


# Structural basis for the interaction of SARS-CoV-2 virulence factor nsp1 with DNA polymerase $\alpha$ -primase

Mairi L. Kilkenny | Charlotte E. Veale | Amir Guppy | Steven W. Hardwick |  
Dimitri Y. Chirgadze | Neil J. Rzechorzek | Joseph D. Maman | Luca Pellegrini 

Department of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, UK

## Correspondence

Luca Pellegrini, Department of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, UK.  
Email: lp212@cam.ac.uk

## Present address

Neil J. Rzechorzek, The Francis Crick Institute, London, NW1 1AT, UK

## Funding information

Wellcome Trust, Grant/Award Number: 104641/Z/14/Z

## Abstract

The molecular mechanisms that drive the infection by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)—the causative agent of coronavirus disease 2019 (COVID-19)—are under intense current scrutiny to understand how the virus operates and to uncover ways in which the disease can be prevented or alleviated. Recent proteomic screens of the interactions between viral and host proteins have identified the human proteins targeted by SARS-CoV-2. The DNA polymerase  $\alpha$  (Pol  $\alpha$ )-primase complex or primosome—responsible for initiating DNA synthesis during genomic duplication—was identified as a target of non-structural protein 1 (nsp1), a major virulence factor in the SARS-CoV-2 infection. Here, we validate the published reports of the interaction of nsp1 with the primosome by demonstrating direct binding with purified recombinant components and providing a biochemical characterization of their interaction. Furthermore, we provide a structural basis for the interaction by elucidating the cryo-electron microscopy structure of nsp1 bound to the primosome. Our findings provide biochemical evidence for the reported targeting of Pol  $\alpha$  by the virulence factor nsp1 and suggest that SARS-CoV-2 interferes with Pol  $\alpha$ 's putative role in the immune response during the viral infection.

## KEYWORDS

cryo-EM, DNA polymerase  $\alpha$ -primase, nsp1, protein-protein interaction, SARS-CoV-2

## 1 | INTRODUCTION

The enormous disruption to public health and the global economy caused by COVID-19 has spurred international research efforts to unravel the molecular biology of SARS-CoV-2 in order to devise effective means of therapeutic and prophylactic intervention.

Coronaviruses comprise a large family of positive-sense single-strand RNA viruses that cause respiratory and enteric

diseases in animals and humans.<sup>1–3</sup> Seven coronaviruses that are capable of infecting humans have been identified to date, belonging to the genera *alphacoronavirus* (HCoV-229E and HCoV-NL63) and *betacoronavirus* (HCoV-OC43, HCoV-HKU1, MERS, SARS-CoV, and SARS-CoV-2).<sup>4,5</sup> Within the *betacoronavirus* genus, HCoV-OC43 and HCoV-HKU1 generally cause only a mild illness; in contrast, infection with MERS-CoV, SARS-CoV, and SARS-CoV-2 results in a severe respiratory disease in some patients, with significantly increased fatality rates.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

SARS-CoV-2 is most closely related to SARS-like coronaviruses isolated from horseshoe bats found in Southeast Asia; it shares ~79% of its sequence with SARS-CoV and ~50% with MERS-CoV. Since its initial identification and characterization,<sup>6,7</sup> genomic sequencing has shown that the SARS-CoV-2 genome encodes—from its 5' to the 3' end—two large polyproteins 1a and 1ab that are cleaved by viral proteases to generate 16 nonstructural proteins (nsp1–nsp16), four structural proteins (S, E, M, and N), and a number of other accessory proteins.

Upon viral infection, the host cell mounts a rapid response of its innate immune system via the interferon (IFN) pathway, resulting in the secretion of type-I and type-III IFNs and subsequent activation of a signaling pathway that leads to the expression of a wide range of IFN-stimulated genes.<sup>8</sup> A hallmark of COVID-19 is the attenuation of the IFN response of the host coupled to high levels of inflammatory cytokine production.<sup>9</sup> SARS-CoV-2 deploys multiple mechanisms to suppress the innate immune response, with participation of several viral proteins including nsp1, nsp6, nsp13, M, ORF3a, ORF6, and ORF7a/b.<sup>10,11</sup>

The nsp1 protein is present in  $\alpha$ - and  $\beta$ -coronaviruses but not in  $\gamma$ - or  $\delta$ -coronaviruses. nsp1 is a major virulence factor for SARS and MERS, due its pleiotropic activities as potent antagonist of virus- and IFN-dependent signaling and suppressor of host mRNA translation.<sup>12–15</sup> While there is poor sequence conservation between the nsp1 proteins from  $\alpha$ - and  $\beta$ -coronaviruses, both have been shown to induce strong suppression of host gene expression, albeit possibly by different mechanisms.<sup>16–19</sup> Despite poor sequence identity,  $\alpha$ -CoV nsp1 and  $\beta$ -CoV nsp1 share structural homology in their core domain, which folds in a 6-stranded  $\beta$ -barrel structure with one  $\alpha$ -helix on the rim of the barrel.<sup>20–24</sup> The importance and role of nsp1's globular domain are not yet fully understood, although it has recently been shown to interact with the 5' UTR of the SARS-CoV-2 mRNA, helping it escape translational inhibition in infected cells.<sup>25</sup> In addition to their globular domain, SARS-CoV and SARS-CoV-2 nsp1 proteins contain a long, unstructured C-terminal extension. A two-helix hairpin motif near the end of the C-terminal tail of SARS-CoV-2 nsp1 has been shown to bind the 40S subunit of the ribosome and block the mRNA entry channel, providing a structural basis for the role of nsp1 in suppressing host translation in infected cells.<sup>16,26–28</sup>

Recently published studies of the human proteins targeted by SARS-CoV-2 reported a physical association between nsp1 and the primosome,<sup>29,30</sup> the complex of DNA polymerase  $\alpha$  (Pol  $\alpha$ ) and primase responsible for initiating DNA synthesis in DNA replication.<sup>31</sup> In addition to its well-characterized nuclear role in genomic duplication, participation of Pol  $\alpha$  in innate immunity

has been suggested, based on the observation that disease-associated mis-splicing of *POLA1*—coding for the catalytic subunit of Pol  $\alpha$ —results in abnormal IFN I response and autoinflammatory manifestations.<sup>32–34</sup> Here, we demonstrate that nsp1 binds directly to the catalytic subunit of Pol  $\alpha$  and elucidate their mode of interaction by determining the cryo-electron microscopy (cryo-EM) structure of SARS-CoV nsp1 bound to the human primosome. Our biochemical and structural characterization of the interaction between the virulence factor nsp1 and the primosome suggests that targeting Pol  $\alpha$ -primase is part of a novel mechanism driving the SARS-CoV-2 infection. Our data further define protein epitopes on nsp1 and Pol  $\alpha$  catalytic subunit that could be exploited for the development of small-molecule inhibitors of their interaction.

