











## Article

# Insights into the SARS-CoV-2 ORF6 Mechanism of Action

Elena Krachmarova<sup>1</sup> , Peicho Petkov<sup>2,\*</sup> , Elena Lilkova<sup>3,\*</sup> , Nevena Ilieva<sup>3</sup> , Miroslav Rangelov<sup>4</sup> , Nadezhda Todorova<sup>5</sup> , Kristina Malinova<sup>1</sup>, Rossitsa Hristova<sup>1</sup> , Genoveva Nacheva<sup>1</sup> , Anastas Gospodinov<sup>1</sup> , Leandar Litov<sup>2</sup> 

<sup>1</sup> Institute of Molecular Biology "Acad. Roumen Tsanev", Bulgarian Academy of Sciences

<sup>2</sup> Faculty of Physics, Sofia University "St. Kliment Ohridski"

<sup>3</sup> Institute of Information and Communication Technologies, Bulgarian Academy of Sciences

<sup>4</sup> Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences

<sup>5</sup> Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences

\* Correspondence: peicho@phys.uni-sofia.bg (P.P.); agg@bio21.bas.bg (A.G.)

**Abstract:** ORF6 is responsible for suppressing the immune response of cells infected by the SARS-CoV-2 virus. It is also the most toxic protein of SARS-CoV-2, and its actions are associated with the viral pathogenicity. Here we study *in silico* and *in vitro* the structure of the protein, its interaction with RAE1, and the mechanism of action behind its high toxicity. We show both computationally and experimentally that SARS-CoV-2 ORF6, embedded in the membranes of the endoplasmic reticulum, binds to RAE1 and sequesters it in the cytoplasm, thus depleting its availability in the nucleus and impairing nucleocytoplasmic mRNA transport. This negatively affects the cellular genome stability by compromising the cell cycle progression into the S-phase and by promoting the accumulation of RNA-DNA hybrids. Understanding the multiple ways in which ORF6 affects DNA replication may also have important implications for elucidating the pathogenesis of SARS-CoV-2 and developing therapeutic strategies to mitigate its deleterious effects on host cells.

**Keywords:** SARS-CoV-2 ORF6 – RAE1 binding, 3D molecular structure modelling, molecular dynamics simulations, cell cycle, cyclin E, R-loops, DNA replication stress

## 1. Introduction

Nearly three years after its outbreak, at the end of 2019, the unprecedented global COVID-19 pandemic is starting to subside due to the successful development of vaccines and antiviral treatments. Nonetheless, COVID-19 still poses an acute global health emergency due to its numerous long-lasting adverse health effects on the world's population. The causative agent of the pandemic is the SARS-CoV-2 beta-coronavirus, which is an enveloped, positive sense, single-stranded RNA virus [1]. The genome of the virus is ~30 kb in size and encodes 29 proteins – 4 structural, 16 nonstructural, and 9 accessory proteins [2]. While latter are not strictly necessary for viral replication and assembly, they play a crucial role in immune evasion and viral pathogenesis [3,4].

Open reading frame 6 (ORF6) is an accessory protein, uniquely expressed in the Sarbecovirus subgenus of the Betacoronaviruses, to which the SARS-CoV and SARS-CoV-2 viruses belong [5]. Its main function is to antagonize the host cell interferon signaling pathways, thus limiting the immune response to the infection [5–9]. SARS-CoV-2 has a high degree of similarity with SARS-CoV, however, a limited degree of similarity is observed in the accessory proteins with ORF6 having the lowest homology of only 69% [1,10,11].

SARS-CoV-2 ORF6 is a 61 amino acids (aa) polypeptide with an amphipathic N-terminal portion (aa 1–40) and a highly polar C terminus. It was shown that ORF6 is mainly localized in the endoplasmic reticulum (ER) and Golgi apparatus (GA) [12]. Its N-terminal aa residues 2–37 form an  $\alpha$ -helix embedded in the membrane [13,14]. Similarly to the other ORF6 proteins, the SARS-CoV-2 ORF6 function is to limit the immune response to the infection. This is achieved by two different mechanisms – on the one hand, it

antagonizes STAT1 nuclear translocation [6–8], and on the other, it preventing mRNA transport via interaction with the nucleopore complex components nucleoporin 98 (NUP98) and ribonucleic acid export 1 (RAE1) [8,9,15,16]. However, the detailed mechanism of the pathogenic immune-suppressive action of ORF6 on the host target proteins is not yet well understood. Furthermore, ORF6 was shown to significantly contribute to the pathogenicity of the virus, having the highest cytotoxicity among the SARS-CoV-2 proteins [12]. All this prompts active search for possible inhibitors of this protein. However, despite the significant research focus and efforts, ORF6 is one of the dozen proteins of SARS-CoV-2, whose 3D structures have not yet been resolved [2], hindering substantially the development of inhibitors.

In this study, we report results from both *in silico* and *in vitro* investigations on the SARS-CoV-2 ORF6 protein mechanism of action. Our data demonstrate that ORF6 accomplishes its effects by binding the RAE1 protein and sequestering it in the cytoplasm. We built a 3D model of the SARS-CoV-2 ORF6 protein embedded in a ER membrane and studied its interaction with RAE1 using molecular dynamics (MD) simulations. RAE1 binds one or more ORF6 molecules very stably, which immobilizes it on the ER membranes in the cytoplasm. As a consequence, the concentration of RAE1 in the cell's nucleus is reduced significantly, strongly suppressing mRNA transport from nucleus to the cytoplasm. These observations were confirmed experimentally by studying the co-localisation of both proteins in ORF6 overexpressing and control cells. Further, we studied the impact of the SARS-CoV-2 ORF6 interaction with RAE1 on the host cell, specifically, on DNA replication. The effects are manifested in at least two ways – by compromising the cell progression into the S-phase, and by promoting the accumulation of RNA–DNA hybrids, which are a major source of genome instability.

## 2. Results

### 2.1. Structure of ORF6 in Water Solvent

For computational studies of the ORF6–RAE1 interaction the 3D structure of ORF6 is a prerequisite. The EXCALATE4COV input model of ORF6 is presented in Suppl. Fig. S1. It contains four somewhat distorted helices (aa Leu<sup>4</sup>–Thr<sup>10</sup>, Leu<sup>16</sup>–Trp<sup>27</sup>, Asp<sup>30</sup>–Leu<sup>40</sup> and Glu<sup>46</sup>–Glu<sup>54</sup>), connected by short loops (2–5 aa long). Starting from this structure, we performed additional 7.6  $\mu$ s of folding MD simulation. The secondary structure plot of the protein along the MD trajectory is shown in Suppl. Fig. S2. The protein remained primarily helical. The distorted helices stabilized and transitioned to very well defined  $\alpha$ -helices, encompassing Leu<sup>4</sup>–Val<sup>9</sup>, Leu<sup>16</sup>–Val<sup>24</sup>, Thr<sup>31</sup>–Ser<sup>43</sup> and Glu<sup>46</sup>–Glu<sup>55</sup>. The most important structural alteration was the change in the helices mutual orientation. The total protein solvent accessible surface area (SASA), as well as its decomposition in hydrophobic, polar and charged contributions are presented in Suppl. Fig. S3. After an initial jump, the total SASA of the protein dropped again to around its initial values and remained stable after the 2nd  $\mu$ s. However, because of tertiary structure changes, the SASA of the hydrophobic aa residues decreased by about 10%, whereas the SASA of the hydrophilic and polar residues increased by 7% and 11% respectively.

A representative conformation of the folded ORF6 for the subsequent simulations was selected based on the cluster analysis of the folding trajectory with a cut-off of 3.8 Å. The results are presented in Suppl. Fig. S4. total of 51 clusters were identified with the chosen cut-off, with the five largest clusters capturing 89% of the conformations. Notably, within the first 1.5  $\mu$ s the protein settled in a local energy minimum and explored conformations that all belonged to the largest identified cluster. Its centroid conformation is shown in Suppl. Fig. S5.

### 2.2. Structure of ORF6 in ER Membrane

ORF6 is considered a membrane protein [8,15–17], therefore the so obtained structure was inserted into a lipid bilayer, modelling an ER membrane, the protein C-terminus

being placed at the outer leaflet of the membrane (Suppl. Fig. S6). The thus obtained conformation was subjected to a 1.4  $\mu$ s production MD simulation.

Throughout the simulation, the ORF6 molecule remained largely organized as an  $\alpha$ -helical bundle (see secondary structure plot in Suppl. Fig. S7), however the different helices underwent some changes in their mutual orientation. Experimental data shows that the part of the ORF6 protein that is responsible for its biological activity is its C-terminus, so it is expected to be flexible and solvent exposed. Hence, a conformation of interest should be one, in which the C-terminal tail of ORF6 is not buried in the membrane, but remains freely exposed to the solvent. Initially, the C-terminus (aa 41-61) was partly structured into an  $\alpha$ -helix, which started to unwind and become disordered after about 1.16  $\mu$ s as it began to stick up from the lipid bilayer. This is manifested in the secondary structure plot Suppl. Fig. S7. The exit of the C-terminus from the membrane is also evident from the evolution of the gyration radius of the protein (Suppl. Fig. S8a): after 1.16  $\mu$ s it almost doubled its initial value of 1.15 nm to nearly 2.1 nm at the end of the simulation as the C-terminal tail unfolded and started to sway around into the solvent bulk. The RMSF analysis (Suppl. Fig. S8b) showed that while the N-terminal part of ORF6 (aa 1-40) remained quite stable, the C-terminus is its most flexible part both prior and after its unfolding and sticking out into the solvent, when the fluctuations of the positions of its residues doubled. About 300 ns additional MD simulation was performed after this point to ensure that a stable conformational state was achieved with only the C-terminal part of ORF6 outside of the membrane and not the whole protein. As evident in the SASA plot (Suppl. Fig. S9), the  $\alpha$ -helical globular part of the protein (aa 1-40) remained firmly submerged in the membrane and not solvent exposed, while the SASA of the C-terminus doubled from about 12.5 nm<sup>2</sup> to about 24.8 nm<sup>2</sup> at the end of the simulation. Thus, using molecular modelling we showed that it is entirely possible that ORF6 is a membrane-bound and not just a membrane-peripheral protein [18]. The final conformation of this simulation was used as an input model for studying the interaction of ORF6 and RAE1 (Suppl. Fig. S10).

