

 Open access • Posted Content • DOI:10.1101/2021.05.11.443555

Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis — [Source link](#)

[Florian Kabinger](#), [Carina Stiller](#), [Jana Schmitzová](#), [Christian Dienemann](#) ...+4 more authors

Institutions: [Max Planck Society](#), [University of Würzburg](#), [University of Göttingen](#)

Published on: 11 May 2021 - [bioRxiv](#) (Cold Spring Harbor Laboratory)

Topics: [Polymerase](#), [RNA](#), [RNA polymerase](#), [Base pair](#) and [Mutagenesis](#)

Related papers:

- [SARS-CoV-2 RdRp is a versatile enzyme with proofreading activity and ability to incorporate NHC into RNA by using diphosphate form molnupiravir as a substrate](#)
- [Denovo designing, retro-combinatorial synthesis, and molecular dynamics analysis identify novel antiviral VTRM1.1 against RNA-dependent RNA polymerase of SARS CoV2 virus.](#)
- [RNA-dependent RNA polymerase \(RdRp\) inhibitors: The current landscape and repurposing for the COVID-19 pandemic.](#)
- [RNA-dependent RNA polymerase: Structure, mechanism, and drug discovery for COVID-19.](#)
- [Allosteric Activation of SARS-CoV-2 RNA-Dependent RNA Polymerase by Remdesivir Triphosphate and Other Phosphorylated Nucleotides.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/mechanism-of-molnupiravir-induced-sars-cov-2-mutagenesis-3y4zxhnwhd>

Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis

Florian Kabinger^{1*}, Carina Stiller^{2*}, Jana Schmitzová^{1*}, Christian Dienemann¹, Hauke S. Hillen^{3,4}, Claudia Höbartner^{2,#} and Patrick Cramer^{1,#}

¹Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Göttingen, Germany. ²Universität Würzburg, Lehrstuhl für Organische Chemie I, Am Hubland, 97074 Würzburg, Germany. ³Department of Cellular Biochemistry, University Medical Center Göttingen, Göttingen, Germany. ⁴Max Planck Institute for Biophysical Chemistry, Research Group Structure and Function of Molecular Machines, Am Fassberg 11, 37077 Göttingen, Germany. *These authors contributed equally.

Molnupiravir is an orally available antiviral drug candidate that is in phase III trials for the treatment of COVID-19 patients^{1,2}. Molnupiravir increases the frequency of viral RNA mutations^{3,4} and impairs SARS-CoV-2 replication in animal models^{4,6} and in patients². Here we establish the molecular mechanisms that underlie molnupiravir-induced RNA mutagenesis by the RNA-dependent RNA polymerase (RdRp) of the coronavirus SARS-CoV-2. Biochemical assays show that the RdRp readily uses the active form of molnupiravir, β -D-N4-hydroxycytidine (NHC) triphosphate, as a substrate instead of CTP or UTP. Incorporation of NHC monophosphate into nascent RNA does not impair further RdRp progression. When the RdRp uses the resulting RNA as a template, NHC directs incorporation of either G or A, leading to mutated RNA products. Structural analysis of RdRp-RNA complexes containing mutagenesis products shows that NHC can form stable base pairs with either G or A in the RdRp active center, explaining how the polymerase escapes proofreading and synthesizes mutated RNA. This two-step mutagenesis mechanism likely applies to various viral polymerases and can explain the broad-spectrum antiviral activity of molnupiravir.

Coronaviruses use an RNA-dependent RNA polymerase (RdRp) for the replication and transcription of their RNA genome⁷⁻¹¹. The RdRp is an important target for the development of antiviral drugs against coronaviruses¹²⁻¹⁵. Structures of the RdRp were reported for SARS-CoV-1¹⁶ and for SARS-CoV-2¹⁷⁻²¹ and provided insights into the mechanisms of RNA-dependent RNA synthesis²². The structures also enable mechanistic studies that can rationalize the molecular processes underlying the antiviral activity of compounds targeting the RdRp.

Antiviral drugs often target viral polymerases and function as nucleoside analogs that terminate RNA chain elongation. However, such chain-terminating antivirals are generally not effective against SARS-CoV-2 because coronaviruses carry an exonucleolytic proofreading activity that can remove misincorporated nucleotides from the nascent RNA 3'-end²³⁻²⁵. The nucleoside analogue remdesivir can circumvent proofreading because its incorporation does not terminate elongation but only stalls RdRp after the addition of three more nucleotides^{19,26-29}. Remdesivir was the first FDA-approved drug for the treatment of COVID-19 patients³⁰⁻³³, but its effectiveness is disputed³⁴, emphasizing the need to develop new antiviral drugs.