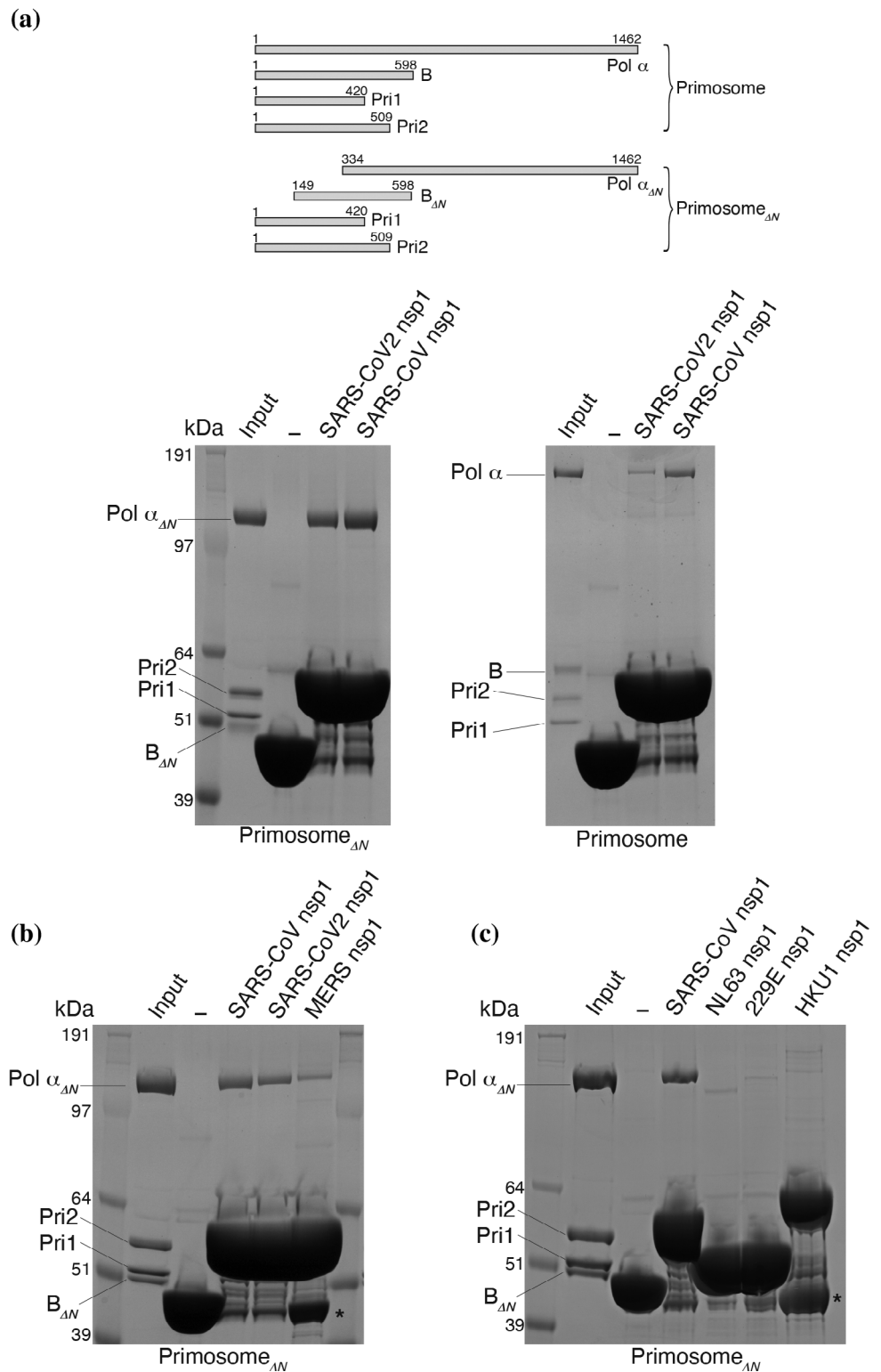
## 2 | RESULTS

### 2.1 | Biochemical analysis of nsp1 binding to the primosome

A proteomic screen of human proteins that interact with SARS-CoV-2 proteins first reported the Pol  $\alpha$ -primase complex as high-confidence interactor of SARS-CoV-2 nsp1.<sup>29</sup> A subsequent study expanded this approach to include the SARS-CoV and MERS proteins and identified the primosome as interactor of the related SARS-CoV nsp1.<sup>30</sup> Recently, a further interactome study between viral and host proteins identified the primosome as a high-confidence target of the SARS-CoV-2 nsp1.<sup>35</sup> As viral–host protein–protein interactions can play essential roles in driving the infection, we sought to confirm the reported interaction using purified, recombinant samples of nsp1 and the primosome.

Maltose-binding protein (MBP)-tagged SARS-CoV-2 and SARS-CoV nsp1 were able to pull down both full-length and truncated forms of purified primosome (Figure 1a). The truncated primosome lacks the flexible and largely unstructured N-terminal regions of both Pol  $\alpha$  (amino acids 1–333) and B subunit (amino acids 1–148) that are known to mediate several primosome interactions,<sup>36–39</sup> thus indicating that the interaction with nsp1 is not dependent on known recognition motifs in Pol  $\alpha$  and the B subunit. MERS nsp1 was also able to bind to the primosome, albeit more weakly than the SARS nsp1 proteins (Figure 1b), in accordance with the proteome-wide interaction studies that identified the primosome as a binding SARS-CoV and SARS-CoV-2 nsp1, but not MERS nsp1.<sup>30</sup> Importantly, nsp1 from human-infective coronavirus strains 229E, NL63, and HKU1 did not interact with the primosome in our pull-

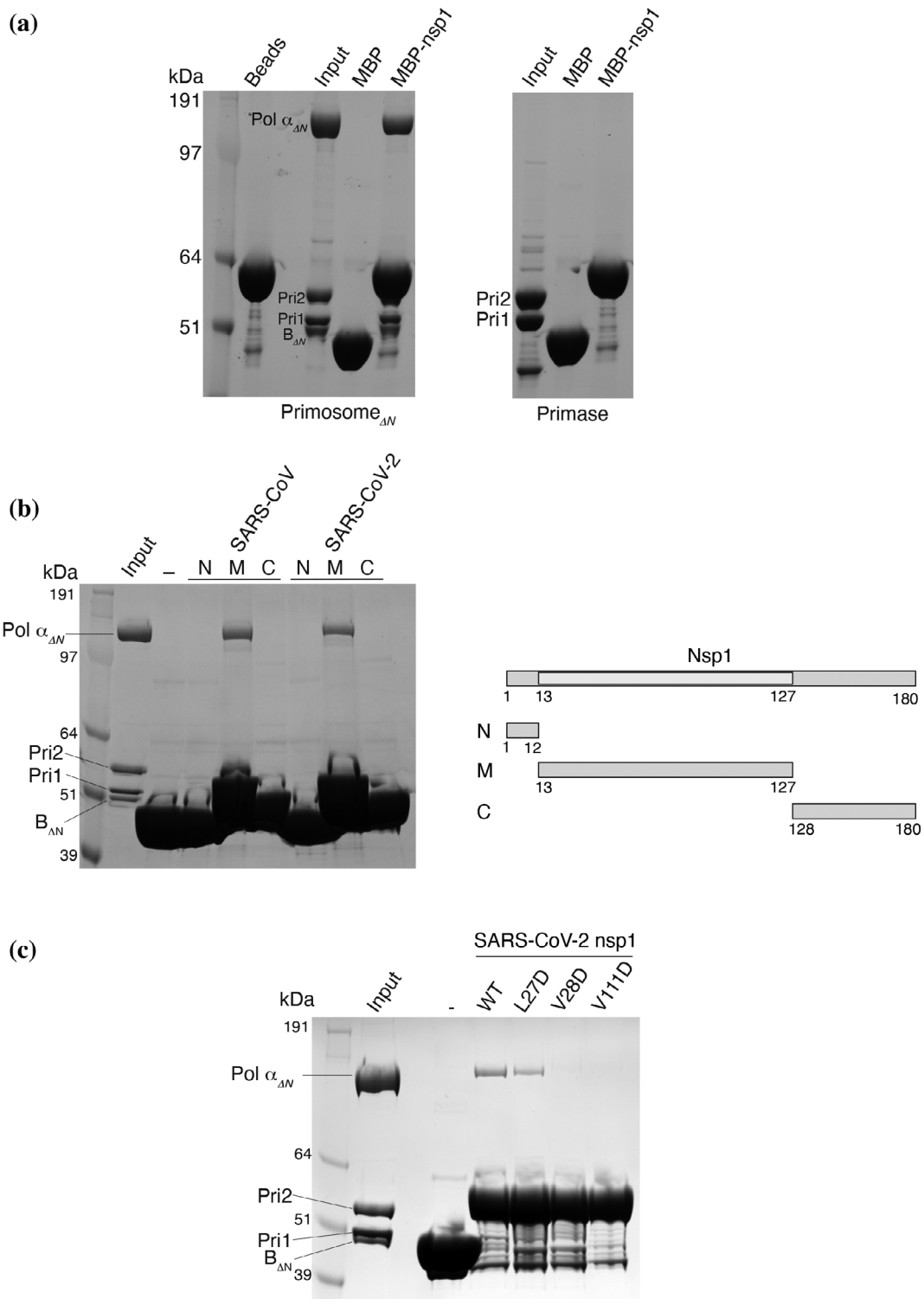
**FIGURE 1** Biochemical reconstitution of the primosome–nsp1 interaction. The panels report the result of pull-down experiments on amylose beads of human primosome by the indicated nsp1 proteins, N-terminally fused to maltose-binding protein (MBP), analyzed by SDS-PAGE. The lane marked with a dash symbol is the negative control for pull-down by MBP only. (a) Pull-down of full-length and truncated primosomes. A schematic drawing of subunit size for full-length and truncated primosomes is shown above the panels; (b) pull-down of truncated primosome with SARS-CoV, SARS-CoV-2, and MERS nsp1; (c) pull-down of truncated primosome with SARS-CoV, NLS63, 229E, and HKU1 nsp1. Bands labeled with an asterisk indicate free His<sub>6</sub>-MBP. nsp1, Nonstructural protein 1; Pol  $\alpha$ , DNA polymerase  $\alpha$ ; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis



down assay (Figure 1c). Thus, the affinity of nsp1 toward Pol  $\alpha$  correlates with the ability of the coronavirus strain to cause serious disease.

In the original interaction study,<sup>29</sup> nsp1 was coprecipitated with all four subunits of Pol  $\alpha$ –primase, which are normally found associated as a constitutive

complex in the cell. To dissect the interaction between primosome components and SARS-CoV-2 nsp1, we performed a pull-down experiment using MBP-tagged nsp1 and the primase heterodimer Pri1–Pri2 (Figure 2a). We detected no interaction between nsp1 and primase which, taken together with the data presented in Figure 1a,



**FIGURE 2** Biochemical dissection of the primosome-nsp1 interaction. The panels report the result of pull-down experiments on amylose beads of human primosome by the indicated nsp1 proteins, N-terminally fused to maltose-binding protein (MBP), analyzed by SDS-PAGE. The lane marked with a dash symbol is the negative control for pull-down by MBP only. (a) Pull-down of truncated primosome and primase; (b) pull-down of truncated primosome by N-, M-, and C-terminal domains of SARS-CoV and SARS-CoV-2 nsp1 proteins. On the right, schematic diagram of the domain structure of nsp1; (c) pull-down of truncated primosome by SARS-CoV-2 wild-type and single-point mutant L27D, V28D, and V111D nsp1 proteins. nsp1, Nonstructural protein 1; Pol α, DNA polymerase α; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis

indicates that the nsp1-binding site resides within the folded domains of Pol  $\alpha$  and the B subunit.

Next, we sought to define the region of nsp1 responsible for the interaction with the primosome. SARS-CoV and SARS-CoV-2 nsp1 fold into a globular middle domain (M; amino acids 13–127) flanked by a short N-terminal region (N; amino acids 1–12) and a long C-terminal extension that is predicted to be largely unstructured (C; amino acids 128–180). Pull-down experiments with the N, M, and C regions of SARS nsp1 showed clearly that the middle domain is responsible for the interaction with the primosome (Figure 2b).

Guided by the published crystal structures of nsp1,<sup>20,23,24</sup> we identified an exposed hydrophobic surface centered around amino acids L27 and V28 in the first interstrand loop of the  $\beta$ -barrel in the middle domain and V111 in the last interstrand loop, as a possible primosome-binding epitope. Point mutations to aspartate—that reversed the chemical nature of the side chain—weakens (L27D) or abolishes (V28D and V111D) the interaction of the MBP-tagged nsp1 mutants with the primosome in our pull-down assay (Figure 2c), thus implicating these residues in the interaction.

## 2.2 | Cryo-EM structure of the primosome–nsp1 complex

Our findings had shown that a direct biochemical interaction exists between nsp1 and the primosome, confirming the result of the proteomic screens.<sup>29,30</sup> We therefore decided to elucidate the structural basis for the interaction by determining the high-resolution structure of the primosome–nsp1 complex by cryo-EM (Figure S1–S3, Table S1). SARS-CoV nsp1 was chosen for the structural analysis as this variant showed somewhat stronger binding to human primosome in our pull-down assays (Figure 1a). We note that there is a high degree of sequence conservation between SARS-CoV and SARS-CoV-2 nsp1 (84.4 and 85.8% identity over the entire sequence and the middle domain only, respectively).

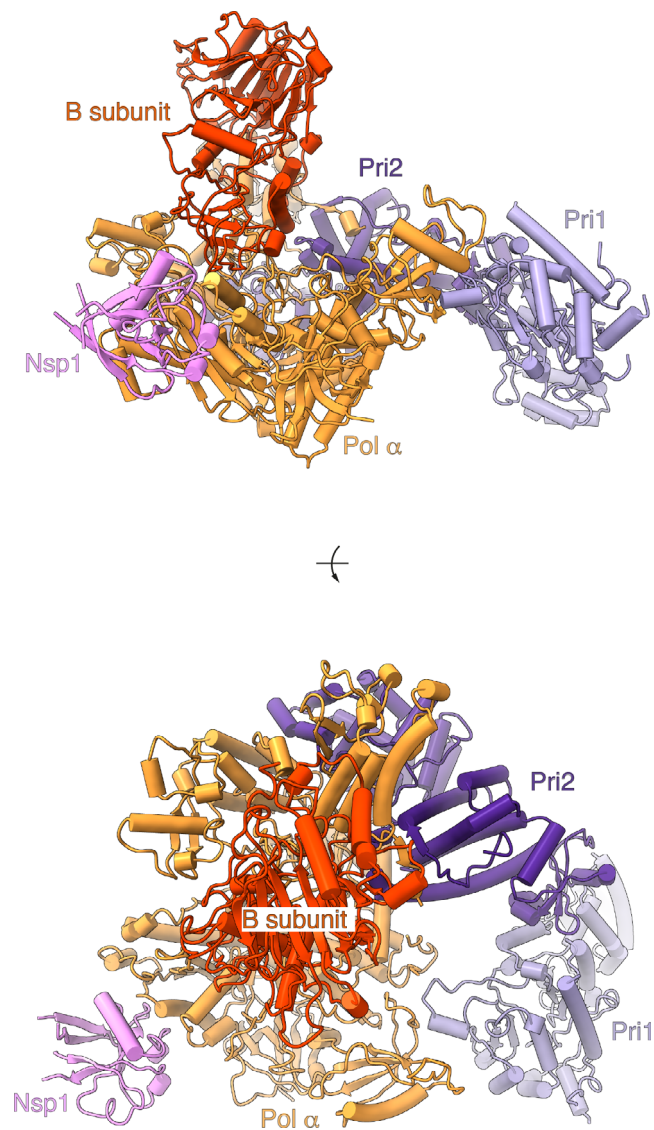
We obtained a 3D reconstruction of the primosome at a global resolution of 3.8 Å after sharpening that clearly showed all components of the complex (Figures S4 and S5). As the portion of the map relative to the primase subunit Pri1 was weaker, a new map was calculated using subtracted particles from which Pri1 had been omitted, yielding a slightly improved map at a global resolution of 3.6 Å after sharpening. Both maps—from full and subtracted particles—were used during model building. Density belonging to the nsp1 middle domain was clearly visible at a lower map threshold than required to visualize the primosome core.

The cryo-EM structure of the human primosome confirmed the presence of the binary interfaces that had been previously defined by X-ray crystallography of yeast and human Pol  $\alpha$ –primase subcomplexes.<sup>40–44</sup> Intriguingly, our cryo-EM analysis of the primosome revealed a quaternary structure that is incompatible with known and presumed conformations required for activity by its two catalytic subunits, Pol  $\alpha$  and primase<sup>42,45–47</sup> (Figure 3). Thus, the B subunit bound to the C-terminal domain of Pol  $\alpha$  is docked within the open “hand” of the polymerase domain, precluding binding of the template DNA–primer RNA duplex. Furthermore, the heterodimeric primase is held in an extended conformation, whereby the Fe-S domain of Pri2—required for initiation of RNA primer synthesis<sup>48</sup>—is sequestered by interactions with both B and Pol  $\alpha$  and unavailable to contact Pri1, which is excluded from the primosome core and largely solvent exposed.

