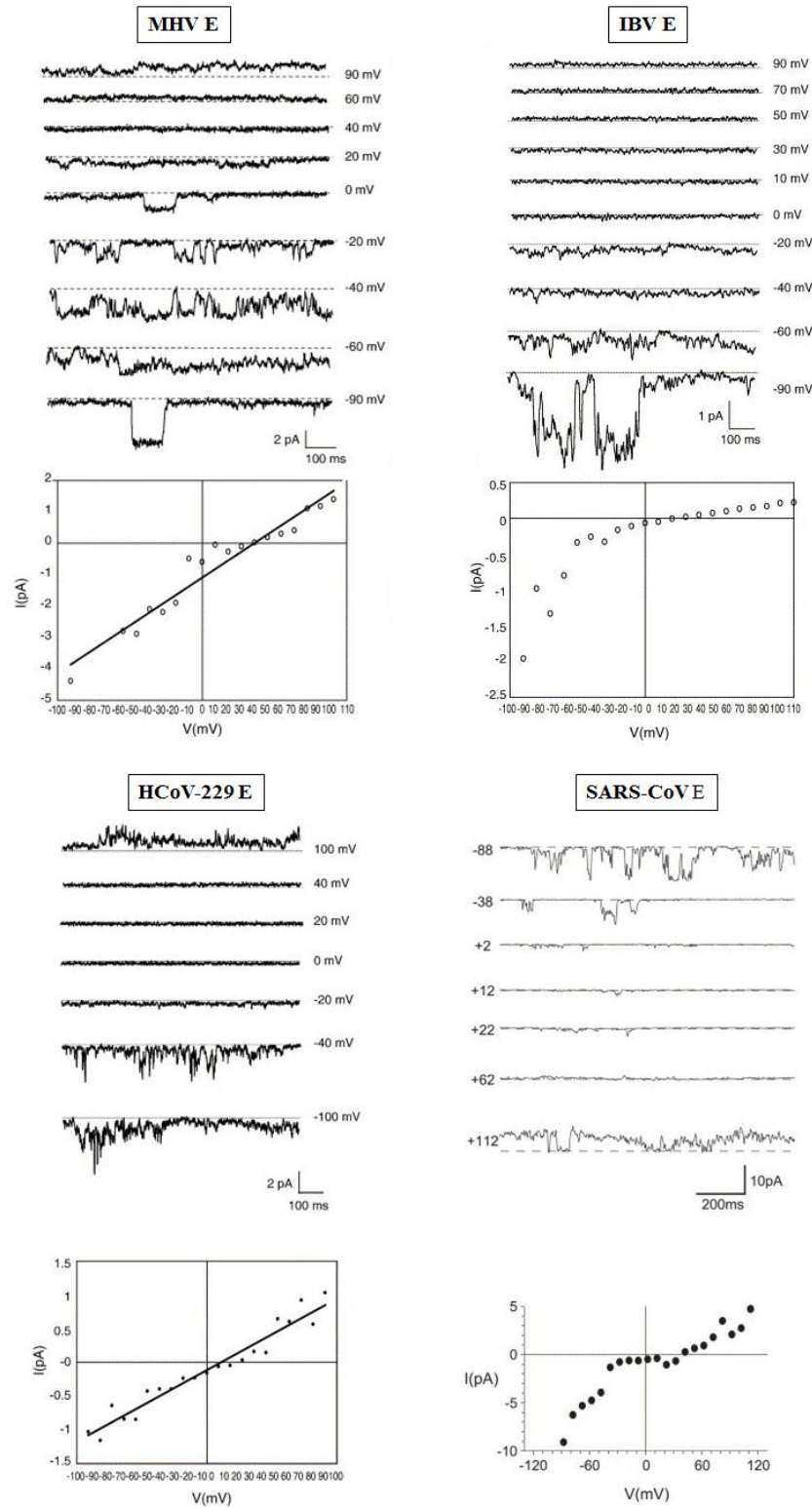
Figure 1.10. Viroporin activity of CoV E proteins. Figure modified from (Wilson, Gage, and Ewart 2006a; Wilson et al. 2004).

15

Table 1.1.

Relative ion permeabilities of coronavirus E protein ion channels

Source of E protein: parent virus	Taxonomic group	Relative ion permeability		
		$P(Na^+/Cl^-)$	$P(K^+/Cl^-)$	Selectivity series
HCoV-229E	1	3	12	$K^+ > Na^+ > Cl^-$
MHV	2a	138	2	$Na^+ \gg K^+ > Cl^-$
SARS-CoV[i]	2b	90	9	$Na^+ \gg K^+ > Cl^-$
IBV	3	10	3	$Na^+ > K^+ > Cl^-$

Reproduced from (Wilson, Gage, and Ewart 2006a).
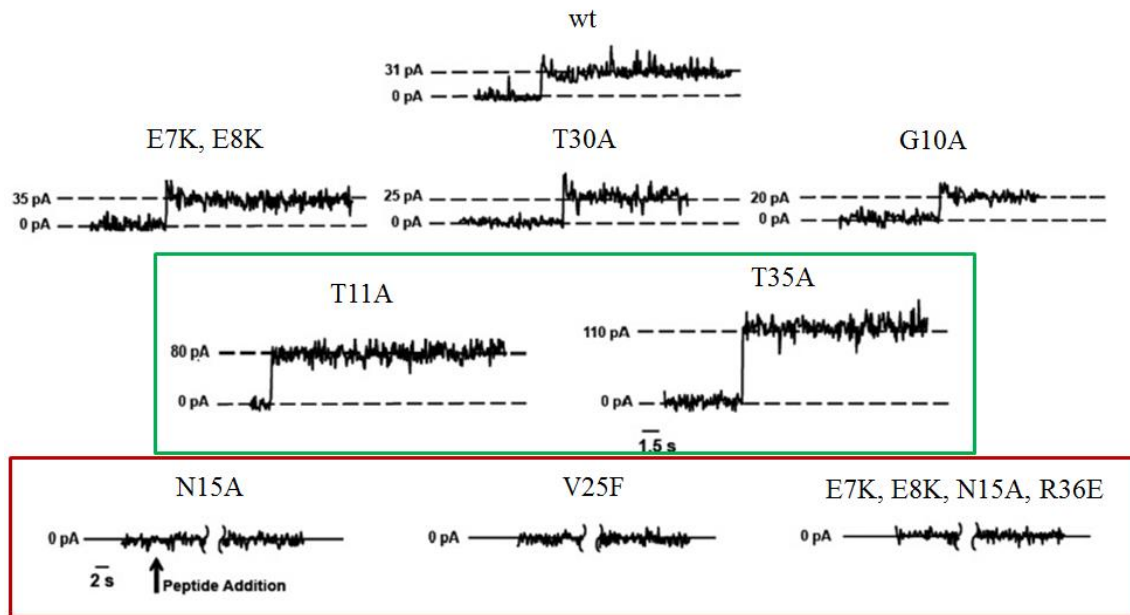
Figure 1.11. SARS-E viroporin mutations. Modified from (Verdia-Baguena et al. 2012).

The PDZ binding motif of CoV E proteins alters cellular pathways

Within the infected host cell, CoV E proteins have been proposed to alter cellular

pathways through a putative PDZ binding motif (PBM) comprised of the last four

residues at the C-terminus (Teoh et al. 2010; Javier and Rice 2011; Schoeman and

Fielding 2019). SARS-CoV E has been found to bind two cellular proteins containing

PDZ domains: PALS1 (Teoh et al. 2010), and syntenin (Jimenez-Guardeno et al. 2014).

CoV E proteins as drug targets

The multiple functional roles of CoV E proteins provide independent drug targets

exploiting either the ion channel activity (Wilson, Gage, and Ewart 2006a) or the PDZ

binding domain. The broad-spectrum ion channel inhibitor hexamethylene amiloride

(HMA) (Fig. 1.12A) was found to inhibit ion channel activity of MHV E in a dose-

dependent manner (Fig. 1.12B-C) and attenuate viral replication in vitro (Fig. 1.12D);

however, the parent compound amiloride (Fig. 1.12A), which lacks the hexamethylene

ring, neither inhibits the ion channel activity nor attenuates viral replication.
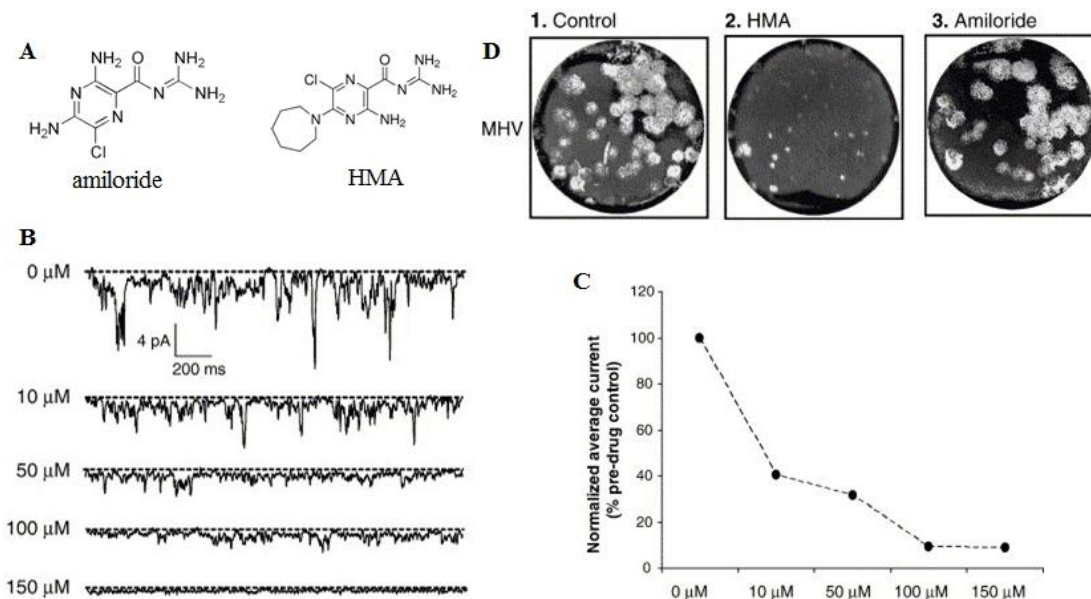
Figure 1.12. HMA inhibits viroporin activity and attenuates viral replication. (**A**)
Structures of amiloride and hexamethylene amiloride (HMA). (**B**) Dose-dependent
inhibition of MHV-E ion channel activity by HMA, and (**C**) corresponding plot of the
normalized average current as a function of HMA concentration. (**D**) HMA greatly

attenuates viral replication, however amiloride treated cells have viral titers comparable to wild type MHV. Reproduced from (Wilson, Gage, and Ewart 2006a).

The drug SB2035805, which targets the cellular pathway altered by the PBM of SARS-CoV-1, has been shown to increase the survival of the host in animal models (Jimenez-Guardeno et al. 2014). Wild-type SARS-CoV infection results in inflammatory cell infiltration of alveolar and bronchiolar airways. However, recombinant virus lacking the PBM shows minimal cell infiltration four days post infection (Fig. 1.13).
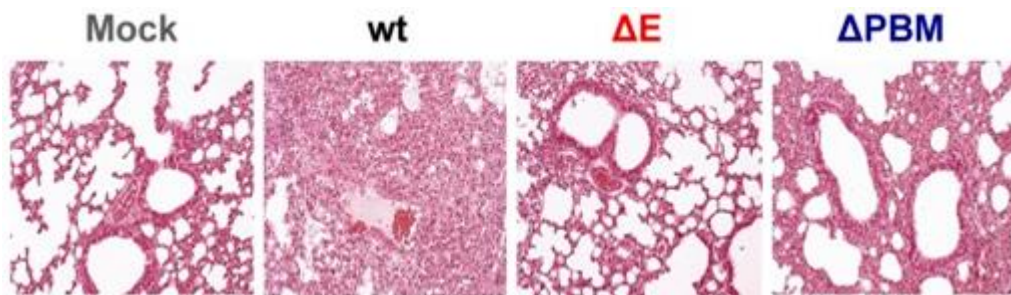
Figure 1.13. SARS-CoV lacking the PBM of E exhibits decreased lung pathology. Reproduced from (Jimenez-Guardeno et al. 2014).

Interestingly, when the E protein was deleted from SARS-CoV, the accessory protein 8a mutated after 10 serial passages to encode an internal PBM (Fig 1.14). The design of a live, attenuated vaccine has been proposed with the PBM of SARS-CoV E mutated in such a way to attenuate function while remaining intact (Jimenez-Guardeno et al. 2015).

A

```
8a-ΔE-p0    MKLLIVLTCISLCSCICTVVQRCASNKPHVLEDPCKVQH                    39
8a-ΔE-p10   MKLLIVLTCISLCSCICTVVQRCAFSLCSCICTVVQRCASNKPHVLEDPCKVQH      54
```

B
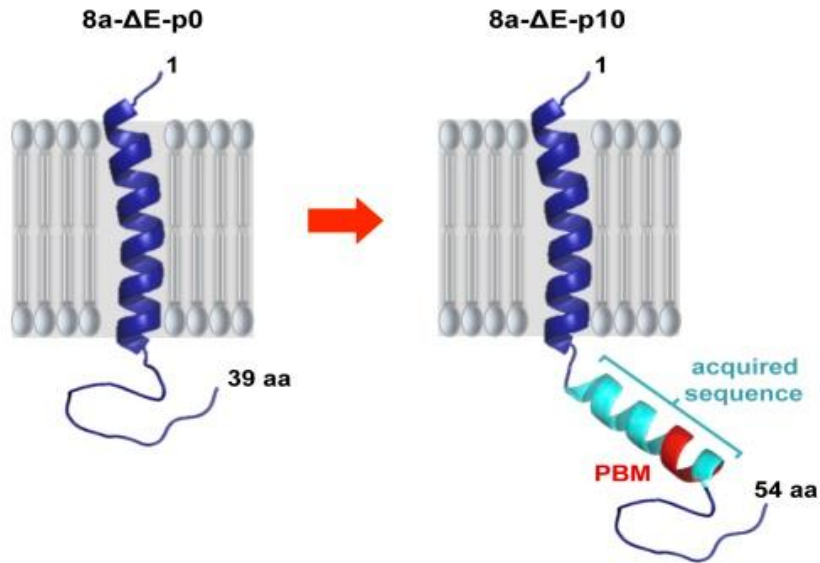
Figure 1.14. SARS-CoV 8a protein acquires a PBM in recombinant SARS-CoV virus lacking the E protein. **(A)** Primary sequences and **(B)** predicted structures of wild-type SARS-CoV 8a protein (8a-ΔE-p0) and mutated SARS-CoV 8a protein following 10 passages (8a-ΔE-p10). Cyan and red indicate the acquired sequence and the PBM, respectively. Reproduced from (Jimenez-Guardeno et al. 2015).

Structures of CoV E proteins

The only known structural information on CoV E proteins are solution NMR structures of SARS-CoV E truncations (Fig. 1.15) which omit the PBM. The first structure of a CoV E protein was solved for residues 1-38 in DPC micelles by Pervushin et al. in which the pentamer was modeled using NOE cross peaks that could not be accounted for by intra-monomer interactions. The structure-function relationships underpinning ion channel conductance and HMA binding were analyzed in detail for the transmembrane construct. The conserved polar residues were found to be lumen facing in the pentameric viroporin

structure. One of the conserved polar residues, Asn15, which is thought to form the selectivity filter, was proposed to be involved in HMA binding (Fig. 1.15A) (Pervushin et al. 2009). The structure of residues 8-65 of SARS-CoV E in DPC/SDS mixed micelles (Fig. 1.15B) revealed a conserved Golgi targeting signal, and the binding sites of HMA and the non-structural protein Nsp3 were mapped (Li, Surya, et al. 2014). Unambiguous inter-monomeric NOE restraints obtained from mixed labelled samples were used to solve the structure of the SARS-CoV E_{8-65} pentamer in LMPG micelles (Fig. 1.15C), which revealed the quaternary contacts in the extra-membrane helices that stabilize the viroporin structure (Surya, Li, and Torres 2018).

A — SARS-E·DPC 1-38
Pervushin et al.

B — SARS-E·DPC/SDS (5:1) 8-65
PDB ID: 2MM4
BMRB entry:19845

C — SARS-E·LMPG 8-65
PDB ID: 5x29
BMRB entry: 36049
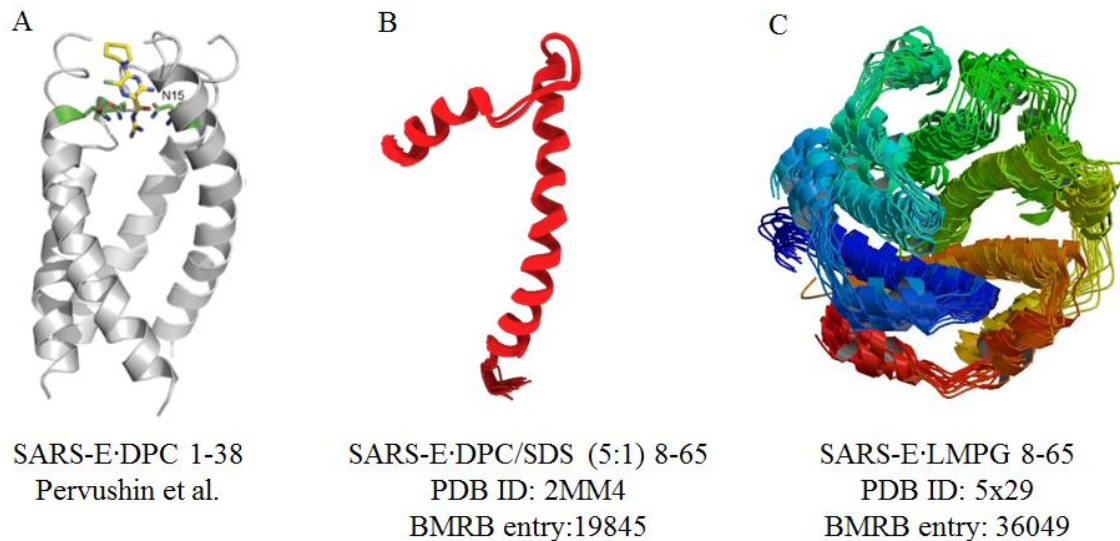
Figure 1.15. Known CoV E structures. **(A)** Pentameric model of the SARS-CoV E_{1-38} viroporin in DPC micelles bound to the ion channel inhibitor HMA, where one of the monomeric subunits has been removed for clarity of viewing HMA interactions with the Asn15 side chain. **(B)** Monomeric structure of SARS-CoV E_{8-65} in DPC/SDS mixed micelles (5:1 molar ratio of DPC:SDS). **(C)** Pentameric model of the SARS-CoV E_{8-65} viroporin in LMPG micelles. Reproduced from (Pervushin et al. 2009; Li, Surya, et al. 2014; Surya, Li, and Torres 2018).

NMR structure determination

Overview

Readers are referenced to (Ernst, Bodenhausen, and Wokaun 1987) for a more comprehensive description of NMR theory and (Rule and Hitchens 2006; Cavanagh 2018) for the application of NMR to protein structure determination. In contrast to X-ray crystallography or cryo-EM, where ionizing radiation interacts with the electron clouds, NMR is based transitions between nuclear spin states and generally relies on spin-1/2 nuclei, ^1H, ^{13}C, and ^{15}N being the most biologically relevant isotopes. Many isotopic labelling strategies are available in protein NMR, some involving deuteration, or stereoselective labeling of specific residues. The inherent low sensitivity of NMR requires concentrated samples (~1 mM) and samples must be stable for extended periods that are required to carry out multi-dimensional experiments. The process of structural determination by NMR begins with the assignment of chemical shifts to their respective atoms within the protein, followed by collecting conformational restraints for the assigned atoms. The types of conformational restraints used in the structure calculation are described in greater detail in following sections. NMR structure determination is an underdetermined problem in terms of the incompleteness of conformational restraints that are used to define atomic coordinates, which can be caused by low signal-to-noise ratio (S/N), unfavorable local dynamics, overlap of peaks, etc. Therefore, a unique solution cannot be determined, but instead an ensemble of conformations are calculated, each one satisfied by the experimental NMR restraints. Following the initial calculation, the restraint violations are analyzed and new restraints are assigned accordingly and these

21

updated restraints are then used to calculate a new ensemble. This process is repeated

iteratively until no more new restraints can be assigned and the number of restraint

violations is minimized. At this stage, refinement is then carried out using the full force

field, which is simplified in the beginning stages of the calculation to reduce

computational time. Refinement is typically done in an implicit bilayer in the case of

membrane proteins (Tian et al. 2015). Typically, the 10-20 lowest energy structures are

reported with associated backbone root mean square deviation (RMSD).

Resonance assignments

A prerequisite for structure determination (more generally for dynamics and binding

studies) is assignment of the resonance frequencies. Triple resonance experiments

developed by Bax and co-workers (Kay et al. 1990) utilize one and two bond scalar

couplings shown in Fig. 1.16 to establish sequential connectivity.
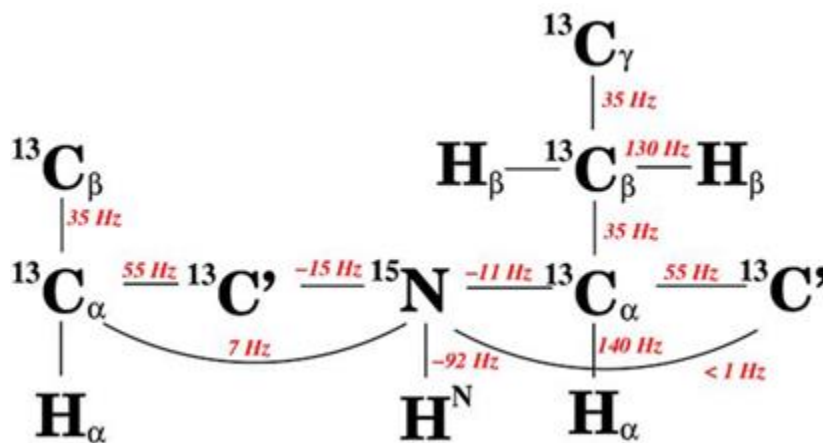
Figure 1.16. One- and two-bond scalar couplings in proteins utilized for magnetization transfer in triple resonance experiments. Reproduced from http://chemsites.chem.rutgers.edu/~skim/nmr_exp_TRE_backbone.html.

Common triple resonance experiments and their associated magnetization transfer

pathways, which are used for resonance assignments, are shown in Fig. 1.17. TROSY

selection is generally utilized in amide groups (Pervushin et al. 1997) and the aromatic

carbon (Pervushin et al.) for membrane proteins to increase sensitivity and resolution.
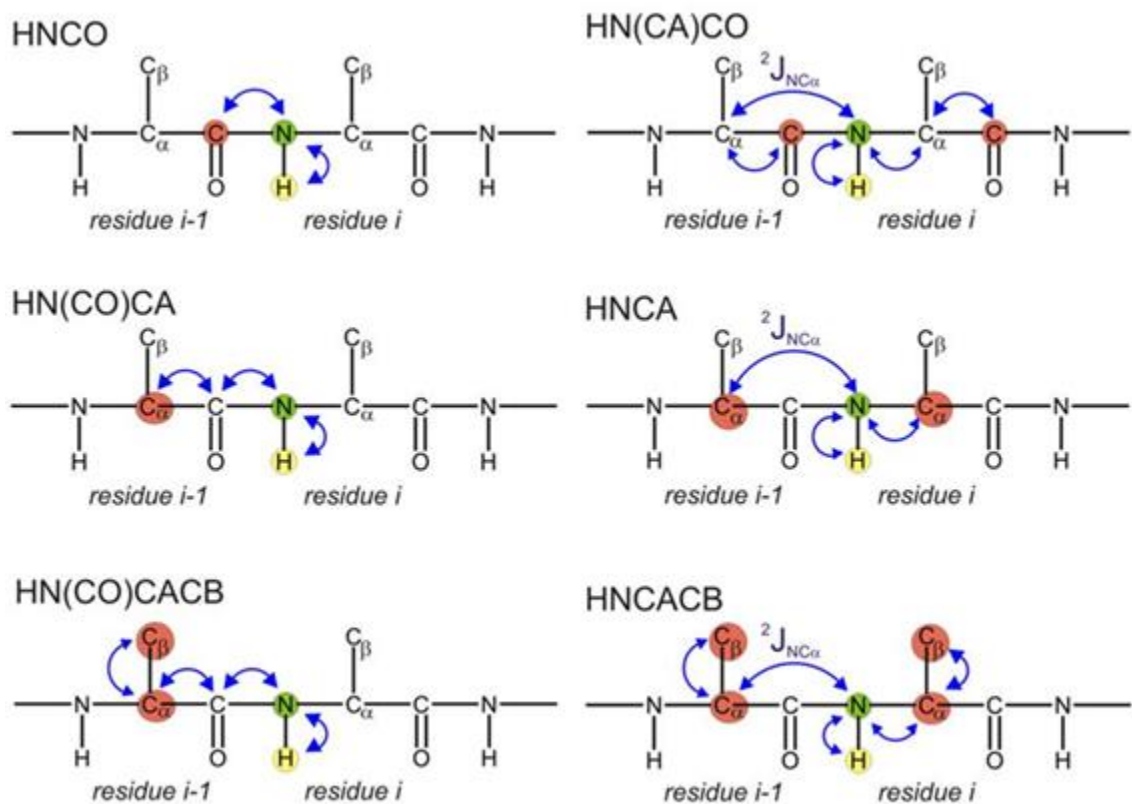
Figure 1.17. Magnetization transfer pathways of common triple resonance experiments. The abbreviations for these experiments denote the order that frequency labeling occurs in the pulse sequence, with parenthesis denoting magnetization transfer through the indicated spin with no frequency labeling. Chemical shifts are recorded for circled atoms, in which each dimension of the 3D experiment is designated by separate colors. Arrows denote magnetization transfers utilizing the one- and two-bond scalar couplings (Fig. 1.16). Reproduced from
http://chemsites.chem.rutgers.edu/~skim/nmr_exp_TRE_backbone.html.

Briefly, the strategy involves performing an experiment that correlates the amide group

of a residue with both the inter- (i-1) and intra-residue (i) backbone or sidechain atoms,

such as the HNCACB (Fig. 1.18), which correlates the amide group of the i residue with both the i and i-1 α and β carbons, to match the inter-residue C_α and C_β chemical shifts associated with the amide group of the i residue with the intra-residue C_α/C_β shifts of the sequential amide. This process is repeated in what has been termed 'walking the backbone' until all sequential connectivities have been established. Placement of these connected stretches within the primary sequence is done through the use of the characteristic C_β chemical shifts of Ala, Gly, Ser, Thr. An example of this strategy is illustrated in Fig. 1.18, which shows the resonance assignments in the putative PDZ binding motif (PBM) of MHV-E. Inversely labelled samples can be used to verify the assignments (Shortle 1994). Application of a combinatorial approach using a set of suitable inversely labelled samples has been proposed as an alternative strategy to obtain resonance assignments at the 2D level which can overcome problems of lower sensitivity for larger molecular weight and membrane proteins (Hiroaki et al. 2011).
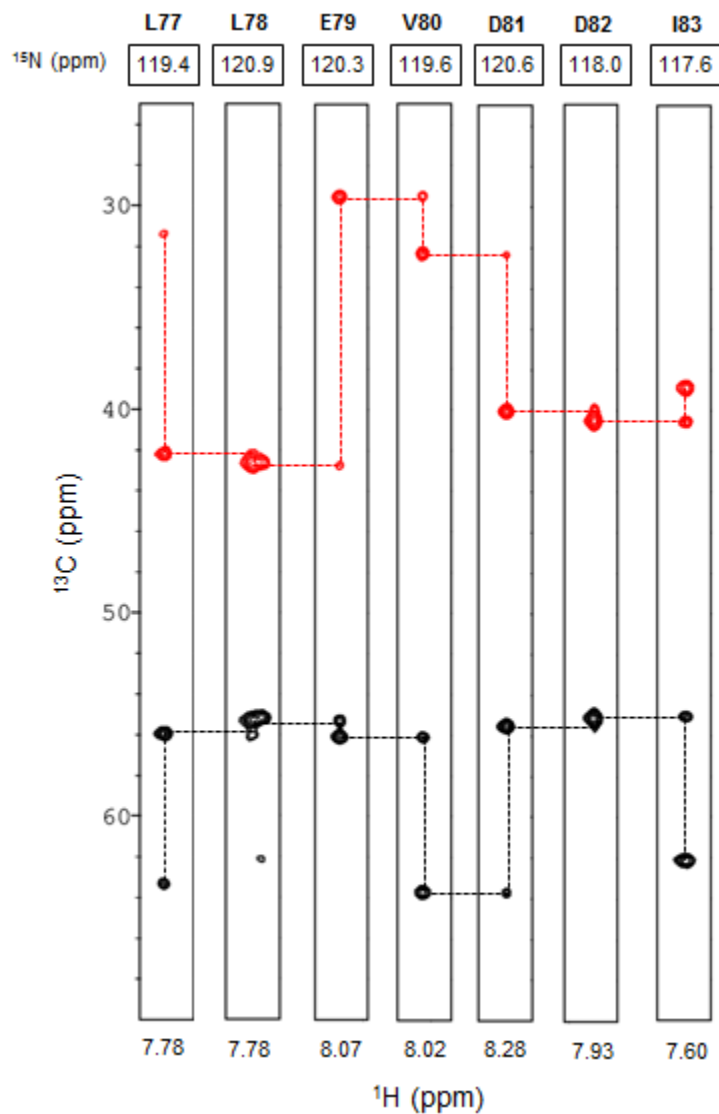
Figure 1.18. Resonance assignment of the MHV-E putative PDZ binding motif. Strips of the TROSY-HNCACB spectrum are shown from L77-I83 (PBM thought to be located at the last 4 amino acids in CoV E proteins (Schoeman and Fielding 2019); in the case of MHV-E, V80-I83) where connectivity between C_α and C_β peaks are indicated by black and red dotted lines, respectively.

25

Structural Restraints from NMR

The following sections summarize the various conformational restraints obtained by NMR that are used in structure calculation.

Torsion Angles

Chemical shifts are converted to restraints for the backbone torsion angles, psi and phi, and more recently for the side chain dihedral angle χ_1. TALOS (Torsion Angle Likelihood Obtained from Shift and sequence similarity) is an extension of the Chemical Shift Index (CSI) method which implements a database of known structures to provide dihedral angle restraints (Shen and Bax 2013). Alternatively, J couplings can be used to obtain dihedral angles, which is more direct than the CSI method (Salvador 2014).

Nuclear Overhauser Effects (NOEs)

NOEs build up through cross-relaxation during a mixing time which is described by the Solomon equation (Ernst, Bodenhausen, and Wokaun 1987). Protons within ~5 Å give rise to a cross peak in the NOESY spectrum, for which assignment provides an inter-proton distance restraint. Because of certain simplifications, such as the isolated spin pair approximation and variation in local dynamics, the distances obtained from NOEs are not exact and are expressed as bounded values. A common approach for the translation   of NOE cross-peak intensities (or volumes) to distances is to categorize them into three groups (strong, medium, weak). Upper distance limits are set based on calibration of NOE cross peak intensities (or volumes) to known distances, and the lower bound is set to the sum of the van der Waals radii (Cavanagh 2018). Fig. 1.19A shows the NMR

ensemble of chicken cofilin (PDB coordinate file: 1TVJ). The dense network of NOEs

used in the structure calculation can be visualized with the stereoview (Fig. 1.19B).
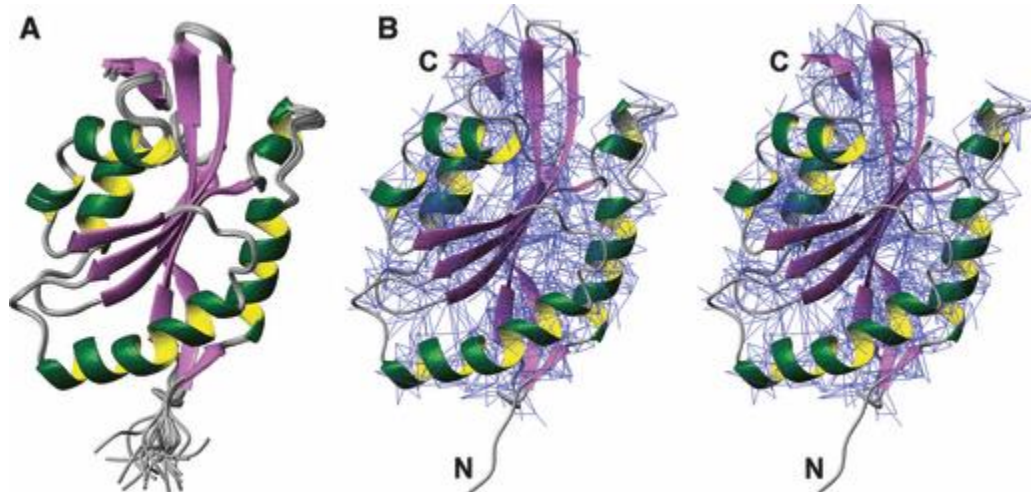
Figure 1.19. Visualizing NOE distance restraints. **(A)** Ensemble of 20 structures of chicken cofilin (PDB 1TVJ). **(B)** Stereoview of the first structure of the ensemble in part **(A)** showing the dense network of NOE restraints used in the structure calculation. Figure reproduced from (Kwan et al. 2011)

.**Hydrogen bonds**

Hydrogen bonding constraints are typically inferred from deuterium exchange rates

(hydrogens that are involved in hydrogen bonding experience a slower rate of exchange)

and NOE patterns of secondary structure. Additionally, there are experiments that

unambiguously identify the acceptor group of the hydrogen bond (Cavanagh 2018). In

the structure calculation, hydrogen bonds are expressed as NOE-type distance restraints

between the H and O that constitute the hydrogen bond. These restraints generally have

significant outcomes on the resulting structure (Cavanagh 2018).

