

# MERS-CoV ORF4b employs an unusual binding mechanism to target IMP $\alpha$ and block innate immunity

Thilini S. Munasinghe<sup>1,7</sup>, Megan R. Edwards<sup>2,7</sup>, Sofiya Tsimbalyuk<sup>1</sup>, Olivia A. Vogel<sup>2</sup>, Kate M. Smith<sup>3</sup>, Murray Stewart<sup>4</sup>, Justin K. Foster<sup>1</sup>, Loretta A. Bosence<sup>1</sup>, David Aragão<sup>5</sup>, Justin A. Roby<sup>1,8</sup>✉, Christopher F. Basler<sup>2,6,8</sup>✉ & Jade K. Forwood<sup>1,8</sup>✉

The MERS coronavirus (MERS-CoV) is a highly pathogenic, emerging virus that produces accessory proteins to antagonize the host innate immune response. The MERS-CoV ORF4b protein has been shown to bind preferentially to the nuclear import adapter IMP $\alpha$ 3 in infected cells, thereby inhibiting NF- $\kappa$ B-dependent innate immune responses. Here, we report high-resolution structures of ORF4b bound to two distinct IMP $\alpha$  family members. Each exhibit highly similar binding mechanisms that, in both cases, lack a prototypical Lys bound at their P2 site. Mutations within the NLS region dramatically alter the mechanism of binding, which reverts to the canonical P2 Lys binding mechanism. Mutational studies confirm that the novel binding mechanism is important for its nuclear import, IMP $\alpha$  interaction, and inhibition of innate immune signaling pathways. In parallel, we determined structures of the nuclear binding domain of NF- $\kappa$ B component p50 bound to both IMP $\alpha$ 2 and  $\alpha$ 3, demonstrating that p50 overlaps with the ORF4b binding sites, suggesting a basis for inhibition. Our results provide a detailed structural basis that explains how a virus can target the IMP $\alpha$  nuclear import adapter to impair immunity, and illustrate how small mutations in ORF4b, like those found in closely related coronaviruses such as HKU5, change the IMP $\alpha$  binding mechanism.

<sup>1</sup>School of Biomedical Sciences, Charles Sturt University, Wagga Wagga, NSW 2678, Australia. <sup>2</sup>Center for Microbial Pathogenesis, Institute for Biomedical Sciences, Georgia State University, Atlanta, GA 30303, USA. <sup>3</sup>Paul Scherrer Institut, Swiss Light Source, 5232 Villigen PSI, Switzerland. <sup>4</sup>MRC Laboratory of Molecular Biology, Francis Crick Ave., Cambridge Biomedical Campus, Cambridge CB2 0QH, UK. <sup>5</sup>Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK. <sup>6</sup>Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. <sup>7</sup>These authors contributed equally: Thilini S. Munasinghe, Megan R. Edwards. <sup>8</sup>These authors jointly supervised: Justin A. Roby, Christopher F. Basler, Jade K. Forwood. ✉email: [jroby@csu.edu.au](mailto:jroby@csu.edu.au); [chris.basler@mssm.edu](mailto:chris.basler@mssm.edu); [jforwood@csu.edu.au](mailto:jforwood@csu.edu.au)

The Middle East respiratory syndrome-related coronavirus (MERS-CoV) is a single-stranded positive-sense RNA virus that is closely related to other important human-infecting coronaviruses including severe acute respiratory syndrome-related coronavirus-1 (SARS-CoV-1) and SARS-CoV-2<sup>1</sup>. MERS-CoV is a highly pathogenic zoonotic agent that belongs to the subgenus *Merbecovirus* of the genus *Betacoronavirus*<sup>2</sup>. It emerged in Saudi Arabia in 2012 and has caused 2494 infections and 858 deaths to date<sup>3</sup>. MERS-CoV has demonstrated a capacity to quickly spread to new geographic areas, causing a major outbreak in South Korea in 2015<sup>4</sup>. Although MERS-CoV has thus far proven less transmissible between people than SARS-CoV-2, it has a markedly higher case fatality rate<sup>5,6</sup>, and so remains a concerning latent pandemic threat. Consequently, generating an improved understanding of the MERS-CoV-host interactions that drive severe pathological outcomes remains a critical prerequisite to understand the disease and to develop novel therapeutic strategies.

The MERS-CoV ORF4b protein is a viral accessory protein that functions as an innate immune inhibitor (Fig. 1A) and strongly localizes to the nucleus, despite virus replication occurring exclusively within the cytoplasmic compartment<sup>7</sup>. This nuclear localization distinguishes ORF4b from other phosphodiesterase (PDE) proteins encoded by RNA viruses, such as the NS2 protein of betacoronaviruses within the subgenus *Embecovirus*, and VP3 of group A rotaviruses; both of which lack an NLS motif<sup>8–10</sup>. Interestingly the mammalian host PDE-domain containing A-kinase anchoring protein 7 gamma (AKAP7γ) (which has been proposed as the common genetic origin of ORF4b, NS2, and VP3<sup>8,9</sup>) does contain an N-terminal NLS motif and localizes strongly to the nucleus<sup>8</sup>, however, the NLS region is not homologous to that of merbecovirus ORF4b. ORF4b nuclear localization has been attributed to two short clusters of basic residues (nuclear localization signal sites NLS1 and NLS2) in the ORF4b N-terminal domain, while the C-terminus harbors a PDE domain (Fig. 1A)<sup>7,10</sup>. The innate immune inhibition functions attributed to ORF4b include inhibition of interferon beta (IFNβ) and IFN lambda (IFNλ) production, as well as inhibition of cytokine production that is mediated by the NF-κB p50/p65 heterodimer<sup>7,11–14</sup>. ORF4b PDE activity inhibits the innate antiviral 2′-5′ oligoadenylate (2–5 A) synthetase (OAS)-RNase L pathway by degrading 2–5 A produced by OAS, thus preventing activation of the antiviral enzyme RNase L via the 2–5A second messenger molecule<sup>10,12</sup>.

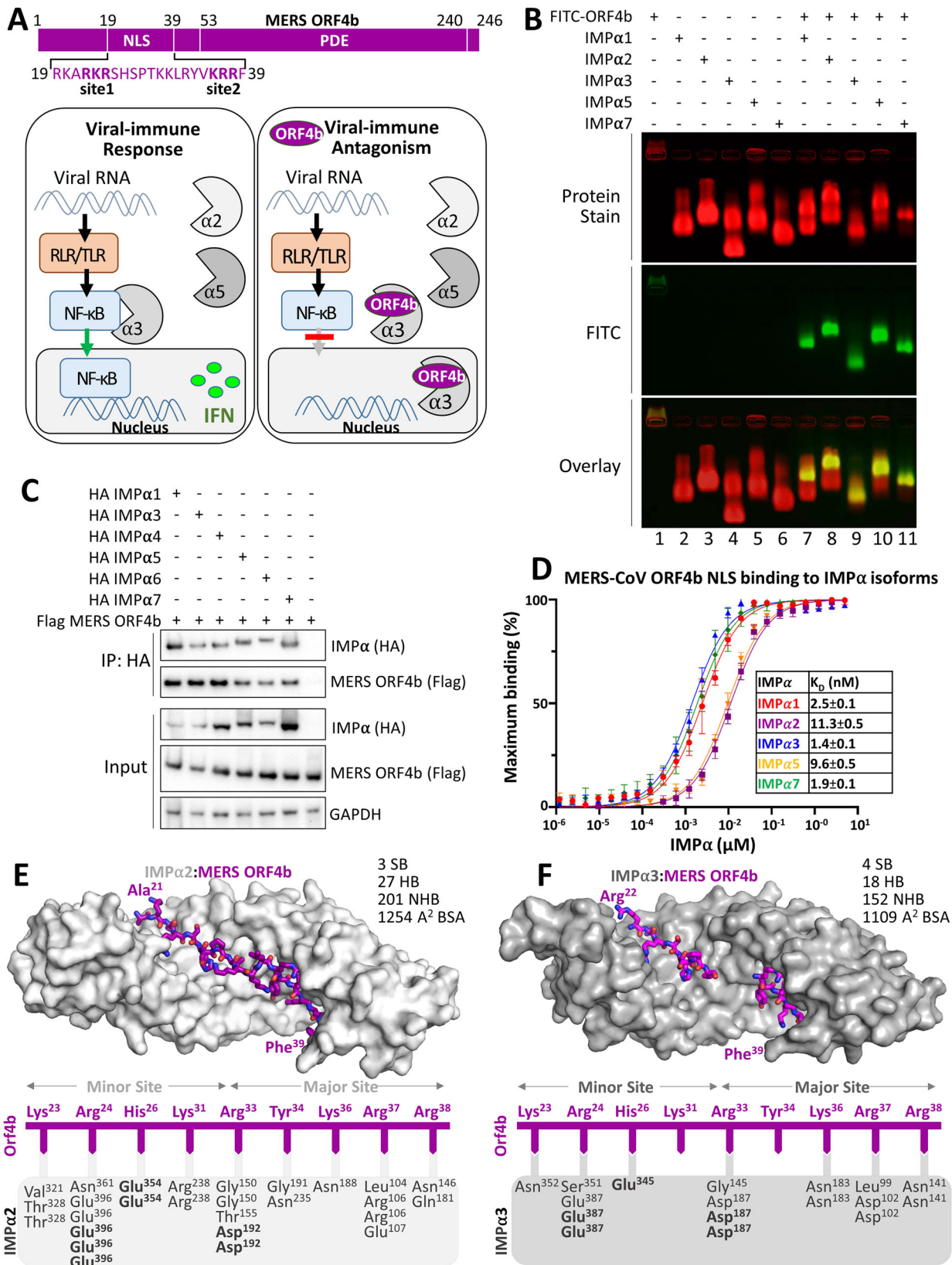
Although several studies demonstrated that ORF4b contributes to innate immune inhibition<sup>7,10–12,15</sup> the mechanisms by which the NLSs facilitate such functions are incompletely understood. One notable study found that, in the context of MERS-CoV infection, ORF4b blocks the nuclear translocation of the p65 subunit of the NF-κB heterodimer, thereby blocking NF-κB-dependent cytokine production<sup>11</sup>. This was attributed to ORF4b binding preferentially (via its NLS sequences) to the host nuclear import protein IMPα3 (also known as KPNA4), which is responsible for the nuclear translocation of NF-κB. Consistent with this hypothesis, wild-type ORF4b, but not ORF4b NLS mutants, outcompete p65 for IMPα3 binding in immunoprecipitation assays<sup>11</sup>. These data place ORF4b among several viral proteins that reportedly block innate immune signaling by disrupting IMPα interactions with key host antiviral transcription factors<sup>16–18</sup>. Other examples include Ebola virus VP24<sup>16,19</sup>, hantavirus nucleoprotein<sup>17</sup>, and Japanese encephalitis virus NS5<sup>18</sup>.

The structural basis for the ORF4b:IMPα interaction and the competition with other IMPα cargo are important prerequisites for understanding how viral proteins can interfere with IMPα nuclear import functions to suppress innate immunity.

Additionally, investigating functional variation within NLS regions among MERS-CoV and related merbecoviruses (including the bat merbecovirus HKU5) has the potential to identify altered interactions with nuclear import receptors with important consequences for host range, virulence and pathogenesis<sup>7</sup>. Here we provide a detailed, structural basis for the interaction between MERS-CoV ORF4b and its innate immune suppression target, the nuclear import adapter IMPα. We show that ORF4b binds to both the major and minor NLS binding sites on the IMPαs through a novel and highly unusual way that is conserved across two IMPα family members. Through structure-guided mutational analysis, we show that this interface is important for MERS ORF4b-mediated inhibition of innate immune signaling. Structural analysis of the NF-κB component p50 bound to IMPα2 and IMPα3 demonstrates that there is a direct overlap between the binding sites for this host transcription factor component and its competitive ORF4b inhibitor. Finally, we observe that, despite harboring conserved elements within these NLS regions, the precise way in which IMPαs interact with ORF4b from related viruses cannot be predicted from sequence alone, with distinct IMPα binding mechanisms existing for MERS-CoV related merbecoviruses including HKU5.