Another promising drug candidate for the treatment of COVID-19 patients is molnupiravir (or EIDD-2801), which also targets the RdRp of SARS-CoV-2. Molnupiravir is an isopropylester prodrug of the nucleoside analogue β -D-N4-hydroxycytidine (NHC or EIDD-1931)^{3,35}. Molnupiravir interferes with the replication of various viruses^{3,35-40} including SARS-CoV-2^{4,41}. Molnupiravir inhibits SARS-CoV-2 replication in human lung tissue⁵, blocks SARS-CoV-2 transmission in ferrets⁶, and reduces SARS-CoV-2

RNA in patients². In contrast to approved drugs such as remdesivir that are administered by infusion, molnupiravir is orally available. Molnupiravir has been tested in phase I trials⁴² for safety, tolerability and pharmacokinetics, and phase II/III studies are currently ongoing² (NCT04405739, NCT04405570, NCT04575597). Available data suggest that molnupiravir acts as a mutagenizing agent that causes an 'error catastrophe' during viral replication^{3,4,43}. Indeed, NHC can introduce mutations into viral RNA, as shown for Venezuelan equine encephalitis virus⁴⁴. Also, sequencing of influenza virus populations indicated that NHC caused G-to-A and C-to-U transitions in viral RNA³, and the same transitions were found for SARS-CoV-2⁴.

Despite this progress, a systematic biochemical and structural analysis of molnupiravir- or NHC-induced RNA mutagenesis by viral RNA polymerases is lacking. Here we quantify the effects of molnupiravir/NHC on the RNA synthesis activity of SARS-CoV-2 RdRp using a purified biochemical system and defined synthetic RNAs. Together with structural analysis we establish the molecular mechanism of molnupiravir-induced RNA mutagenesis. Our results provide detailed insights into the mechanism of action of molnupiravir, which is entirely distinct from that of remdesivir or chain-terminating nucleoside analogues.

RESULTS

SARS-CoV-2 RdRp readily incorporates NHC into RNA

We first tested whether purified SARS-CoV-2 RdRp can use the active form of molnupiravir, NHC triphosphate ('MTP', **Fig. 1a, b**), as a substrate for

