On the Recognition of Natural Substrate CTP and Endogenous Inhibitor ddhCTP of SARS-CoV-2 RNA-Dependent RNA Polymerase: A Molecular Dynamics Study

Angela Parise, Giada Ciardullo, Mario Prejanò, Aurélien de La Lande, Tiziana Marino

On the recognition of natural substrate CTP and endogenous inhibitor ddhCTP of SARS-CoV-2 RNA-dependent RNA-polymerase: a molecular dynamics study

Angela Parise[1,2,§], Giada Ciardullo[1,§], Mario Prejanò[1], Aurélien de la Lande[2], Tiziana Marino[1]*

[1] Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via Pietro Bucci, 87036 Arcavacata di Rende, CS, Italy

[2] Université Paris-Saclay, CNRS, Institut de Chimie Physique UMR8000, Orsay, France

[§] These authors contributed equally.

Abstract

The novel coronavirus SARS-CoV-2 is the causative agent of the COVID-19 outbreak that is affecting the entire planet. As the pandemic is still spreading worldwide, with multiple mutations of the virus, it is of interest and helpful the employment of computational methods for identifying potential inhibitors of the enzymes responsible for viral replication. Attractive antiviral nucleotide analogue RNA-dependent RNA polymerase (RdRp) chain terminator inhibitors are investigated with this purpose. This study, based on molecular dynamics (MD) simulations, addresses important aspects of the incorporation of an endogenously synthesised nucleoside triphosphate, the ddhCTP, in comparison with the natural nucleobase cytidine triphosphate (CTP) in the RdRp. The ddhCTP species is the product of the viperin antiviral protein as part of the innate immune response.

The absence of the ribose 3'-OH in ddhCTP could have important implications in its inhibitory mechanism of RdRp. We built an *in silico* model of RNA strand embedded in RpRd using experimental starting from cryo-electron microscopy structure and exploiting the information obtained by spectrometry on the RNA sequence. We determined that the model was stable during the MD simulation time. Obtained results provide deeper insights into nucleoside triphosphates incorporation whose molecular mechanism by the RdRp active site is still elusive.

1. Introduction

The world is currently in a state of pandemic emergency due to the spread of a disease called COVID-19 caused by a new viral etiologic agent, SARS CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2)[1]. Today SARS-CoV-2 has become probably the fastest-spreading virus in the human history. In fact, despite the vaccination, the occurring continuous specific mutations hide it from the immune system so that the use of antiviral therapy is mandatory to prevent future outbreaks of rapidly spreading coronaviruses.

SARS-CoV-2 is a positive RNA virus, which for its proliferation requires RNA-dependent RNA polymerase (RdRp), a key enzyme that regulates replication and transcription of the viral genome. For this reason, RdRp has become the validated target for the development of drugs against COVID-19 disease that can be hindered by the capacity of human coronavirus to proofread[2] and to remove mismatched nucleotides during genome replication and transcription. In fact, incorporating nucleoside analog molecules into viral RNA can lead to increased mutation rates and to the failure of the virus through "lethal mutagenesis".[3]

The RdRp complex, one of the largest proteins in the viral genome (932 residues), has multiple non-structural protein (nsp) units, and, in detail, is organized in nsp12 core catalytic unit, in nsp7- nsp8 heterodimer, nsp8 subunit, with a stoichiometry of 1:1:2 respectively. [4] These three non-structural proteins along with the nsp13 helicase are the "core" of the replication transcription complex (RTC). The category of antiviral drugs targeting the RdRp includes remdesivir,[5,6] a nucleotide analogue adenosine mimic that showed efficacy against Ebola infection and has been successfully repurposed for the treatment of SARS-CoV-2 infection, such that it is the only drug approved against SARS-CoV-2 by the US Food and Drug Administration (FDA), on August 10, 2020.[7] Several works have proposed remdesivir to induce delayed chain termination [8–12] or to cause the pausing and backtracking of polymerase, paving the way to the use of other nucleotide analogs.[13]

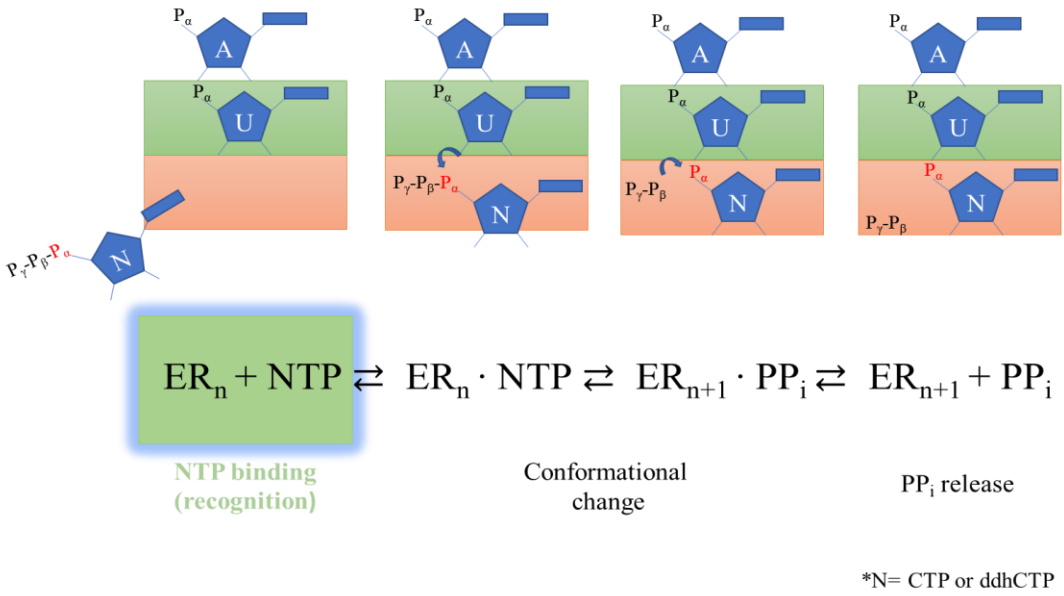
3'-deoxy-3',4'-didehydro-cytidine triphosphate (ddhCTP) is a novel antiviral nucleotide-like compound produced by enzyme viperin as part of the innate immune response, [14,15] and its easy synthesis of ddhCTP has been recently described. [16] Its chain terminator effect for RNA-dependent polymerases of multiple members of the Flavivirus genus has been further reported [17] and it has been recently demonstrated that the SARS-CoV-2 polymerase can incorporate this cytosine analogue well.[13] Drawn inspiration from that evidence on ddhCTP [13,18] and the computational work on Remdesivir, [8,19] a comparative *in silico* molecular modelling study (ddhCTP *versus* CTP) has been carried out on RdRp of SARS-CoV-2, to gain more insights on the incorporation of nucleoside triphosphates (see **Scheme**

1), which is one of the widely supported modes of interfering during viral replication pursuant to the high error rate (or low viral fidelity) of RNA virus replication.[2,20] The comparison between these two species is interesting because the ddhCTP competes with CTP, [21] as confirmed from an intracellular concentration of ~100 µM and lower than purine analogues, [22] despite differ of the -OH group loss in 3' position (see **Scheme 2**).

However, the absence of 3'-OH precludes further nucleotide incorporation from RdRp enzyme, which is consequently stalled when the RNA chain moves through the polymerase and the ddhCTP is processed. Another advantage of ddhCTP, with respect to other nucleoside analogue inhibitors, is that it does not require any particular metabolic transformations, differing from the case of other prodrugs (remdesivir and its derivatives).[13]

The choice of the investigation *in silico* of the endogenous product ddhCTP as antiviral drug represents a good strategy to develop other compounds in more reduced times than those needed to obtain new therapeutic molecules from the scratch. Furthermore, ddhCTP represents an example of how the location of the modification in the ribose differently from other antiviral nucleotide analogs can affect the catalytic pathway used for incorporation.[13]

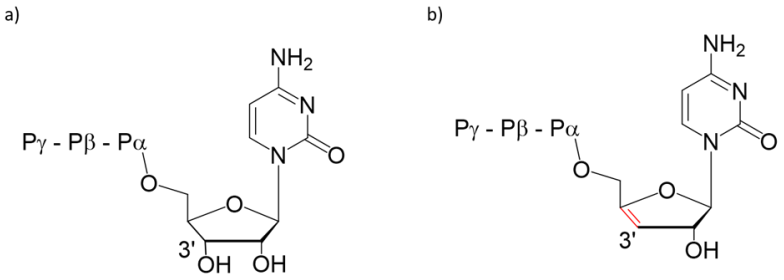
In the present study, classical Molecular Dynamics (MD) simulations are used to investigate the early steps of the RdRp working mechanism (recognition, see **Scheme 1**), including the contribution of possible conformational rearrangements in the substrate selection and binding.



Scheme 1. Canonical schematic reaction of RNA-dependent RNA polymerase.

Molecular docking has been used to test the affinity between polymerase active site and the examined ligands. MDs were performed on RdRp:RNA binary complex, including nsp7, nsp8, nsp12, and a partial double stranded RNA represented by a primer nucleotide strand of 6 units and a template strand of 8 nucleotides, and the ternary complexes formed by RdRp:RNA with ddhCTP and CTP.

The comparative analysis of the dynamic behaviour of RdRp in presence of both endogenously produced nucleotide analog and chain terminator, ddhCTP, and the natural nucleotide CTP provides deeper insights useful to better discriminate the factors underlying their incorporation in RdRp. In addition, the identification of amino acid residues important for the nucleotide recognition, and/or discrimination, represents one of the main focus of the present investigation, coupled to the description at atomistic level outer coordination shell of the two cations of the active site, which can result crucial for the viral replication.