## Results

**MERS-CoV ORF4b bind IMPα subfamily members with high affinity.** MERS-CoV ORF4b plays an important role in innate immune inhibition, where it is believed to block NF-κB-dependent cytokine production through preferential binding to the host nuclear import protein IMPα3 via its NLS sequences (sites 1 and 2 spanning residues 19–39; Fig. 1A)<sup>11</sup>. Understanding the mechanisms through which cargo binds specifically to IMPα isoforms is complicated by the high structural and sequence similarity across IMPα isoforms (particularly in the cargo binding regions) and has only been described recently for a small number of cargo<sup>20</sup>. Because MERS-CoV ORF4b has been reported to bind specifically to IMPα3, we sought to analyze the interaction interface using a combination of biophysical and structural approaches. Using recombinantly expressed IMPα isoforms spanning members from each of the three subfamilies (SF1: IMPα1/2; SF2: IMPα3; SF3: IMPα5/7), agarose gel mobility shift assays were performed in the absence and presence of a FITC-tagged MERS-CoV ORF4b peptide spanning residues 19–39 (henceforth simply referred to as “NLS”). In the absence of binding partners, the FITC-tagged peptide migrated towards the cathode (Fig. 1B, lane 1), whereas the IMPα proteins migrated towards the anode (Fig. 1B, lanes 2–6). In the presence of IMPα isoforms, the FITC-tagged peptide shifted migration towards the anode, forming a distinct band that migrated differently to both IMPα-only and peptide-only samples. These results suggest that the FITC-tagged ORF4b NLS peptide was capable of binding isoforms from each of the three IMPα subfamilies. To compare this result within a human cell expression context, we performed co-immunoprecipitation assays against respective full-length, HA-tagged IMPα isoforms, and probed for the presence of full-length FLAG-tagged MERS-CoV ORF4b. Consistent with the gel mobility shift assays, we found that MERS-CoV ORF4b bound to all IMPα isoforms tested (Fig. 1C and Supplementary Fig. 1). Next, we quantitatively measured the direct binding interaction with MERS-CoV ORF4b NLS to respective IMPα isoforms using fluorescence polarization. We observed very strong binding in the low nM range across all IMPα isoforms (Fig. 1D SF1: IMPα1, 2.5 nM; SF2: IMPα3, 1.4 nM; SF3: IMPα7, 1.9 nM). Indeed, the slight differences in the low nanomolar binding affinities of MERS-CoV ORF4b NLS to IMPα1, IMPα3, and IMPα7 were determined not to be statistically significant (one-way ANOVA



followed by Tukey’s test,  $p > 0.1$ ); therefore, interactions with the NLS region of this viral factor were equally strong across members of the three IMPα families. Because the results indicated that MERS-CoV ORF4b NLS was able to bind IMPα isoforms with high affinity and showed little preference for a specific isoform, we examined the molecular interface with two IMPα isoforms

using high-resolution structural approaches. Specifically, we crystallized the MERS-CoV ORF4b NLS bound to both IMPα2 and IMPα3 and resolved the structures to 2.1 Å and 2.5 Å resolution respectively (see Supplementary Table 1 for data collection and refinement statistics). We identified consistent binding patterns, with MERS-CoV ORF4b residues Lys<sup>23</sup>, Arg<sup>24</sup>, His<sup>26</sup>,



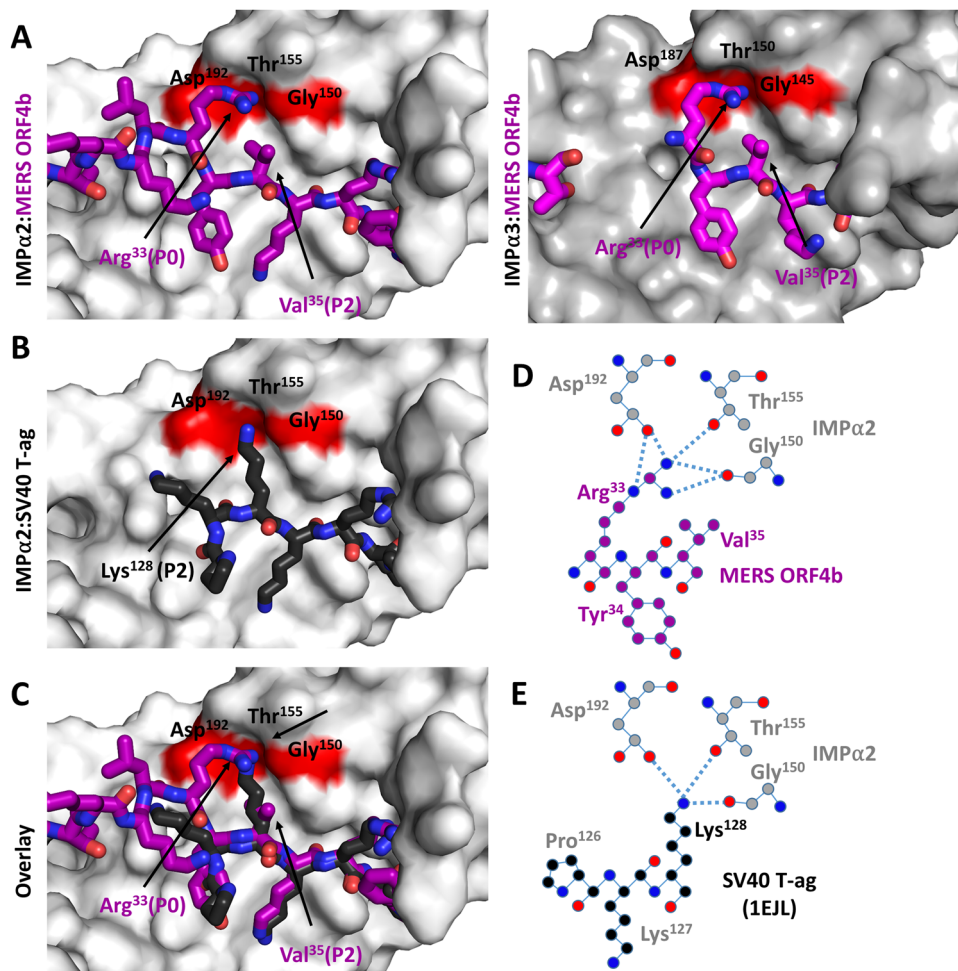
**Fig. 1 MERS ORF4b binding to IMP $\alpha$  is not isoform specific.** **A** Schematic overview of the ORF4b protein. The region responsible for nuclear localization is contained with the N-terminus and spans two sites within residues 19–39. This region has been shown to be required for ORF4b to block the nuclear translocation of the p65 subunit of NF- $\kappa$ B, thereby blocking NF- $\kappa$ B-dependent cytokine production, by binding specifically to IMP $\alpha$ 3. **B** Electro-mobility shift assay (EMSA) showing MERS ORF4b NLS binds with  $\Delta$ IBB-IMP $\alpha$  isoforms spanning members from each of the three subfamilies (SF1: IMP $\alpha$ 1/2; SF2: IMP $\alpha$ 3; SF3: IMP $\alpha$ 5/7). The MERS ORF4b NLS region peptide spans residues 19–39 and contains a FITC and Ahx linker (middle panel). Proteins were stained using Coomassie blue stain (top panel; red), and the overlay is represented in the bottom panel, where the complex produces a yellow color. **C** Co-immunoprecipitation (coIP) assay against HA-tagged IMP $\alpha$  isoforms and probed for the presence of full-length FLAG-tagged MERS-CoV ORF4b. GAPDH was used as an internal control to ensure consistent quantity of cell extracts were used in each assay. CoIP was repeated twice independently with similar results. **D** Fluorescence polarization assay measuring the direct binding between MERS-CoV ORF4b NLS and respective IMP $\alpha$  isoforms. Strong binding in the 1–10 nM range was observed across all IMP $\alpha$  isoforms. Each family member was able to bind in the low nM range (SF1: IMP $\alpha$ 1 = 2.5 nM; SF2: IMP $\alpha$ 3 = 1.4 nM; SF3: IMP $\alpha$ 7 = 1.9 nM). Source data are provided as a Source Data file.  $n = 3$  biologically independent experiments. Error bars represent mean values  $\pm$  SEM. **E** Crystal structure of MERS-CoV ORF4b NLS bound to IMP $\alpha$ 2 at 2.1 Å resolution. The IMP $\alpha$  is shown in light gray and surface mode, and the ORF4b shown in purple in stick mode. Pymol software was used to generate images<sup>45</sup>. A cartoon representation is shown below highlighting the hydrogen bonds and salt bridges (in bold). HB number of hydrogen bonds, SB number of salt bridges, NHB number of non-hydrogen bonds generated in PDBsum, BSA buried surface area calculated in PISA. **F** Crystal structure of MERS-CoV ORF4b NLS bound to IMP $\alpha$ 3 at 2.5 Å resolution. The IMP $\alpha$  is shown in dark gray, and all other parameters are as per 1E. NLS nuclear localization signal, PDE phosphodiesterase domain, MERS Middle East respiratory syndrome-related coronavirus, IMP $\alpha$  importin alpha,  $\alpha$  importin alpha, RLR RIG-I-like receptor, TLR Toll-like receptor, NF- $\kappa$ B Nuclear factor kappa light chain enhancer of activated B cells, IFN interferon, GAPDH Glyceraldehyde 3-phosphate dehydrogenase, HA haemagglutinin epitope tag, FITC fluorescein isothiocyanate.

Arg<sup>33</sup>, Arg<sup>37</sup>, and Arg<sup>38</sup> bound to both IMP $\alpha$ 2 and IMP $\alpha$ 3 at structurally equivalent interfaces (Fig. 1E, F) (see Supplementary Tables 2–3 for hydrogen bond interactions; Supplementary Fig. 2). In the MERS-CoV ORF4b:IMP $\alpha$ 2 structure, we were able to resolve slightly more of the MERS-CoV ORF4b including the region Lys<sup>30</sup> Lys<sup>31</sup>, Leu<sup>32</sup> between the major and minor sites. There were also two additional ORF4b residues contributing to IMP $\alpha$ 2 binding when compared to the IMP $\alpha$ 3 structure (Lys<sup>31</sup> and Tyr<sup>34</sup>), however, these did not appear to contribute significantly to the binding affinity (Fig. 1D). Overall, our structural and biochemical data confirm a direct and high-affinity interaction between MERS-CoV ORF4b and IMP $\alpha$  subfamilies without a marked preference between isoforms.

**MERS-CoV ORF4b interacts with IMP $\alpha$ 2 and IMP $\alpha$ 3 using a novel Arg P0 binding mechanism rather than the canonical Lys P2 at the major NLS binding site.** Our high-resolution crystal structures showed that MERS-CoV ORF4b occupies an extensive interface across both IMP $\alpha$ 2 and IMP $\alpha$ 3. MERS-CoV ORF4b bound both the major NLS binding site, through residues Arg<sup>33</sup>, Lys<sup>36</sup>, Arg<sup>37</sup>, and Arg<sup>38</sup>, as well as the minor site through residues Lys<sup>23</sup>, Arg<sup>24</sup>, and His<sup>26</sup>, in both IMP $\alpha$ 2 and IMP $\alpha$ 3 (Fig. 1E, F). Strikingly, the binding interface of both structures lacked a Lys at the P2 position within the major site (Fig. 2A) that is seen with other NLSs. This was highly unusual because this site has been shown to contribute 50% of the binding energy at the major site, and a Lys at the P2 position is present in >95% of structures, including IMP $\alpha$  bound structures of Hendra and Nipah virus W proteins<sup>21</sup>, HIV Tat<sup>22</sup>, dengue and Zika virus NS5 proteins<sup>23</sup>, SV40 T-ag (Fig. 2B)<sup>24</sup>, and influenza A virus PB2<sup>25</sup> (see also Supplementary Fig. 3). In the structures of MERS-CoV ORF4b bound to IMP $\alpha$ 2 and IMP $\alpha$ 3, this position is instead occupied by Val<sup>35</sup> and neither its side chain nor main chain formed bonded interactions with IMP $\alpha$ 2 or IMP $\alpha$ 3. To compensate for this loss of binding, we observed that the side chain of Arg<sup>33</sup>, positioned two residues before where a P2 Lys would normally be, was bound at this pocket as can be seen in Fig. 2C that presents an overlay of ORF4b and SV40 T-ag NLS structures. This MERS-CoV ORF4b Arg P0 forms near identical bonding patterns to the canonical Lys P2 major site binding, with both forming H-bonds and salt bridge interactions with Gly<sup>150</sup>, Thr<sup>155</sup>, and Asp<sup>192</sup> on IMP $\alpha$ 2 (Fig. 2D, E), and the equivalent Gly<sup>145</sup>, Thr<sup>150</sup>, Asp<sup>187</sup> interactions with IMP $\alpha$ 3 (Fig. 2A right panel). In summary, MERS-CoV ORF4b interacts in a novel way with both IMP $\alpha$ 2 and IMP $\alpha$ 3.

