 doi:10.20944/preprints202206.0201.v1

Article

The Flexible, Extended Alpha Helix Coil of the PDZ-Binding Motif of the Three Deadly Human Coronavirus E Proteins Plays a Role in Pathogenicity

Dewald Schoeman [1,†], Ruben Cloete [2,†] and Burtram C. Fielding [1,*]

[1] Molecular Biology and Virology Research Laboratory, Department of Medical Biosciences, University of the Western Cape, Cape Town, South Africa
[2] South African Medical Research Council Bioinformatics Unit, South African National Bioinformatics Institute, University of the Western Cape, Cape Town, South Africa
* Correspondence: bfielding@uwc.ac.za
† Made an equal contribution.

Abstract: The less virulent human (h) coronaviruses (CoVs) 229E, NL63, OC43, and HKU1 cause mild, self-limiting respiratory tract infections, while the more virulent SARS-CoV-1, MERS-CoV, and SARS-CoV-2 have caused severe outbreaks. The CoV envelope (E) protein, an important contributor to the pathogenesis of severe hCoVs infections, may provide insight into this disparate severity of the disease. We, therefore, generated full-length E protein models for SARS-CoV-1, -2, MERS-CoV, HCoV-229E, and HCoV-NL63 and docked C-terminal peptides of each model to the PDZ domain of the human PALS1 protein. The PDZ-binding motif (PBM) of the SARS-CoV-1, -2, and MERS-CoV models adopted a more flexible, extended coil while the HCoV-229E and HCoV-NL63 models adopted a less flexible alpha helix. All the E peptides docked to PALS1 occupied the same binding site and the more virulent hCoV E peptides generally interacted more stably with PALS1 than the less virulent ones. We propose that the increased flexibility of the PBM in more virulent hCoVs may permit more stable binding to various host proteins, thereby possibly contributing to more severe disease. This is the first paper to model full-length 3D structures for both more virulent and less virulent hCoVs E proteins, providing novel insights for possible drug and/or vaccine development.

Keywords: human coronaviruses; envelope protein; PDZ-binding motif (PBM), homology-based modelling; docking; HADDOCK; protein-protein interaction; PALS1; pathogenesis; SARS-CoV-2

1. Introduction

Of the seven human (h) coronaviruses (CoVs) identified, the recent three, severe acute respiratory syndrome (SARS)-CoV-1, Middle East respiratory syndrome (MERS)-CoV, and SARS-CoV-2, are the most virulent and have caused severe outbreaks in the last two decades [1,2]. The remaining four hCoVs, HCoV-229E, -NL63, -OC43, and -HKU1, are less virulent and circulate continuously within the human population, peaking seasonally in different countries all over the world [3-10]. While these four generally cause less severe acute respiratory tract infections, such as the common cold, in immunocompetent persons, they can lead to more severe outcomes in immunocompromised persons, the elderly, and those with chronic, underlying medical conditions [6,11,12]. The disease severity and immunopathology of SARS-CoV-1 and -2 infections have largely been associated with significant increases in inflammatory cytokine (IL-1β, IL-6, and TNF-α) production, which has been linked to pathological features such as pulmonary oedema, occasional lung epithelial damage, diffuse alveolar damage (DAD), and can culminate in lethal acute respiratory distress syndrome (ARDS) [13-18]. It is unclear, though, why HCoVs 229E, NL63, OC43, and HKU1 generally cause less severe disease in immunocompetent persons than the SARS-CoV-1, -2, and MERS-CoV.

 doi:10.20944/preprints202206.0201.v1

Clinical and pathological features such as DAD, accompanied by occasional extensive damage to lung epithelium, and dissemination of the virus to extrapulmonary tissues reported in SARS patients gave rise to the possibility that SARS-CoV-1 viral proteins could potentially disrupt pulmonary epithelial architecture. Over the last two decades it has become increasingly apparent that such features, often seen in severe SARS-CoV-1 and -2 infections, can be attributed to two properties of the hCoV envelope (E) protein: its ion-channel (IC) activity, and its postsynaptic density protein 95 (PSD95)/Drosophila disc large tumor suppressor (Dlg1)/zonula occludens-1 protein (zo-1) (PDZ)-binding motif (PBM) [18-24]. The PBM of the prototypic SARS-CoV-1 E protein can bind to the PDZ domain of several host cell proteins interacts with the PDZ domain of the intercellular tight junction-associated protein protein-associated with lin-7 1 (PALS1) and redistributes PALS1 to the ER-Golgi intermediate compartment (ERGIC), where the E protein accumulates [23]. This redistribution of PALS1 delayed the formation of tight junctions (TJs) and altered the formation of a uniform polarized monolayer of cells, leading to a model in which the E-PALS1 interaction was proposed to contribute to the observed desquamation of the alveolar wall in *post-mortem* SARS-CoV-1-infected lung tissue.

Intercellular TJs and adherence junctions (AJs) are essential in establishing and maintaining epithelial cell polarity, and by extension tissue architecture [25], which can simultaneously provide physical protection against invading pathogens, such as respiratory viruses [26,27]. Cell polarity is generally maintained by two cell complexes: the Par complex and the Crumbs (Crb) complex, the latter consisting of mammalian Lin-7 isoforms 1, 2, 3 (MALS), PALS1, and PALS1-associated TJ protein (PATJ). [28,29]. Recent studies using PBM-containing C-terminal peptides of SARS-CoV-2 E showed that SARS-CoV-2 E can interact with PALS1 in a similar way to what SARS-CoV-1 E does, also reporting that SARS-CoV-2 E had a higher affinity for PALS1 than SARS-CoV-1 E did [30-34]. Furthermore, Chai, Cai, Pang, Wang, McSweeney, Shanklin and Liu [33] experimentally determined the binding pocket on the PALS1 protein to which the SARS-CoV-2 E peptide binds, and Javorsky, Humbert and Kvansakul [34] experimentally demonstrated the residues involved in the interaction between PALS1 and peptides of the respective SARS-CoV-1, -2, and MERS-CoV E proteins also specifying the types of bonds formed between the interacting residues.

While much progress has been made in experimental structural biology and determining the structure of proteins and protein complexes, factors such as maintaining the native homeostatic environment, limits to current techniques and methodologies, and purification of protein complexes remain a challenge to generating high quality, accurate models. The complexity and largely hydrophobic nature of membrane proteins, such as the CoV E protein, also presents a challenge in studying their structure and dynamics experimentally [35-37]. In cases, such as the hCoV E protein, where experimental data is sparse, molecular modelling and docking are computational techniques that can be employed to generate reasonably accurate structural models and predict interactions between proteins, which can then be explored experimentally. Over the last 50 years, significant progress in proteomics and computational advancements have resulted in more accurate predictions of protein structures and how protein complexes interact

The 3D structural models generated in this study will provide insight into the pathogenicity of the more virulent hCoVs by predicting the secondary and tertiary structural fold of the E protein's PBM, thereby, establishing a structure-function relationship between the pathogenicity of the more virulent hCoVs VS the less virulent ones. The molecular docking will demonstrate whether the less virulent hCoVs can interact with PALS1 in a way similar to what the more virulent hCoVs and could shed light on the disparate virulence of hCoVs, as well as whether the potential PBMs of less virulent hCoVs are actually functional (Figure 1). Finally, the MD simulations may provide insight on whether all hCoVs share a uniform topology when embedded in lipid membranes, while also comparing the dynamic behavior of the E protein between the more virulent and less virulent hCoVs.

 doi:10.20944/preprints202206.0201.v1

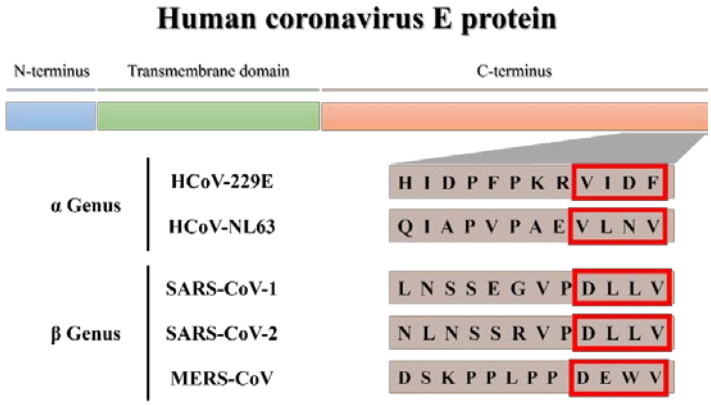

Figure 1. Partial C-terminus amino acid sequences of the human coronavirus (hCoV) envelope (E) proteins of the α- and β-CoV genera. Potential PDZ-binding motifs (PBMs) are indicated in the red blocks.

2. Materials and Methods

Molecular modelling

Template selection and model construction

Full-length three-dimensional (3D) protein models were constructed for the more virulent SARS-CoV-1, -2, MERS-CoV, and the less virulent hCoVs 229E and NL63 E proteins using MODELLER software [38,39]. This was required to perform protein-protein docking studies to better understand the basis of the hCoV E protein PBM and PALS1 interaction and determine a possible disparity between the virulence of these hCoVs.

Two partial nuclear magnetic resonance (NMR)-resolved structures for the SARS-CoV-1 E protein (PDBIDs: 5x29 and 2mm4) were obtained from the protein data bank (PDB) [40,41]. While both structures lacked some of the C-terminus, including the PBM, they spanned residues 8-65 and covered 76% of the total residues. Template 5x29 was used to generate full-length 3D models of the E proteins for SARS-CoV-1, -2, and HCoV-229E, whereas template 2mm4 was used construct the MERS-CoV model as it showed higher sequence identity to MERS-CoV compared to template 5x29, and the predicted structure of HCoV-229E was used as template for modelling the HCoV-NL63 E protein as it shared a higher amino acid sequence identity with HCoV-229E. The routinely used python script, align2d.py, in MODELLER was used to perform an alignment prediction between each of the hCoV E protein sequences and their respective template sequence [38,39]. Thereafter, a full-length 3D model was built for each hCoV E protein using the model-ligand.py script and MODELLER automatically predicted the structure of the missing C-terminal which were similar to the experimentally solved C-terminal E-peptide residues [38,39].

Quality assessment

The phi and psi dihedral angle parameters for the Ramachandran plot was calculated for each of the predicted protein models using PROCHECK webserver [42]. The root mean square deviation (RMSD) analysis was done by performing a structural alignment between the predicted model structure and the homologous template protein. This was done to assess if any structural deviation exists within the main chain atoms of the two protein structures. The predicted structures were visualized using PyMol molecular graphics software (version 2.5.2) [43].

HADDOCK: molecular docking

The experimentally resolved (X-ray) protein complex (PDBID: 7NTK) consists of the last eight residues (8-mer) of the SARS-CoV-2 E peptide (RVPDLLV) in complex with a partial PALS1 protein [34]. While the PALS1 protein contains three domains (PDZ, SH3, GK), the PALS1 protein in the 7NTK complex only contains the PDZ domain (residues

 doi:10.20944/preprints202206.0201.v1

255-336), which has been shown to be sufficient for interaction with the hCoV E protein [23,34]. The pdb file for the 7NTK complex was separated into the partial PALS1 protein and the SARS-CoV-2 E 8-mer experimental peptide. The modelled SARS-CoV-2 E protein was docked to the partial PALS1.