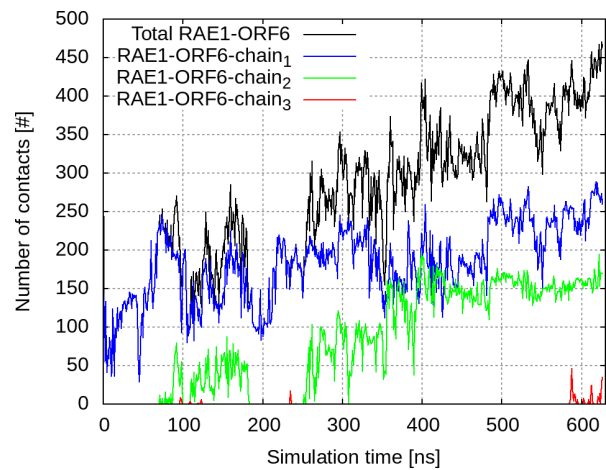
### 2.3. Interaction of RAE1 and ORF6

RAE1 binding propensity with respect to the SARS-CoV-2 ORF6 protein was probed in a simulation with a starting conformation shown in Suppl. Fig. S11. The RAE1 structure was completed and simulated for 100 ns MD as described in Sec. 4.3, the generated trajectory being subjected to cluster analysis with a 0.25 nm cut-off (Suppl. Fig. S12). The centroid of the largest cluster (Suppl. Fig. S12b) was used as an input RAE1 model in the interaction simulation. A 630 ns production MD simulation was carried out of this system.

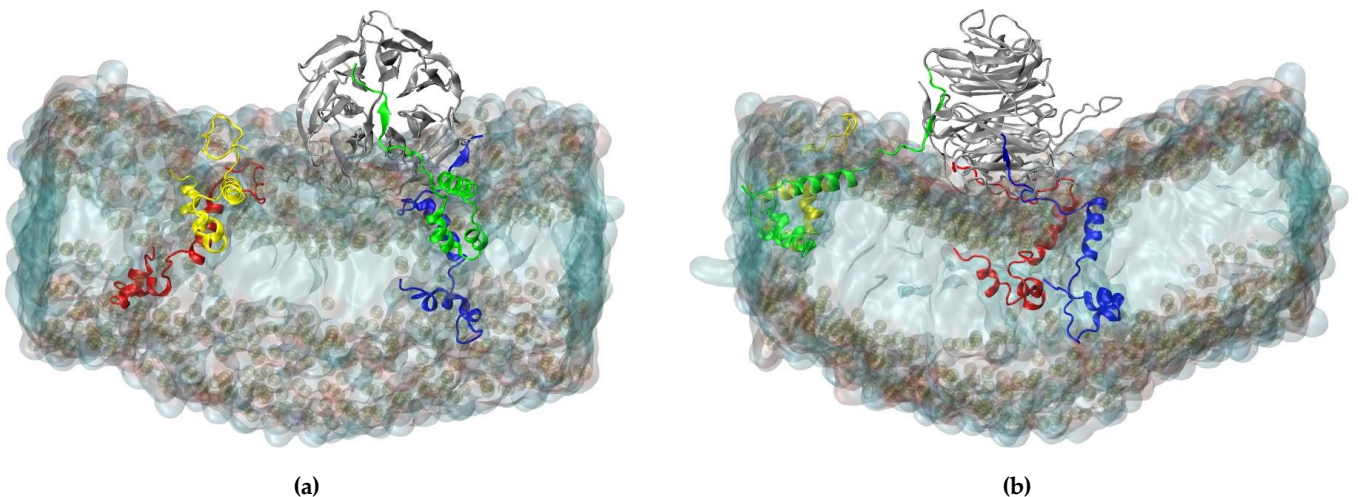
The C-termini of the ORF6 molecules started to interact with the nearby RAE1 almost immediately. Out of the four viral protein molecules, three form close contacts (i.e. heavy atoms from the ORF6 molecules are within a cut-off distance of 6 Å to heavy atoms of the RAE1 molecule) with the mRNA export factor (Fig. 1). The final conformation in the RAE1-ORF6 simulation is presented in Fig. 2.

In the beginning of the simulation, the total number of RAE1-ORF6 contacts was fully determined by the binding of the first ORF6 molecule, represented in blue in Figs. 1 and 2). ORF6-chain<sub>1</sub> bound spontaneously to the RAE1, with some of the binding site residues being Ile<sup>207</sup>-Leu<sup>211</sup>, Gly<sup>239</sup>, Ile<sup>242</sup>-His<sup>243</sup>, Asn<sup>254</sup>-Lys<sup>258</sup>, Trp<sup>300</sup>, Arg<sup>305</sup>-Leu<sup>308</sup> and Thr<sup>310</sup> (Fig. 3a,b). Recently, two experimental structures of the interaction between the end of the ORF6 C-terminus and the RAE1-NUP98 complex became available [9,19]. A comparison of the experimentally and computationally obtained RAE1-ORF6 interaction interfaces is shown in Fig. 3. The identified in our simulation binding site of the first ORF6 molecule and RAE1 agrees excellently with experimentally observed ones.

Shortly after the interaction with the first protein started, a second ORF6 chain bound to the RNA transport protein at residues Phe<sup>4</sup>-Ser<sup>13</sup>, Pro<sup>49</sup>-Thr<sup>51</sup>, Ser<sup>95</sup>-Asp<sup>97</sup>, His<sup>135</sup>-Pro<sup>139</sup>, Ser<sup>142</sup>, Ile<sup>179</sup>-Tyr<sup>180</sup>, Lys<sup>222</sup>-Lys<sup>227</sup>, Pro<sup>283</sup>-His<sup>285</sup>, and His<sup>325</sup>-Asn<sup>326</sup>. This binding surface lay on the entrance of the central tunnel of the RAE1 beta-propeller (Suppl. Fig. S13). The binding occurred right on the opposite side of the NUP98 recruitment site of



**Figure 1.** Number of contacts within 6 Å between RAE1 and the four ORF6 molecules. The total number of contacts between all ORF6 molecules and RAE1 is shown in black, and contacts with individual ORF 6 chains are represented respectively with a, blue, green, and red curves.

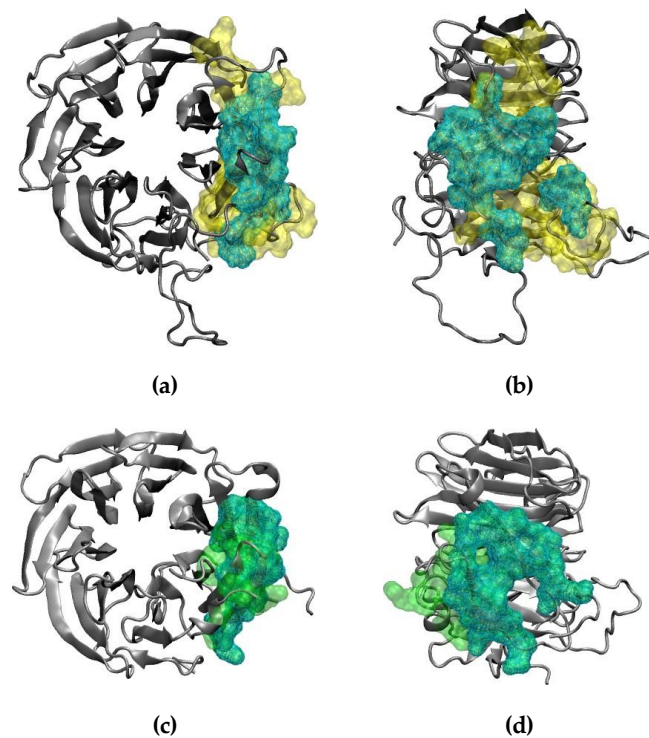


**Figure 2.** (a) X- and (b) Y-projections of RAE1, bound to three ORF6 molecules. The RAE1 protein is in grey cartoon representation and the four ORF6 are in blue, green, red and yellow cartoon representation, respectively.

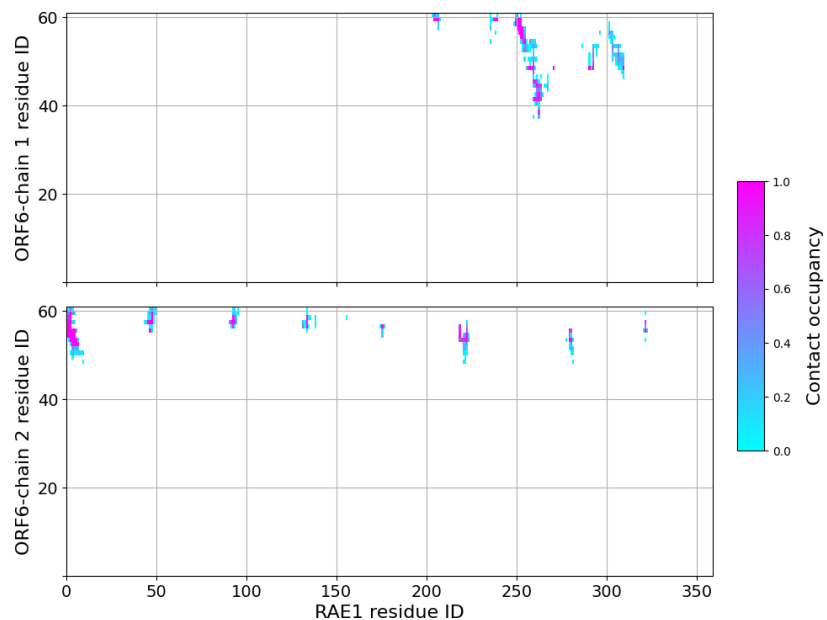
the RAE1 molecule [9,19–21]. The third ORF6 chain interacted sporadically with RAE1 residues Thr<sup>22</sup>-Thr<sup>23</sup>, Thr<sup>35</sup>-Ser<sup>36</sup>, Thr<sup>77</sup>-Pro<sup>79</sup>. The contact interfaces between ORF chains 1 and 2 and RAE1 in the last 100 ns of the simulation are shown in Fig. 4.