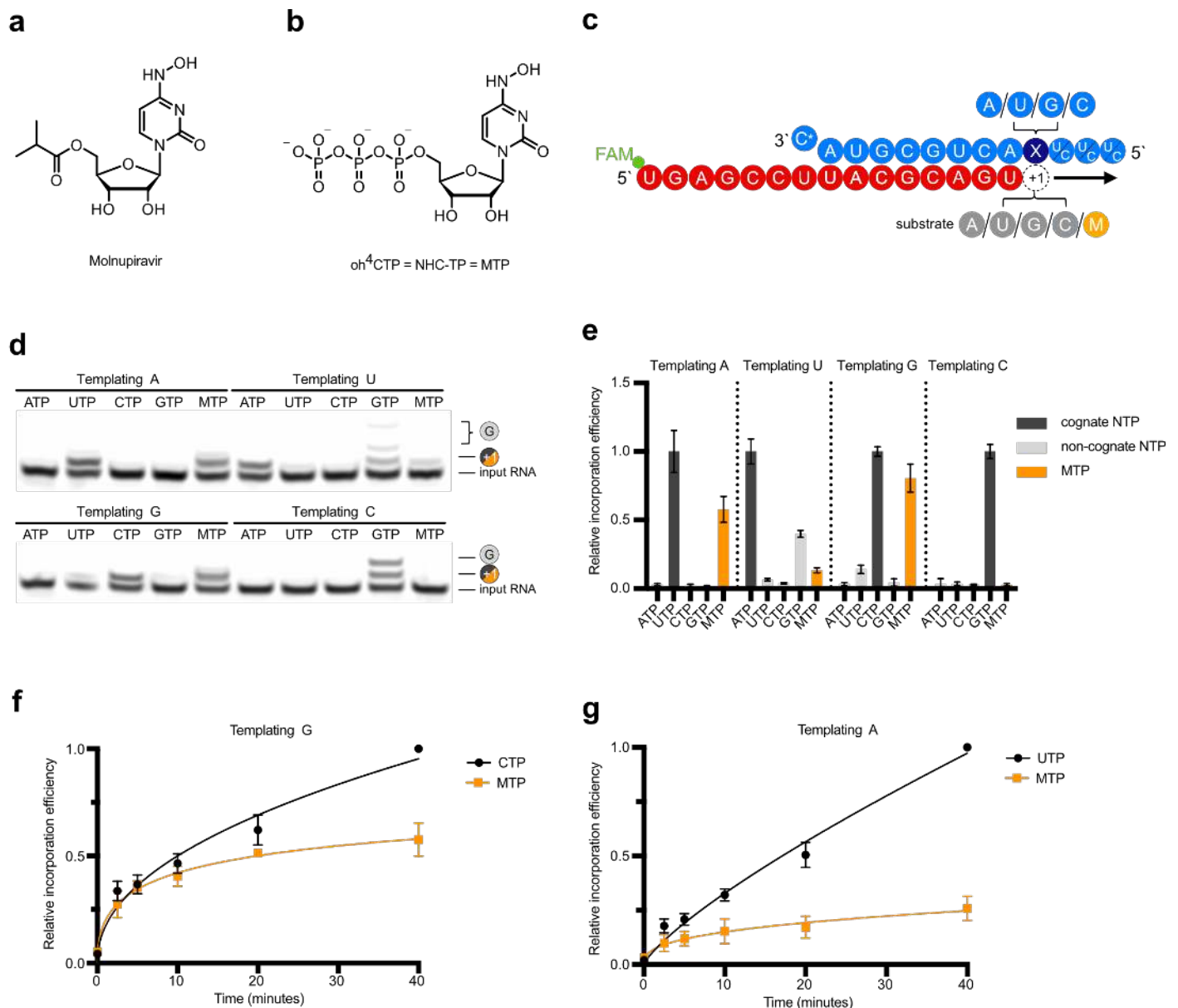


Figure 1 | RdRp incorporates NHC opposite of G and A in the template.

a, Chemical structure of molnupiravir.

b, Chemical structure of NHC triphosphate (MTP).

c, RNA template-product duplex. The direction of RNA extension is shown. Color of depicted circles indicates experimental design: blue, RNA template strand; dark blue, +1 templating nucleotide; red, RNA product strand; gray, NTP substrate; orange: MTP. The 5'-end of the RNA product contains a FAM fluorescent label. C* in 5' end of the template denotes dideoxy-C (ddC).

d, NHC monophosphate is incorporated into growing RNA instead of C or U when G or A are present in the template +1 position.

e, Quantification of nucleotide incorporation efficiency relative to the cognate NTP (dark grey). Non-cognate NTPs and MTP are depicted in light grey and orange, respectively. Standard deviations are shown. Source data are provided as a Source Data file.

f, Quantification of time-dependent M incorporation opposite a templating G residue after triplicate measurements. Incorporation efficiency is calculated relative to cognate C incorporation. Standard deviations are shown.

g, Quantification of time-dependent M incorporation opposite a templating A residue after triplicate measurements. Incorporation efficiency is calculated relative to cognate U incorporation. Standard deviations are shown.

RNA synthesis. We conducted RNA elongation assays in a defined biochemical system using recombinant RdRp and synthetic RNA template-product duplexes (Methods). We used four different RNA duplexes that differed at position +1 of the template strand (**Supplementary Table 1**), which directs binding of the incoming nucleoside triphosphate (NTP) substrate (**Fig. 1c**). The RNA product strand contained a fluorescent label at its 5'-end that allowed us to monitor and quantify RNA elongation.

When the nucleotides G or A were present at the RNA template position +1, NHC monophosphate ('M') was readily incorporated instead of C or U,

respectively (**Fig. 1d, e**). Time-dependent RNA elongation experiments showed that M was slightly less efficiently incorporated as the cognate nucleotide C, but less efficiently than the cognate nucleotide U (**Fig. 1f, g**). These results can be explained by base pairing of an incoming MTP substrate with either G or A in the RNA template strand. Consistent with this model, NHC adopts different tautomeric forms⁴⁵ that were predicted to allow for base pairing with either G or A⁴⁶.

RdRp does not stall after NHC incorporation

We next tested whether the incorporation of NHC monophosphate (M) into nascent RNA interferes

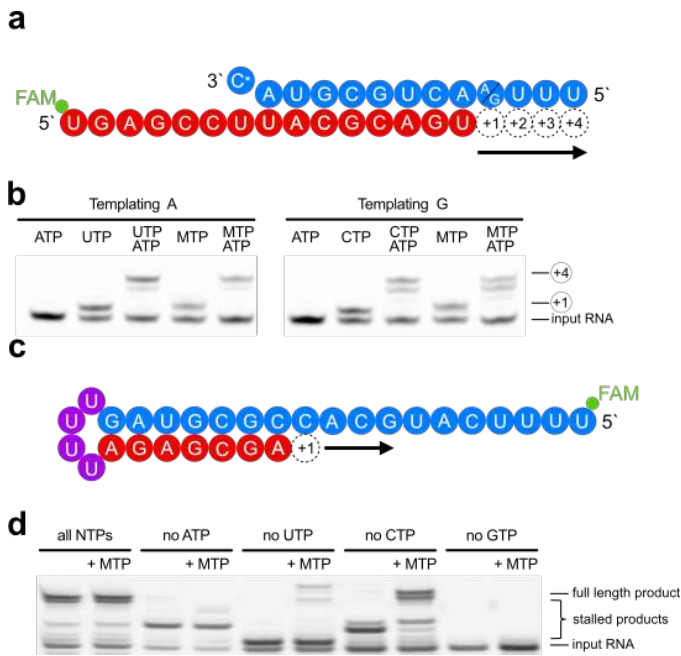


Figure 2 | NHC incorporation does not stall SARS-CoV-2 RdRp.

