

## Structural basis of coronavirus E protein interactions with human PALS1 PDZ domain

Airah Javorsky<sup>1</sup>, Patrick O. Humbert<sup>1,2,3,4</sup>  & Marc Kvensakul<sup>1,2</sup> 

SARS-CoV-2 infection leads to coronavirus disease 2019 (COVID-19), which is associated with severe and life-threatening pneumonia and respiratory failure. However, the molecular basis of these symptoms remains unclear. SARS-CoV-1 E protein interferes with control of cell polarity and cell-cell junction integrity in human epithelial cells by binding to the PALS1 PDZ domain, a key component of the Crumbs polarity complex. We show that C-terminal PDZ binding motifs of SARS-CoV-1 and SARS-CoV-2 E proteins bind the PALS1 PDZ domain with 29.6 and 22.8  $\mu\text{M}$  affinity, whereas the related sequence from MERS-CoV did not bind. We then determined crystal structures of PALS1 PDZ domain bound to both SARS-CoV-1 and SARS-CoV-2 E protein PDZ binding motifs. Our findings establish the structural basis for SARS-CoV-1/2 mediated subversion of Crumbs polarity signalling and serve as a platform for the development of small molecule inhibitors to suppress SARS-CoV-1/2 mediated disruption of polarity signalling in epithelial cells.

<sup>1</sup>Department of Biochemistry & Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Vic, Australia. <sup>2</sup>Research Centre for Molecular Cancer Prevention, La Trobe University, Melbourne, Vic, Australia. <sup>3</sup>Department of Biochemistry & Molecular Biology, University of Melbourne, Melbourne, Vic, Australia. <sup>4</sup>Department of Clinical Pathology, University of Melbourne, Melbourne, Vic, Australia. ✉email: [p.humbert@latrobe.edu.au](mailto:p.humbert@latrobe.edu.au); [m.kvensakul@latrobe.edu.au](mailto:m.kvensakul@latrobe.edu.au)

The emergence of the novel coronavirus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) at the end of 2019 in China led to a global pandemic with currently 120 M infections and over 2.5 M deaths<sup>1</sup>. In addition to the substantial morbidity and mortality associated with SARS-CoV-2 infection<sup>2</sup>, the pandemic has led to major global economic dislocations, spurring a concerted worldwide campaign to develop therapeutics targeting this virus.

Whilst a substantial part of the global effort was aimed at the SARS-CoV-2 spike protein for vaccine development<sup>3</sup>, other proteins have also been examined in detail<sup>4</sup>. Coronaviruses including SARS-CoV-2, the causative agent of COVID-19, comprise four major structural proteins, spike protein (S), nucleocapsid protein (N), membrane protein (M) and envelope protein (E)<sup>5</sup>. E protein is a 75–109 amino acid integral membrane protein and has been predicted to feature a short 7–12 residue hydrophilic N-terminal domain followed by a hydrophobic transmembrane domain of ~25 residues and a long hydrophilic C-terminal domain<sup>6</sup>. In addition to SARS-CoV-2, several other coronaviruses have been shown to be important human pathogens, including the closely related SARS-CoV-1 and MERS-CoV. All three are members of the betacoronavirus family, with SARS-CoV-1 and 2 belonging to clade b, whereas MERS-CoV belongs to clade c<sup>7</sup>.

Functionally, the SARS-CoV-1 E protein has been shown to play a role in virus pathogenesis and may act as a virulence factor<sup>8,9</sup>. NMR studies revealed that peptides covering the central transmembrane region of SARS-CoV-1 E protein oligomerize in lyso-myristoyl phosphatidylglycerol (LMPG) micelles to form pentameric ion channels<sup>10,11</sup>, whereas the C-terminal tail was shown to harbour a binding motif sequence that allows binding to PDZ (postsynaptic density protein (PSD95), *Drosophila* disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1)) domains, an 80–90 amino acid structural domain commonly found in signalling proteins<sup>12</sup>. Importantly, the SARS-CoV-1 E protein's PDZ-binding motif<sup>13</sup> sequence allows interaction with the host cell polarity signalling protein PALS1 (Protein associated with Lin7/MPP5) thus interfering with epithelial function and integrity, a property of the SARS-CoV-1 E protein which is thought to be important for its virulence<sup>13</sup>.

Cell polarity refers to the asymmetric distribution of cellular components including proteins, lipids and carbohydrates, and is a critical process to control animal tissue organization and development<sup>14</sup>. Apico-basal cell polarity is a common form of epithelial polarity that allows spatial segregation of cellular components between the apical and basolateral membrane and is also key for the establishment and maintenance of epithelial junctions<sup>15</sup>. One of the major complexes that regulate apical cell polarity is the Crumbs (CRB) complex (Fig. 1a)<sup>16,17</sup>. The CRB complex consists of CRB (CRB1–3 in humans), PATJ (Pals1-associated tight junction protein) and PALS1<sup>16</sup>. PALS1 is a member of the MAGUK family (membrane-associated guanylate kinase) and is a signalling adaptor protein that consists of a number of protein-protein interaction domains including two L27 domains, a PDZ domain, an Src homology 3 (SH3) domain, a Hook domain, and a Guanylate Kinase (Guk) domain<sup>16</sup>. PALS1 regulates cell polarity through binding to the PDZ-binding domain (PBM) of CRB3 with loss of PALS1 leading to disruption of cell polarity including loss of integrity of tight and adherens junctions<sup>18–23</sup>. The CRB3 protein is the major crumbs isoforms expressed in human epithelial cells with the CRB3A splice form encoding the carboxy-terminal PBM sequence ERLI that enables PALS1 binding<sup>24</sup>.