The HADDOCK 2.4 webserver (https://wenmr.science.uu.nl/haddock2.4/) [44] was used to dock the hCoV E protein models to the respective PALS1 protein. For docking of the E protein models, peptides consisting of eight residues (8-mers) were selected similar to Javorsky, Humbert and Kvansakul [34], while an 18 residue (18-mers) were chosen to determine if a longer peptide might affect its interaction with PALS1. The experimental 8-mer SARS-CoV-2 E peptide (from 7NTK) was docked to the partial PALS1 protein structure to validate our docking parameters to determine if HADDOCK could reproduce the experimental binding pose. We also docked 8-mer and 18-mer peptides from the E protein models. Briefly, the partial PALS1 structure and respective E peptides were uploaded to the HADDOCK webserver as input data, selecting "peptide" or "protein" as the "kind of molecule" when docking the 8-mer or 18-mer E peptides, respectively. Residues L267, G268, A269, V271, R272 (PALS1) and the last four residues of the respective E peptides were selected as input parameters for HADDOCK to specify the active residues to be used in the docking, simulating the same residues experimentally determined by Javorsky, Humbert and Kvansakul [34]. All other docking parameters were kept at default and water (TIP3P) was selected as the solvent to refine all the models.

Interaction analysis

Docked protein clusters generated by HADDOCK were selected, in PyMol, for further analysis based on how closely the cluster resembles the conformation of the SARS-CoV-2 E 8-mer peptide in the 7NTK experimentally resolved protein complex. Those clusters were submitted to PLiP and further selected based on those that reflect the most similar interacting residues and types of interactions that occur between the 7NTK PALS1 protein and SARS-CoV-2 E 8-mer peptide. Final selection of the closest resembling docked protein cluster was based on the following interacting residues and interaction types as reported by Javorsky, Humbert and Kvansakul [34]: hydrogen bonds between L267, G268, A269 (PALS1) and V75 (SARS-CoV-2 E), two hydrogen bonds between V271 (PALS1) and L73 (SARS-CoV-2 E), and an ionic interaction between R272 (PALS1) and D72 (SARS-CoV-2 E). If more than one docked cluster fulfilled these criteria, the E peptide of each cluster were superimposed on the SARS-CoV-2 E 8-mer experimental peptide and the cluster with the lowest RMSD value was selected. The RMSD value is the most commonly accepted measure used to demonstrate the similarity between the backbone atoms of two superimposed structures [45,46]. These criteria formed the basis of our system to determine the docked protein cluster with an E peptide that occupies the same binding site on PALS1, thereby resembling the 7NTK protein complex the closest (Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

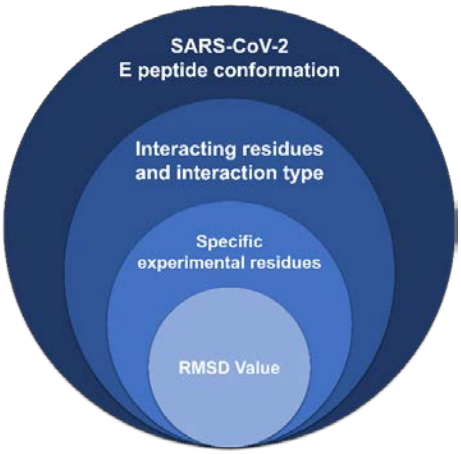

Figure 2. Selection criteria for the most accurately resembling docked protein cluster generated by the HADDOCK webserver.

PyMol and protein-ligand interaction profiler (PLiP): interaction analysis

Docked protein clusters generated by HADDOCK were visualized in PyMol (version 2.5.2) and compared to the experimental protein complex 7NTK. Only clusters that closely resembled the 7NTK complex were further analyzed. Such clusters were subsequently uploaded to protein-ligand interaction profiler (PLiP) webserver (https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index) [47] to identify the residues both involved in the interaction PALS1:hCoV E peptide as well as the type of interaction (hydrophobic, hydrogen bons, or ionic interactions) between the residues.

Molecular dynamic simulations of the hCoV E proteins in POPC lipid bilayer

Five simulation systems comprising of SARS-CoV-1, -2, MERS-CoV, hCoVs 229E and NL63 E proteins were constructed. All five simulations were performed using the GROMACS-2019 package [48] along with the CHARMM36m all-atom force field [49]. All the systems were embedded in a 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) lipid bilayer consisting of 120 lipid molecules (60 in the upper and 60 in the lower leaflet). Each system was solvated with transferable interaction potential with 3 points (TIP3P) water molecules in a cubic box of at least 10 Å of water between the protein and edges of the box at a concentration of 0.15 M. To neutralize the positive and negative charges of the systems for SARS-CoV-1, -2, MERS-CoV, hCoVs 229E and NL63, 22, 16, 29, 20, and 23 potassium (K^+) ions and 22, 18, 30, 23 and 18 chloride (Cl^-) ions were added to each system, respectively. Each of the systems underwent 50000 steps of the steepest descents energy minimization to remove close van der Waals force contacts. All five systems were subsequently subjected to a two-step equilibration phase namely, NVT (constant number of particles, volume and temperature) for 500 ps to stabilize the temperature of the system and a short position restraint NPT (constant number of particles, pressure and temperature) for 500 ps to stabilize the pressure of the system by relaxing the system and keeping the protein restrained. For the NVT simulation, the system was gradually heated by switching on the water bath and the V-rescale temperature-coupling method [50,51] was used, with constant coupling of 0.1 ps at 300 K under a random sampling seed. While for NPT the Parrinello-Rahman pressure coupling [52] was turned on with constant coupling of 0.1 ps at 300 K under conditions of position restraints (all-bonds). For both NVT and NPT, electrostatic forces were calculated using the Particle Mesh Ewald method [53]. All four systems were subjected to a full 100 ns simulation without restraints.

Trajectory analysis and principal component analysis

The analysis of the trajectory files was done using GROMACS utilities. The RMSD was calculated using gmx rmsd for the protein back bone atoms and root mean square

 doi:10.20944/preprints202206.0201.v1

fluctuation (RMSF) analysis for the protein atoms were calculated using gmx rms. The radius of gyration for the backbone atoms was calculated using the gmx gyrate tool, while the solvent accessible surface area (SASA) for the protein atoms was calculated using gmx sasa. Principal component analysis (PCA) was used as a statistical method to reduce the complexity of the data set to identify the most relevant movements of the protein. Briefly, for PCA we first calculated the covariance matrix by removing translated and rotational movement using gmx covar and then extracted the eigenvectors for the first two principal components that contributed to the largest movement of the protein using gmx anaeig.

3. Results

Homology modelling and quality assessment

More virulent hCoVs: SARS-CoV-1, -2, and MERS-CoV

The SARS-CoV-1 E protein shared 91% sequence identity with the amino acid sequence of template 5x29 (Figure S1A). The 3D model showed three α-helices and four coil regions with the C-terminal PBM adopting a coil region (Figure 3A). Quality assessment of the model revealed that 90% of the residues were in the most favored regions of the Ramachandran plot and 1.4% were in the disallowed regions. The RMSD analysis showed a 0.698Å difference between the SARS-CoV-1 E protein model and the template 5x29.

Similarly, the SARS-CoV-2 E protein shared 91% sequence identity with template 5x29 (Figure S1B), and the 3D model also showed three α-helices and four coil regions with the PBM adopting a coil region (Figure 3B). Quality assessment revealed that 88.6% of the residues were in the most favored regions of the Ramachandran plot and 0% were in the disallowed regions. The RMSD analysis showed a 2.155Å difference between the SARS-CoV-2 E protein model and template 5x29.

The MERS-CoV E protein only shared a 35% sequence identity to template 2mm4 (Figure S1C). While the 3D model showed only two α-helices and three coil regions, the PBM also adopted a coil region (Figure 3C). Quality assessment indicated that 95.8% of the residues were in the most favored regions of the Ramachandran plot and 1.4% were in the disallowed regions. The RMSD analysis between MERS-CoV E and the homologous template structure 2mm4 indicated a difference of 1.458Å.

Less virulent hCoVs: HCoV-229E and HCoV-NL63

The HCoV-229E E protein shared a 29% sequence identity with the amino acid sequence of template 5x29 (Figure S1D). The 3D model showed four α-helices and four coil regions where the C-terminal PBM adopted one of the α-helical turn (Figure 3D), and Val74 shared conservation with Val78 from template 5x29. Quality assessment of the model indicated that 91.5% of the residues were in most favored regions of the Ramachandran plot and 0% were in the disallowed regions. The RMSD analysis between the HCoV-229E E protein model and the homologous template structure 5x29 indicated a 1.776Å difference.

The HCoV-NL63 E protein shared a 47% sequence identity with the amino acid sequence of the HCoV-229E E protein and was, therefore, selected as template for model construction (Figure S1E). Similar to the HCoV-229E model, the HCoV-NL63 E 3D model showed four α-helices and four coil regions, where the PBM adopted one α-helical turn conformation (Figure 3E), and conservation of Val74 was shared with Val74 from the E protein of HCoV-229E. Quality assessment indicated that 91% of the residues were in most favored regions of the Ramachandran plot and 0% were in the disallowed regions. The RMSD analysis between the HCoV-NL63 E protein model and the homologous template structure (HCoV-229E E protein model) indicated a 1.217Å difference.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

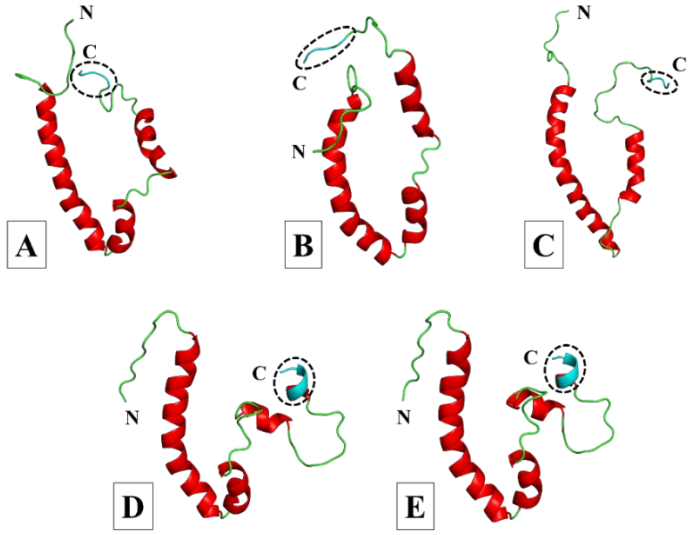

Figure 3. Cartoon representation of the three-dimensional (3D) models of the SARS-CoV-1, -2, MERS-CoV, HCoV-229E, and HCoV-NL63 envelope (E) proteins. A. SARS-CoV-1. **B.** SARS-CoV-2. **C.** MERS-CoV. **D.** HCoV-229E. **E.** HCoV-NL63. Models were generated using MODELLER software and based on the partially resolved nuclear magnetic resonance (NMR) structures for SARS-CoV-1 E (PDBIDs: 5x29 and 2mm4) obtained from the protein data bank (PDB) [38-41]. Models for SARS-CoV-1, -2 and HCoV-229E were generated from the 5x29 template, the MERS-CoV model was constructed from the 2mm4 template, and the HCoV-NL63 model was generated using the HCoV-229E E protein model as template. The amino (N) and carboxy (C)-termini are indicated accordingly, coils are shown in green, α-helices in red, and the C-terminal PDZ-binding motif (PBM) in cyan and is enclosed.