a, RNA template-product duplex (as in Fig. 1c) allows for RNA extension by four nucleotides. The direction of RNA extension is indicated. The 5'-end of the RNA product contains a FAM fluorescent label. C* indicates dideoxy-C (ddC).

b, RNA elongation to the end of the template in panel a is possible when MTP replaces either CTP or UTP in the presence of ATP. Source data are provided as a Source Data file.

c, RNA template-product hairpin duplex allowing for RNA extension by 11 nucleotides.

d, RNA elongation stalls at expected positions when the cognate NTP is withheld from the reaction. Extension to the end of the template is possible when MTP replaces either CTP or UTP in the presence of other substrate NTPs, showing that incorporation of M does not prevent RNA extension. Note that more efficient RNA extension is seen at higher NTP/MTP concentrations, also for MTP replacing UTP (not shown). Source data are provided as a Source Data file.

with further RNA extension. We first conducted RNA elongation assays with a scaffold that allowed for RNA extension by four nucleotides (nt) (Fig. 2a). We observed that incorporation of M instead of the cognate U or C did not prevent incorporation of three subsequent nucleotides (Fig. 2b). Further, we tested RNA extension with a scaffold that allowed for the incorporation of 11 nt (Fig. 2c). Also in this case the RdRp reached the end of the template when UTP or CTP were replaced by MTP, although again incorporation of M instead of U was less efficient than incorporation of M instead of C (Fig. 2d). These results demonstrate that M incorporation into nascent RNA does not prevent further RNA elongation. Thus, longer RNA products containing M nucleotides may be synthesized by the RdRp in the presence of MTP. This posed the question what happens when M-containing RNA is used as a template in a second step.

RdRp uses NHC-containing templates to direct RNA mutagenesis

To investigate the templating properties of NHC, we prepared an M-containing RNA by solid-phase synthesis using the phosphoramidite building block M-PA, which we synthesized by the convertible nucleoside approach from a ribose-protected *O*⁴-chlorophenyluridine (Fig. 3a, Methods, Extended