Scheme 2. 2D representation of a) cytidine triphosphate (CTP) and b) 3'-deoxy-3',4'-didehydro-cytidine triphosphate (ddhCTP). P_α, P_β and P_γ refer to phosphate groups.

2. Computational Methods

2.1. In silico engineered RdRp:RNA model

The investigation started from a cryo-EM[23] crystal complex of RNA polymerase, with a resolution of 2.50 Å (PDB code 7AAP) and with a partial bound template-primer RNA and RDV-TP. The crystallographic structure concerns the subunits nsp7, nsp8 and the catalytic nsp12 bound to a template-primer RNA duplex, in complex with favipiravir ribonucleoside triphosphate (favipiravir-RTP). [23] Due

to different pyrimidine-like nature of the ddhCTP, a painstaking process was also required in the selection of the sequence of the RNA fragment.

Our study additionally includes the RNA template-primer similarly to *Arba et al.* work on the remdesivir,[19] which is crucial for the nucleotide triphosphate binding, differently to other computational investigation on remdesivir -SARS-CoV-2 RdRp interaction. [24–27]

Among the available crystallographic structures our choice fell on the one that presented two magnesium ions essential for the RdRp activity. In fact, it is not always trivial to find out experimentally the number of Mg ions present in the active site as well as the identity and the number of their coordination ligands.[28]

The nsp12 structure contains a right-hand RdRp domain (residues Val398 to Thr929) and a specific N-terminal extension domain (residues from Asp60 to Ala250) that adopts a nidovirus RdRp-associated nucleotidyltransferase (NiRAN) architecture (see **Figure 1**). [29]

The polymerase domain and NiRAN domain are connected by an interface domain (residues Leu251 to Ser397). A further N-terminal β hairpin (residues Val31 to Lys50) inserts into the cleavage blocked by the NiRAN domain and the palm subdomain in the RdRp domain.

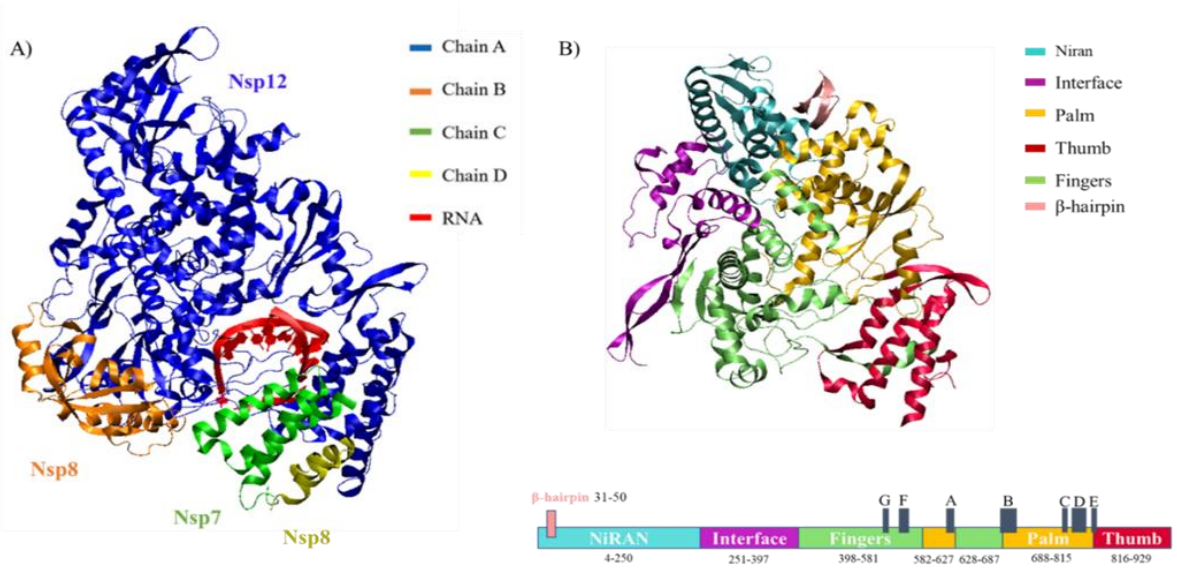


Figure 1. A) Structures of SARS-CoV-2 nsp7(2-68)-nsp8(78-191,84-111)-nsp12(4-929) complex. B) Structure of nsp12 domain organized by color for subdomains and in gray for A-G motifs (Motif A: s 612-626; Motif B: residues 678-710; Motif C: residues 753-767; Motif D: residues 771-796; Motif E: residues 810-820; Motif F: residues 544-560; Motif G: residues 499-511).

The polymerase domain adopts the conserved structure of the viral polymerase family[30] and consists of three subdomains of a "cupped right hand": a finger subdomain (residues Val398 to Ala581 and Asn628 to Thr687), a palm subdomain (residues Thr582 to Pro627 and Ala688 to Gln815) and a thumb subdomain (residues His816 to Thr929).

This right-hand architecture in turn evidences seven conserved structural motifs: motifs A-E located in the palm subdomain and motifs F and G in the fingers subdomain.[31] Most of these motifs are shared with other polymerases indicating the importance of these structural elements in their enzymatic function.[32]

The active site, contained in the palm domain, is formed by residues Ser759, Asp760 and Asp761 of motif C. [33]

The carboxylate groups of these aspartates anchor a pair of divalent metal ions (Mg^{2+}), which play the major role in catalysis, [30] while two other metal ions (Zn^{2+}) play a structural role to stabilize the enzyme. One of the zinc ions is coordinated to amino acid residues His295, Cys301, Cys306, and Cys310 in the N-terminal domain while the second is bonded to Cys487, His642, Cys645, Cys646 residues, in the finger domain. [34] As above-mentioned, the selected RdRp crystal structure incorporates the Favipiravir inhibitor, which is a purine nucleic acid analogue derived from pyrazine carboxamide (6-fluoro3-hydroxy-2-pyrazinecarboxamide).[35]

The design of RNA double strand length and the type of bases for cytosine analogue was then required. The selection of RNA sequence has been based on MALDI TOF experiments on CTP-like RpRd inhibitors. [36] The spectrometry study showed that the required base sequence for ddhCTP and CTP is UAAAAU (5' -> 3') and AUUUUAGU (3' -> 5'), for primer and template, respectively.

To obtain the desired model, the phosphodiester bond strain has been retained from the starting cryoEM structure, while all the already present nucleobases have been scrupulously replaced in silico, as shown in **Figure 2**, in agreement with the sequence for pyrimidine-like RdRp inhibitors. *tleap* tool in Amber16 [37] has been used to get the final model (**Figure 2C**).

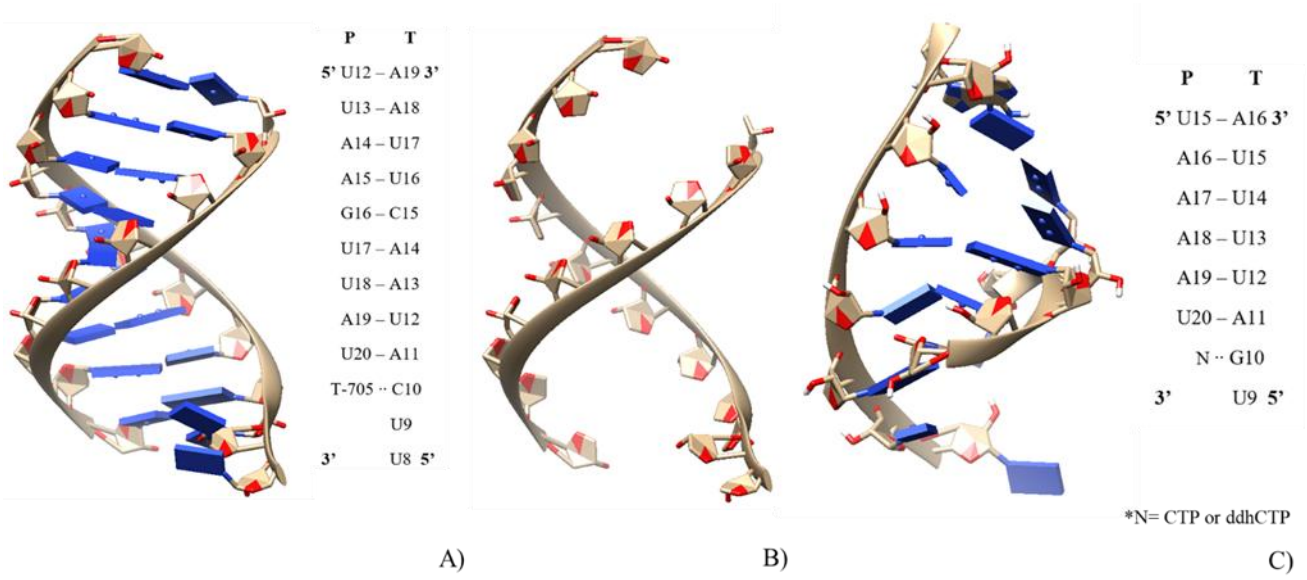


Figure 1. A) Cryo-EM structure of favipiravir-RTP in the catalytic site of the SARS-CoV-2 RdRp, in complex with template:primer RNA and 2D representation of template (T) and primer (P) including T-705 (favipiravir). B) Structure of favipiravir-RTP at the catalytic site of the RdRp in which only the phosphate-ribose scaffold is retained from the pdb 7AAP. C) Computational model used for RdRp in complex with template:primer RNA and 2D representation of template (T) and primer (P) incorporating ddhCTP inhibitor.