Residual dipolar couplings (RDCs)

Dipole-dipole coupling, which is the dominating interaction in the solid state, becomes averaged to zero due to fast molecular reorientation in solution. The dipolar couplings can be re-introduced to a very small extent, on the order of the J-coupling, using suitable weakly orienting media for alignment in order to achieve partial alignment of the proteins on the order of 1 in 1000 such that the dipolar couplings are observed in the low Hz range (reduced from the kHz range in solids). RDCs measured in weakly aligning media provide internuclear bond vector orientations relative to an external alignment tensor. They are generally measured in amide groups. Unique alignment media are required to lift the inherent degeneracy associated with RDCs (Bax and Grishaev 2005).

Paramagnetic resonance enhancements (PREs)

Paramagnetic relaxation enhancement (PRE) can, in principle, detect interactions between the unpaired electron of a paramagnetic label and protons up to 35 Å away (Clore and Iwahara 2009). PRE-based distance restraints are generally applied in structure calculations of multi-pass α-helical membrane proteins to establish tertiary contacts. PREs have also been used to lift the degeneracy associated with RDCs for the single-pass α-helical membrane protein phospholamban as shown in Fig. 1.20A-C (Shi et al. 2011).
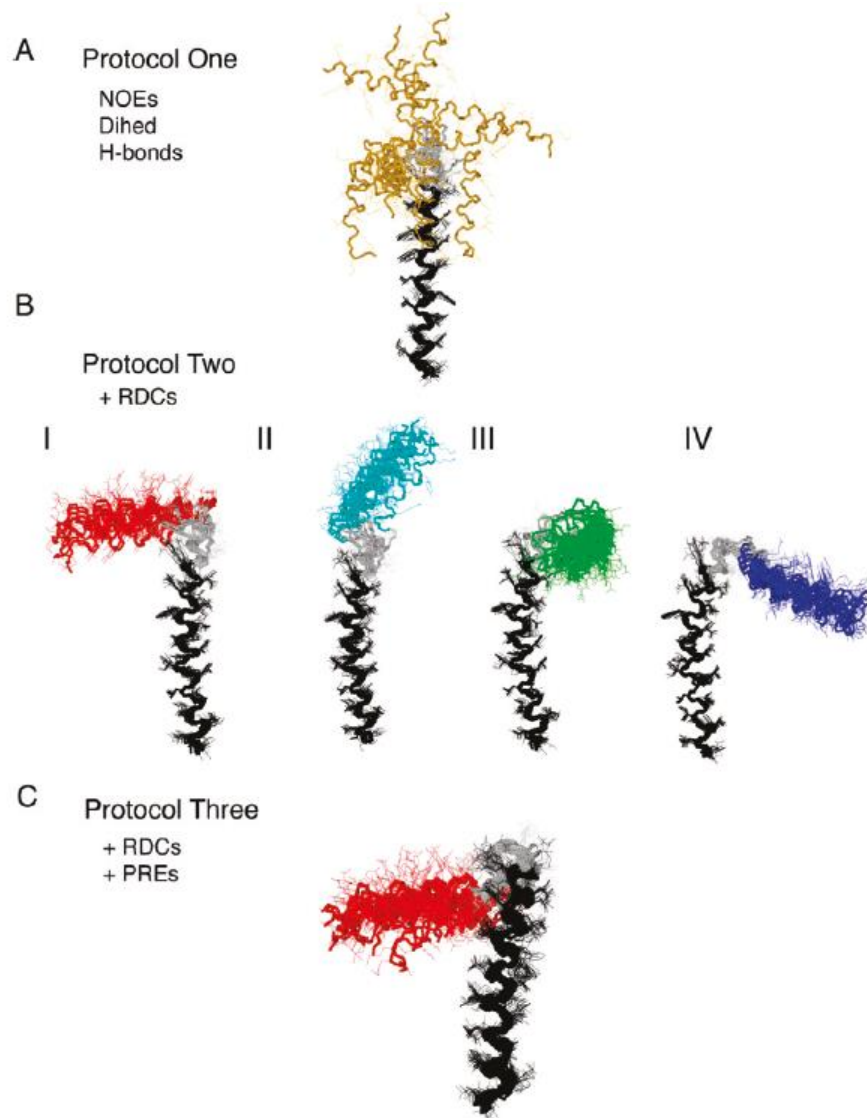
Figure 1.20. PREs lift the degeneracy associated with RDCs in phospholamban. Structure calculation protocols using (A) only NOEs, dihedral angles, and hydrogen bond restraints is insufficient to define the orientational relationship between the extra-membrane helix (colored yellow) and the transmembrane helix. **(B)** Incorporation of RDC restraints yields four degenerate structures I-IV (degenerate in the sense that they all satisfy the RDC constraints.) **(C)** Inclusion of both RDC and PRE restraints uniquely define structure I. Reproduced from (Shi et al. 2011).

Conclusions

Application of NMR to structure determination of single pass alpha helical membrane proteins such as MHV-E requires long-range restraints that were outlined in the preceding sections, and will ultimately expand the known structural information on CoV E proteins.

CHAPTER 2

SOLUTION NMR MEMBRANE MIMETIC SCREENING OF THE MOUSE HEPATITIS CORONAVIRUS A59 ENVELOPE PROTEIN

Bobby Baravati[1,2] *, Jonathan Carillo[3,4] *, James D. Zook[1,2,5], Brian R. Cherry[1,6], Sasha M. Daskalova[1,2,3,7], Felicia M. Craciunescu[2], Michael Goryll[1,8], Wade D. Van Horn[1,9,10], Xu Wang[1], Debra T. Hansen[2,11], Brenda G. Hogue[2,3,4], Petra Fromme[1,2]

[1]School of Molecular Sciences, Arizona State University, Arizona State University, Tempe, Arizona, USA; [2]Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe, Arizona, USA; [3]Biodesign Center for Immunotherapy, Vaccines and Virotherapy, Arizona State University, Tempe, Arizona, USA; [4]School of Life Sciences, Arizona State University, Tempe, Arizona, USA; [5]Present address: LGC, Biosearch Technologies, Middleton, Wisconsin, USA; [6]Magnetic Resonance Resource Center, Arizona State University, Tempe, Arizona, USA; [7]Present address: Biodesign Center for BioEnergetics and School of Molecular Sciences, Arizona State University, Tempe, Arizona, USA; [8]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona, USA; [9]Biodesign Virginia G. Piper Center for Personalized Diagnostics, Arizona State University, Tempe, Arizona, USA; [10]Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, Arizona, USA; [11]Biodesign Center for Innovations in Medicine, Arizona State University, Tempe, Arizona, USA;

*These authors contributed equally to the manuscript

31

Corresponding authors: Petra Fromme E-mail: pfromme@asu.edu and Brenda

Hogue Email: Brenda.hogue@asu.edu

Highlights

- MHV-E was expressed and purified to homogeneity

- Viroporin activity was confirmed from electrophysiology measurements

- TROSY-HSQC screening of membrane mimetics performed

- Deuterium exchange identifies transmembrane residues

- Optimization of spectral resolution and stability for NMR structure and dynamics
  accomplished

Abstract

Coronaviruses are the causative agents of the SARS, MERS and recent COVID-19

epidemics. COVID-19 has been reported to have more than 260,000 cases with over

10,000 deaths at the time of this writing, with numbers continuing to rise. They are

prevalent in humans and animals and can overcome the barrier between species. The

coronavirus envelope proteins have been implicated in various stages of the viral life

cycle, and are therefore promising drug targets. The full-length, wild type mouse hepatitis

coronavirus envelope protein was heterologously expressed as a GFP fusion protein in

Escherichia coli and purified in a single step using an on-column exchange/cleavage

method. The α-helical content as measured by CD spectroscopy was in agreement with

predicted values, and ion channel activity of the viroporin was confirmed from electrophysiology measurements. To find conditions amenable for solution NMR structure and dynamics studies, membrane mimetics were screened by ^{1}H-^{15}N-TROSY-HSQC, which revealed that lysophospholipid micelles yielded the highest quality TROSY spectra. Deuterium exchange was done to identify the transmembrane residues. Spectral resolution and sample stability in lysomyristoylphosphatidylcholine micelles have been optimized for structure determination of the full-length mouse hepatitis coronavirus envelope protein.

Keywords

Solution NMR, viral membrane protein, envelope protein, membrane mimetic screening, deuterium exchange, coronavirus

1. Introduction

Coronaviruses are the causative agents of the SARS, MERS and COVID-19 epidemics. For SARS, approximately 8000 cases were reported with a 10% mortality rate, and MERS has infected 2494 people leading to death in 34% of cases (Peeri et al. 2020). The recent emergence of the novel coronavirus that originated in Wuhan City, China, named SARS-CoV-2, has already been reported to have spread worldwide with 266,073 cases and 11,184 deaths reported by https://coronavirus.jhu.edu/map.html which has prompted the World Health Organization to designate COVID-19 as a pandemic (WHO website, https://www.who.int/emergencies/diseases/novel-coronavirus-2019, accessed 3/21/20). SARS and MERS coronaviruses are thought to have originated from animal reservoirs in which the dynamic landscape of the viral genome allowed the virus to cross the species

33

barrier (Peeri et al. 2020). Because of the potential for coronaviruses to jump the species

barrier, there always exists the possibility for future epidemics. Therefore, it is critical to

undertake basic research of coronaviruses to develop intervention strategies and mitigate

future outbreaks. Considered a prototype beta-coronavirus, Murine Coronavirus Mouse

Hepatitis Virus (MHV) is studied as a model for human disease (Conn 2008).
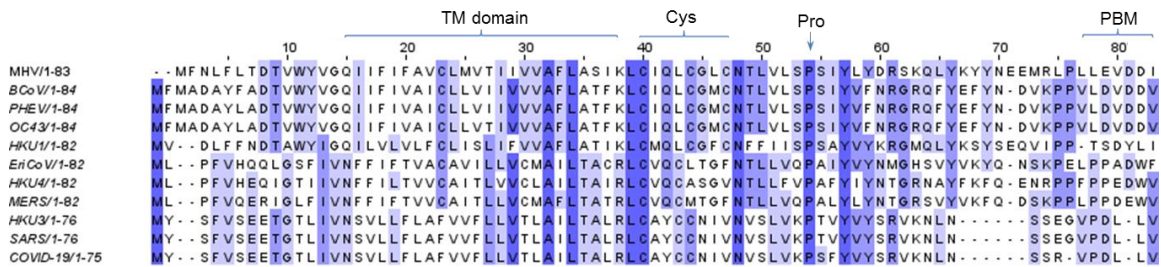
Figure 2.1. Sequence alignment of the betacoronavirus envelope proteins. Residues are
numbered in relation to MHV and shaded by % sequence identity. The transmembrane
domain is denoted by TM domain. Cys denotes the cysteine-rich membrane proximal
region, Pro denotes the conserved proline of the proposed β-turn, and PBM denotes the
PDZ binding motif. Multiple sequence alignment (MSA) was done using T-Coffee
(Notredame, Higgins, and Heringa 2000) and viewed with JalView (Clamp et al. 2004).

The structural proteins comprising the viral envelope are the spike (S), membrane (M),

and envelope (E) proteins (Nieva, Madan, and Carrasco 2012). The MHV-E protein is an

integral membrane protein 83 residues in length that is present in low copy numbers in

the viral envelope. There are ~20 copies per virion, and it is not known to form viroporins

within the viral envelope. Multiple sequence alignment of the betacoronavirus E proteins

(Fig. 2.1) shows large variations in sequence identity, however there are some key

similarities: the size and location of the transmembrane domains, the conserved proline at

the center of a proposed β-turn, a cysteine-rich membrane proximal region, alignment of

polar residues within the lumen of the channel and a putative C-terminal PDZ binding

motif. The coronavirus envelope proteins are promising drug targets because they are

34

implicated in various stages of the viral life cycle. Coronavirus E proteins localize to the Golgi apparatus and endoplasmic reticulum Golgi apparatus intermediate compartments (ERGIC) membranes of the infected host cell (Ye and Hogue 2007), (Machamer and Youn 2006), (Cohen, Lin, and Machamer 2011) where they form pentameric ion channels, called viroporins (Nieva, Madan, and Carrasco 2012), (OuYang and Chou 2014). Additionally, coronavirus E proteins are known to modulate cellular pathways through interactions with host cell proteins (reviewed in (Torres et al. 2015)). In the case of the SARS envelope protein, a PDZ binding motif in the last 4 amino acids of the C-terminus has been shown to bind PALS1, a tight junction-associated protein, which results in alteration of tight junction formation and epithelial morphogenesis (Teoh et al. 2010), (Javier and Rice 2011). Other interactions with cellular proteins have been proposed such as binding to or indirect modification of host cell ion channels (Torres et al. 2015). Coronavirus E proteins also interact with the M protein in the formation of the viral envelope and virus like particles (VLPs) (Ruch and Machamer 2012), (Machamer and Youn 2006). The mechanistic details of these E-M interactions in the formation of the viral envelope remain to be resolved, however it is known that E and M proteins interact at their cytoplasmic tails (Corse and Machamer 2003). Mutations at the C-terminal extramembrane domain of MHV-E have been shown to affect viral assembly and maturation (Fischer et al. 1998). Coronavirus E proteins are similar in their structure and function, as implied from the fact that viral replication is not significantly affected in chimeric viruses in which the MHV-E protein has been swapped with that of bovine coronavirus, infectious bronchitis virus (IBV), or SARS coronavirus (Kuo, Hurst, and Masters 2007).

35

There are no structures reported for MHV-E. Solution NMR structures for truncated

versions of SARS-E encompassing the transmembrane domain (Pervushin et al. 2009)

and residues 8-65 in DPC/SDS mixed micelles (Li, Surya, et al. 2014) and LMPG

micelles (PDB 5X29, BMRB 36049) (Surya, Li, and Torres 2018) have been reported

from which the binding site of the channel inhibitor hexamethylene amiloride (HMA) has

been proposed. For the transmembrane (TM) structure, the authors in (Pervushin et al.

2009) propose that Asn15 is a key residue in mediating the interaction with HMA. HMA

also inhibits channel activity of the MHV-E viroporin (Wilson, Gage, and Ewart 2006a),

in which the corresponding Gln15 residue has been proposed to be involved in binding

(Pervushin et al. 2009). This work outlines the first NMR studies on MHV-E and will

provide a platform for structure based drug design and probing interactions of full-length

wild type E protein, encompassing the putative PDZ binding motif, with cellular and viral

proteins.

2. Materials and Methods

2.1. Expression

MHV-E from strain A59 (accession P0C2R0) was cloned into the pRSET-natGFPHis

vector (Martin-Garcia et al. 2014) as a C-terminal GFP fusion with an intervening

thrombin cleavage site, and a C-terminal His$_6$ tag. The complete DNA and protein

sequences are shown in Supplementary Figures S1A and S1B, respectively. The plasmid

was transformed into *E. coli* Lemo21(DE3) cells (Schlegel et al. 2012). Factorial analysis

(Box, Hunter, and Hunter 2005) was applied to optimize expression using a 2^3 full

factorial screen in which temperature, time and concentration of IPTG were each tested at

2 levels. Whole cell GFP fluorescence was used to measure expression levels and in-gel

fluorescence (Fig. S2) was used to verify that the fusion was intact as outlined in (Drew

et al. 2006). Statistical analysis of the expression screen was done with design of

experiments (DOE) in the JMP software.

Uniformly ^{15}N labeled and ^{13}C/^{15}N double labeled MHV-E-GFP were obtained by

growing cells in M9 minimal media, consisting of 8.5 g/L Na_2PO_4, 3 g/L KH_2PO_4, 0.5

g/L NaCl, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 0.1 mM $FeCl_3$, 1x MEM vitamins (Gibco), 1

g/L $^{15}NH_4Cl$ and 4 g/L ^{12}C- or ^{13}C-glucose (Cambridge Isotope Laboratories) at 37 °C

until the OD_{600} reached 0.4, cooled on ice for a few minutes and grown at 25 °C until the

OD_{600} reached 0.6, at which point protein expression was induced with 0.4 mM IPTG.

After 24 h, cells were harvested by centrifugation at 5,000 x g for 30 min at 4 °C and

stored at -80 °C.

2.2. Purification

2.2.1. Lysis and extraction

Cell pellets were resuspended at a 1:10 ratio (grams of wet cell pellet mass to mL of lysis

buffer) in lysis buffer composed of 50 mM HEPES, pH 7.5, 300 mM NaCl, Sigma

EDTA-free protease inhibitor cocktail (1 tablet/100 mL) and lysed by sonication on ice,

followed by centrifugation at 45,000 x g for 1 h at 4 °C. The supernatant was discarded

and pellet was washed with lysis buffer, followed by centrifugation at 45,000 x g for 1 h

at 4 °C. The supernatant was discarded and the pellet was resuspended in a 1:50 ratio

(extraction buffer volume to culture volume) of extraction buffer composed of lysis

buffer plus 10 mM imidazole and 2% SDS. Following a 1 h incubation with gentle

mixing at room temperature, insoluble particles were removed by centrifugation at 45,000 x g for 1 h at room temperature.

2.2.2. Membrane mimetic screening

A single step exchange and cleavage method based on (Hefti et al. 2001) (Dian et al. 2002) was utilized, in which thrombin cleavage was used to elute the protein instead of the high concentrations of imidazole that may be destabilizing for certain membrane proteins. For the small-scale screening, 250 µL of MHV-E-GFP-His$_6$ (~1 mg) in SDS micelles was immobilized onto a 100 µL His SpinTrap column (GE Healthcare) pre-equilibrated with 50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole and the selected membrane mimetic. The SDS was exchanged with a selected membrane mimetic by extensive washing (40 column volumes [CV]), followed with thrombin cleavage by incubating the resin with 200 µL of 20 mM HEPES, pH 7.5, 150 mM NaCl, 5 mM imidazole, the selected membrane mimetic and 20 units of thrombin (BioPharm Laboratories) for 2 h at room temperature. Following the 2 h incubation, the cleaved MHV-E protein was eluted from the column material by centrifugation. The column was then washed with 5 CV of 50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole containing 2% SDS to elute any remaining E protein. Finally, the GFP and remaining uncleaved E-GFP were eluted with 50 mM HEPES, pH 7.5, 300 mM NaCl, 2% SDS and 0.5 M imidazole. SDS-PAGE analysis was done to identify compatible membrane mimetics before scaling up.

The membrane mimetics which could be exchanged and cleaved at high yield were selected for scale-up on an ÄKTA explorer FPLC (Amersham) to generate enough

material for screening by ^{15}N-TROSY-HSQC. Approximately 40 mg of E-GFP from 0.5 L M9 expression in 10 mL lysis buffer containing 10 mM imidazole and 2% SDS was loaded onto a 5 mL nickel affinity column (HisTrap HP, GE Healthcare) pre-equilibrated with 50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole and the selected membrane mimetic. The flow rate was 0.5 mL/min during sample loading and 5 mL/min for other steps. This was followed by a 20 CV exchange step and cleavage with 500 units of thrombin for 2 h at room temperature to give purified, cleaved E protein in the selected membrane mimetic.

Lipid/detergent mixed micelle stock solutions were prepared using 10 freeze-thaw cycles in liquid nitrogen and a 37 °C heating block. Absorbance at 550 nm was used to monitor turbidity to ensure complete solubilization of the mixed micelles.

MHV-E in bicelles was reconstituted from anionic POPC/POPG (10:1) proteoliposomes following the methods outlined in (Das, Murray, and Cross 2013) and (Morrison and Henzler-Wildman 2012) using a 1:100 molar ratio of lipid to protein. Biobeads SM-2 (Bio-Rad) were washed 3x in methanol and 3x in water directly before use.

Reconstitution of E into the amphipol A8-35 starting from DPC/POPC mixed micelles was done following the protocol outlined in (Zoonens et al. 2005) at a 1:5 ratio of MHV-E:A8-35. The SEC running buffer and NMR buffer were supplemented with 0.01% A8-35, and pH was maintained at pH 7.5 throughout.

2.2.3. SDS-PAGE analysis and cross-linking

SDS-PAGE analysis was done using 4-12% Bis-Tris acrylamide gels (Novex,

Invitrogen), MOPS SDS Running Buffer (Invitrogen) and molecular weight standard

(Bio-Rad). All samples were incubated in Bio-Rad XT Sample Buffer (catalog # 161-

0791) with 50 mM TCEP at 37 °C for 1 h prior to loading the gel. Silver staining was

done using the silver stain kit (Pierce catalog # 24612). For cross-linking experiments

~50 µM of MHV-E in selected membrane mimetic was cross-linked with 20 mM

glutaraldehyde at room temperature for 30 min and the cross-linking reaction was

stopped with addition of 1 M Tris-base pH 8.0 to 50 mM followed by incubation with

Bio-Rad XT Sample Buffer with 50 mM TCEP at 37 °C for 1 hr prior to loading gel

lanes.

2.2.4. Analytical Size Exclusion Chromatography

A 20 µL volume of ~1 mM MHV-E was loaded onto a Superdex 200 Increase 10/300 GL

column (GE Healthcare) connected to an ÄKTA system and run for 1.5 CV at a flow rate

of 0.75 mL/min.

2.3. Ion channel measurements

Following the method outlined in (Wilson et al. 2004), a lipid mixture consisting of 3:1:1

POPE:POPS:POPC in chloroform (Avanti Polar Lipids, Alabaster, AL)] was dried under

a steam of nitrogen and the resulting lipid film was resuspended in n-decane. Bilayers

were painted across an aperture of 150 µm in a 3 mL polystyrene cup (Warner

Instruments, Hamden, CT) separating the cis and trans chambers. To establish a salt

concentration gradient, the cis chamber contained 5 mM HEPES, pH 7.2, 500 mM NaCl and the trans chamber contained 5 mM HEPES, pH 7.2, 50 mM NaCl. Measurements were done with the cis chamber grounded and the transmembrane voltage of -100 mV applied to the trans bath via the current amplifier input. Currents were amplified using a Heka EPC-8 patch-clamp amplifier (Heka Instruments, Southboro, MA) at a gain setting of 1 mV/pA. The Bessel filter of the amplifier was set to 1 kHz and the output was sampled at 50 kHz using a National Instruments PCIe 6251 data acquisition board (National Instruments, Austin, TX). No further digital filtering was employed. The digital recording of the current was performed either using a custom LabView program that controlled the data acquisition as well as the stimulus voltage generation or using WinEDR (John Dempster, University of Strathclyde, Strathclyde, UK). Bilayer capacitance was determined by application of a triangular signal and current measurement and confirmed to be between 40-60 pF to ensure the bilayer thickness does not exceed a single bilayer. Bilayer integrity was evaluated by ensuring that the measured current remained at baseline (i.e., the bilayer is not leaky) at an applied voltage of -100 mV. MHV-E was added to the cis chamber and the applied voltage was cycled between -100 mV to 100 mV to facilitate insertion of the E-channel complex into the bilayer. A voltage sweep was performed from -100 mV to +100 mV in increments of 10 mV.

2.4. CD spectroscopy

CD spectra were acquired on a Jasco J-710 spectrometer in the far UV region (190-260 nm) at 25 °C in 1 nm intervals, with 5 scans averaged per spectrum. Membrane mimetics were screened in 10 mM sodium phosphate pH 7.5, 50 mM NaF, in a 1 mm path length

cell at a protein concentration ~50 µM. Data were baseline subtracted and smoothed

using a Savitzky-Golay filter in Jasco software (Miles and Wallace 2016). Deconvolution

of the data was done using the online server DichroWeb with the Contin-LL algorithm

using the SMP180 reference dataset (Whitmore and Wallace 2004) (Abdul-Gader, Miles,

and Wallace 2011).

2.5. Nuclear magnetic resonance spectroscopy

All NMR experiments were performed at the Magnetic Resonance Resource Center

(MRRC) at Arizona State University. Membrane mimetics were screened by ^{1}H^{15}N-

TROSY-HSQC on the Bruker Avance III 850 MHz spectrometer equipped with a 5 mm

TCI ^{1}H/^{15}N/^{13}C triple resonance cryoprobe. Fine screening and the BEST-TROSY-

HNCO experiment were performed on the Bruker Avance III 600 MHz spectrometer

equipped with a 5 mm TCI ^{1}H/^{15}N/^{13}C triple resonance cryoprobe. Spectra were

processed using NMRPipe (Delaglio et al. 1995) and viewed in NMRFAM-SPARKY

(Lee, Tonelli, and Markley 2015). All of the NMR experiments were done at 45 °C,

except the initial round of membrane mimetic screening, which was done at 25 °C.

For the deuterium exchange experiment, the sample was exchanged for 50 mM MES, pH

5.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP, 0.1% LMPC in 99% D_2O using a spin

desalting column (Thermo Zeba spin desalting columns) and incubated overnight before

collecting a ^{15}N-TROSY-HSQC spectrum.

3. Results

3.1. Membrane mimetic screening

Membrane mimetic screening is largely an empirical process. Selection of membrane mimetics for evaluation was guided by the literature and by databases of membrane proteins solved by solution state NMR (Biological Magnetic Resonance Data Bank [BMRB] (Romero et al. 2020), drorlist [http://www.drorlist.com/nmr.html]). The membrane mimetics screened in this study encompass various classes of mimetics (Warschawski et al. 2011), including micelles, lipid-detergent mixed micelles, lyso-phospholipid micelles (Koehler et al. 2010), bicelles (Durr, Gildenberg, and Ramamoorthy 2012), (Claridge et al. 2013), (Poget, Cahill, and Girvin 2007), (Li, To, et al. 2014) and amphipols (Zoonens et al. 2005). The membrane mimetics used in this screen outlined in Table 1 were also selected to provide adequate coverage of the experimental space in terms of molecular weight/size, head group charge and hydrophobic tail length. In order to initially identify compatible membrane mimetics, a small scale screening format was implemented as described in section 2.2.1.

A representative SDS-PAGE analysis of on-column exchange/cleavage for lipid-detergent mixed micelles is shown in Fig. 2.2. Lanes 2-4 and lanes 10-12 correspond to exchange from SDS micelles to lipid-detergent mixed micelles. Lanes 5 and 13 correspond to purified MHV-E-GFP-His$_6$ in SDS micelles which serve as a reference for evaluation of thrombin cleavage. Elution of MHV-E by thrombin cleavage results in high purity as seen in lanes 6 and 14. The absence of the MHV-E band in lanes 7 and 15, which correspond to exchange back to SDS micelles, indicate that solubility of MHV-E is

maintained in the lipid-detergent mixed micelles following cleavage. The thrombin cleavage runs to completion as indicated by the absence of any uncleaved MHV-E-GFP-His$_6$ band in lanes 8 and 16.

Table 2.1. Membrane mimetics used in the TROSY screen

Membrane Mimetic[a]	Monomer MW (g·mol⁻¹)	CMC[b] (mM)	N[c]	Micellar MW (kDa)	Representative solution NMR structures[d] (NMR conditions)
Micelles					
Non-ionic					
DM (C10-DM)	483	0.2	140	39	
OG (C8-G)	292	20	90	26	PagP (200 mM; 45 °C)
Zwitterionic					
c6-DHPC (D6-PC)	454	15	35	16	M2 (300 mM; 30 °C), OmpX (200 mM; 30 °C), Pf1 (100 mM; 50 °C)
c7-DHPC (D7-PC)	482	1.5	25	12	PR (40 mM; 50 °C), SRII (100 mM; 50 °C)
DPC (C12-PC)	352	1.5	54	25	CD4 (200 mM; 45 °C), DAGK (2 %; 45 °C), p7 (200 mM; 30 °C), PLN (200 mM; 30 °C)
HPC (C14-PC)	380	0.12	108	47	
LMPC (L14-PC)	468	0.04	122	57	
Anionic					
LMPG (L14-PG)	478	0.2	160	26	KCNE1 (80 mM; 40 °C), SARS-E (200 mM; 35 °C), TMEM14A (3 %; 37 °C)
LPPG (L16-PG)	507	0.02	55	81	hGlyR-α1 (100 mM; 40 °C), Ste2p (200 mM; 47 °C)
SDS (C12-S)	288	8	80	23	MerF (500 mM; 60 °C)
Mixed					
DPC/SDS 5:1	—	—	—	—	ZZ (150 mM DPC; 30mM SDS; 30 °C)
Lipid-detergent mixed micelles					
DPC/DMPC 4:1	—	—	—	—	
DPC/POPC 4:1	—	—	—	—	
Bicelles					
DHPC/POPC/POPG q=0.3[e] (isotropic)	—	—	—	—	
Amphipols					
A8-35	4300	0.0005	9	40	

[a]Membrane mimetic abbreviations: DM, decylmaltoside; OG, octylglucoside; c6-DHPC, 1,2-hexanoyl-1-sn-glycero-3-phosphocholine; c7-DHPC, 1,2-heptanoyl-1-sn-glycero-3-phosphocholine; DPC, n-dodecylphosphocholine; HPC, 1LMPC, 1LMPG, 1LPPG, 1-palmitoyl-2-hydroxy-sn-glycero-3-phospho-(1'-rac-glycerol); SDS, sodium dodecyl sulfate; DMPC, 1,2-myristoyl-1-sn-glycero-3-phosphocholine; POPC, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine; POPG, palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol).
[b]CMC, critical micelle concentration.
[c]N, aggregation number.
[d]Protein abbreviations: PagP (Hwang et al. 2002); M2, Influenza A M2 proton channel (Schnell and Chou 2008); OmpX, outer membrane protein X (Fernandez et al. 2004); Pf1, Pf1 coat protein (Park et al. 2009); PR, proteorhodopsin (Reckel et al. 2011); SRII, sensory rhodopsin II (Gautier et al. 2010); CD4, human CD4 (Witlich et al. 2010); DAGK, diacylglycerol kinase (Van Horn et al. 2009); p7, HCV p7 ion channel (OuYang et al. 2013); PLN, phospholamban (Oxenoid and Chou 2005); KCNE1, Potassium voltage-gated channel subfamily E member 1 (Kang et al. 2008); SARS-E, SARS envelope protein (Surya, Li, and Torres 2018); TMEM14A, transmembrane protein 14A (Klammt et al. 2012); hGlyR-α1, human glycine receptor alpha 1 (Mowrey et al. 2013);Ste2p, yeast alpha-factor receptor (Neumoin et al. 2009); MerF, mercury transporter F (Howell, Mesleh, and Opella 2005); ZZ, zeta-zeta transmembrane dimer (Call et al. 2006).

45
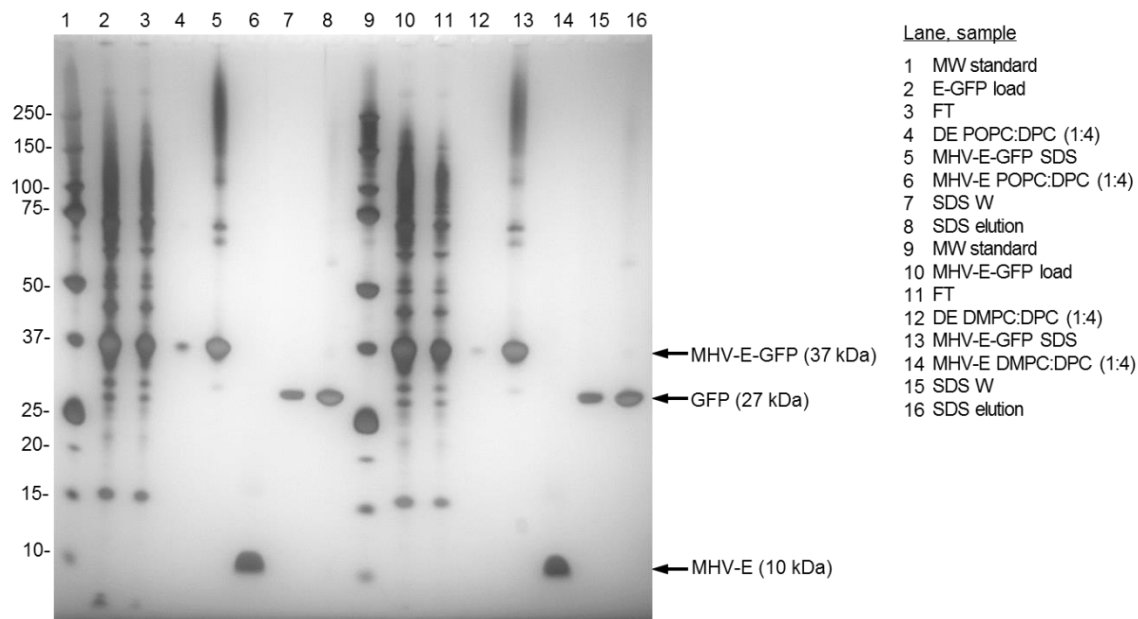
Figure 2.2. SDS-PAGE analysis of on-column detergent exchange and thrombin cleavage. SDS-PAGE analysis of on-column exchange from SDS to lipid/detergent mixed micelles and subsequent thrombin cleavage visualized by silver staining. Lanes 1-8: POPC/DPC mixed micelles (1:4 molar ratio); lanes 9-16: DMPC/DPC mixed micelles (1:4 molar ratio). FT, flowthrough; DE, detergent exchange; W, wash. Load refers to sample injected onto the column.