Overall, the cryo-EM structure is highly similar to a structure of the human primosome determined earlier by X-ray crystallography<sup>49</sup> (Figure S6). Although reasonable doubts existed that the crystal structure might have been a crystallization artifact, these have now been dispelled by our cryo-EM analysis. Altogether, the high-resolution structural data obtained with X-ray and cryo-EM methods indicate that—in its unliganded form—the human primosome exists in a conformation that is incompatible with primer synthesis without the intervention of major conformational rearrangements. Why should the primosome adopt such a distinct quaternary structure in the absence of its DNA and nucleotide substrates and what should be the function of this postulated autoinhibited state<sup>49</sup> remain unclear.

The structure shows that nsp1 binds to the Pol  $\alpha$  subunit of the primosome (Figure 4). The globular domain of nsp1 docks onto the rim of the polymerase fold of Pol  $\alpha$ , and its binding site is entirely contained within the inactive exonuclease domain of the polymerase, in agreement with our biochemical findings (Figures 1 and 2). Nsp1 binding is mediated by residues in the first, second, and fourth interstrand loop of the globular  $\beta$ -barrel domain and in the  $\alpha$ -helix at the C-end of the second interstrand loop. The interface is of a mixed hydrophobic and polar nature (Figure 5). A cluster of hydrophobic nsp1 residues at the center of the interface—including L27, V28, P109, V111, and G112—becomes buried upon complex formation (Figures 5a and S7), validating the observation that mutations of nsp1 residues V28 and V111 abrogate the interaction (Figure 2c).

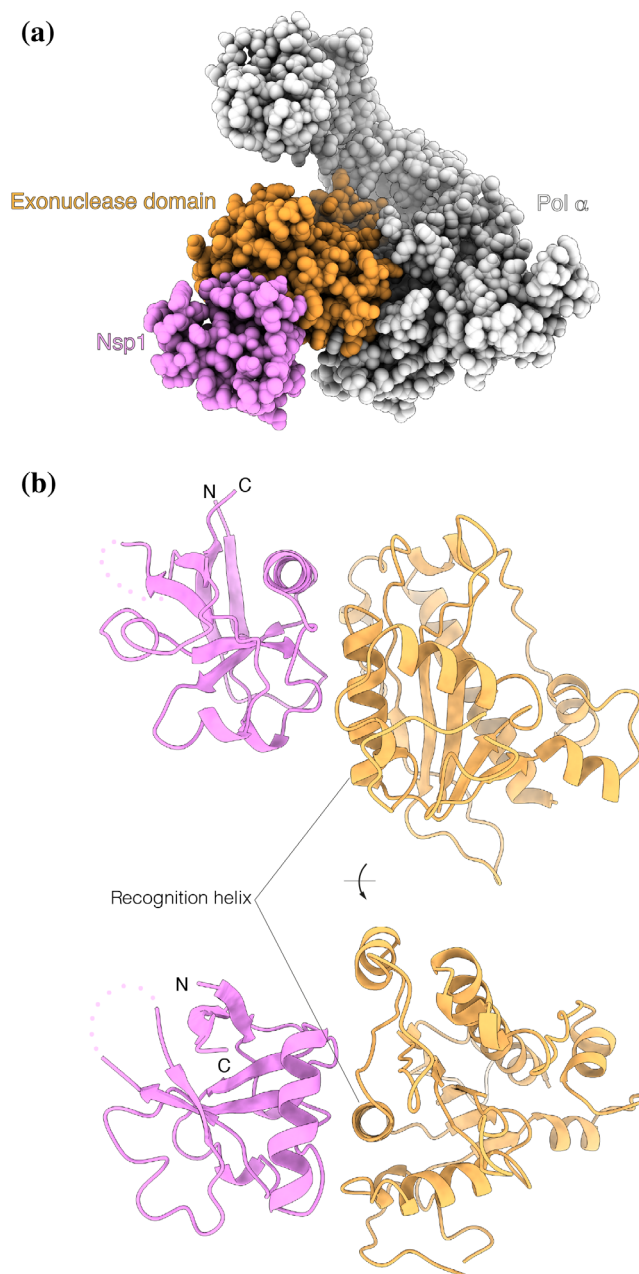
The nsp1-binding area in Pol  $\alpha$  is centered on the second  $\alpha$ -helix of the exonuclease domain, comprising residues 615 to 629 (recognition helix), and extends to include solvent-exposed amino acids flanking the recognition helix,



**FIGURE 3** Cryo-EM structure of the primosome–nsp1 complex. The structure is drawn as a cartoon with helices as solid cylinders and strands as arrows. Each subunit is labeled and uniquely colored, in orange (Pol  $\alpha$ ), red (B subunit), lilac (Pri1), purple (Pri2), and pink (nsp1). The structure is shown in two views, from the side (top panel) and from the top (bottom panel). Cryo-EM, cryo-electron microscopy; nsp1, nonstructural protein 1; Pol  $\alpha$ , DNA polymerase  $\alpha$

from V585 to I662 (Figures 5b and S7). Hydrophobic amino acids in the exonuclease domain that match the hydrophobic surface of nsp1 in the complex are helical residues A624, G620, and F621, the short aliphatic segment 611–VAAT–614 in the loop leading into the recognition helix, and P657 toward the C-end of the binding interface.

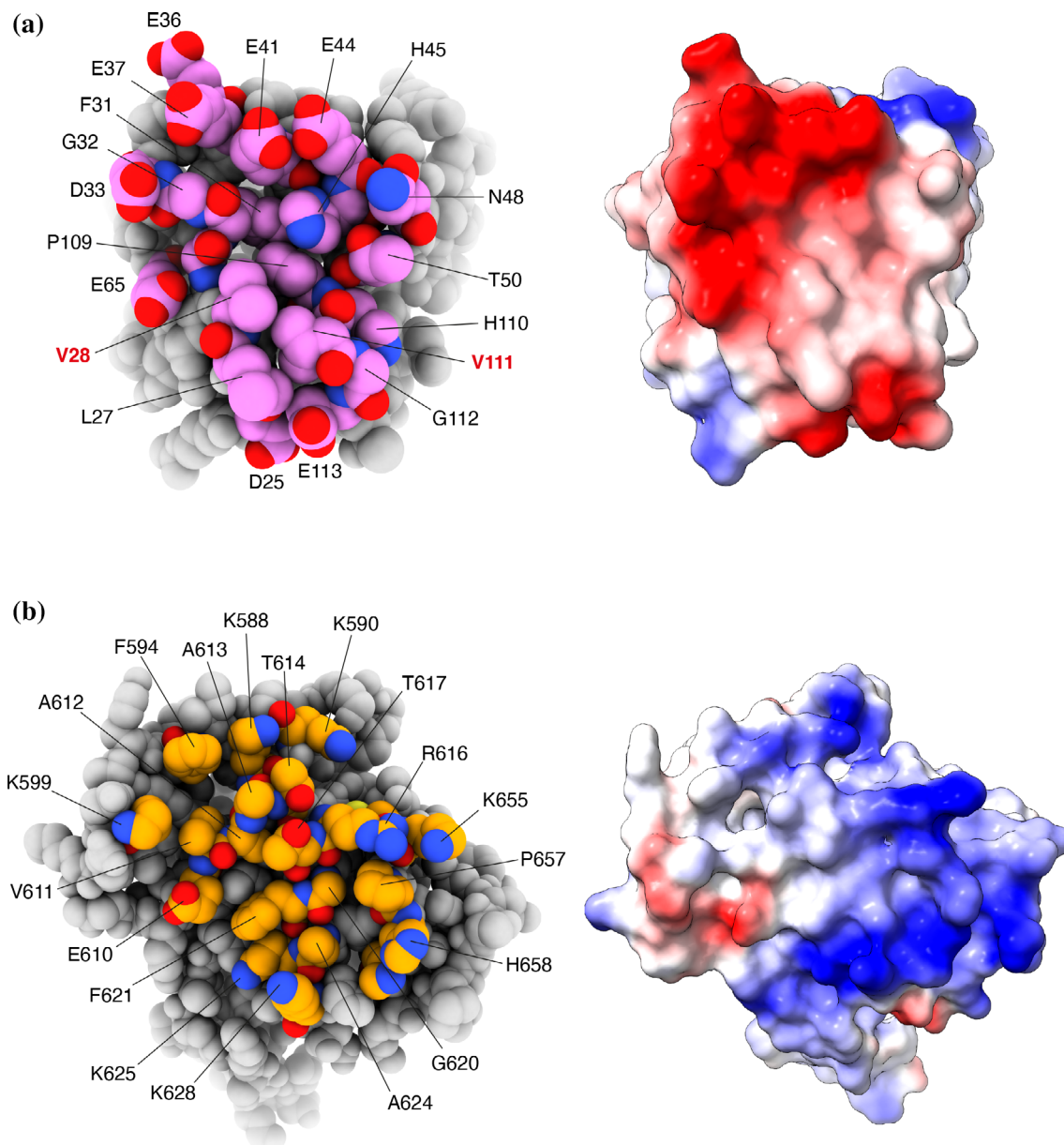
The interaction shows pronounced charge complementarity (Figures 5a,b and S7). A cluster of negatively charged residue in nsp1—D33, E36, E37, E41, and E44 in the helix of the second intrastrand loop, as well as D25, E65, and E113—lines the edge of the interaction area and