The results of this *in silico* study clearly demonstrate that ORF6, embedded in the membrane of the ER and possibly the Golgi apparatus, interacts with RAE1 and binds to it to form a very stable complex. In fact, a single RAE1 molecule is able to simultaneously bind to several ORF6 molecules, which completely immobilizes the transport factor on the ER/GA membrane. Thus, it is reasonable to explain some of the pathogenic effects of SARS-CoV-2 ORF6 by this interaction, which anchors available cytoplasmic RAE1 proteins to the ER and/or GA, thereby depleting RAE1 availability in the nucleus and restricting mRNA transport.





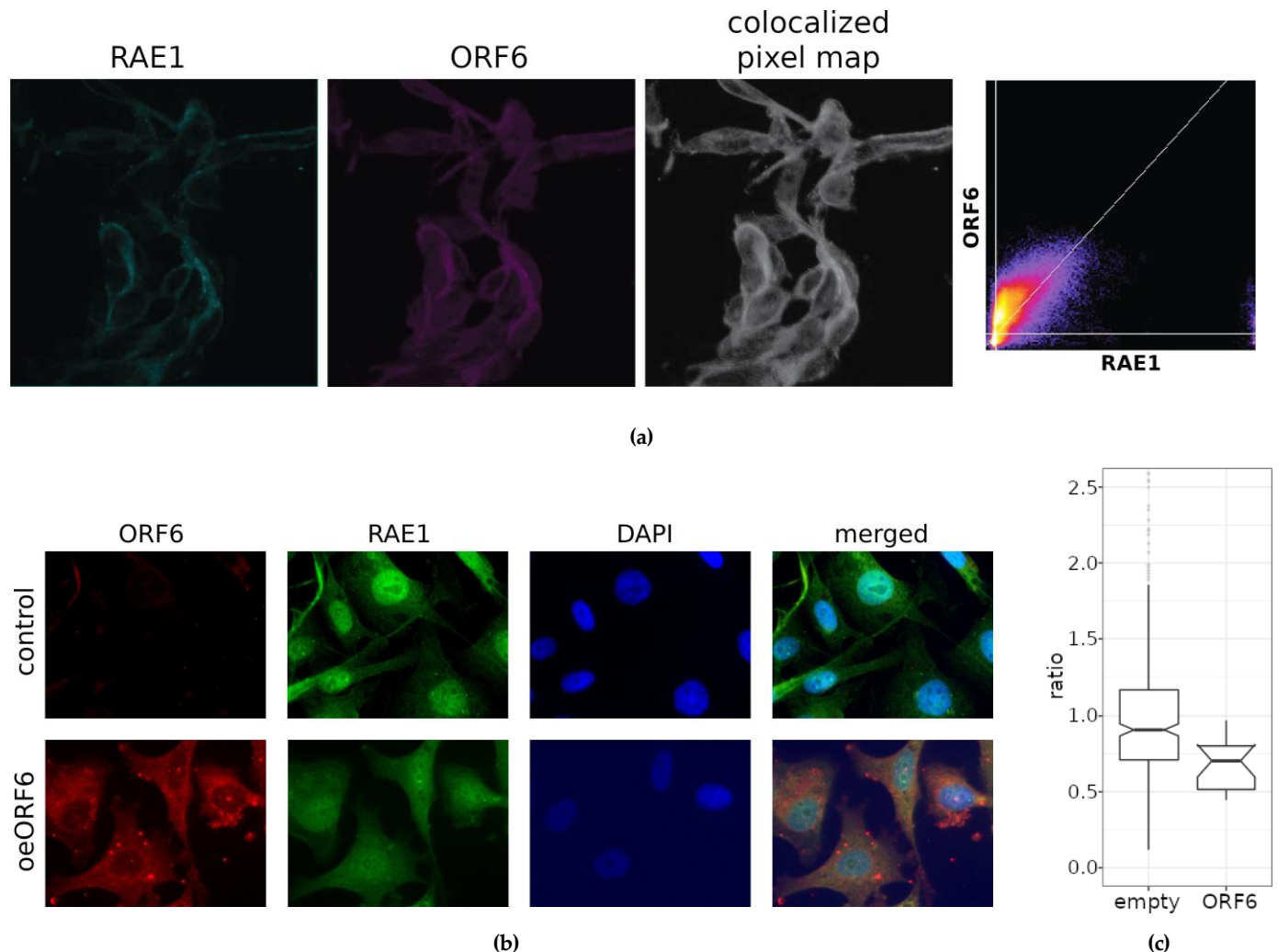
**Figure 3.** Comparison between the RAE1-ORF6 binding interfaces, (a) and (b) obtained in the last 100 ns of the simulation in yellow surface representation, and (c) and (d) experimentally in PDB IDs 7VPH and 7F60 in green surface representation. The RAE1 protein is shown in grey cartoon, and the shared binding interface is represented in cyan wireframe surface.



**Figure 4.** Contact maps in the last 100 ns of RAE1-ORF6 binding simulation. A contact is considered present, if two heavy atoms from the two binding partners are within a cut-off of 6 Å.

#### 2.4. Co-Localization Between RAE1 and ORF6

To test experimentally our model, we transfected PC3 and WISH cell lines with a plasmid, constitutively expressing ORF6. Confocal microscopy (Fig. 5a and Suppl. Fig.14), indicated the formation of the ORF6-RAE1 protein complex in both cell lines, which changes RAE1 localization from mainly nuclear to predominantly cytoplasmic.



**Figure 5.** ORF6 colocalizes with RAE1 and alters its subcellular localization. (a) Confocal images of cells transfected with the ORF6-expressing plasmid stained with antibodies against RAE1 and ORF6. The colocalized pixel map and colocalization scatterplot are given to the right ( $R=0.57$ ). (b) Control (transfected with empty plasmid) and ORF6-overexpressing cells (oeORF6) were stained with antibodies against RAE1 and ORF6. Representative images are shown. (c) Nucleus to cytoplasm ratio of cells from (b) was calculated after measuring fluorescent intensity of RAE1 staining using CellProfiler software. The difference is statistically significant with  $p$ -value  $< 0.0001$ , estimated using Student's test for two independent samples (two-tailed unpaired student test).

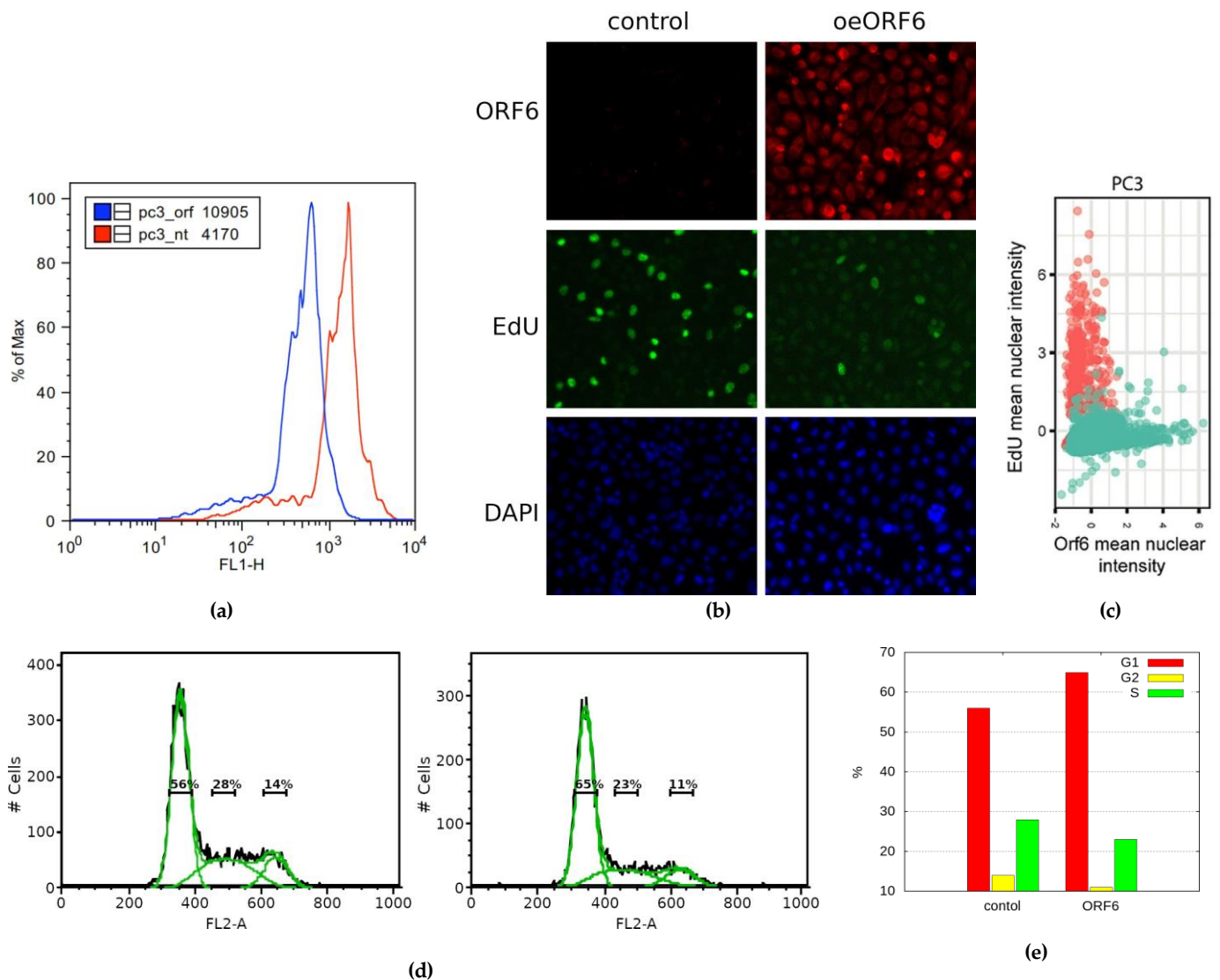
The analysis of the fluorescence intensity of RAE1 (Fig. 5b and c) in the nuclei and the cytoplasm of control and transfected cells showed that the nuclei/cytoplasm ratio of the RAE1 signal is significantly reduced in cells overexpressing ORF6. This supports the hypothesis that alters the cellular localization of RAE1 by anchoring it in the cytoplasm.