The results of the small scale screening are shown in Figure 2.3, in which Lanes 2-10 correspond to MHV-E in the indicated membrane mimetic eluted by thrombin cleavage, as in lanes 6 and 14 of Figure 2.2. For MHV-E cleaved in LPPG micelles (lane 8), there is slight co-elution of GFP which is removed in the downstream concentration steps and size exclusion chromatography. Low recovery of MHV-E protein was observed for detergents having non-ionic head groups (Fig. 2.3, Lanes 9-10). Therefore, subsequent NMR screening was focused on mimetics having anionic or zwitterionic head groups.
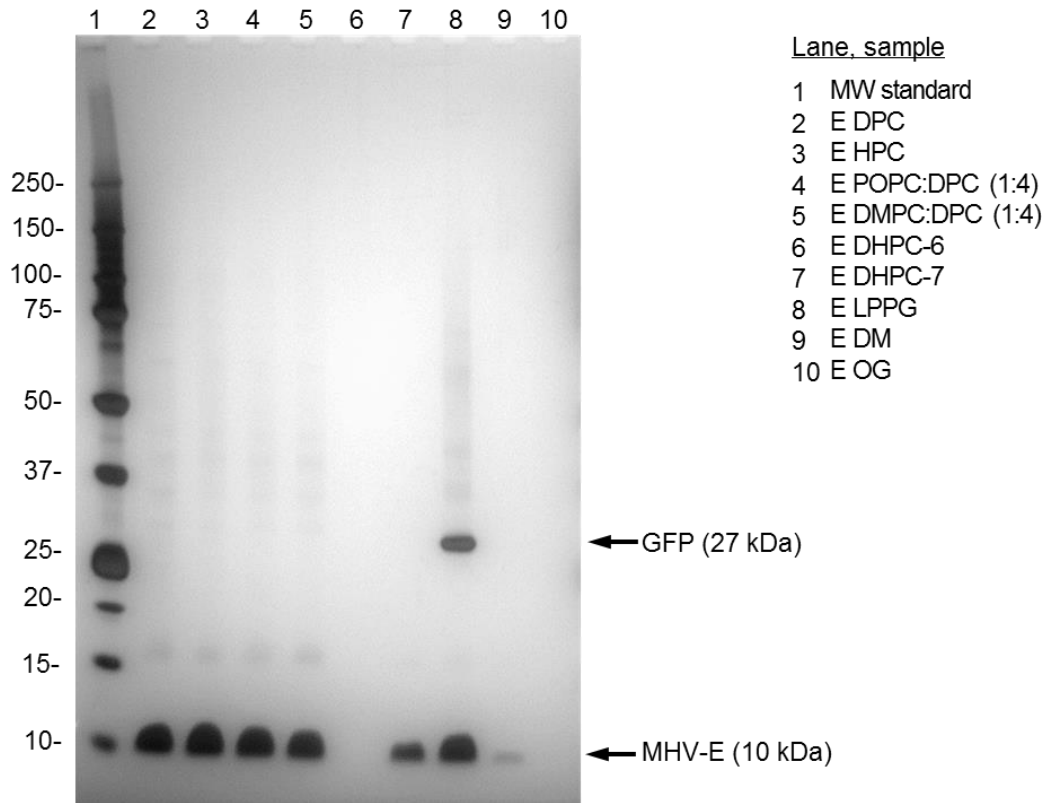
Figure 2.3. SDS-PAGE analysis of membrane mimetic screening. Visualized by silver staining.

3.2. MHV-E is purified on a large scale and cleaved to homogeneity in a single step

A representative chromatogram (Fig. 4) of the scaled-up on-column exchange/cleavage method for the SDS to LMPC exchange followed by thrombin cleavage using a 5 mL nickel affinity column paired to an Amersham ÄKTA FPLC indicates the different steps during the method: flowthrough, exchange, elution of cleaved MHV-E following cleavage of the MHV-E-GFP-His$_6$ fusion by thrombin, and elution of the GFP-His$_6$ with imidazole. The yield of purified, cleaved MHV-E was ~20 mg per L of culture in M9 minimal medium, which is enough material for structural studies. Size exclusion

chromatography of MHV-E in LMPC micelles indicated homogeneity (Fig. 5).

Additional chromatograms are found in Fig. S2.3.
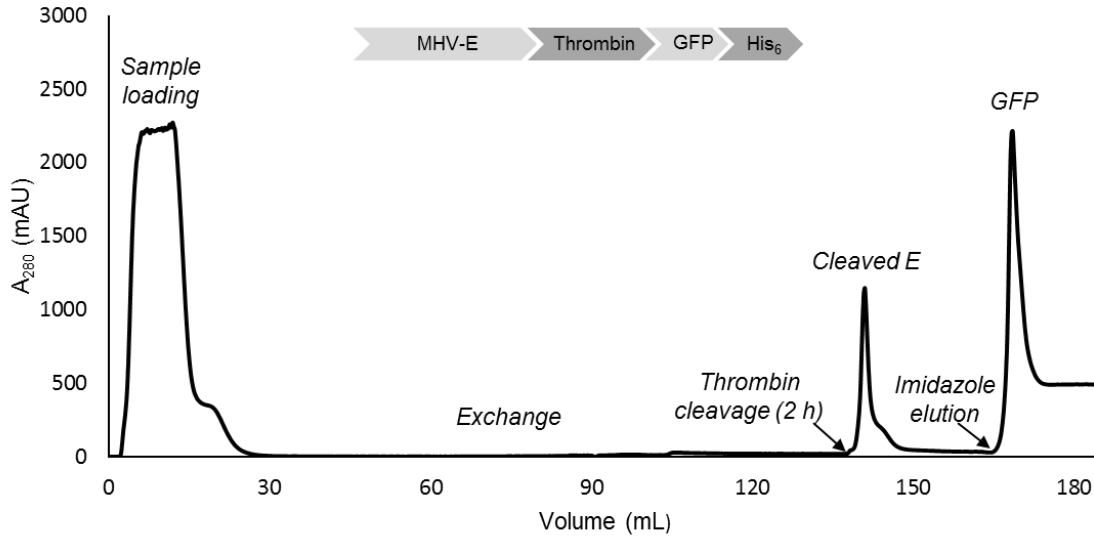
Figure 2.4. On-column detergent exchange/cleavage of MHV-E-GFP-His$_6$ fusion with thrombin cleavage site.
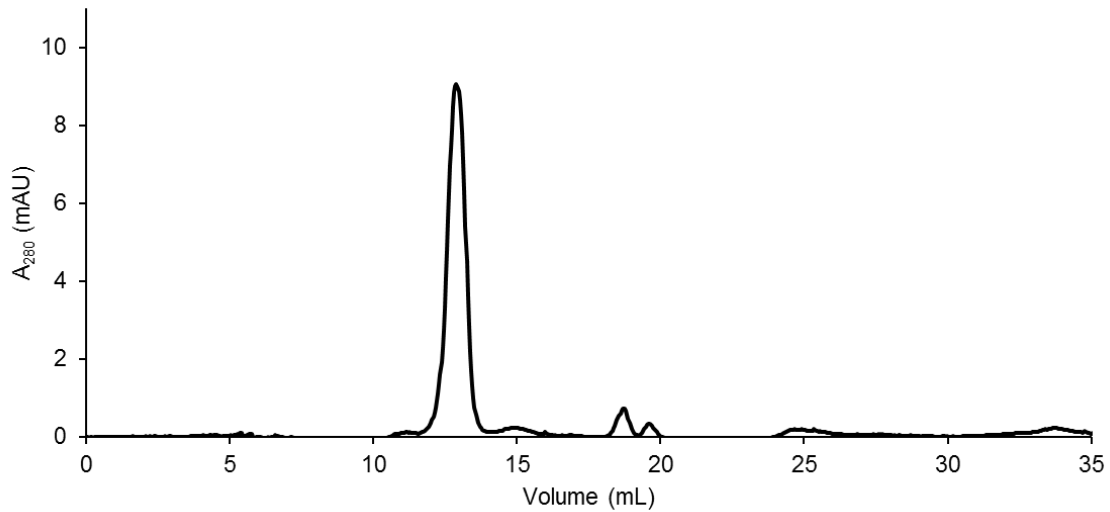
Figure 2.5. Size exclusion chromatography of MHV-E in LMPC micelles. Molecular weight standards were run on this column (data not shown).

3.3. CD spectroscopy verifies α-helical content

CD spectroscopy was used to determine global secondary structure content of MHV-E in the membrane mimetics indicated in Table 1. Selected CD spectra are shown in Fig. 6 (see Fig. S2.4 for individual spectra), and a representative deconvolution is given in Table 2, indicating that MHV-E is predominantly α-helical. The α-helical content is maintained around 60% in the various membrane mimetics as seen in Table 2, however SDS was observed to have slightly lower α-helical content suggesting minor unfolding at the ends of the helix which are restored following exchange into a different membrane mimetic.
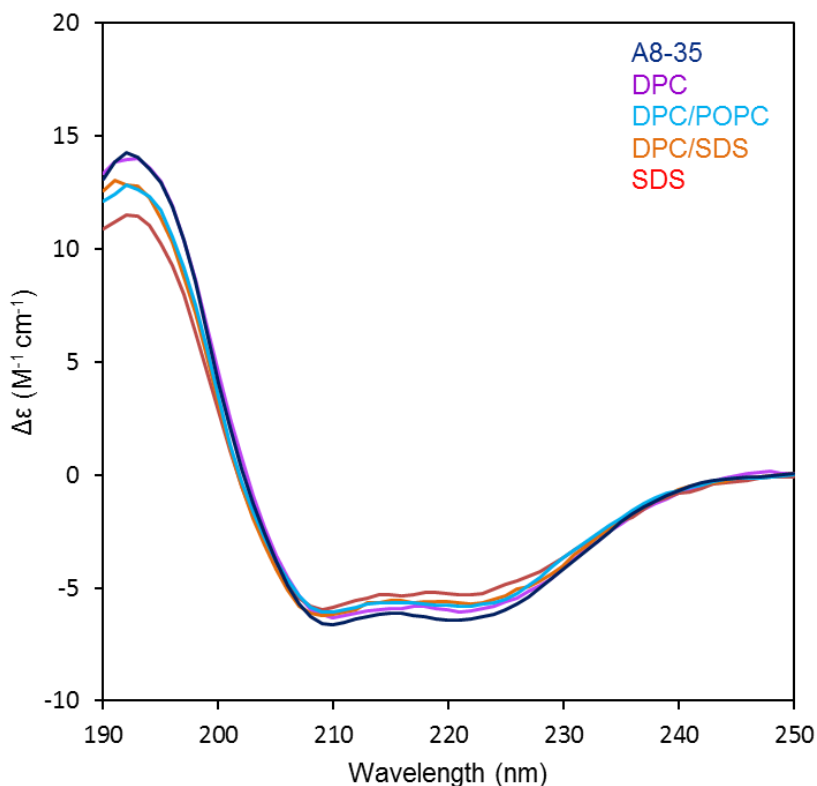
Figure 2.6. CD spectra of MHV-E in specified membrane mimetics. (Individual spectra can be found in Fig. S4). Deconvolutions are shown in Table 2.2.

Table 2.2. Secondary structure deconvolution of CD spectra

Membrane Mimetic	α-helix (%)	β-sheet (%)	Turns (%)	Unordered (%)	RMSD
A8-35	64.9	2.5	9.1	23.5	0.011
DPC	61.2	4.6	9.1	25.2	0.019
DPC/POPC (4:1)	58.8	5.6	9.5	26.1	0.024
DPC/SDS (5:1)	59.7	3.5	9.8	27.1	0.029
SDS	55.0	6.5	10.1	28.5	0.033

3.4. Ion channel activity

Black Lipid Membrane (BLM) electrophysiology measurements of MHV-E in a POPC/POPE/POPG bilayer showed viroporin activity under an asymmetric Na^+ gradient (Fig. 7). A voltage sweep (Fig. 7A) was carried out and the corresponding current-voltage I-V plot (Fig. 7B) gives a reversal potential close to the theoretical equilibrium potential for Na^+, indicating cation selectivity that was previously reported (18).
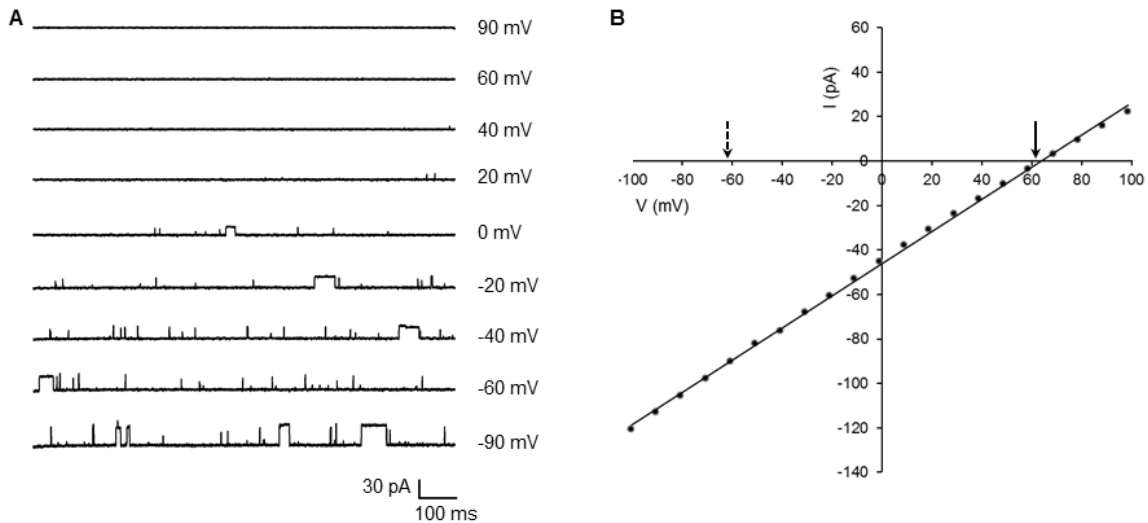
Figure 2.7. Viroporin activity of MHV-E. (A) Voltage sweep showing ion channel activity under a 10-fold gradient (cis-trans) of NaCl. (B) Current-voltage relationship under an asymmetric NaCl gradient. The solid and broken arrows indicate the theoretical equilibrium potentials for Na^+ and Cl^- respectively. Ion channel measurements were

performed by Jonathan Carillo of the Brenda Hogue Lab using the BLM electrophysiology setup in the Goryll Lab located in the ASU Engineering department.

3.5. ^{1}H-^{15}N-TROSY-HSQC screening

Compatible membrane mimetics identified from the screen were scaled up for an initial

TROSY-HSQC screen (Fig. 8). Criteria used to assess TROSY spectra were:

- Number of resonances observed versus expected (peak count)

- Dispersion

- S/N; dynamic range

- Line widths

- Resolution

81 out of 83 expected peaks are observed for MHV-E in SDS micelles (Fig. 2.8A),

having a dispersion consistent with an α-helical protein. Peaks are the most uniform in

linewidths and intensities, and are well resolved with only a few regions showing minor

overlap. For the TROSY-HSQC spectrum of MHV-E in DPC micelles (Fig. 2.8B), a

significant number of peaks are missing and the dispersion of the peaks is narrower, not

extending beyond 8.5 ppm. Two peaks are observed for the indole ^{1}H-^{15}N resonance of

Trp11 indicating conformational heterogeneity. Peak intensities show much greater

variation and line widths are broader in comparison to Fig. 2.7A. Overall, spectral

resolution is lower especially in the central region where many peaks exhibit overlap. The

TROSY-HSQC spectrum of MHV-E in DPC/SDS (5:1) mixed micelles (Fig. 2.8C) is

intermediate, in terms of the criteria, between MHV-E in SDS micelles (Fig. 2.8A) and

MHV-E in DPC micelles (Fig. 2.8B). MHV-E in bicelles (DHPC/POPC/POPG; Fig.

2.8D) is missing a large number of peaks, including the two expected glycines. The

observed peaks show narrow dispersion and broad line widths.
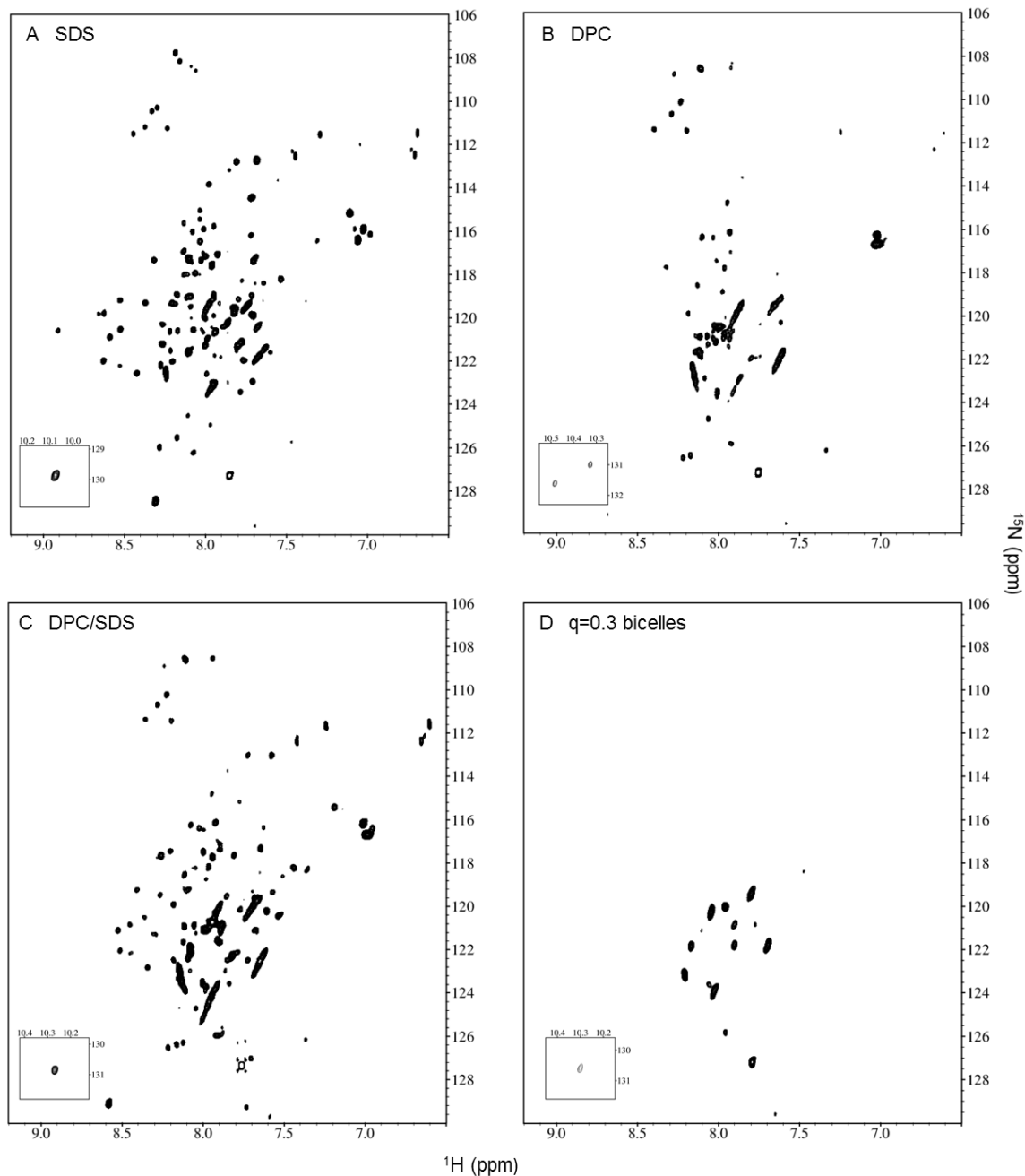
Figure 2.8. Initial TROSY screen of MHV-E. ^{1}H-^{15}N TROSY-HSQC spectra of 0.8 mM MHV-E in 100 mM of each membrane mimetic: (A) SDS, (B) DPC, (C) DPC/SDS (5:1), (D) DMPC/POPC/POPG (50:9:1). This initial round of screening was done at 25 °C.

52

MHV-E in DPC micelles (Fig. 2.8B) shows two peaks for the Trp11 indole N-H, indicating multiple conformations. Addition of lipid in DPC/POPC mixed micelles gave a single peak for Trp11 indole N-H (Fig. 2.9A), indicating that the lipid eliminates conformational heterogeneity found in DPC micelles. Reconstitution of MHV-E into amphipols was done starting from DPC/POPC mixed micelles, and the corresponding TROSY-HSQC spectrum (Fig. 2.9B) shows many peaks are missing. The overlay of the TROSY spectra before and after amphipol reconstitution (Fig. 2.9C) illustrates the pronounced effect that the membrane mimetic can have.
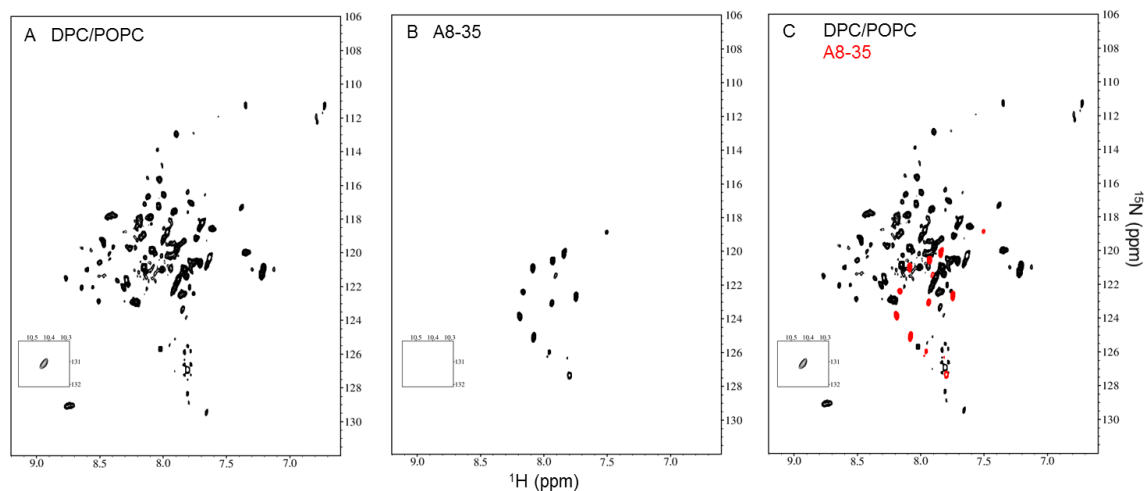
Figure 2.9. Amphipol reconstitution of MHV-E from lipid/detergent mixed micelles. ^{1}H-^{15}N TROSY-HSQC spectra of MHV-E in **(A)** 80 mM DPC/20 mM POPC lipid/detergent mixed micelles and **(B)** A8-35 amphipol. **(C)** is the overlay of (A) and (B).

Lysophospholipid micelles (refer to Fig. 2.10A for LPPG, Fig. 2.10B for LMPG, and Fig. 2.10C for LMPC) yielded the highest quality spectra. LMPC micelles gave the sharpest line widths and highest resolution TROSY spectrum, with 80 out of 83 backbone amide resonances observed (Fig. 2.10C). The effects of lysophospholipid hydrocarbon tail length and head group charge on spectral quality are shown in Fig. 2.11 (refer to Fig. S5

53

for structures of lysophospholipids). In Fig. 11A, LPPG and LMPG have 16 and 14 carbon tails, respectively, while both possess the anionic phosphoglycerol head group. The length of the tail had very little effect on resulting TROSY spectra as seen by nearly identical overlap suggesting a similar conformation is adopted (Fig. 2.11A). In contrast, the head group charge had a large effect on the resulting TROSY spectra (Fig. 2.11B) when comparing the anionic phosphoglycerol head group of LMPG to the zwitterionic head group of LMPC, both of which contain the 14 carbon myristoyl tail. Many of the overlapping resonances in the central region of the spectrum became resolved with the zwitterionic charge of the phosphocholine headgroup (Fig. 2.11C).
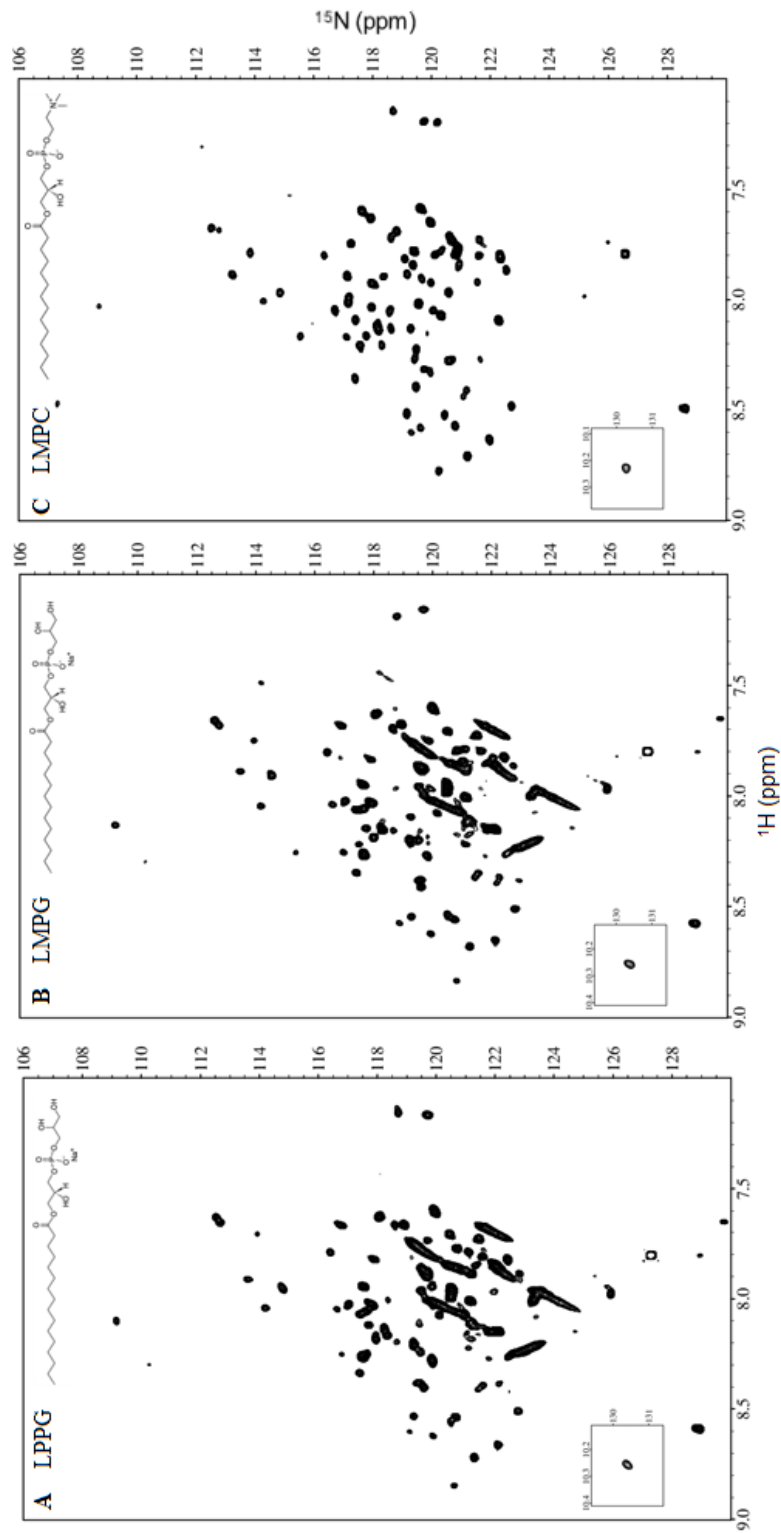
Figure 2.10. ^1H-^{15}N TROSY-HSQC spectra of MHV-E in lysophospholipid micelles. 1 mM MHV-E in (A) 100 mM LPPG, (B) 100 mM LMPG, (C) 100 mM LMPC.
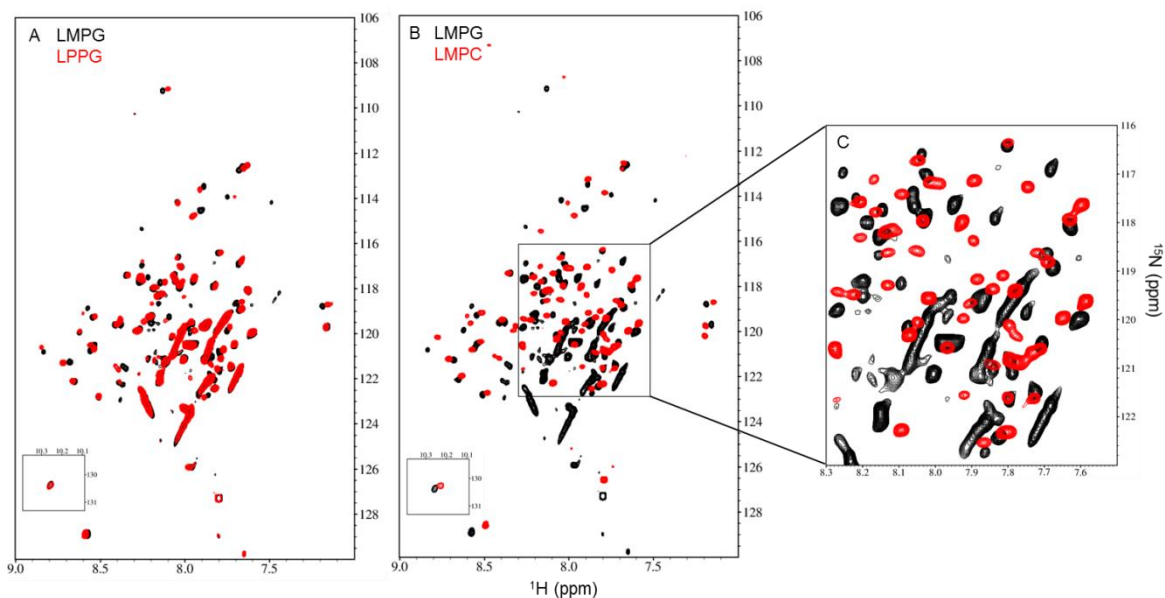
Figure 2.11. Effects of lysophospholipid hydrocarbon tail length **(A)** and head group charge **(B)** on spectral quality. 1 mM MHV-E in **(A)** 100 mM LMPG (black), 100 mM LPPG (red) and **(B)** 100 mM LMPG (black), 100 mM LMPC (red).

3.6. Optimization of spectral quality and stability of MHV-E in LMPC micelles

Fine screening in LMPC micelles was done at pH 5.5, 6, and 6.5 (Fig. 2.12). At pH 5.5, the amide exchange rate is slower which results in higher signal to noise (S/N) ratio. Additionally, certain regions, indicated by the asterisks in Fig. 2.12, show better resolution at pH 5.5 (Fig. 2.12).
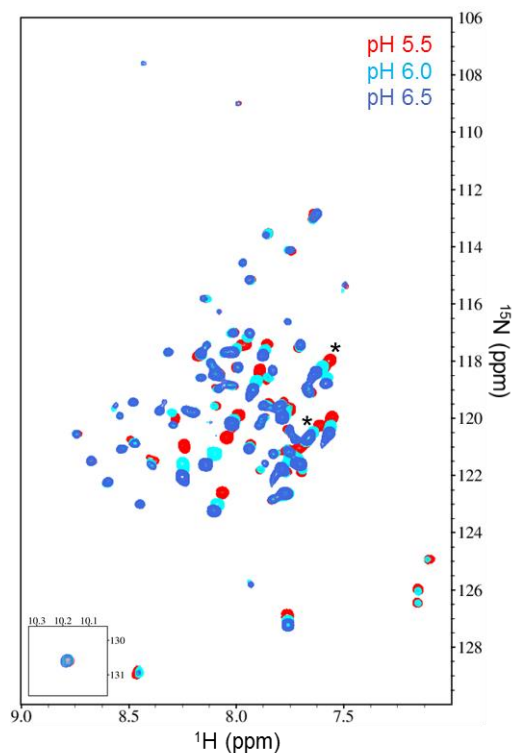
Figure 2.12. Optimization of spectral quality in LMPC micelles. pH: pH 5.5 (red), pH 6 (cyan), pH 6.5 (blue). The two asterisks indicate regions that became better resolved at lower pH.