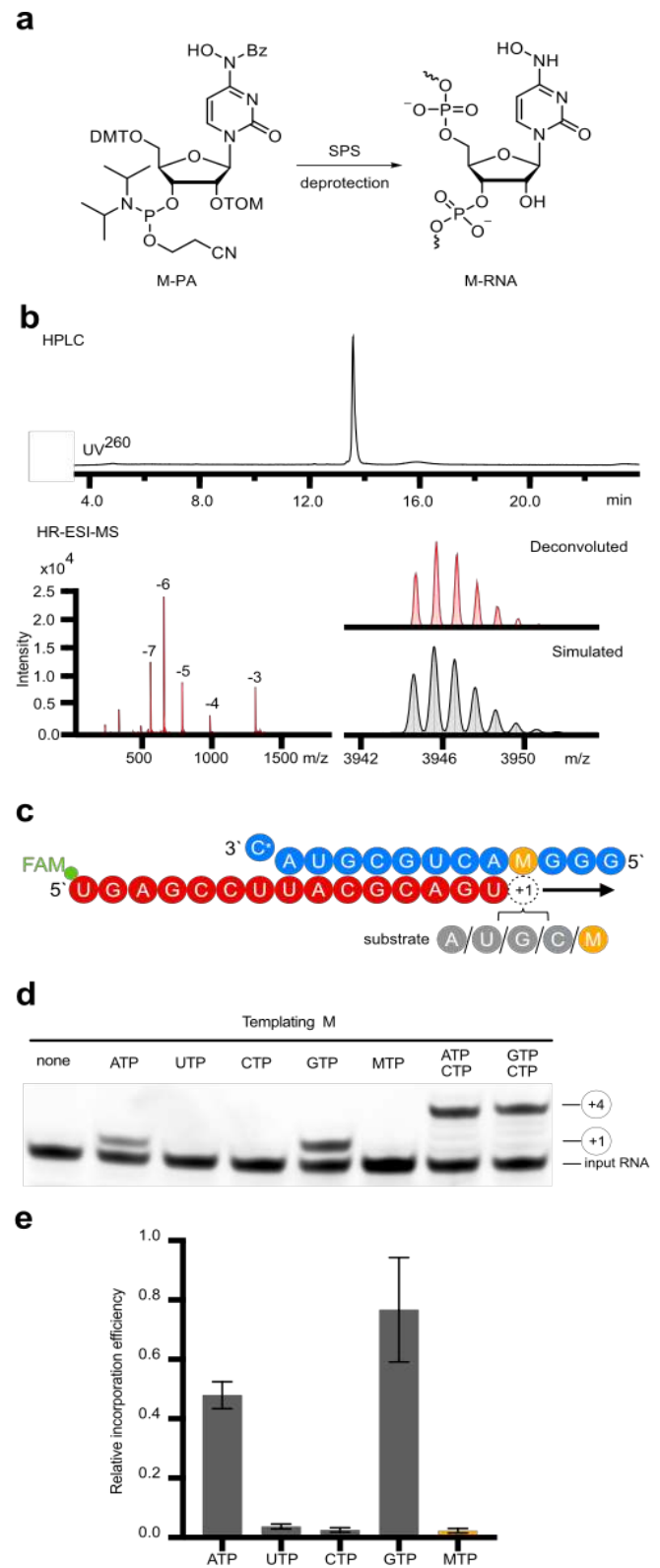


Figure 3 | NHC can direct incorporation of G and A into RNA.

a, Scheme of synthesis of RNA containing NHC monophosphate (M) at a defined position (M-RNA).

b, Analysis of M-containing RNA by denaturing HPLC confirms homogeneity of the synthetic RNA (top). HR-ESI-MS proves the presence of NHC and absence of unmodified RNA (bottom).

c, RNA template-product scaffold with M in the template position +1, where it is used by the RdRp to direct binding of the incoming NTP substrate. The 5'-end of the RNA product contains a FAM fluorescent label. C* in the 5' end of the template is dideoxy-C (ddC).

d, When present at position +1 of the template strand, M can direct the incorporation of G or A into nascent RNA, but not C or U. Source data are provided as a Source Data file.

e, Quantification of the experiment in panel d after triplicate measurements. Incorporation efficiencies are calculated relative to C incorporation opposite of templating G.

