Piece of the Puzzle: Remdesivir disassemble the multimeric SARS-CoV-2 RNA-dependent RNA Polymerase Non-Structural Proteins (RdRp-NSPs) complex

Fisayo A. Olotu  
Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban

Kehinde F. Omolabi  
Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban

Mahmoud E. S Soliman (soliman@ukzn.ac.za)  
Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban

Research Article

Keywords: SARS-CoV-2, RNA-dependent RNA polymerase, Non-Structural Proteins, Accelerated Molecular Dynamics Simulation, Remdesivir, ATP

Posted Date: April 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-23431/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

The recently emerged SARS-like coronavirus (SARS-CoV-2) has continued to spread rapidly among humans with alarming upsurges in global mortality rates. A major key to tackling this virus is to disrupt its RNA replication process as previously reported for Remdesivir ($\text{Rem-P}_3$). For the first time, we modeled the binding of $\text{Rem-P}_3$ to SARS-CoV-2 RdRp-NSPs complex, a multimeric assembly that drives viral RNA replication in human hosts. Findings revealed that while ATP-binding stabilized the replicative tripartite, $\text{Rem-P}_3$ disintegrated the RdRp-NSP complex, starting with the detachment of the NSP7-NSP8 heterodimer followed by minimal displacement of the second NSP8 subunit (NSP8$_{II}$). More so, $\text{Rem-P}_3$ interacted with a relatively higher affinity ($\Delta G_{\text{bind}}$) while inducing high perturbations across the RdRp-NSP domains. D452, T556, V557, S682, and D760 were identified for their crucial roles in stacking the cyano-adenosine and 3,4-dihydroxyoxolan rings of $\text{Rem-P}_3$ while its flexible $P_3$ tail extended towards the palm domain blocking D618 and K798; a residue-pair identified for essential roles in RNA replication. However, ATP folded away from D618 indicative of a more coordinated binding favorable for nucleotide polymerization. We believe findings from this study will significantly contribute to the structure-based design of novel disruptors of the SARS-CoV-2 RNA replicative machinery.

Introduction

Coronaviruses (CoVs) are a representative group of single-stranded, positive-sense RNA-viruses that are known to cause severe respiratory disorders, gastrointestinal and central nervous system (CNS) diseases in both human and animals. $^1$

Severe acute respiratory syndrome CoV-2 is a newly emerged viral strain from the Coronaviridae family currently causing a global pandemic with severe threats to millions, and peradventure, billions of the world’s population. Middle East respiratory syndrome (MERS) CoV and Severe acute respiratory syndrome CoV are other members of the family earlier reported, with mortality rates far lesser than SARS-CoV-2. $^2$-$^4$ Generically referred to as COVID-19, SARS-CoV-2 has recorded more mortality rates globally than MERS and SARS-CoV combined. $^5$ So far, deaths from SARS CoV-2 have been recently estimated at 31,412 with about 667,090 confirmed cases in about 183 countries of the world. (Coronavirus: Euronews)

Presently, there is no FDA approved drug or vaccine for the treatment of SARS CoV-2 (FDA, 2020), the highly urgent need to develop drugs and vaccines that will efficiently curtail its virulence.

SARS CoV-2 replication is very crucial to its pathogenesis even though quite complex. $^8$ Pivotal to the viral establishment in host cells is the assemblage of cytoplasmic and membrane protected replication-transcription complexes (RTCs). $^9$,$^{10}$ The RTCs most importantly coordinate the expression, replication, and amplification of the viral genome. They also create an environment for the virus to evade the hosts’ immune system. $^{11}$ The production of viral sub-genomic mRNAs, synthesis of new genomic molecules and template strand required for replication are hampered without the RTCs. $^{12}$ At the center of the RTCs
is the RNA-dependent RNA polymerase (RdRp) subunit \( ^{11} \). In the CoV genome, the 5'-terminus ORF1a and ORF1b frames encode polyprotein 1a and 1b which are cleaved into 16 non-structural proteins (NSPs) that cumulatively drive the replication and transcription phases of the virus. \(^{13-17}\)

Several viral cycles mediated by the different RdRp-NSPs lead to the full replication and translation of the viral genome. \(^{1}\) One of these steps is the nucleotide polymerization which involves at least three RdRp-NSPs namely NSP12, NSP7, and NSP8.

