

Accepted Manuscript

Structural model of the SARS coronavirus E channel in LMPG micelles

Wahyu Surya, Yan Li, Jaume Torres



PII: S0005-2736(18)30058-0
DOI: [doi:10.1016/j.bbamem.2018.02.017](https://doi.org/10.1016/j.bbamem.2018.02.017)
Reference: BBAMEM 82711

To appear in:

Received date: 1 November 2017
Revised date: 14 February 2018
Accepted date: 16 February 2018

Please cite this article as: Wahyu Surya, Yan Li, Jaume Torres , Structural model of the SARS coronavirus E channel in LMPG micelles. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Bbamem(2018), doi:[10.1016/j.bbamem.2018.02.017](https://doi.org/10.1016/j.bbamem.2018.02.017)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Structural model of the SARS coronavirus E channel in LMPG micelles

Wahyu Surya, Yan Li and Jaume Torres

School of Biological Sciences; Nanyang Technological University; 60 Nanyang Drive, Singapore 637551; Singapore.

*Correspondence: Jaume Torres (E-mail: jtorres@ntu.edu.sg; Tel.: +65-6316-2857; Fax: +65-6791-3856)

ABSTRACT

Coronaviruses (CoV) cause common colds in humans, but are also responsible for the recent Severe Acute, and Middle East, respiratory syndromes (SARS and MERS, respectively). A promising approach for prevention are live attenuated vaccines (LAVs), some of which target the envelope (E) protein, which is a small membrane protein that forms ion channels. Unfortunately, detailed structural information is still limited for SARS-CoV E, and non-existent for other CoV E proteins. Herein, we report a structural model of a SARS-CoV E construct in LMPG micelles with, for the first time, unequivocal intermolecular NOEs. The model corresponding to the detergent-embedded region is consistent with previously obtained orientational restraints obtained in lipid bilayers and *in vivo* escape mutants. The C-terminal domain is mostly α -helical, and extramembrane intermolecular NOEs suggest interactions that may affect the TM channel conformation.

KEYWORDS

Envelope protein; solution NMR; transmembrane α -helices; micelles; oligomerization

ABBREVIATIONS

CoV, coronavirus; SARS, severe acute respiratory syndrome; MERS, Middle East respiratory syndrome; E, envelope; M, membrane; TM, transmembrane; PBM, PDZ-binding motif; IC, ion channel; LAV, live attenuated vaccine; DPC, n-dodecyl-phosphocholine; LMPG, lyso-myristoyl phosphatidylglycerol; PFO, perfluoro octanoic acid; DMPC, dimyristoyl phosphatidylcholine; HMA, hexamethylene amiloride; BN-PAGE, blue-native polyacrylamide gel electrophoresis; AQPZ, aquaporin Z; CSP, chemical shift perturbation.

1. INTRODUCTION

Coronaviruses (CoV) typically affect the respiratory tract and gut of mammals and birds. Approximately 30% of common colds are caused by two human coronaviruses - OC43 and 229E. Of particular interest are the viruses responsible for the severe acute respiratory syndrome (SARS), which produced a near pandemic in 2003 [1], and the recent Middle East respiratory syndrome coronavirus (MERS-CoV) [2].

No effective licensed treatments exist against coronavirus infections [3-5], but live attenuated vaccines (LAVs) [6-10] and fusion inhibitors [11] are promising strategies. One CoV component critical for pathogenesis is the envelope (E) protein, as reported in several coronaviruses, e.g., MERS and SARS-CoVs [12-14]. The CoV envelope (E) proteins are short polypeptides (76-109 amino acids) with a single α -helical transmembrane (TM) domain [15-21] that form homopentameric ion channels (IC) with poor ion selectivity [22, 23]. CoV E proteins are mostly found in the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) [24-29]. In animal models, deletion of SARS-CoV E protein reduced pathogenicity and mortality [30], whereas cellular models displayed up- and down-regulation of stress response and inflammation host genes, respectively [31]. The importance of E protein in pathogenesis has

led to the development of LAVs based on deletion of E protein in SARS- and MERS-CoVs, although this led to compensatory mechanisms that recover virulence [32, 33].

Specific critical features in the SARS-CoV E protein sequence have been identified that determine virulence, e.g., at the C-terminal tail [34] or in the TM domain [30], and precise structural characterization of these regions could help in the design of E protein-based CoV LAVs. However, detailed structural knowledge is still very limited in the case of SARS-CoV E, and non-existent for other CoV E proteins.

A pentameric model for SARS-CoV E was initially proposed by the authors after an *in silico* conformational search [15] of TM domain oligomers. In that report, two pentameric models (termed ‘A’ and ‘B’) that were separated by a $\sim 50^\circ$ rotation of their α -helices were selected. In model A, V25 adopts a more luminal position, whereas in model B, the position of this residue is clearly interhelical (Fig. 1). The pentameric organization of SARS-CoV E has been confirmed experimentally in various detergents: PFO, DPC or C-14 betaine [17, 18], not only for synthetic TM (E_{TM}), but also for an 8-65 (E_{TR}) construct and for full length E protein (E_{FL}).

To confirm experimentally the orientation of the α -helices in the pentameric model, site specific infrared dichroism (SSID) measurements [35] were obtained in hydrated lipid bilayers, with $^{13}C=^{18}O$ isotopically labeled synthetic E_{TM} . However, the orientation of the α -helices turned out to be strongly dependent on the presence of 2 flanking lysine residues at each end of the peptides [16]: with flanking lysine residues, the orientation was a hybrid between models A and B (residues 17–24 were oriented consistent with model B, but from residue 24 onwards, orientation was as expected for model A), consistent with a ‘bend of the α -helices around residues 25–27’ [16]. Without terminal lysines, however, the orientation of the central five labeled consecutive residues, L21 to V25, was entirely consistent with model A [17].