Molecular docking and interaction analysis

Docking and interaction analysis of the SARS-CoV-2 8-mer experimental E peptide to the PALS1 PDZ domain from the 7NTK complex confirmed the interactions reported by Javorsky, Humbert and Kvansakul [34], i.e. hydrogen bonds formed between the terminal PBM residue V75 and PALS1 residues L267, G268, A269, between PBM L73 and PALS1 V271, and PBM D72 and PALS1 residues V271 and S281 (Figure 4A and Table 1). The SARS-CoV-2 E PBM residues V75, L74, and L73 were reportedly involved in hydrophobic interactions but the corresponding PALS1 residues were not specified [34]. Our interaction analysis found hydrophobic interactions between the SARS-CoV-2 E PBM residue L73 and PALS1 residues V314 and F318, E PBM residue L74 and PALS1 residues T270 and V284, and E PBM residue V75 and PALS1 residues V271, F318, and L321. The PBM residues V70 and P71 also formed hydrophobic interactions with PALS1 residues R272 and V314, respectively. A single ionic interaction was found between PBM residue D72 and PALS1 R272 (Figure 4A and Table 1).

Docking of the modelled SARS-CoV-1 E 8-mer peptide largely showed similar interactions with the corresponding PALS1 residues with minor differences in the interacting residues (Figure 4B and Table 1). Hydrogen bonds were found between the modelled SARS-CoV-1 E 8-mer E terminal PBM residue V76 and PALS1 residues L267, G268, A269, between PBM L75 and PALS1 V271. As before, Javorsky, Humbert and Kvansakul [34] stated that SARS-CoV-1 E PBM residues V76, L75, and L74 were involved in hydrophobic interactions but did not specify the corresponding PALS1 residues. Our interaction analysis found hydrophobic interactions between the SARS-CoV-1 E PBM residue L74 and PALS1 residues R272, PBM residue L75 and PALS1 residues V271, F318, and L321, and PBM residue V76 and PALS1 residue P266. A single ionic interaction between PBM resi-

 doi:10.20944/preprints202206.0201.v1

due D73 and PALS1 R272 was also found (Figure 4B and Table 1). Conversely, the modelled SARS-CoV-1 18-mer E peptide adopted a position whereby more upstream residues occupied the PALS1 binding site. The E peptide residue E69 formed hydrogen bonds with PALS1 residues L267, G268, A269, while L75 on SARS-CoV-1 E also interacted with PALS1 V271 through hydrogen bonding (Figure 4C and Table 2). A single ionic interaction was formed between PBM D73 and PALS1 R272. Hydrophobic interactions were formed between the E peptide residues E69, L74, L75, and V76, and PALS1 residues P266, T270, V314, and F318 (Figure 4C and Table 2).

Docking of the modelled SARS-CoV-2 E 8-mer peptide exhibited very similar interactions with the corresponding PALS1 residues to what the experimental SARS-CoV-2 8-mer E peptide showed with PALS1. Interaction analysis revealed that hydrogen bonds were formed between the same PALS1 and modelled SARS-CoV-2 E 8-mer peptides residues as the docked experimental SARS-CoV-2 8-mer E peptide did with PALS1, except for the hydrogen bond between the modelled PBM D72 and PALS1 T270 (Figure 4D and Table 1). Very similar hydrophobic interactions were also found, except for the absence of a hydrophobic interaction between the modelled PBM residue L73 and PALS1 R272, while only the modelled PBM L73 formed a hydrophobic interaction with V314 and F318. A single ionic interaction was found between PBM D72 and PALS1 R282. The modelled SARS-CoV-2 18-mer E peptide formed the same hydrogen bonds with PALS1 as the modelled SARS-CoV-2 8-mer E peptide did (Figure 4E and Table 2). Two ionic interactions were found: one between PBM D72 and PALS1 R272, and the other more upstream between the E peptide R61 and PALS1 E274. Hydrophobic interactions were formed between PBM residue L73 and PALS1 residue V314 and F318, between PBM L74 and PALS1 T270, and PBM V75 and PALS1 V271, F318, and V321 (Figure 4E and Table 2).

The PBM of the modelled MERS-CoV 8-mer E peptide, while different in composition, demonstrated very similar binding tendencies to the docked modelled SARS-CoV-1 and -2 E peptides. It occupied the same binding site and formed hydrogen bonds between the terminal E PBM residue V82 and PALS1 residues L267, G268, A269, as well as two residues upstream between PBM E80 and PALS1 V271 (Figure 4F and Table 1). Like the modelled SARS-CoV-2 8-mer E peptide, a hydrogen bond was also formed between the MERS-CoV PBM residue D79 and PALS1 S281. Hydrophobic interactions were found between PBM E80 and PALS1 V314 and F318, between PBM W81 and PALS1 T270, and between PBM V82 and PALS1 P266, F318, and L321. The modelled 18-mer MERS-CoV E peptide largely formed the same hydrogen bonds as the modelled MERS-CoV 8-mer peptide: between PBM V82 and PALS1 L267, G268, between PBM E80 and PALS1 V271, but E peptide residue P78 also formed a hydrogen bond with R272 (Figure 4G and Table 2). Hydrophobic interactions were less than in the modelled MERS-CoV E 8-mer peptide as only PBM V82 formed hydrophobic interactions with PALS1 P266 and F318. No ionic interaction was detected but a π-stacking interaction was found between PBM W81 and F318 of PALS1 (Figure 4G and Table 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

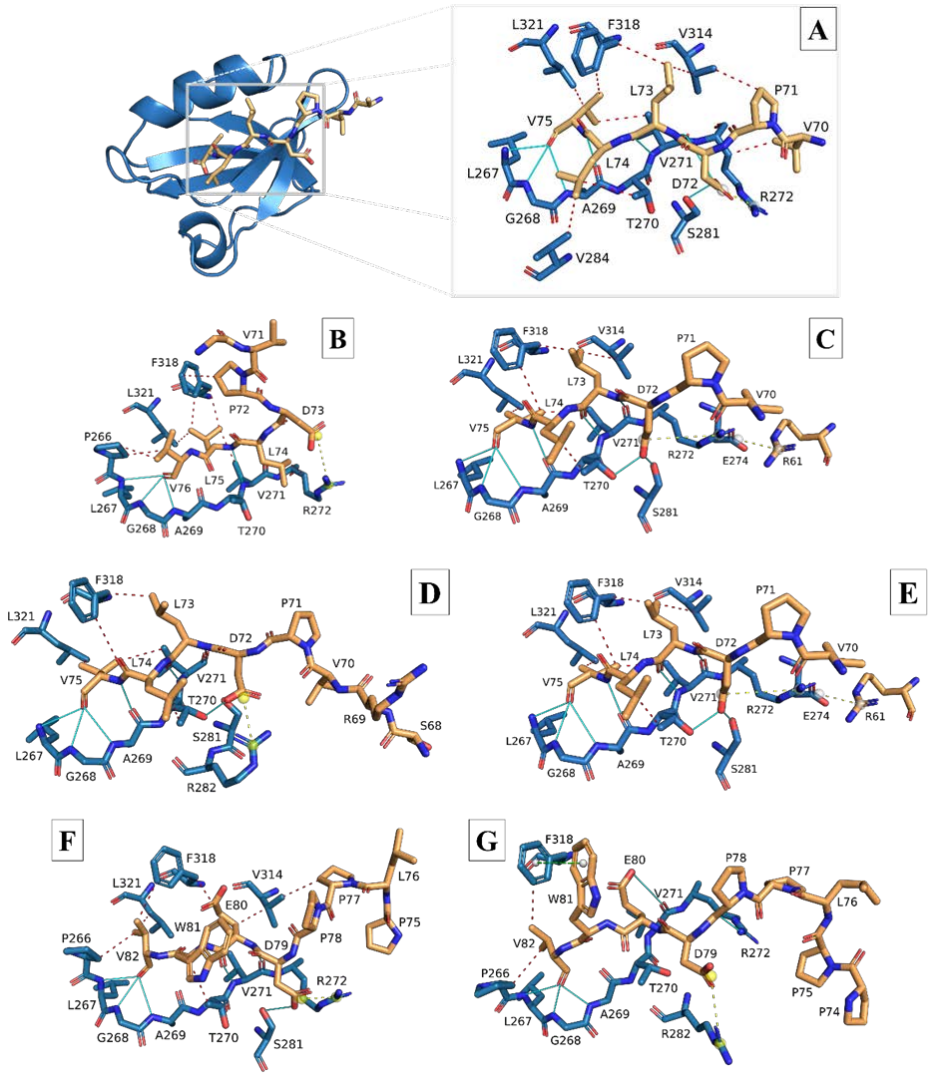

Figure 4. Stick models of the modelled 8-mer and 18-mer envelope (E) peptides of the more virulent hCoV envelope (E) proteins docked to the PALS1 PDZ domain. A. Ribbon and stick model of the SARS-CoV-2 experimental peptide (stick, gold) docked to the PDZ domain of the PALS1 protein (ribbon and stick, blue). **B.** Stick model of the modelled SARS-CoV-1 (8-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **C.** Stick model of the modelled SARS-CoV-1 (18-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **D.** Stick model of the modelled SARS-CoV-2 (8-mer) peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **E.** Stick model of the modelled SARS-CoV-2 (18-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **F.** Stick model of the modelled MERS-CoV (8-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **G.** Stick model of the modelled MERS-CoV (18-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). Peptides in B-G were produced from the homology-modelled E proteins of the respective hCoVs and docked to the PDZ domain of the PALS1 protein (PDBID: 7NTK) using the HADDOCK webserver. Docked E peptides with the lowest RMSD values compared to the 7NTK experimental E peptide were uploaded to the protein-ligand interaction profiler (PLiP) webserver to demonstrate the types of interactions between PALS1 and the E peptides. Red, dashed lines: hydrophobic interactions; Cyan, solid lines: hydrogen bonds; Yellow, dashed lines: ionic interactions (salt bridges).

The PBM of the modelled HCoV-229E 8-mer E peptide occupied the same binding site and hydrogen bonds were found only between the terminal E PBM residue F77 and PALS1 residues L267, G268, A269 (Figure 5A and Table 1). Hydrophobic interactions were found between PBM I75 and PALS1 T270, and PBM F77 and PALS1 P266, F318, and L321. An additional hydrophobic interaction occurred upstream between the E peptide residue

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

F70 and PALS1 R272. While no ionic interaction was found, residue F70, upstream of the PBM, formed a π-cation interaction with PALS1 R272. In the modelled 18-mer HCoV-229E E peptide, hydrogen bonds were found between PBM F77 and PALS1 residues L267, G268, as well as between upstream E peptide residues Y61 and PALS1 V271, and E peptide I60, Y61, and Q62 and PALS1 N315 (Figure 5B and Table 2). Hydrophobic interactions were largely found between PBM F77 and PALS1 P266, L267, F318, and L321, but also between PBM V74 and PALS1 T270, as well as E peptide Y61 and PALS1 V314. No ionic interaction was detected but a π-stacking interaction was found between the E peptide residue Y61 and PALS1 F318 (Figure 5B and Table 2).

The modelled HCoV-NL63 8-mer E peptide had a PBM that occupied the same binding site and hydrogen bonds were found between the terminal E PBM residue V77 and PALS1 residues L267, G268, A269, as well as PBM N76 and PALS1 V271 (Figure 5C and Table 1). Hydrophobic interactions were found between PBM V77 and PALS1 P266 and V284, and PBM V74 and L75 and PALS1 F318. No ionic or π-cation interactions were found. In the modelled 18-mer HCoV-NL63 E peptide, hydrogen bonds were found between PBM N76 and PALS1 residues L267, G268, as well as PBM V77 and PALS1 V271 (Figure 5D and Table 2). The PBM residue V77 formed hydrophobic interactions with PALS1 F318 and L321, and E peptide residues E73 and V74 also formed hydrophobic interactions with PALS1 F318. No ionic interaction was found but a π-stacking interaction was found between the E peptide residue Y61 and PALS1 F318 (Figure 5D and Table 2).