2.2. MD simulations setting

The parameters of cytidine triphosphate (CTP) were obtained by adopting AMBER 94 and AMBER 99 force field[38] data of mono-phosphate cytidine and parameters for the two terminal phosphate group of ATP are available in AMBER parameter database. In order to obtain ddhCTP parameters, gas-phase geometry optimization has been carried out using B3LYP/6-311G*. Atomic charges were derived by fitting the electrostatic potential according to the Merz−Singh−Kollman scheme,[39] using the RESP procedure. Antechamber and parmchk modules of Amber16 have been employed for generating preparatory files to perform MM relaxation of the complexes. 20 Na^+ counter ions were added to neutralize the system for both RdRp:RNA:ddhCTP and RdRp:RNA:CTP cases.

The RdRp-RNA protonation state was prepared at physiological pH by using H^{++} server, [40] with a salt concentration of 0.15 M. [37] The Amber force field FF14SB [41] and RNA.OL3 [42] were used for protein and RNA, respectively. ZAFF force field, [43] specifically parameterized for zinc-containing systems in

its commonly coordination in proteins, was selected for two Zn^{2+} ions, each of them coordinated to three Cys and one His as above specified. These ions only play a structural function. The Mg ions were treated with the Li et al. force field,[44] which allows a good structural description of the coordination sphere in polymerases [45,46] The system was solvated in a orthorhombic box (12 Å from the protein) of TIP3P water molecules.[47]

The solvated structures were first minimized by applying harmonic restraints on all atoms of the enzyme (50 kcal mol-1 $Å^2$), using 5.000 steps of the steepest descent algorithm, followed by 5.000 steps of the conjugate gradient algorithm (CG). In the second minimization step, we released restraint on hydrogen atoms, with third and fourth minimizations being conducted with and without restraints to the protein backbone atoms, respectively. We carried out a progressive heating phase from 0 to 310 K for 200 ps using the Langevin thermostat in NVT ensemble, with a time step of 0.002 ps. The production phase of 300 ns for RdRp:RNA and 300 ns for both RdRp:RNA:ddhCTP and RdRp:RNA:CTP complexes was performed under the following conditions: integration step of 2 fs coupling SHAKE algorithm, NPT ensemble at 1 bar pressure using the barostat. [48] The Particle Mesh Ewald summation method [49] was employed for the electrostatic potential and the long-range electrostatic interactions were calculated with 12 Å cut-off distance. Trajectories for RdRp:RNA and its complexes with ddhCTP and CTP were saved every 0.2 ps and analyzed through the *ptraj* module. [50] Overall, 0.9 μs of simulations were conducted (300 ns for each complex).

MD trajectories so obtained can then be used to assess the magnitude of structural changes (root- mean- square deviation, RMSD), the propensity for a given residue or region to move (root- mean- square fluctuation, RMSF), and the evolution of hydrogen bonding (HB) networks. The secondary structures were assigned with DSSP algorithm.[51]

RMSD-based clustering of the whole trajectories was performed according to the relaxed complex scheme (RCS) protocol as implemented in Amber16,[37] to provide a sampled and energetically accessible conformational ensemble. After removing overall rotations and translations by RMS-fitting the Cα atoms' positions of the trajectory, the average linkage clustering algorithm was applied as implemented in *cpptraj* to identify 10 representative conformations clusters of the protein. The four most populated structures were used as a starting point for the docking procedure. The described procedure has been successfully adopted in a number of investigations on different enzymes. [52-54]

Principal component analysis (PCA) was performed with *cpptraj* module of Ambertools 16 to extract the large-scale collective motions occurring in MD simulations of RdRp:RNA, RdRp:RNA:CTP and

RdRp:RNA:ddhCTP.[37] This procedure usually extracts information on the major conformational changes taking place along the MD trajectories.

Binding free energies between the binary complex RdRp:RNA and ddhCTP or CTP were calculated by solving the linearized Poisson-Boltzman equation using the molecular mechanics-Poisson Boltzman surface area (MM-PBSA) method, as implemented in Amber16 code,[37] selecting per-residue decomposition scheme. The value of *igb* flag equal to 5 associated to a salt concentration of 0.1 M was used. For the calculations 200 frames from each MD trajectory in the last 100 ns were analyzed.

3. Results and Discussions

3.1. Molecular dynamics of RdRp:RNA complex

As initial part of the investigation, 300 ns of molecular dynamics simulation on the RdRp:RNA complex were performed. The RMSD plot calculated for all the backbone atoms of the amino acid residues shows that the system reached the equilibrium after few ns of simulation, as confirmed by the calculated average value 1.49 ± 0.21 Å (see Figure S1). Such observed behaviour can be further evinced by the visual inspection of the protein structures obtained from the hierarchical clustering procedures (see Figure S2).

In addition, the RMSF of each individual residue of the non-structural proteins and RNA was calculated to characterize the local structural fluctuations in each subunit. (Figure S3) High peaks were observed at Val405, Lys430, Thr644, Val737 and Thr850, which corresponded to loop regions of the protein and at Leu895 and Glu903 in the proximity of carboxyl ends. Altogether, the preliminary analysis of the set of results on structural stability of the RdRp:RNA complex, allowed to consider reasonably stable the in silico-designed model.

PCA applied to MD trajectories to disclose the displacement of nsp7 nsp8 and nsp12 along the simulation has been carried out. The domains nsp7 and nsp8 are those that offset more during the simulation, with respect to the nsp12 one (see **Figure 3**). However, the present investigation does not consider such event, which is not therefore object of the study. Indeed, although nsp7 and nsp8 were retained in the simulations, we focused our attention on the catalytic subunit nsp12. The structural stability of the RNA chains is not affected by the above-mentioned movement, as confirmed by the well embedded conformation of RNA in the nsp12 pocket (**Figure 3**). The visual inspection of PCA

calculated for the RNA, furthermore, highlights that a slight shift takes place during the simulation for the nucleic acid, mainly in 3' and 5' positions of the T nucleotides (See **Figure 3).**

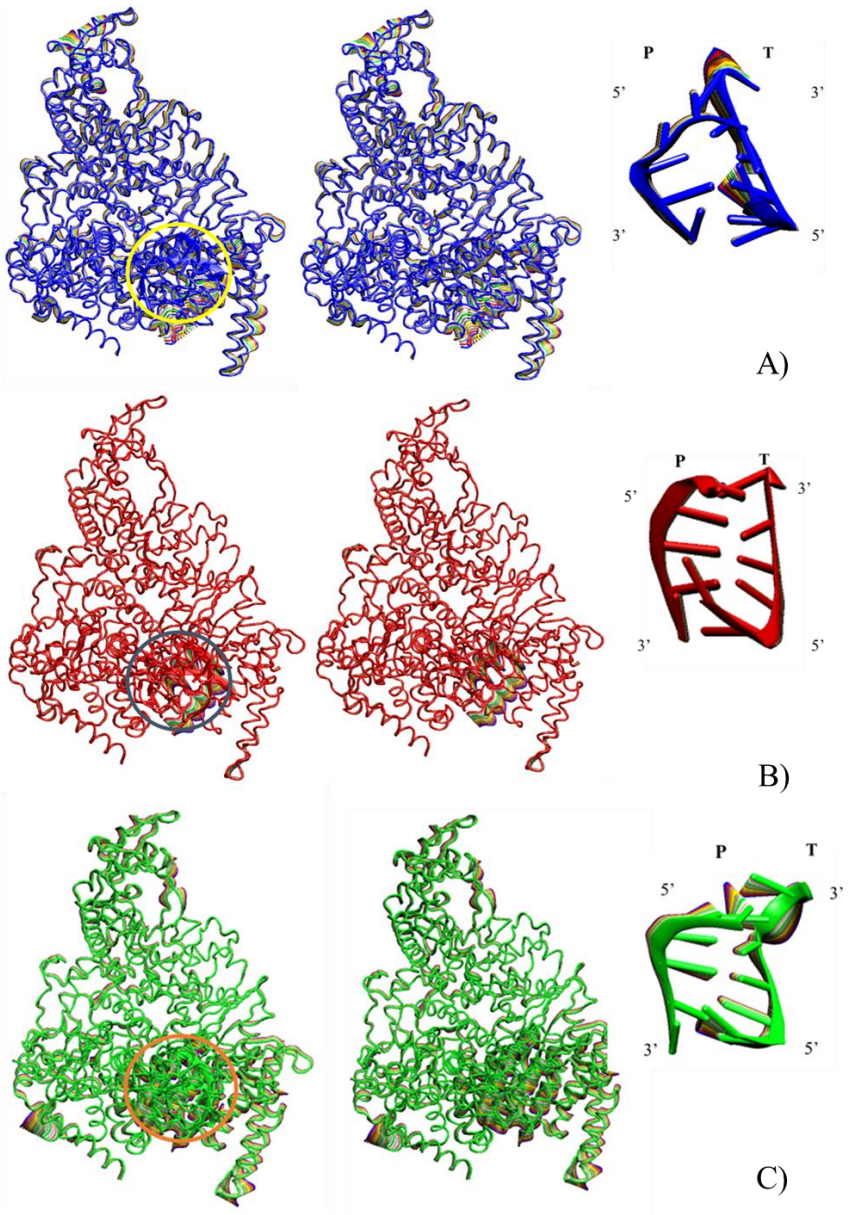


Figure 3. Principal component analysis of the RdRp:RNA binary (A), RdRp:RNA:CTP (B) and RdRp:RNA:ddhCTP (C) ternary complexes. In every system, RNA binding site is circled.

The structural homogeneity of nucleic acid is ensured by hydrogen bond interactions occurring with the positively charged residues of the domains, such as Arg569, Lys849, and Arg858 (Figure S4, Table

S1). Further high interactions occupancies occurred between P:U20 T:G10 T:U12 P:A17 P:A18 and Asp760, Ser682, Arg569, Arg585, Arg836, respectively, of the different subunits with an occupancy range of 77-100, as shown in Table S2 related to native contacts.