These initial results suggested that the conformation of the E_{TM} pentamer may be very sensitive of the presence of extra residues and probably also, extramembrane domains. An NMR study was performed on a synthetic E_{TM} (residues 8-38) in DPC detergent micelles, where E_{TM} was selectively labeled [20]. E_{TM} was ^{15}N -labeled at A22, V24, V25, and $^{13}C, ^{15}N$ -labeled at L18, L19 and L21. Intermonomeric NOEs were

assigned indirectly, i.e., when cross-peaks could not be explained by intramonomer interactions. Of these, derived from difference 2D homonuclear $^1\text{H}^{\text{N}}$, $^1\text{H}^{\text{aromatic}}$ band-selected NOESY, only four NOEs were labeled ‘strong’, and involved the $^1\text{H}^{\delta\epsilon}$ phenyl ring of Phe23, to $^1\text{H}_3^{\delta 1}$ / $^1\text{H}_3^{\delta 2}$ of either Leu18 (two NOEs) or Leu21 (two NOEs). These intermolecular NOEs were insufficient to distinguish between models A and B, and the monomer structure was fit to a model A template.

More recently, recombinant SARS-CoV escape mutants were recovered after introducing a V25F channel-inactivating mutation in the E protein, [36], that led to attenuation in a mouse model [30]. Revertant mutants regained fitness and pathogenicity whereas mutated E protein regained channel activity [30]. Surprisingly, escape mutations in E protein clustered along the helix *interface* opposite to residue V25, consistent with an interhelical orientation of this residue, as found in model B (Fig. 1, cyan).

Figure 1. Comparison of orientation of residue V25 in SARS-CoV E_{TM} pentameric models. Orientation of computational models A (orange) and B (cyan) [15], where the side chain of V25 (F26 is only used to guide the eye) is indicated. The ‘A-like’ model obtained by NMR [20] is shown in red. In model B, the position of V25 is clearly interhelical.

In the present paper, we report a more accurate model of the SARS-CoV E protein pentamer, in LMPG micelles. The construct we have used prolongs the TM domain with another 27 residues in the C-terminal domain (residues 8-65). Following established protocols [37], two types of monomers were mixed, bearing different isotopical labels, that allowed unambiguous identification of ten intermonomeric NOEs. In a nutshell, the results are consistent with a TM model that appears to be a hybrid between models A and B: while overall being closer to model A, residue V25 has a clear ‘model B-like’ interhelical orientation, consistent with the revertant mutants that appeared *in vivo*.

2. MATERIALS AND METHODS

2.1. Protein expression and purification

The expression and purification methods for the truncated SARS-CoV E construct corresponding to residues 8-65 (E_{TR}) have been described previously [19]. This construct does not have cysteines, as these are not required for oligomerization [18, 19, 28, 38]. In the present work, M9 media was supplemented with an appropriate combination of $^{15}\text{NH}_4\text{Cl}$, ^{13}C -glucose, ^2H -glucose, and $^2\text{H}_2\text{O}$ (Cambridge Isotope Laboratories) to produce ^{15}N -, ^{13}C -, $^{15}\text{N}/^{13}\text{C}$ - and $^{15}\text{N}/^2\text{H}$ -labeled E_{TR} samples. For preparation of fully deuterated $^{15}\text{N}/^2\text{H}$ -labeled samples, freshly transformed *E. coli* cells were doubly-selected in LB agar plates and media prepared with 30% and 60% $^2\text{H}_2\text{O}$, successively, and later grown in M9 media prepared with 99.9% $^2\text{H}_2\text{O}$ [39, 40].

2.2. Gel electrophoresis

Blue-native PAGE (BN-PAGE) was performed as described previously [41]. Lyophilized E_{TR} protein was solubilized (0.1 mM) in sample buffer containing LMPG (lyso-myristoyl phosphatidylglycerol, Anatrace) at the indicated concentrations.

2.3. Residue rotational pitch calculations

For α -helical bundle models, the rotational pitch angle of a residue, ω , defined arbitrarily as 0° or 180° when transition dipole moment, helix director, and the z-axis all reside in a single plane, was calculated as described elsewhere [42]. The final result is the average of the ω values calculated in each monomer. For a canonical α -helix, it is expected that $\Delta\omega$ between two consecutive residues is $\sim 100^\circ$.

2.4. NMR sample preparation

Lyophilized E_{TR} protein (0.67 mM) was solubilized in 20 mM sodium phosphate pH 5.5, 50 mM NaCl, and 200 mM LMPG, i.e., a protein:detergent (P/D) molar ratio of 1:300. The same protein concentration and P/D ratio was used for the mixture of ^{15}N -D and ^{13}C -labeled samples. The solution was vortexed and

sonicated several times until a clear solution was obtained, indicating protein reconstitution into detergent micelles.