### 2.5. Overexpression of SARS-CoV-2 ORF6 Inhibits Cell Cycle Progression

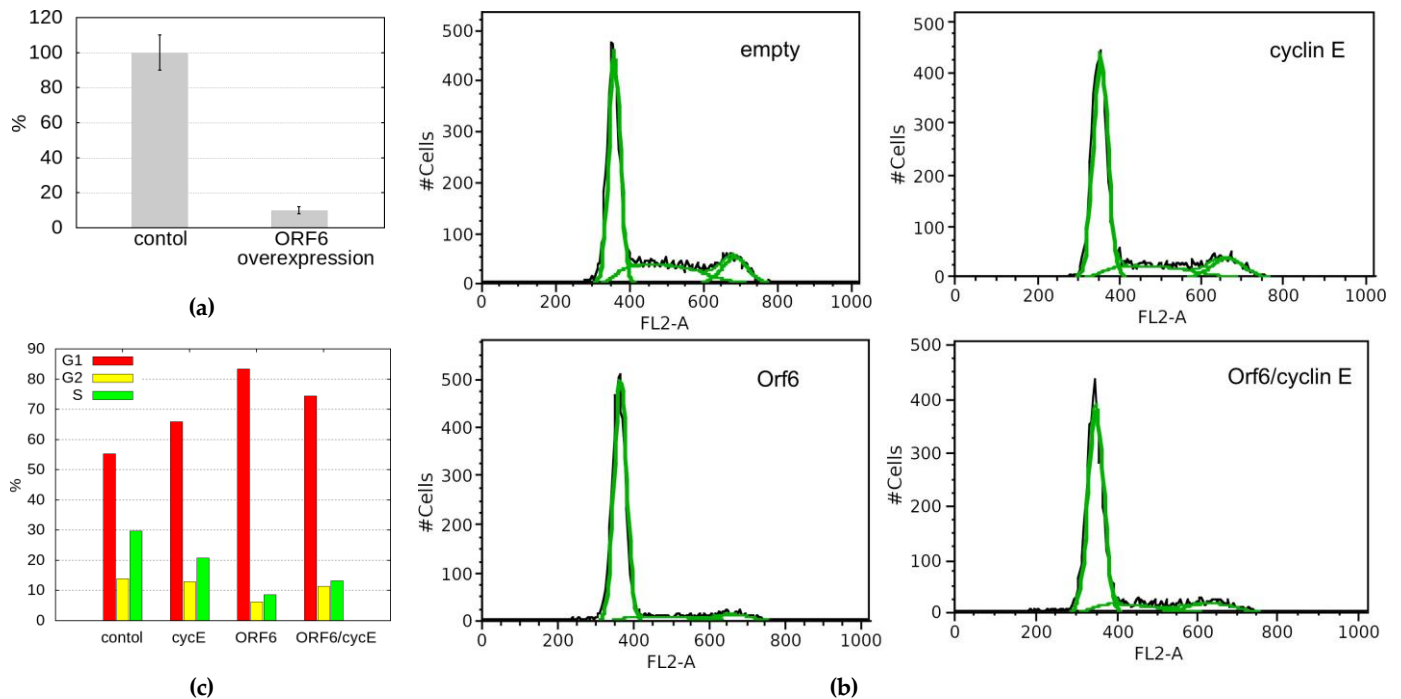
ORF6 was shown to be the most toxic SARS-CoV-2 protein [12]. However, the mechanism of its toxicity is unknown yet because the findings were relying on an assay that quantifies metabolically active cells [12]. To shed light on the processes underlying ORF6 toxicity, we studied the effect of ORF6 overexpression on DNA replication. To this end, ORF6-overexpressing and control cells were labelled with 5-Ethynyl-2'-deoxyuridine (EdU), a thymidine analogue, incorporated into DNA during replication [22]. After staining the incorporated EdU via "click" chemistry and subjecting cells to flow cytometry, we observed a leftward shift of the peak of the ORF6-expressing cells (Fig. 6a). This indicates a weaker

incorporation of the thymidine analogue in these cells. To further confirm that ORF6 causes replication impairment, we co-stained the EdU labelled control and ORF6-expressing cells with AlexaFluor488-azide and an antibody against ORF6 (Fig. 6b). The analysis of the mean immunofluorescence intensity of EdU and ORF6 clearly showed that the level of incorporation of EdU anticorrelates with the levels of ORF6 (Fig. 6c). This suggests that ORF6 impairs DNA replication.

To further investigate the observed effect, we analysed the cell cycle of ORF6 overexpressing cells by flow cytometry after propidium iodide (PI) staining. The cell cycle profiles of the transfected cells (Fig. 6d and e) indicated a significant decrease in the percentage of cells in the S- and the G2-phases of the cell cycle. This is an indication that ORF6 impedes S-phase entry.



**Figure 6.** ORF6 overexpression inhibits cell cycle progression. (a) Control and ORF6-expressing cells were labelled with EdU and stained with AlexaFluor488-azide by “click”-reaction. Fluorescence staining of cells was assessed by flow cytometry. (b) Control and ORF6-expressing cells labelled with EdU were co-stained with an antibody against ORF6 and AlexaFluor488-azide. Representative images are shown. (c) Dot plot to compare fluorescence intensities of EdU and ORF6 for each cell. Red dots- control cells; Green dots – ORF6 overexpressing population. (d) Cell cycle profiles of control and ORF6-expressing cells. (e) Cell cycle distribution of cells from (d).



**Figure 7.** ORF6 overexpression reduces cyclin E levels and S-phase cells. (a) Quantitative RT-PCR analysis of the levels of cyclin E in control and ORF6 overexpressing cells (ORF6 OE). (b) Cell cycle profiles of control cells (transfected with empty vector) and cells transfected with plasmids expressing ORF6, cyclin E or both simultaneously. To maintain the same expression in single and double plasmid transfections, in cyclin E and ORF6 single transfections, half of the plasmid was empty. (c) Cell cycle distribution of cells from (b).

### 2.6. ORF6 Overexpression Inhibits Proliferation by Decreasing Cyclin E Level

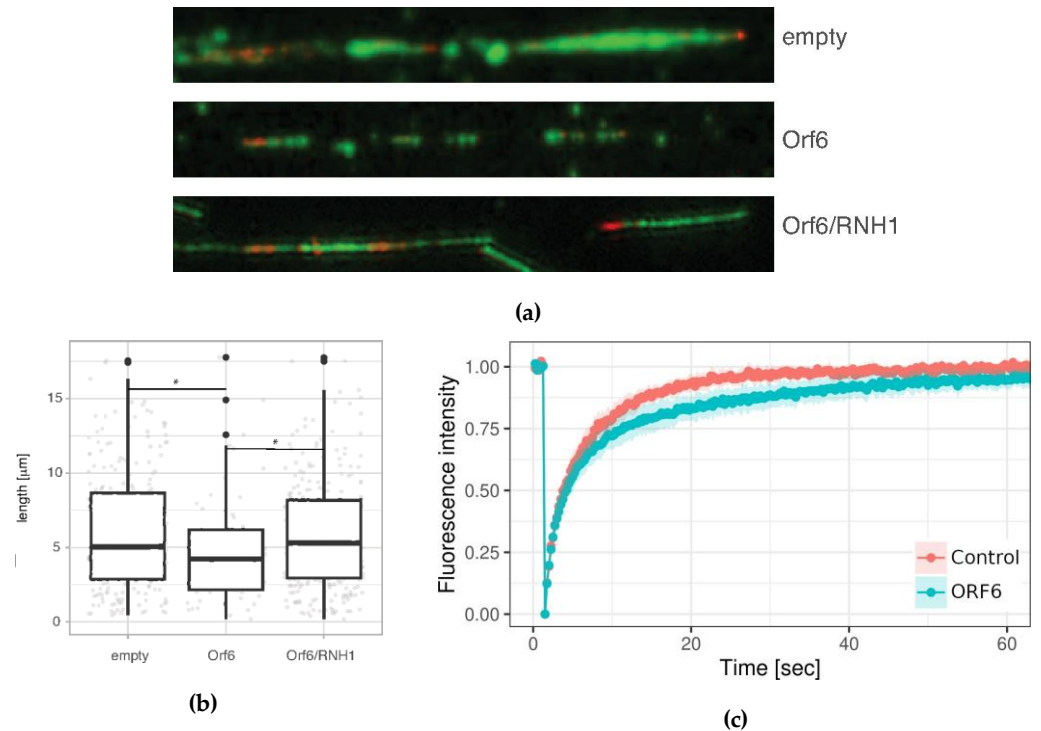
The progression of cells through the different stages of the cell cycle is regulated by a family of regulatory proteins, called cyclins. Cyclin E is a limiting factor for G1 phase progression and S phase entry [23]. It has been shown that *Drosophila melanogaster* RAE1 is required for normal proliferation. Depletion of dmRAE1 inhibited the progression through the G1 phase of the cell cycle and cells failed to enter S-phase due to impaired cyclin E expression [24]. As ORF6 forms a complex with RAE1, it could conceivably, sequester the latter, mimicking RAE1 depletion, reduce the level of cyclin E and prevent cells from entering S-phase. To test this hypothesis, we performed qRT-PCR to monitor the cyclin E levels in cells transfected with ORF6. As a result, we observed a tenfold decrease in the mRNA levels of cyclin E in cells overexpressing ORF6 (Fig. 7a), which could explain the depletion of cells in S-phase.

To functionally test if cyclin E depletion in causes the cell cycle defect ORF6-expressing cells, we checked if it could be rescued by cyclin E overexpression. Cell cycle analysis of cells co-transfected with cyclin E and ORF6 showed that cyclin E expression partially unblocked cell cycle progression (Fig. 7b and c). This indicates that ORF6-induced sequestration of RAE1 and consequent cyclin E decrease significantly contributes to the defective proliferation of cells expressing ORF6.