The NSP12 subunit is a highly essential RdRp of the viral replicative machinery and reportedly interacts with its cofactors; NSP7 and NSP8 which altogether drive genomic replication. \(^{1,18}\) Previous reports have emphasized that NSP12 is unable to perform its function as a single entity \(^{19}\) but must exist in a tripartite complex with NSP7 and nsp8 for activating the replication of long RNAs. \(^{20,21}\) NSP8 has the capability of initiating the replication process, which accounts for its description as a primase. This subunit can \textit{de novo} synthesize about 6 nucleotides in length, which could serve as a primer for RNA synthesis by NSP12 RdRp. On the other hand, NSP7 is integral to the CoV replicase polyprotein which also functions as a primase and binds to NSP12. \(^{18,22,23}\) Together, the NSP7-NSP8 complex enhances the binding of NSP12 to RNA in addition to its enzymatic activity. \(^{8}\)

The functional multiunit complex of SARS-CoV has been structurally resolved, containing one molecule of NSP12, two molecules of NSP8 and one molecule of NSP7 \(^{1}\). According to their findings, NSP12 polymerase was bound to a heterodimer of NSP7-NSP-8 while a second NSP8 subunit was bound at a distinct NSP12 site.

Therefore, since this assembly is crucial for RNA synthesis, we presume that targeting this complex could be an important strategy to interfere with the viral replication process. This could pave way for the design, synthesis, and repurposing of drugs that can disrupt this RdRp enzyme assembly, and consequently, viral replication.

GS-5734/Remdesivir (Rem), a nucleoside analogue, is a prodrug originally developed to combat Ebola virus, and functions by mimicking adenosine structure (Figure 1). \(^{24}\) This drug reportedly converts into a hydrolyzed and tri-phosphorylated (active) metabolite (Rem-P\(_3\)), a form that enhances its activity as a substrate for RdRp thereby replacing ATP. This event results in the termination of the polymerization process; regarded as ‘chain termination’. \(^{25}\)

Efficacies of \textbf{Rem} in COVID-19 treatment is currently been evaluated in clinical trials, in the US and China \(^{26}\), ever since the possibility was proposed \textit{in vitro}. 	extbf{Rem} exhibits broad-spectrum antiviral activities since most RNA viruses exhibit high structural similarities. \(^{24,27}\)

Remdesivir has been experimentally reported to inhibit the replication of SARS-COV-2 \(^{28}\), a feat that could facilitate its adoption and approval for COVID-19 treatment. Although, recent studies have reported its
binding mode and interaction pattern in SARS-CoV-2, there is still a gap in literature with regards to its inhibitory mechanism, most especially, on how it disrupts the replicative machinery of SARS-CoV-2.

Complementary to its competitive ATP binding, we seek to model the modulatory mechanisms of remdesivir on SARS-CoV-2 NSP12-NSP8-NSP7 complex, which is crucial for its replication in host cells.

Recent studies have reported similarity of about 96.4% between amino acid sequences of SARS-CoV-2 and SARS-CoV. More so, 3D structure of the nucleotide polymerization complex (NSP-12-NSP8-NSP7) of SARS-CoV has been resolved using single-particle cryo-electron microscopy (PDB entry 6NUR). This provided an important template for modeling SARS-CoV-2 NSP12-NSP8-NSP7 replicative complex as it applies to the rationale of this study. Pairwise alignment further revealed a shared similarity of 98.8% among the NSP7 of SARS-CoV and SARS-CoV-2, and a 97.5% similarity among their NSP8s.

Summarily, in this study, we seek to unravel the mechanisms by which active remdesivir (Rem-P₃) interrupts RdRp-mediated RNA replication in SARS-CoV-2.

Since the crystal structure of SARS-CoV-2 RdRp is currently unavailable, its nucleoside binding site was determined by structural alignment with the ATP-bound poliovirus RdRp (PDB ID: 2ILY). Unbound, Rem-P₃- and ATP-bound NSP8-NSP12-NSP8-NSP7 multimers were then subjected to GPU-accelerated molecular dynamics (MD) simulations according to protocols used in our previous studies. Conformational events were analyzed over the entire simulation period while relative binding modes of Rem-P₃ and ATP were analyzed. We believe findings from this study would significantly contribute to the design of highly efficient inhibitors for disrupting SARS-CoV-2 replication in human hosts.

**Results**

**Structural modeling of RdRp NSP8-NSP12-NSP7-NSP8 multimer**

We modeled the 3D structures of SARS-CoV-2 NSP7, NSP8, and NSP12 using corresponding proteins from the recently resolved SARS-CoV multimer as templates (PDB ID: 6NUR). Pairwise sequence alignment revealed sequence identities of 98.8% between NSP7, 97.5% NSP8 and 96.4% NSP12 of SARS-CoV and SARS-CoV-2 (Figure 2d). Further analyses identified regions in the SARS-CoV-2 NSP12 with constituent variations as it compares to SARS-CoV. Findings revealed high conservations at the fingers (residues 398-581, 628-687) and thumb (residues 816-919) domain which applies to the RNA/NTP binding grip. Most variations were observed across the N-terminal (residues 1-397) and palm (residues 582-627, 688-815) domains of both proteins. Variations across the NSP7 and NSP8 of SARS-CoV and SARS-CoV-2 were also indicated. More so, the N-terminal of SARS-CoV-2 NSP8 were truncated likewise the NSP7 C-terminal since they were not structurally resolved in 6NUR.