2.5. NMR spectroscopy

NMR experiments were performed at 308K using an Avance-II 700 NMR spectrometer with cryogenic probes. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as the internal reference for ^1H nuclei. The chemical shifts of ^{13}C and ^{15}N nuclei were calculated from the ^1H chemical shifts. The NMR data were processed using TopSpin 3.1 (www.bruker-biospin.com) and analyzed using CARRA (www.nmr.ch). Sequence-specific assignment of backbone $^1\text{H}^{\text{N}}$, ^{15}N , $^{13}\text{C}'$ and $^{13}\text{C}^{\alpha}$ was achieved by using 2D [^1H - ^{15}N]-TROSY-HSQC, 3D HNCA and HN(CO)CA experiments on a $^{15}\text{N}/^{13}\text{C}$ -labeled E_{TR} protein. Side-chain resonances were assigned using 3D ^{15}N -resolved NOESY-HSQC (120 ms mixing time), (H)CCH-TOCSY and ^{13}C -resolved NOESY-HSQC (120 ms mixing time). To identify membrane-embedded residues, the NMR sample was lyophilized overnight and reconstituted in 99% D_2O . Immediately after reconstitution, 2D [^1H - ^{15}N]-TROSY was collected. The titration experiments with 5-(N,N-hexamethylene) amiloride (HMA) were performed with ^{15}N -labeled E_{TR} sample. Chemical shift perturbation (CSP) values and chemical shift differences were calculated using the formula $\text{CSP} = \sqrt{\Delta \delta\text{H}^2 + (0.23 * \Delta \delta\text{N})^2}$.

2.6. Structure calculation

Intra-monomeric NOE distance restraints were obtained from ^{15}N -NOESY-HSQC and ^{13}C -NOESY-HSQC spectra (both with a mixing time of 120 ms). Backbone dihedral angle restraints (ϕ and ψ) were derived from $^{13}\text{C}'$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^1\text{H}^{\alpha}$ and $^1\text{H}^{\beta}$ chemical shift values using TALOS+ [43]. Short-range and medium range NOE connectivities were used to establish sequence-specific ^1H NMR assignments and to identify elements of the regular secondary structure. Hydrogen bonds were derived from the NOE connectivity, and supported by the H/D exchange data. Monomer structure calculations were performed

using CYANA 3.0 [44, 45] and visualized using PyMOL (Delano Scientific). All of the restraints used in the calculations to obtain a total of 10 monomer structures, and all the structure statistics, are summarized in Supplementary Tables S1 and S2.

Inter-monomeric NOE restraints were obtained from 3D ^{15}N -resolved NOESY-HSQC (250 ms mixing time) of two sets of asymmetrically deuterated samples: (1) $^{15}\text{N}/^2\text{H}$ -labeled E_{TR} sample (ND), and (2) an equimolar mixture of $^{15}\text{N}/^2\text{H}$ -labeled and a non-deuterated ^{13}C -labeled E_{TR} sample (ND+C). NOE cross-peaks appearing in sample ND+C but not in sample ND were assigned to inter-monomeric contacts. Conversely, resonances also appearing in the ND sample were attributed to incomplete deuteration, and were assigned to intra-monomeric NOEs.

The pentamer structure was calculated using HADDOCK 2.2 [46] according to standard protocols. Ten inter-monomeric NOE restraints (defined as above) were described as ambiguous and unambiguous 5.0 Å distance restraints. Two segments were described as fully flexible: residues 37-47 and 40-54. A C5 symmetry restraint between all 5 subunits and pairwise non-crystallographic symmetry restraints between neighbouring subunits were applied. Initial rigid-body docking yielded 1000 structures, out of which 200 top-scoring structures (i.e., based on HADDOCK target function score) were selected for refinement by semi-flexible simulated annealing. These were then clustered based on RMSD, and the top-scoring cluster was selected (all 16 structures within the said cluster were grouped to form an ensemble). The corresponding structure statistics are summarized in Supplementary Table S1.

3. RESULTS AND DISCUSSION

3.1. Helical structure and TM domain of SARS-CoV E monomer (E_{TR}) in LMPG micelles

Despite phospholipid isotropic bicelles may have been more membrane-like than detergent micelles, in our hands, phospholipid bicelles did not produce suitable spectra of E_{TR} (not shown). Examples of significant differences observed in bicelles vs micelles have been reported, e.g., in the study of the integrin TM heterodimers [65-70] or in viral channels [71].

Nevertheless, we have shown previously that E_{TR} is pentameric in various detergents [17, 18], although none of them was suitable for NMR studies of E_{TR} or E_{FL} (not shown). E_{TR} only produced reasonably good NMR spectra in DPC when SDS was also present [19], but since SDS disrupts E_{TR} oligomerization, we searched for other micellar environments. Lipid-like LMPG was found to produce good NMR spectra for E_{TR} , although not for E_{FL} . Therefore, E_{TR} in LMPG was used in subsequent experiments. The use of the E_{TR} construct instead of the full-length E protein (E_{FL}) is justified since the $^{13}C\alpha$ chemical shifts of E_{TR} and E_{FL} protein in SDS or SDS/DPC were almost identical for residues 8-65 [19]. In addition, the secondary structure, obtained by CD/FTIR [18], of E_{TR} and E_{FL} is similar and predominantly α -helical, whether in DPC, SDS, mixed (1:2 molar ratio) SDS/DPC micelles or DMPC synthetic membranes [18, 19].

Figure 2. Hydrogen-deuterium exchange protected region and secondary structure of E_{TR} monomer in LMPG. (A) $[^1H-^{15}N]$ -TROSY-HSQC spectra in H_2O (left) and 99% D_2O (right), with cross-peaks labeled by one-letter code and residue number; (B) Secondary structure prediction obtained using TALOS+ [43], comparing E_{TR} in LMPG, SDS, and SDS/DPC [19]. (Layout note: 1 column)

Comparison of the HSQC spectrum of E_{TR} /LMPG before and after exposure to D_2O (Fig. 2A) shows that only 20 residues are protected from hydrogen/deuterium (H/D) exchange. The protected residues correspond to the stretch L18-L37, unequivocally indicating the presence of a *single* TM domain in SARS-CoV E. This result is consistent with the stretch L18-L39 found to be protected in SDS micelles [19]. The chemical shift index (CSI)-based secondary structure of E_{TR} (calculated by using TALOS+) obtained in LMPG (Fig. 2B), has significantly higher helicity in C-terminal residues 52-55, when compared with the data obtained SDS or with a mixture SDS/DPC [19].