Importantly, in addition to being the binding site for the Crumbs PBM, the PDZ domain of PALS1 is also the target of certain viral proteins during viral infections<sup>25</sup>. Indeed, the SARS-CoV-1 E protein features a type II PDZ-binding motif (PBM)

harbouring the X- $\varphi$ -X- $\varphi$ <sub>COOH</sub> consensus sequence (where X represents any amino acid and  $\varphi$  is a hydrophobic residue such as V, I or L)<sup>26</sup> that has been shown to bind to PALS1 PDZ<sup>13</sup>. Following viral infection, PALS1 is recruited to virus-assembly sites at the Golgi where it co-localises with the SARS-CoV-1 E protein<sup>13</sup>. This interaction was shown to be dependent on the PBM of SARS-CoV-1 E protein, which competes with the CRB3 PBM for the PALS1 interaction (Fig. 1a), leading to disrupted localisation of PALS1, altered cell shape and disrupted tight junction assembly of epithelial cells<sup>13</sup>. Conversely, deletion of the PBM in SARS-CoV-1 E protein decreases expression of inflammatory cytokines and resulted in substantial virus attenuation<sup>8,27</sup>, suggesting that E protein plays a prominent role in viral pathogenesis. Hijacking of PALS1 by SARS-CoV-1 E protein in infected lung cells is therefore postulated to lead to the loss of epithelial barrier function resulting in enhanced infiltration of SARS-CoV-1 virions into underlying tissues, thus contributing to the severe alveolar damages observed in post-mortem lung biopsies from SARS-CoV-1-infected patients<sup>13</sup>. However, the structural basis of E protein engagement of human PALS1 PDZ remains to be clarified.

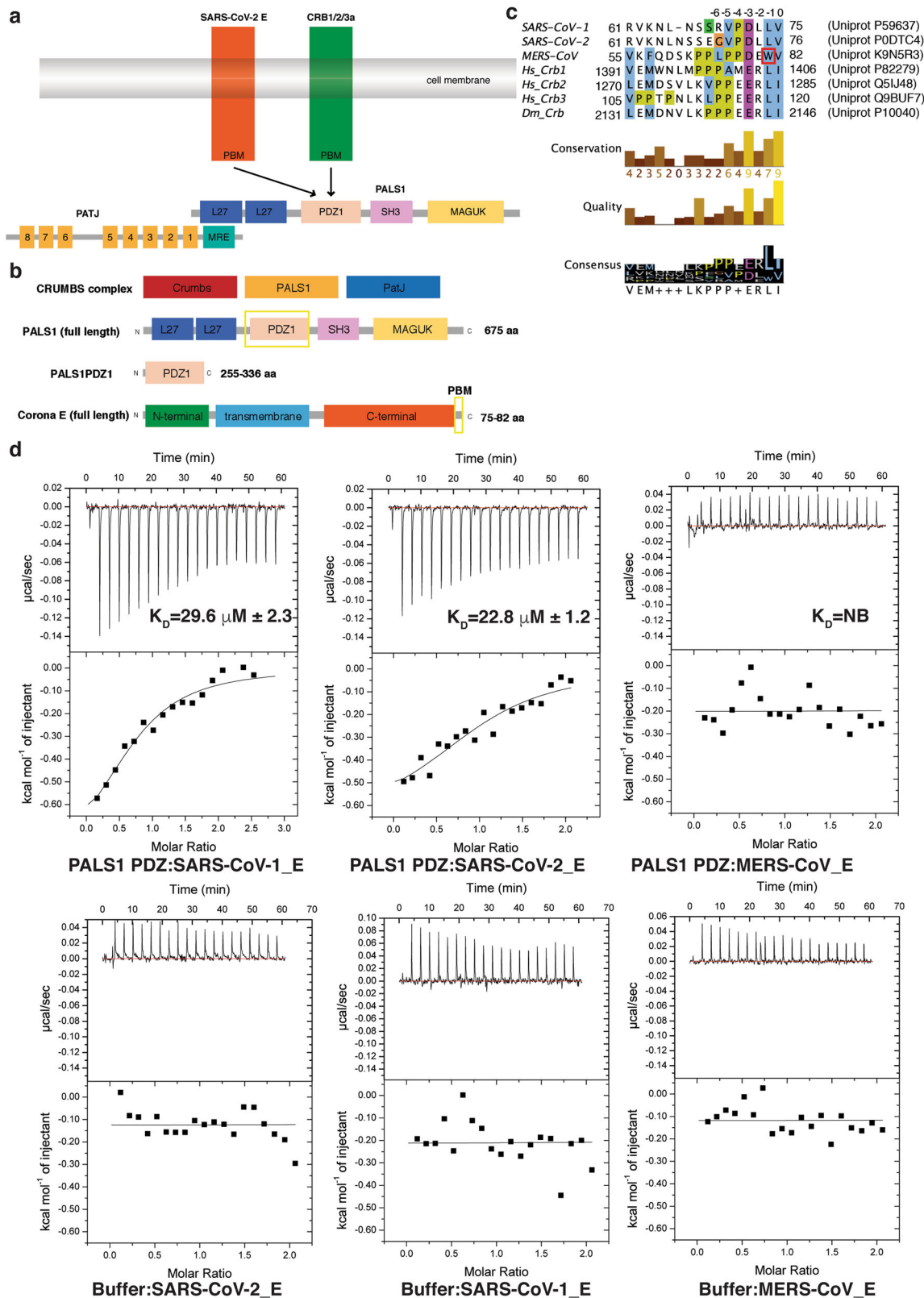
We now report interaction and crystallographic studies examining the binding of Coronavirus E protein PDZ-binding motifs to the human PALS1 PDZ domain. Our results provide a structural basis for the subversion of Crumbs signalling by SARS-CoV-1 and -2, and provide a platform for the development of small molecule antagonists to exploit this for therapeutic intervention.

## Results and discussion

To understand the impact of the Coronavirus E protein binding to PALS1 on polarity signalling, we examined the affinity of recombinant human PALS1 PDZ (Fig. 1b) for 8-mer peptides spanning the C-terminal PBM of SARS-CoV-1\_E, SARS-CoV-2\_E and MERS-CoV\_E (Fig. 1c). Our ITC measurements revealed that whilst PALS1 PDZ did not bind to MERS-CoV\_E, it bound to SARS-CoV-1\_E with an affinity of 29.6  $\mu$ M, and SARS-CoV-2\_E with an affinity of 22.8  $\mu$ M (Fig. 1d). Our measurements differ from previously published findings that revealed an affinity of 130  $\mu$ M for SARS-CoV-1\_E and 40  $\mu$ M for SARS-CoV-2\_E<sup>28</sup>, however, we note that our study used 8-mer peptides whereas Toto et al. used 10-mer peptides together with the PALS1 PDZ F318W mutant. Analysis of the thermodynamic parameters of SARS-CoV-1/2\_E PBM binding to PALS1 PDZ revealed an unfavourable entropy change, with the interaction being driven by enthalpy (Table 1).