Structural alignment of the HADDOCK-derived complexes with ATP-bound poliovirus RdRp helped identify the ATP-bound multimer. Analyses of binding interfaces in the modeled SARS-CoV-2 NSP8-NSP12-NSP7-
NSP8 multimer revealed that the NSP7-NSP8 dimer was bound at loop interfaces of the finger-thumb domain (residues 408-414, 425-447, 552) of NSP12 while regions with residues 257-275, 322-348 and 367-407 served as a binding interface for the second NSP8 subunit. These details agreed with findings from previous structural studies \(^1\) and were used for flexible protein-protein docking.

Molecular docking yielded 10 Rem-P\(_3\) conformers among which the topmost ranked with a docking score of -7.4 kcal mol\(^{-1}\) was selected. This was aligned with the crystallized ATP molecule and yielded an RMSD of 0.2 Å. This further indicated that the inhibitor and substrate were properly aligned at the nucleoside pocket of SARS-CoV-2.

**Remdesivir systematically disintegrates the SARS-CoV-2 RdRp assembly**

To understand possible effects of Rem-P\(_3\) and ATP on the SARS-CoV-2 NSP complex, snapshots were taken from the resulting trajectories at various time-frames as the simulation proceeded. Comparative visual analysis was then carried out for the unbound, Rem-P\(_3\) - and ATP-bound systems.

We observed that the RdRp-NSP multimer remained as a unit in the unbound and ATP-bound systems from the initial to the final frames while dissimilar structural occurrences were observed in the Rem-P\(_3\)-bound complex (Figure 3).

Time-based conformational sampling revealed that the RdRp-NSP assembly was intact in the presence of Rem-P\(_3\) until around 150ns when the NSP8 subunit was firstly detached from its NSP7 counterpart into the solvent environment. This was then followed by the displacement of the NSP7 subunit from the loop interface of the NSP12-thumb domain at the next 20ns time-frame (170ns). Besides, we observed that the second NSP8 subunit was minimally displaced from its binding interface on NSP12. The disintegrated RdRp-NSP complex was maintained until the end of the simulation (Figure 4).

Based on previous studies involving SARS-CoV, NSP7 and NSP8 play important roles in *de novo* initiation, primer extension, RNA synthesis, and replication, particularly when complexed with NSP12. \(^{18,33,34}\). Also, an assemblage of the NSP7-NSP8 heterodimer with NSP12 reportedly enhanced its RNA binding and polymerase activities. \(^8,18\)

Therefore, the ability of Rem-P\(_3\) to disrupt the RdRp-NSP assembly in SARS-CoV-2 could correlate with its inhibitory activity against the virus. This is a major finding that possibly explains its reported efficacy in the treatment of COVID-19 infections.

We further investigated the effects of both compounds on the integrity of the NSP12 multimer over the simulation period by comparing parameters for the unbound, ATP-bound and Rem-P\(_3\)-bound systems. Firstly, we estimated the stability of the complexes using the RMSD metrics and result plotted in Figure 5a. As shown, the unbound and ATP-bound complexes were stable from the beginning to the end of the simulations. On the contrary, the Rem-P\(_3\)-bound NSP complex was structurally stable (RMSD < 2Å) until
～140ns where high instability was observed, characterized by an unusual spike in Cα motions after which the system was restabilized until the end of the run. This observation could correlate with the disintegration of the NSP7-NSP8 heterodimer from the NSP12 polymerase as reported above (Figure 4) thereby explaining high conformational perturbations induced by Rem-P₃ when bound to the SARS-CoV-2 RdRp-NSP complex.

A possible explanation is that the disintegration of the NSP7-NSP8 heterodimer could further destabilize the NSP12 subunit, which could, in turn, affect binding interfaces for other NSPs necessary for RNA replication. Another interesting and important finding was that the ATP-bound NSP-complex was more stable than the unbound system.

This could imply that the binding of ATP to NSP12 further enhanced the integrity of SARS-CoV-2 NSP8-NSP12-NSP7-NSP8 complex. To minimize entropical effects, we defined finally equilibrated RMSDs (FE-RMSDs) from the ultimate time frames (170-200ns) where the systems stabilized. As seen in Figure 5b, huge discrepancies exist between the stabilities of the simulated complexes, with a high FE-RMSD peak for the Rem-P₃ bound system. This could correlate with the systematic disintegration of the RdRp-NSP complex induced by Rem-P₃ relative to ATP which rather stabilized the complex.

Estimated mean FE-RMSDs of the systems were in the order ATP < Unbound < Rem-P₃ as presented in Supplementary Table 1. These stable time-frames were then utilized for subsequent conformational analyses.

Also, we projected the motions of the RdRp-NSP complexes along two principal components (PC₁ vs PC₂). From the PCA plot in Supplementary Figure S1, we observed similar directions of motions among the unbound and ATP-bound RdRp-NSPs while a more dispersed motion pattern was observed for the Rem-P₃-bound complex.