**FIGURE 4** nsp1 binds to the exonuclease domain of Pol  $\alpha$ . (a) Space-fill representation of the interaction between nsp1 (pink) and the inactive exonuclease domain of Pol  $\alpha$  (orange) and (b) two views rotated by 90° of the association mode of nsp1 with the exonuclease domain. The proteins are drawn as ribbons in pink (nsp1) and orange (Pol  $\alpha$ ). A disordered loop of nsp1 facing away from the interface is shown as a dotted line. The position of the recognition helix in the exonuclease domain is indicated. nsp1, Nonstructural protein 1; Pol  $\alpha$ , DNA polymerase  $\alpha$

becomes juxtaposed upon complexation to basic residues K625, K628, and R616 in the recognition helix of the exonuclease domain of Pol  $\alpha$ , as well as to surrounding lysine residues K588, K590, K599, K655, and K661.

The presence at the interface of small amino acids such as alanine and glycine, that are required for the intimate



**FIGURE 5** The nsp1–Pol  $\alpha$  exonuclease domain interface. (a) On the left, a space-fill model of nsp1, with interface residues labeled and colored in pink. On the right, the charge distribution of nsp1 mapped over its molecular surface (red: negative charge; blue: positive charge); (b) on the left, a space-fill model of the exonuclease domain of Pol  $\alpha$ , with interface residues labeled and colored in orange, and on the right, the charge distribution of the exonuclease domain mapped over its molecular surface (red: negative charge; blue: positive charge). nsp1, Nonstructural protein 1; Pol  $\alpha$ , DNA polymerase  $\alpha$

association of rather flat protein surfaces, supports the observed interaction between nsp1 and Pol  $\alpha$ . Notably, the nsp1-binding sequence 611-VAAT-614, as well as the 620-GFxxA-624 motif in the recognition helix, are highly conserved in vertebrate Pol  $\alpha$  sequences (Figure S7).

### 3 | DISCUSSION

Here, we have provided a biochemical and structural basis for the interaction between the SARS-CoV-2

virulence factor nsp1 and the primosome, for which evidence first emerged from protein–protein interaction screens between SARS-CoV-2 and human proteins. Our structure shows that nsp1 uses its globular domain to interact with the primosome, distinct from the C-terminal motif employed to target the ribosome and suppress mRNA translation. Thus, our findings provide the first structural evidence of how nsp1 can deploy its conserved middle domain to engage with a host protein.

The ability of nsp1 to use different regions of its sequence to target different physiological processes in the

cell is in keeping with evidence of its pleiotropic pathological effects during infection. Furthermore, as viral proteins are known to exploit existing binding epitopes on the surface of the target protein, the nsp1-binding site in Pol  $\alpha$  might be indicative of as yet unknown physiological partners of the primosome binding at this site. The discovery of a novel protein–protein interaction surface in nsp1 and Pol  $\alpha$  opens up the possibility of developing small-molecule inhibitors aimed at disrupting their association, which might be of therapeutic value in the treatment of COVID-19.

How the interaction of nsp1 with the primosome might promote virulence is currently unclear. However, a rationale for targeting the primosome by SARS-CoV-2 might be envisaged based on recent evidence that mutations in *POLA1*—coding for the catalytic subunit of Pol  $\alpha$ —cause X-linked reticulate pigmentary disorder (XLPDR), a genetic syndrome that is characterized by activation of type I IFN signaling, a persistent inflammatory state and recurrent lung infections.<sup>32,33</sup> A similar set of symptoms is observed in individuals with inherited haploinsufficiency of the *PRIM1* gene, coding for the PriI subunit of primase.<sup>50</sup> In XLPDR fibroblasts, the ability of Pol  $\alpha$ –primase to synthesize RNA/DNA hybrids in the cytoplasm was found to be impaired.<sup>32</sup> This effect was proposed to be responsible for the disease, although the molecular mechanisms by which cytosolic RNA/DNA hybrids might help regulate the antiviral response are unknown.

To test whether nsp1 binding may directly affect the enzymatic activity of the primosome, we performed biochemical assays to measure the RNA and DNA primer synthesis activities of primase and Pol  $\alpha$ , respectively. However, we could not detect an effect on RNA primer synthesis or DNA extension in the presence of an excess of nsp1 protein (Figure S8). This is consistent with the 3D structure of the primosome–nsp1 complex, in which nsp1 is docked in a position that is not expected to interfere with RNA–DNA primer synthesis.

Given the exclusively cytoplasmic localization of SARS and MERS nsp1 upon infection,<sup>30</sup> the interaction with nsp1 appears aimed at targeting the primosome fraction present in the cytoplasm of infected cells. As nsp1 binding did not seem to impact the enzymatic activity of the primosome, we speculate that any effect on primosome activity might be achieved indirectly. Thus, nsp1 binding might be necessary for recruitment of additional factors that inhibit primosome activity or sequester it in a way that makes it inactive; alternatively, nsp1 binding might target the primosome for ubiquitination and degradation (Figure 6). Further study will be required to establish how primosome targeting by SARS-CoV-2 affects the antiviral response of the host cell.