### 2.7. ORF6 Overexpression Causes Accumulation of R-loops, Which Impedes Progression of Active Replication Forks

Next, we investigated the effect of ORF6 overexpression on the progression of individual replication forks. To this end, 24 hours after transfection with the ORF6-expressing plasmid, cells were subjected to DNA fibre labelling analysis measuring the lengths of second label (green) segments of red-green tracks. The data (Fig. 8a and b) indicated that





**Figure 8.** ORF6 reduces replication fork rate as a consequence of R-loop accumulation. (a) DNA fibres from control (transfected with empty plasmid), ORF6-expressing plasmid and a plasmid expressing ORF6 and RNaseH1 simultaneously. (b) Replication rates cells from (a), The difference is statistically significant (\*) with p-value < 0.05, estimated using Student's test for two independent samples (two-tailed unpaired student test). (c) FRAP analysis of DsRed-RLB in control and ORF6 overexpressing cells.

individual replication fork rates were reduced in cells expressing ORF6. Reduced rate of forks could be explained by impediments that stall their progression.

Defects in mRNA export lead to transcription-dependent genomic instability in yeast and metazoans [25–27] as a consequence of the accumulation of cotranscriptional R loops which interfere with DNA replication [28], ultimately leading to genomic instability [29–31].

Since the interaction of ORF6 with RAE1 disrupts the mRNA transport from the nucleus to the cytoplasm, we asked whether R-loops could be the cause of the DNA replication defect in ORF6-expressing cells. To functionally test this hypothesis, we co-transfected cells with the ORF6-expressing plasmid and a plasmid overexpressing the endonuclease RNase H1, which specifically targets and removes RNA:DNA hybrids from the genome. The assessment of the DNA synthesis rates evaluated using DNA fibre labelling indicated that RNaseH1 rescued fork rates in replicating cells (Fig. 8a and b). This suggests that R-loops are the likely cause of replication impediments in ORF6 expressing cells. To confirm that ORF6 overexpression indeed causes R-loop accumulation, we took advantage of a construct expressing the RNA Binding Domain of RNase H1 fused to DsRed (RBD-DsRed). The binding of this artificial protein to R-loops allows assessment of their abundance in living cells by FRAP analysis. The result (Fig. 8c) indicated that the rate of fluorescence recovery was lower in cells expressing ORF6. The reduction of the mobile fraction of RBD-DsRed observed in these cells indicated that they contain more R-loops than the controls.

### 3. Discussion

SARS-CoV-2 ORF6 is an excellent target for drug design aimed at reducing the pathogenic effects of the virus. Modern rational drug design approaches require detailed knowledge of the structure and interactions of the target protein. However, ORF6 remains

one of the few SARS-CoV-2 proteins that still lack experimentally resolved 3D structures [2]. Currently, there are two entries in the protein data bank (PDB IDs 7VPH and 7F60) that provide experimental data on the interaction of the last dozen of C-terminal residues of ORF6 and its target protein, the mRNA export factor RAE1 [9,19].

Existing data show that ORF6 localizes on cytoplasmic membranes (ER, GA, autophagosomes, and lysosomes) [6,12,18]. Here, we present a full-length structural model of the SARS-CoV-2 ORF6 protein, embedded in a model ER membrane. While the C-terminus of the accessory protein comes out of the lipid bilayer and moves around fairly freely into the solvent, the N-terminal part of the molecule remains quite stably integrated into the membrane. Our results strongly support the hypothesis that ORF6 is a membrane-bound rather than a membrane-peripheral protein [18].

Numerous studies demonstrated that the protein interferes with the RAE1-NUP98 complex, a component of the nuclear pore complex [8,9,16,17], altering the localization of RAE1, which, consequently impairs mRNA transport [17,32]. Hall et al. [33] showed that ORF6 does not interact directly with NUP98 and that mRNA export blockage is dependent only on interaction with RAE1. It is noteworthy that other viruses, such as the vesicular stomatitis virus (VSV) and the herpesvirus, employ the same strategy to disrupt nuclear mRNA export by targeting RAE1 [21,34].

Here, the formation of a stable ORF6–RAE1 protein complex was confirmed both *in silico* and *in vitro* by experiments in transfected PC3 and WISH cell lines. Notably, this process caused the RAE1 to localize predominantly in the cytoplasm, instead of in the nucleus.

A model system, containing four ORF6 molecules, embedded in an ER membrane, with one RAE1 protein placed in close proximity to them was used to elucidate the molecular basis of the SARS-CoV-2 ORF6–RAE1 binding and the changes in RAE1 subcellular localization by means of MD simulations. They revealed that RAE1 immediately engages with at least one ORF6 protein. The contact surface of this interaction coincides to a great extent with the binding interface of RAE1 to the M-protein of the VSV [21] (see Suppl. Fig. S15). A similar binding mode had been observed for the herpesvirus ORF10 protein and mouse RAE1 [34]. This surface groove on the side of its beta-propeller was found to actually be the RNA binding site of the RAE1 molecule [20,21]. Importantly, Met<sup>58</sup> in the C-terminus of ORF6-chain<sub>1</sub> interacts with residues Phe<sup>255</sup>, Phe<sup>257</sup>, Trp<sup>300</sup> and Arg<sup>305</sup> that lie in a deep side pocket of the RAE1 beta-propeller in a manner, similar to VSV M-protein. The M-cavity formed by these residues selects strictly for a methionine residue and favours binding of sequences containing a methionine, flanked by acidic residues that form salt-bridges with positively charged residues (lysine and arginine) in the RNA-binding groove of RAE1 [19]. The same contact interface (239-240, 242, 254-259, 261, 271, 300, 305-312) was experimentally found to be the binding site for both the SARS-CoV ORF6 and SARS-CoV-2 ORF6 C-terminal peptides [9,19]. Thus our modelling results are in excellent agreement with the available experimental data.

In addition to the first ORF6, in our simulation system, RAE1 also interacts with two more ORF6 molecules. The second molecule binds the transport protein at the entrance of the central tunnel of the RAE1 beta-propeller that is opposite to the NUP98 binding site. Occasionally the RAE1 also forms some transient contacts with a third ORF6 C-terminus involving residues in its N-terminus that is positioned in close proximity to the ER membrane.

Our *in vitro* data fully supports the observations from the MD simulations and they clearly show the co-localization of ORF6 and RAE1, which leads to significant depletion of RAE1 in the nuclei of cells transfected with ORF6. Moreover, our data suggests that ORF6 interacts with RAE1 in the cytoplasm fixing it at the surface of the corresponding membrane (ER or GA) thus preventing RAE1 from binding to mRNA and possibly to NUP98 at the nuclear pores.

It is very probable that the toxicity of ORF6 – the highest of all SARS-CoV2 proteins with a reduction in cell viability of over 50% [12] – is rooted in the observed sequestering

of RAE1 into the cytoplasmic membranes. RAE1 sequestration also underlies the effects we observed on cell cycle progression and replication fork speeds contributing to ORF6 cytotoxicity. ORF6 overexpression impairs DNA replication, as evidenced by the weaker incorporation of the thymidine analogue EdU and the correlation between EdU incorporation and ORF6 levels. The cell cycle analysis reveals that ORF6 hinders cell cycle progression, with a significant decrease in the percentage of cells in the S- and G2-phases, suggesting that ORF6 interferes with S-phase entry.

A possible explanation for ORF6-mediated cell cycle defects lies in its effect on cyclin E levels. As cyclin E is a key regulator of G1 phase progression and S phase entry, the observed tenfold decrease in mRNA levels of cyclin E in ORF6-overexpressing cells could be responsible for the observed depletion of cells in the S-phase. The partial rescue of cell cycle progression upon co-transfection of cyclin E and ORF6 supports the hypothesis that ORF6-induced sequestration of RAE1 and the resulting decrease in cyclin E underlie the defective proliferation of cells expressing ORF6. These data are in line with the report by Sitterlin [24] that the presence of dmRAE1 is required for normal proliferation and, more importantly, for normal cyclin E expression, suggesting that the human homologue, hsRAE1, may also play a similar role during the cell cycle.

Impaired nuclear export has been linked to accumulation of R-loops [35–37], which could stall DNA replication forks [38–40]. The obstruction of RNA export is likely the cause of R-loops accumulation, known to cause replication stress, which we observe in FRAP experiments as a lower mobile fraction of RBD-DsRed in ORF6-expressing cells. The link between R-loops accumulated in ORF6-expressing cells and impeded replication is provided by the observation that RNaseH1 overexpression rescued replication fork rates. A similar effect has been described in which Kaposi's sarcoma herpesvirus sequesters transcription and export factor TREX, leading to R-loops and genome instability [41].

Thus, our study provides evidence for two ways in which SARS-CoV-2 ORF6 protein could affect cellular proliferation: by inhibiting cell cycle progression, and by accumulation of R-loops. A very recent paper has demonstrated that SARS-CoV-2 proteins ORF6 and Nsp13 cause degradation of the DNA damage response kinase CHK1 through proteasome and autophagy, respectively. CHK1 loss leads to deoxynucleoside triphosphate shortage, causing impaired S-phase progression, DNA damage, pro-inflammatory pathways activation and cellular senescence [42]. Using DNA fibre labelling analysis we observed that the supplementation of deoxynucleoside triphosphates rescued the progression of replication forks in cells overexpressing ORF6 (data not shown) thus confirming the quoted above research. These and reported here results indicate the broad range of mechanisms by which SARS-CoV-2 influences the cell proliferation and genome integrity.

## 4. Materials and Methods

### 4.1. MD Simulation Protocol

All MD simulations were performed using the GROMACS simulation package [43], version 2019.6 and later. The CHARMM36 force field was used for parameterization of the proteins and lipids in the membrane [44,45]. All simulations were performed in explicit water solvent with a concentration of NaCl of 0.15 mol/l. The energy of the systems was minimised using the steepest descent algorithm with a maximum force tolerance of 100 kJ/(mol.nm). Then, a short 50 ps position-restraint simulations was done to equilibrate the solvent, followed by a 10 ns isothermal-isobaric simulation, equilibrating the temperature at 310 K and the pressure at 1 atm using the Berendsen thermo- and barostat [46]. For production simulations in the NTP ensemble, temperature was maintained constant at 310 K using v-rescale thermostat [47] with a coupling constant of 0.2 ps, and the pressure – at 1 atm, using the Parrinello-Rahman barostat [48] with a coupling constant of 2 ps. The leap-frog integrator was used with a timestep of 2 fs, allowed for by constraining bonds between heavy atoms and hydrogens with the PLINCS algorithm [49]. The PME method with a direct cut-off of 1.2 nm was used for calculations of the electrostatic interactions. Van

der Waals forces were smoothly switched off from a distance of 1.0 nm and truncated at 1.2 nm. Neighbourlists were updated every 10 timesteps.