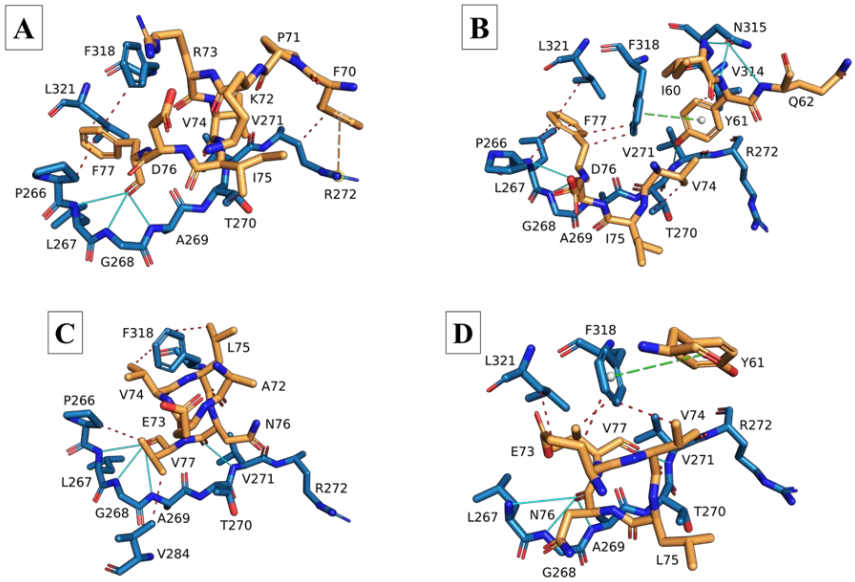

Figure 5. Stick models of the modelled 8-mer and 18-mer envelope (E) peptides of the less virulent HCoV-229E and HCoV-NL63 envelope (E) proteins docked to the PALS1 PDZ domain. A. Stick model of the modelled HCoV-229E (8-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **B.** Stick model of the modelled HCoV-229E (18-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **C.** Stick model of the modelled HCoV-NL63 (8-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). **D.** Stick model of the modelled HCoV-NL63 (18-mer) E peptide (gold) docked to the PDZ domain of the PALS1 protein (blue). Peptides in A-D were produced from the homology-modelled E proteins of the respective hCoVs and docked to the partial PALS1 protein (PDBID: 7NTK) using the HADDOCK webserver. Docked E peptides with the lowest RMSD values compared to the 7NTK experimental E peptide were uploaded to the protein-ligand interaction profiler (PLiP) webserver to demonstrate the types of interactions between PALS1 and the E peptides. Red, dashed lines: hydrophobic interactions; Cyan, solid lines: hydrogen bonds; Yellow, dashed lines: ionic interactions (salt bridges); Green, dashed lines: cation-π interactions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

Table 1. Types of interactions that occurred between the docked 8-mer hCoV E peptides and the PALS1 protein. The PALS1 PDZ domain residues to which the docked hCoV E peptides bound are shown at the top (second row) along with the corresponding PBM residue(s) of each docked hCoV E peptide. Residues are grouped according to the type of interaction that is formed between them.

HCoVs	Hydrophobic								Hydrogen						Ionic		π-Cation
	P266	T270	V271	R272	V284	V314	F318	L321	L267	G268	A269	T270	V271	S281	R272	R282	R272
7NTK (8-mer)	none	L74	V75	V70	L74	L73 P71	L73 V75	V75	V75	V75	V75 (×2)	none	L73 (×2) D72	D72	D72	none	none
SARS-CoV-1	V76 (×2)	none	L75	L74	none	none	L75 (×2) P72	L75	V76	V76	V76	none	L75	none	D73	none	none
SARS-CoV-2	none	L74	V75	none	L74	L73 (×2)	L73	V75	V75	V75	V75 (×2)	D72	L73 (×2)	D72	none	D72	none
MERS-CoV	V82	W81	none	none	none	P77 E80	E80 V82	V82	V82	V82	V82	none	E80 (×2)	D79 (×2)	D79	none	none
HCoV-229E	F77	I75	none	F70	none	none	F77	F77	F77	F77	F77	none	none	none	none	none	F70
HCoV-NL63	V77	none	none	none	V77	none	V74 L75	none	V77	V77	V77	none	N76	none	none	none	none

Table 2. Types of interactions that occurred between the docked 18-mer hCoV E peptides and the PALS1 protein. The PALS1 PDZ domain residues to which the docked hCoV E peptides bound are shown at the top (second row) along with the corresponding PBM residue(s) of each docked hCoV E peptide. Residues are grouped according to the type of interaction that is formed between them.

HCoVs	Hydrophobic									Hydrogen								Ionic			π-Stacking
	P266	L267	T270	V271	R272	V284	V314	F318	L321	P266	L267	G268	A269	T270	V271	R272	S281	N315	R272	E274	R282	F318
7NTK (8-mer)	none	none	L74	V75	V70	L74	L73 P71	L73 V75	V75	none	V75	V75	V75 (×2)	none	L73 (×2) D72	none	D72	none	D72	none	none	none
SARS-CoV-1	E69 (×2)	none	L74	none	none		L75	V76	none	S68	E69	E69	E69	none	L75	none	none	none	none	none	D73	none
SARS-CoV-2	none	none	L74	V75	none		L73	L73 V75	V75	none	V75	V75	V75 (×2)	D72	L73 (×2)	none	D72	none	D72	R61	none	none
MERS-CoV	V82	none	none	none	none		none	V82	none	none	V82	V82	V82	none	E80	P78	none	none	none	none	none	W81
HCoV-229E	F77	F77 (×2)	V74	none	none		Y61 (×2)	F77 (×2)	F77	none	F77	F77	F77	none	Y61	none	none	I60 Y61 Q62	none	none	none	Y61
HCoV-NL63	none	none	none	none	none		none	E73 V74 V77	V77	none	N76	N76	N76	none	V77	none	none	none	none	none	none	Y61

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 June 2022 doi:10.20944/preprints202206.0201.v1

Molecular dynamic simulation analysis

The RMSD analysis of the five hCoVs systems indicated that all five systems reached equilibrium after 60 ns (Figure 6A). The mean and standard deviation values for the backbone RMSD values was the largest for HCoV-229E followed by SARS-CoV-1, MERS-CoV, HCoV-NL63, and SARS-CoV-2 each having 0.94±0.21, 0.71±0.12, 0.65±0.12, 0.56±0.11 and 0.40±0.05 nm, respectively. For the RMSF analysis the largest flexibility was observed for HCoV-229E, then SARS-CoV-1, HCoV-NL63 and MERS-CoV, and while the lowest flexibility was observed for SARS-CoV-2 (Figure 6B). Additionally, the mean and standard deviation values for the protein residues was the lowest for SARS-CoV-2 followed by HCoV-NL63, MERS-CoV, SARS-CoV-1, and HCoV-229E with the largest each having 0.15±0.06, 0.27±0.13, 0.27±0.33, 0.44±0.21 and 1.27±1.48 nm, respectively. Plotting the 3D structure for the E protein of each hCoV using PyMol indicated that the PBM was flexible for SARS-CoV-1, -2, and MERS-CoV, while for the HCoV-229E and HCoV-NL63 the PBM was stable (Figure 3A-E). Furthermore, the 3D conformations of the PBM of HCoV-NL63 and HCoV-229E adopted a single α-helical conformation compared to SARS-CoV-1, -2, and MERS-CoV that has an extended PBM C-terminal coil (Figure 3A-E). The mean and standard deviation values for the radius of gyration (Rg) was the highest for systems MERS-CoV and HCoV-229E each having 2.07±0.05 and 2.04±0.10 nm followed by HCoV-NL63 and SARS-CoV-1 having 1.72±0.06 and 1.83±0.05 nm compared to SARS-CoV-2 having the lowest value 1.75±0.03 nm, respectively. Similarly, the mean and standard deviation SASA values was the highest for MERS-CoV and HCoV-229E, followed by SARS-CoV-1 and HCoV-NL63, and the lowest for SARS-CoV-2 each having 87.14±2.32, 83.95±2.34, 79.06±1.74, 77.91±1.87 nm² and 75.42±1.44 nm², respectively. The PCA analysis indicated that the SARS-CoV-2 had the smallest covariance matrix value after diagonalization of 5.93 nm² compared to HCoV-NL63, MERS-CoV, SARS-CoV-1, and HCoV-229E with the largest having 14.09, 25.49, 32.59 and 54.58 nm², respectively (Figure S2).

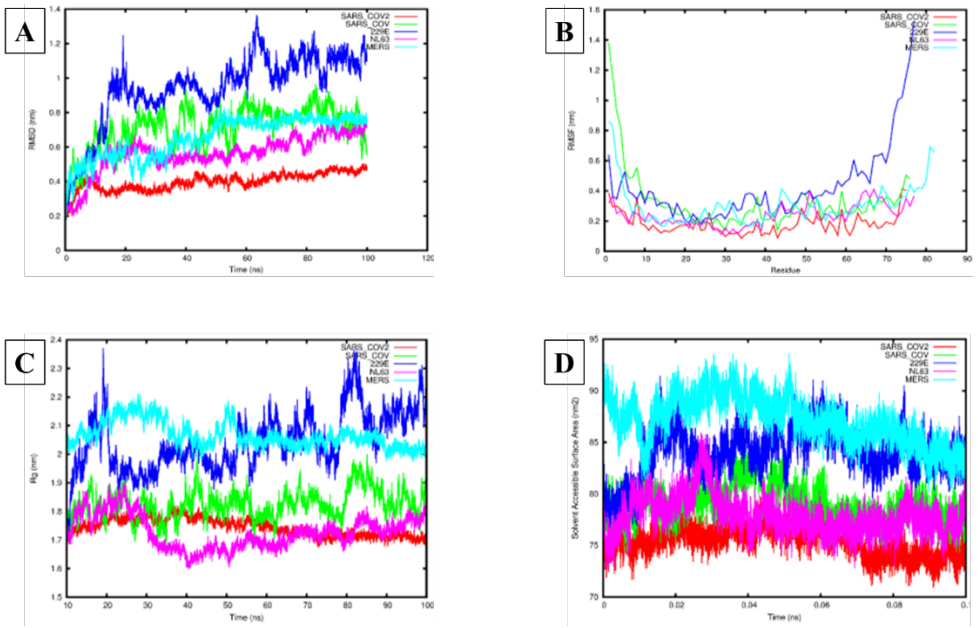

Figure 6. Analyses of molecular dynamics simulations of the human (h) coronavirus (CoV) envelope (E) proteins for SARS-CoV-1, -2, MERS-CoV, HCoV-229E, and HCoV-NL63. A. The root-mean-square deviation (RMSD) (nm) of each of the hCoV E proteins over 100 ns. **B.** The root-mean-square fluctuation (RMSF) (nm) of each residue for the hCoV E proteins. **C.** The radius of gyration (Rg) (nm) of each of the hCoV E proteins over 100 ns. **D.** The solvent-accessible surface area (SASA) (nm²) of each of the hCoV E proteins over 0.1 ns. Red: SARS-CoV-2 E protein, Green: SARS-CoV-1 E protein, Blue: HCoV-229E E protein, Pink: HCoV-NL63 E protein, Cyan: MERS-CoV E protein.