Figure 3. E_{TR} consists of three α -helical segments in LMPG. (A) Ensemble of 10 calculated E_{TR} monomer structures in LMPG showing the backbone as line representation; (B) for clarity, the helical segments shown in (A) are superimposed locally and the side chains are shown as line representation (local RMSD values are listed in Supplementary Table S1); (C) graphical comparison of α -helical stretches and H/D protection (showing the TM domain) in LMPG obtained herein and in SDS/DPC environments [19]. Structure statistics in LMPG are summarized in Supplementary Table S1. (Layout note: 1.5 columns)

The structure of E_{TR} was calculated from 10 E_{TR} monomer structures (Fig. 3A) and the structure statistics are summarized in Supplementary Table S1. The E_{TR} monomer in LMPG consists of three helical segments: the one encompassing the TM domain (H1, residues 12–37), a juxtamembrane middle helical segment (H2, residues 39–47), and a C-terminal helix (H3, residues 52–65) (Fig. 3B). In contrast, E_{TR} in DPC/SDS [19] was formed by only two helical segments separated by a long flexible link (Fig. 3C). Compared to the results in SDS or SDS/DPC [19], in LMPG helix H3 is extended by 3 residues on its N-terminal side, whereas a new helical segment, H2, is formed.

3.2. Oligomeric state of SARS-CoV E in LMPG

Figure 4. Oligomeric state of SARS-CoV E in LMPG. BN-PAGE of E_{TR} in lipid-like LMPG detergent (peptide-to-detergent ratio is indicated). A ladder of oligomeric sizes is indicated by stars (*). The membrane protein aquaporin Z from *E. coli* (AqpZ) is used as reference, in monomeric and tetrameric forms (AqpZ:1 and AQPZ:4, respectively). (Layout note: 1 column)

To assess the oligomerization of E_{TR} in LMPG micelles, its migration in a BN-PAGE gel was analyzed at various protein-to-detergent (P/D) ratios (Fig. 4). At the lowest P/D molar ratio (1:1000), E_{TR} migrates as a ladder of increasingly larger oligomers where the fastest migrating band is assumed to correspond to

monomers (lower star), ~8 kDa, whereas at a high P/D ratio (1:125), E_{TR} migrates with an apparent molecular weight of ~150 kDa. These results are almost identical to those obtained previously for MERS-CoV E, for which a pentameric oligomer was determined using analytical ultracentrifugation in C-14 betaine. In that case, migration in BN-PAGE gels was also observed as a single ~150 kDa band in detergents DPC, DHPC and LMPG [21], and the ladder observed at higher detergent concentration conveniently provided an internal reference that served as a oligomeric size marker. Similar to E_{TR} , by comparison with that ladder, we confidently assigned the single band observed for MERS-CoV E to pentameric oligomers. It should be noted that in BN-PAGE gels of membrane proteins, molecular weights can appear up to 80% higher due to a contribution of the dye [72]. We have shown this for tetrameric AQPZ, which migrated at ~170 kDa instead of the expected ~100 kDa, and with a viroporin, the SH protein pentamer [41], which migrated as ~66 kDa instead of ~40 kDa. In the case of envelope E proteins, the effect is even more pronounced. In both SARS-CoV E_{TR} and MERS-CoV E, the pentameric form appears at ~150 kDa, therefore the monomer should appear at > 30 kDa. This is consistent with its migration above the AQPZ monomer (~25 kDa). The ladder ends with a pentamer, which is the predominant band at high P/D ratios. The proportion of large oligomers naturally decrease at low P/D ratios, but a significant amount of pentamer species is still present even at the 1:1000 P/D ratio. The NMR data was collected at a P/D molar ratio of 1:300, which should mostly be formed by pentamers.

3.3. HMA binding to oligomeric E_{TR}

In a previous paper, we showed that monomeric E_{TR} in SDS micelles was not affected by addition of the drug HMA [19]. However, after addition of DPC to SDS, to a SDS/DPC 1:4 molar ratio, HMA induced clear chemical shift perturbations (CSPs), concomitant with E_{TR} oligomerization. The oligomerization in DPC/SDS was not homogeneous, which precluded a more detailed study, whereas in LMPG a predominant oligomeric size is observed at a high protein-detergent ratio (Fig. 4). Therefore, in LMPG the changes observed after HMA addition should more reliably represent the binding of HMA to E_{TR} .

HMA-induced CSPs were detected herein after addition of 7.75 mM HMA to 0.25 mM E_{TR} in 200 mM LMPG micelles (P/D molar ratio 1:800) (Fig. 5).

Figure 5. E_{TR} oligomer chemical shifts perturbation (CSP) by HMA. (A) Superposition of TROSY-HSQC spectra of uniformly ¹⁵N-labeled E_{TR} protein (0.25 mM monomer concentration) in the absence (red) and presence (blue) of 7.75 mM HMA. Peaks that undergo significant shifts upon HMA addition are highlighted; (B) selected regions in the TROSY-HSQC spectrum at varying HMA concentration: 0 (red), 0.25 (pink), 0.75 (purple), 1.75 (yellow), 3.75 (light green), 7.75 (green), 9.75 (light blue), 11.75 mM HMA (blue); (C) chemical shift perturbation (CSP) of the backbone amide resonances of 0.25 mM ¹⁵N-labeled E_{TR} protein upon titration with 7.75 mM HMA. Mean CSP value across all residues and the standard deviation are shown by dashed and dotted line, respectively. Missing/overlapping residues are omitted. (Layout note: 1 column)

The average CSP value was 0.019 ppm, and several residues showed CSP more than 1 S.D. from the average value, notably Thr-9, Leu-12, Ile-13, Ala-36 and Val-47. These results suggest the presence of two binding sites located at both ends of the TM domain. Given the long distance between Ala-36 and Val-47, the two HMA-interacting residues may be located in different monomers.