To understand the structural basis of the PALS1 PDZ and SARS-CoV-1/2 E PBM interaction, we determined the crystal structures of PALS1 PDZ bound to the SARS-CoV-1\_E (SRVPDLLV) and SARS-CoV-2\_E (EGVPDLLV) PBM peptides to resolutions of 1.74 and 1.9  $\text{\AA}$ , respectively (Table 2). As previously shown the PALS1 PDZ domains adopts a compact globular fold comprising five to six  $\beta$ -strands and two  $\alpha$ -helices that form a  $\beta$ -sandwich structure<sup>23,29</sup>. Examination of both the PALS1 PDZ:SARS-CoV-1\_E and PALS1 PDZ:SARS-CoV-2\_E complexes showed that the E protein PBM peptides are bound in the canonical ligand-binding groove formed by the  $\beta$ 2 strand and helix  $\alpha$ 2 (Fig. 2b–e). PALS1 PDZ:SARS-CoV-1\_E and PALS1 PDZ:SARS-CoV-2\_E superimposed with one another yielded an overall root-mean-square deviation (RMSD) of 0.72  $\text{\AA}$  (Fig. 2f). There was no discernible difference between PALS1 PDZ domains from the PDZ:SARS-CoV-2\_E complex and PALS1 PDZ from an endogenous complex PALS1 PDZ:Crumbs (PDB ID 4UU5), which were superimposed to produce an overall RMSD of 0.80  $\text{\AA}$  (Fig. 2g).

In the PALS1 PDZ:SARS-CoV-1\_E complex, SARS-CoV-1\_E Val75 is docked in the conserved hydrophobic pocket formed by



PALS1 PDZ Leu267, Gly268 and Ala269 (Fig. 3a). In addition, two hydrogen bonds are formed by Val271<sup>PALS1 PDZ</sup>–Leu74<sup>SARS-CoV-1\_E</sup> located at the second β-sheet of PALS1 PDZ (Fig. 3a). Lastly, an ionic interaction is found between Arg272<sup>PALS1 PDZ</sup>–Asp73<sup>SARS-CoV-1\_E</sup> (Fig. 3a), and SARS-CoV-2\_E Val75, Leu74 and Leu 73 are involved in hydrophobic interactions.

Key hydrogen bond and ionic interactions between the PALS1 PDZ:SARS-CoV-2\_E are identical with SARS-CoV-1\_E (Fig. 3b). Major differences are found among the hydrophobic contacts, with the rotamer of SARS-CoV-1\_E Leu73 differing from SARS-CoV-2\_E Leu74. Furthermore, in SARS-CoV-1\_E only Val75, Leu74 and Leu 73 are involved in hydrophobic contacts, whereas

**Fig. 1 Interactions of PALS1 PDZ domain with SARS-CoV-1, SARS-CoV-2 and MERS-CoV E protein PBM.** **a** Schematic outline of Crumb's complex interactions. **b** Domain organization of PALS1 and SARS-CoV-2 E protein. **c** Sequence alignment of C-terminal PBM motifs of coronavirus E proteins with human (hs) crumbs isoforms and *D. melanogaster* (*dm*) crumbs. Trp81 in the MERS CoV E protein is boxed in red. PBM sequences are numbered above the alignment with the first C-terminal residue designated as 0, with preceding residues numbered  $-1$ ,  $-2$ ,  $-3$  etc. Uniprot accession numbers and sequence boundaries are listed for all sequences. Residue colouring is as follows: blue for hydrophobic, magenta for a negative charge, red for positive charge, yellow for prolines and white for unconserved. Three different conservation metrics calculated by Jalview are shown. Quality score is calculated for each column in an alignment by summing, for all mutations, the ratio of the two BLOSUM 62 scores for a mutation pair and each residue's conserved BLOSUM62 score, and then plotted on a scale from 0 to 1. The consensus displayed below the alignment is the percentage of the modal residue per column. The alignment was generated using Clustal Omega<sup>44</sup> followed by analysis using Jalview<sup>45</sup>. **d** Binding profiles of isolated PALS1 PDZ domain interaction with coronavirus E protein PBM peptides are displayed. Each profile is represented by a raw thermogram (top panel) and a binding isotherm fitted with a one-site binding model (bottom panels).  $K_D$  dissociation constant,  $\pm$  standard deviation, NB no binding. Each of the values was calculated from at least three independent experiments. Top panels show peptide into protein titrations, bottom panels show peptide into buffer control titrations.

**Table 1 Summary of affinity and thermodynamic binding parameters for PALS1 PDZ domain interactions with SARS CoV-1 and -2 PBM peptides.**

	$K_D$ ( $\mu\text{M}$ )	$\Delta H$ (kcal/mol)	$\Delta G$ (kcal/mol)	$-\Delta S$ (kcal/mol/K)	$N$
SARS-CoV-1 PBM	$29.6 \pm 2.3$	$-0.67 \pm 0.01$	$-6.15 \pm 0.03$	$5.49 \pm 0.04$	$1.0 \pm 0.02$
SARS-CoV-2 PBM	$22.8 \pm 1.2$	$-0.68 \pm 0.01$	$-6.34 \pm 0.02$	$5.66 \pm 0.15$	$0.8 \pm 0.1$
MERS-CoV PBM	NB	NB	NB	NB	NB

NB denotes no binding. Each of the value was calculated from at least three independent experiments. Errors are  $\pm$ SD.