4. Discussion

 doi:10.20944/preprints202206.0201.v1

Membrane proteins, such as the hCoV E protein, play an important part in viral assembly and release of virions – aspects reliant on the protein structure and topology. [54-56]. In addition to these functions, the hCoV E protein is also a major contributor to coronaviral pathogenesis and clinical pathology, owing to its IC activity and ability to interact with host cell proteins through its PBM [18-24,57-59]. Of all the hCoV E proteins, only structures for SARS-CoV-1 and -2 have been resolved experimentally and deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) protein data bank (PDB). The PDB IDs 5x29 and 2mm4 were used as templates in this study because they are the longest experimentally resolved structure (spanning residues 8-65) but are both from SARS-CoV-1 E [60,61]. The longest SARS-CoV-2 E experimentally resolved structure only spans the transmembrane domain (TMD), residues 8-38 [62]. Therefore, 5x29 and 2mm4 were selected as templates to construct a 3D full-length model of the SARS-CoV-1 E protein, which was subsequently used to construct the 3D full-length models for the other hCoV E proteins.

Both SARS-CoV-1 and -2 E proteins shared a 91% sequence identity with the amino acid sequence of template 5x29, suggesting high sequence homology (Figures S1A and S1B). This is to be expected given the very high amino acid sequence similarity between SARS-CoV-1 and -2 E, and since the template 5x29 spans residues 8-65 of the SARS-CoV-1 E UniProt reference sequence (UniProt accession number: P59637) [17,63]. The RMSD analysis showed a 0.698Å and 2.155Å difference between SARS-CoV-1 and -2 E and the template 5x29, respectively, suggesting very little structural deviation between the respective two structures and the template. The 3D structures predicted for the SARS-CoV-1 and -2 E proteins successfully satisfied the Ramachandran plot dihedral angle distributions and based on the low RMSD values the correct protein folds were assigned to the respective protein sequences. We are, therefore, confident in the 3D full-length E proteins models constructed for SARS-CoV-1 and -2. The MERS-CoV E protein shared a 35% sequence identity to template 2mm4, suggesting a medium amount of sequence homology (Figure S1C) but this is also to be expected since the amino acid sequence similarity between SARS-CoV-1 E and MERS-CoV E has been reported to be much lower than between SARS-CoV-1 and -2 E [22,64]. Nevertheless, the RMSD value of 1.458Å calculated between the MERS-CoV 3D structure and 2mm4 showed a low deviation value suggesting high structural similarity, and still allows for reliable domain prediction. We are, therefore, confident that the correct protein folds were also assigned to this model.

Other studies have also modelled the hCoV E proteins but have focused exclusively on the more virulent hCoV E proteins and excluded the less virulent ones. Some studies have only constructed partial models of SARS-CoV-2 E using 5x29 as template [65], or SARS-CoV-1, -2, and MERS-CoV E proteins from the 2mm4 template [66]. One study, interestingly, used a full-length SARS-CoV-2 E model that was constructed from machine learning and physics-based refinement [67] but the model was not published. Only one other study constructed full-length models of SARS-CoV-1 and -2 E similar to our models and used only 2mm4 as template [30]. Thus, according to our knowledge, our study is the first to construct full-length 3D models of the E protein for both the more virulent and less virulent hCoVs. And, bar the missing residues from 5x29 and 2mm4, all five of our E protein models share exactly the same structural features as those found in these previous studies, i.e. the N-terminus forms a coiled structure, the TMD adopts an α-helical conformation, and a second α-helix (H2) after the TMD [30,66,67], further validating the accuracy of our models.

All our models adopted an α-helix in the region corresponding to the TMD, which inserts into the membrane of the ERGIC where the E protein accumulates, and coronaviral assembly occurs [68-72]. Yet, neither our MERS-CoV E model, nor the partial 2mm4-based MERS-CoV E model generated by Aldaais, Yegnaswamy, Albahrani, Alsowaiket and Alramadan [66] adopts the amphipathic H2 α-helix immediately adjacent to the TMD. This H2 α-helix is, nevertheless, also present in the SARS-CoV-1, -2, HCoV-229E and

HCoV-NL63 E proteins, suggesting that H2 might not play a role in the E protein-mediated pathogenesis but could be linked to other E protein functions, such as viral assembly. While the scarcity of MERS-CoV data and lack of an experimentally resolved E protein structure makes it difficult to confirm the significance of the missing helix at this position, *in silico* coarse-grain (CG) and atomistic simulations do suggest that the second and third α-helices (H2 and H3) of the SARS-CoV-2 E protein are important in inducing membrane curvature [67]. And since amphipathic helices (AHs) in both cellular and viral proteins have been shown to be important in membrane curvature [73-76], the lack of an H2 in the MERS-CoV E model suggests that MERS-CoV E could be structurally defective in its ability to induce membrane curvature, which would impede the assembly and release of MERS-CoV virions. Since MERS-CoV was much less effective at human-human transmission than the two other more virulent hCoVs [77-80], it's possible that this could be linked to its limited transmissibility. While the lack of this H2 would not render it incapable of infection and replication, it could impede the ability of MERS-CoV to assemble and release new virions effectively, possibly leading to fewer viruses shed by infected persons, thereby limiting its transmissibility to small, close-contact groups [79,81]. However, CoV nonstructural proteins (nsps), some of which may have overlapping functions, could also potentially compensate, at least partially, for the lack of an H2 α-helix in the MERS-CoV E protein [82], or MERS-CoV E could simply induce membrane curvature by a different mechanism. Without experimental evidence, this remains to be determined.

Despite the moderate sequence identity between the target sequence of HCoV-229E E and the 5x29 template, the protein model from this pairwise sequence alignment will provide a medium accuracy protein model [83] which is still useful for domain comparisons. The RMSD analysis between HCoV-229E E and the homologous template structure 5x29 indicated a 1.776Å difference, suggesting very little deviation between the two structures. The RMSD analysis between HCoV-NL63 E and the homologous template structure (229E protein model) similarly indicated a 1.217Å difference, also suggesting very little deviation between the generated model and the homologous template structure. Thus, the 3D structures predicted for the E proteins of HCoV-229E and HCoV-NL63 successfully passed quality assessment and based on the low RSMD values, the correct protein folds were assigned to the respective protein sequences. Therefore, despite the absence of experimentally resolved structures for the HCoV-229E and HCoV-NL63 E proteins, we are confident in the accuracy of our 3D full-length models. We also attempted to construct full-length models of the HCoV-OC43 and HCoV-HKU1 E proteins but excluded them from this study since these two E proteins shared a lower sequence identity with the other hCoV E proteins, which resulted in lower accuracy and less reliable models of the HCoV-OC43 and HCoV-HKU1 E proteins.

Comparing the structural aspects of the E protein models, the PBM of SARS-CoV-1, -2 and MERS-CoV E proteins adopted a more flexible extended coil region, while the HCoV-229E and HCoV-NL63 E PBM is characterized by a less flexible α-helical turn, suggesting that the flexibility of the E protein PBM region might be a feature distinguishing the more virulent hCoVs from the less virulent ones. So far, several host cell proteins have been reported to interact with the E protein of all three more virulent hCoVs [14,84-86] , and more may yet be identified. The conformation of the E protein PBM could quite possibly be a structural feature that allows the more virulent hCoVs to interact with a wider range of host proteins which could explain their increased pathogenicity and disease severity. By comparison, the less flexible α-helical turn of the less virulent hCoV E protein PBM could likely be an impeding characteristic, limiting or even possibly preventing interaction(s) with host proteins, which might explain the limited pathogenic capabilities of these hCoVs. This would certainly be interesting to confirm and warrants exploring experimentally.

While our system is based *in silico*, the generated 3D models and docking was based on experimentally determined, published results [34,60,61], and we validated our system by first docking the experimental SARS-CoV-2 8-mer E peptide to the PALS1 PDZ domain

 doi:10.20944/preprints202206.0201.v1

to replicate the experimentally determined interaction between SARS-CoV-2 E and PALS1. Although we only used the experimentally determined residues by Javorsky, Humbert and Kvansakul [34] to perform docking, it is very interesting that our interaction analysis detected residues P266, L267, V284, F318, L321 to varying degrees in our docked models (Figures 4 and 5) (Tables 1 and 2). We find this interesting because residues P266, L267, V284, F318, L321, L369, L403 are reported to be involved in the forming the hydrophobic pocket on PALS1 to which the SARS-CoV-2 E protein binds, and the positioning of F318 between the SARS-CoV-2 E protein's L73 and V75 is considered an important recognition feature [33]. More importantly, though, all our docked E peptides (8-mers and 18-mers) occupied the same binding site on the PALS1 PDZ domain as the experimentally determined interaction between the PALS1 PDZ domain and the 8-mer SARS-CoV-2 E peptide (Figures 4 and 5) (Tables 1 and 2) [34]. This is based on our interaction analysis in which we found overlapping features: (1) hydrogen bonds between the PALS1 residues L267, G268, A269 and the last PBM residue of each docked hCoV E peptide, as well as (2) hydrophobic interaction(s) between F318 on PALS1 and either the last PBM residue and/or two residues upstream of the PBM. This demonstrates that our docked E peptides closely resemble the experimental PALS1-E protein interaction, which supports the docking protocol for our system.

Another notable aspect is the ionic interaction formed between the aspartic acid residue at the start of the PBM of all three of the more virulent hCoV E peptides and the PALS1 residue R272/R282. Such an interaction is absent from the less virulent hCoV E peptides and is likely be due to their PBM starting with the nonpolar residue valine, as opposed to the negatively charged aspartic acid on the E protein of more virulent hCoVs. Given that ionic interactions are generally considered stronger than both hydrogen bonds and hydrophobic interactions, the presence of the negatively charged aspartic acid at this position in the PBM might also be something that allows the more virulent hCoV E proteins to form a more stable interaction with PALS1 to facilitate their increased viral pathogenicity. Indeed, the native binding partner of PALS1, Crumbs (Crb), also forms an ionic interaction with R282 on PALS1 through its negatively charged E1403 residue [29,34], similar to what we found in our docked SARS-CoV-2 E 8-mer peptide. The ERLI motif (E1403, R1404, L1405, I1406) of Crb is functionally similar to the E protein's PBM and allows Crb to occupy the same binding site on the PALS1 PDZ domain [29,32,34]. Furthermore, the last residue of the Crb ERLI motif (I1406), forms hydrogen bonds with PALS1 residues L267, G268, and A269, and its E1403 interacts with PALS1 residues T270 and S281 through hydrogen bonds [29]. Similarly, the last residue of each of our modelled E peptide PBMs also hydrogen bond with PALS1 residues L267, G268, and A269. Moreover, the aspartic acid of the more virulent SARS-CoV-2 (D72) and MERS-CoV (D79) E peptides also shows hydrogen bonding to the PALS1 S281, whereas the SARS-CoV-1 (D76) E peptide does not, and the SARS-CoV-2 E peptide (D72) is the only one to form a hydrogen bond with PALS1 T270. Using the number of bonds formed between the E peptide and PALS1 PDZ domain as a measure of stability, this could potentially explain why SARS-CoV-1 was less severe than MERS and COVID-19, and why COVID-19 appears to be more severe than both MERS and SARS diseases were. Overall, the similarity between previous studies and the types of interactions between the specific E protein and PALS1 residues further validates the accuracy of the docked E peptides of our system.