## 4 | MATERIALS AND METHODS

### 4.1 | Cloning, expression, and protein purification

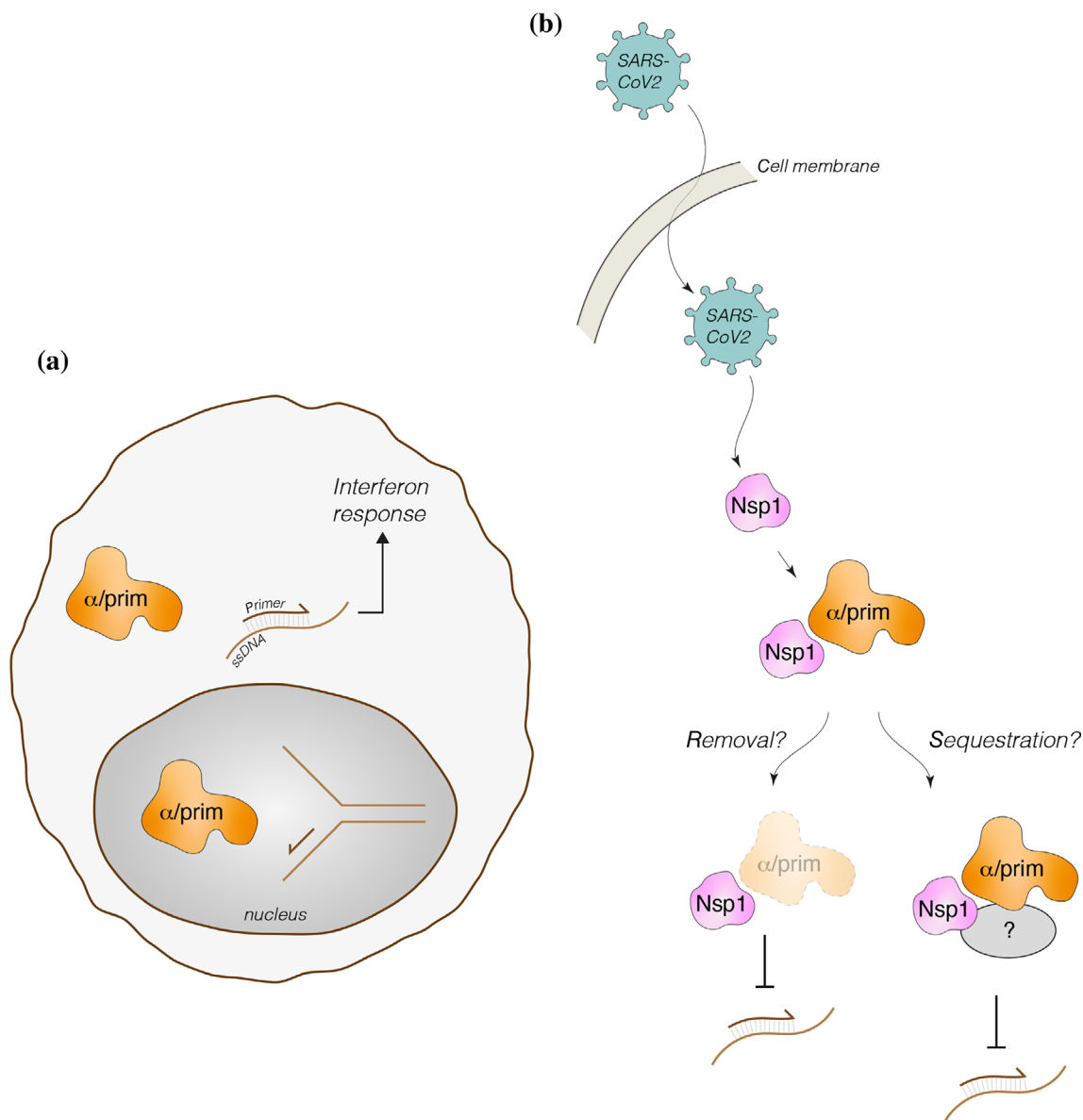
Full-length human primosome (Pol  $\alpha$ : 1–1,462; B subunit: 1–598; PriS: 1–420; PriL: 1–509) and truncated human primosome (Pol  $\alpha$ : 334–1,462; B subunit: 149–598; PriS: 1–420; PriL: 1–509) were expressed in *Sf9* insect cells and purified as described previously.<sup>38</sup> Full-length coronavirus nsp1 genes (SARS-CoV-2, SARS-CoV, MERS, HKU1, NL63, and 229E) were ordered as synthetic gBlocks (IDT) and cloned into bacterial expression vector pMAT11.<sup>51</sup> nsp1 proteins were expressed as His<sub>6</sub>-MBP-tagged fusion proteins from the Rosetta2 (DE3) *Escherichia coli* strain (Novagen). nsp1 fragments (SARS-CoV N-terminus [1–12], middle domain [13–127], and C-terminus [128–180]; SARS-CoV-2 N-terminus [1–12], middle domain [13–127], and C-terminus [128–180]) were cloned and expressed in the same way. Full-length SARS-CoV-2 nsp1 point mutants (V111D, L27D, and V28D) were ordered as synthetic gBlocks (IDT) and cloned and expressed as above.

All His<sub>6</sub>-MBP-fusion proteins were purified by metal affinity chromatography using Ni-NTA agarose (Qiagen). A high-salt wash (500 mM KCl) was performed on-column to remove any traces of nucleic acid from the protein preparation. Proteins were eluted using 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.2, 300 mM KCl, and 200 mM imidazole.

### 4.2 | Cryo-EM sample preparation and grid freezing

SARS-CoV His<sub>6</sub>-MBP-nsp1 (13–127) was TEV-cleaved overnight at 4°C, and the His<sub>6</sub>-MBP tag was subsequently recaptured using amylose resin (NEB). The flow-through was loaded onto a Q-HP anion exchange column (Cytiva) and eluted with a KCl gradient (25 mM HEPES pH 7.2, 80–750 mM KCl, and 2 mM Dithiothreitol [DTT]). Nsp1 and the truncated primosome were subsequently individually buffer-exchanged into cryo-EM buffer (25 mM HEPES pH 7.2, 150 mM KCl, and 1 mM DTT) using a NAP5 column (Cytiva). The two proteins were mixed at final concentrations of 1  $\mu$ M primosome and 10  $\mu$ M nsp1. 1 mM BS3 cross-linker was added, and the sample was incubated on ice before grid freezing using a Vitrobot set at blot force –10 (Thermo Fisher Scientific). The samples were diluted 1:2 with cryo-EM buffer immediately before the application of the protein sample to the grid; 3  $\mu$ l sample was applied to each side of the UltrAuFoil grid (300 mesh, R 1.2/1.3, Quantifoil), which had been glow





**FIGURE 6** Possible functional significance of primosome targeting by SARS-CoV-2 nsp1. (a) In addition to its well-established role in chromosomal DNA duplication during S-phase, a cytoplasmic role for Pol  $\alpha$ -primase has been postulated to synthesize RNA/DNA duplexes that might mediate the interferon response<sup>32</sup>; (b) speculative modes of action of nsp1-dependent SARS-CoV-2 interference with cytoplasmic Pol  $\alpha$ -primase. nsp1, Nonstructural protein 1; Pol  $\alpha$ , DNA polymerase  $\alpha$

discharged on both sides (25 mA, 1 min each side). The grids used for the eventual structure determination were plunge frozen in liquid ethane approximately 45 min after the addition of the BS3 cross-linker.

### 4.3 | Cryo-EM data processing and structure refinement

Data collection parameters, structure refinement details, and statistics are reported in Table S1 and Figure S1; 393 movies (1.11 e/Å<sup>2</sup>/fraction, 40 fractions) were

collected on a Talos Arctica at 200 keV, 92,000 $\times$  magnification, and 1.13 px/Å, with a Falcon III detector in counting mode. Movie processing was carried out in Relion 3.1.<sup>52</sup> The movies were motion corrected using Relion's implementation and CTFs were estimated using CTFFIND4-1.<sup>53</sup> An initial set of 165,011 particles was auto-picked and used to generate 2D templates for template-based picking of a new set of 183,301 particles. The template-picked particle set was reduced to 97,680 particles by rounds of 2D classification and used to reconstruct an initial 3D volume. 3D classification into four models generated one model with 47,542 particles

(48.6%) which was refined to 7.29 Å. The 3D volume showed clear evidence of the four primosome subunits and of the bound nsp1 protein; the initial 3D model was used as starting point for 3D classification during processing of the higher resolution data set.

A total of 2,919 movies (0.98 e/Å<sup>2</sup>/fraction, 48 fractions) were collected on Titan Krios at 300 keV, 130,000 magnification, and 0.326 Å px/Å with a K3 detector in super-resolution counting mode. Movie processing was carried out in Relion 3.1.<sup>52</sup> The movies were motion corrected and 2× binned using Relion's own implementation, and CTFs were estimated using CTFFIND4-1.<sup>53</sup> An initial set of 314,987 particles was auto-picked and improved by rounds of 2D classification, to generate 2D templates for template-based picking of a new set of 709,068 particles. The particle set was improved by multiple rounds of selection by 2D classification, and a first 3D classification into four models was performed using the initial 3D model obtained for the Talos data set. Class #1 (251,901 particles) was selected and a second round of 3D classification was performed, leading to a new 3D Class #4 with 233,476 particles. 3D refinement of this class yielded a map at 4.12 Å, which was sharpened in Phenix<sup>54</sup> using “phenix.local\_aniso\_sharpen” to a resolution of 3.8 Å (0.143 FSC) and used for model interpretation and model building.