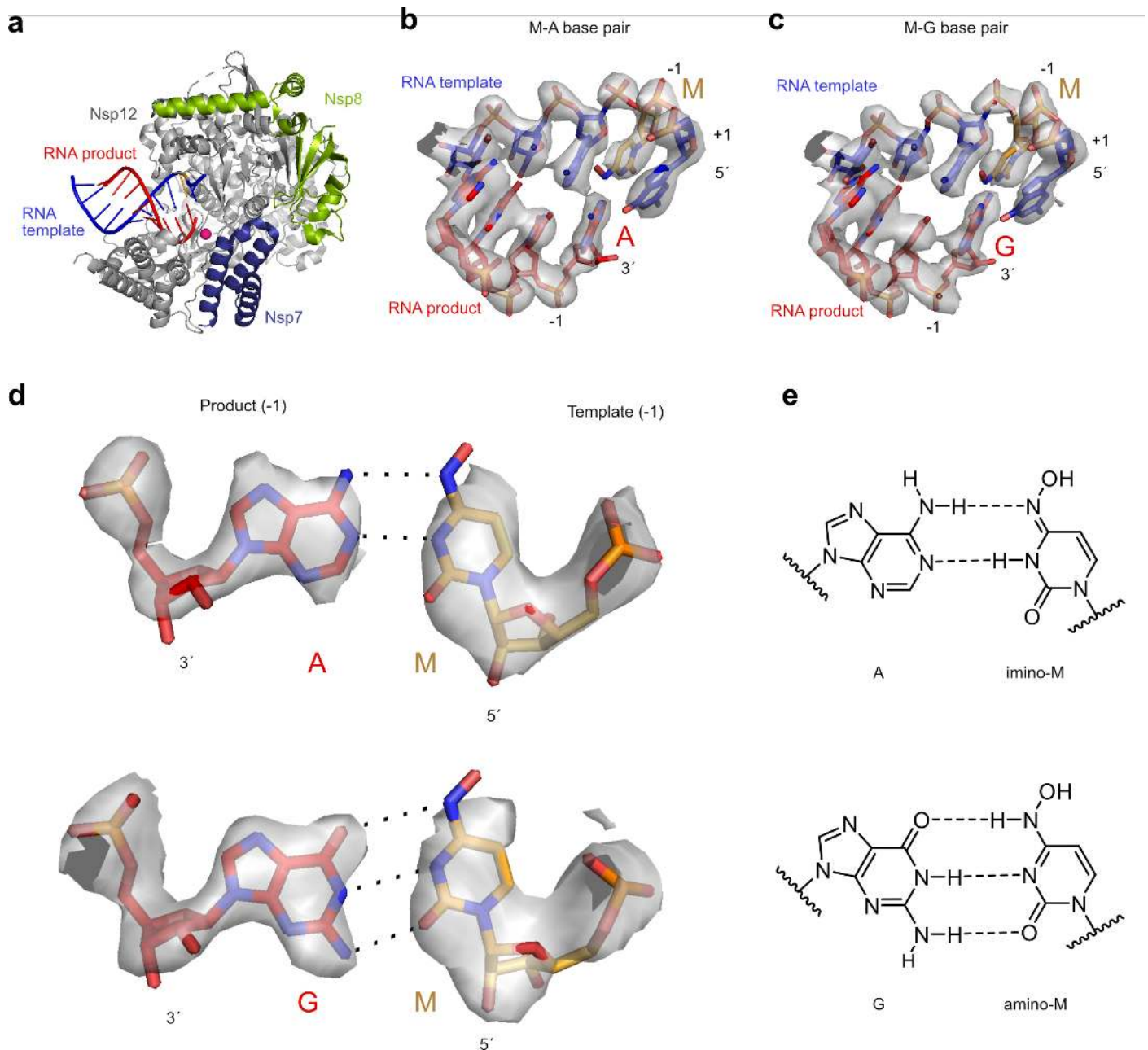


Figure 4 | Structures of RdRp-RNA product complexes after NHC-induced mutagenesis.

a, Overview of RdRp-RNA structure with M residue (orange) at position -1 in the RNA template strand. RdRp subunits nsp7, nsp8, and nsp12 are in dark blue, green and grey, respectively. The RNA template and product are in blue and red, respectively. The active site is indicated with a magenta sphere. Depicted is the structure containing the M-A base pair.

b, RNA duplex containing the M-A base pair in the RdRp active center. The +1 position (templating nucleotide, NTP substrate site) and the -1 position (post-translocation position of nascent base pair) are indicated.

c, RNA duplex containing the M-G base pair in the RdRp active center.

d, Cryo-EM density for the nascent M-A (top) and M-G (bottom) base pairs in position -1, viewed along the RNA duplex axis in the direction of RNA translocation.

e, M-A (top) and M-G (bottom) base pairing relies on different tautomeric forms of NHC45 as predicted⁴⁶.

Data Fig. 1, Supplementary information). The presence of M in the obtained RNA and RNA purity were confirmed by denaturing HPLC and HR-ESI-MS (**Fig. 3b**). The M-containing RNA oligo was annealed with a fluorescently labeled product RNA such that the M nucleotide occupied the templating position +1 (**Fig. 3c, Supplementary Table 1**).

Elongation assays showed that the M residue at the +1 position of the template strand directed incorporation of either G or A into nascent RNA, but not C or U (**Fig. 3d, e**). This can be explained by the formation of M-GTP or M-ATP base pairs in the RdRp active center. Consistent with this, thermal melting experiments with RNA duplexes containing

M-G or M-A base pairs located at terminal or internal positions showed similar RNA duplex stabilities that were slightly lower than for duplexes containing a C-G base pair (**Extended Data Fig. 2, Supplementary Table 2**). Thus, when the RdRp uses RNA containing NHC monophosphate as a template, either the correct or the incorrect nucleotide is incorporated into the RNA product, and thus mutagenesis will occur.

Structural basis of NHC-induced RNA mutagenesis

The above data indicate that the key aspect of the mutagenesis mechanism is the formation of stable

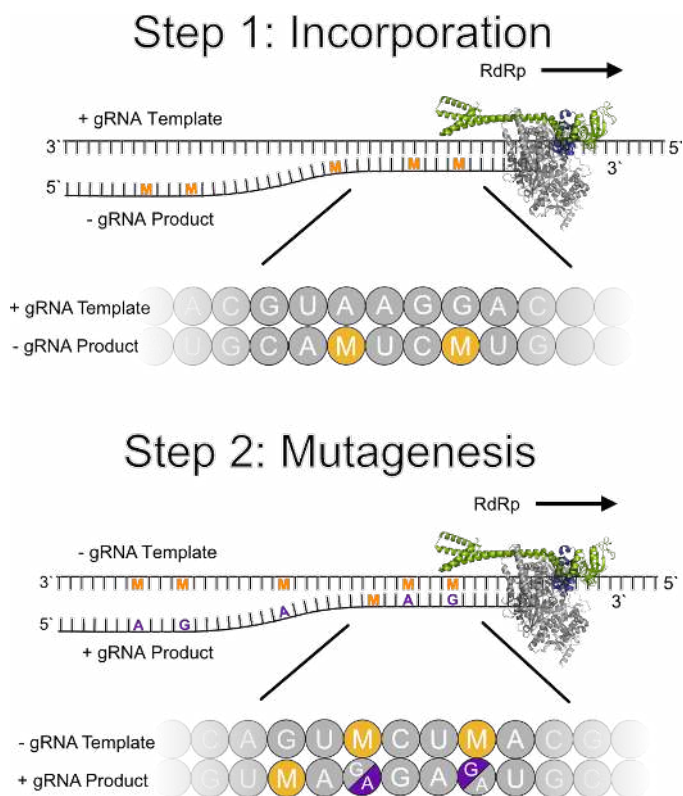


Figure 5 | Two-step model of molnupiravir-induced RNA mutagenesis.