In order to assess stability, samples with protein:LMPC molar ratios of 1:100 (Fig. 2.13A) and 1:200 (Fig. 2.13B) were held at 50 °C for 1 week and TROSY-HSQC spectra were collected. Reduction in signal intensity was observed for several peaks at a 1:100 protein:LMPC ratio, most notably those of the TM region (Fig. 2.13A; refer to Sec. 3.7 for identification of the TM residues). The reduction in signal intensity is largely mitigated at the lower protein:LMPC ratio. Fig. S6 shows the TROSY-HNCO spectrum collected under the optimized conditions, which exhibits high S/N and is well resolved in the carbon dimension.
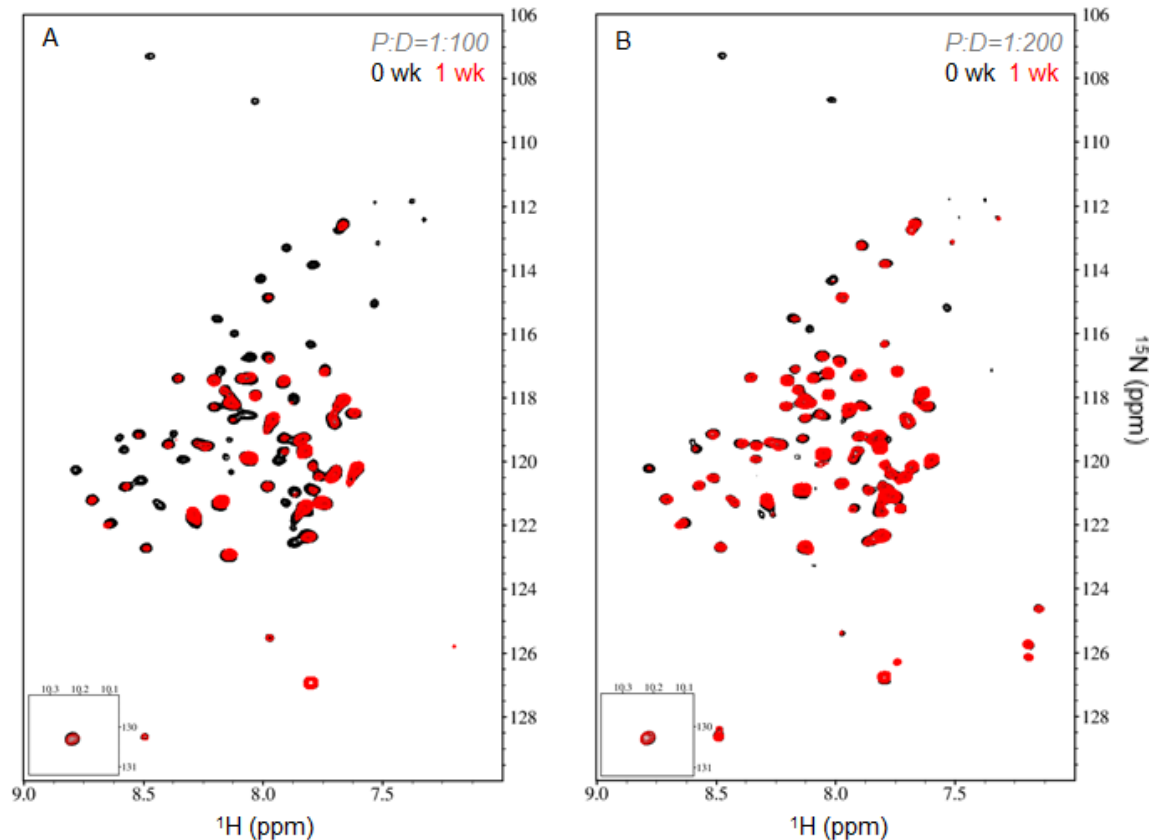
Figure 2.13. Stability of MHV-E in LMPC micelles. MHV-E:LMPC ratios of (A) 1:100 and (B) 1:200 assessed by TROSY spectra collected before (black), and after 1 week incubation at 50 °C (red). P:D, protein-to-detergent ratio, in this case denotes MHV-E:LMPC

3.7. Deuterium exchange identifies transmembrane residues

MHV-E is predicted from the TMHMM algorithm (Krogh et al. 2001) to have one transmembrane helix with 23 residues spanning from Gln15 to Ile37. Figure 2.14A shows the TROSY spectrum of MHV-E·LMPC in 95% H_2O:5% D_2O. Following exchange of MHV-E·LMPC into 95% D_2O:5% H_2O buffer (Fig. 2.14B), 19 of the 23 predicted TM residues were observed to be protected from deuterium exchange corresponding to the micelle embedded TM residues (Fig. 2.14B), providing adequate coverage for structure and dynamics studies. Figure 2.14C shows an overlay of the two spectra identifying the

corresponding transmembrane residues, which will be used to aid in resonance assignments.
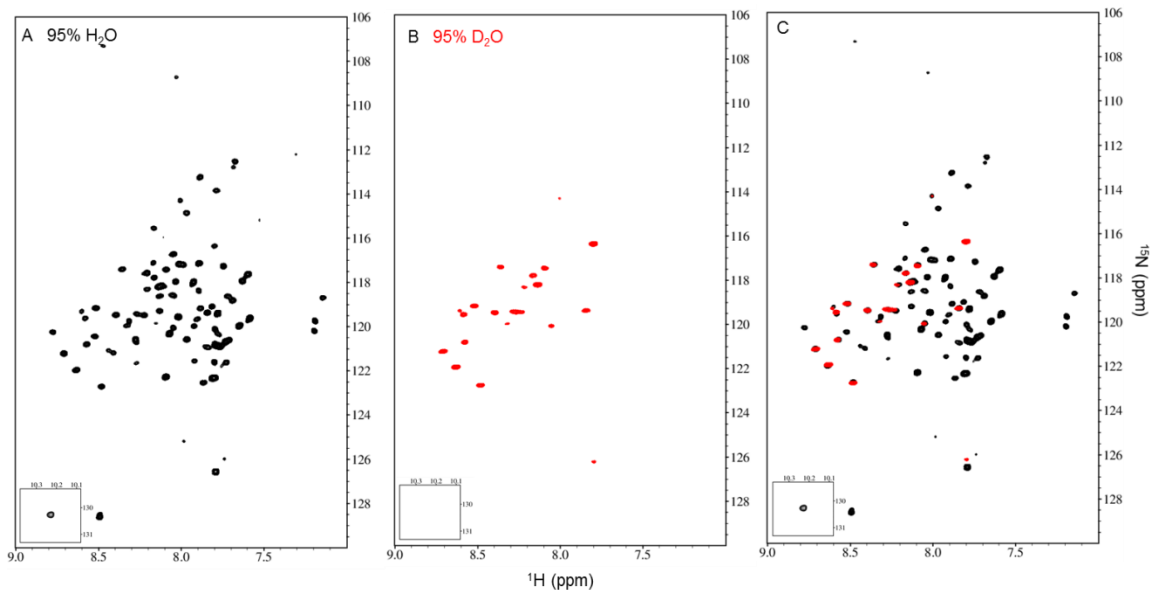
Figure 2.14. Determination of transmembrane residues from deuterium exchange. ^{1}H-^{15}N TROSY-HSQC spectra of 1 mM MHV-E in 100 mM LMPC in (A) 95% H$_2$O, (B) 95% D$_2$O and (C) overlay. Spectra were collected at 45 °C.

4. Discussion/Conclusions

4.1. Importance of the membrane mimetic

The choice of membrane mimetic is critical for structural determination and functional studies of membrane proteins. The membrane mimetic should be selected to permit resonance assignments and be conducive to future screening of drug targets of MHV-E such as inhibitors and probing for interactions with viral and cellular proteins, the latter encompassing proteins which interact with the C-terminal PDZ binding domain. Bicelles are a closer approximation to a lipid bilayer than micellar systems; however, a large number of residues were missing (Fig. 2.8D) very likely due to the larger molecular weight of the bicelles. Observation of only the dynamic terminal residues is similarly

59

observed in p7 from Hepatitis C virus and Vpu from the Human immunodeficiency virus 1 at higher q values (Cook et al. 2011), where q is defined as the ratio of lipid to detergent (q values of ~0.2-0.4 are routinely used in solution state NMR). As the ratio of lipid to detergent (q) is increased in a q-titration (Son et al. 2012), only the mobile loop and terminal regions are observed as the planar bilayer region increases, which correspondingly increases the rotational correlation time, the effect of which is most pronounced locally for the transmembrane resonances which vanish at the higher q values. The high affinity of amphipols to membrane proteins may restrict local dynamics of the TM region and, similar to bicelles, only the dynamic terminal residues are observed (Fig. 2.9B). Additionally, amphipols require slightly basic conditions which results in unfavorable amide exchange rates, further deteriorating the signal to noise (Zoonens et al. 2005).

The results of the small-scale screen suggest that membrane mimetics having a non-ionic head group are incompatible with MHV-E. The α-helical content was found to be maintained in several membrane mimetics, in close agreement with the predicted value from the TMHMM prediction algorithm (Krogh et al. 2001) of around 53% with a single transmembrane helix. The TROSY screen identified lysophospholipid micelles as giving the highest quality spectra (Fig. 9). This is not unexpected given that the lipid composition of the ERGIC/Golgi membranes are predominantly PC lipids. Deuterium exchange of MHV-E in LMPC micelles identified ~19 residues corresponding to the TM helix.

4.2. Quaternary structure depends on choice of membrane mimetic

MHV-E is largely monomeric in SDS micelles (Fig. S1.7). The lower molecular weight of the MHV-E•SDS complex results in narrow line widths (Fig. 2.8A) and α helical content was maintained in SDS (Fig. 2.6, Table 2.2). However, it would be very difficult to use SDS as a model system for binding studies because of the propensity of SDS to denature proteins. Also the fact that the protein is α helical does not mean that all parts are properly folded, as it has been shown before that while the transmembrane helices of membrane proteins can be maintained in SDS the membrane extrinsic parts of the proteins may be disordered (Dutta et al. 2010) therefore it was necessary to select a membrane mimetic that maintains the native fold of MHV-E. For DPC, broader line widths were observed (Fig. 2.8B). CoV E proteins have been predicted to form pentamers in the host cell lipid membranes (Ref). We therefore conducted cross-linking experiments of MHV-E in both SDS and DPC micelles. The results indicate that multiple oligomeric forms may coexist as multiple bands are observed from tetramer to octamer (Supplementary Fig. S2.7). Such heterogeneity in the quaternary structure (distribution of oligomers) may make the channel structure problematic to solve, as further work has to be done to stabilize the oligomeric form of the viroporin.
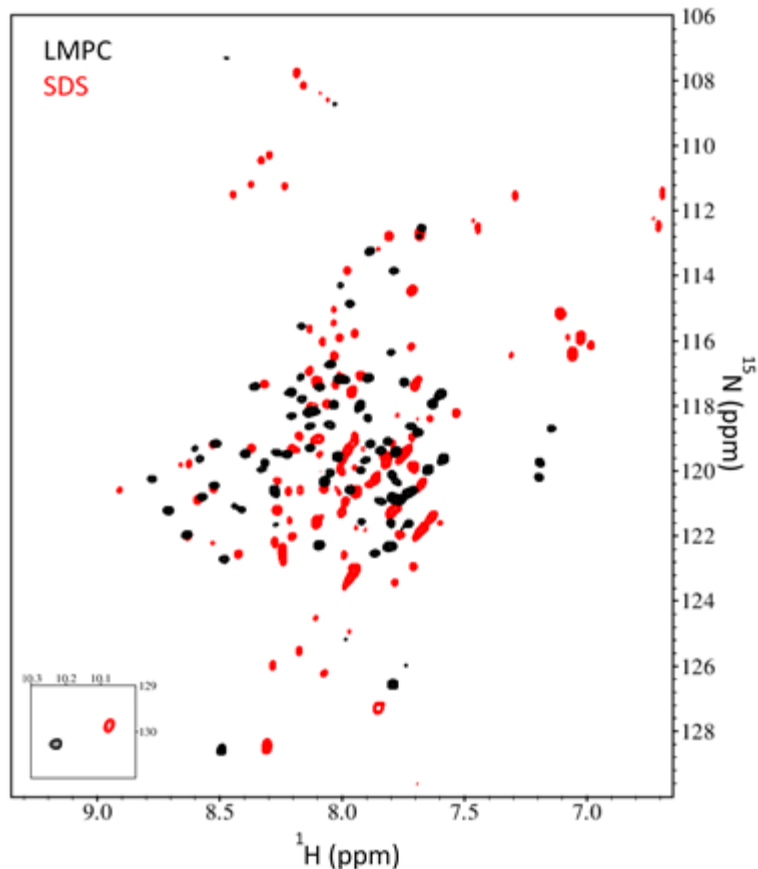
Figure 2.15. Overlay of MHV-E in SDS and LMPC micelles. 1 mM E in 100 mM LMPC (black) at pH 5.5 and 45 °C, and 100 mM SDS (red) at pH 6.5 and 25 °C.

4.3. Conclusions

We explored the influence of different membrane mimics for the NMR analysis of the MHV-E protein. Lysophospholipids gave the highest NMR spectral quality for the full-length envelope protein from mouse hepatitis coronavirus strain A59. Several high-resolution NMR structures of membrane proteins have been solved in lysophospholipid micelles (Table 2.1), which are closely related to lipids by removal of one of the acyl chains. MHV-E in LMPC micelles adopted a different conformation or differences in dynamics than in SDS micelles following exchange (Fig. 2.15). This work optimized the

spectral quality of MHV-E in LMPC micelles for future high resolution solution NMR structure determination with resonance assignments in progress.

Author contributions

BB, JC and JZ have developed the expression and isolation of the MHV-E protein and performed single channel conductivity experiments with MG. FC and DH have performed the cloning and vector construction. DH, BH and PF planned and supervised the expression and isolation optimization. SD was involved in the early stage of the project and supported the initial expression and purification efforts. All the membrane mimic exchange experiments were performed by BB. NMR data were collected and evaluated by BB and BC with support and guidance of from WV and XW. BH and PF conceived the experiments. BB, PF, and DH wrote the paper.

SUPPLEMENTARY FIGURES

A

```
LOCUS        pRSET-MHV-E-GFP\wt        3833 bp    DNA    circular    13-NOV-2019
DEFINITION
SOURCE
  ORGANISM
COMMENT
FEATURES           Location/Qualifiers
    promoter       20..39
                   /vntifkey="30"
                   /label=T7\promoter
    rep_origin     1392..1847
                   /vntifkey="33"
                   /label=f1\ori
    CDS            1978..2838
                   /vntifkey="4"
                   /label=AMP-R
    promoter       1879..1977
                   /vntifkey="30"
                   /label=bla\pr
    rep_origin     2983..3656
                   /vntifkey="33"
                   /label=pUC\ori
    RBS            85..92
                   /vntifkey="32"
                   /label=RBS
    CDS            403..1119
                   /vntifkey="4"
                   /label=folding\reporter\GFP
    terminator     1192..1321
                   /vntifkey="43"
                   /label=T7\term
    CDS            367..402
                   /vntifkey="4"
                   /label=GFP\linker
    CDS            1120..1137
                   /vntifkey="4"
                   /label=His6-tag
    CDS            100..366
                   /vntifkey="4"
                   /label=MHV\E\protein
    3'clip         3678..1341
                   /vntifkey="49"
                   /label=SEQUENCE\VERIFIED\(DTH750\FC167\FC168)
    CDS            100..1137
                   /vntifkey="4"
                   /label=EXPRESSED\PROTEIN
BASE COUNT       999 a       903 c       913 g       1018 t
ORIGIN
        1 gatctcgatc ccgcgaaatt aatacgactc actatagggg gaccacaacg gtttccctct
       61 agaaataatt ttgtttaact ttaagaagga gatatacata tgtttaattt attccttaca
      121 gacacagtat ggtatgtggg gcagattatt tttatattcg cagtgtgttt gatggtcacc
      181 ataattgtgg ttgccttcct tgcgtctatc aaactttgta ttcaactttg cggtttatgt
      241 aatactttgg tgctgtcccc ttctatttat ttgtatgata ggagtaagca gctttataag
      301 tattataatg aagaaatgag actgcccta ttagaggtgg atgatatcct ggtgccacgc
      361 ggctcgggat cagcaggttc cgctgctggt tctggcgaat tcatggctag caaaggagaa
      421 gaacttttca ctggagttgt cccaattctt gttgaattag atggtgatgt taatgggcac
      481 aaattttctg tcagtggaga gggtgaaggt gatgctacat acggaaagtt gacccttaaa
      541 tttatttgca ctactggaaa actacctgtt ccatggccaa cacttgtcac tactctgact
      601 tatggtgttc aatgcttttc ccgttatccg gatcacatga aacggcatga cttttttcaag
      661 agtgccatgc ccgaaggtta tgtacaggaa cgcactatat ctttcaaaga tgacgggaac
      721 tacaagacgc gtgctgaagt caagtttgaa ggtgatacc ttgttaatcg tatcgagtta
      781 aaaggtattg attttaaaga agatggaaac attctcggac acaaactgga gtacaactat
      841 aactcacaca atgtatacat cacggcagac aaacaaaaga atggaatcaa agctaacttc
      901 aaaattcgcc acaacattga agatggatcg gttcaactag cagaccatta tcaacaaaat
      961 actccaattg gcgatggccc tgtcctttta ccagacaacc attacctgtc gacacaatct
     1021 gcccttttcga aagatcccaa cgaaaagcgt gaccacatgg tccttcttga gtttgtaact
```

(continued)

Figure S2.1. Sequences for plasmid pRSET-MHV-E-GFP wt. (**A**) DNA sequence. (**B**) Protein sequence. (**C**) Plasmid map.

```
      1081 gctgctggga ttacacatgg catggatgag ctctacaaac accaccacca ccaccactaa
      1141 taataaaagg gcgaattcca gcacactggc ggccgttact agtggatccg gctgctaaca
      1201 aagcccgaaa ggaagctgag ttggctgctg ccaccgctga gcaataacta gcataacccc
      1261 ttggggcctc taaacgggtc ttgaggggtt ttttgctgaa aggaggaact atatccggat
      1321 ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa
      1381 tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg
      1441 cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc
      1501 ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcggggc tccctttagg
      1561 gttccgattt agtgctttac ggcacctcga ccccaaaaa cttgattagg gtgatggttc
      1621 acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt
      1681 ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc
      1741 ttttgattta taggatttt tgccgatttc ggcctattgg ttaaaaaatg agctgattta
      1801 acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaatttagg tggcactttt
      1861 cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat
      1921 ccgctcatga caataaccc ctgataaatg cttcaataat attgaaaaag gaagagtatg
      1981 agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt
      2041 tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga
      2101 gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa
      2161 gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt
      2221 attgacgccg gcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt
      2281 gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc
      2341 agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga
      2401 ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat
      2461 cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct
      2521 gtagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc
      2581 cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg
      2641 gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc
      2701 ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg
      2761 acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca
      2821 ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta
      2881 aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc
      2941 aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa
      3001 ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca
      3061 ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta
      3121 actggcttca gcagagcgca gataccaaat actgttcttc tagtgtagcc gtagttaggc
      3181 caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca
      3241 gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta
      3301 ccggataagg cgcagcggtc gggctgaacg ggggttcgt gcacacagcc cagcttggag
      3361 cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt
      3421 cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc
      3481 acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac
      3541 ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac
      3601 gccagcaacg cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc
      3661 tttcctgcgt tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat
      3721 accgctcgcc gcagccgaac gaccgagcgc agcgagtcag tgagcgagga gcggaagag
      3781 cgcccaatac gcaaaccgcc tctccccgcg cgttggccga ttcattaatg cag
//
```

B

```
LOCUS       Translation\of\pRSET-MHV-E-GFP\wt\DTH       346 aa                    13-NOV-2019
DEFINITION  Translation of a fragment of RECONSTRUCTION DTH.
KEYWORDS    TRANSLATED.
SOURCE
  ORGANISM
COMMENT
FEATURES             Location/Qualifiers
     Site            341..346
                     /vntifkey="264"
                     /label=His-tag
     Region          1..83
                     /vntifkey="1000"
                     /label=MHV\A59\E\(ACO72895.1)
     Region          102..340
                     /vntifkey="1000"
                     /label=folding\reporter\GFP\(ABC47319)
     Site            84..89
                     /vntifkey="221"
                     /label=thrombin\cleavage\LVPR*GS
ORIGIN
        1 mfnlfltdtv wyvgqiifif avclmvtiiv vaflasiklc iqlcglcntl vlspsiylyd
       61 rskqlykyyn eemrlpllev ddilvprgsg sagsaagsge fmaskgeelf tgvvpilvel
      121 dgdvnghkfs vsgegegdat ygkltlkfic ttgklpvpwp tlvttltygv qcfsrypdhm
      181 krhdffksam pegyvqerti sfkddgnykt raevkfegdt lvnrielkgi dfkedgnilg
      241 hkleynynsh nvyitadkqk ngikanfkir hniedgsvql adhyqqntpi gdgpvllpdn
      301 hylstqsals kdpnekrdhm vllefvtaag ithgmdelyk hhhhhh
//
```

65

C

Lane, sample
1 MHV-E-GFP-His$_6$ 4 h
2 MHV-E-GFP-His$_6$ 8 h
3 MHV-E-GFP-His$_6$ 12 h
4 MHV-E-GFP-His$_6$ 16 h
5 MHV-E-GFP-His$_6$ 20 h
6 MHV-E-GFP-His$_6$ 24 h
7 MHV-E-GFP-His$_6$ 28 h
8 GFP-His$_6$
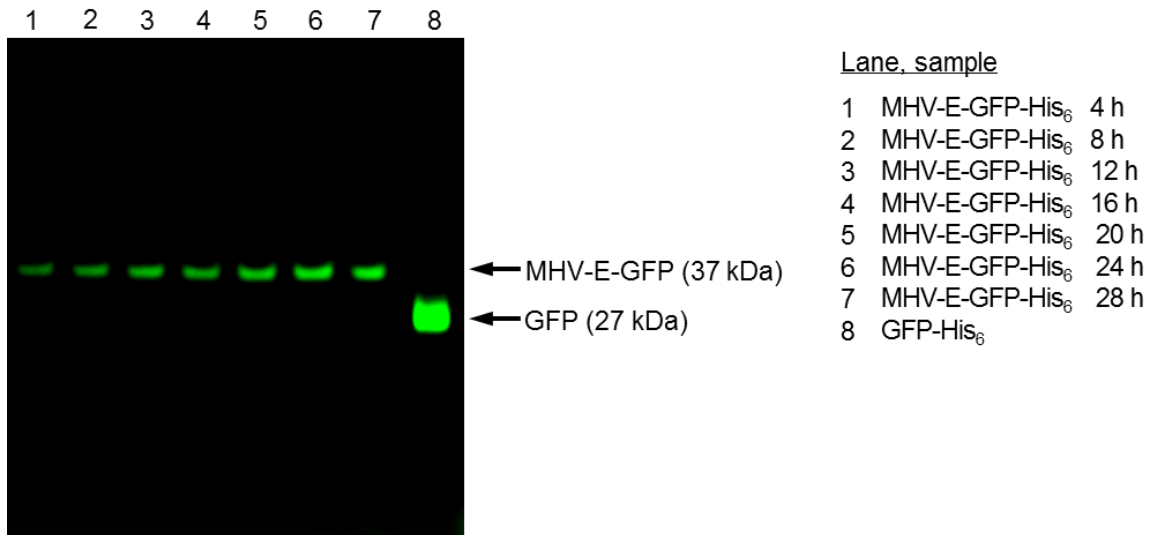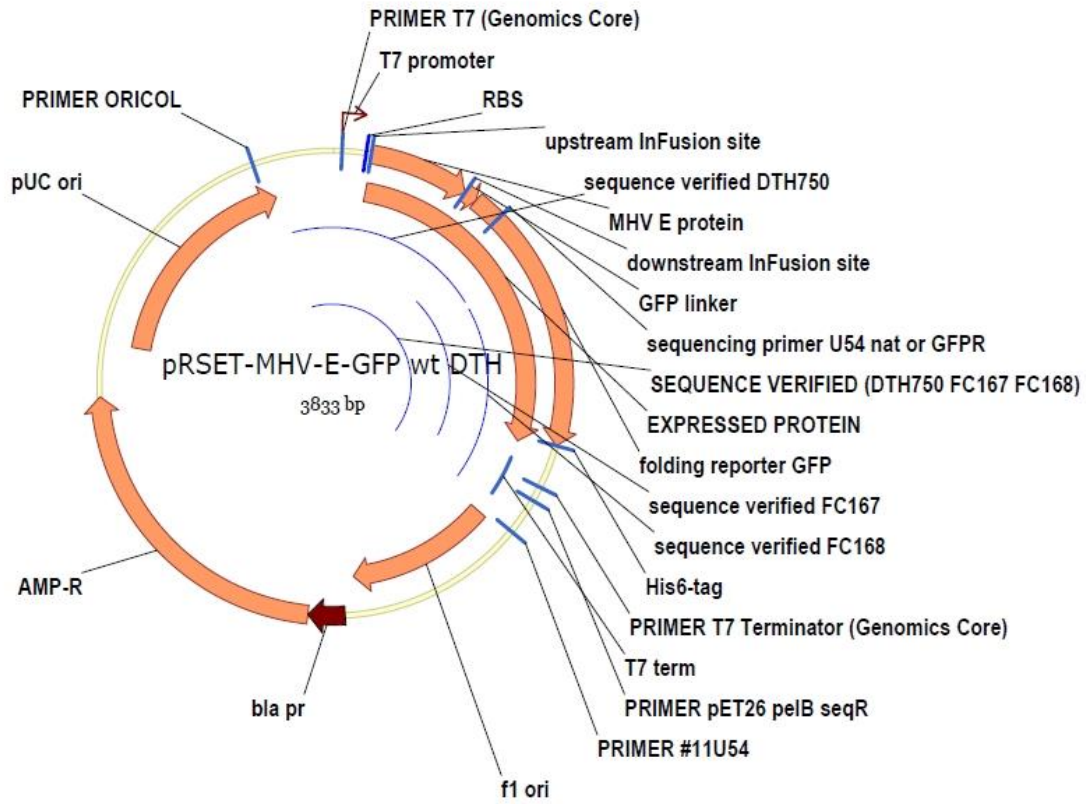
Supplementary Figure S2.2. Time course of GFP-His$_6$-tagged MHV-E expression monitored by in-gel fluorescence.
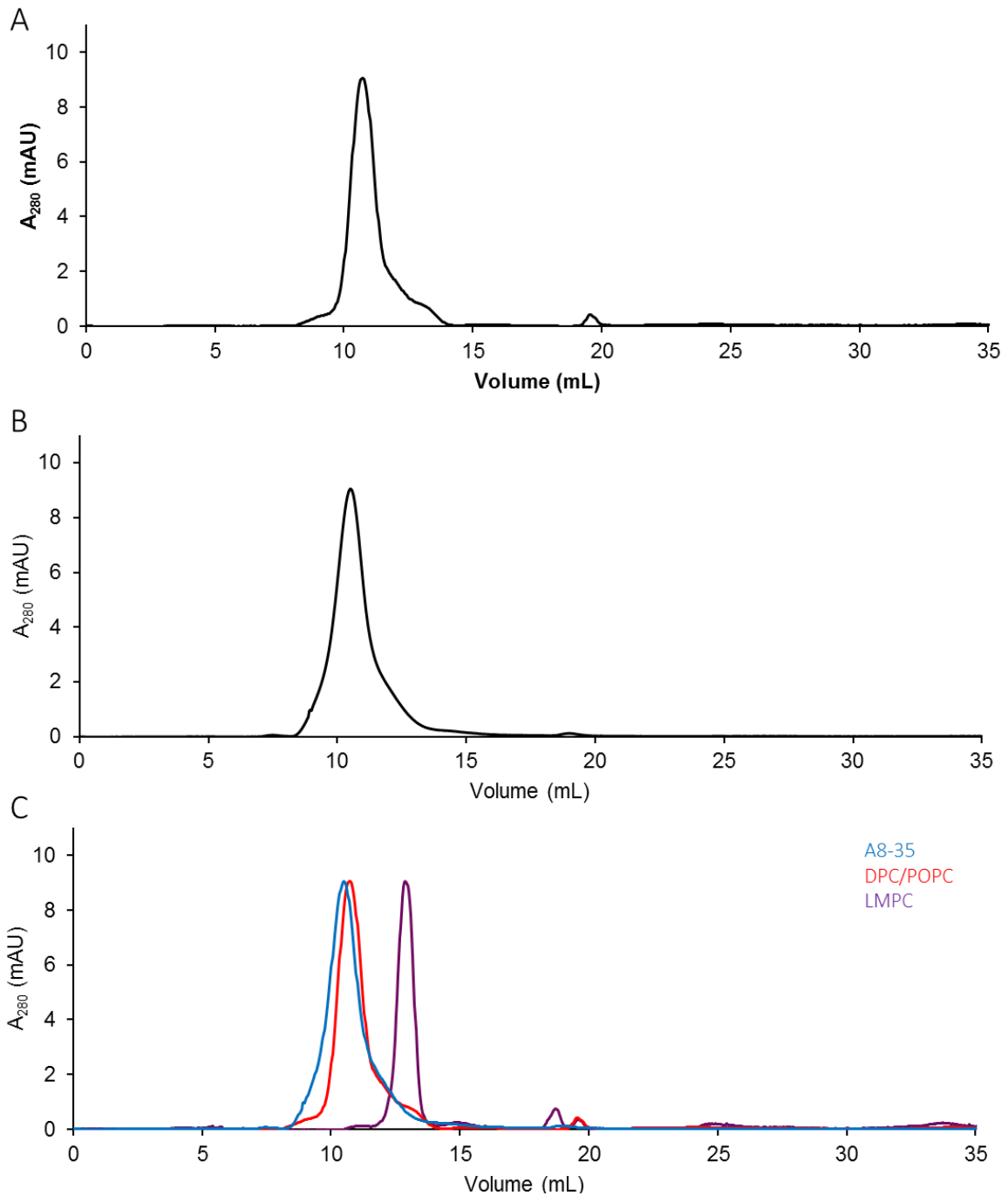
Supplementary Figure S2.3. SEC of MHV-E in **(A)** DPC/POPC and **(B)** A8-35. **(C)** Overlay of **(A)**, **(B)**, and MHV E·LMPC (refer to Fig. 4 for separate chromatogram) for comparison of retention volumes.
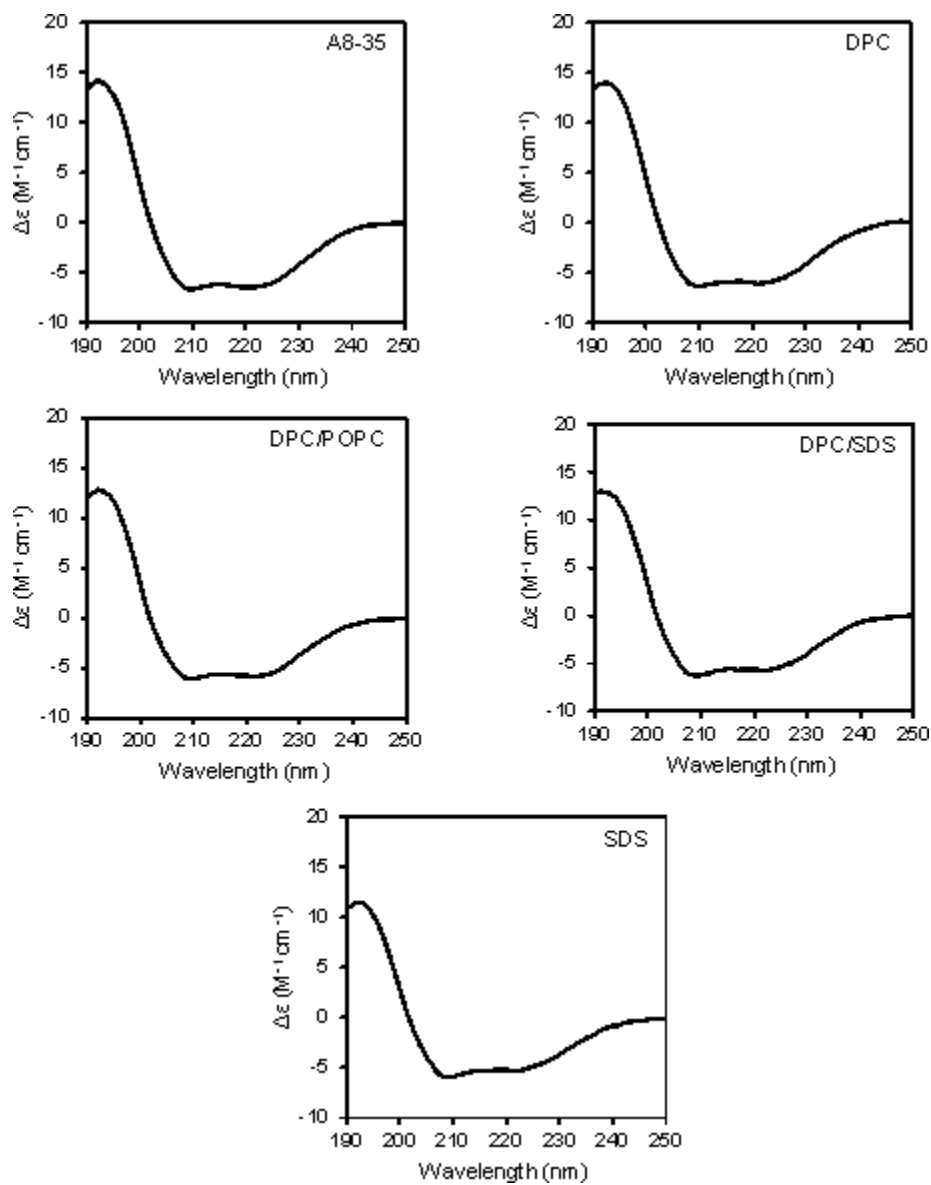
Supplementary Figure S2.4. CD spectra of MHV-E in the specified membrane mimetic. Individual spectra from the overlay in Fig. 2.5.
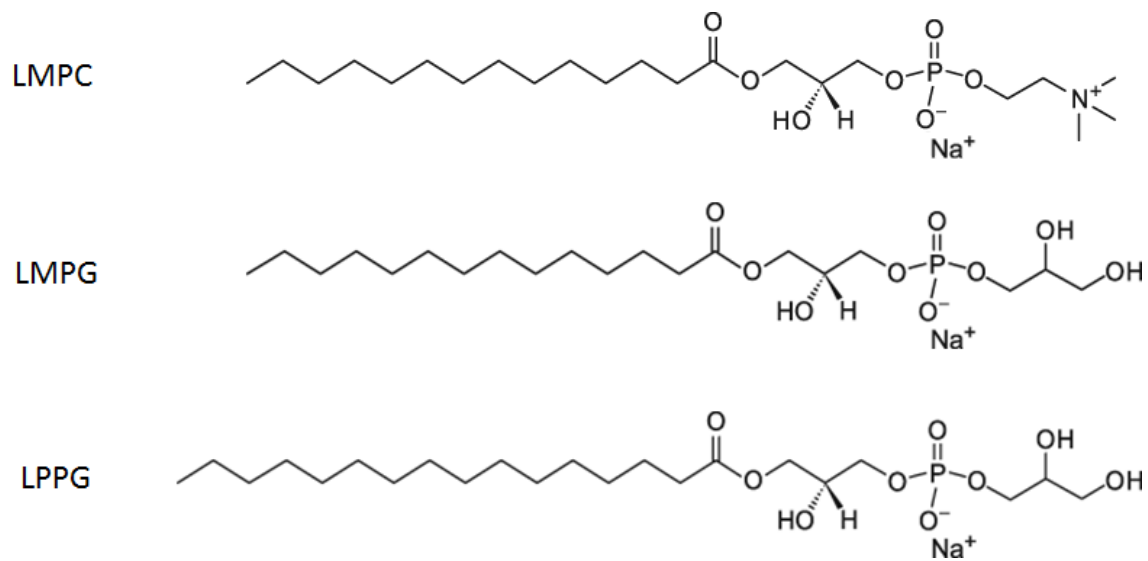
Supplementary Figure S2.5. Structures of lysophospholipids used in the TROSY screen.
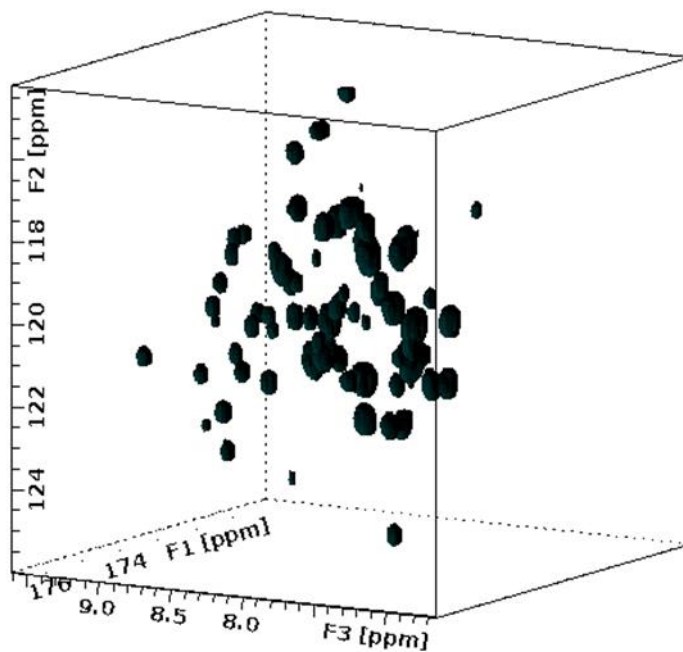
Supplementary Figure S2.6. TROSY-HNCO of MHV-E in LMPC micelles.
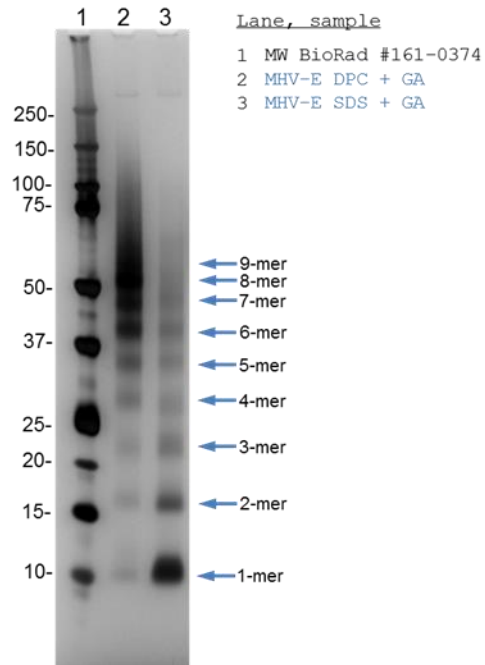
Supplementary Figure S2.7. Assessing oligomeric state of MHV-E using glutaraldehyde (GA) crosslinking. MHV-E in DPC micelles (Lane 2) and SDS micelles (Lane 3).