**Table 2 X-ray data collection and refinement statistics.**

	PALS1 PDZ:SARS-CoV-1_E	PALS1 PDZ:SARS-CoV-2_E
<i>Data collection</i>		
Space group	P 1	P 2 <sub>1</sub>
No. of molecules in AU	2	4
<i>Cell dimensions</i>		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	28.15, 39.96, 40.83	28.69, 103.76, 59.08
$\alpha$ , $\beta$ , $\gamma$ (°)	94.90, 108.78, 100.84	90.00, 103.98, 90.00
Wavelength (Å)	0.9537	0.9537
Resolution (Å)*	38.18-1.74 (1.8-1.74)	27.84-1.9 (1.94-1.9)
$R_{\text{sym}}$ or $R_{\text{merge}}$ *	0.040 (0.150)	0.129 (0.725)
$I / \sigma I$ *	12.9 (4.4)	3.4 (0.8)
CC(1/2)	0.998 (0.965)	0.987 (0.431)
Completeness (%)*	96.4 (90.5)	99.4 (98.6)
Redundancy*	3.6 (3.4)	2.9 (2.9)
Wilson B-factor	14.62	19.33
<i>Refinement</i>		
Resolution (Å)	38.18-1.736	27.84-1.9
No. of reflections	16,267	26,171
$R_{\text{work}}/R_{\text{free}}$	0.162/0.188	0.196/0.229
<i>No. of non-hydrogen atoms</i>		
Protein	1378	2763
Ligand/ion	0	33
Water	306	359
<i>B-factors</i>		
Protein	19.66	25.55
Ligand/ion	0	45.71
Water	31.01	30.33
<i>R.m.s. deviations</i>		
Bond lengths (Å)	0.003	0.006
Bond angle (°)	0.61	0.59
<i>Ramachandran plot (%)</i>		
Favoured	98.29	98.87
Allowed	1.71	1.13
Disallowed	0	0

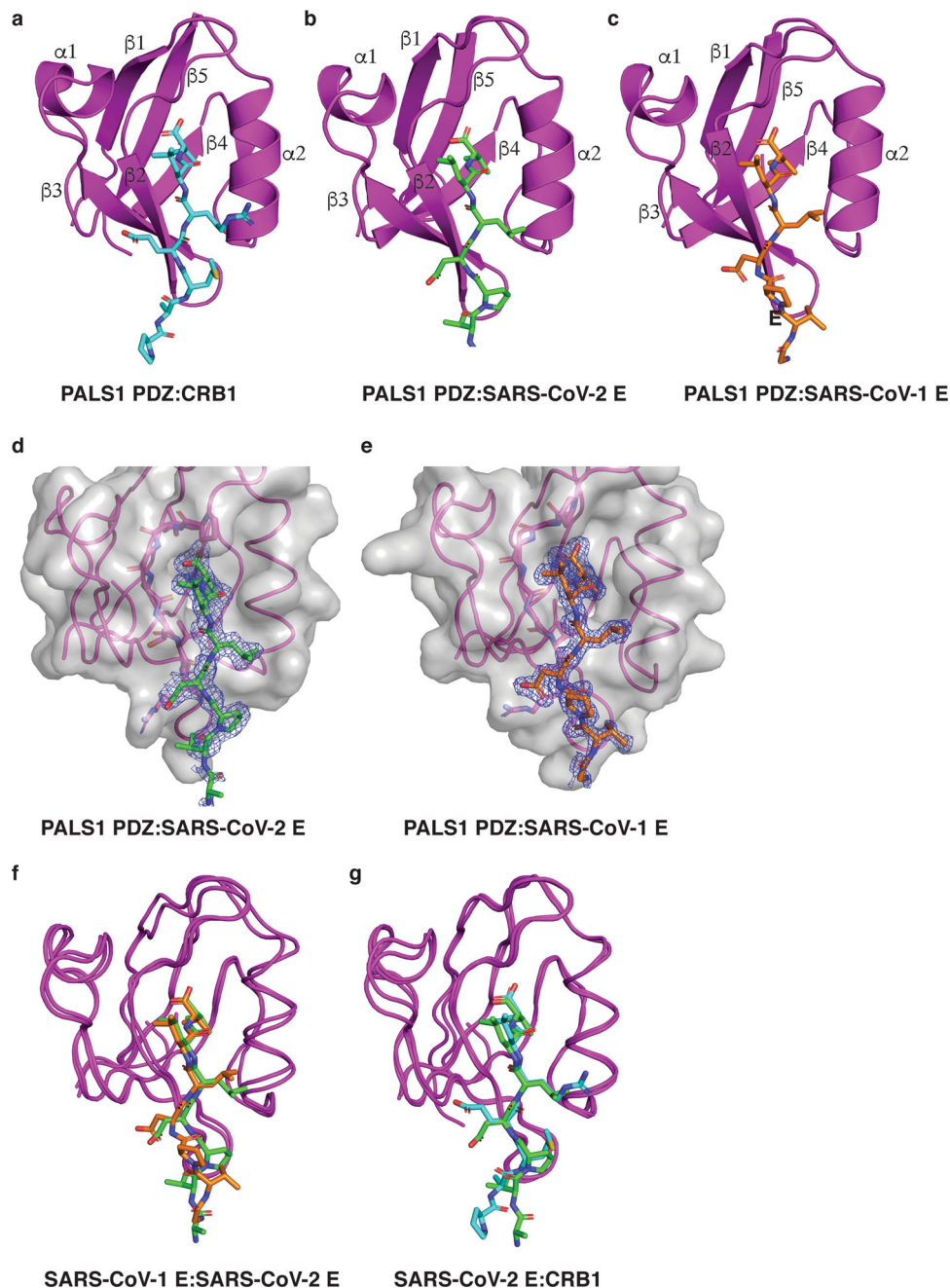
\*Values in parentheses are for the highest resolution shell.

in SARS-CoV-2\_E Pro72, Val71 are also involved in hydrophobic contacts. However, despite the sequence differences between SARS-CoV-1\_E and SARS-CoV-2\_E PBM in positions  $-5$  and  $-6$ , affinities for PALS1 PDZ are comparable, indicating that positions 0,  $-1$ ,  $-2$  and  $-3$  are required for CoV\_E PBM interaction with PALS1 PDZ and define binding, whereas positions  $-4$ ,  $-5$ ,  $-6$  and  $-8$  do not play a major role. In contrast, CRB1 engagement of PALS1 PDZ differs from SARS-CoV\_E binding due to notable differences in the CRB1 PBM sequence (Fig. 1c) such as the  $-2$  position where Arg1404<sup>CRB1</sup> contacts  $\alpha 2$  to form a hydrogen bond with Asn315<sup>PALS1 PDZ</sup> and  $-3$  position Glu1403<sup>CRB1</sup> forms an ionic interaction and hydrogen bond to the  $\beta 3$  with Arg282<sup>PALS1 PDZ</sup> (Fig. 3c)<sup>29</sup>. The ionic interaction between Glu1403<sup>CRB1</sup> and Arg282<sup>PALS1 PDZ</sup> mimics the interaction seen in the E protein complexes, however, due to the presence of the shorter side chain in Asp in E PBM compared to the Glu found in CRB1 the acceptor residue in PALS1 PDZ is Arg272 (Fig. 3c). It was previously shown that binding of CRB1 PBM to PALS1 required movement of the gatekeeper Phe318, which in the unbound form blocks access to the canonical ligand-binding groove in PALS1 PDZ<sup>29</sup>. A comparison of our structures with apo-PALS1 PDZ reveals that similarly, both SARS-CoV-1\_E and SARS-CoV-2\_E binding to PALS1 PDZ trigger rotation of PALS1 Phe318 out of the ligand-binding groove to allow binding of SARS-CoV-1/2\_E PBM (Fig. 3d).