There is a rather clear difference in the number of bonds formed between the E peptides of the more virulent hCoVs and PALS1 and the bonds formed between PALS1 and the less virulent hCoV E peptides. Generally, more bonds are formed between PALS1 and the E peptides of the more virulent hCoVs, possibly producing a more stable interaction, especially since many are hydrogen bonds and even hydrophobic interactions. In comparison, the E peptides of the less virulent HCoV-229E and HCoV-NL63 form fewer hydrogen bonds and fewer hydrophobic interactions, suggesting that while they might be capable of binding to PALS1, the interaction might be less stable or more transient than in the case of the more virulent hCoV E peptides. Nevertheless, our system still suggests that

the SARS-CoV-1, -2, MERS-CoV, HCoV-229E, and HCoV-NL63 E proteins are capable of occupying the same binding site on the PALS1 PDZ domain, albeit to different extents. This could explain why the less virulent hCoVs tend to be less likely to spread to extra-pulmonary sites, thereby being less likely to cause severe disease. If the HCoV-229E and HCoV-NL63 E proteins cannot bind to PALS1 as effectively and stably as those of SARS-CoV-1, -2, and MERS-CoV can, they might not be able to displace PALS1 from the Crb-tight-junction-complex and would consequently not be able to compromise the pulmonary epithelial barrier to the severity that the more virulent hCoV E proteins can.

Two recent studies used a different technique to measure the binding affinity between E peptides of the more virulent hCoVs and the PALS1 PDZ domain and reported conflicting results. One study used C-terminal E peptides of 12 amino acids long (12-mer) and found that SARS-CoV-1 E did not bind to the PALS1 PDZ domain, while SARS-CoV-2 E only bound weakly to PALS1 [87]; the other study used 8-mer C-terminal E peptides and found that both SARS-CoV-1 and -2 E bound to the PALS1 PDZ domain with dissociation constants of 29.6±2.3 μM and 22.8±1.2 μM, respectively [34]. Both studies, however, reported that the MERS-CoV E peptide did not bind to the PALS1 PDZ domain [34,87]. Javorsky, Humbert and Kvansakul [34] proposed that the W81 of the MERS-CoV E PBM is too large and bulky to accommodate binding to the PDZ binding groove, whereas SARS-CoV-1 and -2 E, and Crb feature the smaller leucine residue that could facilitate binding to PALS1. Conversely, Caillet-Saguy, Durbesson, Rezelj, Gogl, Tran, Twizere, Vignuzzi, Vincentelli and Wolff [87] suggested that the differences in binding of SARS-CoV-1 and 2, and MERS-CoV E proteins to PALS1 are due to their difference in PBM classes. Class II PBMs, as seen in the SARS-CoV-1 and -2 E proteins, are characterized by the motif: Φ-X-Φ, where X is any amino acid and Φ is a hydrophobic residue, whereas and class III PBMs, as seen in MERS-CoV E protein, have the motif: D/E-X- Φ. The authors attributed the hydrophobic character of class II PBMs to having sharp binding profiles but lower affinities for their target host proteins, while the presence of the tryptophan residue in the class III PBM of MERS-CoV E likely strongly contributed to the binding affinity of more of its host target proteins but at the expense of specificity. While the PBMs of the E proteins for the less virulent hCoVs have not specifically been classified or characterized, the HCoV-229E and HCoV-NL63 E proteins appear to adopt a class II PBM, similar to SARS-CoV-1 and -2, despite demonstrating a lower viral pathogenicity. Although the E proteins of HCoV-229E and HCoV-NL63 potentially also have a class II PBM, their lower viral pathogenicity could also, at least in part, be due to the lower number of hydrogen bonds and hydrophobic interactions, and absence of ionic interactions between their E proteins and PALS1. Both hydrogen bonds and hydrophobic interactions have been shown to be important in protein stability and at the interface of protein-protein interactions (PPIs) [88,89]. This combination of a less flexible alpha helix in their E protein PBM could possibly reduce the number of hydrogen bonds and hydrophobic interactions, while the lack of a negatively charged residue, such as aspartic acid, prevents the formation of an ionic interaction with the PALS1 protein, ultimately negating the ability of these hCoV E proteins from binding to PALS1 [89]. While perhaps still capable of interaction with host proteins, our results suggest that the E proteins of HCoV-229E and HC-oV-NL63 is unable to bind to PALS1 as effectively as the E proteins of the more virulent hCoVs. But this should be validated experimentally.

The MD analysis of the RMSF average structures showed agreement with the predicted 3D structures for the more virulent SARS-CoV-1, -2, and MERS-CoV, and the less virulent HCoV-229E and HCoV-NL63 E proteins. Here, SARS-CoV-1, -2 and MERS-CoV had an extended coil PBM motif at the C-terminal end, while HCoV-229E and HCoV-NL63 had a stable alpha helical PBM motif domain. Furthermore, MD analysis of the entire protein dynamics showed that SARS-CoV-2 were more stable than the SARS-CoV-1, MERS-CoV and HCoV-NL63, while HCoV-229E adopted higher flexibility values suggesting that it is the least stable of the four hCoVs. This is based on RMSD, RMSF, Radius of gyration and SASA parameter calculations. The PCA analysis also provided additional

support for the protein dynamics showing reduced randomized movements of SARS-CoV-2 compared to SARS-CoV-1, MERS-CoV, HCoV-NL63 and HCoV-229E showing more concerted movements throughout the phase space.

The SARS-CoV-2 E protein displayed overall reduced dynamics with a flexible extended coil at its PBM region that could allow improved interaction with human host proteins to facilitate an increased viral pathogenicity. Therefore, based on the key interacting residues formed between PALS1 and the E proteins of the more virulent hCoVs, and the absence of many of these interactions in the E proteins of the less virulent hCoVs, we propose that the flexibility of the E protein PBM contributes to how it interacts with the PALS1 protein by virtue of the types of interactions which, in turn, facilitates how stable the PALS1-E interaction is. In this way, the PBM can be considered a potential determinant of the virulence of the hCoV, distinguishing more virulent hCoVs from less virulent ones. Although the structure of the SARS-CoV-2 E protein has been resolved experimentally, this was only a truncated portion, consisting of the TMD, which lacked the C-terminal and did not allow us to make meaningful interpretations about the full-length protein. However, this study did attempt to illuminate on the proposed mechanism of pathogenicity for a full-length protein structure including the PDB motif albeit a predictive model. Future studies could benefit from longer simulations to obtain more sampled conformations of the E protein structure in complex with human host partners. Furthermore, the experimental determination of the full-length structure of the E proteins will be a significant contribution in understanding human host protein interaction with hCoV E proteins. This study only provides a partial picture of the dynamics of the hCoV E proteins but shows good agreement between predictive modelling and simulation studies.

There are two limitations to our study. One is that docking is static and, thereby, simulates a singular pose of the interaction between the hCoV E peptide and the PALS1 protein at any given time. It does not account for any possible conformational changes that the E protein might undergo during its interaction with PALS1. The second is the length of the E peptides. Docking the full-length hCoV E protein to the full-length PALS1 protein would simulate a more accurate interaction. Nevertheless, since our docking is validated and based on the experimentally determined PALS1 interaction with E peptides, we are confident in the data that we generated. We recommend the use of the full-length E proteins for future research.

5. Conclusions

This is the first study, to predict and generate full length 3D models of the E protein for five hCoVs. This study also attempted to determine a distinguishing feature between the E proteins of the more virulent and the less virulent hCoVs to provide a possible explanation for the difference in pathogenicity and disease severity. Interaction partners have only been identified for the more virulent hCoV E proteins [87], and it remains to be seen if the less virulent hCoV E proteins can also interact with host cell proteins through their proposed PBMs, to what extent this could happen, and host cell proteins they can interact with. Therefore, we docked 8-mer and 18-mer E protein C-terminal domains of five hCoV proteins to the host PALS-1 protein to understand protein-protein interaction. We observed that SARS-CoV-2 PBM 8-mer and 18-mer peptides had stronger affinity and number of interactions with PALS-1 possibly explaining its ability to associate with human host proteins resulting in severe pathogenicity. This needs to be verified experimentally to validate *in silico* findings.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: **Pairwise sequence alignments between the virulent SARS-CoV-1, -2 and MERS-CoV, and nonvirulent HCoV-229E and HCoV-NL63 human (h) coronavirus (CoV) envelope (E) proteins and the respective template used to generate each three-dimensional (3D) model.**; S2: **Principal component analysis (PCA) indicating the covariance matrix value of**

the human (h) coronavirus (CoV) envelope (E) proteins for SARS-CoV-1, -2, MERS-CoV, HCoV-229E, and HCoV-NL63 based on the projection of two eigenvectors.

Author Contributions: BCF and DS conceptualized the research; DS performed the molecular docking and PLiP interaction analysis; RC performed the molecular modeling and MD simulations; DS and RC drafted the manuscript; All authors read, edited, and approved the final manuscript.

Funding: DS was funded by the National Research Foundation of South Africa, the Poliomyelitis Research Foundation (PRF) of South Africa [17/53 and 19/06]. RC was funded by the Department of Higher Education, next Generation of Academic Programme (nGAP) in the form of full-time academic positions and salaries. BCF received funding from the UWC Senate Research Fund and the National Research Foundation of South Africa. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors, and therefore, the NRF does not accept any liability in regard thereto.

Data Availability Statement: Not applicable.

Acknowledgments: We apologize to any author whose work has been inadvertently omitted from this article. Special thanks to the Centre for High Performance Computing (CHPC) in Rondebosch for allowing us to run our computationally intensive simulations.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Broadbent, L. Coronaviruses - a brief history. The Conversation: 2020.
2. Hewings-Martin, Y. How do SARS and MERS compare with COVID-19? Medical News Today: 2020.
3. Aldridge, R.W.; Lewer, D.; Beale, S.; Johnson, A.M.; Zambon, M.; Hayward, A.C.; Fragaszy, E.B. Seasonality and immunity to laboratory-confirmed seasonal coronaviruses (HCoV-NL63, HCoV-OC43, and HCoV-229E): results from the Flu Watch cohort study. *Wellcome Open Research* **2020**, *5*, 52.
4. Cui, L.-J.; Zhang, C.; Zhang, T.; Lu, R.-J.; Xie, Z.-D.; Zhang, L.-L.; Liu, C.-Y.; Zhou, W.-M.; Ruan, L.; Ma, X.-J. Human coronaviruses HCoV-NL63 and HCoV-HKU1 in hospitalized children with acute respiratory infections in Beijing, China. *Adv. Virol.* **2011**, *2011*.
5. Gaunt, E.R.; Hardie, A.; Claas, E.C.; Simmonds, P.; Templeton, K.E. Epidemiology and clinical presentations of the four human coronaviruses 229E, HKU1, NL63, and OC43 detected over 3 years using a novel multiplex real-time PCR method. *J. Clin. Microbiol.* **2010**, *48*, 2940-2947, doi:https://doi.org/10.1128/JCM.00636-10.
6. Killerby, M.E.; Biggs, H.M.; Haynes, A.; Dahl, R.M.; Mustaquim, D.; Gerber, S.I.; Watson, J.T. Human coronavirus circulation in the United States 2014–2017. *J. Clin. Virol.* **2018**, *101*, 52-56.
7. Lau, S.K.; Woo, P.C.; Yip, C.C.; Tse, H.; Tsoi, H.-w.; Cheng, V.C.; Lee, P.; Tang, B.S.; Cheung, C.H.; Lee, R.A. Coronavirus HKU1 and other coronavirus infections in Hong Kong. *J. Clin. Microbiol.* **2006**, *44*, 2063-2071.
8. Su, S.; Wong, G.; Shi, W.; Liu, J.; Lai, A.C.; Zhou, J.; Liu, W.; Bi, Y.; Gao, G.F. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* **2016**, *24*, 490-502, doi:https://doi.org/10.1016/j.tim.2016.03.003.
9. Zeng, Z.-Q.; Chen, D.-H.; Tan, W.-P.; Qiu, S.-Y.; Xu, D.; Liang, H.-X.; Chen, M.-X.; Li, X.; Lin, Z.-S.; Liu, W.-K., et al. Epidemiology and clinical characteristics of human coronaviruses OC43, 229E, NL63, and HKU1: a study of hospitalized children with acute respiratory tract infection in Guangzhou, China. *European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology* **2018**, *37*, 363-369, doi:10.1007/s10096-017-3144-z.
10. Edridge, A.W.D.; Kaczorowska, J.; Hoste, A.C.R.; Bakker, M.; Klein, M.; Loens, K.; Jebbink, M.F.; Matser, A.; Kinsella, C.M.; Rueda, P., et al. Seasonal coronavirus protective immunity is short-lasting. *Nat Med* **2020**, *26*, 1691, doi:10.1038/s41591-020-1083-1.
11. Trombetta, H.; Faggion, H.Z.; Leotte, J.; Nogueira, M.B.; Vidal, L.R.; Raboni, S.M. Human coronavirus and severe acute respiratory infection in Southern Brazil. *Pathogens and global health* **2016**, *110*, 113-118.