CHAPTER 3

SOLUTION NMR BACKBONE RESONANCE ASSIGNMENTS AND SECONDARY STRUCTURE OF THE MURINE CORONAVIRUS MHV-A59 ENVELOPE PROTEIN

Bobby Baravati[1,2], , Brian R. Cherry[1,3], Jonathan Carillo[4], James D. Zook[1,2,5], Sasha M. Daskalova[1,2,4,6], Felicia M. Craciunescu[2], Michael Goryll[1,7], Debra T. Hansen[2,8], Brenda G. Hogue[2,3], Wade D. Van Horn[1,9,10], Xu Wang[1], Petra Fromme[1,2]

[1]School of Molecular Sciences, Arizona State University, Arizona State University, Tempe, Arizona, USA; [2]Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe, Arizona, USA; [3]Magnetic Resonance Resource Center, Arizona State University, Tempe, Arizona, USA; [4]Biodesign Center for Immunotherapy, Vaccines and Virotherapy, Arizona State University, Tempe, Arizona, USA; [5]Present address: LGC, Biosearch Technologies, Middleton, Wisconsin, USA; [6]Present address: Biodesign Center for BioEnergetics and School of Molecular Sciences, Arizona State University, Tempe, Arizona, USA; [7]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona, USA; [8]Biodesign Center for Innovations in Medicine, Arizona State University, Tempe, Arizona, USA; [9]Biodesign Virginia G. Piper Center for Personalized Diagnostics, Arizona State University, Tempe, Arizona, USA; [10]Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, Arizona, USA

Abstract

Sequence specific backbone resonance assignments for the murine coronavirus envelope protein (MHV-E) in LMPC micelles have been obtained using TROSY triple resonance experiments and verified with inverse labelling. The assignments have been deposited to the BMRB, which will provide a basis for structure determination and probing interactions of the full-length MHV envelope protein. The full length MHV-E protein encompasses the putative protein interaction motif (the PDZ binding motif) thought to be involved in altering cellular pathways.

Keywords

Declarations

Conflicts of interest/Competing interests

All authors declare that they have no conflicts of interests/competing interests.

Ethics approval

These studies do not require ethics approval

Consent for publication

All authors have given consent for publication

Availability of data and material

The data has been submitted to the BMRB database

Code availability

The data has been evaluated with the following programs:

Strathclyde Electrophysiology Software

(https://pureportal.strath.ac.uk/en/datasets/strathclyde-electrophysiology-software-winwcp-winedr)

nmrPipe (Delaglio et al. 1995)

SMILE (Ying et al. 2017)

NMRFAM-SPARKY (Lee, Tonelli, and Markley 2015)

TALOS-N (Shen and Bax 2013)

Authors' contributions

BB, JC and JZ have developed the expression and isolation of the MHV-E protein and performed single channel conductivity experiments with MG. FC and DH have performed the cloning and vector construction. DH, BH and PF planned and supervised

the expression and isolation optimization. SD was involved in the early stage of the project and supported the initial expression and purification efforts. All the membrane mimic exchange experiments were performed by BB. NMR data were collected and evaluated by BB and BC with support and guidance of from WVH and XW. BH and PF conceived the experiments. BB and PF wrote the paper.

Biological context

Coronaviruses (CoV) are the causative agents of the SARS, MERS and COVID-19 epidemics. SARS was reported in approximately 8000 cases with a 10% mortality rate, and MERS has infected 2494 people leading to death in 34% of cases (Peeri et al. 2020). The recent emergence of the novel coronavirus CoV-2 in Wuhan City, China, named SARS-CoV-2, was reported in 520 000 cases as of 3/26/2020 with an approximate 3.3% mortality rate, which prompted the World Health Organization to designate COVID-19 as a pandemic.

Coronaviruses are prevalent in several vertebrates and can cross species barriers. The recent COVID-19 outbreak has been suspected to be caused by such a cross-species transmission in a food market, even if the original host species has not yet been identified. In this study, here we investigated the envelope protein from the Murine Coronavirus Mouse Hepatitis Virus (MHV) as a model for human coronavirus disease.

The structural proteins comprising the viral envelope of coronaviruses are the spike (S), matrix (M), and envelope (E) proteins. Coronavirus envelope (CoV E) proteins are implicated in various stages of the viral life cycle, and are therefore promising drug targets (Schoeman and Fielding 2019). CoV E proteins have been found to localize to the

74

Golgi apparatus and endoplasmic reticulum Golgi apparatus intermediate compartments

(ERGIC) membranes of the infected host cell (Ye and Hogue 2007; Cohen, Lin, and

Machamer 2011), where they form pentameric ion channels, called viroporins, thought to

be involved in budding and scission of virions (Nieva, Madan, and Carrasco 2012)

(OuYang and Chou 2014). Electrophysiology measurements have shown that the small

molecule hexamethylene amiloride (HMA) inhibits ion channel activity in the CoV E

viroporins of the β-coronaviruses (Wilson, Gage, and Ewart 2006a). Additionally, CoV E

proteins are known to modulate cellular pathways through interactions with host cell

proteins. In the case of SARS-E, identification of a PDZ binding domain in the last four

amino acids of the C-terminus has been shown to bind PALS1, a tight junction-associated

protein, and alter tight junction formation and epithelial morphogenesis (Teoh et al. 2010;

Javier and Rice 2011). Other interactions with cellular proteins have been proposed such

as binding to or indirect modification of host cell ion channels (Torres et al. 2015). CoV

E proteins interact with the M protein in the formation of the viral envelope and virus like

particles (VLPs), the mechanistic details of which remain unclear (Schoeman and

Fielding 2019). Mutations at the C-terminal extramembrane domain of MHV-E affects

viral assembly and maturation and resulted in aberrations in morphology where some

exhibited an ovoid rather than spherical morphology implying that the E protein induces

membrane curvature in the viral envelope (Fischer et al. 1998). Studies of chimeric

viruses in which the MHV-E protein has been swapped with that of bovine coronavirus

(BCoV), infectious bronchitis virus (IBV), and SARS coronavirus (SARS-CoV) have

shown that viral replication is not significantly affected, suggesting that CoV E proteins

are similar in their structure and function (Kuo, Hurst, and Masters 2007). Recombinant

viruses in which the E protein has been deleted have exhibited reduced virus titers and crippled virus maturation indicating the crucial role of the E protein in viral production and maturation (Kuo and Masters 2003; DeDiego et al. 2007). The varied roles of the E protein in assembly, budding, scission, trafficking and release of virions and altering cellular pathways (as an example, through protein interaction involving the PDZ binding motif at the C-terminal end of MHV-E) is not surprising due to the inherent minimalistic design of viruses in which the limited set of encoded accessory proteins must carry out multiple functions. These multiple functional roles provide independent drug targets/intervention strategies.

Solution NMR structures have been solved for truncations of SARS-E encompassing only the transmembrane domain (Pervushin et al. 2009), as well as residues 8-65 in DPC/SDS mixed micelles (PDB 2MM4) (Li, Surya, et al. 2014) and LMPG micelles (PDB 5X29, BMRB 36049) (Surya, Li, and Torres 2018). There are no experimental structures reported for full length MHV-E or any of the other coronavirus envelope proteins. The assignments reported here can be used in structure determination and probing interactions of full length MHV-E protein (encompassing the putative PDZ binding motif) with cellular and viral proteins that play a role in the infection cycle of coronaviruses, and allow for design of improved inhibitors of viroporin activity,

Methods and experiments

Protein expression and purification

All chemicals were obtained from Sigma unless otherwise indicated. $^{15}NH_4Cl$ and ^{13}C-glucose were obtained from Cambridge Isotope Laboratories.

Full length MHV-E from strain A59 (accession P0C2R0) was cloned into the pRSET-natGFPHis$_6$ vector as a C-terminal GFP fusion with an intervening thrombin cleavage site, and a C-terminal His$_6$ tag (Martin-Garcia et al. 2014). The plasmid was transformed into *Escherichia coli* Lemo21(DE3) cells. Rhamnose was omitted as it was found to have negligible effects on the expression level. Uniformly ^{15}N labeled and ^{13}C/^{15}N double labeled MHV-E-GFP were obtained by growing cells in M9 minimal media, consisting of 8.5 g/L Na$_2$PO$_4$, 3 g/L KH$_2$PO$_4$, 0.5 g/L NaCl, 0.1 mM CaCl$_2$, 2 mM MgSO$_4$, 0.1 mM FeCl$_3$, 1x MEM vitamins (Gibco), 1 g/L ^{15}NH$_4$Cl and 4 g/L ^{12}C- or ^{13}C-glucose, at 37 °C until the OD$_{600}$ reached 0.4, cooled on ice for a few minutes and grown at 25 °C until the OD$_{600}$ reached 0.6, at which point protein expression was induced with 0.4 mM IPTG. After 24 h, cells were harvested by centrifugation at 5,000 x g for 30 min at 4 °C and stored at -80 °C. Cell pellets were resuspended at a 1:10 ratio (grams of wet cell pellet mass to mL of lysis buffer) in lysis buffer composed of 50 mM HEPES, pH 7.5, 300 mM NaCl, Sigma EDTA-free protease inhibitor cocktail (1 tablet/100 mL) and lysed by sonication on ice, followed by centrifugation at 45,000 x g for 1 h at 4 °C. The supernatant was discarded and pellet was washed with lysis buffer, followed by centrifugation at 45,000 x g for 1 h at 4 °C. The supernatant was discarded and the pellet was resuspended in a 1:50 ratio (extraction buffer volume to culture volume) of extraction buffer composed of lysis buffer plus 10 mM imidazole and 2% SDS. Following a 1 h incubation with gentle mixing at room temperature, insoluble particles were removed by centrifugation at 45,000 x g for 1 h at room temperature.

An on-column exchange and cleavage method based on (Hefti et al. 2001) was utilized, in which thrombin cleavage is used to elute the protein instead of the high concentrations

of imidazole that may be destabilizing for certain membrane proteins. Approximately 40 mg of MHV-E-GFP obtained from 0.5 L M9 expression in 10 mL lysis buffer containing 10 mM imidazole and 2% SDS was loaded onto a 5 mL nickel affinity column (HisTrap HP, GE Healthcare) pre-equilibrated with 50 mM HEPES, pH 7.5, 300 mM NaCl, 10 mM imidazole and 0.1% LMPC. The flow rate was 0.5 mL/min during sample loading and 5 mL/min for other steps. Following a 20 column volumes (CV) SDS to LMPC exchange step, cleavage was carried out by incubating the immobilized MHV-E-GFP with 500 units of thrombin (BioPharm) in 0.5 CV running buffer for 2 h at room temperature to give purified, cleaved E protein. Size exclusion chromatography was used for further purification and to exchange to the NMR buffer. The 30 kDa concentrated pooled IMAC eluate was loaded onto a Superdex 200 Increase 10/300 GL column (GE Healthcare) connected to an ÄKTA system pre-equilibrated with 50 mM MES, pH 5.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP and 5% D_2O and run for 1.5 CV at a flow rate of 0.75 mL/min. SDS-PAGE analysis was done using 4-12% acrylamide gels (Novex, Invitrogen), MOPS SDS Running Buffer (Invitrogen) and molecular weight standard (Bio-Rad). All samples were incubated in Bio-Rad XT Sample Buffer (catalog # 161-0791) supplemented with 50 mM TCEP at 37 °C for 1 h prior to loading the gel. Silver staining was done using the silver stain kit (Pierce catalog # 24612), which indicated a purity >95%. NMR samples contained ~1 mM MHV-E in 200 mM LMPC, 50 mM MES, pH 5.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP and 5% D_2O. For the deuterium exchange experiment, a spin desalting column (Zeba, Pierce) pre-equilibrated in 50 mM MES, pH 5.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP, 200 mM LMPC and 95% D_2O was used to exchange the sample buffer.

For verification of assignments, five inversely labeled samples were prepared: [U-^{15}N, ^{14}N Ala, ^{14}N Arg], [U-^{15}N, ^{14}N Ile, ^{14}N Leu, ^{14}N Val], [U-^{15}N, ^{14}N Phe, ^{14}N Tyr], [U-^{15}N, ^{14}N Lys], [U-^{15}N, ^{14}N Met]. The following modifications were made to the [U-^{15}N]/[U-^{13}C,^{15}N] expression protocol: M9 minimal media was supplemented with 1 g/L of the natural abundance amino acid(s) to be unlabeled, and expression time was shortened from 24 hours to 6 h to minimize metabolic scrambling/isotopic dilution (Shortle 1994).

NMR spectroscopy

All experiments were carried out at 318 K on Bruker and Varian spectrometers equipped with cryogenic probes operating at field strengths of 600, 800 and 850 MHz at the Magnetic Resonance Resource Center (MRRC) at Arizona State University. The following experiments were performed: 2D [^{1}H,^{15}N]-TROSY-HSQC (Pervushin et al. 1997), 3D NUS-TROSY-HNCA, 3D NUS-TROSY-HN(CO)CA, 3D TROSY-HNCO, 3D TROSY-HN(CA)CO, 3D TROSY-HNCACB, 3D TROSY-CBCA(CO)NH, 3D [^{1}H,^{1}H]-NOESY-^{15}N-TROSY (Salzmann et al. 1998). BEST type experiments were utilized for ^{1}H,^{15}N-TROSY-HSQC, 3D TROSY-HNCO and 3D TROSY-HN(CA)CO experiments (Schanda, Van Melckebeke, and Brutscher 2006). Spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using NMRFAM-SPARKY (Lee, Tonelli, and Markley 2015). Non-uniformly sampled data (Hyberts, Arthanari, and Wagner 2012) were reconstructed using the SMILE algorithm (Ying et al. 2017).

Assignments and data deposition
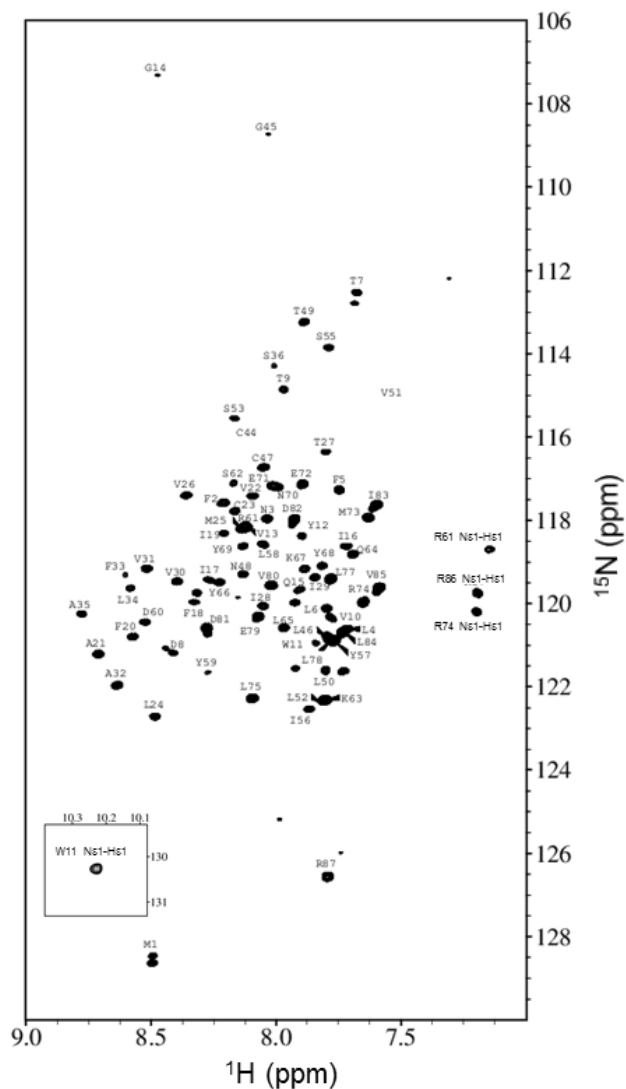
Figure 3.1. MHV-E·LMPC backbone amide assignments. ^{1}H,^{15}N-TROSY-HSQC spectrum of MHV-E in LMPC micelles showing amide assignments.

The ^{1}H,^{15}N-TROSY-HSQC spectrum of MHV-E in 200 mM LMPC micelles exhibits a

modest dispersion consistent with the α-helical structure (Fig. 3.1). Assignments were

obtained using TROSY-type triple resonance experiments. A total of 91% of the

backbone resonances were assigned, with 97% of the transmembrane region assigned

(Fig. 3.2). Assignments could not be obtained for the stretch Ile37-Cys44, which is

80

located at the protein-LMPC micelle interface (residues I17-S36 are micelle embedded as determined from the deuterium exchange experiment). Inverse labeling was used to confirm the assignments (Fig. 3.3a-b). Resonance assignments in the transmembrane region were additionally confirmed using deuterium exchange (Fig. 3.3c). TALOS-N (Shen and Bax 2013) analysis of the chemical shifts indicates MHV-E contains 3 α-helices (Fig. 3.4), with the N-proximal helix encompassing the transmembrane region, a short membrane-proximal middle helix and a C-proximal extramembrane helix. The secondary structure of MHV-E is similar to that reported for SARS-E in LMPG micelles (Surya, Li, and Torres 2018) as determined from TALOS-N. The chemical shifts were referenced relative to DSS for ^{1}H and indirectly for ^{13}C and ^{15}N (Wishart et al. 1995). The ^{1}H, ^{15}N, and ^{13}C chemical shifts of the backbone resonances have been deposited in the BMRB.

Conclusions

There are no experimental structures reported for full length MHV-E or any of the other coronavirus envelope proteins. The secondary structure of MHV-E was found to be similar to SARS E, both possessing a single helical transmembrane domain and two C-terminal extra-membrane domains. The assignments reported here can be used in structure determination and probing interactions of full length MHV-E protein (encompassing the putative PDZ binding motif) with cellular and viral proteins that play a role in the infection cycle of coronaviruses, as well as allow for design of improved inhibitors of viroporin activity.
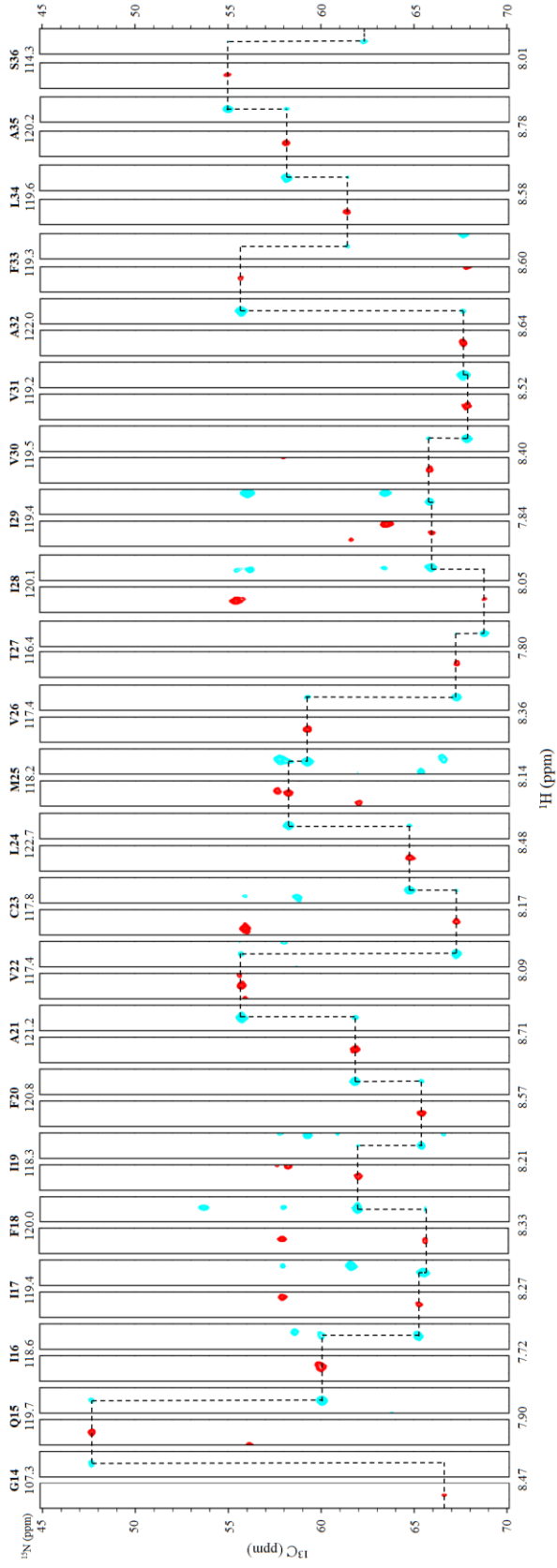
Figure 3.2. Resonance assignments in the transmembrane region. HNCA/HN(CO)CA strip plot showing C$_\alpha$ connectivity.
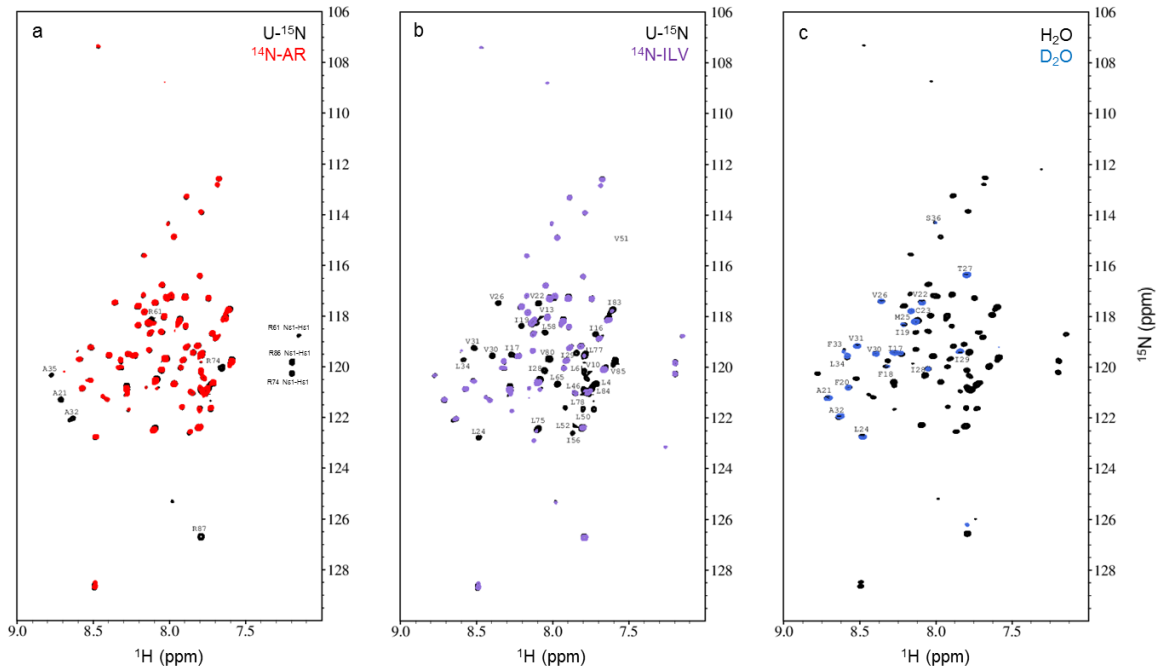
Figure 3.3. Verification of resonance assignments by inverse labeling (**a,b**) and deuterium exchange (**c**).
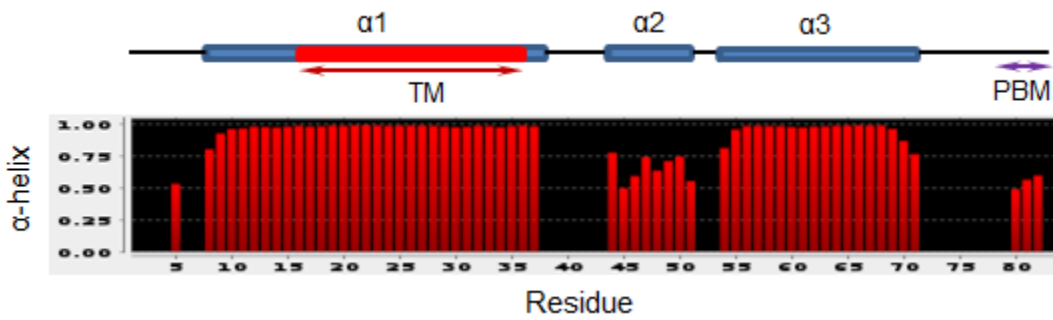
Figure 3.4. Predicted helicity of MHV-E from TALOS-N (Shen and Bax 2013). Abbreviations: TM, transmembrane; PBM, PDZ binding motif.

CHAPTER 4

MAPPING THE BINDING SITE OF THE CHANNEL INHIBITOR HEXAMETHYLENE AMILORIDE TO THE VIROPORIN FORMED BY THE MURINE CORONAVIRUS MHV-A59 ENVELOPE PROTEIN BY AMIDE AND AROMATIC TROSY NMR

Bobby Baravati[1,2], Brian R. Cherry[1,5] , Jonathan Carillo[3,4], James D. Zook[1,2,5], Sasha M. Daskalova[1,2,3,7], Felicia M. Craciunescu[2] , Michael Goryll[1,8], Debra T. Hansen[2,9], Brenda G. Hogue[2,3,4] Wade D. Van Horn[1,10,11], Xu Wang[1], Petra Fromme[1,2]

[1]School of Molecular Sciences, Arizona State University, Arizona State University, Tempe, Arizona, USA; [2]Biodesign Center for Applied Structural Discovery, Arizona State University, Tempe, Arizona, USA; [3]Biodesign Center for Immunotherapy, Vaccines and Virotherapy, Arizona State University, Tempe, Arizona, USA; [4]School of Life Sciences, Arizona State University; [5]Present address: LGC, Biosearch Technologies, Middleton, Wisconsin, USA;

[6]Magnetic Resonance Resource Center, Arizona State University, Tempe, Arizona, USA; [7]Present address: Biodesign Center for BioEnergetics and School of Molecular Sciences, Arizona State University, Tempe, Arizona, USA; [8]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona, USA; [9]Biodesign Center for Innovations in Medicine, Arizona State University, Tempe, Arizona, USA; [10]Biodesign Virginia G. Piper Center for Personalized Diagnostics, Arizona State University, Tempe, Arizona, USA; [11]Biodesign Center for Mechanisms of Evolution, Arizona State University, Tempe, Arizona, USA.

Highlights

- Viroporin activity and cation selectivity was confirmed from electrophysiology measurements

- Mapping of the binding site of HMA to the MHV-E viroporin

- Application of chemical shift mapping using aromatic ^{1}H-^{13}C TROSY to interrogate membrane protein interactions

Abstract

One of the functions of the Coronavirus envelope (E) proteins is the oligomerization (assembly) in the host cell membrane to form ion channels, called viroporins, which are thought to be involved in budding and scission of nascent virions. Hexamethylene amiloride, a broad-spectrum ion channel inhibitor, has been previously shown to inhibit the ion channel activity of several of the coronavirus envelope proteins, including MHV-E and SARS-E, and attenuate viral replication in plaque assays (in vitro). Here we propose the binding site of this inhibitor leading to the inactivation of the murine coronavirus MHV-A59 E viroporin. The assignment of the residues involved in inhibitor binding is based on analysis of chemical shift perturbations in 2D [^{1}H,^{15}N]-TROSY-HSQC and 2D aromatic [^{1}H,^{13}C]-TROSY-HSQC NMR spectra, and intrinsic tryptophan fluorescence spectroscopy. The combined data suggests that the binding site is located at the N-terminal opening of the channel involving a hydrogen bonding network between the guanidine group of HMA and side chain amides of Gln15, and potential π-π stacking interactions between the aromatic side chain (indole ring) of Trp11 (on the same face of

the transmembrane helix of MHV-E) and the aromatic pyrazine heterocyclic ring of HMA. This binding configuration orients the hexamethylene ring at the opening of the channel, suggesting steric effects may have a role in the mechanism of channel inactivation. These findings will be important in guiding the design of novel improved channel inhibitors that are more selective, show less side effects, and have an improved therapeutic index-safety profile.

Keywords

Coronavirus envelope protein, viroporin, ion channel inhibitors (hexamethylene amiloride), solution NMR, aromatic [^{1}H,^{13}C]-TROSY-HSQC chemical shift perturbations

1. Introduction

1.1. Coronavirus envelope proteins

Coronaviruses, the largest known RNA viruses, are classified according to genetic criteria into three groups: α, β, and γ. The β coronaviruses are further classified into the four lineages, A, B, C, and D: A includes Murine hepatitis virus (MHV) and the Human coronaviruses OC43 (Vabret et al. 2003) and HKU1 (Woo et al. 2005); B includes SARS-CoV (Drosten et al. 2003) and SARS-CoV-2 (Zhou et al. 2020); and C includes MERS-CoV (Zaki et al. 2012). SARS-CoV-2 is responsible for the ongoing COVID-19 pandemic (Zaki et al. 2012), which has been reported in 1,226,644 cases with 66,560 deaths as of 4/5/2020 (https://coronavirus.jhu.edu/map.html).

In the structural organization of the virion, RNA is in a complex with the nucleocapsid (N) protein which together are surrounded by a lipid envelope which contains three

86

structural proteins: the spike (S), membrane (M) and small envelope (E) proteins (Fehr and Perlman 2015).

The E protein is perhaps the most "mysterious" of the three structural proteins that are embedded in the viral envelope in terms of the varied functional roles that have been proposed (Schoeman and Fielding 2019). Despite large variations in sequence identity between coronavirus E proteins (Fig. 1.8), there are conserved structural features: the size and location of the transmembrane domains, the conserved proline at the center of a proposed β-turn (Li, Surya, et al. 2014), a cysteine-rich membrane proximal region which is subject to palmitoylation (Lopez et al. 2008), alignment of polar residues within the lumen of the channel (Ye and Hogue 2007) and a putative C-terminal PDZ binding motif (Teoh et al. 2010; Jimenez-Guardeno et al. 2014). In the context of the viral envelope, the E protein is found in low copy numbers and is involved in interactions with the M protein where E is thought to induce curvature in the viral envelope. Within the infected host cell, CoV E proteins have been proposed to alter cellular pathways through a putative PDZ binding motif (PBM) (Teoh et al. 2010; Jimenez-Guardeno et al. 2014), and through indirect modification of host cell channels (Torres et al. 2015). Additionally, CoV E proteins assemble into pentameric channels called viroporins (OuYang and Chou 2014; Nieva, Madan, and Carrasco 2012) which are thought to be implicated in viral budding and scission. These multiple functional roles provide independent drug targets exploiting either the ion channel activity (Wilson, Gage, and Ewart 2006a) or the PDZ binding domain. The drug SB2035805, which targets the PBM of SARS-CoV-1, has been shown to increase the survival of the host in animal models (Jimenez-Guardeno et al. 2014). Therefore, CoV E proteins are promising drug targets.

1.2 Viroporins and their inhibitors

Inhibition of CoV E viroporins by HMA

Several of the CoV E proteins form cation selective channels (Wilson et al. 2004; Wilson, Gage, and Ewart 2006a; Surya et al. 2015) that have been shown to be inhibited by hexamethylene amiloride (HMA), an amiloride analogue (Wilson, Gage, and Ewart 2006a; Li, Surya, et al. 2014). Of note, amiloride (Fig. 4.1), which features a chemical structure where the hydrogens are located on N5 instead of the cyclic hexamethylene ring (Fig. 4.1), does not inhibit channel activity (Wilson, Gage, and Ewart 2006a). The only structural information on HMA binding to CoV E proteins are on truncated SARS-E constructs (Pervushin et al. 2009; Li, Surya, et al. 2014; Surya, Li, and Torres 2018), where two potential binding sites were proposed in which HMA localizes at both ends of the transmembrane helix on the basis of backbone amide chemical shift perturbations. It was speculated by Pervushin et al (Pervushin et al. 2009) that the binding of HMA to SARS-E near Asn15 should be similar in other CoV E proteins since this position, which was found to be lumen-facing from the solution NMR structure, is a conserved polar residue in the CoV E proteins and is thought to form the selectivity filter. The authors further propose that the length of the side chain could play a role in mediating hydrogen bonding interactions with HMA that lead to channel inactivation, which may account for the HMA-insensitivity of avian coronavirus infectious bronchitis virus (IBV) (Wilson, Gage, and Ewart 2006a), which has a the shorter side chain of threonine at this position.
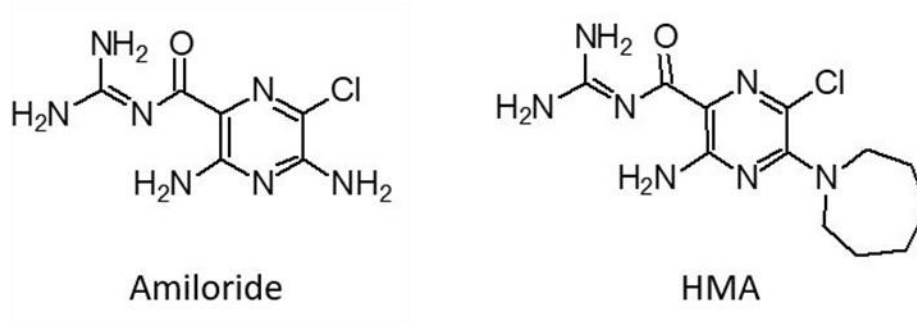
Figure 4.1. Structures of amiloride and hexamethylene amiloride.