We further monitored systemic fluctuations among constituent residues of each NSP subunits using the FE-RMSF metrics, derived from the stable time-frames. Our findings revealed that per-residual motions were relatively high in the Rem-P₃-bound NSP complex compared to ATP-bound and unbound systems (Supplementary Figure S2). Collatively, the mean FE-RMSF was lowest in the ATP-system further indicative of the importance of ATP in stabilizing the RdRp-NSP multimeric complex. Estimated mean FE-RMSFs of the systems also followed in the ATP < Unbound < Rem-P₃ order.

We then mapped out the distinct subunits of the RdRp-NSP complex (including NSP12 sub-domains) and measured their fluctuations with respect to the binding of the compounds.

Our findings further emphasized the perturbative effects of Rem-P₃ across the NSP8-NSP12-NSP7-NSP8 assembly. The disruptive effects of Rem-P₃ binding were most pronounced in the NSP7-bound NSP8 heterodimer (FE-RMSF = 41.6Å) and could have led to their detachment from the finger loop interfaces, which was also highly perturbed (FE-RMSF = 9.0Å) according to our calculations (Figure 6). Cumulatively,
high Ca motions induced by \textit{Rem-P$_3$} at the NSP12 subdomains affected the binding and stability of associated NSPs at the interfaces (Supplementary Table 2). More so, high fluctuations induced by \textit{Rem-P$_3$} among residues of the NSP12 N-terminal and finger subdomains could have disrupted interface interactions with the second NSP8 subunit (NSP8$_{II}$), which also exhibited high structural perturbations (FE-RMSF = 11.9 Å, 9.0Å).

Interestingly, the NSP12 subdomains, in addition to other NSPs (7 and 8) exhibited minimal residual motions in the presence of ATP, indicative of the stability of the NSP complex, even compared to the unbound system. Hence, we can presume, that the binding of ATP further enhanced structural integrity across the RdRp-NSP assembly.

Differential stabilities of the NSP12-nucleoside sites were further evaluated, and as observed (Figure 5c), \textit{Rem-P$_3$}-bound nucleoside site was more unstable compared to the ATP-bound and unbound sites. This could suggest a possible correlation between the activities of \textit{Rem-P$_3$} at the nucleoside pocket and its disruptive effects on the entire RdRp-NSP complex.

Also, \textit{Rem-P$_3$} induced high Ca motions at the nucleoside site as further estimated using the FE-RoG metrics, which is also an indicator of structural compactness. Relatively, the nucleoside side was more compact when bound by ATP, even compared to the unbound NSP12-pocket (Fig 5d).

Furthermore, we projected trajectorial motions of the compounds at the active pockets with respect to their binding activities. From the PCA plot, we could observe that \textit{Rem-P$_3$} exhibited highly unstable motions at the nucleoside pocket while ATP was coordinately bound with more compact motions indicative of its systemic stability.

Complementarily, we masked the compounds for FE-RMSD calculations to further determine their binding stability. As shown in Figure 7, \textit{Rem-P$_3$} was highly unstable (FE-RMSD = 3.5Å ± 0.5 ) while ATP demonstrated stable motions at the binding pocket (FE-RMSD = 1.5Å ± 0.8).

Also, we observed that ATP assumed a folded conformation at the hydrophobic pocket of the NSP12 subunit, which could have accounted for the compact site architecture (Figure 7d). On the contrary, \textit{Rem-P$_3$} exhibited an extended conformation with its cyano-adenosine ring stacked in between a visible hydrophobic patch (Figure 7b). The effect of this \textit{Rem-P$_3$} binding mode could be seen in the pocket which was less compact.

Using average structures for the \textit{Rem-P$_3$}– and ATP-bound complexes, we closely analyzed relative orientations at the NSP12-nucleoside site and complementary interactions (Figure 8). As mentioned above, ATP ultimately assumed a compact conformation, with an inwardly folded P$_3$ tail while \textit{Rem-P$_3$} rather exhibited a linearly stretched conformation.
As observed, its cyano-nucleotide portion extended into the pocket created by D452, T556, V557, T680, S681, and S682 (finger domain). This ring orientation could allow for high-affinity (NH–O) interactions with D452 and T556, which could, in turn, hold the cyano-adenosine ring in place. More so, its $P_3$ tail extends into the palm domain towards K798, D618, K621, R553, and D623. Also, the 3,4 dihydroxyoxolan ring was coordinated by N691, S759, and D760. These findings, altogether, indicate an orientation that supports its inhibitory activities. Relatively, we could observe that the $P_3$ tail of ATP was folded away from D618 and K798, an important residue-pair responsible SARS-CoV RdRp polymerase activity as previously reported. 36

**Binding free energy calculations and decomposition**

The extent to which the compounds bind to the RdRp-NSP complex was determined using the MM/PBSA method. 37,38 Stable time-frames (170-200ns) were also selected to minimize entropial effects and our results are presented in Table 1. Accordingly, $Rem-P_3$ was strongly bound than ATP with $\Delta G_{bind}$ difference of -13.2 kcal mol$^{-1}$.