In the presence of NTPs and MTP, M nucleotides can be incorporated by SARS-CoV-2 RdRp instead of C or U into the negative strand genomic (-gRNA) or subgenomic RNA (-sgRNA) during copying of the positive strand genomic RNA template (+gRNA). The obtained M-containing negative-strand RNAs can then be used as a template for the production of mutagenized +gRNA and positive strand subgenomic mRNA (+sgmRNA). These RNA products are predicted to be mutated and not to support formation of functional viruses. RNA of random sequence is shown with M and mutated residues indicated as orange and violet letters, respectively.

M-G and M-A base pairs in the RdRp active center. To investigate this, we solved two structures of RdRp-RNA complexes that correspond to mutagenesis products after M-templated incorporation of either G or A (Methods). We formed RdRp-RNA complexes containing M in the template strand and either G or A at the 3'-end of the product strand. This was predicted to result in the formation of nascent M-G or M-A base pairs in position -1, which is occupied after successful M-templated nucleotide incorporation and RdRp translocation. We prepared RNA duplex scaffolds with M-containing oligonucleotides (Extended Data Fig. 3), formed RdRp-RNA scaffold complexes and subjected these to cryo-EM analysis as described²⁰. We indeed obtained RdRp-RNA structures that contained either a M-A or a M-G base pair at position -1 (Fig. 4, Table 1). The structures showed an overall resolution of 3.3 Å and 3.2 Å, respectively, with the active center region resolved at ~2.9 Å in both cases (Extended Data Fig. 4). As expected from the scaffold design, the structures showed the post-translocation state with a free NTP-binding site at position +1 (Fig. 4b, c). Comparison of the two structures with each other and with our original RdRp-RNA structure²⁰ and remdesivir-containing RdRp-RNA structures²⁸ did not reveal major differences, neither in the protein subunits

nor in the nucleic acids, except that the protruding, second turn of RNA and the sliding poles of the nsp8 subunits were poorly ordered and not retained in the final model.

The cryo-EM densities at the -1 position of the structures could readily be interpreted with the expected M-A or M-G base pairs (Fig. 4b, c). The densities were so detailed that we could clearly distinguish G and A bases (Fig. 4d), and were consistent with the proposed base pairing⁴⁶ that is enabled by different tautomeric forms of NHC⁴⁵ (Fig. 4e). However, the observed hydrogen bonding geometries were not optimal, possibly explaining our biochemical observations that suggest that M mimics C or U well, but not perfectly (Figs. 1, 2). These results represent the first direct visualization of NHC in a polymerase enzyme and show that stable M-G and M-A base pairs can be formed and accommodated in the RdRp active center, readily explaining our biochemical results.

DISCUSSION

Our systematic biochemical analysis suggests a two-step model for the mechanism of molnupiravir-induced coronavirus RNA mutagenesis (Fig. 5). When the molnupiravir prodrug enters the cell, it is converted to NHC triphosphate (MTP), which can be used by the RdRp of SARS-CoV-2 as a substrate instead of CTP or UTP. Therefore, in a first step, the RdRp is predicted to frequently incorporate M instead of C or U when it uses the positive strand genomic RNA (+gRNA) as a template to synthesize negative strand genomic (-gRNA) and subgenomic RNA (-sgRNA). In a second step, the resulting M-containing RNA can be used as a template for the synthesis of +gRNA or positive strand subgenomic mRNA (+sgmRNA). The presence of M in the -gRNA then leads to mutations in the positive strand RNA products, which do not support formation of intact new viruses, as predicted by the 'error catastrophe' model^{3,4,43}.

Our structural studies confirm the key aspect of this model, namely the formation of stable M-G and M-A base pairs in the RdRp active center. These base pairs mimic natural C-G and U-A base pairs, respectively, and do not impair RdRp progression. This antiviral mechanism is conceptually similar to the recently suggested mutagenesis mode of action of favipiravir^{47,48}, but is entirely distinct from that of remdesivir, which impairs RdRp progression²⁸. However, like remdesivir, molnupiravir escapes viral RNA proofreading because M incorporation and M-directed misincorporation are apparently not recognized by the viral exonuclease^{23,24}. Such proofreading escape may also be due to the stability of the M-G and M-A base pairs that are predicted not to induce or favor backtracking of RdRp, which is likely required for exposing the RNA 3'-end to the proofreading exonuclease^{49,50}.