Inhibition of other viroporins by HMA

Additionally, Vpu from HIV-1 (Ewart et al. 1996) and p7 from HCV (Premkumar et al. 2004) form viroporins which are inhibited by HMA (Ewart et al. 2004). BIT225, which is structurally similar to HMA, also inhibits p7 viroporin activity and underwent clinical trials as an anti-HIV-1 drug candidate (Wilkinson et al. 2016). HMA is one of the most effective and well characterized Vpu ion channel inhibitors while other amiloride analogues such as 5-(N,N-dimethyl) amiloride (DMA) and 5-(N-methyl-N-isobutyl) amiloride (MIA) have been shown to be less effective (Rosenberg, Weaver, and Casarotto 2016).

1.3. Experimental approaches for studying MHV-E·HMA interactions

NMR spectroscopy is well suited to study the binding of HMA to MHV-E. In chemical shift perturbation experiments, residue specific information can be obtained due to the sensitivity of spins to local changes in the environment as a result of either direct binding or allosteric changes that manifest as perturbations in peak positions (Williamson 2013). Chemical shift perturbations of amide groups are most commonly used to identify

possible binding sites. Perturbations of Ile, Leu, Val methyl ^{1}H-^{13}C groups have also been utilized to study binding (Wiesner and Sprangers 2015). Aromatic residues in membrane proteins are often found clustered at the interface of the membrane where they play important functional roles (Zhou and Cross 2013; Howard et al. 2008). The aromatic residues in MHV-E are fairly well distributed throughout the primary sequence serving as additional reporters of HMA binding by monitoring chemical shift perturbations in the aromatic ^{1}H-^{13}C groups, for which TROSY selection in the aromatic carbon dimension gives increased sensitivity for application to membrane proteins (Pervushin et al.).

Additionally, MHV-E contains a single Trp residue located near the N-terminal start of the transmembrane helix (as inferred from deuterium exchange) making it well suited for probing interactions with HMA through fluorescence spectroscopy. The position and intensity of the emission maximum of the Trp fluorescence emission is generally correlated with polarity of the local environment, with a blue shift corresponding to a change in conformation to a more hydrophobic environment (Ward 1985). Additionally, HMA fluoresces in the blue region with an excitation maximum near 360 nm, which overlaps with the emission bandwidth of Trp. In principle, FRET should occur between Trp11 and HMA, if HMA binds close in distance to the Trp11 with a suitable orientation.

These experiments allow comprehensive characterization of the HMA binding site in MHV-E, which is expected to bind at the selectivity filter, located at position 15, a conserved polar residue in all βCoV E proteins.

2. Materials and Methods

2.1. Sample preparation

All materials were obtained from Sigma unless otherwise noted. $^{15}NH_4Cl$ and ^{13}C-glucose were obtained from Cambridge Isotope Laboratories. Expression and isolation of MHV A59 envelope protein are outlined in detail in Ch. 3. NMR samples contained ~1 mM MHV-E in 200 mM LMPC, 50 mM MES, pH 5.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP and 5% D_2O. For NMR binding experiments, a 1 M HMA stock solution in DMSO was added to a final concentration of 10 mM HMA (1:10 molar ratio MHV-E:HMA). DMSO was added to 1 % in the MHV-E only sample as a control to account for perturbations arising from the solvent.

2.2. Ion Channel Measurements

Black lipid bilayer electrophysiology measurements were performed following the methods outlined in Chapter 2

2.3. Nuclear magnetic resonance spectroscopy

All NMR experiments were performed at 318 K on a Bruker Avance III 850 MHz spectrometer equipped with a 5 mm TCI $^1H/^{15}N/^{13}C$ triple resonance cryoprobe at the Magnetic Resonance Resource Center (MRRC) at Arizona State University. Resonance assignments of the full length MHV-E protein from strain A59 were obtained using TROSY triple resonance experiments and verified with inverse labelling as described in detail in Chapter 3, and the ^1H, ^{15}N, and ^{13}C chemical shifts of the backbone resonances have been deposited in the BMRB. Aromatic ^1H and ^{13}C resonance assignments were

obtained using the following experiments: 2D aromatic [^{1}H,^{13}C]-TROSY-HSQC (Pervushin et al. 1998), 3D [^{1}H,^{1}H]-NOESY-^{15}N-TROSY (Pervushin et al. 1997) and 3D aromatic [^{1}H,^{1}H]-NOESY-^{13}C-TROSY (Pervushin et al. 1998). 3D experiments utilized a 50 % non-uniform sampling schedule. Spectra were processed using NMRPipe (Delaglio et al. 1995) and analyzed using NMRFAM-SPARKY (Lee, Tonelli, and Markley 2015). Non-uniformly sampled data (Hyberts, Arthanari, and Wagner 2012) were reconstructed using the SMILE algorithm (Ying et al. 2017). Chemical shift perturbations in the ^{15}N-TROSY caused by addition of HMA were calculated using the following equation from (Williamson 2013):

$$\Delta\delta = \sqrt{\frac{(\Delta\delta_H)^2 + (\Delta\delta_N/5)^2}{2}}$$

Chemical shift perturbations in the aromatic ^{13}C-TROSY caused by addition of HMA were calculated using the following equation from (Williamson 2013):

$$\Delta\delta = \sqrt{\frac{(\Delta\delta_H)^2 + (\Delta\delta_C/5)^2}{2}}$$

(These equations are directly from (Williamson 2013). Alternatively, the Euclidean distance of the combined perturbations can be calculated, omitting the scaling factor. B-factor of 1/5 reflects differences in chemical shift ranges of H vs C, and is generally set to 0.2-0.3)

2.4 Fluorescence spectroscopy

Fluorescence measurements were performed on a Spectramax M5 microplate reader using black-walled fluorescence 96-well plates (Corning). The parameters used for monitoring intrinsic Trp fluorescence in response to HMA titration were as follows. An excitation wavelength of 280 nm with a cutoff filter of 325 nm was used for excitation and a fluorescence emission spectrum was collected from 300-500 nm in 5 nm intervals. The fluorescence spectrum of 50 µM HMA in 50 mM sodium phosphate, pH 6.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP and 8 mM DPC • 2mM DMPC was acquired using an excitation wavelength of 350 nm with a cutoff filter of 420 nm and emission was recorded from 400-600 nm in 5 nm intervals. Protein samples for fluorescence measurements contained 50 µM MHV-E in 50 mM sodium phosphate, pH 6.5, 50 mM NaCl, 1 mM EDTA, 5 mM TCEP and either: 8 mM DPC • 2 mM DMPC, 8 mM DPC • 2 mM POPC, or 20 mM SDS as the membrane mimetic. For the HMA titrations, a 1 mM HMA stock was added to 200 µL aliquots of 50 µM MHV-E to give final HMA concentrations of 10, 20, 30, 40 and 50 µM, and the total DMSO concentration was adjusted to 1 % DMSO for all samples.

3. Results

3.1. Viroporin activity

Viroporin activity of MHV-E was confirmed with electrophysiology measurements in a POPC/POPE/POPG bilayer (Fig. 2.7). A voltage sweep (Fig. 2.7A) was carried out and the corresponding current-voltage I-V plot (Fig. 2.7B) gives a reversal potential close to

the theoretical equilibrium potential for Na$^+$, indicating cation selectivity that was previously reported (Wilson, Gage, and Ewart 2006a).

3.2. Amide ^{1}H-^{15}N chemical shift perturbations by HMA binding

A large number of perturbations were observed in the ^{1}H,^{15}N-TROSY HSQC spectrum (Fig. 4.2). The corresponding perturbation plot including secondary structure predicted from TALOS-N (Shen and Bax 2013) is shown in Fig 4.3. The backbone amide chemical shifts experiencing the greatest perturbations were in the regions of the N-terminal end of transmembrane helix (Tyr12, Val13, Gly14), near the proposed β-turn (Ser51), the C-term end of helix 3 (the second extramembrane helix) (72-75) and within the putative PBM (Asp81, Asp82). Perturbations were additionally seen for selected side chain amides, however many of these resonances could not be assigned using the ^{15}N-NOESY-TROSY. ^{1}H chemical shift perturbations (Fig. S4.1) in the transmembrane region exhibited a pattern of essentially monotonically decreasing values from residues Ile17-Ala35 with the directions of the ^{1}H perturbations reversing in the center of the helix. Minimal perturbations at the center of the transmembrane helix (residues Leu24-Ile29) were observed.
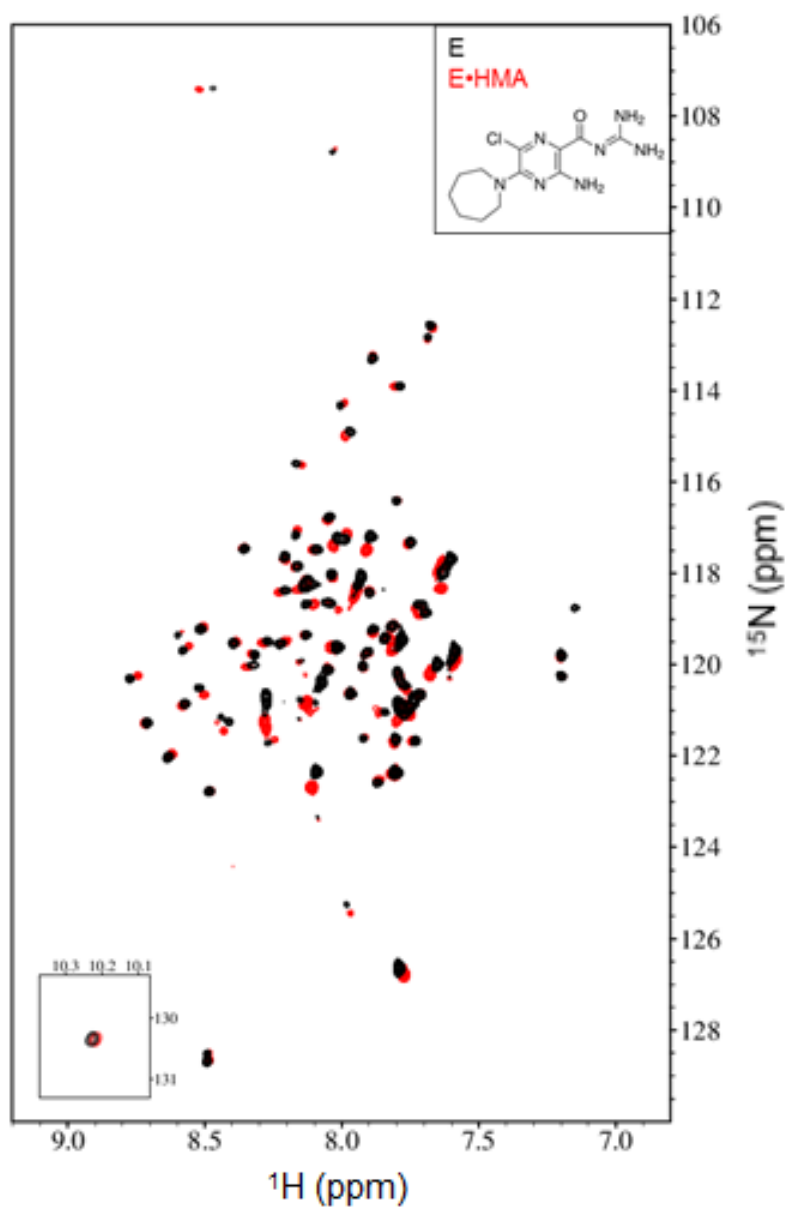
Figure 4.2. Amide TROSY perturbations of MHV-E caused by HMA. ^{1}H-^{15}N TROSY-HSQC spectra of ~1 mM MHV-E in the absence (black) and presence of a 10-fold molar excess of HMA in DMSO (red) collected at 45 °C.
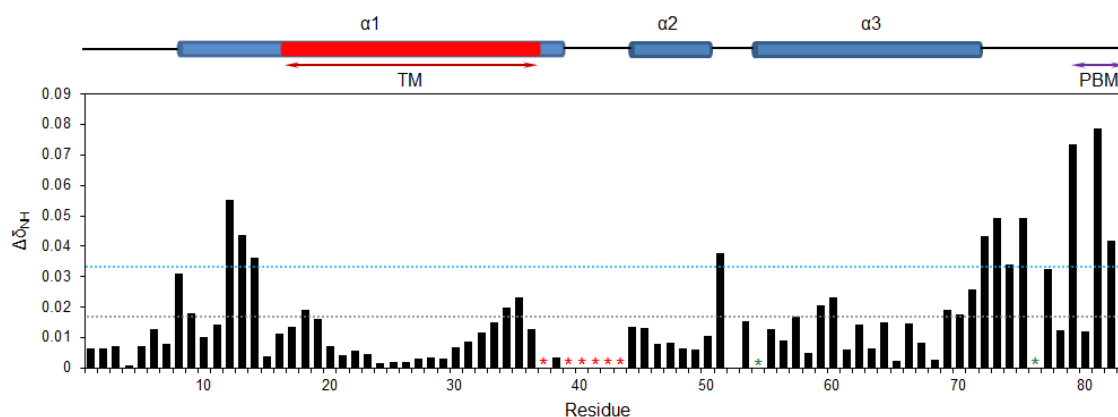
Figure 4.3. Perturbations of amide groups by HMA. Dotted gray line designates the average of the combined ^{1}H and ^{15}N chemical shift perturbations $\Delta\delta_{NH}$ for all residues and the cyan dotted line designates the threshold, taken as one standard deviation above the average of the perturbations. Prolines are indicated by green asterisks. Residues for which assignments could not be obtained are indicated by red asterisks. Shown above the perturbation plot are the positions of the 3 helices of MHV-E based on TALOS-N analysis (Shen and Bax 2013), where TM designates the transmembrane region (the micelle-embedded residues are Ile17-Ser36, as determined by deuterium exchange (Fig. 3.3C) and PBM denotes the putative PDZ binding motif (PDZ stands for **P**ost-synaptic density protein 95 (PSD-95), **D**rosophila disc large tumor suppressor (Dlg1), **Z**ona occludens 1 (ZO-1), the first proteins for which this domain (5-6 β strands and 2 α helices, one short and one long) was described).

3.3. Aromatic ^{1}H-^{13}C chemical shift perturbations by HMA binding

3.3.1. Aromatic ^{1}H-^{13}C resonance assignments

Assignment of aromatic ^{1}H and ^{13}C chemical shifts was obtained through the ^{15}N-NOESY-TROSY and aromatic ^{13}C NOESY-TROSY spectra. Fig. 4.4 outlines the approach used to obtain assignments for the ^{1}H$_\delta$ and ^{13}C$_\delta$ chemical shifts of the transmembrane Phe residues Phe18, Phe20, and Phe33. Representative strips from the 3D aromatic ^{13}C-NOESY-TROSY in Fig. 4.5 illustrates intra-aromatic NOE cross peaks of Trp11, which were used to obtain assignments. Inter-aromatic NOEs (between Trp 11

and neighboring Tyr12, and the more distant F20) were additionally observed which will serve to more accurately define the side chain conformations in this region (at the opening of the channel). The assigned aromatic TROSY spectrum is shown in Fig. 4.6.
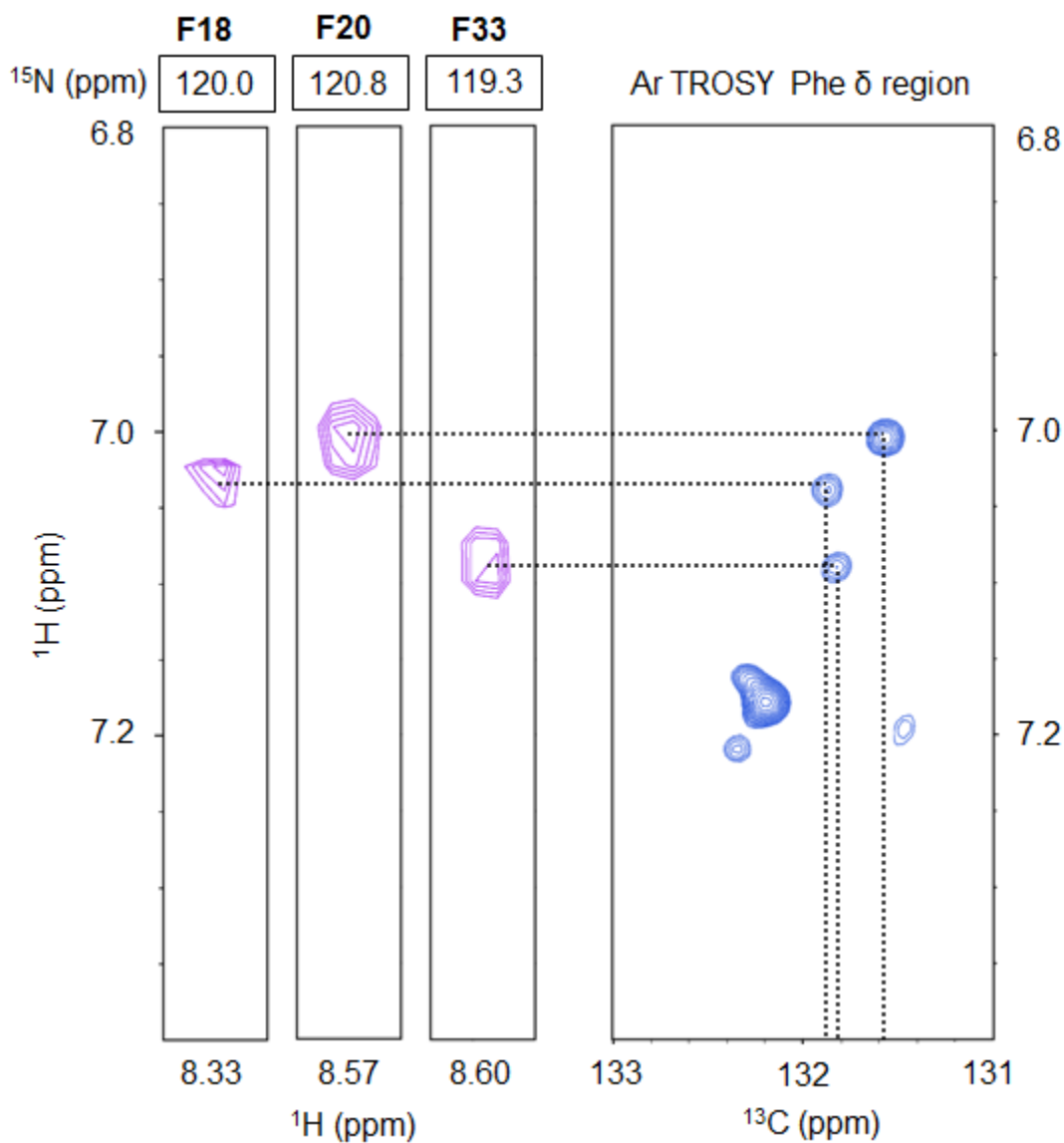
Figure 4.4. Assignment of the transmembrane Phe δ ^{13}C chemical shifts. There are a total of 5 Phe residues in MHV-E, 3 of which in the transmembrane domain were assigned based on matching the ^{1}H resonance frequencies from the corresponding ^{15}N-NOESY strips of Phe18, Phe20, and Phe33 to the Ar TROSY spectrum (horizontal broken lines) from which the ^{13}C chemical shifts could be obtained as shown by the dotted vertical lines.
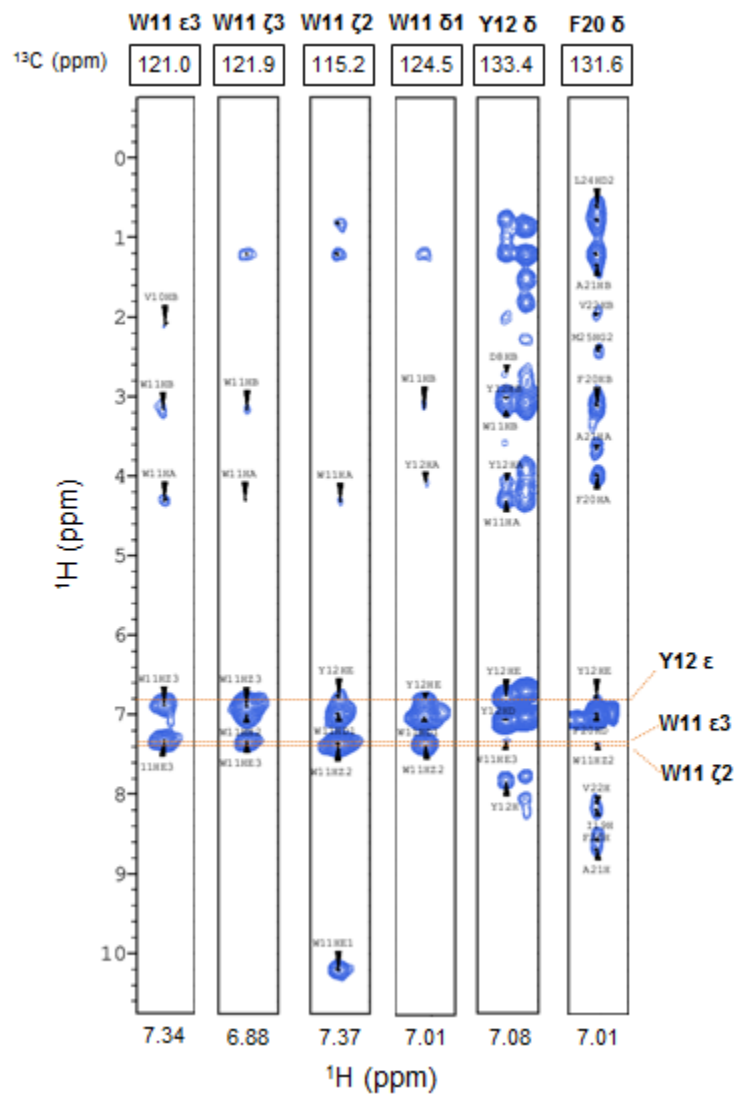
Figure 4.5. 3D aromatic ^{13}C-NOESY-TROSY of MHV-E. Representative strips from the 3D aromatic ^{13}C-NOESY-TROSY showing intra- and inter-aromatic NOE cross peaks. NOE cross peaks from W11 ζ2 and Y12 δ to F20 δ, which is located in the transmembrane region.
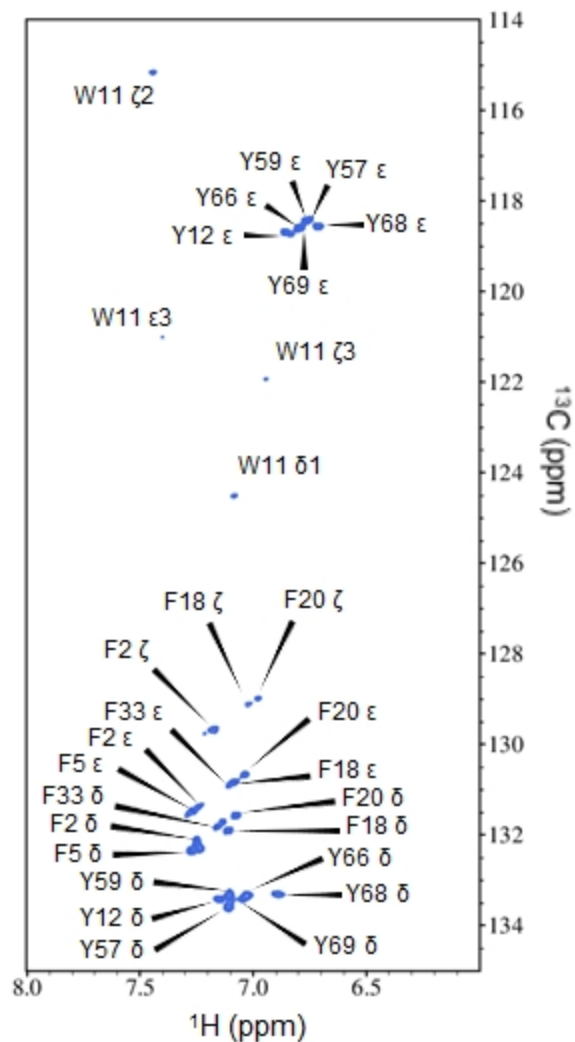
Figure 4.6. Assigned aromatic ^{1}H,^{13}C-TROSY-HSQC of MHV-E in LMPC micelles.

3.3.2. Aromatic ^{1}H-^{13}C perturbations

Aromatic ^{1}H-^{13}C TROSY-HSQC spectra of MHV-E in the absence (black) and presence of a 10-fold molar excess of HMA (red) are shown in Fig. 4.7A. The corresponding perturbation plot (Fig. 4.7B) indicates significant perturbations of the Trp11 ε3, ζ3, δ1, and neighboring Tyr12 δ ^{1}H-^{13}C groups by HMA. Tyr68 δ also exhibits a slight perturbation; however, this was below the calculated threshold.
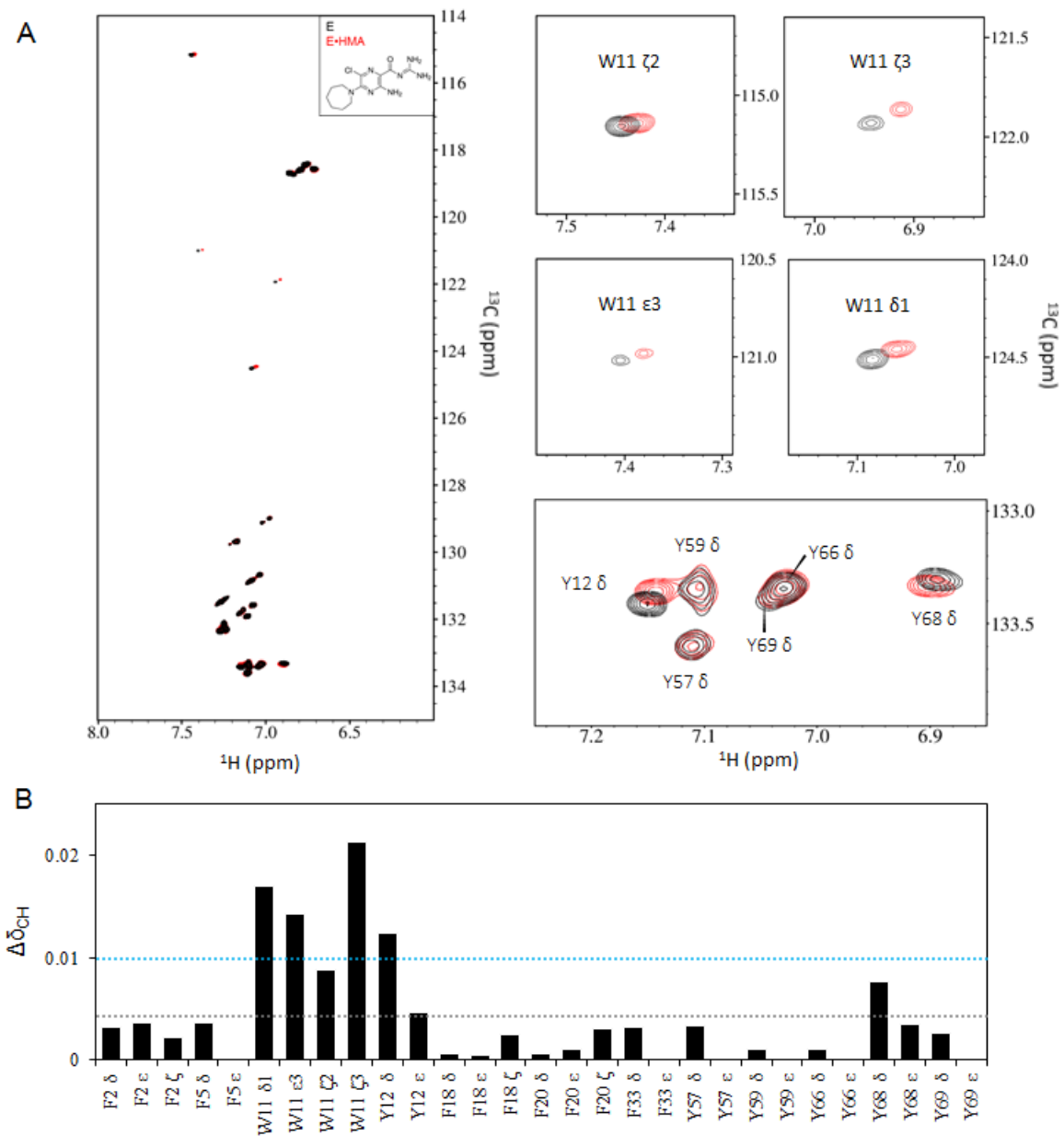
Figure 4.7. Aromatic ^{1}H-^{13}C perturbations. **(A)** Aromatic ^{1}H-^{13}C TROSY-HSQC spectra of MHV-E in the absence (black) and presence of a 10-fold molar excess of HMA (red). **(B)** Corresponding perturbation plot in which the threshold (dotted cyan line) is set as one standard deviation above the average of all perturbations (dotted gray line).

3.4 Tryptophan fluorescence

The intrinsic fluorescence (stemming primarily from the single Trp at position 11) of MHV-E in DPC/DMPC mixed detergent lipid micelles has a maximum emission at 355 nm in the absence of HMA. DMPC is the parent lipid from which LMPC is derived. The MHV-E Trp11 emission peak experienced a reduction in intensity and accompanying blue shift (~10 nm) with increasing concentrations of HMA (Fig. 4.8). At higher concentrations of HMA, FRET was observed between Trp11 and HMA indicated by the cyan arrow in Fig. 4.8. Figure 4.9 examines the effect of membrane mimetics on HMA binding. FRET is observed for the lipid-detergent mixed micelles, DPC/DMPC and DPC/POPC, with an accompanying blue shift of the HMA fluorescence (HMA fluorescence in the absence of MHV-E is shown for reference in cyan). However, there was essentially no FRET observed between Trp11 and HMA for MHV-E in SDS micelles.
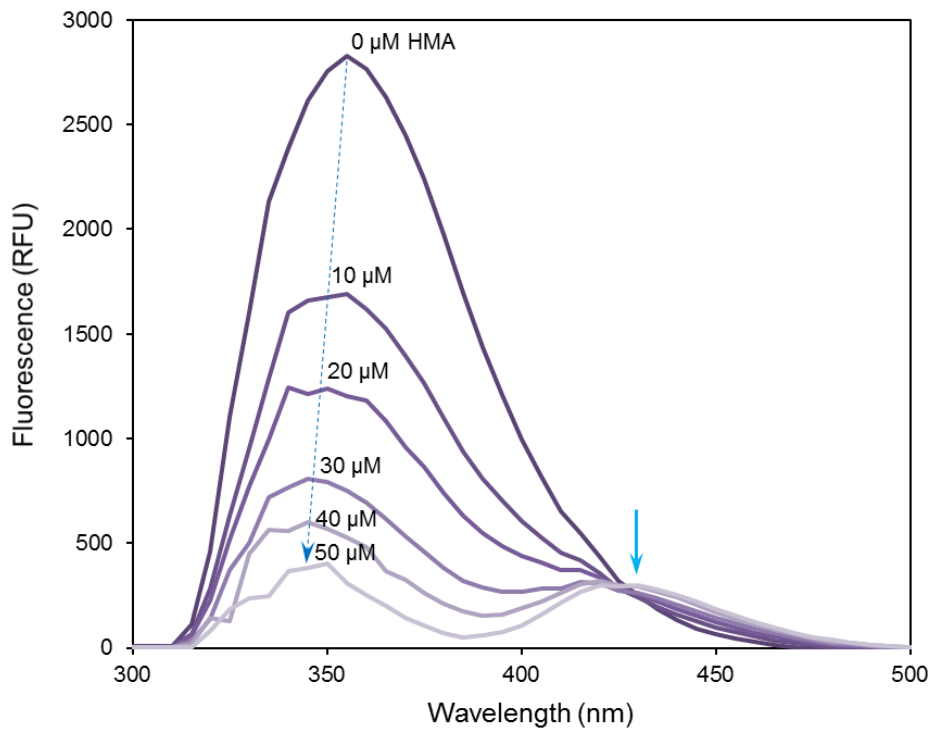
Figure 4.8. HMA titration monitored by intrinsic Trp fluorescence. MHV-E Trp11 in response to HMA titration at the indicated concentrations of HMA. A decrease in Trp11 fluorescence emission with corresponding blue shift was observed as indicated with the dotted blue line. Additionally, at higher HMA concentrations a FRET peak was observed as indicated by the solid blue arrow.
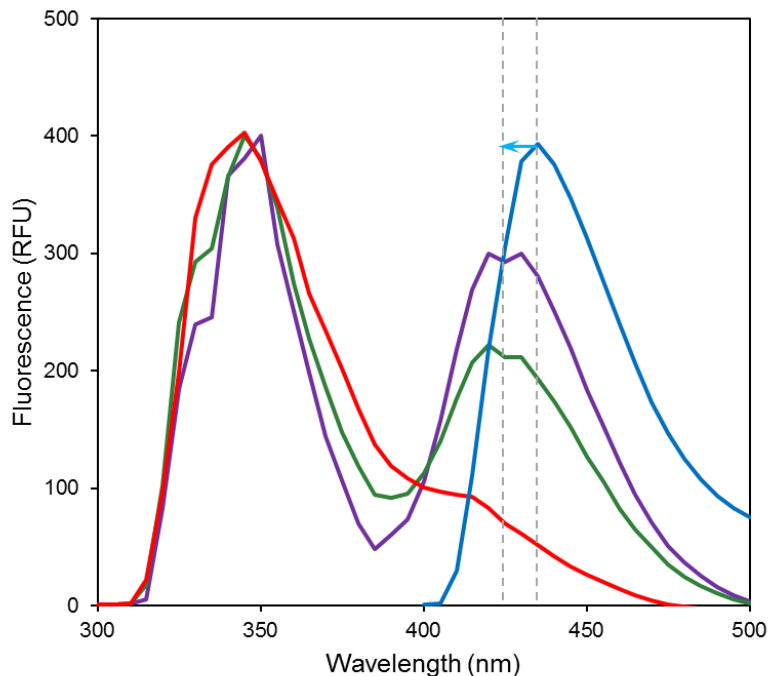
Figure 4.9. FRET between Trp11 of MHV-E and HMA. Blue shifting of both Trp11 and HMA emission peaks is observed at higher HMA concentrations (blue shifting of the HMA peak is indicated by the cyan arrow). Emission maxima were normalized to Trp emission maximum of MHV-E in DPC/DMPC detergent-lipid mixed micelles + 50 μM HMA. DPC/DMPC (purple trace); DPC/POPC (green trace); SDS (red trace); HMA (cyan trace)

4. Discussion/Conclusions

4.1. Proposed binding site. HMA binds the N-terminal opening of the MHV-E viroporin

The pattern of amide chemical shift perturbations (Figs. 4.4 and 4.5) suggests that MHV-E is either interacting with HMA near the start of the transmembrane helix at the channel opening or at the C-terminal extramembrane region encompassing the PBM. The large perturbations observed for the C-terminal residues, encompassing the putative PDZ binding motif in the ^{1}H,^{15}N-TROSY-HSQC (Figs. 4.2 and 4.3) may result from interactions between HMA and side chains of polar residues, which are more prevalent in

this region. However, they may also underlie a more complex binding interaction between HMA, the residues at the end of helix 3 (2nd extramembrane helix) and the PBM, which may have implications in the mechanism of attenuation of viral replication, since there is lack of evidence to establish a correlation between the attenuation of viral replication with the inhibition of channel activity. Another possibility is conformational rearrangement transmitted from the N-terminal to the C-terminal end upon binding of HMA, leading to disruption of the oligomer interface at the C-terminal end of the channel. The proton shift pattern in transmembrane region (Fig. S4.1) cannot be correlated to conformational changes (such as tilting or rotation of the TM helix) without the structure of HMA bound to MHV-E, however the center of the TM helix experiences negligible perturbations suggesting these residues are near the oligomer interface forming the viroporin.