In the selected crystal structure, the active site Mg_A and Mg_B cations are far from putative ligands, like aspartate or glutamate residues, making hardly the establishment of metal first coordination sphere.[23] However, the analysis of trajectory revealed that Mg_A and Mg_B interacted mainly with Asp618, Asp760 and Asp 761, as will be shown later in detail. In general, both cations resulted in an octahedral geometry further consisting of solvent molecules, as interestingly evidenced by the analysis of radial distribution functions (RDFs) for the Mg_A-O_W and Mg_B-O_W pairs. In detail, first and second hydration sphere for the two metal ions were identified, with higher peaks observed in the case of Mg_B (see Figure S5). The most frequent presence of waters in proximity of Mg_B, with respect to Mg_A, can be ascribable to P:U20 nucleotide located in proximity of the latter, which occupies the room for solvent molecules. The presence of H_2O molecules has been further confirmed by the visual inspection of structures collected from the hierarchical clustering analysis, supplying an average number of 4 and 5 molecules in proximity of Mg_A and Mg_B, respectively (Table S3).

3.2. Molecular dynamics of RdRp:RNA:CTP and RdRp:RNA:ddhCTP complexes

The average binding affinities (in kcal/mol) obtained for the best docking poses of ddhCTP and CTP into the SARS-CoV-2 RdRp of are reported in Table S4. The most energetically favored conformers, -11.4 and -11.3 kcal/mol for of RdRp:RNA:CTP and RdRp:RNA:ddhCTP respectively, confirm that the introduction of ddhCTP can be considered feasible and, thus, can be allowed.

After the docking of the ligands, 300 ns of MDs were further performed on both RdRp:RNA:CTP and RdRp:RNA:ddhCTP ternary complexes. The RMSD plot reveals that the system reached the equilibrium, with average values slightly higher than the respective RdRp:RNA system (2.18 ± 0.28 Å and 2.20 ± 0.17 Å, for CTP and ddhCTP respectively vs 1.49 ± 0.21 Å for the binary, see **Figure 4**). Consistently with such behaviour, also the calculated RMSF values of ternary complexes were higher than the binary one (see **Figure 4A**).

In particular, in the case of the RdRp complexed to CTP, more intense fluctuations were detected in proximity of NiRAN (Asn198 and Pro227), fingers (Ser425) and thumb (Met902) subdomains.

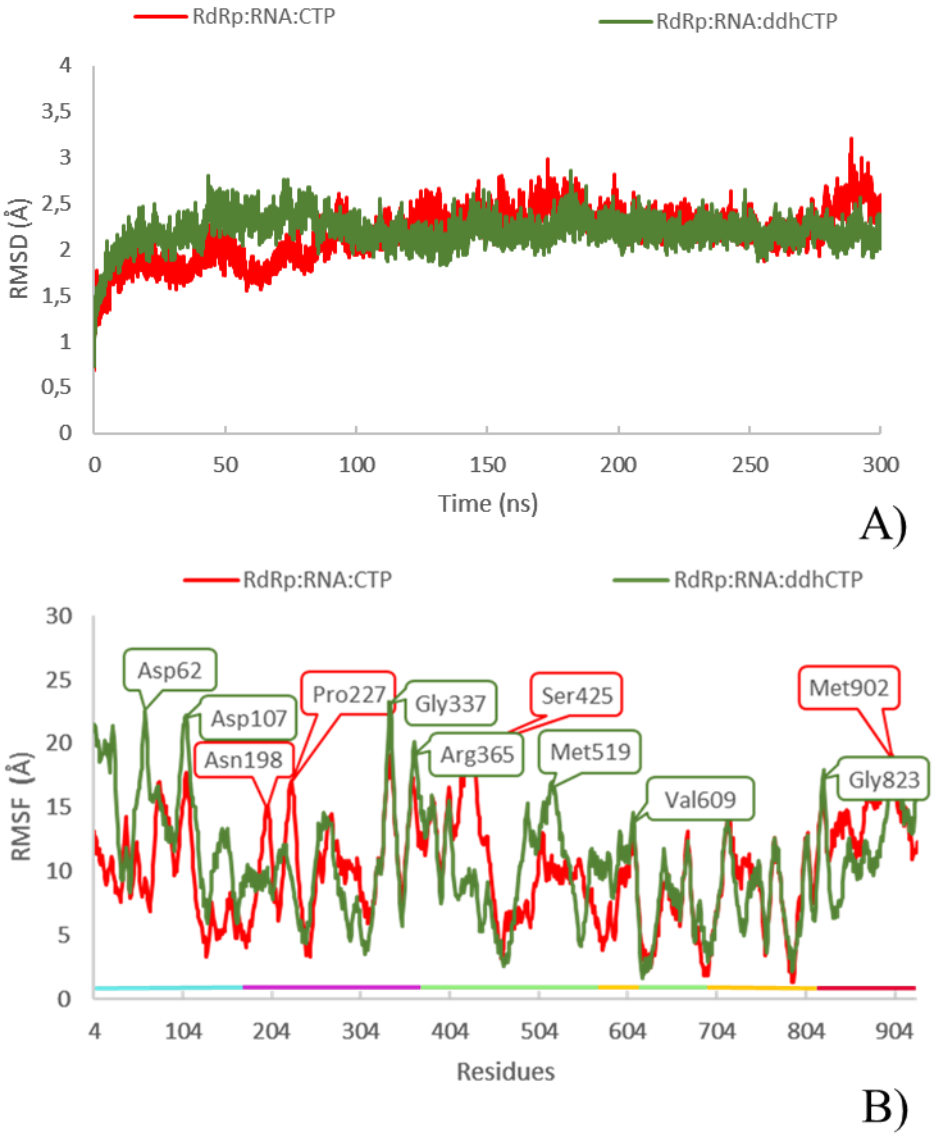


Figure 4. RMSD and RMSF plots of RdRP:RNA:CTP (red) and RdRP:RNA:ddhCTP (green) complexes.

In the MDs of RdRp:RNA:ddhCTP the major fluctuations were further observed in the interface subdomain, as evidenced by the peaks obtained for Gly337 and Arg365 residues exposed to solvent, in addition to fluctuations for the NiRAN (Asp62 and Asp107) and fingers (Met519 and Val 609) subdomains (see **Figure 4B**).

In analogy to the observation obtained for the RdRp:RNA complex, PCA revealed that for both ternary complexes nsp12 and nsp8 are the regions of the system involved in more detectable movements

during the MDs and that the RNA molecule conformation is not affected by any important structural rearrangement (see **Figure 3**).

In the binding site, CTP and ddhCTP ligands are stably located during the entire simulation, as can be observed by the analysis of their RMSD plots reported in Figure S6. The orientation of the substrates is mainly ensured by the presence of cation-phosphate interactions, occurring with different combination between the phosphate arms of CTP and ddhCTP with Mg_A and Mg_B. The distances of each Mg ions with P_α, P_β and P_γ were analyzed to characterize the binding mode of the phosphate moieties of ligands during the simulation and the relative distributions are displayed in Figure S7. In the case of CTP, the plot analysis revealed that oxygens of P_α group bridge both cations, while those of P_γ coordinate mainly the Mg_B (see Figure S8 for further details). A different behaviour of P_α group can be observed in ddhCTP as far as the Mg_A is concerned. In fact, it is shifted at distances around 5 Å.

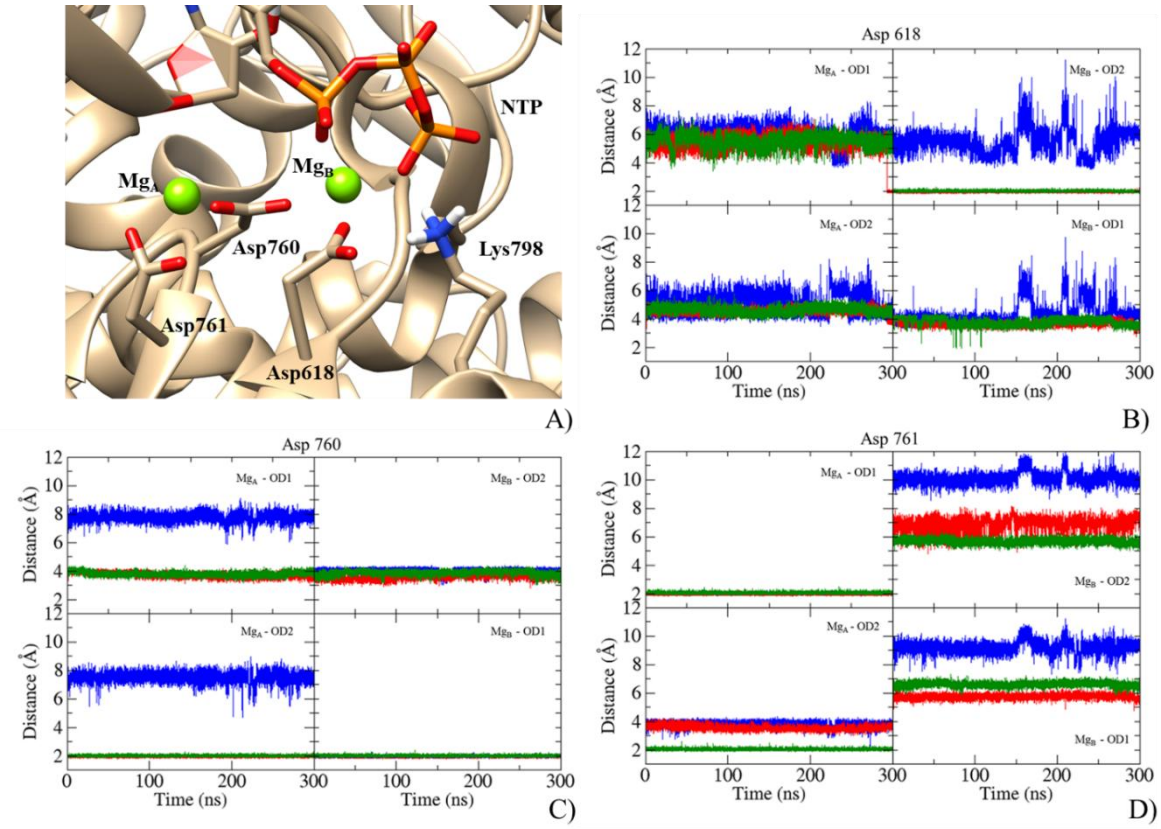


Figure 5. A) Active site of RdRp:RNA:NTP complex (NTP=CTP or ddhCTP). Distances calculated between Mg_A or Mg_B and B) Asp618, C) Asp760 and D) Asp761, for the RdRp:RNA (blue), RdRp:RNA:CTP (red) and RdRp:RNA:ddhCTP (green) system.