For model building, the crystal structure of human Pol α-primase (PDB ID 5EXR) was fitted manually in the map with ChimeraX<sup>55</sup> and real-space refined in Phenix,<sup>54</sup> together with local manual rebuilding in Coot.<sup>56</sup> Additional density corresponding to the known structure of nsp1 was clearly detectable, although at a lower map threshold, bound to Pol α. A model of SARS nsp1 was built based on PDB entries 7K3N and 2HSX and docked into the map in ChimeraX.<sup>55</sup> The Pol α-primase-nsp1 complex was further real-space refined in Phenix.<sup>54</sup>

Inspection of the 2D/3D classes and of the 3D refined map for the Pol α-primase-nsp1 complex showed that the density corresponding to the Pri1 subunit of primase was weaker. This observation was explained by the looser contact of Pri1 with the rest of the Pol α-primase, which is primarily mediated by the known interface with Pri242. To improve the map of the rest of Pol α-primase and of the bound nsp1, a mask was prepared in ChimeraX<sup>55</sup> that excluded the Pri1 lobe and used to generate a set of subtracted particles from the 233,476 particles set used to generate the original 3D refined map. A new initial 3D model was first generated, and after a round of 3D classification, Class #3 (103,763 particles) was selected. 3D refinement of this class generated a 4.40 Å map, that was sharpened using “phenix.local\_aniso\_sharpen” to a resolution of 3.6 Å (0.143 FSC) and used for model interpretation and model building.

The map showed slightly better high-resolution features for the primosome core and clearer features for the nsp1 fold and was therefore used as a visual aid in the refinement of the primosome-nsp1 complex structure.

#### 4.4 | Pull-downs

A total of 150 μl amylose resin (NEB) was equilibrated in pull-down buffer (1× phosphate-buffered saline, 5% glycerol, 0.1% Igepal CA-630, and 1 mM tris(2-carboxyethyl) phosphine or DTT). An excess of His<sub>6</sub>-MBP-nsp1 was added to saturate the beads with bait protein. The His<sub>6</sub>-MBP tag protein was used as negative control. Beads were washed twice with pull-down buffer to remove excess, unbound bait protein. BSA was added to a final concentration of 2.5 mg/ml<sup>-1</sup>. Primosome (full length or truncated) was added to the beads, which were subsequently incubated at 4°C for 1 hr with rolling. The beads were washed four times with 1 ml pull-down buffer. Bound protein was eluted using 150 μl pull-down buffer supplemented with 50 mM maltose. Samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

#### 4.5 | Primosome activity assay

Full-length His<sub>6</sub>-MBP-nsp1 (from SARS-CoV-2 and SARS-CoV) was TEV-cleaved overnight at 4°C. The His<sub>6</sub>-MBP tag was recaptured using amylose resin (NEB). The flow-through was loaded onto a Q-HP anion exchange column (Cytiva), and nsp1 was eluted with a KCl gradient (25 mM HEPES pH 7.2, 80–750 mM KCl, and 2 mM DTT). Pure nsp1 was subsequently buffer exchanged into 25 mM HEPES pH 7.2, 100 mM KCl, and 1 mM DTT using a NAP5 column (Cytiva).

An activity assay, to monitor the ability of Pol α to extend an RNA primer, was performed using a pre-annealed DNA-RNA template (DNA45: 5'- TTCTTT ATCTCTTATTTCTTCTATTTTCCACGCGCCTTCTAT TT-3'; RNA13: 5'-[6FAM]-GGCGGUGGAAAA-3', Sigma Aldrich). The reaction buffer comprised 25 mM HEPES pH 7.2, 100 mM KCl, 10 mM Mg(OAc)<sub>2</sub>, and 1 mM DTT. Each 100 μl reaction contained 2 μM DNA-RNA template, 1 mM dNTP, 10 μM nsp1, and 0.1 μM full-length primosome. Reactions were incubated at 25°C, and 30-μl aliquots were removed and quenched (by adding an equal volume of 95% formamide and 25 mM EDTA) at the indicated time points. Samples were heated to 70°C for 2 min, then loaded onto a 19% urea-polyacrylamide gel (running conditions: 450 V, 90 min, 0.5× TBE buffer). Gels were initially imaged by laser scanning (473 nm laser, Typhoon FLA 9000, GE

Healthcare), to detect only the 6FAM-labeled products. Gels were subsequently stained using a 1:10000 dilution of SYBR Gold Stain (Thermo Fisher Scientific) in 0.5× TBE buffer and reimaged in the same way to detect all nucleic acid.

## ACKNOWLEDGMENTS

We would like to thank Ed Greenwood, Thomas Crozier, and Paul Lehner (Cambridge Institute of Therapeutic Immunology and Infectious Disease) and Marc de la Roche (Biochemistry Department) for helpful discussions and for sharing unpublished results. This work was funded by a Wellcome Trust Investigator award (104641/Z/14/Z) to LP.

## CONFLICT OF INTERESTS

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Mairi L Kilkenny:** Conceptualization (equal); data curation (lead); investigation (lead); supervision (equal); writing – original draft (lead); writing – review and editing (equal). **Charlotte E Veale:** Data curation (supporting); investigation (supporting). **Amir Guppy:** Data curation (supporting); investigation (supporting). **Steven W Hardwick:** Data curation (supporting). **Dimitri Y Chirgadze:** Data curation (supporting). **Neil J Rzechorzek:** Data curation (supporting); investigation (supporting); writing – review and editing (supporting). **Joseph D Maman:** Data curation (supporting); investigation (supporting); writing – review and editing (supporting). **Luca Pellegrini:** Conceptualization (equal); funding acquisition (lead); project administration (lead); supervision (lead); writing – original draft (equal); writing – review and editing (lead).

## DATA AVAILABILITY STATEMENT

Coordinates and 3D reconstructions for the primosome–nsp1 complex have been deposited under accession codes 7OPL (PDB), 13020 (EMDB), and 13021 (EMDB).

## ORCID

Luca Pellegrini  <https://orcid.org/0000-0002-9300-497X>

## REFERENCES

- Artika IM, Dewantari AK, Wiyatno A. Molecular biology of coronaviruses: Current knowledge. *Heliyon*. 2020;6:e04743.
- Wang Y, Grunewald M, Perlman S. Coronaviruses, methods and protocols. *Methods Mol Biol*. 2020;2(203):1–29.
- Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. The molecular virology of coronaviruses. *J Biol Chem*. 2020;295:12910–12934.
- Cui J, Li F, Shi Z-L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol*. 2019;17:181–192.
- Fung TS, Liu DX. Human coronavirus: Host-pathogen interaction. *Annu Rev Microbiol*. 2019;73:1–29.
- Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265–269.
- Zhou P, Yang X-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;579:270–273.
- Lazear HM, Schoggins JW, Diamond MS. Shared and distinct functions of type I and type III interferons. *Immunity*. 2019;50:907–923.
- Blanco-Melo D, Nilsson-Payant BE, Liu W-C, et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell*. 2020;181:1036–1045.
- Xia H, Cao Z, Xie X, et al. Evasion of type I interferon by SARS-CoV-2. *Cell Rep*. 2020;33:108234.
- Lei X, Dong X, Ma R, et al. Activation and evasion of type I interferon responses by SARS-CoV-2. *Nat Commun*. 2020;11:3810.
- Wathelet MG, Orr M, Frieman MB, Baric RS. Severe acute respiratory syndrome coronavirus evades antiviral signaling: Role of nsp1 and rational design of an attenuated strain. *J Virol*. 2007;81:11620–11633.
- Narayanan K, Huang C, Lokugamage K, et al. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J Virol*. 2008;82:4471–4479.
- Kamitani W, Narayanan K, Huang C, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. *Proc Natl Acad Sci USA*. 2006;103:12885–12890.
- Terada Y, Kawachi K, Matsuura Y, Kamitani W. MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. *Virology*. 2017;511:95–105.
- Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat Struct Mol Biol*. 2009;16:1134–1140.
- Lokugamage KG, Narayanan K, Nakagawa K, et al. Middle East respiratory syndrome coronavirus nsp1 inhibits host gene expression by selectively targeting mRNAs transcribed in the nucleus while sparing mRNAs of cytoplasmic origin. *J Virol*. 2015;89:10970–10981.
- Narayanan K, Ramirez SI, Lokugamage KG, Makino S. Coronavirus nonstructural protein 1: Common and distinct functions in the regulation of host and viral gene expression. *Virus Res*. 2015;202:89–100.
- Shen Z, Wang G, Yang Y, et al. A conserved region of non-structural protein 1 from alphacoronaviruses inhibits host gene expression and is critical for viral virulence. *J Biol Chem*. 2019;294:13606–13618.
- Almeida MS, Johnson MA, Herrmann T, Geralt M, Wuthrich K. Novel  $\beta$ -barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J Virol*. 2007;81:3151–3161.
- Jansson AM. Structure of alphacoronavirus transmissible gastroenteritis virus nsp1 has implications for coronavirus nsp1 function and evolution. *J Virol*. 2013;87:2949–2955.
- Shen Z, Ye G, Deng F, et al. Structural basis for the inhibition of host gene expression by porcine epidemic diarrhea virus nsp1. *J Virol*. 2018;92:e01896-17.