**Mutations within the MERS-CoV ORF4b NLS alter its binding to IMP $\alpha$ s, switching to a canonical P2 Lys.** To examine the effect of mutations within the key IMP $\alpha$  binding determinants of MERS-CoV ORF4b, we also crystallized MERS-CoV ORF4b wild-type (WT) and mutant NLS peptides in complex with IMP $\alpha$ 2 (Fig. 3A–E; Supplementary Figs. 2 and 4; Supplementary Tables 4–5 for data collection and refinement statistics and Supplementary Tables 6–10 for hydrogen bond interactions). We found that, compared with WT (Fig. 3A), a MERS-CoV ORF4b mutation of Arg<sup>24</sup> to Ala (R24A) bound only at the major site of IMP $\alpha$ 2. Strikingly, with the lack of minor site binding, there was a distinctly different binding register at the major site (Fig. 3B). Here, the major site was occupied with a canonical Lys at the P2 position (MERS-CoV ORF4b Lys<sup>30</sup>), and indeed, the <sup>28</sup>PTKCLR bound at the major site represents a very typical major site NLS. We next examined the effect of a MERS-CoV ORF4b His<sup>26</sup> to Ala (H26A) mutation within the minor site and found that the structure was almost identical to the MERS-CoV WT ORF4b NLS (Fig. 3C). The structure of a MERS-CoV ORF4b Arg<sup>33</sup> to Ala (R33A) mutant bound to IMP $\alpha$ 2 showed a change in the major site binding register similar to that of R24A, with Lys<sup>30</sup> positioned at the P2 site, with minor site binding retained (Fig. 3D). Finally, we examined the crystal structure of a MERS-CoV ORF4b Arg<sup>37</sup> to Ala (R37A) mutant bound to IMP $\alpha$ 2 and found that the major site was also disrupted and formed a Lys<sup>30</sup> P2 register similar to the R24A and R33A mutants (Fig. 3E). Overall, Arg<sup>24</sup> and Arg<sup>33</sup> are key binding determinants within the minor and major sites respectively, with Arg<sup>33</sup> mediating a novel P0 binding site interaction, whereas point mutations alter the binding register to a canonical P2 Lys binding mechanism. Possible implications for merbecovirus evolution are discussed below.

**Mutations to NLS binding sites reduce binding affinities across IMP $\alpha$  family members.** To further understand the effect of key MERS-CoV ORF4b binding determinants at the IMP $\alpha$  interface, we examined the strength of IMP $\alpha$ 2 binding for FITC-tagged ORF4b wild-type (WT) and mutant NLS peptides using a fluorescence polarization assay (Fig. 3F). From this quantitative experiment, we determined that R24A at the minor site binding region and R33A at the major site binding region both reduced IMP $\alpha$ 2 binding to the greatest extent, with binding affinities of 238 nM (significant ~20-fold reduction compared to WT; one-way ANOVA followed by Tukey's test,  $p = 0.0002$ ) and 471 nM (significant ~40-fold reduction compared to WT; one-way ANOVA followed by Tukey's test,  $p < 0.0001$ ), respectively. In



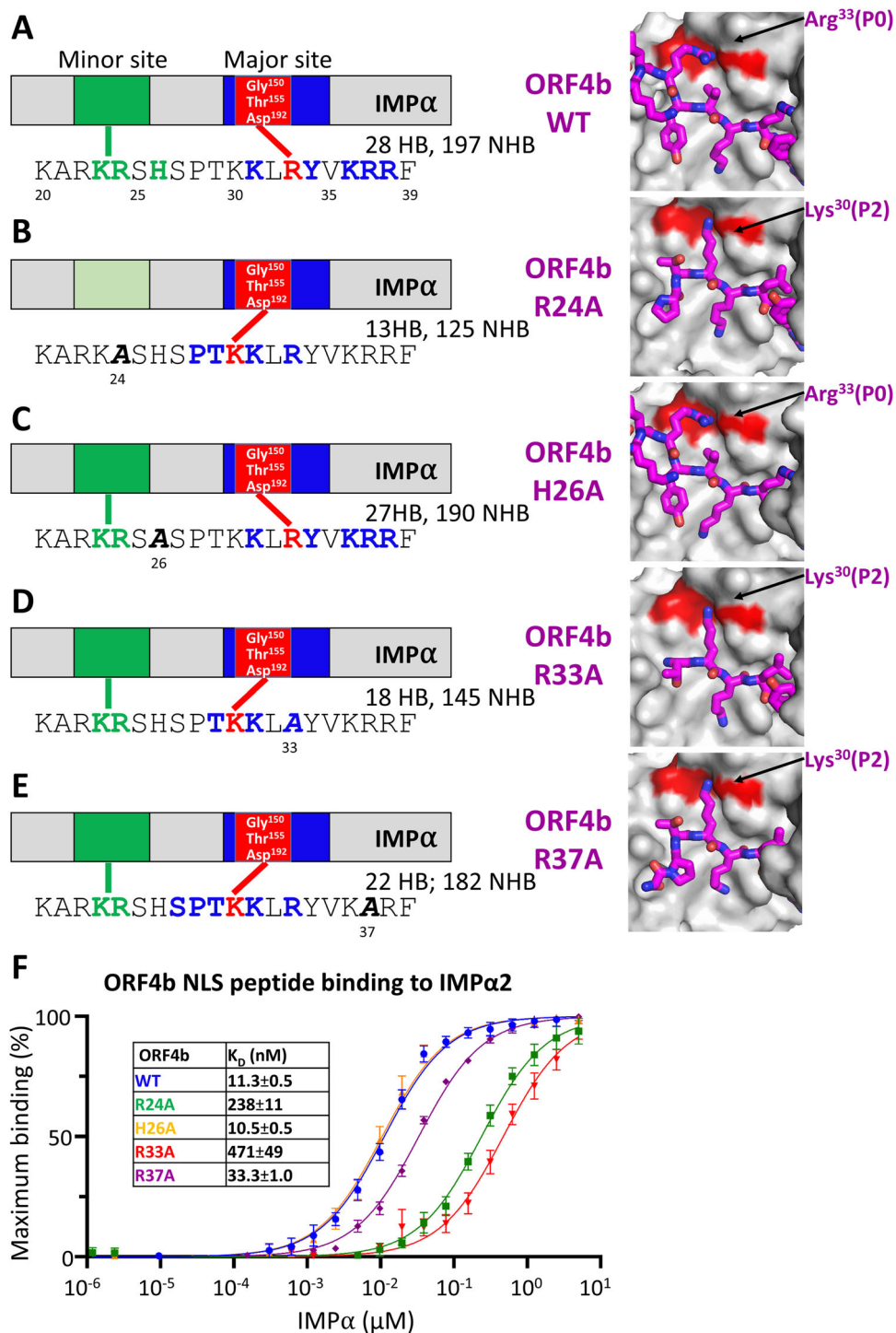
**Fig. 2** MERS ORF4b interacts with IMPα2 via an unusual P0 Arg binding mechanism at the major site. **A** Detailed view of MERS ORF4b bound to the major site of IMPα2 (left) and IMPα3 (right). IMPα2 and α3 are colored light and dark gray respectively and shown in surface view, and the MERS ORF4b is colored magenta and shown in stick view. Pymol software was used to generate images<sup>45</sup>. MERS-CoV ORF4b binds IMPα without the use of a Lys at the P2 position, and instead has a Val at this position. This is compensated by an Arg at the P0 position. **B** Comparison view of the well-characterized SV40 T-ag NLS bound to IMPα<sup>24</sup>, showing a well-characterized Lys at the P2 site. Other viral NLSs using this canonical Lys P2 site are displayed in Supplementary Fig. 2. **C** Overlay of MERS ORF4b and SV40 T-ag binding at the P2 site on IMPα<sup>2</sup>. **D** Detailed hydrogen bonding mediating the MERS ORF4b Arg<sup>33</sup> P0 at the Gly, Thr, Asp bonding pocket of IMPα<sup>2</sup>. Bonds are denoted as dotted lines between atoms, and coloring is preserved within **A**. **E** Detailed hydrogen bonding mediating the SV40 T-ag Lys P2 at the Gly, Thr, Asp bonding pocket of IMPα<sup>2</sup>. Bonds are denoted as dotted lines between atoms, and coloring is preserved within **B**. MERS Middle East respiratory syndrome-related coronavirus, IMPα importin alpha, SV40 T-ag Simian virus 40 large T antigen.

contrast, the R37A mutation reduced IMPα2 binding to a smaller extent (33 nM; ~3-fold reduction, not significant), whereas H26A exhibited IMPα2 binding affinity that was effectively unchanged compared to WT (~10 nM, not significant). Because such highly similar binding patterns were seen for both IMPα2 and IMPα3, we also examined the effect of these interactions across multiple IMPα subfamilies including IMPα1, IMPα3, and IMPα5. We observed a similar trend to that observed with IMPα2, where mutations R24A and R33A resulted in large reductions in binding affinity (Fig. 4A–C). These interactions were also examined in a cellular context through immunoprecipitation assays, where we found that R24A and R33A, as well as a control mutant derived from the literature comprised of R22A/K23A/R24A (RKRA)<sup>7,11</sup>, reduced full-length ORF4b interaction across all IMPα subfamilies (Fig. 4D–F and Supplementary Figs. 5–7).

**MERS-CoV ORF4b mutations affect nuclear localization.** Based on our crystal structure and biochemical data indicating that Arg<sup>24</sup> and Arg<sup>33</sup> play key roles in binding to the minor and