<table>
<thead>
<tr>
<th>Energy components (kcal mol$^{-1}$)</th>
<th>Complexes</th>
<th>$Rem-P_3$</th>
<th>ATP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta E_{vdW}$</td>
<td>-44.4±0.3</td>
<td>-19.0±0.5</td>
<td></td>
</tr>
<tr>
<td>$\Delta E_{ele}$</td>
<td>-38.7±0.8</td>
<td>-39.2±0.8</td>
<td></td>
</tr>
<tr>
<td>$\Delta G_{gas}$</td>
<td>-83.2±1.0</td>
<td>-58.2±0.9</td>
<td></td>
</tr>
<tr>
<td>$\Delta E_{GB}$</td>
<td>54.9±0.7</td>
<td>41.1±0.4</td>
<td></td>
</tr>
<tr>
<td>$\Delta E_{np,sol}$</td>
<td>-5.4±0.04</td>
<td>-3.4±0.1</td>
<td></td>
</tr>
<tr>
<td>$\Delta G_{solv}$</td>
<td>49.5±0.7</td>
<td>37.7±0.3</td>
<td></td>
</tr>
<tr>
<td>$\Delta H$</td>
<td>-33.7±0.6</td>
<td>-20.5±0.8</td>
<td></td>
</tr>
<tr>
<td>$-T\Delta S$</td>
<td>0.3±0.02</td>
<td>0.01±0.02</td>
<td></td>
</tr>
<tr>
<td>$\Delta G_{bind}$</td>
<td>-33.4±0.4</td>
<td>-20.5±0.8</td>
<td></td>
</tr>
</tbody>
</table>

$\Delta E_{ele} = $ electrostatic energy; $\Delta E_{vdW} = $ van der Waals energy; $\Delta G_{bind} = $ total binding free energy; $\Delta G_{solv} = $ solvation free energy $\Delta G_{gas} = $ gas phase free energy; $\Delta G_{GB} = $ polar desolvation energy; $\Delta E_{np,sol} = $ non-polar solvation energy

This relatively higher affinity for the NSP12 nucleoside site further reflects its ability to competitively impede and replace the natural substrate. 25

Analysis of the binding components further revealed that the binding of $Rem-P_3$ was more unfavorable in the polar region as evidenced by higher $\Delta G_{np,sol}$ value. This could suggest that $Rem-P_3$ does not only block ATP binding but binds deeply in the hydrophobic nucleoside pocket of the NSP12 subunit.
This could favor its retention and involvement with more residues of the pocket. Also, vdW contributions to \textit{Rem-P}_3 binding was relatively higher, which may compensate electrostatic ($\Delta E_{ele}$) effects that were negated by unfavorable polar solvation energies ($\Delta G_{solv}$). Cumulatively, $\Delta E_{vdW}$ and $\Delta E_{ele}$ highly favored gas-phase interactions ($\Delta G_{gas}$) of \textit{Rem-P}_3 relative to ATP.

Also, we calculated the energies of individual residues and their respective contributions to the disparate binding of both compounds. Energy contributions > -1 kcal mol$^{-1}$ were considered favorable and results are presented in Figure 9.

As shown, total energy contributions to \textit{Rem-P}_3 were highly favorable for D452 (-4.1 kcal mol$^{-1}$), D760 (-3.6 kcal mol$^{-1}$), S682 (-2.8 kcal mol$^{-1}$), R555 (-2.0 kcal mol$^{-1}$), T556 (-1.6 kcal mol$^{-1}$), T687 (-1.1 kcal mol$^{-1}$), V557 (-1.0 kcal mol$^{-1}$) and T680 (-1.1 kcal mol$^{-1}$) in that order.

This could further emphasize the importance of D452 in enhancing \textit{Rem-P}_3 binding via interaction with its cyano-adenosine portion, and D760 via interactions with its 3,4 dihydroxyoxolan ring (Figure 8a). Relative to ATP, most of these energies were reduced and highly unfavorable. For instance, D425 and R553 contributed unfavorable energies (+ve) of 1.0 and 5.3 kcal mol$^{-1}$ to ATP which could minimize its interaction affinity compared to \textit{Rem-P}_3. However, S549 (-1.1 kcal mol$^{-1}$), R555 (-5.9 kcal mol$^{-1}$), K621 (-2.5 kcal mol$^{-1}$) and D623 (-2.5 kcal mol$^{-1}$) all contributed favorably to ATP binding.

The roles of D618 and K798 to the binding of \textit{Rem-P}_3 were further defined with electrostatic contributions of -11.4 and -1.2 kcal mol$^{-1}$ respectively. On the contrary, electrostatic effects by D618 were highly unfavorable for ATP (10.1 kcal mol$^{-1}$).