A comparison of MERS-CoV E PBM with SARS-CoV-1\_E and SARS-CoV-2\_E reveals that the major difference is the presence of a large bulky Trp in MERS-CoV E PBM position  $-1$  (Fig. 1b), where SARS-CoV-1\_E and SARS-CoV-2\_E, as well as Crumbs, feature the smaller Leu. With MERS-CoV E PBM displaying no detectable affinity for PALS1 PDZ we surmise that the Trp in the  $-1$  position is too bulky to be accommodated in the PALS1 PDZ binding groove and particularly the pocket accommodating the  $-1$  Leu from Crumbs PBM (Fig. 3e), thus completely ablating the interaction.

A 6-mer peptide spanning the CRB1 PBM motif bound to PALS1 PDZ with an affinity of  $16.8 \mu\text{M}$ <sup>29</sup>, comparable to the affinities we determined for our E protein PBM 8-mer peptides. This suggests that CoV\_E proteins do not outcompete CRB from

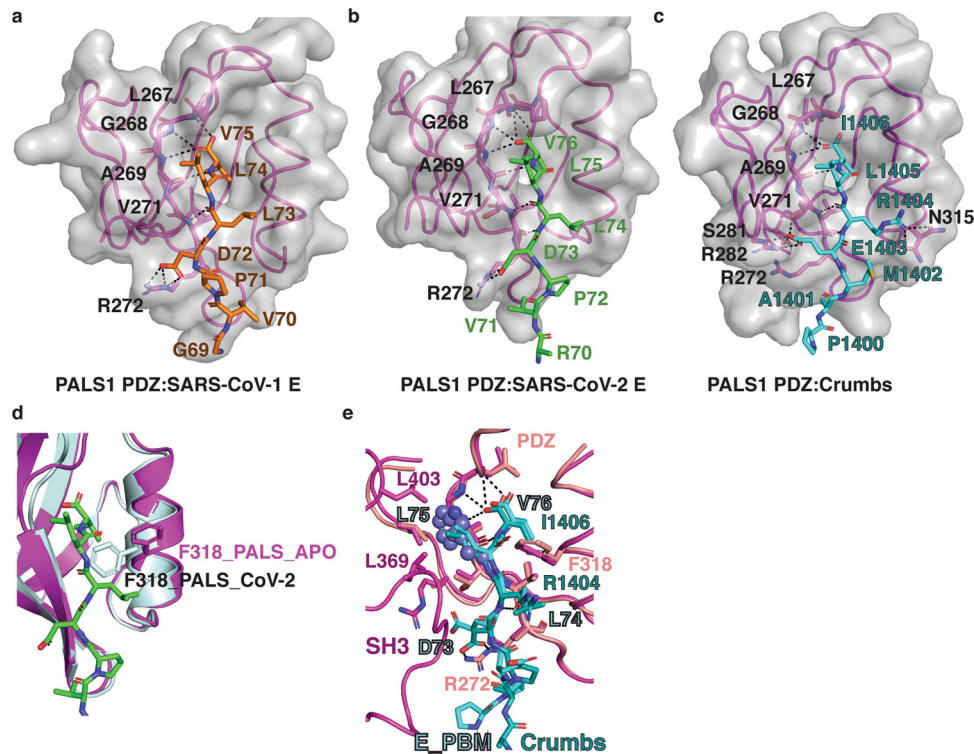




**Fig. 2** The crystal structures of PALS1 PDZ bound to SARS-CoV-1 and SARS-CoV-2 E protein PBM. The E protein PBM peptide engages PALS1 PDZ via a shallow groove located between the  $\beta 2$  and  $\alpha 2$ . **a** Human PALS1 PDZ (magenta) is shown as a cartoon with a human CRB1 peptide (cyan) represented as sticks (PDB ID 4UU5)<sup>29</sup>. **b** PALS1 PDZ (magenta) is shown as a cartoon with SARS-CoV-2 E protein PBM (green) represented as sticks. **c** PALS1 PDZ (magenta) is shown as a cartoon with SARS-CoV-1 E protein PBM (orange) represented as sticks. **d**, **e** Electron density maps centred on the SARS-CoV-2 E protein PBM peptide (green) or SARS-CoV-1 protein PBM peptide (orange) in the binding groove of PALS1 PDZ. The electron density map is a 2Fo-Fc map shown as a blue mesh contoured at 1.5 $\sigma$ . **f** Overlay of ribbon traces of PALS1 PDZ with SARS-CoV-2 E protein PBM (green) and PALS1 PDZ with SARS-CoV-1 E protein PBM peptide (orange). **g** Overlay of ribbon traces of PALS1 PDZ with SARS-CoV-2 E protein PBM (green) and human Crumbs peptide (cyan) peptide (PDB ID 4UU5)<sup>29</sup>.