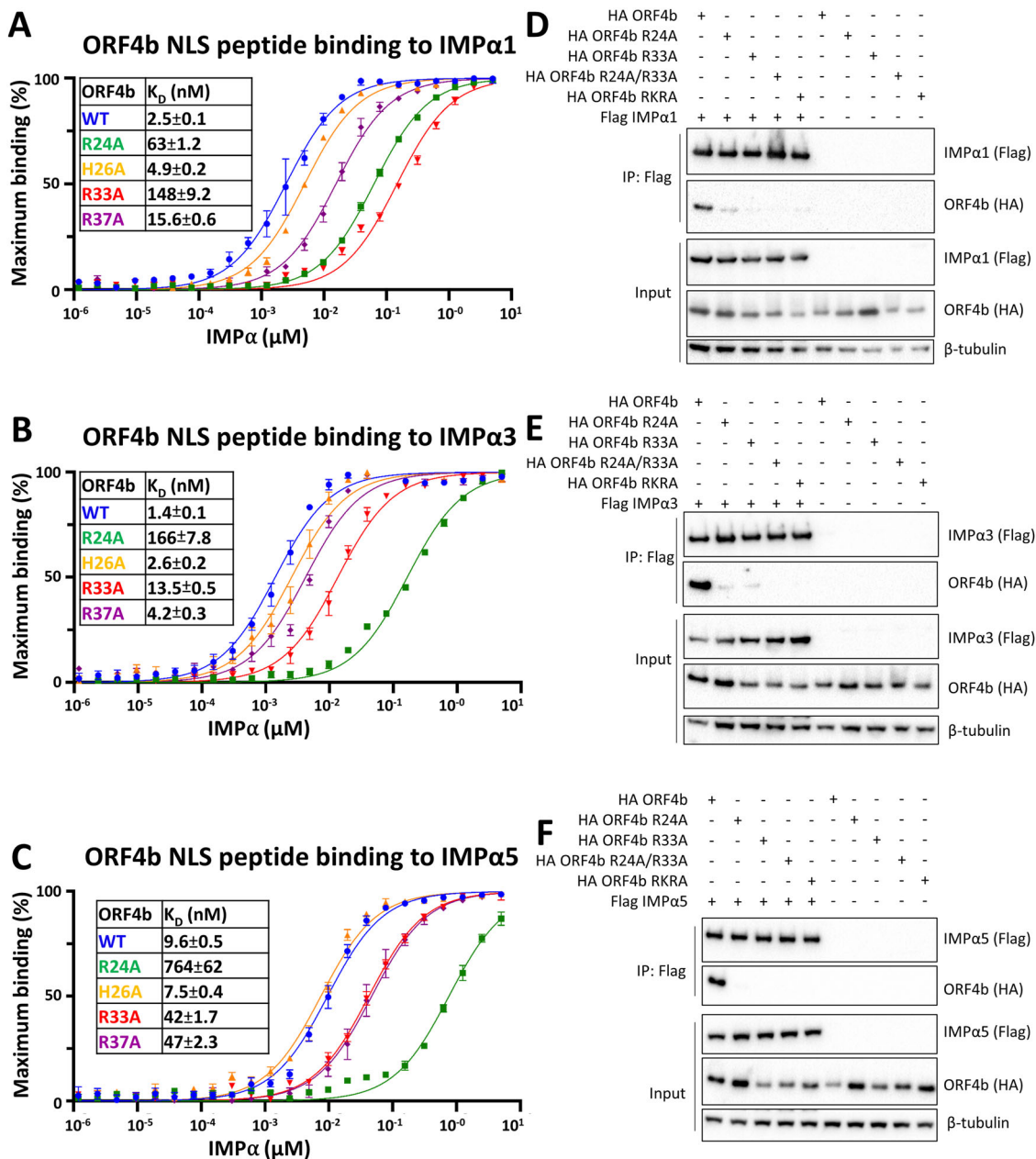
major sites respectively, we examined how these mutations altered nuclear localization of MERS-CoV ORF4b. We found that mutation of R24A alone had a small effect on the nuclear import of ORF4b, whereas mutation of R33A alone, or in combination with R24A (R24A/R33A), significantly decreased the nuclear localization of the mutant ORF4b proteins (Fig. 5A, B). The R24A/R33A double mutant demonstrated a loss of nuclear import similar to the previously described RKRA mutant (R22A/K23A/R24A)<sup>7,11</sup> and an N-terminal deletion mutant (delta38) lacking the first 38 amino acids of ORF4b (Fig. 5A, B).

**NF-κB p50 and MERS-CoV ORF4b bind at overlapping sites on IMPαs.** Previous reports indicated that MERS-CoV ORF4b can inhibit the innate immune response by blocking NF-κB signaling<sup>11</sup>. Since structural data of the binding site of NF-κB on IMPα remains elusive, we tested whether these sites overlapped. We crystallized the well-characterized NLS of p50<sup>26</sup> in complex with both IMPα2 and IMPα3, and found that it bound at the major site of both IMPα2 and IMPα3 through highly similar



**Fig. 3** Mutations to key residues at the MERS ORF4b binding interface change the binding mechanism to canonical Lys P2 site. **A** A wild-type peptide of MERS ORF4b was used as a control and crystallized with IMP $\alpha$ 2. As expected, the Arg33 bound at the Gly150, Thr155, and Asp192 pocket. A summary cartoon is presented, highlighting the minor and major site interactions and the structure presented in the right panel. A list of hydrogen-bonded interactions is presented in Supplementary Table 5. **B** An Arg24Ala mutation within the minor site, denoted in bold and italics with the numbering below the mutation is presented, showing a loss of minor site binding and change in binding mechanism at the major site. A list of hydrogen-bonded interactions is presented in Supplementary Table 6. **C** A His26Ala mutation, depicted as above, had little effect on the binding. A list of hydrogen-bonded interactions is presented in Supplementary Table 7. **D** An Arg33Ala mutation, depicted as above, caused a change in binding at the major site to the canonical P2 Lys binding. A list of hydrogen-bonded interactions is presented in Supplementary Table 8. **E** An Arg37Ala mutation, depicted as above, caused a change in binding at the major site to the canonical P2 Lys binding. A list of hydrogen-bonded interactions is presented in Supplementary Table 9. **F** Fluorescence polarization measuring the binding interaction of IMP $\alpha$ 2 with the crystallized mutants ( $n = 3$  biologically independent experiments). Error bars represent mean values  $\pm$  SEM. R24A and R33A mutants had the greatest effect on binding. Source data are provided as a Source Data file. NLS nuclear localization signal, IMP $\alpha$  importin alpha, WT wild-type, HB number of hydrogen bonds, NHB number of non-hydrogen bonds.

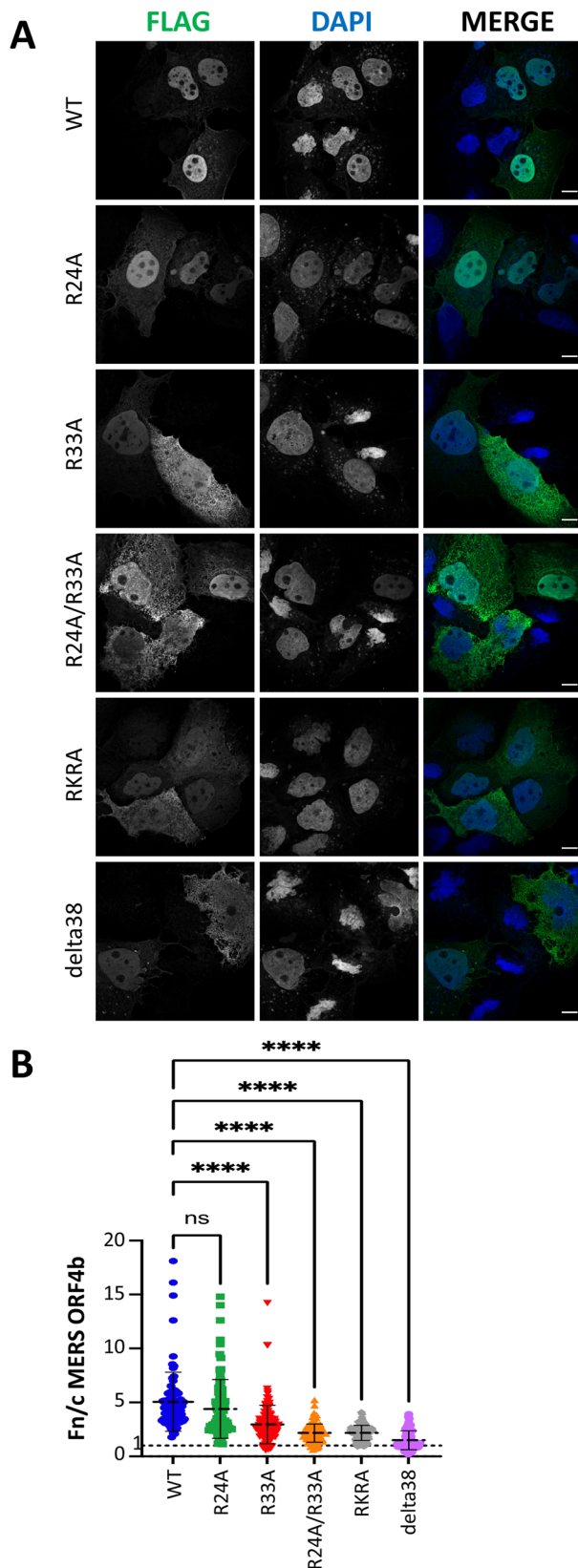




**Fig. 4 Mutations within key binding determinants at the MERS ORF4b binding interface alter binding efficiency to IMPα subfamilies.** Fluorescence polarization measurements were undertaken to determine the affinity of interaction of MERS ORF4b and mutants across the IMPα family members SF1 (A), SF2 (B), SF3 (C). *n* = 3 biologically independent experiments and error bars represent mean values ± SEM for A–C. Source data are provided as a Source Data file. Respective immunoprecipitation assays were performed, measuring the interaction in the strongest mutants (R24A, R33A), and a control mutant RKRA (RKR24 to AAA24) in SF1 (D), SF2 (E), and SF3 (F). Assays were performed against HA-tagged IMPα isoforms and probed for the presence of full-length FLAG-tagged MERS-CoV ORF4b. β-tubulin was used as an internal control to ensure consistent quantity of cell extracts were used in each assay. ColP’s were repeated twice for IMPα1 and α5 and three times for IMPα3. NLS nuclear localization signal, IMPα importin alpha, HA haemagglutinin epitope tag.

interactions involving p50 residues <sup>361</sup>RKRQKM (Fig. 6A, B) in which Lys<sup>362</sup> bound at the P2 site (see Supplementary Fig. 4 and Supplementary Table 11 for data collection and refinement statistics, and Supplementary Tables 12 and 13 for hydrogen bond interactions). Overall, there was clear overlap in the binding sites of MERS-CoV ORF4b and NF-κB on both IMPα2 and IMPα3 (Fig. 6C, D). To examine whether the key binding determinants identified at the MERS-CoV ORF4b:IMPα interface could affect ORF4b inhibition of immune signaling, we tested the ORF4b R24A and R33A mutants for their ability to reduce IFNβ (SeV induction) and NF-κB (TNFα induction) responses. Consistent with the other data, we found that Arg<sup>33</sup>

was the most potent single point mutation in reducing the ability of ORF4b to inhibit immune effects in both HEK293T cells (Fig. 6E, F and Supplementary Figs. 8 and 9) and a human airway epithelial cell-derived cell line A549 (Supplementary Figs. 10 and 11). Because we were unable to obtain similar structural data for an IMPα:p65 complex, we examined the ability of MERS ORF4b to inhibit NF-κB p65 nuclear import during a TNFα response and found a marked reduction in nuclear import of this subunit. Consistent with the above data, the R33A mutant was significantly impaired in its inhibition of p65 nuclear import compared to WT ORF4b (Supplementary Figs. 12 and 13).



**Related merbecoviruses show different IMP $\alpha$  binding mechanisms.** Based on our observation that mutations within the MERS-CoV ORF4b dramatically changed the binding register on IMP $\alpha$ , we examined the conservation of the NLS binding region in related merbecoviruses (Fig. 7A). Because HKU5 ORF4b had been studied previously and shown to traffic to the nucleus<sup>7</sup>, and

**Fig. 5 MERS ORF 4b protein localizes to the nucleus while mutations at the IMP $\alpha$  interface impair nuclear import efficiency.** **A** Confocal microscopy images captured using a Zeiss LSM 800 confocal microscope of Huh7 cells transfected with Flag-tagged MERS ORF4b plasmid and stained using an anti-FLAG tag monoclonal antibody conjugated to Alexa Fluor 488. DAPI stain was used to identify the nuclear boundary within the cell, and merged panels represent an overlay of FLAG and DAPI panels. Significant reductions in nuclear import were observed for the R33A, R24A/R33A, RKR24/AAA24 (RKRA), and delta38 ORF4b constructs. The experiment was performed independently twice, with similar results. The scale bar is 10  $\mu$ m. **B** Quantitation of the nuclear to cytoplasmic fluorescence signal (Fn/c) was measured in 88 cells for the wild-type protein; 85 cells for the R24A mutant protein; 139 cells for the R33A mutant protein; 118 cells for the R24A/R33A; 121 cells for the RKRA mutant protein; and 60 cells for the delta38 truncated protein, with error bars indicating mean values  $\pm$  standard deviation. Statistical significance for the indicated comparisons was determined by a one-way ANOVA, followed by the Dunnett correction: ns no significant difference, \*\*\*\* $p < 0.0001$ . Source data are provided as a Source Data file. MERS Middle East respiratory syndrome-related coronavirus, WT wild-type, DAPI 4',6-diamidino-2-phenylindole, Fn/c fluorescence intensity in the nucleus divided by fluorescence intensity in the cytoplasm.