For Mg_B the distribution of P_α and P_γ results very similar for both ligands, while P_β in ddhCTP is longer than that in CTP. In addition to the phosphate groups of CTP and ddhCTP, Mg_A and Mg_B complete their hexa-coordination (Figure S8) with the side chain of two consecutive Asp residues in motif C, Asp 760 in a μ(1,3) fashion, as confirmed by a measured distance of ca 2 Å of metal-O_{Asp760} distances (see **Figure 5**), and Asp761. Asp618, at the beginning of motif A, preferentially coordinates the Mg_B one.

The above-described orientation of P_α, P_β and P_γ differs from that observed for the remdesivir, in the active site of RdRp. [8] It can be ascribable to a more advanced step of the recognition of remdesivir, considering that the OH of the primer is deprotonated and already coordinated to Mg_A. However, the orientation here-described of ddhCTP is encountered in other inhibitors, generally identified as "chain terminators".[13]

The oxygen atoms of P_β interacted during the whole simulation with the positively charged Lys551 and Arg553 residues of motif F for both examined ternary complexes, thus contributing to the binding of CTP and ddhCTP. Our finding is in good agreement with observations on other viral nucleic acid polymerases.[8,18,51,55,56]

Very interestingly, the RdRp:RNA:CTP and RdRp:RNA:ddhCTP analysis of the trajectories of salt-bridges established by Asp618, revealed that it is frequently implicated in the interaction with Lys798, while the same interaction was not frequently observed in the binary complex (**Figure 6**).

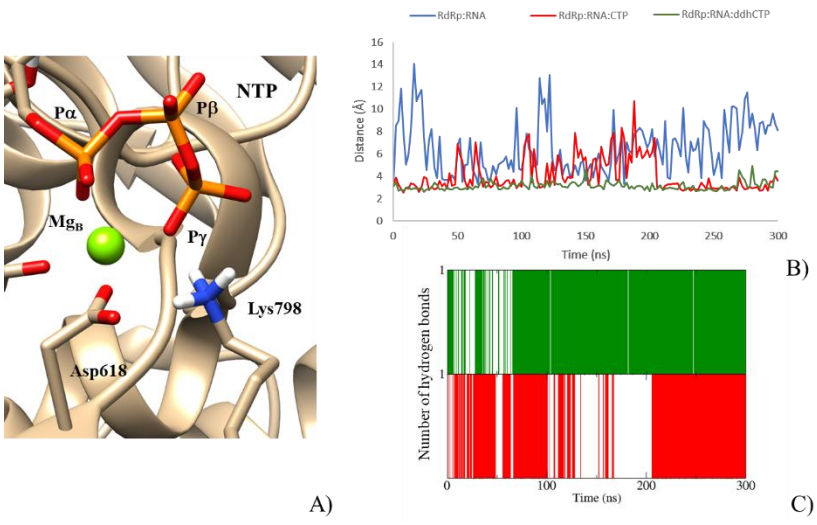


Figure 6. A) Focus on Lys798 orientation in active site of RdRP:RNA:NTP complex. B) Asp618-Lys798 salt-bridge distance and C) $O_{P\gamma}$-N_{Lys798} hydrogen bond calculated in during the MDs of RdRP:RNA:NTP complexes (NTP=CTP or ddhCTP) .

This behavior highlights the role played by Lys798 as "needed counterion" in orienting the P_γ towards Mg_B. Lys798 is indeed located within interactive hydrogen-bond distance to the P_γ of incoming ligands, thereby suggesting that such residue is involved in the binding of phosphorylated substrates and thus in the release of PPi after the chemical reaction has been completed (**Figure 6**).

This residue indeed represents the highly conserved lysine, critical in the control of the activity and fidelity of transcription and in the action act as a general acid to protonate the pyrophosphate leaving group, as widely emphasized in literature.[3,51]

Finally, a lower number of water molecules coordinating Mg_A and Mg_B was observed during the simulations due to the presence of phosphate groups, as can be noted from the count of solvent molecules coordinating the cations in the clustered geometries (see Table S3).

During the simulations, Mg_A and Mg_B lie at a distance of 4.48 ± 0.12 Å and 4.96 ± 0.15 Å for CTP and ddhCTP, respectively, (Figure S9). The observed value is consistent with that required by the two metal ions reaction mechanism (< 4 Å). It was suggested that during the reaction a shrinking of this distance may occur, leading O3′ closer to Pα and consequently facilitating the nucleophilic attack. The bit longer distance in the ddhCTP compared with CTP could affect the O3′−P_α alignment for catalysis during the nucleophilic attack, in the reactive Michaelis-Menten complex.[57-61]

In order to analyze the interaction involving the cytosine nucleobase, which is the same for both ligands, attention was focused on three main geometrical parameters involving the pairing with guanosine nucleobase, like $C1'_{CTP/ddhCTP}$-$C1'_{G:T}$, $N1_{CTP/ddhCTP}$ -$N1_{G:T}$ and $N3_{CTP/ddhCTP}$ -$C1'_{G:T}$. Outcomes on the analysis of these distances are linked to the fidelity of DNA polymerases[46,62] and can be extended to RNA polymerase, due to many similarities between the two classes of enzymes[63-65].

In the case of CTP ligand, the analysis of the distribution distances revealed that the nucleobases lie in an unpaired-like conformation, which can be linked to the low fidelity behavior of the RNA-polymerases accordingly to the mechanism. Indeed, the calculated average values of 11.73 ± 0.47 Å, 6.72 ± 0.59 Å and 5.98 ± 0.10 Å, for $C1'_{CTP}$ -$C1'_{T:G}$, $N1_{CTP}$-$N1_{T:G}$ and $N3_{CTP}$-$N1_{T:G}$ fit results described in previous work[46]. More interestingly, the inclusion of ddhCTP, increased of ca 2 Å all the considered distances (see Figure S10), thus suggesting that in presence of the endogenous inhibitor the fidelity of the enzyme is further affected and that errors in the course of the duplication can take place already in this step, thus concurring to the stalling of the enzyme.

Due to the modification present at the 3’ position of the ribose ring of ddhCTP, the analysis of radial distribution function and visual inspection of clustered structures were conducted on the interactions involving the ribose group, focusing on the hydroxyl groups at C2’ and C3’ positions of CTP in comparison with ddhCTP. In the case of CTP molecule, O3’ atom established hydrogen bonds with phosphate group of U20 and water molecules, while O2’ resulted mainly involved in interactions with solvent molecules. In absence of O3’-H group in the ddhCTP containing system, the interaction with U20 phosphate was observed occurring with the O2’ atoms. (Figure S11)

Contact analysis effected for both CTP and ddhCTP evinced as residues (618-621) of motif A, 760 of motif C and 551 of motif F are implicated in interactions with the ligands (see Table S5, S6, S7) suggesting a similar network of interaction inside the binding site. However, the analysis did not provide any relevant information about the interactions involving ribose hydroxyl groups and the amino acid residues.

The calculation of ΔG_{bind}, carried out according to the MM-PBSA scheme, revealed some differences for the ligands. In detail, the binding free energy of ddhCTP resulted 15 kcal/mol higher than CTP (Table S8). Energy decomposition analysis of every residue involved in non-covalent interaction with CTP and ddhCTP showed that, for the former, a stronger stabilization of tertiary complex is due to the electrostatic interaction of Arg553, Arg555, Lys621 and Asp761 sidechains (see Supporting information for details, Figure S12) with the active site environment.