#### 4.2. Input Structural Models

##### 4.2.1. SARS-CoV-2 ORF6 Protein

The 3D structure of the SARS-CoV-2 ORF6 has not yet been resolved. Therefore, we used as input a model structure of the protein, developed by the EXCALATE4COV project and available in their open access repository [50]. This model was developed using 10  $\mu$ s long MD folding simulations, starting from a molten globule model conformation that consisted of three  $\alpha$ -helical segments. We further performed additional 7.6  $\mu$ s of folding simulation of this structure using the simulation protocol, described in sec. 4.1. The protein was solvated in a cubic box with a 1.2 nm minimal distance to the box walls. Trajectory frames were written every nanosecond.

##### 4.2.2. SARS-CoV-2 ORF6 Embedded in a Model ER Membrane

As an input structure for the development of this model, the centroid of the largest cluster of the trajectory from sec. 4.2 (with a cut-off of 3.8  $\text{\AA}$ ) was embedded in a lipid bilayer with the membrane composed of 54% POPC, 20% POPE, 11% POPI and 8% cholesterol, as to model the content of endoplasmic reticulum membranes, according to [51] and references therein. For the lipid positions generation and ORF6 insertion in the bilayer the Membrane/Bilayer Builder module [52] of the CHARMM-GUI server [53] was used. The lipid bilayer had an area of  $76.3 \times 76.3 \text{ \AA}^2$  and contained 192 molecules in total. A 1.4  $\mu$ s long production MD run was performed, with trajectory frames put out every nanosecond.

#### 4.3. Modelling the Interaction of SARS-CoV-2 ORF6 and RAE1

The obtained stable structure of the ORF6 protein in a model ER membrane was duplicated in both directions in the plane of the lipid bilayer, which resulted in a membrane of size  $150 \times 150 \text{ \AA}^2$  containing a total of 769 lipids and four inserted into them ORF6 molecules. As an input structure for the RAE1 protein we used chain A of the crystallographic structure with PDB ID 3MMY [20]. The missing loop residues (aa 19-22 and 264-267) were reconstructed using the macromolecular crystallography and model-building toolkit COOT [54]. This RAE1 initial model was simulated for 100 ns. The centroid of the largest cluster of the equilibration trajectory was placed at a distance of about 7  $\text{\AA}$  above the ER membrane between the fluttering C-termini of the four ORF6 molecules. A 630 ns production MD simulation was carried out, in order to study the interaction of the RNA carrier protein and the viral accessory proteins.

#### 4.4. MD Data Analysis

The MD trajectories were postprocessed and analysed using the standard GROMACS postprocessing and analysis tools for calculations of RMSD, RMSF, SASA, Rg, and cluster analysis. Least-square fitting to the initial conformation was performed prior to all analyses in order to remove all global translational and rotational movements. The gromos algorithm [55] was used for cluster analysis. Secondary structure for the proteins was assigned by the STRIDE algorithm [56] as implemented in the visualisation and manipulation package VMD [57]. All structural figures were also generated with VMD.

#### 4.5. Cell culture, Plasmids, and Antibodies

WISH (ATCC®CCL-25™, ATCC, Manassas, VA, USA) and PC3 (ATCC®CRL-1435™) cell lines were propagated in the MEM and DMEM, respectively, supplemented with 10% fetal bovine serum (Gibco™, Waltham, MA, USA) and penicillin-streptomycin (10 000 U/mL, Gibco™, Waltham, MA, USA). Cells were cultured in a humidified incubator at 37°C and 5% CO<sub>2</sub>.



For the overexpression of ORF6 SARS-CoV-2 protein, the plasmid pLVX-EF1alpha-SARS-CoV-2-orf6-2xStrep-IRES-Puro (Addgene plasmid # 141387 from Nevan Krogan) [15] was used. The cyclin E-expressing vector was a gift from Bob Weinberg (Addgene # 8963).

The RBD-DsRed plasmid was constructed by PCR cloning the HB domain of RNase H1 into the pDsRed-Express-C1 vector (Clontech) as earlier described [58].

#### 4.6. Quantitative Real-Time PCR Analysis

Total RNA from cells, overexpressing ORF6, was extracted by using TRIzol reagent (Invitrogen, Carlsbad, CA). The concentration and purity of extracted RNA were determined by Nanodrop-1000 (Thermo Fisher). RNA integrity and quality was assessed using 1% agarose gel electrophoresis in TAE buffer (40 mM Tris-acetate, 1 mM EDTA). Subsequently, 1 µg total RNA from each sample was subjected to cDNA synthesis by RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific™) according to the manufacturer's recommendations. Relative expression levels of target genes (RAE1 and Cyclin E) were assessed by qRT-PCR analysis using the SYBR™ Select Master Mix (Thermo Scientific™). Housekeeping gene  $\beta$ -actin was used as an endogenous control to normalize genes expression.

Primer oligonucleotide sequences of the studied genes are listed in Suppl. Table 1. The analysis was performed on Rotor-Gene 6000 thermal cycler (Corbett, QIAGEN). Gene expression data were analyzed using Rotor-Gene 6000 Software (QIAGEN) and the relative expression levels of the genes of interest were normalized to the endogenous control for each sample. Each qRT-PCR reaction was performed in at least three replicates in different PCR runs. Statistical differences were evaluated using t-test and values at  $< 0.05$  were considered as statistically significant.

#### 4.7. 5-Ethynyl-2'-deoxyuridine (EdU) Labelling

Cells were incubated with 25 µM EdU for 20 minutes immediately before fixation. "Click" reaction was carried out using Click-iT™ EdU AlexaFluor™ 488 kit according to the manufacturer's protocol (Thermo Scientific™). When combined with immunostaining, "click" reaction was performed immediately after the secondary antibody incubation.

#### 4.8. Flow Cytometry

To analyse cell cycle profiles, approximately  $5 \times 10^5$  cells were trypsinized, harvested by centrifugation for 10 min at 400 g and fixed in 70% ethanol. Before analysis cells were resuspended in 1×PBS, treated with RNase A (20 µg/ml) and stained with propidium iodide (20 µg/ml). Analysis was carried out by a FACScalibur apparatus with Cellquest (Becton Dickinson) and FlowJo software.

#### 4.9. Immunofluorescent Microscopy

WISH and PC3 cells were cultured on Ø 12 mm coverslips (EpreDia) and transfected with plasmids, relevant to the experiment being performed. Cells were fixed with either 3,7% formaldehyde in 1×PBS for 10 min at room temperature or with methanol for 10 min at -20°C. Fixed cells were permeabilised with 0,5% Triton X-100 in PBS for 5 min and then blocked in blocking buffer (5% BSA and 0,1% Tween 20 in 1×PBS) for 1 h. Cells were then incubated overnight at 4°C with primary antibody (in blocking buffer). After washes cells were stained with a secondary antibody for 1 h at room temperature.

The nuclei were counterstained with DAPI (Cell Signaling Technology) to final concentration of 0,5 µg/ml for 2 min at room temperature. The coverslips were washed and mounted using ProLong™ Gold Antifade mounting media (Invitrogen). The cells were imaged with Zeiss Axiovert 200M fluorescence inverted microscope and images were analyzed by CellProfiler software [59].

#### 4.10. Fluorescence Recovery After Photobleaching (FRAP) Analysis

Cells were transfected with the RBD-DsRed plasmid expressing the RNA Binding Domain of RNase H1 fused to DsRed. FRAP analysis was carried out using Andor Revolution XDI spinning disk confocal system, with a heated chamber in CO<sub>2</sub>-independent medium. Imaging was carried out in 1 s intervals for 150 s with the bleaching pulse applied at the fifth second. Intensity measurements and analysis were carried out with the CellTool software [60].

#### 4.11. DNA Fibre Labeling

DNA fibre analyses were performed following standard protocol [61] with slight modifications. Briefly, exponentially growing PC3 cells were first incubated with 25 µM chlorodeoxyuridine (CldU) for 10 min and then with 250 µM iododeoxyuridine (IdU) for 25 min both at 37°C and 5% CO<sub>2</sub>. Spreads were prepared from 2500 cells, suspended in 1×PBS at 1 × 10<sup>6</sup> cells/ml. Cell lysis was carried out in fibre lysis solution (50 mM EDTA and 0.5% SDS in 200 mM Tris-HCl, pH 7.5). DNA fibres were spread by tilting the slides 25° until the drop of the fibre solution reached the bottom of the slide and letting it dry. Dried slides were immersed in 2.5 M HCl for 80 min, washed in PBS, and blocked for 40 min in 5% BSA in 1×PBS. Primary antibodies — mouse anti-BrdU antibody (Becton Dickinson, cat # 347580) to detect IdU and rat anti-BrdU antibody (Abcam cat # Ab6326) to detect CldU — were diluted in blocking buffer and applied overnight. After washes slides were incubated with secondary antibodies goat anti-rat DyLight®594 (Abcam 96889) and goat anti-mouse DyLight®488 (Abcam, 96879) – for 60 min. Slides were mounted with ProLong Gold anti-fade reagent (Invitrogene). Images were acquired with Axiovert 200 M microscope (Carl Zeiss) equipped with an AxioCam MR3 camera (Carl Zeiss). Fibre length measurements were carried out using the DNA size finder software [62].