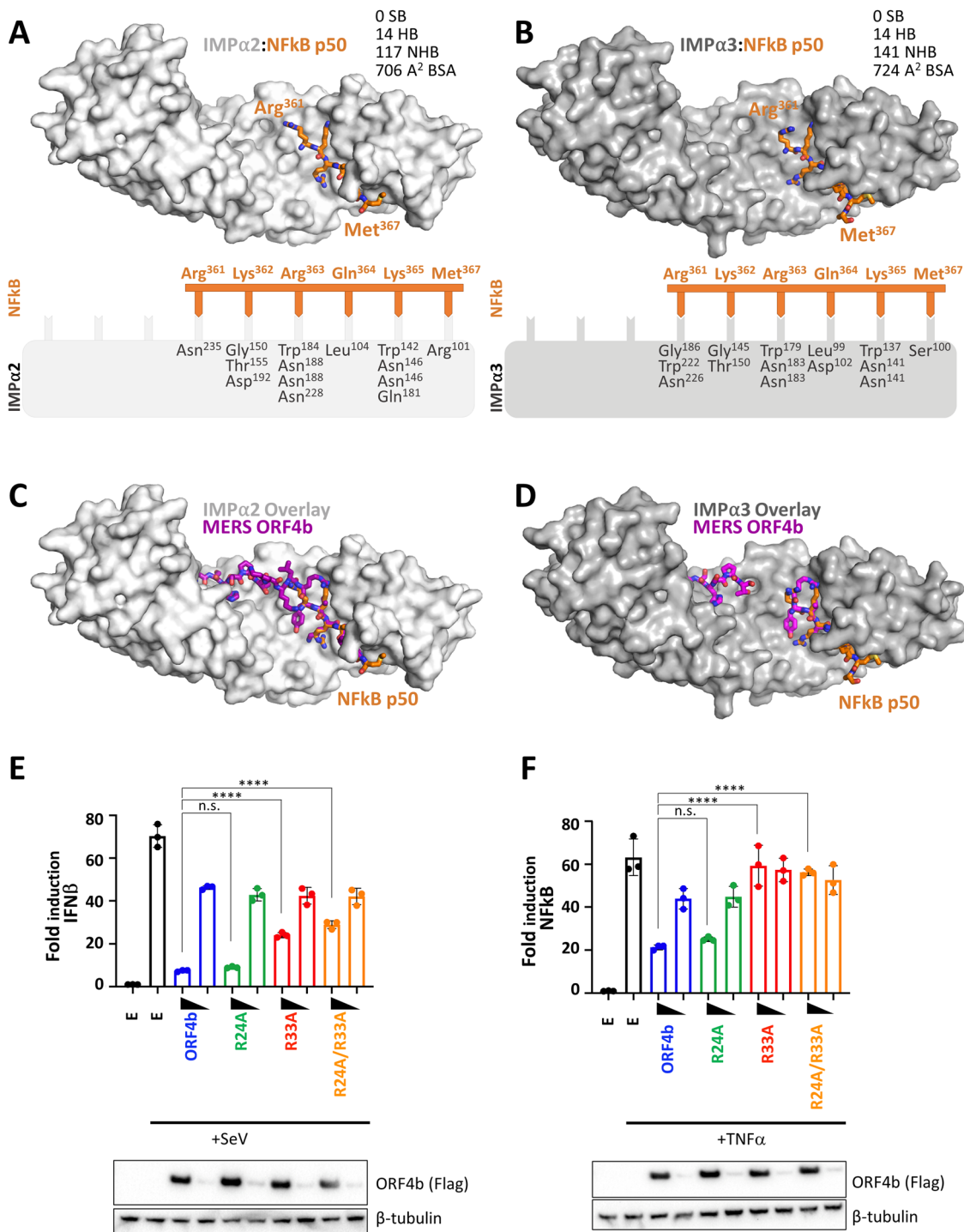
an analysis of the two NLS domains between the two viruses revealed only minor differences (Fig. 7B and Supplementary Fig. 14), we examined how differences could potentially result in changes in the mode of IMP $\alpha$  binding. Accordingly, we crystallized IMP $\alpha$ 2 in complex with HKU5 ORF4b and found that, rather than utilizing the Arg/Val binding register motif of MERS-CoV ORF4b, HKU5 ORF4b bound IMP $\alpha$ 2 using the canonical Lys P2 site (Fig. 7C, D; Supplementary Fig. 4; see Supplementary Table 11 for data and refinement statistics, and Supplementary Table 14 for hydrogen bond interactions). The NLS site also shifted, with HKU5 binding at the major and minor sites with the same sequence, RKRRRHP<sup>37</sup>. The equivalent sequence in MERS-CoV ORF4b, RKRSHSP<sup>28</sup>, lacks two Arg residues. These results highlight that small changes within ORF4b of viruses within the subgenus *Merbecovirus* may result in significant changes in IMP $\alpha$  binding, indicating that a thorough investigation of the binding mechanisms of each would be needed to fully understand how they might influence immune evasion properties.

## Discussion

MERS-CoV ORF4b is a viral accessory protein that inhibits NF- $\kappa$ B-dependent innate immune responses in infected cells by binding preferentially to the nuclear import adapter IMP $\alpha$ 3, one of seven isoforms in humans<sup>11</sup>. The structural data presented here showed that ORF4b binds to IMP $\alpha$ s using a novel interaction geometry and also provides insight into how MERS-CoV ORF4b inhibits innate immunity by impacting the nuclear import of NF- $\kappa$ B.

ORF4b has been reported to bind preferentially to the host nuclear import protein IMP $\alpha$ 3, one of seven isoforms in humans<sup>11</sup>. Mechanisms through which cargoes are able to bind specifically to different IMP $\alpha$  isoforms are poorly understood and this question is complicated by the high degree of structural and sequence homology among IMP $\alpha$  isoform cargo binding regions<sup>21</sup>. To date, only three studies have provided a structural basis for specificity<sup>21,27,28</sup>. Here, by using a combination of structural, biophysical, and cellular data, we have examined the ORF4b interaction with different IMP $\alpha$  isoforms. Our structural data show an extensive interaction interface that involves residues within both of the previously identified sites 1 and 2 of the ORF4b NLS (docked at the minor and major NLS binding sites of IMP $\alpha$ , respectively), together with critical residues between





these sites that had not previously been recognized as important (Fig. 1E, F). Paradoxically, this interaction interface was very similar in both the nuclear import adapter IMP $\alpha$ 2 and in IMP $\alpha$ 3, the latter of which ORF4b reportedly exhibited a binding preference<sup>11</sup>. Probing this interaction interface using in vitro binding studies showed a direct and high-affinity interaction of ~1–10 nM between the ORF4b NLS and representative members from all of the IMP $\alpha$  families (Fig. 1D), consistent with the extensive interface identified in the crystal structures. Moreover, examining these interactions within a cellular context, our immunoprecipitation experiments showed that IMP $\alpha$  isoforms from each subfamily were able to interact with MERS ORF4b (Figs. 1C and 4D–F), however these results

do not necessarily dispute the hypothesis, nor previous studies, that full-length ORF4b may bind IMP $\alpha$ 3 specifically. These same approaches have been used to determine mechanisms of isoform specificity for both viral proteins from our laboratory, including Hendra and Nipah virus W proteins<sup>21</sup>, and host transcription factors including SOX2 protein<sup>27</sup>, as well as studies from other laboratories including RCC1<sup>28</sup>. Additional specificity may be provided by other regions of ORF4b, or it may be that certain cell types or the context of an infected cell may alter ORF4b IMP $\alpha$  isotype specificity. Although further studies will be required to address this possibility, the structural and functional data presented here provide an important platform for such investigations.

**Fig. 6 NF- $\kappa$ B p50 binding on IMP $\alpha$ 2 and IMP $\alpha$ 3 directly overlap with MERS ORF4b. A** Crystal structure of NF- $\kappa$ B p50 bound to IMP $\alpha$ 2 at 2.6 Å resolution (see Supplementary Table 11 for data collection and refinement statistics). IMP $\alpha$ 2 is shown in light gray and surface mode, and the p50 is shown in orange in stick mode. Pymol software was used to generate images<sup>45</sup>. A cartoon representation is shown below highlighting the hydrogen bonds (see Supplementary Table 12 for a full list of hydrogen bond interactions). HB number of hydrogen bonds, SB number of salt bridges, NHB number of non-hydrogen bonds generated in PDBsum, BSA buried surface area calculated in PISA. **B** Crystal structure of NF- $\kappa$ B p50 ORF4b NLS bound to IMP $\alpha$ 3 at 2.15 Å resolution (see Supplementary Table 11 for data collection and refinement statistics). The IMP $\alpha$ 3 is shown in dark gray and surface mode, and the p50 is shown in orange in stick mode. Pymol software was used to generate images<sup>45</sup>. A cartoon representation is shown below highlighting the hydrogen bonds (see Supplementary Table 13 for a full list of hydrogen bond interactions). **C** IMP $\alpha$ 2 bound to ORF4b (as per Fig. 1E), and overlaid with the crystal structure of IMP $\alpha$ 2 bound to p50 (as per A) showing a direct overlap at the major site. **D** IMP $\alpha$ 3 bound to ORF4b (as per Fig. 1F), and overlaid with the crystal structure of IMP $\alpha$ 3 bound to p50 (as per B) showing a direct overlap at the major site. **E** Reporter assay testing inhibition of SeV-induced activation of the IFN $\beta$  promoter by ORF4b and the indicated ORF4b mutants. Empty vector (E) served as a control and samples were infected with SeV as indicated. Error bars represent the mean values  $\pm$  standard deviation for triplicate experiments. Statistical significance was determined by a one-way ANOVA followed by Tukey's test; n.s. no significance; \*\*\*\* $p < 0.0001$ .  $n = 3$  biologically independent experiments. Source data are provided as a Source Data file. **F** Reporter assay testing inhibition of TNF $\alpha$ -induced activation of an NF- $\kappa$ B-responsive promoter by ORF4b and the indicated ORF4b mutants. Empty vector (E) served as a control and samples were treated with TNF $\alpha$  as indicated. Error bars represent the mean values  $\pm$  standard deviation for triplicate experiments. Statistical significance was determined by a one-way ANOVA followed by Tukey's test; n.s. no significance; \*\*\*\* $p < 0.0001$ .  $n = 3$  biologically independent experiments. Source data are provided as a Source Data file. MERS Middle East respiratory syndrome-related coronavirus, IMP $\alpha$  importin alpha, NF $\kappa$ B Nuclear factor kappa light chain enhancer of activated B cells, IFN $\beta$  interferon  $\beta$ , SeV Sendai virus, TNF $\alpha$  Tumor necrosis factor  $\alpha$ .

Our structural data showed that the ORF4b NLS bound to IMP $\alpha$  using a novel interaction geometry (Fig. 2). In the classical nuclear import pathway, cargo binding the major site of IMP $\alpha$  is highly dependent on a Lys at the P2 position that interacts with Gly<sup>150</sup>, Thr<sup>155</sup>, and Asp<sup>192</sup> within the IMP $\alpha$  pocket. The terminal NZ nitrogen of the highly conserved P2 Lys is able to mediate an array of H-bonding interactions with all three residues that comprise the IMP $\alpha$  P2 binding pocket, including H-bonds with the main chain carbonyl group of Gly<sup>150</sup> and the hydroxyl group of Thr<sup>155</sup>, as well as a salt bridge with the negatively charged Asp<sup>192</sup> side chain. Numerous viral and host proteins have been shown to shuttle into the nucleus using this canonical P2 Lys site (Supplementary Fig. 3). Surprisingly, MERS ORF4b does not contain a Lys at the P2 site, but rather a Val (position 35), which did not form any interactions with the Gly<sup>150</sup>, Thr<sup>155</sup>, and Asp<sup>192</sup> pocket. Rather, MERS ORF4b bound IMP $\alpha$ 2 and IMP $\alpha$ 3 through an Arg residue (position 33, the equivalent of the P0 position) at this site and formed similar interactions to those made by a Lys NZ nitrogen at the P2 site. Because of this binding mechanism was so unusual, we examined all IMP $\alpha$  structures deposited to the Protein Data Bank, where we found that only one other cargo bound to IMP $\alpha$  using a P0 Arg. The Epstein-Barr virus nuclear antigen leader protein (EBNA-LP) was shown to bind using a highly similar interaction, in which a P0 Arg binds at the canonical P2 site (Supplementary Fig. 3). Interestingly, both proteins contain a Val at the P2 site, and future studies are required to examine whether an RxV motif is important for this binding mechanism, and the extent to which other viruses or host nuclear cargo utilize this binding motif. One notable difference between MERS-CoV ORF4b and EBNA-LP is that the latter bound only at the major site, contrasting with MERS-CoV ORF4b that bound at both the major and minor sites of IMP $\alpha$ . Together these data suggest that the RxV motif can mediate mono- and bi-partite NLS interactions.