This could suggest that ATP does not interfere with the roles mediated by D618 during nucleotide polymerization. More so, while D452 electrostatically favored \textit{Rem-P}_3 binding (-5.3 kcal mol$^{-1}$), an unfavorable contribution of +12.5 kcal mol$^{-1}$ was estimated for ATP. Similar \textit{Rem-P}_3 $\rightarrow$ ATP ($\Delta E_{ele}$) transitions were observed for T556 (-5.4 $\rightarrow$ 1.9 kcal mol$^{-1}$), D623 (-10.1 $\rightarrow$ 9.0 kcal mol$^{-1}$), D760 (-11.2 $\rightarrow$ 9.2 kcal mol$^{-1}$) and D761 (-8.3 $\rightarrow$ 7.4 kcal mol$^{-1}$). Presumably, residues that favored \textit{Rem-P}_3 binding at the nucleotide site of RdRp-NSP12 rather contributed unfavorably to ATP binding, which could limit ATP affinity in the region. Also, in \textit{Rem-P}_3, unfavorable $\Delta E_{ele}$ contributions were highly compensated for by per-residue $\Delta E_{vdW}$ energies which were favorable in R555 (-2.9 kcal mol$^{-1}$), T556 (-1.4 kcal mol$^{-1}$), V557 (1.0 kcal mol$^{-1}$), D623 (-2.3 kcal mol$^{-1}$), R624 (-1.4 kcal mol$^{-1}$), T680 (-1.1 kcal mol$^{-1}$), S681 (-1.4 kcal mol$^{-1}$) and S682 (-2.8 kcal mol$^{-1}$). These were decreased to -1.7 kcal mol$^{-1}$ (R555), -0.4 kcal mol$^{-1}$ (T556), +0.4 kcal mol$^{-1}$ (D623) and -0.7 kcal mol$^{-1}$ (R624) in the ATP-complex.

Taken together, we could presume that \textit{Rem-P}_3 binds uncoordinatedly at the NSP12 nucleoside site, interacting with a high number of residues along its path. This binding pattern favorably enhanced its
affinity and, perhaps, longer retention at the NSP12 site until it exerts maximum destabilizing effects sufficient to disrupt the RdRp-NSP multimeric assembly.

Relatively, the ability of ATP to fold its flexible P₃ tail away from D618 and K798 in a compact conformation indicates a selective and coordinated binding, favorable for the nucleotide polymerization process.

In this study, we investigated the differential binding of Rem-P₃ and ATP to the RdRp-NSP assembly. This involved subunit binding of an NSP7-NSP8 heterodimer to NSP12 while a second NSP8 (NSP8₁₁) was bound at a distant site. The complex, although not currently available for SARS-CoV-2 at the time of filing this report, was modeled using a structurally similar SARS-CoV RdRp-NSP complex (PDB ID: 6NUR). The Rem-P₃, ATP- and Unbound complexes were subjected to GPU-accelerated molecular dynamics simulation of 600ns after which the trajectories were sampled conformationally. Findings revealed that the RdRp-NSP complex was more stable in the presence of ATP, even compared to the unbound system while on the other hand, Rem-P₃ appeared to disintegrate the complex. The NSP7-NSP8 heterodimer was firstly detached at ~150ns from the NSP12 while the second NSP8 (NSP8₁₁) subunit was seen to be minimally displaced at the ultimate time-frame. Conformational analysis performed with RMSD and RoG parameters further revealed that Rem-P₃ induced a high degree of structural instability compared to the ATP and unbound system. RMSF analysis revealed relatively higher subdomain perturbations in the Rem-P₃-bound RdRp-NSP complex. Also, analysis of ligand motions revealed that ATP exhibited a more stable and compact motion compared to Rem-P₃, which could have also reflected on the active site, which was less deviated and more compact in the ATP-system. Analysis of binding dynamics revealed that Rem-P₃ exhibited an extended conformation that allowed it to interact with D542, T556, V557, T680, S681 and S682 of the NSP12-finger domain via its cyano-adenosine ring while its 3,4 dihydroxyoxolan ring oriented towards D760 and N961 forming high-affinity H-bonds. Also, its P₃ tail was extended towards the palm domain, interacting with D618 and K798. This binding pattern clearly interferes with the roles mediated by D618 and K798 in RNA replication process. On the contrary, ATP exhibited a more selective and coordinated binding wherein it assumed a folded (compact) conformation away from these residues. Binding energy value of -33.4 kcal mol⁻¹ was estimated for Rem-P₃ while ATP had a ΔGbind estimate of -20.5 kcal mol⁻¹. This further reflected the ability of Rem-P₃ to competitively replace ATP at the NSP12-nucleoside pocket. Per-residue energy decomposition further emphasized the roles of D542, T556, V557, T680, S681, S682, D618 and K798 to the high-affinity binding of Rem-P₃. We believe findings from this study will significantly contribute to drug design endeavors, particularly structure-based approaches. This could yield novel compounds that can bind selectively and strongly to the NSP12-nucleoside site and disrupt interactions with other non-structural proteins.