12. Liu, D.X.; Liang, J.Q.; Fung, T.S. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). *Encyclopedia of Virology* **2021**, 10.1016/B978-0-12-809633-8.21501-X, 428-440, doi:10.1016/B978-0-12-809633-8.21501-X.
13. Fett, C.; DeDiego, M.L.; Regla-Nava, J.A.; Enjuanes, L.; Perlman, S. Complete protection against severe acute respiratory syndrome coronavirus-mediated lethal respiratory disease in aged mice by immunization with a mouse-adapted virus lacking E protein. *J. Virol.* **2013**, *87*, 6551-6559, doi:https://doi.org/10.1128/JVI.00087-13.
14. Jimenez-Guardeño, J.M.; Nieto-Torres, J.L.; DeDiego, M.L.; Regla-Nava, J.A.; Fernandez-Delgado, R.; Castaño-Rodriguez, C.; Enjuanes, L. The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS Pathog.* **2014**, *10*, e1004320, doi:https://doi.org/10.1371/journal.ppat.1004320.
15. Nieto-Torres, J.L.; DeDiego, M.L.; Verdiá-Báguena, C.; Jimenez-Guardeño, J.M.; Regla-Nava, J.A.; Fernandez-Delgado, R.; Castaño-Rodriguez, C.; Alcaraz, A.; Torres, J.; Aguilella, V.M. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathog.* **2014**, *10*, e1004077, doi:https://doi.org/10.1371/journal.ppat.1004077.
16. Regla-Nava, J.A.; Nieto-Torres, J.L.; Jimenez-Guardeño, J.M.; Fernandez-Delgado, R.; Fett, C.; Castaño-Rodríguez, C.; Perlman, S.; Enjuanes, L.; DeDiego, M.L. SARS coronaviruses with mutations in E protein are attenuated and promising vaccine candidates. *J. Virol.* **2015**, https://doi.org/10.1128/JVI.03566-14, JVI. 03566-03514, doi:https://doi.org/10.1128/JVI.03566-14.
17. Schoeman, D.; Fielding, B.C. Is There a Link Between the Pathogenic Human Coronavirus Envelope Protein and Immunopathology? A Review of the Literature. *Front. Microbiol.* **2020**, *11*, 2086, doi:10.3389/fmicb.2020.02086.
18. Xia, B.; Shen, X.; He, Y.; Pan, X.; Liu, F.L.; Wang, Y.; Yang, F.; Fang, S.; Wu, Y.; Duan, Z., et al. SARS-CoV-2 envelope protein causes acute respiratory distress syndrome (ARDS)-like pathological damages and constitutes an antiviral target. *Cell Research* **2021**, *31*, 847-860, doi:10.1038/s41422-021-00519-4.
19. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X., et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497-506, doi:10.1016/S0140-6736(20)30183-5.
20. Nieto-Torres, J.L.; DeDiego, M.L.; Verdia-Baguena, C.; Jimenez-Guardeno, J.M.; Regla-Nava, J.A.; Fernandez-Delgado, R.; Castano-Rodriguez, C.; Alcaraz, A.; Torres, J.; Aguilella, V.M., et al. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathog* **2014**, *10*, e1004077, doi:10.1371/journal.ppat.1004077.
21. Nieto-Torres, J.L.; Verdia-Baguena, C.; Jimenez-Guardeno, J.M.; Regla-Nava, J.A.; Castano-Rodriguez, C.; Fernandez-Delgado, R.; Torres, J.; Aguilella, V.M.; Enjuanes, L. Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology* **2015**, *485*, 330-339, doi:10.1016/j.virol.2015.08.010.
22. Schoeman, D.; Fielding, B.C. Is There a Link Between the Pathogenic Human Coronavirus Envelope Protein and Immunopathology? A Review of the Literature. *Front Microbiol* **2020**, *11*, 2086, doi:10.3389/fmicb.2020.02086.
23. Teoh, K.T.; Siu, Y.L.; Chan, W.L.; Schluter, M.A.; Liu, C.J.; Peiris, J.S.; Bruzzone, R.; Margolis, B.; Nal, B. The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Mol Biol Cell* **2010**, *21*, 3838-3852, doi:10.1091/mbc.E10-04-0338.
24. Wang, D.; Hu, B.; Hu, C.; Zhu, F.; Liu, X.; Zhang, J.; Wang, B.; Xiang, H.; Cheng, Z.; Xiong, Y., et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA* **2020**, *323*, 1061-1069, doi:10.1001/jama.2020.1585.
25. Roignot, J.; Peng, X.; Mostov, K. Polarity in Mammalian Epithelial Morphogenesis. *Csh Perspect Biol* **2013**, *5*, doi:10.1101/cshperspect.a013789.
26. Linfield, D.T.; Raduka, A.; Aghapour, M.; Rezaee, F. Airway tight junctions as targets of viral infections. *Tissue Barriers* **2021**, *9*, doi:10.1080/21688370.2021.1883965.

27. Paradis, T.; Begue, H.; Basmaciyan, L.; Dalle, F.; Bon, F. Tight Junctions as a Key for Pathogens Invasion in Intestinal Epithelial Cells. *Int J Mol Sci* **2021**, *22*, doi:10.3390/ijms22052506.
28. Bilder, D.; Schober, M.; Perrimon, N. Integrated activity of PDZ protein complexes regulates epithelial polarity. *Nat Cell Biol* **2003**, *5*, 53-58, doi:10.1038/ncb897.
29. Ivanova, M.E.; Fletcher, G.C.; O'Reilly, N.; Purkiss, A.G.; Thompson, B.J.; McDonald, N.Q. Structures of the human Pals1 PDZ domain with and without ligand suggest gated access of Crb to the PDZ peptide-binding groove. *Acta Crystallogr D* **2015**, *71*, 555-564, doi:10.1107/S139900471402776x.
30. De Maio, F.; Lo Cascio, E.; Babini, G.; Sali, M.; Della Longa, S.; Tilocca, B.; Roncada, P.; Arcovito, A.; Sanguinetti, M.; Scambia, G., et al. Improved binding of SARS-CoV-2 Envelope protein to tight junction-associated PALS1 could play a key role in COVID-19 pathogenesis. *Microbes Infect* **2020**, *22*, 592-597, doi:10.1016/j.micinf.2020.08.006.
31. Toto, A.; Ma, S.; Malagrino, F.; Visconti, L.; Pagano, L.; Stromgaard, K.; Gianni, S. Comparing the binding properties of peptides mimicking the Envelope protein of SARS-CoV and SARS-CoV-2 to the PDZ domain of the tight junction-associated PALS1 protein. *Protein Sci* **2020**, *29*, 2038-2042, doi:10.1002/pro.3936.
32. Lo Cascio, E.; Toto, A.; Babini, G.; De Maio, F.; Sanguinetti, M.; Mordente, A.; Della Longa, S.; Arcovito, A. Structural determinants driving the binding process between PDZ domain of wild type human PALS1 protein and SLiM sequences of SARS-CoV E proteins. *Comput Struct Biotechnol J* **2021**, *19*, 1838-1847, doi:10.1016/j.csbj.2021.03.014.
33. Chai, J.; Cai, Y.; Pang, C.; Wang, L.; McSweeney, S.; Shanklin, J.; Liu, Q. Structural basis for SARS-CoV-2 envelope protein recognition of human cell junction protein PALS1. *Nat Commun* **2021**, *12*, 3433, doi:10.1038/s41467-021-23533-x.
34. Javorsky, A.; Humbert, P.O.; Kvansakul, M. Structural basis of coronavirus E protein interactions with human PALS1 PDZ domain. *Commun Biol* **2021**, *4*, 724, doi:10.1038/s42003-021-02250-7.
35. Arbely, E.; Khattari, Z.; Brotons, G.; Akkawi, M.; Salditt, T.; Arkin, I.T. A highly unusual palindromic transmembrane helical hairpin formed by SARS coronavirus E protein. *J. Mol. Biol.* **2004**, *341*, 769-779.
36. Wu, Q.; Zhang, Y.; Lü, H.; Wang, J.; He, X.; Liu, Y.; Ye, C.; Lin, W.; Hu, J.; Ji, J. The E protein is a multifunctional membrane protein of SARS-CoV. *Genomics, proteomics & bioinformatics* **2003**, *1*, 131-144.
37. Latek, D.; Trzaskowski, B.; Niewieczerzał, S.; Miszta, P.; Młynarczyk, K.; Dębiński, A.; Puławski, W.; Yuan, S.; Sztyler, A.; Orzeł, U. Modeling of Membrane Proteins. In *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*, Springer: 2019; pp. 371-451.
38. Eswar, N.; Eramian, D.; Webb, B.; Shen, M.Y.; Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **2008**, *426*, 145-159, doi:10.1007/978-1-60327-058-8_8.
39. Sali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779-815, doi:10.1006/jmbi.1993.1626.
40. Li, Y.; Surya, W.; Claudine, S.; Torres, J. Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins. *J. Biol. Chem.* **2014**, *289*, 12535-12549, doi:10.1074/jbc.M114.560094.
41. Surya, W.; Li, Y.; Torres, J. Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim Biophys Acta Biomembranes* **2018**, *1860*, 1309-1317, doi:10.1016/j.bbamem.2018.02.017.
42. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography* **1993**, *26*, 283-291.
43. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on protein crystallography* **2002**, *40*, 82-92.
44. Dominguez, C.; Boelens, R.; Bonvin, A.M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **2003**, *125*, 1731-1737, doi:10.1021/ja026939x.