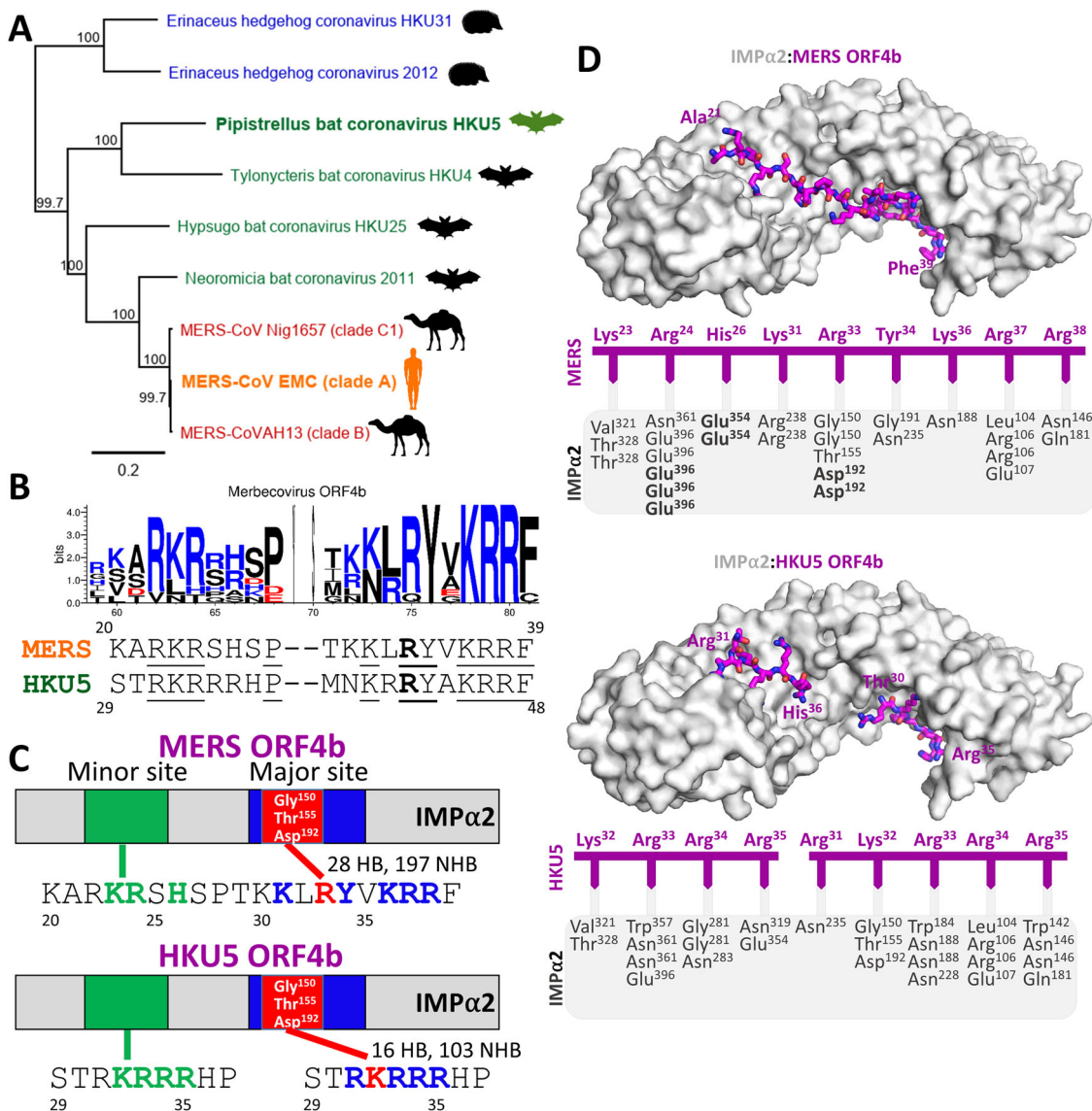
Intriguingly, mutations with the NLS of MERS-CoV ORF4b resulted in a dramatic shift towards canonical Lys P2 binding (Fig. 3). For example, mutations within both the minor site (Arg24Ala), and the major site (Arg33Ala and Arg37Ala) resulted in the MERS-CoV ORF4b binding using a canonical Lys P2 involving <sup>29</sup>TKKLR<sup>33</sup>, with Lys<sup>30</sup> binding at the P2 site. These NLSs all bound with reduced affinity compared to the Arg P0 site. Consistently, although mutations still allowed interaction with IMP $\alpha$  at the interface using the Lys P2, both the level of nuclear import and the ability to interfere with innate immunity were reduced (Figs. 5 and 6E, F and Supplementary Figs. 10 and 11).

This illustrates how small mutations within viruses can result in different binding geometries to IMP $\alpha$  with associated functional consequences, and that sequence alignment alone cannot be used with great confidence to predict IMP $\alpha$  binding interfaces.

Our data also provides a structural basis for understanding how the MERS-CoV ORF4b inhibits innate immunity by impacting the nuclear import of NF- $\kappa$ B. Both the p50 and p65 components of the NF- $\kappa$ B heterodimer contain NLSs and it has been proposed that the p50 NLS binds to the major NLS-binding site on IMP $\alpha$ s whereas the p65 NLS binds to the minor site<sup>26</sup>. Our crystal structures confirm that the binding site for the NLS region of p50 is located within the major site of IMP $\alpha$ , overlapping directly with the high-affinity P0 Arg binding site of ORF4b (Fig. 6A–D). We confirm that ORF4b reduces innate immune responses including NF- $\kappa$ B and IFN $\beta$  induction, and that mutations within the ORF4b interface reduce this activity. This finding supports the published work of Canton et al. (2018) who observed an inhibitory mechanism dependent upon direct competition between MERS-CoV-expressed ORF4b and endogenous NF- $\kappa$ B subunits for binding to IMP $\alpha$  and subsequent nuclear import. Whether ORF4b inhibition of IRF3 and IRF7-mediated IFN $\beta$  induction (as suggested within the literature<sup>7,14</sup>) occurs via an analogous mechanism remains to be explored.

Future studies will be aimed towards understanding how ORF4b proteins of related merbecoviruses bind to IMP $\alpha$ . In this regard, we have examined the binding mechanism of the HKU5 ORF4b NLS bound to IMP $\alpha$ , which indeed uses a canonical Lys P2 binding mechanism (Fig. 7C, D). Our data confirms the results from an earlier study which demonstrated that downstream “site 2” basic residues are unimportant for HKU5 ORF4b nuclear localization<sup>7</sup>. Indeed, despite the complete conservation of the downstream KRR motif among merbecovirus ORF4b proteins (Fig. 7B and Supplementary Fig. 12), these are not utilized for HKU5 ORF4b interactions with IMP $\alpha$ 2; rather the functional NLS comprises the mono-partite sequence <sup>31</sup>RKRRR<sup>35</sup> which was capable of binding at both the major and minor sites on IMP $\alpha$ . Intriguingly, this monopartite “site 1”-only NLS has been shown to still be sufficient for the inhibition of NF- $\kappa$ B-mediated innate immune signaling<sup>7</sup>.

An earlier study suggested that an altered ORF4b:IMP $\alpha$  interaction modulates the efficiency of PDE domain-mediated inhibition of the OAS/RNase L pathway<sup>10</sup>. When each were expressed recombinantly from an NS2-inactive mouse hepatitis virus (MHV) vector, HKU5 ORF4b localized to the nucleus to a greater extent than MERS-CoV ORF4b<sup>10</sup> (for which a small



**Fig. 7 Related merbecovirus HKU5 binds IMPα through a different mechanism to MERS ORF4b.** **A** Maximum likelihood tree of complete genome sequences of select viruses within subgenus *Merbecovirus*. Virus names are colored in accordance with the host from which the sequences were derived (blue = hedgehogs, green = bats, red = camels, and orange = humans). Numbers next to nodes indicate bootstrap proportion support, and the scale indicates the number of nucleotide substitutions per site. The Bat norbecovirus (HM211099.1) outgroup was used to root the tree. **B** Sequence logo demonstrating conservation of positively charged amino acids (blue) within the NLS region of ORF4b from the merbecoviruses in panel A (see Supplementary Fig. 10 for the full multiple sequence alignment). Below the sequence logo, the NLS region amino acids of MERS-CoV and HKU5 ORF4b are aligned. Conserved residues are underlined and the P0 Arg identified in MERS-CoV is in bold. **C** Cartoon representation highlighting the minor and major site interactions within MERS ORF4b and HKU5 ORF4b binding to IMPα2. **D** Crystal structure of HKU5 ORF4b bound to IMPα2 at 2.2 Å resolution (see Supplementary Table 11 for data collection and refinement statistics) and a comparison of MERS ORF4b bound to IMPα2 as per Fig. 1E. The IMPα is shown in light gray and surface mode, and the ORF4b shown in purple in stick mode. Pymol software was used to generate images<sup>45</sup>. A cartoon representation is shown below highlighting the hydrogen bonds and salt bridges (in bold) and the structure presented in the right panel. A list of hydrogen-bonded interactions is presented in Supplementary Table 14. MERS/MERS-CoV Middle East respiratory syndrome-related coronavirus, HKU5 Pipistrellus bat coronavirus HKU5, IMPα importin alpha.

proportion of protein remained cytoplasmic, as has been observed in the current study (Fig. 5)). However, the authors reported that RNase L-dependent RNA cleavage was still observed in MHV-HKU5 ORF4b infected cells, contrasting with inhibition of this pathway seen when WT and NLS-deleted ( $\Delta$ 1-52) MERS-CoV ORF4b were expressed<sup>10</sup>. This suggests that the activity of the ORF4b NLS region may have to strike a balance between nuclear import to competitively inhibit host immune transcription factors, and retention within the cytoplasm to allow PDE domain inhibition of OAS/RNase L. Further studies are required to

determine whether the non-canonical Arg P0 IMPα-binding mechanism somehow mediates this balanced innate immune antagonism phenotype.

Finally, it should be noted that both SARS-CoV-1 and 2 (genus *Betacoronavirus*, subgenus *Sarbecovirus*), although related to MERS-CoV (genus *Betacoronavirus*, subgenus *Merbecovirus*), lack the ORF4b gene. Instead, SARS-CoV-1 and 2 inhibit nuclear import of innate immune signaling proteins using ORF6. Although ORF4b and ORF6 are unrelated, both are tasked with the function of preventing innate immune signals from reaching



the nucleus. In the case of SARS-CoV-1, ORF6 was demonstrated to antagonize STAT1 functioning by tethering IMP $\alpha$ 1 to the rough endoplasmic reticulum/Golgi membrane<sup>29</sup>. This activity mapped to the C-terminal region of ORF6 and bears no similarity to ORF4b. In SARS-CoV-2, it was demonstrated ORF6 localized to the nuclear pore complex, and disrupted cargo-receptor docking<sup>30</sup>. Together this suggests that although betacoronaviruses have evolved different mediators of innate immune pathway inhibition, mechanisms that suppress IMP $\alpha$  transport of transcription factor cargo are a common virulence determinant among pathogenic viruses within this group.