## 5. Conclusions

In this report we studied the interaction of the SARS-CoV-2 ORF6 and the RAE1 proteins. Our *in silico* data demonstrate that RAE1 is able to bind simultaneously multiple C-termini of ORF6 molecules. These interactions anchor the transport protein to the ER-membrane thus sequestering it in the cytoplasm of the host cell and depleting its availability in the nucleus, disrupting nucleocytoplasmic trafficking. These results were confirmed experimentally by the observed formation of the ORF6–RAE1 protein complex and the change in RAE1 localization from mainly nuclear to predominantly cytoplasmic. Our results suggest for the first time a mechanisms by which the interaction of ORF6 with RAE1 leads to genome instability. Firstly, this complex formation interferes with the cell cycle progression. Secondly, R-loops accumulate due to deficient mRNA transportation, leading to stalled DNA replication forks and eventually causing replication stress. Both of these effects are partially revisable by overexpression of either cyclin E, which helps the progression of the cell from G1 to S phase, or RNase H1 which has the ability to remove RNA from the RNA:DNA hybrids.

This action of the SARS-CoV-2 ORF6 aims at hindering the synthesis of antiviral molecules and slowing down the immune response of the host cells, but does not affect significantly the viral replication, since coronaviruses do not rely on nuclear export for their replication and transcribe their genome in the cytoplasm. Therefore, targeting the interaction of the SARS-CoV-2 ORF6 C-terminus and RAE1 can potentially lessen the pathogenic effects of the virus, but also facilitate earlier antiviral response, thus reducing viral replication in infected host cells. Understanding the multiple ways in which ORF6 affects DNA replication might also have important implications in clarifying the pathogenesis of SARS-CoV-2 and for the development of therapeutic strategies to counteract its deleterious effects on host cells.

**Author Contributions:** Conceptualization, L.L., G.N. and A.G.; methodology, L.L., P.P., E.K. and M.R.; software, P.P. and M.R.; validation, E.L., N.T. and R.H.; formal analysis, L.L., P.P., N.I., E.L., G.N., E.K.

and A.G.; investigation, L.L., P.P., N.I., E.L., M.R., N.T., G.N., E.K., K.M., R.H. and A.G.; resources, L.L., N.I., G.N. and A.G.; data curation, E.K., E.L., R.H.; writing—original draft preparation, E.K. E.L. A.G.; writing—contributing parts to the original draft, P.P., M.R., N.T., K.M., R.H.; writing—review and editing, L.L., N.I., G.N. and A.G.; visualization, E.L., E.K. and A.G.; supervision, L.L., N.I. and G.N.; project administration, N.I.; funding acquisition, L.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially funded by the Bulgarian National Science Fund under Grant KP-06-DK1/5/2021 SARSIMM.

**Acknowledgments:** Computational resources were provided by the Discoverer supercomputer thanks to Discoverer PetaSC and EuroHPC JU, as well as the BioSim HPC cluster at the Faculty of Physics, Sofia University “St. Kliment Ohridski”.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

aa	amino acids
CHOL	cholesterol
ER	endoplasmic reticulum
GA	Golgi apparatus
MD	molecular dynamics
NUP98	nucleoporin 98
oeORF6	ORF6-overexpressing cells
ORF6	Open reading frame 6
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POPI	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol
RAE1	ribonucleic acid export 1
$R_g$	radius of gyration
RMSD	root mean square deviation
RMSF	root mean square fluctuations
SASA	solvent accessible surface area

## References

1. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *Nature*, 265–269. <https://doi.org/10.1038/s41586-020-2008-3>.
2. Bai, C.; Zhong, Q.; Gao, G. Overview of SARS-CoV-2 genome-encoded proteins. *Science China Life Sciences* **2022**, *65*, 280–294. <https://doi.org/10.1007/s11427-021-1964-4>.
3. Zandi, M.; Shafaati, M.; Kalantar-Neyestanaki, D.; Pourghadamyari, H.; Fani, M.; Soltani, S.; Kaleji, H.; Abbasi, S. The role of SARS-CoV-2 accessory proteins in immune evasion. *Biomedicine & Pharmacotherapy* **2022**, *156*, 113889. <https://doi.org/10.1016/j.biopha.2022.113889>.
4. Redondo, N.; Zaldívar-López, S.; Garrido, J.J.; Montoya, M. SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. *Frontiers in Immunology* **2021**, *12*. <https://doi.org/10.3389/fimmu.2021.708264>.
5. Kimura, I.; Konno, Y.; Uriu, K.; Hopfensperger, K.; Sauter, D.; Nakagawa, S.; Sato, K. Sarbecovirus ORF6 proteins hamper induction of interferon signaling. *Cell Reports* **2021**, *34*, 108916. <https://doi.org/10.1016/j.celrep.2021.108916>.
6. Frieman, M.; Yount, B.; Heise, M.; Kopecky-Bromberg, S.A.; Palese, P.; Baric, R.S. Severe Acute Respiratory Syndrome Coronavirus ORF6 Antagonizes STAT1 Function by Sequestering Nuclear Import Factors on the Rough Endoplasmic Reticulum/Golgi Membrane. *Journal of Virology* **2007**, *81*, 9812–9824. <https://doi.org/10.1128/JVI.01012-07>.
7. Kopecky-Bromberg, S.A.; Martínez-Sobrido, L.; Frieman, M.; Baric, R.A.; Palese, P. Severe Acute Respiratory Syndrome Coronavirus Open Reading Frame (ORF) 3b, ORF 6, and Nucleocapsid Proteins Function as Interferon Antagonists. *Journal of Virology* **2007**, *81*, 548–557. <https://doi.org/10.1128/JVI.01782-06>.
8. Miorin, L.; Kehrer, T.; Sanchez-Aparicio, M.T.; Zhang, K.; Cohen, P.; Patel, R.S.; Cupic, A.; Makio, T.; Mei, M.; Moreno, E.; et al. SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. *Proceedings of the National Academy of Sciences* **2020**, *117*, 28344–28354. <https://doi.org/10.1073/pnas.2016650117>.

9. Li, T.; Wen, Y.; Guo, H.; Yang, T.; Yang, H.; Ji, X. Molecular Mechanism of SARS-CoVs Orf6 Targeting the Rae1–Nup98 Complex to Compete With mRNA Nuclear Export. *Frontiers in Molecular Biosciences* **2022**, *8*. <https://doi.org/10.3389/fmolb.2021.813248>.
10. Yuen, C.K.; Lam, J.Y.; Wong, W.M.; Mak, L.F.; Wang, X.; Chu, H.; Cai, J.P.; Jin, D.Y.; To, K.K.W.; Chan, J.F.W.; et al. SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerging Microbes & Infections* **2020**, *9*, 1418–1428. <https://doi.org/10.1080/22221751.2020.1780953>.
11. Miyamoto, Y.; Itoh, Y.; Suzuki, T.; Tanaka, T.; Sakai, Y.; Koido, M.; Hata, C.; Wang, C.X.; Otani, M.; Moriishi, K.; et al. SARS-CoV-2 ORF6 disrupts nucleocytoplasmic trafficking to advance viral replication. *Communications Biology* **2022**, *5*, 483. <https://doi.org/10.1038/s42003-022-03427-4>.
12. Lee, J.G.; Huang, W.; Lee, H.; van de Leemput, J.; Kane, M.A.; Han, Z. Characterization of SARS-CoV-2 proteins reveals Orf6 pathogenicity, subcellular localization, host interactions and attenuation by Selinexor. *Cell & Bioscience* **2021**, *11*, 58. <https://doi.org/10.1186/s13578-021-00568-7>.
13. Netland, J.; Ferraro, D.; Pewe, L.; Olivares, H.; Gallagher, T.; Perlman, S. Enhancement of Murine Coronavirus Replication by Severe Acute Respiratory Syndrome Coronavirus Protein 6 Requires the N-Terminal Hydrophobic Region but Not C-Terminal Sorting Motifs. *Journal of Virology* **2007**, *81*, 11520–11525. <https://doi.org/10.1128/JVI.01308-07>.
14. Zhao, J.; Falcón, A.; Zhou, H.; Netland, J.; Enjuanes, L.; Breña, P.P.; Perlman, S. Severe Acute Respiratory Syndrome Coronavirus Protein 6 Is Required for Optimal Replication. *Journal of Virology* **2009**, *83*, 2368–2373. <https://doi.org/10.1128/JVI.02371-08>.
15. Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O’Meara, M.J.; Rezelj, V.V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**, *583*, 459–468. <https://doi.org/10.1038/s41586-020-2286-9>.
16. Overexpression of SARS-CoV-2 protein ORF6 dislocates RAE1 and NUP98 from the nuclear pore complex. *Biochemical and Biophysical Research Communications* **2021**, *536*, 59–66. <https://doi.org/10.1016/j.bbrc.2020.11.115>.
17. Addetia, A.; Lieberman, N.A.P.; Phung, Q.; Hsiang, T.Y.; Xie, H.; Roychoudhury, P.; Shrestha, L.; Loprieno, M.A.; Huang, M.L.; Gale, M.; et al. SARS-CoV-2 ORF6 Disrupts Bidirectional Nucleocytoplasmic Transport through Interactions with Rae1 and Nup98. *mBio* **2021**, *12*, e00065–21. <https://doi.org/10.1128/mBio.00065-21>.
18. Wong, H.T.; Cheung, V.; Salamango, D.J. Decoupling SARS-CoV-2 ORF6 localization and interferon antagonism. *Journal of Cell Science* **2022**, *135*. <https://doi.org/10.1242/jcs.259666>.
19. Gao, X.; Tian, H.; Zhu, K.; Li, Q.; Hao, W.; Wang, L.; Qin, B.; Deng, H.; Cui, S. Structural basis for Sarbecovirus ORF6 mediated blockage of nucleocytoplasmic transport. *Nature Communications* **2022**, *13*, 4782. <https://doi.org/10.1038/s41467-022-32489-5>.
20. Ren, Y.; Seo, H.S.; Blobel, G.; Hoelz, A. Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. *Proceedings of the National Academy of Sciences* **2010**, *107*, 10406–10411. <https://doi.org/10.1073/pnas.1005389107>.
21. Quan, B.; Seo, H.S.; Blobel, G.; Ren, Y. Vesiculoviral matrix (M) protein occupies nucleic acid binding site at nucleoporin pair (Rae1•Nup98). *Proceedings of the National Academy of Sciences* **2014**, *111*, 9127–9132. <https://doi.org/10.1073/pnas.1409076111>.
22. Salic, A.; Mitchison, T.J. A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proceedings of the National Academy of Sciences* **2008**, *105*, 2415–2420. <https://doi.org/10.1073/pnas.0712168105>.
23. Sunwoo, H.H.; Suresh, M.R. Chapter 9.13 - Cancer Markers. In *The Immunoassay Handbook (Fourth Edition)*, Fourth Edition ed.; Wild, D., Ed.; Elsevier: Oxford, 2013; pp. 833–856. <https://doi.org/10.1016/B978-0-08-097037-0.00067-1>.
24. Characterization of the Drosophila Rae1 protein as a G1 phase regulator of the cell cycle. *Gene* **2004**, *326*, 107–116. <https://doi.org/10.1016/j.gene.2003.10.024>.
25. Jimeno, S.; Rondón, A.G.; Luna, R.; Aguilera, A. The yeast THO complex and mRNA export factors link RNA metabolism with transcription and genome instability. *The EMBO Journal* **2002**, *21*, 3526–3535. <https://doi.org/10.1093/emboj/cdf335>.
26. Zenklusen, D.; Vinciguerra, P.; Wyss, J.C.; Stutz, F. Stable mRNP Formation and Export Require Cotranscriptional Recruitment of the mRNA Export Factors Yra1p and Sub2p by Hpr1p. *Molecular and Cellular Biology* **2002**, *22*, 8241–8253. <https://doi.org/10.1128/MCB.22.23.8241-8253.2002>.
27. Domínguez-Sánchez, M.S.; Barroso, S.; Gímez-González, B.; Luna, R.; Aguilera, A. Genome Instability and Transcription Elongation Impairment in Human Cells Depleted of THO/TREX. *PLOS Genetics* **2011**, *7*, 1–17. <https://doi.org/10.1371/journal.pgen.1002386>.
28. Gómez-González, B.; García-Rubio, M.; Bermejo, R.; Gaillard, H.; Shirahige, K.; Marín, A.; Foiani, M.; Aguilera, A. Genome-wide function of THO/TREX in active genes prevents R-loop-dependent replication obstacles. *The EMBO Journal* **2011**, *30*, 3106–3119. <https://doi.org/10.1038/emboj.2011.206>.
29. Sollier, J.; Cimprich, K.A. Breaking bad: R-loops and genome integrity. *Trends in Cell Biology* **2015**, *25*, 514–522. <https://doi.org/10.1016/j.tcb.2015.05.003>.
30. Gaillard, H.; Aguilera, A. Transcription as a Threat to Genome Integrity. *Annual Review of Biochemistry* **2016**, *85*, 291–317. <https://doi.org/10.1146/annurev-biochem-060815-014908>.
31. Tuduri, S.; Crabbé, L.; Conti, C.; Tourrière, H.; Holtgreve-Grez, H.; Jauch, A.; Pantescio, V.; De Vos, J.; Thomas, A.; Theillet, C.; et al. Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature Cell Biology* **2009**, *11*, 1315–1324. <https://doi.org/10.1038/ncb1984>.