PALS1, and that the effect of perturbed polarity signalling is due to the forced mislocalization of PALS1 after binding to CoV\_E. In SARS-CoV-1 infected cells, PALS1 is specifically recruited to the virus budding compartment (ERGIC/Golgi region)<sup>13</sup>. As correct localisation of cell polarity proteins is essential for their function, mislocalisation of PALS1 would thus result in the altered cell shape and disrupted tight junction assembly of epithelial cells observed following SARS-CoV-1 infection of epithelial cells<sup>13</sup>. Interestingly, mislocalisation is specific to PALS1 as another PDZ

domain-containing protein, ZO-1 remained localised to tight junctions<sup>13</sup>. The differential binding of MERS vs. SARS-CoV-1/2 is also consistent with differential virulence. Interestingly, SARS-CoV infects polarised epithelium specifically through the apical cell membrane<sup>30</sup>, whereas MERS enters both through apical and basolateral membranes, and viral budding occurs via apoptosis<sup>31</sup>, which raises the possibility that the differences in viral entry could also reflect the differential ability to perturb polarity signalling. As PALS1 is also expressed in endothelial cells where it



**Fig. 3 Detailed view into the interfaces of PALS 1 PDZ bound to SARS-CoV-1 and SARS-CoV-2 E protein PBM.** **a–c** Detailed view of interactions in the binding groove of PALS1 PDZ in complex with **a, b** SARS-CoV-2 E protein PBM peptide (green) or SARS-CoV-1 E protein PBM peptide (orange) or **c** Crumbs PBM peptide (cyan) (PDB ID 4UU5)<sup>29</sup>. Residues involved in hydrogen bonds or ionic interactions (shown as black dotted lines) are labelled. **d** Superimposition of PALS1 PDZ:SARS-CoV-2\_E\_PBM (magenta cartoon and green sticks) and apo-PALS1 PDZ (light blue). The orientation of the gatekeeper F318 is shown as sticks. **e** Superimposition of PALS1 PDZ:SARS-CoV-2\_E\_PBM with PALS-PDZ-SH3-MAGUK:Crumbs\_PBM complex. PALS1 PDZ:SARS-CoV-2\_E\_PBM is shown as pink cartoon and sticks or cyan sticks, respectively. PALS-PDZ-SH3-MAGUK (PDB ID 4WS1)<sup>23</sup> is shown as light pink cartoon and sticks. Dm Crumbs\_PBM is shown as dark cyan sticks. Key residues are labelled, and key hydrogen or ionic bonds are denoted by black dashed lines. The modelled Trp 81 from MERS CoV E PBM is shown as steel blue spheres.

regulates vasculature morphogenesis through binding and modulation of VE-cadherin function<sup>32</sup>, CoV\_E interference with PALS1 function could also play a direct role in compromising endothelial cell–cell junctions and inducing vascular leakage observed in SARS-CoV-2 patients<sup>33</sup>. Of note, VE-cadherin levels are also downregulated as a result of COVID-19 infection<sup>34</sup> suggesting endothelial cells could be already sensitised to further disruption of junctional regulators.

Subversion of polarity signalling is a feature associated with a number of viral infections, including those by human papillomavirus (HPV), human T lymphotropic virus-1 (HTLV), Tick-Borne encephalitis virus (TBEV) and influenza A virus (IAV)<sup>25</sup>. For tumorigenic viruses such as HPV, the ability to manipulate host cell polarity pathways has been shown to be strongly associated with the development of subsequent malignancies, whereas for non-malignant viruses including TBEV and IAV subversion of polarity signalling is associated with increased virulence and evasion of immune defences<sup>25</sup>. Evidence for SARS-CoV-1\_E suggests that similar to other non-malignant viruses the ability to modulate polarity signalling is associated with increased virulence<sup>13,27</sup>.

In summary, we have shown that SARS-CoV-1\_E and SARS-CoV-2\_E PBM bind to PALS1 PDZ with comparable affinities, and established the structural basis for CoV\_E PBM binding to PALS1 PDZ. These findings provide a mechanistic basis for coronavirus induced perturbation of polarity signalling and may serve as a platform for the design of small-molecule inhibitors to

target the envelope protein interaction with PALS1 for antiviral therapy.

## Methods

**Protein expression and purification.** The codon optimised cDNA encoding PALS1 (Uniprot accession number: Q8N3R9) PDZ domain (residues 255–336) was cloned into the bacterial expression vector pGex-6p-1 (GenScript). Target proteins were expressed using *Escherichia coli* condon+ in 2YT media (1.6% w/v Tryptone, 1.0% w/v yeast extract and 0.5% w/v sodium chloride (NaCl)) supplemented with 100 µg/ml ampicillin at 37 °C for 20 h. Protein expression was induced using the auto-induction protocol<sup>35</sup>. Cells were harvested by centrifugation at 4000g for 20 min using the JLA 9.1000 rotor (Beckman Coulter Avanti J-E), then resuspended in 100 ml buffer A (50 mM Tris pH 8.0, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA)). Lysis was carried out via sonication at the amplitude of 50 for 4 min with a 5-s gap after every 30-s sonication interval, using the Qsonica Q700 Sonicator. The resultant lysate was transferred into SS34 tubes for further centrifugation at 18,000 rpm for 30 min using the JA-25.50 rotor (Beckman Coulter Avanti J-E). The supernatant obtained after centrifugation was then filtered through a 0.45 µm pore size filter unit (Merck Millipore). The filtered supernatant was loaded onto 5 mL of glutathione sepharose 4B (GE Healthcare) equilibrated with buffer A. The column was washed with 100 ml of buffer A and protein on-column cleavage was achieved by adding 3 mg human rhinovirus 3 C protease and incubating for 16 h at 4 °C. 100 ml of buffer A was then added to collect cleaved protein, with the flowthrough concentrated at 4 °C using 3 kDa molecular weight cut-off centrifugal concentrator (Amicon® Ultra 15 Millipore) with a benchtop centrifuge (Allegra X-15, Beckman Coulter) at 4000g. PALS1 PDZ was subjected to size-exclusion chromatography using a Superdex S75 increase 10/300 column mounted on an ÄKTA Pure system (GE Healthcare) equilibrated in TRIS buffer (25 mM TRIS pH 8, 200 mM NaCl) or HEPES buffer (20 mM HEPES, 150 NaCl) and fractions were analysed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). The final sample purity was estimated to be higher than 95% based on SDS-PAGE analysis. Appropriate fractions were collected and concentrated using a centrifugal concentrator with 3 kDa molecular weight cut-off (Amicon® Ultra 15) to a final

concentration of 1.5 mg/ml with protein in TRIS buffer subsequently used for ITC analysis, and at 3.5 mg/ml in HEPES buffer used for crystallisation. The concentration of proteins was quantified using a Nanodrop UV spectrophotometer (Thermo Scientific) at wavelength 205 nm via the Scopes method<sup>36</sup>.