Aromatic TROSY was used to delineate the potential binding locations inferred from the amide TROSY perturbations. Only the Trp11 $\varepsilon3$, $\zeta3$, $\delta1$ and Tyr12 δ ^{1}H, ^{13}C chemical shifts experienced significant perturbations suggesting that HMA binds at or near this region. Fluorescence spectroscopy was used to further characterize HMA binding, since Trp11 is the only tryptophan found in MHV-E. FRET efficiency depends on distance and orientation (Ward 1985). Essentially no FRET signal was detected for MHV-E in SDS micelles (Fig. 4.10). This result fits well with the assumption from the NMR linewidth analysis that suggests MHV-E in SDS micelles is predominantly monomeric and shows very little tendency to oligomerize, precluding formation of the pentameric viroporin structure.

Taken together with the results of the NMR experiments (Figs. 4.2, 4.3 and 4.7), the fluorescence measurements indicates that Trp11 is located at or near the binding pocket for HMA. Very likely the binding is stabilized through π-π interactions where the pyrazine ring of HMA, or cation-pi interactions between Trp11, Tyr12 and the guanidine moiety of HMA however, the guanidine moiety is more likely to localize at the conserved polar residue Gln15 (Asn15 in SARS-CoV E) would position the amiloride (guanidine group) moiety of HMA to be encircled (surrounded) by the 5-fold symmetric Gln15 amide side chains (position 15 is a conserved polar residue at in CoV E proteins) which can engage in a hydrogen bonding network with the guanadino group of the amiloride moiety, thereby orienting the hexamethylene ring of HMA towards the N-terminal opening of the channel, which suggests inhibition by a steric mechanism. This interpretation is supported by electrophysiology measurements and plaque assays showing no inhibition of either channel conductance or viral replication of MHV-E treated with amiloride (Fig. 4.1), the parent compound from which HMA is derived and lacking the seven-membered ring (Wilson, Gage, and Ewart 2006a). This result implies that the bulky seven-membered ring is essential for inhibition of viroporin activity.

4.2 Comparison with known binding sites/mechanisms of other CoV E viroporins

SARS-E

Of the CoV E viroporins for which HMA has been shown to inhibit viroporin activity, SARS E has been most extensively characterized. The mutation N15A abolished viroporin activity (Verdia-Baguena et al. 2012) and was found to be lumen-facing in the NMR structure of the transmembrane domain of SARS-E in DPC micelles (Pervushin et

al. 2009). Alignment of polar residues within the lumen of the channel is a general

feature of CoV E viroporins. Another mutation, Thr11Ala, was actually found to increase

peak channel currents (Fig. 1.11) (Verdia-Baguena et al. 2012) suggesting that position

11 on the same face of the helix as Asn15 could be implicated in the mechanism of HMA

channel inactivation in other β CoV E proteins. It has been proposed that the shorter side

chain of Ala11 vs Thr11 is thought to widen the entrance to the channel pore (Verdia-

Baguena et al. 2012). The role of Trp11 in HMA binding can be further characterized

through fluoroindole labeling and ^{19}F TROSY spectroscopy.
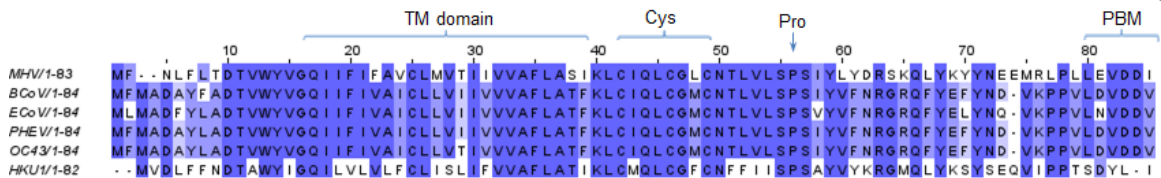
Figure 4.10. Sequence alignment of beta coronavirus subtype A (MHV lineage) envelope proteins. Residues are shaded by % sequence identity. The transmembrane domain is denoted by TM domain. Cys denotes the cysteine-rich membrane proximal region, Pro denotes the conserved proline of the proposed β-turn, and PBM denotes the PDZ binding motif. Multiple sequence alignment (MSA) was done using T-Coffee (Notredame, Higgins, and Heringa 2000) and viewed with JalView (Clamp et al. 2004).

Future mutagenesis studies (Gln15Thr, Trp11Ala) are needed in conjunction with surface

plasmon resonance (SPR) and electrophysiology measurements to determine the effect of

π-π stacking on the binding affinity and to determine to what extent π-π interactions

stabilize the binding of HMA, if at all. As opposed to the requirement of a polar residue

with a long polar side chain (Asn, Gln) at position 15, tryptophan at position 11 does not

appear to be essential for binding in other CoV E -proteins, since CoV E proteins (Fig.

1.8) along the SARS-related phylogenetic branch (group β B) do not have a Trp at this

106

position. However, Trp11 is conserved among subgroup A of the βCoV E proteins, including the human coronaviruses OC43 and HKU1 (Fig. 4.10). Although viroporin activity of SARS-E has been experimentally shown to be inhibited by HMA (Li, Surya, et al. 2014), there are no reported binding affinities for comparison to the values reported in (Wilson, Gage, and Ewart 2006a). The Gln15Thr mutation is expected to confer HMA resistance to MHV-E on the basis of the aforementioned considerations. There is some evidence for interaction of the Gln15 side chain amide from perturbation in this region of the ^{15}N-TROSY spectrum; however, we are not confident on the assignment of the side chain amide region to propose this at this time. Further experiments are necessary to investigate hydrogen bonding interactions between the Gln15 side chain and HMA.

4.3 Conclusions

We hypothesize this WxxxQ motif found near the N-terminal start of the transmembrane helix in the E proteins of human CoVs, OC43 and HKU1, would bind HMA similarly to MHV E and result in inhibition of both channel activity and viral replication; however experiments have to be performed in the future to prove this hypothesis.

Comprehensive measurements such as ion channel measurements, plaque assays and SPR (Rosenberg, Weaver, and Casarotto 2016) for additional βCoVs within subtypes A and B are needed to quantify to what extent the aromatic doublet Trp11-Tyr12 stabilizes the binding to HMA and if there is a trend among betacoronavirus type A and B CoVs.

Direct NOEs to the drug await an economical source for deuterated LMPC in order to use longer mixing times in order to observe longer-range NOEs between HMA and MHV-E. A deuterated LMPC would also eliminate potential overlap of LMPC and HMA NOEs,

so that the NOEs are not obscured by the aliphatic proton resonances of the LMPC aliphatic tail. In the case of the p7 viroporin, NOEs were detected only through the use of deuterated detergents (OuYang et al. 2013). Ultimately, confirmations from mutagenesis studies and NOEs between MHV-E and HMA in conjunction with MD simulations are needed to achieve a more detailed picture of the interaction of MHV-E with the HMA inhibitor at the atomic level.

Supplementary Figures
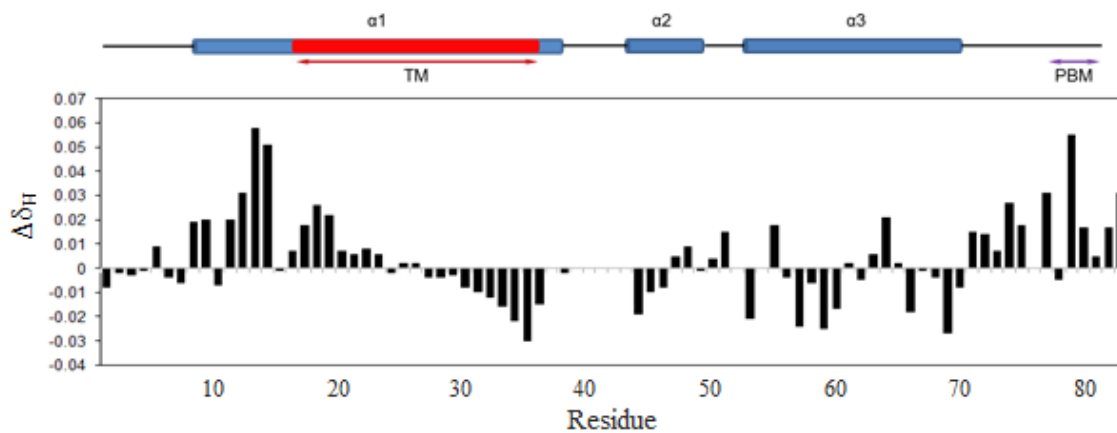
Figure S4.1. [1]H chemical shift perturbations of MHV-E by HMA. A pattern of essentially monotonically decreasing ^{1}H perturbations in the transmembrane region was observed with negligible perturbations in the center of the transmembrane helix.

CHAPTER 5

CONCLUSIONS AND OUTLOOK

Conclusions

This work outlined the first NMR studies on MHV E, which provide a foundation for
structure based drug design and probing interactions. Following extensive assignment of
backbone and side-chain resonances of MHV E, the binding site of the ion channel
inhibitor HMA was mapped to MHV E using chemical shift perturbations in both amide
and aromatic TROSY spectra. The combined data suggests the inhibitor binding site is
located at the N-terminal opening of the channel, which is in accord with one of the
proposed HMA binding sites in the SARS-CoV E. These studies have provided a
foundation for drug design and probing the interactome of the MHV E protein, which can
be extended to other coronavirus envelope proteins.

Outlook

Structure of the transmembrane region of MHV-E

Structure calculation of residues Met1-Lys38 of MHV-E, encompassing the
transmembrane region, is currently in progress using dihedral angle restraints obtained
from isotropic chemical shifts (Shen and Bax 2013) and distance restraints obtained from
manually assigned NOE cross-peaks in the ^{15}N NOESY-TROSY, aliphatic ^{13}C NOESY-
HSQC, and aromatic ^{13}C NOESY-TROSY experiments. Currently, ~75 % of all atoms
are assigned in the transmembrane domain, including aromatic side chain atoms. Figure

5.1 shows assignment of the isoleucine, leucine, and valine (ILV) side-chain methyl resonances. Representative strips of the ^{13}C NOESY-HSQC in the transmembrane region used to obtain inter-proton distance restraints for the structure calculation are shown in Figure 5.2. This structure will be used in generating a model of the pentameric sodium-selective viroporin in complex with the ion channel inhibitor HMA, which together with the reported SARS E structures (Pervushin et al. 2009; Li, Surya, et al. 2014; Surya, Li, and Torres 2018) will expand the known structural information on the coronavirus envelope proteins.
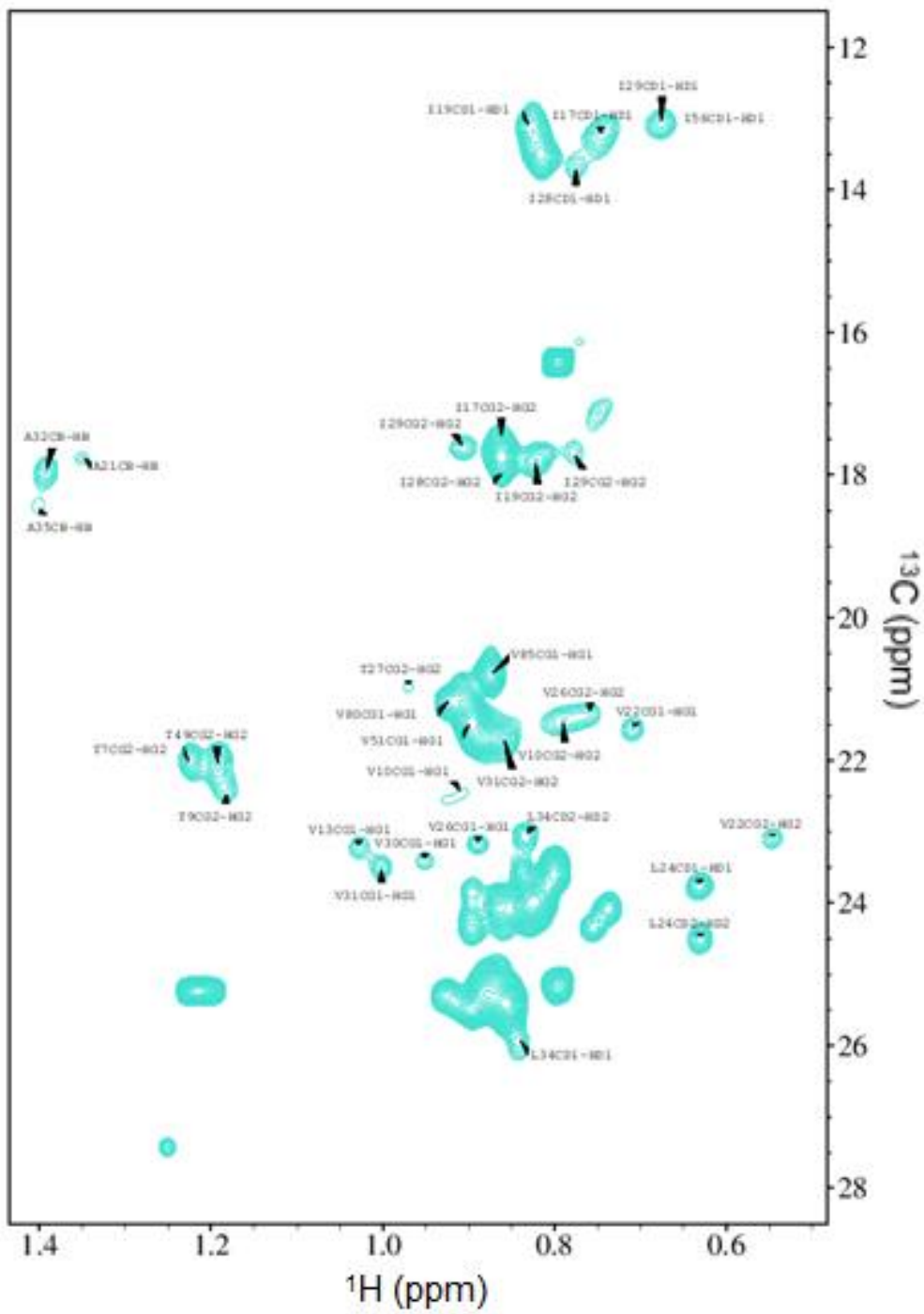
Figure 5.1. Ile Leu Val methyl assignments.
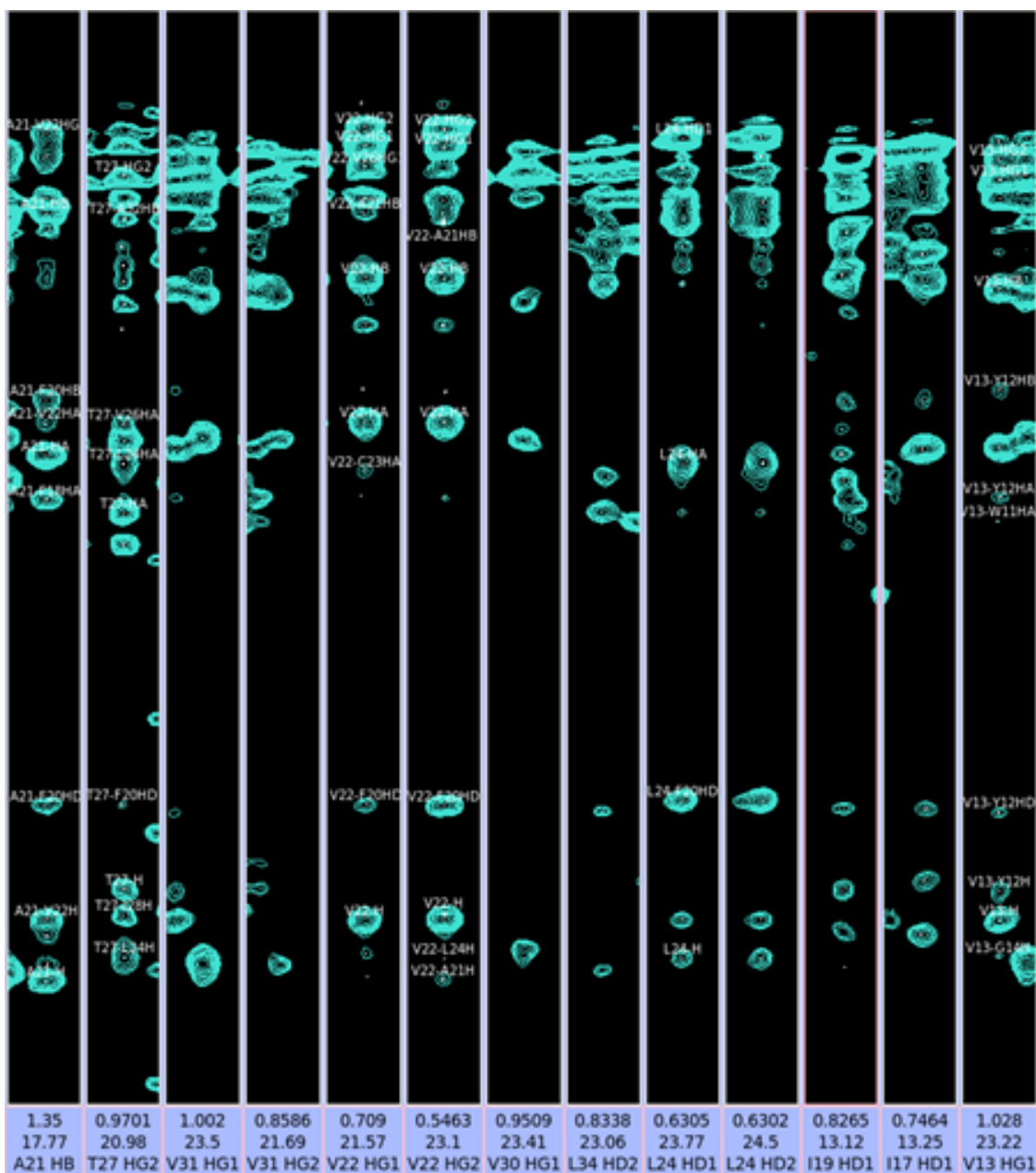
Figure 5.2. Ile Leu Val methyl NOEs in the transmembrane region.

Towards the Full length structure of MHV-E

Obtaining long-range distance restraints from NOE data is very difficult for single-pass

α-helical membrane proteins (Opella and Marassi 2017), therefore, long-range structural

113

restraints are needed for MHV-E to determine the relative orientation between the transmembrane helix and the two extra-membrane helices. Residual dipolar couplings (RDCs), which provide orientations of the internuclear bond vectors with respect to a global reference frame (Tolman et al. 2001), and paramagnetic relaxation enhancements (PREs), which report on distances up to ~35 Å (Liang, Bushweller, and Tamm 2006; Gottstein et al. 2012), are two types of long-range restraints which are commonly used to solve NMR structures of α-helical membrane proteins. Two attempts were made to obtain RDCs on MHV-E in both 3.8% neutral and 4% positively charged strained polyacrylamide gels (Jones and Opella 2004; Cierpicki and Bushweller 2004); however, a large background signal was observed in both cases, possibly caused by free acrylamide that was not removed during wash steps, or incomplete polymerization, which leads to high local dynamics in the individual polymer chains. Intense background peaks can be seen in the BEST-TROSY spectrum of ~300 μM MHV-E in a 4% positively charged gel (Figure 5.3). Polyacrylamide gels have been cast using longer polymerization times and washed more extensively in order to eliminate the observed background, which await evaluation.
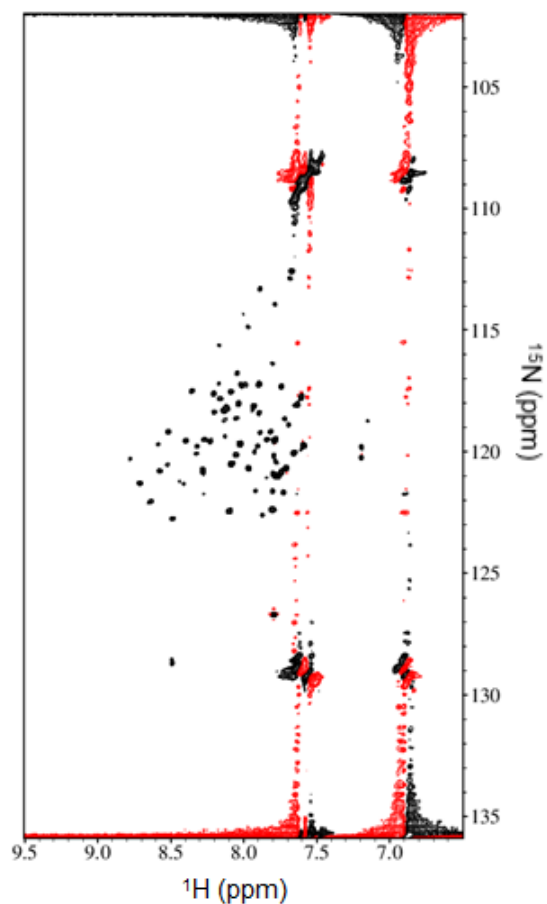
Figure 5.3. Background in gels used for RDCs. ^{1}H,^{15}N-TROSY-HSQC spectrum of ~300 μM MHV-E in a 4% positively charged polyacrylamide gel.

A single cysteine mutant of MHV-E, Cys23Ala/Cys40Ser/Cys44Ser, in which the

remaining sole cysteine (Cys47) is located in the short central helix, was generated in

order to obtain PREs. The expression level of this triple mutant was ~40% of that of the

wild type MHV-E. Nevertheless, enough material was obtained to collect a TROSY

spectrum. Shown in Fig. 5.4 is an overlay of wild-type MHV-E (black) and MHV-E

Cys23Ala/Cys40Ser/Cys44Ser (red) TROSY spectra. Many of the resonances exhibit

considerable shifts, most notably those in the transmembrane region as a result of the

Cys23Ala mutation, indicating significant conformational change. Therefore, restraints

obtained from this mutant are not likely to reflect the structure of wild type MHV-E.
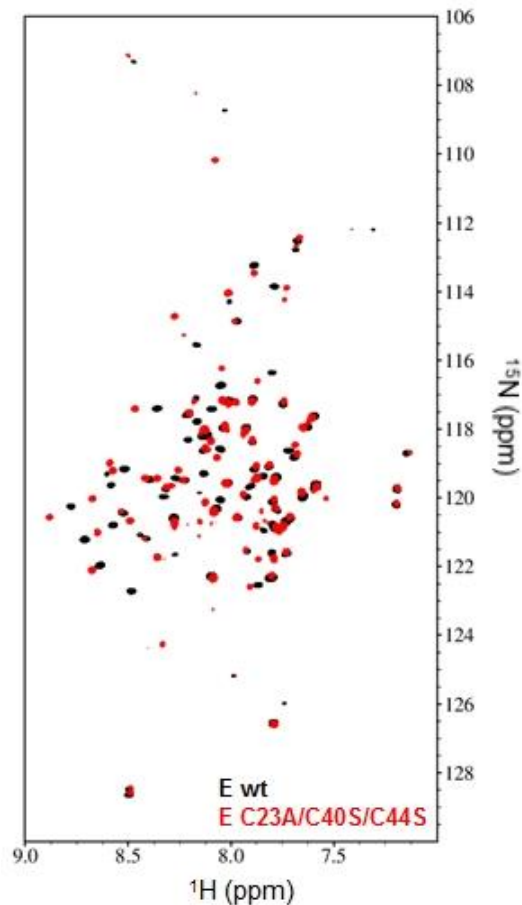
Figure 5.4. ^{1}H,^{15}N-TROSY-HSQC spectra of single cysteine mutant. Wild-type MHV-E (black) and MHV-E Cys23Ala/Cys40Ser/Cys44Ser single cysteine mutant (red).

Proteoliposome reconstitution of MHV-E for solid-state NMR studies

Solid-state NMR of membrane proteins incorporated in oriented bilayers or

proteoliposomes affords a more accurate representation of the native environment than

other mimetics, such as micelles, which are commonly used in solution-state NMR. For

example, solid-state NMR was used to solve the structure of the amantadine binding site

of the influenza M2 proton channel in lipid bilayers (Cady et al. 2010).

116
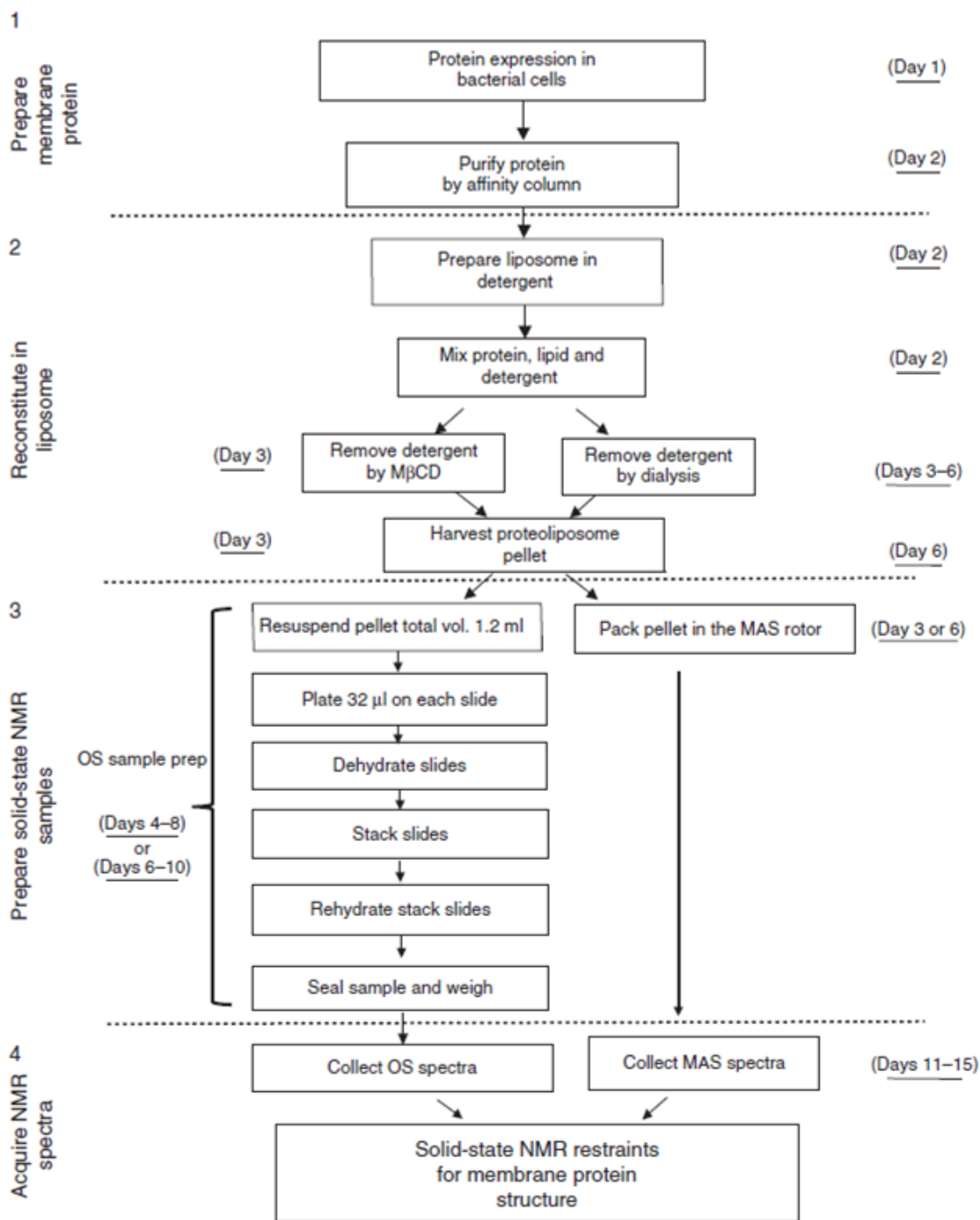
Figure 5.5. Proteoliposome reconstitution workflows for solid-state NMR. Flowchart reproduced from (Das, Murray, and Cross 2013) outlining the steps in proteoliposome reconstitution. Detergent removal is carried out either via methyl-β-cyclodextrin (MβCD) or by dialysis.
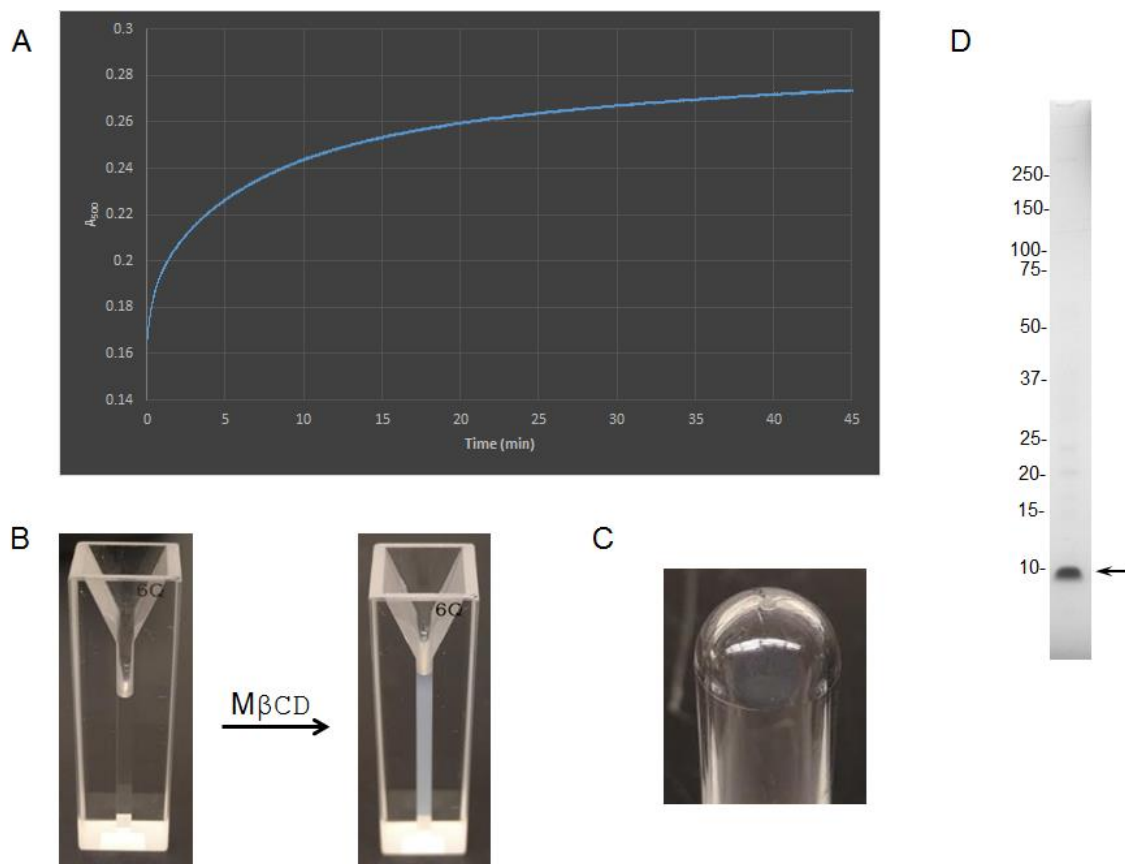
Figure 5.6. MHV-E·DMPC proteoliposome reconstitution. (A) Formation of proteoliposomes upon MβCD mediated detergent removal as monitored by turbidity measurements at A_{500}. (B) Cuvette from part (A) before and 45 minutes following addition of MβCD. (C) Final proteoliposome pellet following ultracentrifugation. (D) SDS-PAGE analysis of the proteoliposome pellet, where the arrow indicates MHV-E.

To this end, following the methods described in detail in (Das, Murray, and Cross 2013) and outlined in Figure 5.5, reconstitution of MHV-E into DMPC liposomes was carried out through methyl-β-cyclodextrin mediated detergent removal from the protein:detergent:lipid ternary complex. A summary of the results of the MHV-E·DMPC proteoliposome reconstitution are given in Figure 5.6, which demonstrated that enough material can be obtained to collect an initial 2D dipolar-assisted rotational resonance ($^{13}C–^{13}C$ DARR) spectrum (McDermott 2009). Ultimately, solid-state NMR studies will allow structure determination and elucidation of the binding site in a more physiological

118

membrane mimetic that can be tuned (by changing lipid composition, for example) to optimize stability (or emphasize different aspects of channel conductance) of the viroporin.