45. Burra, P.V.; Zhang, Y.; Godzik, A.; Stec, B. Global distribution of conformational states derived from redundant models in the PDB points to non-uniqueness of the protein structure (vol 106, pg 10505, 2009). *P Natl Acad Sci USA* **2009**, *106*, 12560-12560, doi:10.1073/pnas.0907397106.
46. Kufareva, I.; Abagyan, R. Methods of protein structure comparison. *Methods Mol Biol* **2012**, *857*, 231-257, doi:10.1007/978-1-61779-588-6_10.
47. Salentin, S.; Schreiber, S.; Haupt, V.J.; Adasme, M.F.; Schroeder, M. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res* **2015**, *43*, W443-447, doi:10.1093/nar/gkv315.
48. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19-25, doi:https://doi.org/10.1016/j.softx.2015.06.001.
49. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmuller, H.; MacKerell, A.D., Jr. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017**, *14*, 71-73, doi:10.1038/nmeth.4067.
50. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J Chem Phys* **2007**, *126*, doi:10.1063/1.2408420.
51. Wong-ekkabut, J.; Karttunen, M. Assessment of Common Simulation Protocols for Simulations of Nanopores, Membrane Proteins, and Channels. *Journal of Chemical Theory and Computation* **2012**, *8*, 2905-2911, doi:10.1021/ct3001359.
52. Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, *52*, 7182-7190, doi:10.1063/1.328693.
53. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995**, *103*, 8577-8593, doi:10.1063/1.470117.
54. Ruch, T.R.; Machamer, C.E. A single polar residue and distinct membrane topologies impact the function of the infectious bronchitis coronavirus E protein. *PLoS Pathog.* **2012**, *8*, e1002674.
55. Seppälä, S.; Slusky, J.S.; Lloris-Garcerá, P.; Rapp, M.; von Heijne, G. Control of Membrane Protein Topology by a Single C-Terminal Residue. *Science* **2010**, *328*, 1698-1700, doi:10.1126/science.1188950.
56. White, C.; Nixon, A.; Bradbury, N.A. Determining Membrane Protein Topology Using Fluorescence Protease Protection (FPP). *JoVE* **2015**, doi:10.3791/52509, e52509, doi:doi:10.3791/52509.
57. Nieto-Torres, J.L.; Dediego, M.L.; Alvarez, E.; Jimenez-Guardeno, J.M.; Regla-Nava, J.A.; Llorente, M.; Kremer, L.; Shuo, S.; Enjuanes, L. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology* **2011**, *415*, 69-82, doi:10.1016/j.virol.2011.03.029.
58. Jimenez-Guardeno, J.M.; Nieto-Torres, J.L.; DeDiego, M.L.; Regla-Nava, J.A.; Fernandez-Delgado, R.; Castano-Rodriguez, C.; Enjuanes, L. The PDZ-Binding Motif of Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Is a Determinant of Viral Pathogenesis. *Plos Pathogens* **2014**, *10*, doi:10.1371/journal.ppat.1004320.
59. Yang, Y.; Xiong, Z.; Zhang, S.; Yan, Y.; Nguyen, J.; Ng, B.; Lu, H.; Brendese, J.; Yang, F.; Wang, H., et al. Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *Biochem J* **2005**, *392*, 135-143, doi:10.1042/BJ20050698.
60. Li, Y.; Surya, W.; Claudine, S.; Torres, J. Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins. *J Biol Chem* **2014**, *289*, 12535-12549, doi:10.1074/jbc.M114.560094.
61. Surya, W.; Li, Y.; Torres, J. Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim Biophys Acta Biomembr* **2018**, *1860*, 1309-1317, doi:10.1016/j.bbamem.2018.02.017.
62. Mandala, V.S.; McKay, M.J.; Shcherbakov, A.A.; Dregni, A.J.; Kolocouris, A.; Hong, M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat Struct Mol Biol* **2020**, *27*, 1202-U1270, doi:10.1038/s41594-020-00536-8.

63. Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R.H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host & Microbe* **2020**, *27*, 671-680.e672, doi:https://doi.org/10.1016/j.chom.2020.03.002.

64. Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R.H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **2020**, *27*, 671-680 e672, doi:10.1016/j.chom.2020.03.002.

65. Sun, S.J.; Karki, C.; Aguilera, J.; Hernandez, A.E.L.; Sun, J.J.; Li, L. Computational Study on the Function of Palmitoylation on the Envelope Protein in SARS-CoV-2. *Journal of Chemical Theory and Computation* **2021**, *17*, 6483-6490, doi:10.1021/acs.jctc.1c00359.

66. Aldaais, E.A.; Yegnaswamy, S.; Albahrani, F.; Alsowaiket, F.; Alramadan, S. Sequence and structural analysis of COVID-19 E and M proteins with MERS virus E and M proteins-A comparative study. *Biochem Biophys Rep* **2021**, *26*, 101023, doi:10.1016/j.bbrep.2021.101023.

67. Kuzmin, A.; Orekhov, P.; Astashkin, R.; Gordeliy, V.; Gushchin, I. Structure and dynamics of the SARS-CoV-2 envelope protein monomer. *Proteins* **2022**, *90*, 1102-1114, doi:10.1002/prot.26317.

68. Boson, B.; Legros, V.; Zhou, B.J.; Siret, E.; Mathieu, C.; Cosset, F.L.; Lavillette, D.; Denolly, S. The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. *Journal of Biological Chemistry* **2021**, *296*, doi:10.1074/jbc.RA120.016175.

69. Hogue, B.G.; Machamer, C.E. Coronavirus structural proteins and virus assembly. *Nidoviruses* **2007**, 179-200.

70. Siu, Y.L.; Teoh, K.T.; Lo, J.; Chan, C.M.; Kien, F.; Escriou, N.; Tsao, S.W.; Nicholls, J.M.; Altmeyer, R.; Peiris, J.S.M., et al. The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. *Journal of Virology* **2008**, *82*, 11318-11330, doi:10.1128/Jvi.01052-08.

71. Torres, J.; Maheswari, U.; Parthasarathy, K.; Ng, L.F.; Liu, D.X.; Gong, X.D. Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein. *Protein Science* **2007**, *16*, 2065-2071, doi:10.1110/ps.062730007.

72. Verdia-Baguena, C.; Nieto-Torres, J.L.; Alcaraz, A.; DeDiego, M.L.; Torres, J.; Aguilella, V.M.; Enjuanes, L. Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. *Virology* **2012**, *432*, 485-494, doi:10.1016/j.virol.2012.07.005.

73. Farsad, K.; Ringstad, N.; Takei, K.; Floyd, S.R.; Rose, K.; De Camilli, P. Generation of high curvature membranes mediated by direct endophilin bilayer interactions. *J Cell Biol* **2001**, *155*, 193-200, doi:DOI 10.1083/jcb.200107075.

74. Martyna, A.; Bahsoun, B.; Badham, M.D.; Srinivasan, S.; Howard, M.J.; Rossman, J.S. Membrane remodeling by the M2 amphipathic helix drives influenza virus membrane scission. *Sci Rep-Uk* **2017**, *7*, doi:10.1038/srep44695.

75. Varkey, J.; Zhang, J.T.; Kim, J.; George, G.; He, G.J.; Belov, G.; Langen, R.; Wang, X.F. An Amphipathic Alpha-Helix Domain from Poliovirus 2C Protein Tubulate Lipid Vesicles. *Viruses-Basel* **2020**, *12*, doi:10.3390/v12121466.

76. Bhatia, V.K.; Madsen, K.L.; Bolinger, P.Y.; Kunding, A.; Hedegard, P.; Gether, U.; Stamou, D. Amphipathic motifs in BAR domains are essential for membrane curvature sensing. *Embo J* **2009**, *28*, 3303-3314, doi:10.1038/emboj.2009.261.

77. Wang, Z.; Ma, W.; Zheng, X.; Wu, G.; Zhang, R. Household transmission of SARS-CoV-2. *J Infect* **2020**, *81*, 179-182, doi:10.1016/j.jinf.2020.03.040.

78. Wilson-Clark, S.D.; Deeks, S.L.; Gournis, E.; Hay, K.; Bondy, S.; Kennedy, E.; Johnson, I.; Rea, E.; Kuschak, T.; Green, D., et al. Household transmission of SARS, 2003. *CMAJ* **2006**, *175*, 1219-1223, doi:10.1503/cmaj.050876.

79. Drosten, C.; Meyer, B.; Muller, M.A.; Corman, V.M.; Al-Masri, M.; Hossain, R.; Madani, H.; Sieberg, A.; Bosch, B.J.; Lattwein, E., et al. Transmission of MERS-coronavirus in household contacts. *N Engl J Med* **2014**, *371*, 828-835, doi:10.1056/NEJMoa1405858.

80. Grijalva, C.G.; Rolfes, M.A.; Zhu, Y.; McLean, H.Q.; Hanson, K.E.; Belongia, E.A.; Halasa, N.B.; Kim, A.; Reed, C.; Fry, A.M., et al. Transmission of SARS-COV-2 Infections in Households - Tennessee and Wisconsin, April-September 2020. *MMWR Morb Mortal Wkly Rep* **2020**, *69*, 1631-1634, doi:10.15585/mmwr. mm6944e1.
81. Killerby, M.E.; Biggs, H.M.; Midgley, C.M.; Gerber, S.I.; Watson, J.T. Middle East Respiratory Syndrome Coronavirus Transmission. *Emerg Infect Dis* **2020**, *26*, 191-198, doi:10.3201/eid2602.190697.
82. Alsaadi, E.A.J.; Jones, I.M. Membrane binding proteins of coronaviruses. *Future Virol* **2019**, *14*, 275-286, doi:10.2217/fvl-2018-0144.
83. Rost, B. Twilight zone of protein sequence alignments. *Protein Engineering* **1999**, *12*, 85-94, doi:DOI 10.1093/protein/12.2.85.
84. Nieto-Torres, J.L.; DeDiego, M.L.; Álvarez, E.; Jiménez-Guardeño, J.M.; Regla-Nava, J.A.; Llorente, M.; Kremer, L.; Shuo, S.; Enjuanes, L. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology* **2011**, *415*, 69-82, doi:https://doi.org/10.1016/j.virol.2011.03.029.
85. Teoh, K.-T.; Siu, Y.-L.; Chan, W.-L.; Schlüter, M.A.; Liu, C.-J.; Peiris, J.M.; Bruzzone, R.; Margolis, B.; Nal, B. The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Mol. Biol. Cell* **2010**, *21*, 3838-3852, doi:https://doi.org/10.1091/mbc.e10-04-0338.
86. Yang, Y.; Xiong, Z.; Zhang, S.; Yan, Y.; Nguyen, J.; Ng, B.; Lu, H.; Brendese, J.; Yang, F.; Wang, H. Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. *Biochem. J.* **2005**, *392*, 135-143.
87. Caillet-Saguy, C.; Durbesson, F.; Rezelj, V.V.; Gogl, G.; Tran, Q.D.; Twizere, J.C.; Vignuzzi, M.; Vincentelli, R.; Wolff, N. Host PDZ-containing proteins targeted by SARS-CoV-2. *FEBS J* **2021**, *288*, 5148-5162, doi:10.1111/febs.15881.
88. Pace, C.N.; Scholtz, J.M.; Grimsley, G.R. Forces stabilizing proteins. *Febs Letters* **2014**, *588*, 2177-2184, doi:10.1016/j.febslet.2014.05.006.
89. Kangueane, P.; Nilofer, C. Principles of Protein-Protein Interaction. In *Protein-Protein and Domain-Domain Interactions*, Springer Singapore: Singapore, 2018; 10.1007/978-981-10-7347-2_8pp. 93-111.