**Measurement of dissociation constants.** Binding affinities were measured for PALS1 PDZ in TRIS buffer at a final concentration of 90  $\mu$ M. Peptide ligands were equilibrated in TRIS buffer and used at a concentration of 900  $\mu$ M, and all affinity measurements were performed in triplicate. Peptides of E protein used were SARS-CoV-1: SRVPDLLV (UniProt accession code P59637, residues 68–75) (Mimotopes), SARS-CoV-2: EGVPELLV (UniProt accession code P0DTC4, residues 69–76) (Mimotopes) and MERS-CoV: PLPPEWV (UniProt accession code K9N5R3, residues 75–82) (Mimotopes). Peptides were dissolved in water to prepare 5 mM stock solutions and subsequently diluted with buffer for the final experiment. ITC experiments were conducted using a MicroCal iTC200 System (GE Healthcare) at 25 °C with a stirring speed of 750 rpm using 20 injections as previously described<sup>37</sup>. Specifically, we used an initial 0.4  $\mu$ L injection followed by nineteen 2.0  $\mu$ L injections to determine affinities and thermodynamic parameters. Data were analysed using the Origin 7.0 software (OriginLab Corporation) using a one-site binding model.

**Crystallisation and structure determination.** Complexes of PALS1 PDZ domain in HEPES buffer with coronavirus peptides were reconstituted by mixing protein and peptide were mixed at a molar ratio of 1:4 and then concentrated to 4 mg/ml using 3 kDa cut-off (Amicon® Ultra 0.5) at 4 °C. Concentrated protein complex samples were subjected to high-throughput crystallisation screening using a Gryphon LCP (Art Robbins Instruments) with 0.2  $\mu$ L of protein sample and 0.2  $\mu$ L reservoir solution per drop using 96-well sitting drop trays (swiss, Neuheim, Switzerland) and a range of commercial screens using the sitting drop vapour diffusion method at 20 °C in-house.

Crystals were mounted on nylon and copper loops (MiTGen). All data were collected at the Australian Synchrotron using the MX2 beamline equipped with the Eiger 16 M detector (Dectris) with an oscillation range of 0.1° per frame with a wavelength of 0.9537, integrated using XDS<sup>38</sup> and scaled using AIMLESS<sup>39</sup>.

PALS1 PDZ:SARS-CoV-1\_E crystals were grown in 0.1 M Bis-Tris pH 6.5, 25 % w/v PEG 3350 and flash cooled at –173 °C in the mother liquor. The PALS1 PDZ:SARS-CoV-1 complex formed plate-shaped crystals belonging to space group P1 with  $a = 28.15$  Å,  $b = 39.96$  Å,  $c = 40.83$  Å,  $\alpha = 94.90^\circ$ ,  $\beta = 108.78^\circ$ ,  $\gamma = 100.84^\circ$ . Molecular replacement was carried out using PHASER<sup>40</sup> with the previously solved structure of apo PALS1 PDZ domain (PDB ID: 4UU6)<sup>29</sup> as a search model. PALS1 PDZ:SARS-CoV-1\_E crystals contained two molecules each of PALS1 PDZ and SARS-CoV-1 peptide in the asymmetric unit, with 43.5% solvent content. The final TFZ and LLG values after molecular replacement were 11.9 and 113.72, respectively. The final model was built manually over several cycles using Coot<sup>41</sup> and refined using PHENIX with a final Rwork/Rfree of 0.16/0.19, with 98.3% of residues in the favoured region of the Ramachandran plot and 0% of rotamer outliers.

PALS1 PDZ:SARS-CoV-2\_E crystals were grown in 0.2 M Sodium acetate trihydrate, 0.1 M TRIS pH 8.5, 30% w/v PEG 4000 and flash cooled at –173 °C in mother liquor supplemented with 25% ethylene glycol. The PALS1 PDZ:SARS-CoV-2 complex formed prism-shaped crystals belonging to space group P2<sub>1</sub> with  $a = 28.69$  Å,  $b = 103.77$  Å,  $c = 59.07$  Å,  $\alpha = 90.00^\circ$ ,  $\beta = 103.97^\circ$ ,  $\gamma = 120^\circ$ . Molecular replacement was carried out using PHASER<sup>40</sup> with the previously solved structure of apo PALS1 PDZ domain (PDB ID: 4UU6)<sup>29</sup> as a search model. PALS1 PDZ:SARS-CoV-2\_E crystals contained 4 molecules each of PALS1 PDZ and SARS-CoV-2 peptide in the asymmetric unit, with 47% solvent content. The final TFZ and LLG values after molecular replacement were 15.3 and 454.1, respectively. The final model was built manually over several cycles using Coot<sup>41</sup> and refined using PHENIX<sup>42</sup> with a final Rwork/Rfree of 0.19/0.23, with 99.4% of residues in the favoured region of the Ramachandran plot and 0% of rotamer outliers.

**Statistics and reproducibility.** All isothermal titration calorimetry experiments were performed in triplicates from discrete samples, with errors calculated as  $\pm$ SD.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data supporting the findings of this manuscript are available from the corresponding authors upon reasonable request. Coordinate files were deposited at the Protein Data Bank (<https://www.rcsb.org/>) using accession codes 7NTJ and 7NTK for PALS1 PDZ:SARS-CoV-1\_E and PALS1 PDZ:SARS-CoV-2\_E, respectively. The raw X-ray diffraction data were deposited at the SGrid Data Bank<sup>43</sup> (<https://data.sgrid.org/data/>) using their PDB accession code 7NTJ and 7NTK for PALS1 PDZ:SARS-CoV-1\_E and PALS1 PDZ:SARS-CoV-2\_E, respectively.