3.4. Pentameric model of E_{TR}

Figure 6. Inter-monomeric NOEs in E_{TR} pentamer. (A) List of inter-monomeric NOE contacts, with those located in the extramembrane C-terminal region in bold; (B) a representative example of NOE E_{TR} inter-monomer connectivity (green lines). Selected strips correspond to a ¹⁵N-NOESY-HSQC spectrum and NH protons of V14 for samples ¹⁵N/²H-labeled (ND), ¹⁵N/²H-labeled + ¹³C-labeled (ND+C), and ¹⁵N/¹³C-labeled (NC). The NOE strips from the NC sample are shown as reference, as they contain both intra and inter-monomer contacts. Strips corresponding to the remaining NOE connectivity are shown in Figure S3. (Layout note: 1 column)

A pentameric model was obtained by docking the monomeric form of E_{TR} using HADDOCK 2.2 [46], which incorporated 10 inter-monomeric NOE restraints (Fig. 6A). We note that 2 inter-monomeric NOEs are located at the extramembrane C-terminal tail: L39 HN - Y57 HB and V47 HN - N64 HN. The same figure shows a representative example of NOE E_{TR} inter-monomer connectivity (Fig. 6B). The remaining plots of inter-monomeric NOEs are shown in Fig. S1. Structure statistics are summarized in Supplementary Table S2.

The E_{TR} pentamer is a right handed α -helical bundle where the C-terminal tails coil around each other (Fig. 7A) likely owing to the 2 inter-monomeric restraints between the two C-terminal helices. Each pentamer subunit (Fig. 7B) has better defined structure compared to the monomer alone (Fig. 3A). This is

mainly due to decreased flexibility at the inter-helical segments, which were kept flexible during the docking, as the two C-terminal helices now adopt a relatively fixed conformation. This is also apparent from the RMSD values; the pentamer subunit RMSD values are significantly reduced as compared to the monomer (Fig. 7C).

Figure 7. Structure of the E_{TR} pentamer. (A) Top view of the E_{TR} pentamer showing an ensemble of 16 structures obtained using HADDOCK and 10 inter-monomeric NOE restraints (see Materials and Methods); (B) Side view of one subunit of the pentamer showing the backbone as line representation; (C) RMSD values (per-residue) of the monomer ensemble (see Fig. 3A, black), structured helical segments of the monomer (see Fig. 3B, blue), and the pentamer ensemble (Fig. 7A, red). The average RMSD value of the monomer (dashed line) and ± 1 S.D. values (grey band) are indicated. (Layout note: 1.5 column)

Figure 8. Orientation of the E_{TR} pentamer. (A) Top view and (B) side view of average structure of the E_{TR} pentamer bundle in cartoon representation. The N- and C-terminus of one monomeric unit is indicated; (C) top view of a monomer-monomer TM interaction, showing the distances between the side chain of V25 and those of residues appearing in SARS-CoV E V25F revertant mutants [30]; (D) differences in TM residue rotational orientation, ω , between the experimental model proposed here (LMPG) versus computational models A and B [15] and that of E_{TM} obtained by NMR in DPC micelles [20]. The region with larger differences between the present model (LMPG) and model A (residues 25-28) is highlighted. (Layout note: 1.5 columns)

Notably, in this pentameric model, the location of V25 is interhelical (Fig. 8B-C), whereas in the previously proposed model it was closer to a luminal orientation [16]. The rotational pitch of the residues in the TM domain of this pentameric model were measured individually [35] and compared to those from the computational models A and B [15] (Fig. 8D). While values for residues 25-27 are closer to model B,

the rest of the sequence is similar to model A, except at residue 28 which deviates from both models. For comparison, the rotational pitch close to model A for residues in E_{TM} obtained previously by NMR in DPC micelles [20] is also shown. The present model has been constructed independently from A and B model templates, and the result appears to be a hybrid between the two [15]. This is not surprising since the *in silico* study assumed a certain rigidity in the TM α -helices [15]. Most of the residues in the model we report have an orientation consistent with model A. This is not surprising, since model A had the lowest energy value for each individual E protein homologs [15]. However, the model gets closer to model B in the turn that contains V25 (Fig. 8D). This enables V25 to adopt a more interhelical orientation consistent with the revertant mutants that appeared *in vivo* [30]. Additionally, the helix kink region suggested by infrared dichroism data in lipid bilayers [16] is also observed, which supports the validity of the membrane-mimic environment used herein.

Finally, in LMPG micelles, the C-terminal tail of SARS-CoV E protein is α -helical, more so than observed in mixed DPC/SDS micelles [19], and the presence of extramembrane NOEs suggest interactions between the C-terminal domains that may affect the pentameric conformation.

ACCESSION NUMBERS

The atomic coordinates have been deposited in the Protein Data Bank (PDB ID: 5X29). Assigned chemical shifts have been deposited at the Biological Magnetic Resonance Bank (BMRB ID: 36049).

ACKNOWLEDGEMENTS

J.T. acknowledges the funding from Singapore MOE Tier 1 grant RG 51/13. The authors declare no conflict of interest.

REFERENCES

[1] K.V. Holmes, SARS coronavirus: a new challenge for prevention and therapy, *J Clin Invest*, 111 (2003) 1605-1609.