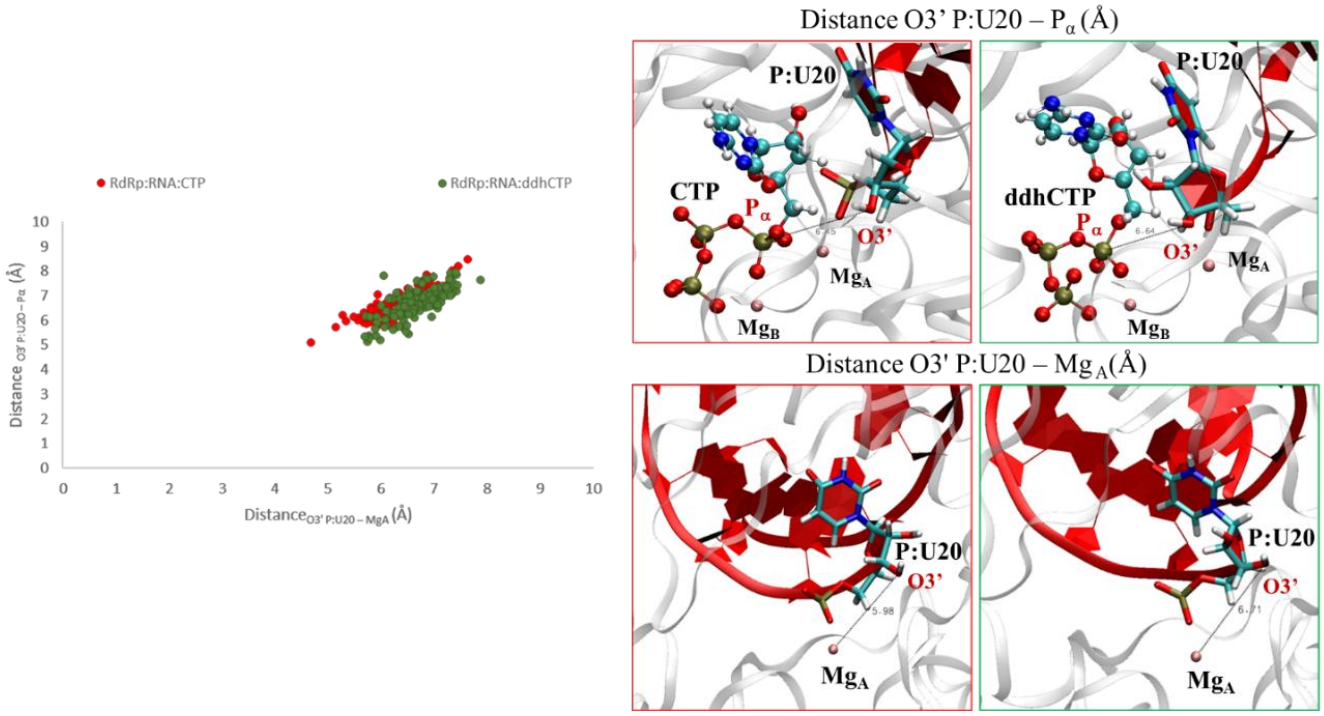


Figure 7. Distribution of O3'P:U20-P_α vs O3'P:U20-Mg_A distances calculated for RdRp:RNA:CTP and RdRp:RNA:ddhCTP complexes.

In addition, the higher ΔG_{bind} (lower affinity) can be linked to the absence of hydroxyl group at C3' position, which causes a loss of 3 hydrogen bond interactions from ddhCTP species.For both CTP and ddhCTP, finally, it can be noted a reacting like conformation in the course of the dynamics. This can be evinced, in particular, analyzing the Mg_A-O3'P:U20 and the O3'P:U20-$P_{\alpha NTP}$ distances, which are the reaction coordinates involved in the reaction catalyzed by the enzyme.

A very similar distribution of values was observed, see **Figure 7**, which can be further linked to experimental evidences concerning the incorporation and reactivity in presence of ddhCTP ligand.[13]

In addition, considering the importance of the distances Mg_A-$O3'_{P:U20}$, $P_{\alpha NTP}$-$O3'_{P:U20}$ and $OD1_{Asp761}$-$HO3'_{P:U20}$ we run umbrella sampling simulations, to restrain a reaction window corresponding to the transition slot from non-bonded-like to bonded-like conformations (Figure S13). [66]

The analysis of the obtained potential of mean force (PMF) showed that the approach of the $O3'_{P:U20}$ to the Mg_A generates two almost equal in energy minima, in both CTP and ddhCTP (Figure S13a). The free energy difference is the driving force of the process of unbound/bound, which indicates possible incorporation of a new base into the RNA strand. The analysis of the distance $P_{\alpha\ NTP}$-$O3'_{P:U20}$ (Figure S13b) revealed PMF values reasonably compatible with other theoretically studied polymerases.[31,45,46]

Following the reaction coordinate of the H-transfer from the ribose O3' (last reacting uracil of elongating strand) to the general base Asp761 (Figure S13c), finally, it was possible to establish a slightly favored thermodynamic process in the case of ddhCTP. In general, the presence of different local accessible minima reflects the complexity of the process, confirming anyway the potential of ddhCTP as an inhibitor of RdRp.

Conclusions

Nucleotide analogs represent a powerful tool to fight the viral infections and their consequences. In the present work, the interaction of RdRp protein, involved in the spreading of SARS-CoV-2 disease, and the endogenously synthesized ddhCTP, the product of the viperin antiviral protein, has been investigated. Such molecule differs from CTP, its natural precursor, by the absence of 3'-OH group. However, it is reputed that the missing hydroxyl group can favor the stalling of the enzyme and, consequently, the replication of viral RNA. The study has been conducted in the mean of all-atom MD

simulation, on the RdRp:RNA system and on each RdRp:RNA:NTP complex, with NTP equal to CTP or ddhCTP, and focused on the recognition step of RdRp catalytic mechanism.

From the analysis of binary complex, it turned out that the RNA is kept in active site by interaction with a number of residues, such as Ser682 and Arg585, and that, in absence of other substrates, Mg_A and Mg_B present a hexa-coordination sphere, comprised of Asp618, Asp760, Asp761 and water molecules.

The comparative analysis of conformational behavior of ddhCTP and CTP, overall, indicated high similarity between the two molecules. In detail, phosphate groups P_α and P_γ of CTP and ddhCTP are directly involved in coordination of Mg_A and Mg_B, while P_β mainly interacted with positively charged Lys551 and Arg553, facilitating the binding of the ligands in the active site. Importantly, Lys789 has been identified as crucial residue bridging Asp618 and P_γ group of both ddhCTP and CTP.

The hydroxyls groups of CTP at C2' and C3' positions established hydrogen-bond interaction with solvent molecules and U20 of RNA, respectively, while in the case of ddhCTP the interaction with nucleic acid is ensured by the O2'-H group.

The interactions of the cytidine group of both nucleosides with guanine of RNA chain were evaluated in term of $C1'_{CTP}$ -$C1'_{T:G}$, $N1_{CTP}$-$N1_{T:G}$ and $N3_{CTP}$-$N1_{T:G}$ distances. The analyzed results are in agreement with low fidelity of RdRp, which resulted slightly more affected by the presence of ddhCTP molecule. Finally, both CTP and ddhCTP showed pro-reactive conformation and umbrella sampling simulations revealed high energy similarity reacting-like and non-reacting like conformations of the ligands.

The results above-described shed light on relevant aspects of the incorporation of CTP and ddhCTP species in the RdRp, laying the foundation for the rational design of the SARS-CoV2 vaccine and antiviral targets and providing further insights that were not accessible to experiments. Altogether, the available results from experiments and the rationalization from computational investigation can thus be very helpful to stimulate and drive future experimental investigations.

Acknowledgments

We thank the Università degli Studi della Calabria-Dipartimento di Chimica e Tecnologie Chimiche (CTC) for the support. Computer simulations were carried out adopting resources of supercomputer Marconi 100 from ISCRA (Project code: IsC92).

Data and Software Availability

AutoDock version 4.2 was adopted to effect molecular docking calculations. All MD simulations were carried out adopting the AMBER16 package. AmberTools package was further adopted in the preparation of parameters and input for MD simulations. DFT calculations, to parametrize ddhCTP, were carried out using Gaussian 09 ver. D01. Trajectory analysis was done adopting AMBER cpptraj module. The preparation of the structures and the images were carried out using VMD software, version 1.9.3, and Chimera software. The full workflow is reported in the "Computational Methods" section of the manuscript.

References

1. Riva, L.; Yuan, S.; Yin, X.; Martin-Sancho, L.; Matsunaga, N.; Pache, L.; Burgstaller-Muehlbacher, S.; De Jesus, P. D.; Teriete, P.; Hull, M. V.; Chang, M. W.; Chan, J. F.-W.; Cao, J.; Poon, V. K.-M.; Herbert, K. M.; Cheng, K.; Nguyen, T.-T. H.; Rubanov, A.; Pu, Y.; Nguyen, C.; Choi, A.; Rathnasinghe, R.; Schotsaert, M.; Miorin, L.; Dejosez, M.; Zwaka, T. P.; Sit, K.-Y.; Martinez-Sobrido, L.; Liu, W.-C.; White, K. M.; Chapman, M. E.; Lendy, E. K.; Glynne, R. J.; Albrecht, R.; Ruppin, E.; Mesecar, A. D.; Johnson, J. R.; Benner, C.; Sun, R.; Schultz, P. G.; Su, A. I.; García-Sastre, A.; Chatterjee, A. K.; Yuen, K.-Y.; Chanda, S. K. Discovery of SARS-CoV-2 Antiviral Drugs through Large-Scale Compound Repurposing. *Nature* **2020**, *586* (7827), 113–119, DOI: 10.1038/s41586-020-2577-1
2. Robson, F.; Khan, K. S.; Le, T. K.; Paris, C.; Demirbag, S.; Barfuss, P.; Rocchi, P.; Ng, W. L. Coronavirus RNA proofreading: molecular basis and therapeutic targeting. *Molecular cell* **2020**, *79*(5), 710-727, DOI: 10.1016/j.molcel.2020.07.027
3. Boehr, D. D.; Arnold, J. J.; Moustafa, I. M.; Cameron, C. E. Structure, dynamics, and fidelity of RNA-dependent RNA polymerases. *Nucleic Acid Polymerases* **2014**, chapter 14, pp. 309-333. Springer, Berlin, Heidelberg, DOI: 10.1007/978-3-642-39796-7_14
4. Jiang, Y.; Yin, W.; Xu, H. E. RNA-Dependent RNA Polymerase: Structure, Mechanism, and Drug Discovery for COVID-19. *Biochem. Biophys. Res. Commun.* **2021**, *538*, 47–53, DOI: 10.1016/j.bbrc.2020.08.116
5. Wang, M.; Cao, R.; Zhang, L.; Yang, X.; Liu, J.; Xu, M.; Shi, Z.; Hu, Z.; Zhong, W.; Xiao, G. Remdesivir and Chloroquine Effectively Inhibit the Recently Emerged Novel Coronavirus (2019-NCoV) in Vitro. *Cell Res* **2020**, *30* (3), 269–271, DOI: 10.1038/s41422-020-0282-0
6. Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; Chang, S.; Xie, Y.-C.; Tian, G.; Jiang, H.-W.; Tao, S.-C.; Shen, J.; Jiang, Y.; Jiang, H.; Xu,