Received: 25 March 2021; Accepted: 20 May 2021;

Published online: 11 June 2021

## References

- Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
- Polverino, F. & Kheradmand, F. COVID-19, COPD, and AECOPD: immunological, epidemiological, and clinical aspects. *Front Med (Lausanne)* **7**, 627278 (2020).
- Tumban, E. Lead SARS-CoV-2 candidate vaccines: expectations from phase III trials and recommendations post-vaccine approval. *Viruses* <https://doi.org/10.3390/v13010054> (2020).
- Dai, L. & Gao, G. F. Viral targets for vaccines against COVID-19. *Nat. Rev. Immunol.* **21**, 73–82 (2021).
- Yao, H. et al. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730–738 (2020). e713.
- Ruch, T. R. & Machamer, C. E. The coronavirus E protein: assembly and beyond. *Viruses* **4**, 363–382 (2012).
- Corman, V. M. et al. Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat. *J. Virol.* **88**, 11297–11303 (2014).
- Netland, J. et al. Immunization with an attenuated severe acute respiratory syndrome coronavirus deleted in E protein protects against lethal respiratory disease. *Virology* **399**, 120–128 (2010).
- Schoeman, D. & Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virol. J.* **16**, 69 (2019).
- Pervushin, K. et al. Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS Pathog.* **5**, e1000511 (2009).
- Surya, W., Li, Y. & Torres, J. Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim. Biophys. Acta Biomembr.* **1860**, 1309–1317 (2018).
- Lee, H. J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* **8**, 8 (2010).
- Teoh, K. T. et al. The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Mol. Biol. Cell* **21**, 3838–3852 (2010).
- Nelson, W. J. Adaptation of core mechanisms to generate cell polarity. *Nature* **422**, 766–774 (2003).
- Dow, L. E. & Humbert, P. O. Polarity regulators and the control of epithelial architecture, cell migration, and tumorigenesis. *Int. Rev. Cytol.* **262**, 253–302 (2007).
- Assemat, E., Bazellieres, E., Pallesi-Pocachard, E., Le Bivic, A. & Massey-Harroche, D. Polarity complex proteins. *Biochim. Biophys. Acta* **1778**, 614–630 (2008).
- McCaffrey, L. M. & Macara, I. G. Widely conserved signaling pathways in the establishment of cell polarity. *Cold Spring Harb. Perspect. Biol.* **1**, a001370 (2009).
- Roh, M. H., Liu, C. J., Laurinec, S. & Margolis, B. The carboxyl terminus of zona occludens-3 binds and recruits a mammalian homologue of discs lost to tight junctions. *J. Biol. Chem.* **277**, 27501–27509 (2002).
- Hurd, T. W., Gao, L., Roh, M. H., Macara, I. G. & Margolis, B. Direct interaction of two polarity complexes implicated in epithelial tight junction assembly. *Nat. Cell Biol.* **5**, 137–142 (2003).
- Roh, M. H., Fan, S., Liu, C. J. & Margolis, B. The Crumbs3-Pals1 complex participates in the establishment of polarity in mammalian epithelial cells. *J. Cell Sci.* **116**, 2895–2906 (2003).
- Straight, S. W. et al. Loss of PALS1 expression leads to tight junction and polarity defects. *Mol. Biol. Cell* **15**, 1981–1990 (2004).
- Wang, Q., Chen, X. W. & Margolis, B. PALS1 regulates E-cadherin trafficking in mammalian epithelial cells. *Mol. Biol. Cell* **18**, 874–885 (2007).
- Li, Y. et al. Structure of Crumbs tail in complex with the PALS1 PDZ-SH3-GK tandem reveals a highly specific assembly mechanism for the apical Crumbs complex. *Proc. Natl Acad. Sci. USA* **111**, 17444–17449 (2014).
- Margolis, B. The Crumbs3 Polarity Protein. *Cold Spring Harb. Perspect. Biol.* <https://doi.org/10.1101/cshperspect.a027961> (2018).
- Thomas, M. & Banks, L. Upsetting the balance: when viruses manipulate cell polarity control. *J. Mol. Biol.* **430**, 3481–3503 (2018).
- Songyang, Z. et al. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275**, 73–77 (1997).
- Jimenez-Guardeno, J. M. et al. The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS Pathog.* **10**, e1004320 (2014).
- Toto, A. et al. 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.* **29**, 2038–2042 (2020).
- Ivanova, M. E. et al. 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 Biol. Crystallogr.* **71**, 555–564 (2015).
- Tseng, C. T. et al. Apical entry and release of severe acute respiratory syndrome-associated coronavirus in polarized Calu-3 lung epithelial cells. *J. Virol.* **79**, 9470–9479 (2005).



31. Tao, X. et al. Bilateral entry and release of Middle East respiratory syndrome coronavirus induces profound apoptosis of human bronchial epithelial cells. *J. Virol.* **87**, 9953–9958 (2013).
32. Brinkmann, B. F. et al. VE-cadherin interacts with cell polarity protein Pals1 to regulate vascular lumen formation. *Mol. Biol. Cell* **27**, 2811–2821 (2016).
33. Bosmuller, H., Matter, M., Fend, F. & Tzankov, A. The pulmonary pathology of COVID-19. *Virchows Arch.* **478**, 137–150 (2021).
34. Ackermann, M., Mentzer, S. J. & Jonigk, D. Pulmonary Vascular Pathology in Covid-19. *Reply N. Engl. J. Med.* **383**, 888–889 (2020).
35. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
36. Scopes, R. K. Measurement of protein by spectrophotometry at 205 nm. *Anal. Biochem.* **59**, 277–282 (1974).
37. Maddumage, J. C., Stewart, B. Z., Humbert, P. O. & Kvensakul, M. Crystallographic studies of PDZ domain-peptide interactions of the Scribble polarity module. *Methods and Protocols* **2256**, 125–135 [https://doi.org/10.1007/978-1-0716-1166-1\\_7](https://doi.org/10.1007/978-1-0716-1166-1_7) (2021).
38. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
39. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
40. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 32–41 (2007).
41. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
42. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
43. Meyer, P. A. et al. Data publication with the structural biology data grid supports live analysis. *Nat. Commun.* **7**, 10882 (2016).
44. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
45. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

## Acknowledgements

We thank staff at the MX beamlines at the Australian Synchrotron for help with X-ray data collection. We thank the ACRF for their support of the Eiger MX detector at the Australian Synchrotron MX2 beamline and the Comprehensive Proteomics Platform at La Trobe University for core instrument support. This research was funded by La Trobe University (Scholarship to A.J.).

## Author contributions

A.J.: Experimental design, acquisition of data; analysis and interpretation of data; drafting and revising the article; P.O.H.: Conception and design; analysis and interpretation of data; drafting and revising the article; M.K.: Conception and design; acquisition of data; analysis and interpretation of data; drafting and revising the article.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-02250-7>.

**Correspondence** and requests for materials should be addressed to P.O.H. or M.K.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Jung-Eun Lee and George Inglis. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021