32. Pritchard, C.E.; Fornerod, M.; Kasper, L.H.; van Deursen, J.M. RAE1 Is a Shuttling mRNA Export Factor That Binds to a GLEBS-like NUP98 Motif at the Nuclear Pore Complex through Multiple Domains. *Journal of Cell Biology* **1999**, *145*, 237–254. <https://doi.org/10.1083/jcb.145.2.237>.
33. Hall, R.; Guedán, A.; Yap, M.W.; Young, G.R.; Harvey, R.; Stoye, J.P.; Bishop, K.N. SARS-CoV-2 ORF6 disrupts innate immune signalling by inhibiting cellular mRNA export. *PLOS Pathogens* **2022**, *18*, 1–24. <https://doi.org/10.1371/journal.ppat.1010349>.
34. Feng, H.; Tian, H.; Wang, Y.; Zhang, Q.; Lin, N.; Liu, S.; Yu, Y.; Deng, H.; Gao, P. Molecular mechanism underlying selective inhibition of mRNA nuclear export by herpesvirus protein ORF10. *Proceedings of the National Academy of Sciences* **2020**, *117*, 26719–26727. <https://doi.org/10.1073/pnas.2007774117>.
35. Skourti-Stathaki, K.; Proudfoot, N.; Gromak, N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination. *Molecular Cell* **2011**, *42*, 794–805. <https://doi.org/10.1016/j.molcel.2011.04.026>.
36. Crossley, M.P.; Bocek, M.; Cimprich, K.A. R-Loops as Cellular Regulators and Genomic Threats. *Molecular Cell* **2019**, *73*, 398–411. <https://doi.org/10.1016/j.molcel.2019.01.024>.
37. Santos-Pereira, J.M.; Aguilera, A. R loops: new modulators of genome dynamics and function. *Nature Reviews Genetics* **2015**, *16*, 583–597. <https://doi.org/10.1038/nrg3961>.
38. Gan, W.; Guan, Z.; Liu, J.; Gui, T.; Shen, K.; Manley, J.L.; Li, X. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes & Development* **2011**, *25*, 2041–56. <https://doi.org/10.1101/gad.17010011>.
39. Helmrich, A.; Ballarino, M.; Tora, L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Molecular Cell* **2011**, *44*, 966–977. <https://doi.org/10.1016/j.molcel.2011.10.013>.
40. Hamperl, S.; Cimprich, K.A. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair* **2014**, *19*, 84–94. Cutting-edge Perspectives in Genomic Maintenance, <https://doi.org/10.1016/j.dnarep.2014.03.023>.
41. Jackson, B.R.; Noerenberg, M.; Whitehouse, A. A Novel Mechanism Inducing Genome Instability in Kaposi's Sarcoma-Associated Herpesvirus Infected Cells. *PLOS Pathogens* **2014**, *10*, 1–15. <https://doi.org/10.1371/journal.ppat.1004098>.
42. Gioia, U.; Tavella, S.; Martínez-Orellana, P.; Cicio, G.; Colliva, A.; Ceccon, M.; Cabrini, M.; Henriques, A.C.; Fumagalli, V.; Paldino, A.; et al. SARS-CoV-2 infection induces DNA damage, through CHK1 degradation and impaired 53BP1 recruitment, and cellular senescence. *Nature Cell Biology* **2023**, *25*, 550–564. <https://doi.org/10.1038/s41556-023-01096-x>.
43. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19 – 25. <https://doi.org/10.1016/j.softx.2015.06.001>.
44. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B.L.; Grubmüller, H.; MacKerell Jr, A.D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature Methods* **2016**, *14*, 71–73. <https://doi.org/10.1038/nmeth.4067>.
45. Klauda, J.B.; Venable, R.M.; Freites, J.A.; O'Connor, J.W.; Tobias, D.J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A.D.J.; Pastor, R.W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B* **2010**, *114*, 7830–7843. <https://doi.org/10.1021/jp101759q>.
46. Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; DiNola, A.; Haak, J.R. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics* **1984**, *81*, 3684. <https://doi.org/10.1063/1.448118>.
47. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **2007**, *126*, 014101. <https://doi.org/10.1063/1.2408420>.
48. Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* **1980**, *45*, 1196. <https://doi.org/10.1103/physrevlett.45.1196>.
49. Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 116–122. <https://doi.org/10.1021/ct700200b>.
50. [https://ds-814.cr.cnaf.infn.it:8443/ex4cov-public/Molecular\\_Dynamics/ORF6/](https://ds-814.cr.cnaf.infn.it:8443/ex4cov-public/Molecular_Dynamics/ORF6/), 2020. Accessed: .....
51. Casares, D.; Escribá, P.V.; Rosselló, C.A. Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *International Journal of Molecular Sciences* **2019**, *20*. <https://doi.org/10.3390/ijms20092167>.
52. Wu, E.L.; Cheng, X.; Jo, S.; Rui, H.; Song, K.C.; Dávila-Contreras, E.M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R.M.; et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *Journal of Computational Chemistry* **2014**, *35*, 1997–2004. <https://doi.org/10.1002/jcc.23702>.
53. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **2008**, *29*, 1859–1865. <https://doi.org/10.1002/jcc.20945>.
54. Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **2004**, *60*, 2126–2132. <https://doi.org/10.1107/S0907444904019158>.
55. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W.F.; Mark, A.E. Peptide Folding: When Simulation Meets Experiment. *Angewandte Chemie International Edition* **1999**, *38*, 236–240. [https://doi.org/10.1002/\(SICI\)1521-3773\(19990115\)38:1/2<236::AID-ANIE236>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1521-3773(19990115)38:1/2<236::AID-ANIE236>3.0.CO;2-M).
56. Frishman, D.; Argos, P. Knowledge-based secondary structure assignment. *Proteins: structure, function and genetics* **1995**, *23*, 566–579. <https://doi.org/10.1002/prot.340230412>.

- 
57. Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
  58. Prendergast, L.; McClurg, U.L.; Hristova, R.; Berlinguer-Palmini, R.; Greener, S.; Veitch, K.; Hernandez, I.; Pasero, P.; Rico, D.; Higgins, J.M.G.; et al. Resolution of R-loops by INO80 promotes DNA replication and maintains cancer cell proliferation and viability. *Nature Communications* **2020**, *11*, 4534. <https://doi.org/10.1038/s41467-020-18306-x>.
  59. Stirling, D.R.; Swain-Bowden, M.J.; Lucas, A.M.; Carpenter, A.E.; Cimini, B.A.; Goodman, A. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **2021**, *22*, 433. <https://doi.org/10.1186/s12859-021-04344-9>.
  60. Danovski, G.; Dyankova, T.; Stoyanov, S. CellTool: an open source software combining bio-image analysis and mathematical modeling. *bioRxiv* **2018**. <https://doi.org/10.1101/454256>.
  61. Schwab, R.A.; Niedzwiedz, W. Visualization of DNA Replication in the Vertebrate Model System DT40 using the DNA Fiber Technique. *Journal of Visualized Experiments* **2011**, *56*, e3255. <https://doi.org/10.3791/3255>.
  62. Kirilov, T.; Gospodinov, A.; Kirilov, K. An algorithm and application to efficiently analyse DNA fibre data. *Biotechnology & Biotechnological Equipment* **2023**, *37*, 2206488. <https://doi.org/10.1080/13102818.2023.2206488>.