23. Clark LK, Green TJ, Petit CM. Structure of nonstructural protein 1 from SARS-CoV-2. *J Virol.* 2020;95:e02029-20.
24. Semper C, Watanabe N, Savchenko A. Structural characterization of nonstructural protein 1 from SARS-CoV-2. *iScience.* 2021;24:101903.
25. Shi M, Wang L, Fontana P, et al. SARS-CoV-2 Nsp1 suppresses host but not viral translation through a bipartite mechanism. *Ssrn Electron J.* 2020. <https://doi.org/10.1101/2020.09.18.302901>
26. Schubert K, Karousis ED, Jomaa A, et al. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat Struct Mol Biol.* 2020;27:959–966.
27. Thoms M, Buschauer R, Ameisemeier M, et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science.* 2020;369:1249–1255.
28. Lapointe CP, Grosely R, Johnson AG, Wang J, Fernández IS, Puglisi JD. Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. *Proc Natl Acad Sci USA.* 2021;118:e2017715118.
29. Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature.* 2020;583:459–468.
30. Gordon DE, Hiatt J, Bouhaddou M, et al. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science.* 2020;370:eabe9403.
31. Pellegrini L. The Pol  $\alpha$ -primase complex. *Subcell Biochem.* 2012;62:157–169.
32. Starokadomskyy P, Gemelli T, Rios JJ, et al. DNA polymerase  $\alpha$  regulates the activation of type I interferons through cytosolic RNA:DNA synthesis. *Nat Immunol.* 2016;17:495–504.
33. Starokadomskyy P, Perez-Reyes AE, Burstein E. Immune dysfunction in Mendelian disorders of POLA1 deficiency. *J Clin Immunol.* 2021;41:285–293.
34. Esch HV, Colnaghi R, Freson K, et al. Defective DNA polymerase  $\alpha$ -primase leads to X-linked intellectual disability associated with severe growth retardation, microcephaly, and hypogonadism. *Am J Hum Genet.* 2019;104:957–967.
35. Chen Z, Wang C, Feng X, et al. Interactomes of SARS-CoV-2 and human coronaviruses reveal host factors potentially affecting pathogenesis. *EMBO.* 2021;40:e107776.
36. Simon AC, Zhou JC, Perera RL, et al. A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. *Nature.* 2014;510:293–297.
37. Kolinjivadi AM, Sannino V, Antoni AD, et al. Smarcal1-mediated fork reversal triggers Mre11-dependent degradation of nascent DNA in the absence of Brca2 and stable Rad51 nucleofilaments. *Mol Cell.* 2017;67:867–881.
38. Kilkenny ML, Simon AC, Mainwaring J, Wirthensohn D, Holzer S, Pellegrini L. The human CTF4-orthologue AND-1 interacts with DNA polymerase  $\alpha$ /primase via its unique C-terminal HMG box. *Open Biol.* 2017;7:170217.
39. Evrin C, Maman JD, Diamante A, Pellegrini L, Labib K. Histone H2A-H2B binding by pol  $\alpha$  in the eukaryotic replisome contributes to the maintenance of repressive chromatin. *EMBO.* 2018;37:e99021.
40. Klinge S, Núñez-Ramírez R, Llorca O, Pellegrini L. 3D architecture of DNA Pol alpha reveals the functional core of multi-subunit replicative polymerases. *EMBO.* 2009;28:1978–1987.
41. Kilkenny ML, Piccoli GD, Perera RL, Labib K, Pellegrini L. A conserved motif in the C-terminal tail of DNA polymerase  $\alpha$  tethers primase to the eukaryotic replisome. *J Biol Chem.* 2012;287:23740–23747.
42. Kilkenny ML, Longo MA, Perera RL, Pellegrini L. Structures of human primase reveal design of nucleotide elongation site and mode of pol  $\alpha$  tethering. *Proc Natl Acad Sci USA.* 2013;110:15961–15966.
43. Suwa Y, Gu J, Baranovskiy AG, Babayeva ND, Pavlov YI, Tahirov TH. Crystal structure of the human pol  $\alpha$  B subunit in complex with the C-terminal domain of the catalytic subunit. *J Biol Chem.* 2015;290:14328–14337.
44. Baranovskiy AG, Zhang Y, Suwa Y, et al. Crystal structure of the human primase. *J Biol Chem.* 2015;290:5635–5646.
45. Perera RL, Torella R, Klinge S, Kilkenny ML, Maman JD, Pellegrini L. Mechanism for priming DNA synthesis by yeast DNA Polymerase  $\alpha$ . *eLife.* 2013;2:e00482.
46. Baranovskiy AG, Babayeva ND, Suwa Y, Gu J, Pavlov YI, Tahirov TH. Structural basis for inhibition of DNA replication by aphidicolin. *Nucleic Acids Res.* 2014;42:14013–14021.
47. Coloma J, Johnson RE, Prakash L, Prakash S, Aggarwal AK. Human DNA polymerase  $\alpha$  in binary complex with a DNA:DNA template-primer. *Sci Rep.* 2016;6:23784.
48. Klinge S, Hirst J, Maman JD, Krude T, Pellegrini L. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nature Structural & Molecular Biology.* 2007;14:875–877.
49. Baranovskiy AG, Babayeva ND, Zhang Y, et al. Mechanism of concerted RNA–DNA primer synthesis by the human primosome. *J Biol Chem.* 2016;291:10006–10020.
50. Parry DA, Tamayo-Orrego L, Carroll P, et al. PRIM1 deficiency causes a distinctive primordial dwarfism syndrome. *Gene Dev.* 2020;34:1520–1533.
51. Peranen J, Rikkonen M, Hyvonen M, Kaariainen L. T7 vectors with modified T7lac promoter for expression of proteins in *Escherichia coli*. *Anal Biochem.* 1996;236:371–373.
52. Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol.* 2012;180:519–530.
53. Rohou A, Grigorieff N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol.* 2015;192:216–221.
54. Adams PD, Afonine PV, Bunkoczi G, et al. PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Cryst.* 2010;D66:213–221.
55. Goddard TD, Huang CC, Meng EC, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* 2018;27:14–25.
56. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Cryst.* 2004;D60:2126–2132.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Kilkenny ML, Veale CE, Guppy A, Hardwick SW, Chirgadze DY, Rzechorzek NJ, et al. Structural basis for the interaction of SARS-CoV-2 virulence factor nsp1 with DNA polymerase  $\alpha$ -primase. *Protein Science.* 2021;1–12. <https://doi.org/10.1002/pro.4220>