- [2] V.S. Raj, A.D.M.E. Osterhaus, R.A.M. Fouchier, B.L. Haagmans, MERS: emergence of a novel human coronavirus, *Curr Opin Virol*, 5 (2014) 58-62.
- [3] Z. Lou, Y. Sun, Z. Rao, Current progress in antiviral strategies, *Trends Pharmacol Sci*, 35 (2014) 86-102.
- [4] A. Kilianski, S.C. Baker, Cell-based antiviral screening against coronaviruses: developing virus-specific and broad-spectrum inhibitors, *Antiviral Res*, 101 (2014) 105-112.
- [5] A. Kilianski, A.M. Mielech, X. Deng, S.C. Baker, Assessing activity and inhibition of Middle East respiratory syndrome coronavirus papain-like and 3C-like proteases using luciferase-based biosensors, *J Virol*, 87 (2013) 11955-11962.
- [6] Y.Q. Wei, H.C. Guo, H. Dong, H.M. Wang, J. Xu, D.H. Sun, S.G. Fang, X.P. Cai, D.X. Liu, S.Q. Sun, Development and characterization of a recombinant infectious bronchitis virus expressing the ectodomain region of S1 gene of H120 strain, *Appl Microbiol Biotechnol*, 98 (2014) 1727-1735.
- [7] L. Lv, X. Li, G. Liu, R. Li, Q. Liu, H. Shen, W. Wang, C. Xue, Y. Cao, Production and immunogenicity of chimeric virus-like particles containing the spike glycoprotein of infectious bronchitis virus, *J Vet Sci*, 15 (2014) 209-216.
- [8] R.L. Graham, M.M. Becker, L.D. Eckerle, M. Bolles, M.R. Denison, R.S. Baric, A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease, *Nat Med*, 18 (2012) 1820-+.
- [9] L. Enjuanes, J.L. Nieto-Torres, J.M. Jimenez-Guardeno, M.L. DeDiego, Recombinant Live Vaccines to Protect Against the Severe Acute Respiratory Syndrome Coronavirus, in: P.R.D.e. al. (Ed.) *Replicating vaccines*, Birkhauser Advances in Infectious Diseases, Springer Basel 2011, pp. 73-97.
- [10] M.W. Jackwood, Review of Infectious Bronchitis Virus Around the World, *Avian Dis*, 56 (2012) 634-641.
- [11] T. Heald-Sargent, T. Gallagher, Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence, *Viruses*, 4 (2012) 557-580.
- [12] E.W. Lamirande, M.L. DeDiego, A. Roberts, J.P. Jackson, E. Alvarez, T. Sheahan, W.J. Shieh, S.R. Zaki, R. Baric, L. Enjuanes, K. Subbarao, A live attenuated severe acute respiratory syndrome coronavirus is immunogenic and efficacious in golden Syrian hamsters, *J Virol*, 82 (2008) 7721-7724.
- [13] J. Netland, M.L. DeDiego, J. Zhao, C. Fett, E. Alvarez, J.L. Nieto-Torres, L. Enjuanes, S. Perlman, Immunization with an attenuated severe acute respiratory syndrome coronavirus deleted in E protein protects against lethal respiratory disease, *Virology*, 399 (2010) 120-128.
- [14] F. Almazán, M.L. DeDiego, I. Sola, S. Zuñiga, J.L. Nieto-Torres, S. Marquez-Jurado, G. Andrés, L. Enjuanes, Engineering a Replication-Competent, Propagation-Defective Middle East Respiratory Syndrome Coronavirus as a Vaccine Candidate, *mBio*, 4 (2013).

- [15] J. Torres, J. Wang, K. Parthasarathy, D.X. Liu, The transmembrane oligomers of coronavirus protein E, *Biophys J*, 88 (2005) 1283-1290.
- [16] J. Torres, K. Parthasarathy, X. Lin, R. Saravanan, A. Kukol, X.L. Ding, Model of a putative pore: The pentameric α -helical bundle of SARS coronavirus E protein in lipid bilayers, *Biophys J*, 91 (2006) 938-947.
- [17] K. Parthasarathy, L. Ng, X. Lin, D.X. Liu, K. Pervushin, X. Gong, J. Torres, Structural flexibility of the pentameric SARS coronavirus envelope protein ion channel, *Biophys J*, 95 (2008) L39-41.
- [18] K. Parthasarathy, H. Lu, W. Surya, A. Vararattanavech, K. Pervushin, A. Torres, Expression and purification of coronavirus envelope proteins using a modified beta-barrel construct, *Protein Express Purif*, 85 (2012) 133-141.
- [19] Y. Li, W. Surya, S. Claudine, J. Torres, Structure of a Conserved Golgi Complex-targeting Signal in Coronavirus Envelope Proteins, *J Biol Chem*, 289 (2014) 12535-12549.
- [20] K. Pervushin, E. Tan, K. Parthasarathy, X. Lin, F.L. Jiang, D. Yu, A. Vararattanavech, W.S. Tuck, X.L. Ding, J. Torres, Structure and inhibition of the SARS coronavirus envelope protein ion channel, *PLoS Pathogens*, 5 (2009).
- [21] W. Surya, Y. Li, C. Verdià-Bàguena, V.M. Aguilera, J. Torres, MERS coronavirus envelope protein has a single transmembrane domain that forms pentameric ion channels, *Virus research*, 201 (2015) 61-66.
- [22] C. Verdia-Baguena, J.L. Nieto-Torres, A. Alcaraz, M.L. DeDiego, J. Torres, V.M. Aguilera, L. Enjuanes, Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids, *Virology*, 432 (2012) 485-494.
- [23] L. Wilson, P. Gage, G. Ewart, Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication, *Virology*, 353 (2006) 294-306.
- [24] Y. Liao, Q. Yuan, J. Torres, J.P. Tam, D.X. Liu, Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein, *Virology*, 349 (2006) 264-275.
- [25] F.Y.T. Tung, S. Abraham, M. Sethna, S.-L. Hung, P. Sethna, B.G. Hogue, D.A. Brian, The 9-kDa hydrophobic protein encoded at the 3' end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated, *Virology*, 186 (1992) 676-683.
- [26] E. Corse, C.E. Machamer, The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction, *Virology*, 312 (2003) 25-34.
- [27] M.J.B. Raamsman, J. Krijnse Locker, A. De Hooge, A.A.F. De Vries, G. Griffiths, H. Vennema, P.J.M. Rottier, Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E, *J Virol*, 74 (2000) 2333-2342.