**Methods**

**Structural modeling of ligand and protein complexes**
Based on the rationale of this study, it was expedient to model the NSP8-NSP12-NSP7-NSP8 multi-subunit complex for SARS-CoV-2 to which the inhibitor ($\text{Rem-P}_3$) and substrate ($\text{ATP}$) will be differentially bound. Homology models of each protein were built by MODELLER 9.18, a structural refinement tool, for which SARS-CoV NSP12, NSP7 and NSP8 crystal structures (PDB ID: 6NUR)\(^1\) were used as templates. Amino acid sequences for the SARS-CoV-2 NSPs were retrieved from the NCBI with entries YP_009725303.1 (NSP7), YP_009725304.1 (NSP8) and YP_009725307.1 (NSP12). From the modeled structures, the ones with the best DOPE scores were selected. To obtain the tetramer assembly similar to the crystallized complex in\(^1\), modeled SARS-CoV-2 NSP12 was first superimposed with a 6NUR_multimeric complex (SARS-CoV) after which active residues involved in NSP12, NSP7 and NSP8 interactions on SARS-CoV-2 were mapped out (Figure 2). These coordinates were then used to generate the SARS-CoV-2 multimer in a three-stepped flexible protein-protein docking procedure on the High Ambiguity Driven protein-protein DOCKing (HADDOCK 2.2) tool; step 1: NSP7-NSP8, step 2: the NSP7-NSP8-NSP12 and finally, step 3: NSP7-NSP8-NSP12-NSP8 multimeric complex after which 1000 energy minimization steps were performed to relax the model.

We then defined the nucleoside region in the modeled SARS-CoV-2 multimer for binding simulation studies with $\text{ATP}$ and $\text{Rem-P}_3$. This was achieved by aligning the modeled complex with ATP-bound poliovirus RdRp (PDB ID: 2ILY).\(^{29}\) Upon generation of the $\text{ATP}$-bound SARS-CoV-2 NSP7-NSP8-NSP12-NSP8 multimer, a grid box was centered on the $\text{ATP}$ site to which optimized $\text{Rem-P}_3$ was docked using the Vina module.\(^{39}\)

Prior to this, $\text{Rem-P}_3$ was prepared using Gaussview after which the 2D structure was minimized at the B3LYP/6-311++G(d,p) theory level on Gaussian16 program package.\(^{40}\)

The best binding pose was aligned with the crystallized nucleoside in 2ILY, and was assumed as the starting structure for the $\text{ATP}$ molecule and $\text{Rem-P}_3$ in the SARS-CoV-2 RdRp multimer.

**GPU-accelerated molecular dynamics (MD) simulation**

Having modeled the SARS-CoV-2 multimer and docked the compounds at the nucleoside binding site, the systems: NSP8-NSP12-NSP7-NSP8-ATP, NSP8-NSP12-NSP7-NSP8-$\text{Rem-P}_3$ and unbound NSP8-NSP12-NSP7-NSP8, were subjected to simulation runs of 600ns on Amber 18 Graphical Processor Unit\(^{41}\) which enabled accelerated production runs. Parametrization of the individual receptor was performed on FF14SB forcefield. Also, Antechamber and Parmchk modules were used to generate .frcmod files for $\text{ATP}$ and $\text{Rem-P}_3$. The LEAP module was then used to generate topology and parameter files for the complexes in addition to system neutralization and explicit solvation.\(^{42}\) These complexes were then minimized for 2500 steps with a 500kcal/mol Å\(^2\) restraint potential and also for 5000 steps with no restraints. Simultaneous heating (0-300k) and equilibration steps were performed followed by production runs that were restarted subsequently. Resulting trajectories were saved at every 1ns after which they were analyzed with the integrated CPPTRAJ module.\(^{43}\)
Relative stabilities of the SARS-CoV-2 multimers were determined by measuring the Cα-root mean square deviations (RMSDs) while other metrics such as the root mean square fluctuation (RMSF), radius of gyration (RoG) and solvent accessibility surface area (SASA) were used to measure per-residual motions, structural compactness, and solvent-surface motions.\textsuperscript{30–32,35,44,45}

Trajectorial motions of the molecules were also projected along two principal components (PC1 and PC2) for insights into their dynamics and motion patterns.\textsuperscript{46–48} Data plots and snapshots for time-based visual analysis were obtained using the Origin analytical software and UCSF Chimera GUI\textsuperscript{49}, Biovia Discovery Studio\textsuperscript{50} and GIMP 2.0.