References

Abdul-Gader, A., A. J. Miles, and B. A. Wallace. 2011. 'A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy', *Bioinformatics*, 27: 1630-6. 10.1093/bioinformatics/btr234.

Andersen, K. G., A. Rambaut, W. I. Lipkin, E. C. Holmes, and R. F. Garry. 2020. 'The proximal origin of SARS-CoV-2', *Nat Med*, 26: 450-52. 10.1038/s41591-020-0820-9.

Bax, A., and A. Grishaev. 2005. 'Weak alignment NMR: a hawk-eyed view of biomolecular structure', *Curr Opin Struct Biol*, 15: 563-70. 10.1016/j.sbi.2005.08.006.

Bergmann, C. C., T. E. Lane, and S. A. Stohlman. 2006. 'Coronavirus infection of the central nervous system: host-virus stand-off', *Nat Rev Microbiol*, 4: 121-32. 10.1038/nrmicro1343.

Box, George E. P., J. Stuart Hunter, and William Gordon Hunter. 2005. *Statistics for experimenters : design, innovation, and discovery* (Wiley-Interscience: Hoboken, N.J.).

Cady, S. D., K. Schmidt-Rohr, J. Wang, C. S. Soto, W. F. Degrado, and M. Hong. 2010. 'Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers', *Nature*, 463: 689-92. 10.1038/nature08722.

Call, M. E., J. R. Schnell, C. Xu, R. A. Lutz, J. J. Chou, and K. W. Wucherpfennig. 2006. 'The structure of the zetazeta transmembrane dimer reveals features essential for its assembly with the T cell receptor', *Cell*, 127: 355-68. 10.1016/j.cell.2006.08.044.

Cavanagh, John. 2018. *Protein nmr spectroscopy : principles and practice* (Elsevier: Waltham, MA).

Cierpicki, T., and J. H. Bushweller. 2004. 'Charged gels as orienting media for measurement of residual dipolar couplings in soluble and integral membrane proteins', *J Am Chem Soc*, 126: 16259-66. 10.1021/ja046054g.

Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. 2004. 'The Jalview Java alignment editor', *Bioinformatics*, 20: 426-7. 10.1093/bioinformatics/btg430.

Claridge, J. K., J. Aittoniemi, D. M. Cooper, and J. R. Schnell. 2013. 'Isotropic bicelles stabilize the juxtamembrane region of the influenza M2 protein for solution NMR studies', *Biochemistry*, 52: 8420-9. 10.1021/bi401035m.

Clore, G. M., and J. Iwahara. 2009. 'Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes', *Chem Rev*, 109: 4108-39. 10.1021/cr900033p.

Cohen, J. R., L. D. Lin, and C. E. Machamer. 2011. 'Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein', *J Virol*, 85: 5794-803. 10.1128/JVI.00060-11.

Conn, P. Michael. 2008. *Sourcebook of models for biomedical research* (Humana Press: Totowa, N.J.).

Cook, G. A., H. Zhang, S. H. Park, Y. Wang, and S. J. Opella. 2011. 'Comparative NMR studies demonstrate profound differences between two viroporins: p7 of HCV and Vpu of HIV-1', *Biochim Biophys Acta*, 1808: 554-60. 10.1016/j.bbamem.2010.08.005.

Corse, E., and C. E. Machamer. 2003. 'The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction', *Virology*, 312: 25-34. 10.1016/s0042-6822(03)00175-2.

Das, N., D. T. Murray, and T. A. Cross. 2013. 'Lipid bilayer preparations of membrane proteins for oriented and magic-angle spinning solid-state NMR samples', *Nat Protoc*, 8: 2256-70. 10.1038/nprot.2013.129.

DeDiego, M. L., E. Alvarez, F. Almazan, M. T. Rejas, E. Lamirande, A. Roberts, W. J. Shieh, S. R. Zaki, K. Subbarao, and L. Enjuanes. 2007. 'A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo', *J Virol*, 81: 1701-13. 10.1128/JVI.01467-06.

Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. 1995. 'NMRPipe: a multidimensional spectral processing system based on UNIX pipes', *J Biomol NMR*, 6: 277-93. 10.1007/bf00197809.

Dian, C., S. Eshaghi, T. Urbig, S. McSweeney, A. Heijbel, G. Salbert, and D. Birse. 2002. 'Strategies for the purification and on-column cleavage of glutathione-S-transferase fusion target proteins', *J Chromatogr B Analyt Technol Biomed Life Sci*, 769: 133-44. 10.1016/s1570-0232(01)00637-7.

Drew, D., M. Lerch, E. Kunji, D. J. Slotboom, and J. W. de Gier. 2006. 'Optimization of membrane protein overexpression and purification using GFP fusions', *Nat Methods*, 3: 303-13. 10.1038/nmeth0406-303.

Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. Osterhaus, H.

Schmitz, and H. W. Doerr. 2003. 'Identification of a novel coronavirus in patients with severe acute respiratory syndrome', *N Engl J Med*, 348: 1967-76. 10.1056/NEJMoa030747.

Durr, U. H., M. Gildenberg, and A. Ramamoorthy. 2012. 'The magic of bicelles lights up membrane protein structure', *Chem Rev*, 112: 6054-74. 10.1021/cr300061w.

Dutta, A., T. Y. Kim, M. Moeller, J. Wu, U. Alexiev, and J. Klein-Seetharaman. 2010. 'Characterization of membrane protein non-native states. 2. The SDS-unfolded states of rhodopsin', *Biochemistry*, 49: 6329-40. 10.1021/bi100339x.

Ernst, Richard R., Geoffrey Bodenhausen, and Alexander Wokaun. 1987. *Principles of nuclear magnetic resonance in one and two dimensions* (Clarendon Press; Oxford University Press: Oxford Oxfordshire New York).

Ewart, G. D., N. Nasr, H. Naif, G. B. Cox, A. L. Cunningham, and P. W. Gage. 2004. 'Potential new anti-human immunodeficiency virus type 1 compounds depress virus replication in cultured human macrophages', *Antimicrob Agents Chemother*, 48: 2325-30. 10.1128/AAC.48.6.2325-2330.2004.

Ewart, G. D., T. Sutherland, P. W. Gage, and G. B. Cox. 1996. 'The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels', *J Virol*, 70: 7108-15.

Fehr, A. R., and S. Perlman. 2015. 'Coronaviruses: an overview of their replication and pathogenesis', *Methods Mol Biol*, 1282: 1-23. 10.1007/978-1-4939-2438-7_1.

Fernandez, C., C. Hilty, G. Wider, P. Guntert, and K. Wuthrich. 2004. 'NMR structure of the integral membrane protein OmpX', *J Mol Biol*, 336: 1211-21. 10.1016/j.jmb.2003.09.014.

Fischer, F., C. F. Stegen, C. A. Koetzner, and P. S. Masters. 1997. 'Analysis of a recombinant mouse hepatitis virus expressing a foreign gene reveals a novel aspect of coronavirus transcription', *J Virol*, 71: 5148-60.

Fischer, F., C. F. Stegen, P. S. Masters, and W. A. Samsonoff. 1998. 'Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly', *J Virol*, 72: 7885-94.

Gautier, A., H. R. Mott, M. J. Bostock, J. P. Kirkpatrick, and D. Nietlispach. 2010. 'Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy', *Nat Struct Mol Biol*, 17: 768-74. 10.1038/nsmb.1807.

Gottstein, D., S. Reckel, V. Dotsch, and P. Guntert. 2012. 'Requirements on paramagnetic relaxation enhancement data for membrane protein structure determination by NMR', *Structure*, 20: 1019-27. 10.1016/j.str.2012.03.010.

Hamre, D., and J. J. Procknow. 1966. 'A new virus isolated from the human respiratory tract', *Proc Soc Exp Biol Med*, 121: 190-3. 10.3181/00379727-121-30734.

Haring, J., and S. Perlman. 2001. 'Mouse hepatitis virus', *Curr Opin Microbiol*, 4: 462-6. 10.1016/s1369-5274(00)00236-8.

Hefti, M. H., C. J. Van Vugt-Van der Toorn, R. Dixon, and J. Vervoort. 2001. 'A novel purification method for histidine-tagged proteins containing a thrombin cleavage site', *Anal Biochem*, 295: 180-5. 10.1006/abio.2001.5214.

Hiroaki, H., Y. Umetsu, Y. Nabeshima, M. Hoshi, and D. Kohda. 2011. 'A simplified recipe for assigning amide NMR signals using combinatorial 14N amino acid inverse-labeling', *J Struct Funct Genomics*, 12: 167-74. 10.1007/s10969-011-9116-0.

Hoffmann, M., H. Kleine-Weber, S. Schroeder, N. Kruger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N. H. Wu, A. Nitsche, M. A. Muller, C. Drosten, and S. Pohlmann. 2020. 'SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor', *Cell* 10.1016/j.cell.2020.02.052.

Howard, M. W., E. A. Travanty, S. A. Jeffers, M. K. Smith, S. T. Wennier, L. B. Thackray, and K. V. Holmes. 2008. 'Aromatic amino acids in the juxtamembrane domain of severe acute respiratory syndrome coronavirus spike glycoprotein are important for receptor-dependent virus entry and cell-cell fusion', *J Virol*, 82: 2883-94. 10.1128/JVI.01805-07.

Howell, S. C., M. F. Mesleh, and S. J. Opella. 2005. 'NMR structure determination of a membrane protein with two transmembrane helices in micelles: MerF of the bacterial mercury detoxification system', *Biochemistry*, 44: 5196-206. 10.1021/bi048095v.

Hwang, P. M., W. Y. Choy, E. I. Lo, L. Chen, J. D. Forman-Kay, C. R. Raetz, G. G. Prive, R. E. Bishop, and L. E. Kay. 2002. 'Solution structure and dynamics of the outer membrane enzyme PagP by NMR', *Proc Natl Acad Sci U S A*, 99: 13560-5. 10.1073/pnas.212344499.

Hyberts, S. G., H. Arthanari, and G. Wagner. 2012. 'Applications of non-uniform sampling and processing', *Top Curr Chem*, 316: 125-48. 10.1007/128_2011_187.

Javier, R. T., and A. P. Rice. 2011. 'Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses', *J Virol*, 85: 11544-56. 10.1128/JVI.05410-11.

Jimenez-Guardeno, J. M., J. L. Nieto-Torres, M. L. DeDiego, J. A. Regla-Nava, R. Fernandez-Delgado, C. Castano-Rodriguez, and L. Enjuanes. 2014. 'The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein

is a determinant of viral pathogenesis', *PLoS Pathog*, 10: e1004320. 10.1371/journal.ppat.1004320.

Jimenez-Guardeno, J. M., J. A. Regla-Nava, J. L. Nieto-Torres, M. L. DeDiego, C. Castano-Rodriguez, R. Fernandez-Delgado, S. Perlman, and L. Enjuanes. 2015. 'Identification of the Mechanisms Causing Reversion to Virulence in an Attenuated SARS-CoV for the Design of a Genetically Stable Vaccine', *PLoS Pathog*, 11: e1005215. 10.1371/journal.ppat.1005215.

Jones, D. H., and S. J. Opella. 2004. 'Weak alignment of membrane proteins in stressed polyacrylamide gels', *J Magn Reson*, 171: 258-69. 10.1016/j.jmr.2004.08.022.

Kang, C., C. Tian, F. D. Sonnichsen, J. A. Smith, J. Meiler, A. L. George, Jr., C. G. Vanoye, H. J. Kim, and C. R. Sanders. 2008. 'Structure of KCNE1 and implications for how it modulates the KCNQ1 potassium channel', *Biochemistry*, 47: 7999-8006. 10.1021/bi800875q.

Kay, L. E., M. Ikura, R. Tschudin, and A. Bax. 1990. 'Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. 1990', *J Magn Reson*, 213: 423-41. 10.1016/j.jmr.2011.09.004.

King, A.M.Q., M.J. Adams, E.B. Carstens, and E.J. Lefkowitz. 2012. *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses.* (Academic Press: London).

Klammt, C., I. Maslennikov, M. Bayrhuber, C. Eichmann, N. Vajpai, E. J. Chiu, K. Y. Blain, L. Esquivies, J. H. Kwon, B. Balana, U. Pieper, A. Sali, P. A. Slesinger, W. Kwiatkowski, R. Riek, and S. Choe. 2012. 'Facile backbone structure determination of human membrane proteins by NMR spectroscopy', *Nat Methods*, 9: 834-9. 10.1038/nmeth.2033.

Koehler, J., E. S. Sulistijo, M. Sakakura, H. J. Kim, C. D. Ellis, and C. R. Sanders. 2010. 'Lysophospholipid micelles sustain the stability and catalytic activity of diacylglycerol kinase in the absence of lipids', *Biochemistry*, 49: 7089-99. 10.1021/bi100575s.

Krogh, A., B. Larsson, G. von Heijne, and E. L. Sonnhammer. 2001. 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes', *J Mol Biol*, 305: 567-80. 10.1006/jmbi.2000.4315.

Kuo, L., K. R. Hurst, and P. S. Masters. 2007. 'Exceptional flexibility in the sequence requirements for coronavirus small envelope protein function', *J Virol*, 81: 2249-62. 10.1128/JVI.01577-06.

Kuo, L., and P. S. Masters. 2003. 'The small envelope protein E is not essential for murine coronavirus replication', *J Virol*, 77: 4597-608. 10.1128/jvi.77.8.4597-4608.2003.

Kwan, A. H., M. Mobli, P. R. Gooley, G. F. King, and J. P. Mackay. 2011. 'Macromolecular NMR spectroscopy for the non-spectroscopist', *FEBS J*, 278: 687-703. 10.1111/j.1742-4658.2011.08004.x.

Lai, M. M., and S. A. Stohlman. 1981. 'Comparative analysis of RNA genomes of mouse hepatitis viruses', *J Virol*, 38: 661-70.

Lee, W., M. Tonelli, and J. L. Markley. 2015. 'NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy', *Bioinformatics*, 31: 1325-7. 10.1093/bioinformatics/btu830.

Li, F., W. Li, M. Farzan, and S. C. Harrison. 2005. 'Structure of SARS coronavirus spike receptor-binding domain complexed with receptor', *Science*, 309: 1864-8. 10.1126/science.1116480.

Li, W., Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, and L. F. Wang. 2005. 'Bats are natural reservoirs of SARS-like coronaviruses', *Science*, 310: 676-9. 10.1126/science.1118391.

Li, Y., W. Surya, S. Claudine, and J. Torres. 2014. 'Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins', *J Biol Chem*, 289: 12535-49. 10.1074/jbc.M114.560094.

Li, Y., J. To, C. Verdia-Baguena, S. Dossena, W. Surya, M. Huang, M. Paulmichl, D. X. Liu, V. M. Aguilella, and J. Torres. 2014. 'Inhibition of the human respiratory syncytial virus small hydrophobic protein and structural variations in a bicelle environment', *J Virol*, 88: 11899-914. 10.1128/JVI.00839-14.

Liang, B., J. H. Bushweller, and L. K. Tamm. 2006. 'Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy', *J Am Chem Soc*, 128: 4389-97. 10.1021/ja0574825.

Lopez, L. A., A. Jones, W. D. Arndt, and B. G. Hogue. 2006. 'Subcellular localization of SARS-CoV structural proteins', *Adv Exp Med Biol*, 581: 297-300. 10.1007/978-0-387-33012-9_51.

Lopez, L. A., A. J. Riffle, S. L. Pike, D. Gardner, and B. G. Hogue. 2008. 'Importance of conserved cysteine residues in the coronavirus envelope protein', *J Virol*, 82: 3000-10. 10.1128/JVI.01914-07.

Machamer, C. E., and S. Youn. 2006. 'The transmembrane domain of the infectious bronchitis virus E protein is required for efficient virus release', *Adv Exp Med Biol*, 581: 193-8. 10.1007/978-0-387-33012-9_33.

Martin-Garcia, J. M., D. T. Hansen, J. Zook, A. V. Loskutov, M. D. Robida, F. M. Craciunescu, K. F. Sykes, R. M. Wachter, P. Fromme, and J. P. Allen. 2014. 'Purification and biophysical characterization of the CapA membrane protein FTT0807 from Francisella tularensis', *Biochemistry*, 53: 1958-70. 10.1021/bi401644s.

Matsuyama, S., N. Nao, K. Shirato, M. Kawase, S. Saito, I. Takayama, N. Nagata, T. Sekizuka, H. Katoh, F. Kato, M. Sakata, M. Tahara, S. Kutsuna, N. Ohmagari, M. Kuroda, T. Suzuki, T. Kageyama, and M. Takeda. 2020. 'Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells', *Proc Natl Acad Sci U S A*, 117: 7001-03. 10.1073/pnas.2002589117.

Maxmen, A. 2017. 'Bats are global reservoir for deadly coronaviruses', *Nature*, 546: 340. 10.1038/nature.2017.22137.

McDermott, A. 2009. 'Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR', *Annu Rev Biophys*, 38: 385-403. 10.1146/annurev.biophys.050708.133719.

Miles, A. J., and B. A. Wallace. 2016. 'Circular dichroism spectroscopy of membrane proteins', *Chem Soc Rev*, 45: 4859-72. 10.1039/c5cs00084j.

Morrison, E. A., and K. A. Henzler-Wildman. 2012. 'Reconstitution of integral membrane proteins into isotropic bicelles with improved sample stability and expanded lipid composition profile', *Biochim Biophys Acta*, 1818: 814-20. 10.1016/j.bbamem.2011.12.020.

Mowrey, D. D., T. Cui, Y. Jia, D. Ma, A. M. Makhov, P. Zhang, P. Tang, and Y. Xu. 2013. 'Open-channel structures of the human glycine receptor alpha1 full-length transmembrane domain', *Structure*, 21: 1897-904. 10.1016/j.str.2013.07.014.

Neumoin, A., L. S. Cohen, B. Arshava, S. Tantry, J. M. Becker, O. Zerbe, and F. Naider. 2009. 'Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles', *Biophys J*, 96: 3187-96. 10.1016/j.bpj.2009.01.012.

Nieva, J. L., V. Madan, and L. Carrasco. 2012. 'Viroporins: structure and biological functions', *Nat Rev Microbiol*, 10: 563-74. 10.1038/nrmicro2820.

Notredame, C., D. G. Higgins, and J. Heringa. 2000. 'T-Coffee: A novel method for fast and accurate multiple sequence alignment', *J Mol Biol*, 302: 205-17. 10.1006/jmbi.2000.4042.

Opella, S. J., and F. M. Marassi. 2017. 'Applications of NMR to membrane proteins', *Arch Biochem Biophys*, 628: 92-101. 10.1016/j.abb.2017.05.011.

OuYang, B., and J. J. Chou. 2014. 'The minimalist architectures of viroporins and their therapeutic implications', *Biochim Biophys Acta*, 1838: 1058-67. 10.1016/j.bbamem.2013.09.004.

OuYang, B., S. Xie, M. J. Berardi, X. Zhao, J. Dev, W. Yu, B. Sun, and J. J. Chou. 2013. 'Unusual architecture of the p7 channel from hepatitis C virus', *Nature*, 498: 521-5. 10.1038/nature12283.

Oxenoid, K., and J. J. Chou. 2005. 'The structure of phospholamban pentamer reveals a channel-like architecture in membranes', *Proc Natl Acad Sci U S A*, 102: 10870-5. 10.1073/pnas.0504920102.

Park, S. H., W. S. Son, R. Mukhopadhyay, H. Valafar, and S. J. Opella. 2009. 'Phage-induced alignment of membrane proteins enables the measurement and structural analysis of residual dipolar couplings with dipolar waves and lambda-maps', *J Am Chem Soc*, 131: 14140-1. 10.1021/ja905640d.

Peeri, N. C., N. Shrestha, M. S. Rahman, R. Zaki, Z. Tan, S. Bibi, M. Baghbanzadeh, N. Aghamohammadi, W. Zhang, and U. Haque. 2020. 'The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned?', *Int J Epidemiol* 10.1093/ije/dyaa033.

Peiris, J. S., Y. Guan, and K. Y. Yuen. 2004. 'Severe acute respiratory syndrome', *Nat Med*, 10: S88-97. 10.1038/nm1143.

Peng, G., D. Sun, K. R. Rajashankar, Z. Qian, K. V. Holmes, and F. Li. 2011. 'Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor', *Proc Natl Acad Sci U S A*, 108: 10696-701. 10.1073/pnas.1104306108.

Perlman, Stanley, Thomas Gallagher, and Eric J. Snijder. 2008. *Nidoviruses* (ASM Press: Washington, DC).

Pervushin, K., R. Riek, G. Wider, and K. Wuthrich. 1997. 'Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution', *Proc Natl Acad Sci U S A*, 94: 12366-71. 10.1073/pnas.94.23.12366.

Pervushin, K., E. Tan, K. Parthasarathy, X. Lin, F. L. Jiang, D. Yu, A. Vararattanavech, T. W. Soong, D. X. Liu, and J. Torres. 2009. 'Structure and inhibition of the SARS coronavirus envelope protein ion channel', *PLoS Pathog*, 5: e1000511. 10.1371/journal.ppat.1000511.

Pervushin, Konstantin, Roland Riek, Gerhard Wider, and Kurt Wüthrich. 1998. 'Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in13C-Labeled Proteins', *Journal of the American Chemical Society*, 120: 6394-400. 10.1021/ja980742g.

Poget, S. F., S. M. Cahill, and M. E. Girvin. 2007. 'Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies', *J Am Chem Soc*, 129: 2432-3. 10.1021/ja0679836.

Premkumar, A., L. Wilson, G. D. Ewart, and P. W. Gage. 2004. 'Cation-selective ion channels formed by p7 of hepatitis C virus are blocked by hexamethylene amiloride', *FEBS Lett*, 557: 99-103. 10.1016/s0014-5793(03)01453-4.

Reckel, S., D. Gottstein, J. Stehle, F. Lohr, M. K. Verhoefen, M. Takeda, R. Silvers, M. Kainosho, C. Glaubitz, J. Wachtveitl, F. Bernhard, H. Schwalbe, P. Guntert, and V. Dotsch. 2011. 'Solution NMR structure of proteorhodopsin', *Angew Chem Int Ed Engl*, 50: 11942-6. 10.1002/anie.201105648.

Romero, P. R., N. Kobayashi, J. R. Wedell, K. Baskaran, T. Iwata, M. Yokochi, D. Maziuk, H. Yao, T. Fujiwara, G. Kurusu, E. L. Ulrich, J. C. Hoch, and J. L. Markley. 2020. 'BioMagResBank (BMRB) as a Resource for Structural Biology', *Methods Mol Biol*, 2112: 187-218. 10.1007/978-1-0716-0270-6_14.

Rosenberg, M. R., L. M. Weaver, and M. G. Casarotto. 2016. 'Probing interactions of Vpu from HIV-1 with amiloride-based compounds', *Biochim Biophys Acta*, 1858: 733-9. 10.1016/j.bbamem.2015.12.028.

Ruch, T. R., and C. E. Machamer. 2012. 'The coronavirus E protein: assembly and beyond', *Viruses*, 4: 363-82. 10.3390/v4030363.

Rule, Gordon S., and T. Kevin Hitchens. 2006. *Fundamentals of protein NMR spectroscopy* (Springer: Dordrecht).

Salvador, Pedro. 2014. "Annual reports on NMR spectroscopy." In *Chapter Five - Dependencies of J-Couplings upon Dihedral Angles on Proteins*, 185-227. London: Academic Press. 10.1016/B978-0-12-800185-1.00005-X.

Salzmann, M., K. Pervushin, G. Wider, H. Senn, and K. Wuthrich. 1998. 'TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins', *Proc Natl Acad Sci U S A*, 95: 13585-90. 10.1073/pnas.95.23.13585.

Sawicki, S. G., D. L. Sawicki, D. Younker, Y. Meyer, V. Thiel, H. Stokes, and S. G. Siddell. 2005. 'Functional and genetic analysis of coronavirus replicase-transcriptase proteins', *PLoS Pathog*, 1: e39. 10.1371/journal.ppat.0010039.

Schanda, P., H. Van Melckebeke, and B. Brutscher. 2006. 'Speeding up three-dimensional protein NMR experiments to a few minutes', *J Am Chem Soc*, 128: 9042-3. 10.1021/ja062025p.

Schlegel, S., J. Lofblom, C. Lee, A. Hjelm, M. Klepsch, M. Strous, D. Drew, D. J. Slotboom, and J. W. de Gier. 2012. 'Optimizing membrane protein overexpression

in the Escherichia coli strain Lemo21(DE3)', *J Mol Biol*, 423: 648-59. 10.1016/j.jmb.2012.07.019.

Schnell, J. R., and J. J. Chou. 2008. 'Structure and mechanism of the M2 proton channel of influenza A virus', *Nature*, 451: 591-5. 10.1038/nature06531.

Schoeman, D., and B. C. Fielding. 2019. 'Coronavirus envelope protein: current knowledge', *Virol J*, 16: 69. 10.1186/s12985-019-1182-0.

Shen, Y., and A. Bax. 2013. 'Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks', *J Biomol NMR*, 56: 227-41. 10.1007/s10858-013-9741-y.

Shi, L., N. J. Traaseth, R. Verardi, M. Gustavsson, J. Gao, and G. Veglia. 2011. 'Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins', *J Am Chem Soc*, 133: 2232-41. 10.1021/ja109080t.

Shortle, D. 1994. 'Assignment of amino acid type in 1H-15N correlation spectra by labeling with 14N-amino acids', *J Magn Reson B*, 105: 88-90. 10.1006/jmrb.1994.1106.

Snijder, E. J., E. Decroly, and J. Ziebuhr. 2016. 'The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing', *Adv Virus Res*, 96: 59-126. 10.1016/bs.aivir.2016.08.008.

Son, W. S., S. H. Park, H. J. Nothnagel, G. J. Lu, Y. Wang, H. Zhang, G. A. Cook, S. C. Howell, and S. J. Opella. 2012. ''q-Titration' of long-chain and short-chain lipids differentiates between structured and mobile residues of membrane proteins studied in bicelles by solution NMR spectroscopy', *J Magn Reson*, 214: 111-8. 10.1016/j.jmr.2011.10.011.

Surya, W., Y. Li, and J. Torres. 2018. 'Structural model of the SARS coronavirus E channel in LMPG micelles', *Biochim Biophys Acta Biomembr*, 1860: 1309-17. 10.1016/j.bbamem.2018.02.017.

Surya, W., Y. Li, C. Verdia-Baguena, V. M. Aguilella, and J. Torres. 2015. 'MERS coronavirus envelope protein has a single transmembrane domain that forms pentameric ion channels', *Virus Res*, 201: 61-6. 10.1016/j.virusres.2015.02.023.

Teoh, K. T., Y. L. Siu, W. L. Chan, M. A. Schluter, C. J. Liu, J. S. Peiris, R. Bruzzone, B. Margolis, and B. Nal. 2010. 'The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis', *Mol Biol Cell*, 21: 3838-52. 10.1091/mbc.E10-04-0338.

Tian, Y., C. D. Schwieters, S. J. Opella, and F. M. Marassi. 2015. 'A Practical Implicit Membrane Potential for NMR Structure Calculations of Membrane Proteins', *Biophys J*, 109: 574-85. 10.1016/j.bpj.2015.06.047.

Tolman, J. R., H. M. Al-Hashimi, L. E. Kay, and J. H. Prestegard. 2001. 'Structural and dynamic analysis of residual dipolar coupling data for proteins', *J Am Chem Soc*, 123: 1416-24. 10.1021/ja002500y.

Torres, J., W. Surya, Y. Li, and D. X. Liu. 2015. 'Protein-Protein Interactions of Viroporins in Coronaviruses and Paramyxoviruses: New Targets for Antivirals?', *Viruses*, 7: 2858-83. 10.3390/v7062750.

Vabret, A., T. Mourez, S. Gouarin, J. Petitjean, and F. Freymuth. 2003. 'An outbreak of coronavirus OC43 respiratory infection in Normandy, France', *Clin Infect Dis*, 36: 985-9. 10.1086/374222.

van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout. 2004. 'Identification of a new human coronavirus', *Nat Med*, 10: 368-73. 10.1038/nm1024.

Van Horn, W. D., H. J. Kim, C. D. Ellis, A. Hadziselimovic, E. S. Sulistijo, M. D. Karra, C. Tian, F. D. Sonnichsen, and C. R. Sanders. 2009. 'Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase', *Science*, 324: 1726-9. 10.1126/science.1171716.

Vennema, H., G. J. Godeke, J. W. Rossen, W. F. Voorhout, M. C. Horzinek, D. J. Opstelten, and P. J. Rottier. 1996. 'Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes', *EMBO J*, 15: 2020-8.

Verdia-Baguena, C., J. L. Nieto-Torres, A. Alcaraz, M. L. DeDiego, J. Torres, V. M. Aguilella, and L. Enjuanes. 2012. 'Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids', *Virology*, 432: 485-94. 10.1016/j.virol.2012.07.005.

Walls, A. C., M. A. Tortorici, B. J. Bosch, B. Frenz, P. J. M. Rottier, F. DiMaio, F. A. Rey, and D. Veesler. 2016. 'Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer', *Nature*, 531: 114-17. 10.1038/nature16988.

Wang, N., J. Shang, S. Jiang, and L. Du. 2020. 'Subunit Vaccines Against Emerging Pathogenic Human Coronaviruses', *Front Microbiol*, 11: 298. 10.3389/fmicb.2020.00298.

Ward, L. D. 1985. 'Measurement of ligand binding to proteins by fluorescence spectroscopy', *Methods Enzymol*, 117: 400-14. 10.1016/s0076-6879(85)17024-2.

Warschawski, D. E., A. A. Arnold, M. Beaugrand, A. Gravel, E. Chartrand, and I. Marcotte. 2011. 'Choosing membrane mimetics for NMR structural studies of transmembrane proteins', *Biochim Biophys Acta*, 1808: 1957-74. 10.1016/j.bbamem.2011.03.016.

Whitmore, L., and B. A. Wallace. 2004. 'DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data', *Nucleic Acids Res*, 32: W668-73. 10.1093/nar/gkh371.

Wiesner, S., and R. Sprangers. 2015. 'Methyl groups as NMR probes for biomolecular interactions', *Curr Opin Struct Biol*, 35: 60-7. 10.1016/j.sbi.2015.08.010.

Wilkinson, J., G. Ewart, C. Luscombe, K. McBride, W. Ratanasuwan, M. Miller, and R. L. Murphy. 2016. 'A Phase 1b/2a study of the safety, pharmacokinetics and antiviral activity of BIT225 in patients with HIV-1 infection', *J Antimicrob Chemother*, 71: 731-8. 10.1093/jac/dkv389.

Williamson, M. P. 2013. 'Using chemical shift perturbation to characterise ligand binding', *Prog Nucl Magn Reson Spectrosc*, 73: 1-16. 10.1016/j.pnmrs.2013.02.001.

Wilson, L., P. Gage, and G. Ewart. 2006a. 'Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication', *Virology*, 353: 294-306. 10.1016/j.virol.2006.05.028.

Wilson, L., P. Gage, and G. Ewart. 2006b. 'Validation of coronavirus E proteins ion channels as targets for antiviral drugs', *Adv Exp Med Biol*, 581: 573-8. 10.1007/978-0-387-33012-9_104.

Wilson, L., C. McKinlay, P. Gage, and G. Ewart. 2004. 'SARS coronavirus E protein forms cation-selective ion channels', *Virology*, 330: 322-31. 10.1016/j.virol.2004.09.033.

Wishart, D. S., C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. 1995. '1H, 13C and 15N chemical shift referencing in biomolecular NMR', *J Biomol NMR*, 6: 135-40. 10.1007/bf00211777.

Wittlich, M., P. Thiagarajan, B. W. Koenig, R. Hartmann, and D. Willbold. 2010. 'NMR structure of the transmembrane and cytoplasmic domains of human CD4 in micelles', *Biochim Biophys Acta*, 1798: 122-7. 10.1016/j.bbamem.2009.09.010.

Woo, P. C., S. K. Lau, C. M. Chu, K. H. Chan, H. W. Tsoi, Y. Huang, B. H. Wong, R. W. Poon, J. J. Cai, W. K. Luk, L. L. Poon, S. S. Wong, Y. Guan, J. S. Peiris, and K. Y. Yuen. 2005. 'Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia', *J Virol*, 79: 884-95. 10.1128/JVI.79.2.884-895.2005.

Wrapp, D., N. Wang, K. S. Corbett, J. A. Goldsmith, C. L. Hsieh, O. Abiona, B. S. Graham, and J. S. McLellan. 2020. 'Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation', *Science*, 367: 1260-63. 10.1126/science.abb2507.

Ye, Y., and B. G. Hogue. 2007. 'Role of the coronavirus E viroporin protein transmembrane domain in virus assembly', *J Virol*, 81: 3597-607. 10.1128/JVI.01472-06.

Ying, J., F. Delaglio, D. A. Torchia, and A. Bax. 2017. 'Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data', *J Biomol NMR*, 68: 101-18. 10.1007/s10858-016-0072-7.

Zaki, A. M., S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus, and R. A. Fouchier. 2012. 'Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia', *N Engl J Med*, 367: 1814-20. 10.1056/NEJMoa1211721.

Zhou, H. X., and T. A. Cross. 2013. 'Influences of membrane mimetic environments on membrane protein structures', *Annu Rev Biophys*, 42: 361-92. 10.1146/annurev-biophys-083012-130326.

Zhou, P., X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, H. R. Si, Y. Zhu, B. Li, C. L. Huang, H. D. Chen, J. Chen, Y. Luo, H. Guo, R. D. Jiang, M. Q. Liu, Y. Chen, X. R. Shen, X. Wang, X. S. Zheng, K. Zhao, Q. J. Chen, F. Deng, L. L. Liu, B. Yan, F. X. Zhan, Y. Y. Wang, G. F. Xiao, and Z. L. Shi. 2020. 'A pneumonia outbreak associated with a new coronavirus of probable bat origin', *Nature*, 579: 270-73. 10.1038/s41586-020-2012-7.

Zoonens, M., L. J. Catoire, F. Giusti, and J. L. Popot. 2005. 'NMR study of a membrane protein in detergent-free aqueous solution', *Proc Natl Acad Sci U S A*, 102: 8893-8. 10.1073/pnas.0503750102.