- [28] L.A. Lopez, A.J. Riffle, S.L. Pike, D. Gardner, B.G. Hogue, Importance of conserved cysteine residues in the coronavirus envelope protein, *J Virol*, 82 (2008) 3000-3010.
- [29] J.L. Nieto-Torres, M.L. DeDiego, E. Alvarez, J.M. Jimenez-Guardeno, J.A. Regla-Nava, M. Llorente, L. Kremer, S. Shuo, L. Enjuanes, Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein, *Virology*, 415 (2011) 69-82.
- [30] M.L. DeDiego, J.L. Nieto-Torres, J.A. Regla-Nava, J.M. Jimenez-Guardeño, R. Fernandez-Delgado, C. Fett, C. Castaño-Rodríguez, S. Perlman, L. Enjuanes, Inhibition of NF- κ B-mediated inflammation in severe acute respiratory syndrome coronavirus-infected mice increases survival, *J Virol*, 88 (2014) 913-924.
- [31] M.L. DeDiego, J.L. Nieto-Torres, J.M. Jiménez-Guardeño, J.A. Regla-Nava, E. Álvarez, J.C. Oliveros, J. Zhao, C. Fett, S. Perlman, L. Enjuanes, Severe acute respiratory syndrome coronavirus envelope protein regulates cell stress response and apoptosis, *PLoS Pathogens*, 7 (2011).
- [32] J.M. Jimenez-Guardeno, J.A. Regla-Nava, J.L. Nieto-Torres, M.L. DeDiego, C. Castano-Rodriguez, R. Fernandez-Delgado, S. Perlman, L. Enjuanes, Identification of the Mechanisms Causing Reversion to Virulence in an Attenuated SARS-CoV for the Design of a Genetically Stable Vaccine, *PLoS Pathog*, 11 (2015) e1005215.
- [33] J.M. Jimenez-Guardeno, J.L. Nieto-Torres, M.L. DeDiego, J.A. Regla-Nava, R. Fernandez-Delgado, C. Castano-Rodriguez, L. Enjuanes, The PDZ-Binding Motif of Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Is a Determinant of Viral Pathogenesis, *PLoS Pathogens*, 10 (2014).
- [34] J.A. Regla-Nava, J.L. Nieto-Torres, J.M. Jimenez-Guardeño, R. Fernandez-Delgado, C. Fett, C. Castaño-Rodríguez, S. Perlman, L. Enjuanes, M.L. De Diego, Severe acute respiratory syndrome coronaviruses with mutations in the E protein are attenuated and promising vaccine candidates, *J Virol*, 89 (2015) 3870-3887.
- [35] J. Torres, J.A. Briggs, I.T. Arkin, Multiple site-specific infrared dichroism of CD3-zeta, a transmembrane helix bundle, *J Mol Biol*, 316 (2002) 365-374.
- [36] J. Torres, U. Maheswari, K. Parthasarathy, L. Ng, X.L. Ding, X. Gong, Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein, *Protein Science*, 16 (2007) 2065-2071.
- [37] B. OuYang, S. Xie, M.J. Berardi, X. Zhao, J. Dev, W. Yu, B. Sun, J.J. Chou, Unusual architecture of the p7 channel from hepatitis C virus, *Nature*, 498 (2013) 521-525.
- [38] M. Godet, R. L'Haridon, J.-F. Vautherot, H. Laude, TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions, *Virology*, 188 (1992) 666-675.

- [39] A. Sivashanmugam, V. Murray, C.X. Cui, Y.H. Zhang, J.J. Wang, Q.Q. Li, Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*, *Protein Science*, 18 (2009) 936-948.
- [40] V. Tugarinov, V. Kanelis, L.E. Kay, Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy, *Nat. Protocols*, 1 (2006) 749-754.
- [41] S.W. Gan, A. Vararattanavech, N. Nordin, S. Eshaghi, J. Torres, A cost-effective method for simultaneous homo-oligomeric size determination and monodispersity conditions for membrane proteins, *Anal Biochem*, 416 (2011) 100-106.
- [42] I.T. Arkin, K.R. MacKenzie, A.T. Brunger, Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers, *J Am Chem Soc*, 119 (1997) 8973-8980.
- [43] G. Cornilescu, F. Delaglio, A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR*, 13 (1999) 289-302.
- [44] P. Guntert, C. Mumenthaler, K. Wuthrich, Torsion angle dynamics for NMR structure calculation with the new program DYANA, *Journal of molecular biology*, 273 (1997) 283-298.
- [45] T. Herrmann, P. Guntert, K. Wuthrich, Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA, *Journal of molecular biology*, 319 (2002) 209-227.
- [46] C. Dominguez, R. Boelens, A.M.J.J. Bonvin, HADDOCK: A Protein-Protein Docking Approach Based on Biochemical or Biophysical Information, *J Am Chem Soc*, 125 (2003) 1731-1737.
- [47] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics*, (2013) btt055.
- [48] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R.A. Broglia, PLUMED: A portable plugin for free-energy calculations with molecular dynamics, *Computer Physics Communications*, 180 (2009) 1961-1972.
- [49] J.P. Jämbeck, A.P. Lyubartsev, An extension and further validation of an all-atomistic force field for biological membranes, *Journal of chemical theory and computation*, 8 (2012) 2938-2948.
- [50] J.P. Jämbeck, A.P. Lyubartsev, Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids, *The journal of physical chemistry B*, 116 (2012) 3164-3179.
- [51] J. Domański, P.J. Stansfeld, M.S. Sansom, O. Beckstein, Lipidbook: a public repository for force-field parameters used in membrane simulations, *The Journal of membrane biology*, 236 (2010) 255-258.
- [52] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J.L. Klepeis, R.O. Dror, D.E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field, *Proteins: Structure, Function, and Bioinformatics*, 78 (2010) 1950-1958.