**Binding free energy calculation**

The Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) method\textsuperscript{51} was used to measure the binding of ATP and Rem-P\textsubscript{3} at the NSP12-nucleoside pocket of the SARS-CoV-2 multimer (NSP8-NSP12-NSP7-NSP8). Estimated $\Delta G_{\text{binds}}$ were decomposed into various energy components [$\Delta E_{\text{ele}} = \text{electrostatic energy}; \Delta E_{\text{vdW}} = \text{van der Waals energy}; \Delta G_{\text{solv}} = \text{solvation free energy} \ \Delta G_{\text{gas}} = \text{gas phase free energy}; \Delta G_{\text{GB}} = \text{polar desolvation energy}; -T\Delta S = \text{entropy}, \Delta E_{\text{np,sol}} = \text{non-polar solvation energy}$] based on the expression: $\Delta G_{\text{bind}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}} + \Delta E_{\text{GB}} + \Delta E_{\text{np,sol}} - T\Delta S$ to further understand the differential binding mechanisms of the compounds. Also, we tracked the interaction energies of the compounds over the simulation period to gain insights into the binding dynamics. Per-residue energies were also decomposed to identify residues that significantly contribute to inhibitor/substrate binding and stability at the NSP12 active pocket.

**Declarations**

**Data Availability**

All data generated or analyzed during this study are included in this published article (and its Supplementary Information file).

**Acknowledgment**

The authors would like to thank Centre for High-Performance Computing (CHPC), Cape Town, South Africa, for making computational resources available.

**Authors Contributions**

FAO conceptualized the study, analyzed, discussed results and drafted the manuscript. KFO assisted with the system set up and performed accelerated molecular dynamics simulations while MES supervised the study, proof-read, and approved for submission to the Scientific Reports Journal.

**Additional information**
Competing interests statement

The authors declare no competing interests.

References

5. Mahase, E. Coronavirus: covid-19 has killed more people than SARSFig and MERS combined, despite lower case fatality rate. **641**, 2020 (2020).
7. Coronavirus Disease 2019 (COVID-19) | FDA


31. Abdullahi, M., Olotu, F. A. & Soliman, M. E. Solving the riddle: Unraveling the mechanisms of blocking the binding of leukotoxin by therapeutic antagonists in periodontal diseases. Cell. Biochem. 119,


**Supplemental Information Note**

Supplementary Figure S1. Principal component plot showing projection of the RdRp-NSP complex in the unbound (black), ATP-bound (yellow) and Rem-P3-bound states.

Supplementary Figure S2. RMSF plot showing residual fluctuations across the NSP12 subdomains in the unbound (black), ATP-bound (yellow) and Rem-P3-bound states. a. NSP12- finger domain b. NSP12 N-terminal domain c. NSP12-Pal domain d. NSP12-Thumb domain

**Figures**

![Chemical structures of Remdesivir in its prodrug and active forms](image)

**Figure 1**

Chemical structures of Remdesivir in its prodrug and active forms
Figure 2

Structural modeling steps of SARS-CoV-2 NSP8-NSP12-NSP7-NSP8 complex from SARS-CoV (PDB ID – 6NUR). Sequence alignment of SARS-CoV-2 NSP7, NSP8 and NSP12 proteins with the template protein SARS-CoV. Points of variations across all subunits are highlighted.
Figure 3

Time-based conformational sampling of unbound and ATP-bound NSP8-NSP12-NSP7-NSP8 multimer. Insets show time-based orientations of ATP at the NSP12 nucleotide site.
Figure 4

Structural analysis showing time-based disintegration of the RdRp-NSP assembly in the presence of Rem-P3. Different orientations assumed by Rem-P3 at each time-frames are also shown (insets).
Figure 5

Conformational analyses for whole structural and nucleoside site stability using Cα-RMSD and RoG parameters. a. shows huge structural instability for Rem-P3 system. b. Finally equilibrated (FE)-RMSD using more stable time-frames (170-200ns). c. active site stability – FE-RMSD d. FE-RoG plot showing degree of active site compactness.
Figure 6

RMSF analysis showing perturbations across the NSP7-NSP8 heterodimer (a and b) in addition to the second NSP8 subunit designated NSP8II (c). 3D depiction of the structural events is shown in d. Unbound-, ATP- and RemP3-bound NSP subunits are colored black, yellow and green respectively.
Figure 7

Analysis of Rem-P3 and ATP motions, and orientations at the active site region of RdRp-NSP12.  

a. Ligand FE-RMSD  
b. Binding modes of Rem-P3 and d. ATP at the hydrophobic NSP12 pocket. Interacting residues  
are also shown in addition to the degree of hydrophobicity color coded from least → highly hydrophobic;  
-3.00 → 3.00. c. PCA clustering of ligand motions over the simulation period.
**Figure 8**

Relative orientations and interaction analyses of a. Rem-P3 and b. ATP at the NSP12 nucleoside site. Complementary H-bond interactions are also pin-pointed.
Figure 9

Per-residue decomposition plots showing energy contributions by nucleoside-site residues. a. Total energy plots for Rem-P3 and b. ATP c. Electrostatic and vdW energy contributions for Rem-P3 and d. ATP.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supportinginformation.pdf