## Methods

**Protein expression and purification for structural and binding studies.** A gene fragment encoding the MERS-CoV ORF4b NLS region (NC\_019843.3 residues 19–39) and an N-terminal TEV cleavage site was codon optimized for bacterial expression and synthesized by Genscript and cloned into the pGEX4T vector at the BamHI site. The NF- $\kappa$ B p50 NLS region (UniProtKB P19838; residues 331–370) was codon optimized for expression in *E. coli*, synthesized with an added N-terminal TEV cleavage site, and cloned into pGEX4T-1 vector at BamHI and EcoRI sites (Genscript, Piscataway, NJ). Constructs encoding IMP $\alpha$ s lacking the IBB domain and an N-terminal TEV cleavage site, were codon optimized and synthesized by Genscript, and cloned into pET30a using the BamHI site, with the following accession numbers: IMP $\alpha$ 1 (UniProtKB P52292; residues 70–529); IMP $\alpha$ 2 (UniProtKB P52293; residues 70–529), IMP $\alpha$ 3 (UniProtKB 000629; residues 64–521), IMP $\alpha$ 5 (UniProtKB p52294, residues 73–538), and IMP $\alpha$ 7 (UniProtKB encoding residues 73–536). Plasmids were transformed into BL21 (DE3) pLysS cells and expressed as described previously<sup>31</sup>. Purification of 6xHis tagged IMP $\alpha$  proteins were performed by injecting clarified cell lysate onto a HisTrap 5 mL column using binding buffer (50 mM phosphate buffer pH 8.0, 300 mM NaCl, 20 mM imidazole). The column was washed with 15 column volumes of binding buffer then eluted using a gradient elution over 5 column volumes to reach 100% elution buffer (50 mM phosphate pH 8.0, 300 mM NaCl, 500 mM imidazole). Samples were pooled, and the affinity tag removed by TEV proteolysis. Size exclusion purification of pooled samples was performed on a Superdex 200 pg 26/600 column using gel filtration buffer (50 mM Tris, 125 mM sodium chloride, pH 8.0). Eluted proteins were pooled and concentrated to 10–20 mg/ml using 10 kDa MW centrifuge filters.

Complexes of MERS-CoV ORF4b NLS and IMP $\alpha$  isoforms were achieved by mixing lysed *E. coli* cells expressing GST-tagged MERS-CoV ORF4b NLS with *E. coli* cells expressing IMP $\alpha$  isoforms. Purification of complexed proteins were performed by passing cleared lysate over a HisTrap 5 mL column using binding buffer (50 mM phosphate buffer pH 8.0, 300 mM NaCl, 20 mM imidazole), washing with 15 column volumes of binding buffer, and elution using a gradient elution over 5 column volumes to reach 100% elution buffer (50 mM phosphate pH 8.0, 300 mM NaCl, 500 mM imidazole). Samples were pooled, and the His and GST affinity tag removed by TEV proteolysis. Size exclusion purification of pooled samples was performed on a Superdex 200 pg 26/600 column using gel filtration buffer (50 mM Tris, 125 mM sodium chloride, pH 8.0). The protein was then passed through a 5 mL GSTrap FF column to remove GST, and the unbound protein complex was concentrated to 10–20 mg/ml using 10 kDa MW centrifuge filters, aliquoted, and flash-frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

**Crystallization, data collection, and processing.** All crystals were obtained using the hanging drop vapor diffusion method over a 300  $\mu\text{L}$  reservoir solution. IMP $\alpha$ 2: MERS-CoV NLS was crystallized in 700 mM sodium citrate (pH 7.0) and 10 mM DTT. Single rod-shaped crystals forming within 3 days. IMP $\alpha$ 3: MERS-CoV NLS crystallized in 0.2 M magnesium chloride, 0.1 M Tris pH 8.0, 20% w/v PEG 6000 with plate-shaped crystals forming within 14 days. Crystallization of IMP $\alpha$ 2 in the presence of MERS ORF4b NLS wild-type and mutant peptides was achieved by mixing the peptide in a 2:1 molar ratio, and crystallization conditions were the same as the protein complex above. IMP $\alpha$ 2 bound to NF- $\kappa$ B p50 was crystallized in 700 mM sodium citrate, 10 mM DTT, 100 mM sodium HEPES pH 7.0, with rod-shaped crystals appearing in 2–3 days. IMP $\alpha$ 3 bound to NF- $\kappa$ B p50 was crystallized in 0.2 M potassium thiocyanate and 20% w/v PEG 3350, with rod-shaped crystals appearing in 2–3 days. X-ray diffraction data were collected at the Australian Synchrotron on the MX1 and MX2 macromolecular beam lines using Eiger2 9 M and Eiger 16 M detectors, respectively<sup>32,33</sup>. Data reduction and integration was performed using iMosflm<sup>34</sup> and XDS<sup>35</sup>. Merging, space group assignment, scaling and selection of 5% reflections for Rfree calculations were performed using Aimless<sup>36</sup> within the CCP4 suite<sup>37</sup>. Phasing was performed using molecular replacement in Phaser MR, with PDB codes 7J1M and 6BVZ used as a search model for IMP $\alpha$ 2 and IMP $\alpha$ 3 respectively<sup>21,38</sup>. Models were built and refined using iterative cycles of cool<sup>39</sup> and maximum likelihood phenix refine<sup>40</sup>. The final models have been validated and deposited to the PDB with accession numbers detailed in Supplementary Tables.

**Electrophoretic mobility shift assays.** FITC labeled MERS-CoV ORF4b NLS peptides (8.5  $\mu\text{l}$  of 0.5 mg/ml) were mixed with 40  $\mu\text{M}$  of each IMP $\alpha$  isoform and incubated in a total volume 17  $\mu\text{l}$  for 15 min at room temperature. Samples were supplemented with 3  $\mu\text{l}$  of 50% glycerol and run on a 1% agarose TB gel for 1.5 h at 70 V in TB running buffer. The images were recorded using a SYBR green filter within a Gel Doc BioRad Gel Doc imaging system. The gel was then stained using Coomassie stain and destained in 10% ethanol and 10% glacial acetic acid.

**Fluorescence polarization assays.** FITC-tagged MERS ORF4b peptides (5 nM) were incubated with two-fold serially diluted IMP $\alpha$  concentrations (starting concentration 5  $\mu\text{M}$ ) across 23 wells to a total volume of 200  $\mu\text{L}$ . Fluorescence polarization measurements were recorded using a CLARIOstar Plus plate reader (BMG Labtech) in 96-well black Fluotrac microplates (Greiner Bio-One; Kremsmünster, Austria). Assays were repeated in triplicate and contained a negative control (no binding partner) and blank (Tris buffered saline, pH 8). Triplicate data was fitted to a single binding curve using GraphPad Prism.

**Cell culture and plasmids.** HEK293T (ATCC—CRL-3216) and Huh7 cells (a generous gift from the Gordan lab at University of California at San Francisco) were maintained in Dulbecco's Modified Eagle Medium, supplemented with 10% fetal bovine serum and cultured at 37  $^{\circ}\text{C}$  and 5%  $\text{CO}_2$ .

The codon-optimized sequence for MERS ORF4b (NC\_019843.3), was synthesized (Genscript, Piscataway, NJ) and cloned with an N-terminal Flag- or HA-tag into the mammalian expression plasmid pCAGGS. Flag-tagged MERS ORF4b mutants R24A, R33A, R24A/R33A and RKRA (R22A/K23A/R24A) were generated by overlapping PCR. MERS ORF4b delta38 was cloned using PCR to remove the first 38 amino acids. pCAGGS Flag- and HA-tagged IMP $\alpha$ 1, 3, 4, 5, 6, and 7 have previously been described<sup>16</sup>.

**Immunofluorescence assays.** MERS ORF4b nuclear translocation: Huh7 cells ( $3 \times 10^4$ ) grown on glass coverslips were transfected with indicated wt and mutant Flag-tagged MERS ORF4b plasmids (500 ng) using Lipofectamine 2000 (Thermo Fisher Scientific, MA). At 24 h post transfection, cells were fixed using 4% paraformaldehyde and permeabilized using 0.1% Triton X-100. Cells were stained using anti-Flag (DYKDDDDK) tag monoclonal antibody conjugated to Alexa Fluor 488 (Invitrogen, cat# MA1-142-A488, lot# UB276601) (dilution 1:1000). To determine the ratio of nuclear to cytoplasmic fluorescence signal (Fn/c), coverslips were imaged using a BioTek Cytation 5 Cell Imaging Multi-Mode reader. Images were analyzed using Gen5 Image Prime software to determine Fn/c, using the calculation  $\text{Fn/c} = (\text{Fn-background})/(\text{Fc-background})$ , where Fn is nuclear fluorescence and Fc is cytoplasmic fluorescence. Fn/c was determined for  $\geq 60$  cells in each condition; error bars indicate the mean values  $\pm$  standard deviation. Images of the same coverslips were taken using a Zeiss LSM 800 confocal microscope at  $\times 64$ .

p65 nuclear translocation: Huh7 cells were transfected as above with the indicated Flag-tagged MERS ORF4b plasmids or empty vector. At 24 h post-transfection, media was changed to DMEM without FBS for four hours, after which coverslips were treated with 20 ng/ml of TNF $\alpha$  (Peprotech) for 30 min. Coverslips were fixed and permeabilized as above and cells were stained with anti-p65 (Santa Cruz, cat# sc-109, lot# F2304) (dilution 1:100) and goat anti-rabbit IgG Alexa Fluor 647 (Invitrogen, cat# A-21244, lot# 1871168) (dilution 1:1000), and anti-Flag Alexa Fluor 488 (Invitrogen, cat# MA1-142-A488, lot# UB276601) (dilution 1:1000). Fn/c for p65 was determined as above using a BioTek Cytation 5 Cell Imaging Multi-Mode reader and Gen5 Image Prime software, selecting for cells expressing MERS ORF4b or mutants in transfected cells. Fn/c was determined across five fields of view for  $\geq 24$  cells expressing MERS ORF4b or mutants, and across three fields of view for  $\geq 139$  cells for  $-\text{TNF}\alpha$  and  $+\text{TNF}\alpha$  controls.

**Co-immunoprecipitation assay.** HEK293T cells ( $1 \times 10^6$ ) were transfected with the indicated plasmids using Lipofectamine 2000 (Thermo Fisher Scientific). At 24 h post transfection, cells were lysed in NP-40 lysis buffer (50 mM Tris pH 7.5, 280 mM NaCl, 0.5% NP-40, 0.2 mM EDTA, 2 mM EGTA, 10% glycerol, protease inhibitor (complete; Roche, Indianapolis, IN)). Anti-FLAG M2 magnetic beads or EZview Red anti-HA agarose affinity gel (Sigma-Aldrich) were incubated as indicated with lysates for 1 h at 4  $^{\circ}\text{C}$ , washed five times in NP-40 lysis buffer, and eluted using 3X FLAG or HA peptide (Sigma-Aldrich) at 4  $^{\circ}\text{C}$  for 30 min or by boiling in sample buffer for five minutes. Whole cell lysates (4% of sample) and co-precipitation samples (25% of sample) were analyzed by western blot.

**Western blots.** Lysates were run on 10% Bis-Tris Plus polyacrylamide gels (Thermo Fisher) and transferred to PVDF membrane (Bio-Rad). Membranes were probed with the indicated antibodies and were developed by Western Lightning Plus ECL (Perkin Elmer) and imaged on a ChemiDoc MP Imaging System (Bio-Rad). Antibodies, rabbit anti-Flag (cat# F7425), and mouse anti- $\beta$ -tubulin (cat# T8328) were purchased from Sigma-Aldrich. Rabbit anti-HA (cat# 71-5500) and mouse anti-GAPDH (GA1R) (cat# MA5-15738) were purchased from Invitrogen. Anti-rabbit and anti-mouse IgG, HRP-linked antibodies were purchased from Cell Signaling (cat# 7074 and 7076).

**Reporter assays in HEK293T cells.** For the Sendai virus (SeV)-induced IFN $\beta$  promoter luciferase assay, HEK293T cells ( $7.5 \times 10^4$  cells/well) were transfected with 30 ng of the IFN $\beta$  firefly luciferase reporter, 30 ng of a constitutively expressed Renilla luciferase reporter (pRLTK), and Flag-tagged MERS ORF4b and mutants (25 ng and 2.5 ng), using Lipofectamine 2000 (Thermo Fisher Scientific). Twenty-four hours post transfection, cells were infected with 200 HA units (HAU) of SeV for 18 h. Luciferase activity was assessed using a dual-luciferase assay (Promega) and read on an EnVision multilabel plate reader (PerkinElmer, Inc.). Firefly luciferase values were normalized to Renilla luciferase values. The assay was performed in triplicate; error bars indicated the mean values  $\pm$  standard deviation for the triplicate. For the NF- $\kappa$ B promoter luciferase assay, HEK293T cells ( $7.5 \times 10^4$  cells/well) were transfected with 30 ng of the NF- $\kappa$ B firefly luciferase reporter, 30 ng of a constitutively expressed Renilla luciferase reporter (pRLTK), and Flag-tagged MERS ORF4b and mutants (25 ng and 2.5 ng), using Lipofectamine 2000 (Thermo Fisher Scientific). Twenty-four hours post transfection, cells were treated with 10 ng/ml of TNF $\alpha$  (PeproTech, cat# 300-01A) for six hours. Luciferase activity was assessed and analyzed as described above.