- [53] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins: Structure, Function, and Bioinformatics*, 65 (2006) 712-725.
- [54] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *The Journal of chemical physics*, 79 (1983) 926-935.
- [55] J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics*, 23 (1977) 327-341.
- [56] B. Hess, H. Bekker, H.J. Berendsen, J.G. Fraaije, LINCS: a linear constraint solver for molecular simulations, *Journal of computational chemistry*, 18 (1997) 1463-1472.
- [57] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems, *The Journal of chemical physics*, 98 (1993) 10089-10092.
- [58] D.J. Evans, B.L. Holian, The nose–hoover thermostat, *The Journal of chemical physics*, 83 (1985) 4069-4074.
- [59] M. Parrinello, A. Rahman, Crystal structure and pair potentials: A molecular-dynamics study, *Physical Review Letters*, 45 (1980) 1196.
- [60] F. Pietrucci, A. Laio, A collective variable for the efficient exploration of protein beta-sheet structures: application to SH3 and GB1, *Journal of chemical theory and computation*, 5 (2009) 2197-2201.
- [61] W.J. Allen, J.A. Lemkul, D.R. Bevan, GridMAT-MD: A grid-based membrane analysis tool for use with molecular dynamics, *Journal of computational chemistry*, 30 (2009) 1952-1958.
- [62] N. Kučerka, M.A. Kiselev, P. Balgavý, Determination of bilayer thickness and lipid surface area in unilamellar dimyristoylphosphatidylcholine vesicles from small-angle neutron scattering curves: a comparison of evaluation methods, *European Biophysics Journal*, 33 (2004) 328-334.
- [63] N. Kučerka, M.-P. Nieh, J. Katsaras, Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature, *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1808 (2011) 2761-2771.
- [64] J. Rubenstein, B.A. Smith, H.M. McConnell, Lateral diffusion in binary mixtures of cholesterol and phosphatidylcholines, *Proceedings of the National Academy of Sciences*, 76 (1979) 15-18.
- [65] T.L. Lau, C. Kim, M.H. Ginsberg, T.S. Ulmer, The structure of the integrin α IIb β 3 transmembrane complex explains integrin transmembrane signalling, *The EMBO journal*, 28 (2009) 1351-1361.
- [66] W. Surya, Y. Li, O. Millet, T. Diercks, J. Torres, Transmembrane and Juxtamembrane Structure of α L Integrin in Bicelles, *PLoS One*, 8 (2013) e74281.

- [67] T.L. Lau, V. Dua, T.S. Ulmer, Structure of the integrin α IIb transmembrane segment, *J Biol Chem*, 283 (2008) 16162-16168.
- [68] R. Li, C.R. Babu, J.D. Lear, A.J. Wand, J.S. Bennett, W.F. DeGrado, Oligomerization of the integrin α IIb β 3: roles of the transmembrane and cytoplasmic domains, *Proceedings of the National Academy of Sciences of the United States of America*, 98 (2001) 12462-12467.
- [69] J.-E. Suk, A.J. Situ, T.S. Ulmer, Construction of Covalent Membrane Protein Complexes and High-Throughput Selection of Membrane Mimics, *J Am Chem Soc*, 134 (2012) 9030-9033.
- [70] C. Lai, X. Liu, C. Tian, F. Wu, Integrin α 1 has a long helix, extending from the transmembrane region to the cytoplasmic tail in detergent micelles, *PLoS One*, 8: e62954. doi:10.1371/journal.pone.0062954 (2013).
- [71] S.W. Gan, E. Tan, X. Lin, D. Yu, J. Wang, G. Tan, A. Vararattanavech, C.Y. Yeo, C.H. Soon, T.W. Soong, K. Pervushin, J. Torres, The Small Hydrophobic Protein Of The Human Respiratory Syncytial Virus Forms Pentameric Ion Channels, *J Biol Chem*, 287 (2012) 24671-24689.
- [72] E.H. Heuberger, L.M. Veenhoff, R.H. Duurkens, R.H. Friesen, B. Poolman, Oligomeric state of membrane transport proteins analyzed with blue native electrophoresis and analytical ultracentrifugation, *J Mol Biol*, 317 (2002) 591-600.

HIGHLIGHTS

- SARS-CoV E protein is a pentameric ion channel
- The SARS-CoV E protein (8-65) is almost completely α -helical in LMPG micelles
- Ten inter-monomeric NOEs have been identified
- The SARS-CoV E protein (8-65) pentameric model has been obtained
- Orientation of key residues, e.g. Val25 is consistent with previous *in vivo* results

ACCEPTED MANUSCRIPT