Y.; Zhang, S.; Zhang, Y.; Xu, H. E. Structural Basis for Inhibition of the RNA-Dependent RNA Polymerase from SARS-CoV-2 by Remdesivir. *Science* **2020**, *368* (6498), 1499–1504, DOI: 10.1126/science.abc1560

7. Lamb, Y. N. Remdesivir: First Approval. *Drugs* **2020**, *80* (13), 1355–1363, DOI: 10.1007/s40265-020-01378-w
8. Aranda, J.; Wieczór, M.; Terrazas, M.; Brun-Heath, I.; Orozco, M. Mechanism of reaction of RNA-dependent RNA polymerase from SARS-CoV-2. *Chem catalysis* **2022**, *2*(5), 1084-1099, DOI: 10.1016/j.checat.2022.03.019
9. Gordon, C. J.; Tchesnokov, E. P.; Feng, J. Y.; Porter, D. P.; Götte, M. The Antiviral Compound Remdesivir Potently Inhibits RNA-Dependent RNA Polymerase from Middle East Respiratory Syndrome Coronavirus. *J. Biol. Chem.* **2020**, *295* (15), 4773–4779, https://doi.org/10.1074/jbc.AC120.013056
10. Tchesnokov, E. P.; Obikhod, A.; Schinazi, R. F.; Götte, M. Delayed Chain Termination Protects the Anti-Hepatitis B Virus Drug Entecavir from Excision by HIV-1 Reverse Transcriptase. *J. Biol. Chem.* **2008**, *283* (49), 34218–34228, DOI: 10.1074/jbc.M806797200
11. Salgado-Benvindo, C.; Thaler, M.; Tas, A.; Ogando, N. S.; Bredenbeek, P. J.; Ninaber, D. K.; Wang, Y.; Hiemstra, P. S.; Snijder, E. J.; Hemert, M. J. van. Suramin Inhibits SARS-CoV-2 Infection in Cell Culture by Interfering with Early Steps of the Replication Cycle. *Antimicrobial Agents and Chemotherapy* **2020**, DOI: 10.1128/AAC.00900-20
12. Ko, W.-C.; Rolain, J.-M.; Lee, N.-Y.; Chen, P.-L.; Huang, C.-T.; Lee, P.-I.; Hsueh, P.-R. Arguments in Favour of Remdesivir for Treating SARS-CoV-2 Infections. *Int. J. Antimicrob. Agents* **2020**, *55* (4), 105933, DOI: 10.1016/j.ijantimicag.2020.105933
13. Seifert, M.; Bera, S. C.; Van Nies, P.; Kirchdoerfer, R. N.; Shannon, A.; Meng, X.; Xia, H.; Wood, J. M.; Harris, L. D.; Papini, F. S.; Arnold, J. J.; Almo, S.; Grove, T. L.; Shi, P.-Y.; Xiang, Y.; Canard, B.; Depken, M.; Cameron, C. E.; Dulin, D. Inhibition of SARS-CoV-2 polymerase by nucleotide analogs from a single-molecule perspective. *eLife* **2021**, *10*, e70968, DOI: 10.7554/eLife.70968
14. Carlton-Smith, C.; Elliott, R. M. Viperin, MTAP44, and Protein Kinase R Contribute to the Interferon-Induced Inhibition of Bunyamwera Orthobunyavirus Replication. *J. Virol.* **2012**, *86* (21), 11548–11557, DOI: 10.1128/JVI.01773-12

15. Chin, K.-C.; Cresswell, P. Viperin (Cig5), an IFN-Inducible Antiviral Protein Directly Induced by Human Cytomegalovirus. *Proc. Natl. Acad. Sci.* U.S.A. **2001**, *98* (26), 15125–15130, DOI: 10.1073/pnas.011593298
16. Wood, J. M.; Evans, G. B.; Grove, T. L.; Almo, S. C.; Cameron, S. A.; Furneaux, R. H.; Harris, L. D. Chemical Synthesis of the Antiviral Nucleotide Analogue DdhCTP. *J. Org. Chem.* **2021**, 86 (13), 8843–8850, DOI: 10.1021/acs.joc.1c00761
17. Gizzi, A. S.; Grove, T. L.; Arnold, J. J.; Jose, J.; Jangra, R. K.; Garforth, S. J.; Du, Q.; Cahill, S. M.; Dulyaninova, N. G.; Love, J. D.; Chandran, K.; Bresnick, A. R.; Cameron, C. E.; Almo, S. C. A Naturally Occurring Antiviral Ribonucleotide Encoded by the Human Genome. *Nature* **2018**, 558 (7711), 610–614, DOI: 10.1038/s41586-018-0238-4
18. Shi, W.; Ye, H.-Q.; Deng, C.-L.; Li, R.; Zhang, B.; Gong, P. A Nucleobase-Binding Pocket in a Viral RNA-Dependent RNA Polymerase Contributes to Elongation Complex Stability. *Nucleic Acids Research* **2020**, *48* (3), 1392–1405, DOI: 10.1093/nar/gkz1170
19. Arba, M.; Wahyudi, S. T.; Brunt, D. J.; Paradis, N.; Wu, C. Mechanistic Insight on the Remdesivir Binding to RNA-Dependent RNA Polymerase (RdRp) of SARS-Cov-2. *Comput. Biol. Med.* **2021**, *129*, 104156, DOI: 10.1016/j.compbiomed.2020.104156
20. Denison, M. R.; Graham, R. L.; Donaldson, E. F.; Eckerle, L. D.; Baric, R. S. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA biology* **2011**, *8*(2), 270-279, DOI: 10.4161/rna.8.2.15013
21. Gordon, C. J.; Tchesnokov, E. P.; Woolner, E.; Perry, J. K.; Feng, J. Y.; Porter, D. P.; Götte, M. Remdesivir Is a Direct-Acting Antiviral That Inhibits RNA-Dependent RNA Polymerase from Severe Acute Respiratory Syndrome Coronavirus 2 with High Potency. *J. Biol. Chem.* **2020**, 295 (20), 6785–6797, DOI: 10.1074/jbc.RA120.013679
22. RECOVERY Collaborative Group; Horby, P.; Lim, W. S.; Emberson, J. R.; Mafham, M.; Bell, J. L.; Linsell, L.; Staplin, N.; Brightling, C.; Ustianowski, A.; Elmahi, E.; Prudon, B.; Green, C.; Felton, T.; Chadwick, D.; Rege, K.; Fegan, C.; Chappell, L. C.; Faust, S. N.; Jaki, T.; Jeffery, K.; Montgomery, A.; Rowan, K.; Juszczak, E.; Baillie, J. K.; Haynes, R.; Landray, M. J. Dexamethasone in Hospitalized Patients with Covid-19. *N. Engl. J. Med.* **2021**, 384 (8), 693–704, DOI: 10.1056/NEJMoa2021436
23. Naydenova, K.; Muir, K. W.; Wu, L.-F.; Zhang, Z.; Coscia, F.; Peet, M. J.; Castro-Hartmann, P.; Qian, P.; Sader, K.; Dent, K.; Kimanius, D.; Sutherland, J. D.; Löwe, J.; Barford, D.; Russo, C. J.

Structure of the SARS-CoV-2 RNA-Dependent RNA Polymerase in the Presence of Favipiravir-RTP. *Proc. Natl. Acad. Sci.* U.S.A. **2021**, *118* (7), DOI: 10.1073/pnas.2021946118