**Reporter Assays in A549 cells.** For the IFN $\beta$  reporter assay, A549 cells ( $4 \times 10^4$  cells/well) were transfected with 30 ng of the IFN $\beta$  firefly luciferase reporter and 30 ng of a constitutively expressed Renilla luciferase reporter plasmid (pRLTK). Additionally, 25 ng or 2.5 ng of Flag-tagged MERS WT ORF4b (ORF4b), ORF4b R24A (R24A), ORF4b R33A (R33A), or ORF4b R24A/R33A (R24A/R33A) were transfected using TransIT-LT1 (Mirus). Twenty-four hours post-transfection, cells were infected with 200 hemagglutinating units (HAU) of SeV for 18 h. Luciferase activity was assessed using a dual-luciferase assay (Promega) and read on an EnVision multilabel plate reader (PerkinElmer, Inc.). Firefly luciferase values were normalized to Renilla luciferase values. The assay was performed in triplicate; error bars indicate the mean values  $\pm$  standard deviation for the triplicate. For the NF $\kappa$ B reporter assay, the transfection was performed as above, using 30 ng of the NF $\kappa$ B firefly luciferase reporter. Twenty-four hours post-transfection, cells were treated with 50 ng/ml of TNF $\alpha$  (PeproTech, cat# 300-01 A) for 18 h. Luciferase activity was assessed and analyzed as described above.

**Phylogenetic analyses.** Phylogenetic analysis of betacoronavirus sequences was performed using Geneious Prime Java Version 11.0.4 + 11 (64 bit). Complete genome nucleotide sequences of 10 viruses (GenBank accession numbers HM211099.1, NC\_039207.1, KC869678.4, MK907287.1, KX442565.1, NC\_009019.1, NC\_009020.1, NC\_019843.3, MG923475.1, and KJ650295.1) were aligned using MAFFT v7.450<sup>41</sup>, with the resulting alignment subsequently used to build a maximum likelihood PhyML (v3.3.20180621) tree using the HKY85 substitution model and bootstrapping of 1,000 replicates<sup>42</sup>. The resulting tree was rooted on the Bat norbecovirus (HM211099.1) outgroup. Full-length amino acid sequences of the ORF4b protein from 9 merbecoviruses corresponding to the full-length genomes (GenBank accession numbers AHX00714.1, AIG13099.1, ASL68956.1, AVN89394.1, QGA70705.1, YP\_001039956.1, YP\_001039965.1, YP\_009047207.1, and YP\_009513014.1) were also aligned using MAFFT v7.450. The aligned amino acids 59–81 were subsequently used to generate a sequence logo using WebLogo 3<sup>43</sup>, and colored based upon amino acid charge.

**Statistical analyses.** Statistical analysis was performed using GraphPad Prism 9<sup>44</sup>. Data points were considered significantly different if the *P* value was  $<0.05$ . Further details can be found in the figure legends.

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

## Data availability

Protein Data Bank files associated with the structures generated in this study have been deposited to the Protein Data Bank and were released prior to review of the manuscript. The codes are IMPa2:MERS ORF4b 7RFX; IMPa3:MERS ORF4b 7RFY; IMPa2:MERS ORF4b wt peptide 7RFZ; IMPa2:MERS ORF4b R24A peptide 7RG0; IMPa2:MERS ORF4b H26A peptide 7RG1; IMPa2:MERS ORF4b R33A peptide 7RG2; IMPa2:MERS ORF4b R37A peptide 7RG3; IMPa2:NF- $\kappa$ B p50 7RG4; IMPa3:NF- $\kappa$ B p50 7RG5; and IMPa2:HKU5 ORF4b 7RG6. Source data are provided with this paper.

Received: 16 July 2021; Accepted: 11 February 2022;

Published online: 25 March 2022

## References

- Zhang, Y. Y., Li, B. R. & Ning, B. T. The comparative immunological characteristics of SARS-CoV, MERS-CoV, and SARS-CoV-2 coronavirus infections. *Front. Immunol.* **11**, 2033 (2020).
- Lefkowitz, E. J. et al. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* **46**, D708–D717 (2018).
- WHO. Middle East respiratory syndrome coronavirus (MERS-CoV), [https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-\(mers-cov\)](https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)) (2021).
- Majumder, M. S., Brownstein, J. S., Finkelstein, S. N., Larson, R. C. & Bourouiba, L. Nosocomial amplification of MERS-coronavirus in South Korea, 2015. *Trans. R. Soc. Trop. Med. Hyg.* **111**, 261–269 (2017).
- Willman, M., Kobasa, D. & Kindrachuk, J. A Comparative analysis of factors influencing two outbreaks of middle eastern respiratory syndrome (MERS) in Saudi Arabia and South Korea. *Viruses* <https://doi.org/10.3390/v11121119> (2019).
- Mobaraki, K. & Ahmadzadeh, J. Current epidemiological status of Middle East respiratory syndrome coronavirus in the world from 1.1.2017 to 17.1.2018: a cross-sectional study. *BMC Infect. Dis.* **19**, 351 (2019).
- Matthews, K. L., Coleman, C. M., van der Meer, Y., Snijder, E. J. & Frieman, M. B. The ORF4b-encoded accessory proteins of Middle East respiratory syndrome coronavirus and two related bat coronaviruses localize to the nucleus and inhibit innate immune signalling. *J. Gen. Virol.* **95**, 874–882 (2014).
- Gusho, E. et al. Murine AKAP7 has a 2',5'-phosphodiesterase domain that can complement an inactive murine coronavirus ns2 gene. *mBio* **5**, e01312–e01314 (2014).
- Forni, D., Cagliani, R., Clerici, M. & Sironi, M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* **25**, 35–48 (2017).
- Thornbrough, J. M. et al. Middle east respiratory syndrome coronavirus NS4b protein inhibits host RNase L activation. *mBio* **7**, e00258 (2016).
- Canton, J. et al. MERS-CoV 4b protein interferes with the NF-kappaB-dependent innate immune response during infection. *PLoS Pathog.* **14**, e1006838 (2018).
- Comar, C. E. et al. Antagonism of dsRNA-Induced Innate Immune Pathways by NS4a and NS4b Accessory Proteins during MERS Coronavirus Infection. *mBio* <https://doi.org/10.1128/mBio.00319-19> (2019).
- Lee, J. Y., Bae, S. & Myoung, J. Middle east respiratory syndrome coronavirus-encoded accessory proteins impair MDA5- and TBK1-mediated activation of NF-kappaB. *J. Microbiol. Biotechnol.* **29**, 1316–1323 (2019).
- Yang, Y. et al. Middle East respiratory syndrome coronavirus ORF4b protein inhibits type I interferon production through both cytoplasmic and nuclear targets. *Nucl. Rep.* **5**, 17554 (2015).
- Yang, Y. et al. The structural and accessory proteins M, ORF 4a, ORF 4b, and ORF 5 of Middle East respiratory syndrome coronavirus (MERS-CoV) are potent interferon antagonists. *Protein Cell* **4**, 951–961 (2013).
- Reid, S. P., Valmas, C., Martinez, O., Sanchez, F. M. & Basler, C. F. Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. *J. Virol.* **81**, 13469–13477 (2007).
- Taylor, S. L., Frias-Staheli, N., Garcia-Sastre, A. & Schmaljohann, C. S. Hantaan virus nucleocapsid protein binds to importin alpha proteins and inhibits tumor necrosis factor alpha-induced activation of nuclear factor kappa B. *J. Virol.* **83**, 1271–1279 (2009).
- Ye, J. et al. Japanese encephalitis virus NS5 inhibits type I interferon (IFN) production by blocking the nuclear translocation of IFN regulatory factor 3 and NF-kappaB. *J. Virol.* <https://doi.org/10.1128/JVI.00039-17> (2017).
- Xu, W. et al. Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1. *Cell Host Microbe* **16**, 187–200 (2014).
- Pumroy, R. A. & Cingolani, G. Diversification of importin-alpha isoforms in cellular trafficking and disease states. *Biochem. J.* **466**, 13–28 (2015).
- Smith, K. M. et al. Structural basis for importin alpha 3 specificity of W proteins in Hendra and Nipah viruses. *Nat. Commun.* **9**, 3703 (2018).
- Smith, K. M., Himiari, Z., Tsimbaluk, S. & Forwood, J. K. Structural basis for importin-alpha binding of the human immunodeficiency virus tat. *Sci. Rep.* **7**, 1650 (2017).
- Ng, I. H. W. et al. Zika virus NS5 forms supramolecular nuclear bodies that sequester importin-alpha and modulate the host immune and pro-inflammatory response in neuronal cells. *ACS Infect. Dis.* **5**, 932–948 (2019).
- Fontes, M. R., Teh, T. & Kobe, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.* **297**, 1183–1194 (2000).
- Pumroy, R. A., Ke, S., Hart, D. J., Zachariae, U. & Cingolani, G. Molecular determinants for nuclear import of influenza A PB2 by importin alpha isoforms 3 and 7. *Structure* **23**, 374–384 (2015).
- Fagerlund, R., Kinnunen, L., Kohler, M., Julkunen, I. & Melen, K. NF- $\kappa$ B is transported into the nucleus by importin  $\alpha$ 3 and importin  $\alpha$ 4. *J. Biol. Chem.* **280**, 15942–15951 (2005).
- Jagga, B. et al. Structural basis for nuclear import selectivity of pioneer transcription factor SOX2. *Nat. Commun.* **12**, 28 (2021).
- Sankhala, R. S. et al. Three-dimensional context rather than NLS amino acid sequence determines importin alpha subtype specificity for RCC1. *Nat. Commun.* **8**, 979 (2017).

29. Frieman, M. et al. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. *J. Virol.* **81**, 9812–9824 (2007).
30. Miorin, L. et al. SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. *Proc. Natl Acad. Sci. USA* **117**, 28344–28354 (2020).
31. Roman, N., Christie, M., Swarbrick, C. M., Kobe, B. & Forwood, J. K. Structural characterisation of the nuclear import receptor importin alpha in complex with the bipartite NLS of Prp20. *PLoS ONE* **8**, e82038 (2013).
32. Aragao, D. et al. MX2: a high-flux undulator microfocussing beamline serving both the chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **25**, 885–891 (2018).
33. Cowieson, N. P. et al. MX1: a bending-magnet crystallography beamline serving both chemical and macromolecular crystallography communities at the Australian Synchrotron. *J. Synchrotron Radiat.* **22**, 187–190 (2015).
34. Batty, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
35. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
36. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).
37. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
38. Tu, W. J. et al. Targeting nuclear LSD1 to reprogram cancer cells and reinvigorate exhausted T cells via a novel LSD1-EOMES switch. *Front. Immunol.* **11**, 1228 (2020).
39. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
40. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
41. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
42. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
43. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
44. Prism. GraphPad Software. <https://www.graphpad.com/> (2021).
45. The PyMOL Molecular Graphic System v. 2.5.0 (2010).

## Acknowledgements

J.K.F. acknowledges support from the NHMRC grant APP2003636. C.F.B. acknowledges support from NIH grants R21AI144880, R21AI164080, and P01AI120943. This research

was undertaken in part using the MX2 beamline at the Australian Synchrotron, part of ANSTO, and made use of the Australian Cancer Research Foundation (ACRF) detector. M.S. acknowledges support from MRC grant MC-A025-5PL41.

## Author contributions

T.S.M. acquired and analyzed the data, M.R.E. acquired and analyzed the data, S.T. acquired and analyzed the data, O.A.V. acquired and analyzed the data, K.M.S. acquired and analyzed the data, M.S. assisted with manuscript preparation, Justin K Foster acquired and analyzed the data, L.A.B. analyzed the data, D.A. acquired and analyzed the data, J.A.R. analyzed the data and revised the manuscript, C.F.B. analyzed the data and revised the manuscript, Jade K Forwood analyzed the data and revised the manuscript.

## 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/s41467-022-28851-2>.

**Correspondence** and requests for materials should be addressed to Justin A. Roby, Christopher F. Basler or Jade K. Forwood.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

**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) 2022