24. Elfiky, A. A. SARS-CoV-2 RNA Dependent RNA Polymerase (RdRp) Targeting: An in Silico Perspective. *J. Biomol. Struct. Dyn.* **2021**, *39* (9), 3204–3212, DOI: 10.1080/07391102.2020.1761882
25. Elfiky, A. A. Anti-HCV, Nucleotide Inhibitors, Repurposing against COVID-19. *Life Sci* **2020**, *248*, 117477, DOI: 10.1016/j.lfs.2020.117477
26. Kato, K.; Honma, T.; Fukuzawa, K. Intermolecular Interaction among Remdesivir, RNA and RNA-Dependent RNA Polymerase of SARS-CoV-2 Analyzed by Fragment Molecular Orbital Calculation. *J. Mol. Graph. Model.* **2020**, *100*, 107695, DOI: 10.1016/j.jmgm.2020.107695
27. Zhang, L.; Zhou, R. Structural Basis of the Potential Binding Mechanism of Remdesivir to SARS-CoV-2 RNA-Dependent RNA Polymerase. *J. Phys. Chem. B.* **2020**, *124* (32), 6955–6962, DOI: 10.1021/acs.jpcb.0c04198
28. Palermo, G.; Cavalli, A.; Klein, M. L.; Alfonso-Prieto, M.; Dal Peraro, M.; De Vivo, M. Catalytic metal ions and enzymatic processing of DNA and RNA. *Acc. Che. Res.* **2015**, *48 (2)*, 220-228. DOI: 10.1021/ar500314j
29. Lehmann, K. C.; Gulyaeva, A.; Zevenhoven-Dobbe, J. C.; Janssen, G. M. C.; Ruben, M.; Overkleeft, H. S.; van Veelen, P. A.; Samborskiy, D. V.; Kravchenko, A. A.; Leontovich, A. M.; Sidorov, I. A.; Snijder, E. J.; Posthuma, C. C.; Gorbalenya, A. E. Discovery of an Essential Nucleotidylating Activity Associated with a Newly Delineated Conserved Domain in the RNA Polymerase-Containing Protein of All Nidoviruses. *Nucleic Acids Res.* **2015**, *43* (17), 8416–8434, DOI: 10.1093/nar/gkv838
30. McDonald, S. M. RNA Synthetic Mechanisms Employed by Diverse Families of RNA Viruses. *Wiley Interdiscip. Rev.: RNA* **2013**, *4* (4), 351–367, DOI: 10.1002/wrna.1164
31. Geronimo, I.; Vidossich, P.; Donati, E.; De Vivo, M. Computational investigations of polymerase enzymes: Structure, function, inhibition, and biotechnology. *Wiley Interdiscip. Rev.: Computational Molecular Science* **2021**,*11*(6), e1534, DOI: 10.1002/wcms.1534
32. Itoh, S. G.; Tanimoto, S.; Okumura, H. Dynamic properties of SARS-CoV and SARS-CoV-2 RNA-dependent RNA polymerases studied by molecular dynamics simulations. *Chemical Physics Letters* **2021**, *778*, 138819, DOI: 10.1016/j.cplett.2021.138819
33. Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; Ge, J.; Zheng, L.; Zhang, Y.; Wang, H.; Zhu, Y.; Zhu, C.; Hu, T.; Hua, T.; Zhang, B.; Yang, X.;

Li, J.; Yang, H.; Liu, Z.; Xu, W.; Guddat, L. W.; Wang, Q.; Lou, Z.; Rao, Z. Structure of the RNA-Dependent RNA Polymerase from COVID-19 Virus. *Science* **2020**, *368* (6492), 779–782, DOI: 10.1126/science.abb7498

34. Kirchdoerfer, R. N.; Ward, A. B. Structure of the SARS-CoV Nsp12 Polymerase Bound to Nsp7 and Nsp8 Co-Factors. *Nat. Commun.* **2019**, *10* (1), 2342, DOI: 10.1038/s41467-019-10280-3
35. Shannon, A.; Selisko, B.; Le, N.-T.-T.; Huchting, J.; Touret, F.; Piorkowski, G.; Fattorini, V.; Ferron, F.; Decroly, E.; Meier, C.; Coutard, B.; Peersen, O.; Canard, B. Rapid Incorporation of Favipiravir by the Fast and Permissive Viral RNA Polymerase Complex Results in SARS-CoV-2 Lethal Mutagenesis. *Nat. Commun.* **2020**, *11* (1), 4682, DOI: 10.1038/s41467-020-18463-z
36. Chien, M.; Anderson, T. K.; Jockusch, S.; Tao, C.; Li, X.; Kumar, S.; Russo, J. J.; Kirchdoerfer, R. N.; Ju, J. Nucleotide Analogues as Inhibitors of SARS-CoV-2 Polymerase, a Key Drug Target for COVID-19. *J. Proteome Res.* **2020**, *19* (11), 4690–4697, DOI: 10.1021/acs.jproteome.0c00392
37. Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017; University of California, San Francisco, **2017**.
38. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65* (3), 712-25, DOI: 10.1002/prot.21123
39. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280, DOI: 10.1021/j100142a004
40. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res* **2012**, *40*, W537-541, DOI: 10.1093/nar/gks375

41. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.* **2015**, 11, 8, 3696–3713, DOI: 10.1021/acs.jctc.5b00255
42. Zgarbová, M.; Otyepka, M.; Šponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurečka, P. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.* **2011**, *7* (9), 2886–2902, DOI: 10.1021/ct200162x
43. Yu, Z.; Li, P.; Merz, K. M. Extended Zinc AMBER Force Field (EZAFF). *J. Chem. Theory Comput.* **2018**, *14* (1), 242–254, DOI: 10.1021/acs.jctc.7b00773
44. Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9* (6), 2733–2748, DOI: 10.1021/ct400146w
45. Genna, V., Donati, E., De Vivo, M. The catalytic mechanism of DNA and RNA polymerases. *ACS Catalysis* **2018**, *8*(12), 11103-11118, DOI: 10.1021/acscatal.8b03363
46. Geronimo, I., Vidossich, P., De Vivo, M. Local Structural Dynamics at the Metal-Centered Catalytic Site of Polymerases is Critical for Fidelity. *ACS Catalysis* **2021**, *11*(22), 14110-14121, DOI: 10.1021/acscatal.1c03840
47. Vassetti, D.; Pagliai, M.; Procacci, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1983–1995, DOI: 10.1021/acs.jctc.8b01039
48. Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. van; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1998**, *81* (8), 3684, DOI: 10.1063/1.448118
49. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092, DOI: 10.1063/1.464397
50. Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095, DOI: 10.1021/ct400341p
51. Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637, DOI: 10.1002/bip.360221211

52. Prejanò, M.; Romeo, I.; Sgrizzi, L.; Russo, N.; Marino, T. Why hydroxy-proline improves the catalytic power of the peptidoglycan N-deacetylase enzyme: insight from theory. *Phys. Chem. Chem. Phys.* **2019**, *21*, 23338-23345. DOI: 10.1039/C9CP03804C
53. Prejanò, M.; Romeo, I.; La Serra, M. A.; Russo, N.; Marino, T. Computational Study Reveals the Role of Water Molecules in the Inhibition Mechanism of LAT1 by 1,2,3-Dithiazoles. *J. Chem. Inf. Model.* **2021**, *61(12)*, 5883–5892. DOI: 10.1021/acs.jcim.1c01012
54. Prejanò, M.; Romeo, I.; Russo, N.; Marino, T. On the Catalytic Activity of the Engineered Coiled-Coil Heptamer Mimicking the Hydrolase Enzymes: Insights from a Computational Study. *Int. J. Mol. Sci.* **2020**, *21(12)*, 4551. DOI: 10.3390/ijms21124551
55. Xu, X.; Zhang, L.; Chu, J. T. S.; Wang, Y.; Chin, A. W. H.; Chong, T. H.; Dai, Z.; Poon, L. L. M.; Cheung, P. P.-H.; Huang, X. A novel mechanism of enhanced transcription activity and fidelity for influenza A viral RNA-dependent RNA polymerase. *Nucleic acids research* **2021**, *49*(15), 8796-8810, DOI: 10.1093/nar/gkab660
56. Wilson, K. A.; Fernandes, P. A.; Ramos, M. J.; Wetmore, S. D. Exploring the identity of the general base for a DNA polymerase catalyzed reaction using QM/MM: the case study of human translesion synthesis polymerase η. *ACS Catalysis* **2019**, *9*(3), 2543-2551, DOI: 10.1021/acscatal.8b04889
57. Ng, K. K. S.; Arnold, J. J.; Cameron, C. E. Structure-function relationships among RNA-dependent RNA polymerases. *RNA interference* **2008**, 137-156, DOI: 10.1007/978-3-540-75157-1_7
58. Sgrignani, J.; Magistrato, A. The structural role of Mg^{2+} ions in a class I RNA polymerase ribozyme: a molecular simulation study. *The Journal of Physical Chemistry B* **2012**, *116*(7), 2259-2268, DOI: 10.1021/jp206475d
59. Boero, M.; Tateno, M.; Terakura, K.; Oshiyama, A. J. Double-metal-ion/single-metal-ion mechanisms of the cleavage reaction of ribozymes: First-principles molecular dynamics simulations of a fully hydrated model system. *Chem. Theory Comput.* **2005**, 1, 925– 934, DOI: 10.1021/ct050066q
60. Chaudret, R.; Piquemal, J.-P.; Andres Cisneros, G. Correlation between electron localization and metal ion mutagenicity in DNA synthesis from QM/MM calculations. *Phys. Chem. Chem. Phys.* **2011**, 13, 11239–11247, DOI: 10.1039/C0CP02550J
61. Gleghorn, M. L.; Davydova, E. K.; Basu, R.; Rothman-Denes, L. B.; Murakami, K. S. X-ray crystal structures elucidate the nucleotidyl transfer reaction of transcript initiation using two

nucleotides. *Proc. Natl. Acad. Sci.* U.S.A. **2011**, 108, 3566–3571, DOI: /10.1073/pnas.1016691108

62. Yang, W.; Lee, J. Y.; Nowotny, M. Making and breaking nucleic acids: two-Mg^{2+}-ion catalysis and substrate specificity. *Mol. Cell.* **2006**, 22(1): 5-13, DOI: 10.1016/j.molcel.2006.03.013
63. Yoon, H., Warshel, A. Simulating the fidelity and the three Mg mechanism of pol η and clarifying the validity of transition state theory in enzyme catalysis. *Proteins: Structure, Function, and Bioinformatics* **2017**, *85*(8), 1446-1453, DOI: 10.1002/prot.25305
64. Otsuka, J.; Kikuchi, N.; Kojima, S. Similarity relations of DNA and RNA polymerases investigated by the principal component analysis of amino acid sequences. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* **1999**, *1434*(2), 221-247, DOI: 10.1016/S0167-4838(99)00187-9
65. Yin, Y. W.; Steitz, T. A. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **2004**, *116*(3), 393-404, DOI: 10.1016/S0092-8674(04)00120-5
66. Hub, J. S.; De Groot, B. L.; Van Der Spoel, D. g_wham- A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J. Chem. Theor. Comput.* **2010** *6*(12), 3713-3720, DOI: 10.1021/ct100494z