Finally, the two-step model can explain how molnupiravir or NHC monophosphate leads to RNA mutagenesis by polymerases of other viruses. For influenza, a possible two-step mutagenesis mechanism had been inferred from sequencing a

molnupiravir-experienced virus population³. Also consistent with the model, MTP does not inhibit RNA synthesis by hepatitis C polymerase⁴⁰ or the RdRp of respiratory syncytial virus³⁶. Further, the reverse transcriptase of human immunodeficiency virus can incorporate G or A opposite of NHC that is located in the template⁵¹. The two-step mutagenesis model resides on the base-pairing properties of NHC that we structurally defined here and can explain why molnupiravir and NHC exhibit broad-spectrum antiviral activity against a wide variety of RNA viruses. In the future, it will be important to characterize the effects of molnupiravir and NHC on cellular polymerases and understand possible side effects of the antiviral agent.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Protein preparation

Preparation of SARS-CoV-2 RdRp, composed of nsp12, nsp7 and two copies of the nsp8 subunits, was carried out as described²⁰, with some modifications. Nsp12 protein was expressed from pFastBac vector 438C (Addgene #154759) in Hi5 insect cells. Production of the bacmid, V0 and V1 viruses was carried out as described²⁰. 60 hours after transfection with corresponding V1 virus, the cells producing nsp12 with a 6xHis-MBP N-terminal tag were harvested by centrifugation (3000 rpm, 10 min at 4°C) and lysed by sonication in lysis buffer A1 (400 mM NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 30 mM imidazole pH 8.0, 5 mM β-mercaptoethanol, 0.284 μg ml⁻¹ leupeptin, 1.37 μg ml⁻¹ pepstatin, 0.17 mg ml⁻¹ PMSF and 0.33 mg ml⁻¹ benzamidine). Lysate was clarified by centrifugation at 74,766 g for 60 min and ultracentrifugation at 100,000 g at 4°C for 60 min, followed by filtration through 0.45 μm filter (Amicon Ultra Centrifugal Filter; Merck). The protein was bound to HisTrap HP prepacked columns (GE Healthcare) pre-equilibrated in lysis buffer, washed with high salt buffer B2 (1 M NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 5 mM β-mercaptoethanol) followed by washing with buffer A1, and eluted with a gradient of 0-80% A1-B1 buffers over 30 column volumes (CV) (B1 buffer: 200 mM NaCl, 25 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 400 mM imidazole pH 8.0, 3 mM MgCl₂ and 5 mM β-mercaptoethanol). Fractions containing nsp12 were pooled and the tag was cleaved with His-tagged TEV protease overnight during dialysis against buffer D (200 mM NaCl, 25 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 5 mM β-mercaptoethanol). The protein solution was applied to a HisTrap HP pre-packed column to remove uncleaved protein, tag and TEV protease. The flow-through containing nsp12 was further purified by ion exchange chromatography using Hi TRAP Q HP and SP HP pre-packed columns (GE Healthcare) equilibrated with buffer D. The unbound protein

was concentrated and further purified via size exclusion chromatography on HiLoad S200 16/60 column (GE Healthcare) in buffer A2 (300 mM NaCl, 20 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 1 mM MgCl₂, 1 mM TCEP). The peak contained monomeric fractions of nsp12 that were pooled, concentrated to 55 μM, aliquoted, flash-frozen in liquid nitrogen and stored at -80 °C.

Nsp7 and nsp8 were prepared as described²⁰. Briefly, both proteins were expressed in *E. coli* BL21(DE3) RIL from the pET-derived vector 14-B (a gift from S. Gradić; Addgene 48308) in LB medium individually. Cells were grown to an optical density at 600 nm of 0.4 at 30 °C and protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at 18 °C for 16 h. After harvesting, cells were resuspended in lysis buffer A1 (400 mM NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 30 mM imidazole pH 8.0, 5 mM β-mercaptoethanol, 0.284 μg ml⁻¹ leupeptin, 1.37 μg ml⁻¹ pepstatin, 0.17 mg ml⁻¹ PMSF and 0.33 mg ml⁻¹ benzamidine). Nsp8 and nsp7 were purified separately using the same purification procedure. The cells were lysed using a French press and three cycles. Lysates were subsequently cleared by centrifugation (74,766 g, 4 °C, 30 min). The supernatant was applied to a HisTrap HP column (GE Healthcare), pre-equilibrated in lysis buffer. The column was washed with high-salt buffer B2 (1 M NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 5 mM β-mercaptoethanol), and with buffer A1 (400 mM NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 30 mM imidazole pH 8.0 and 5 mM β-mercaptoethanol). The protein sample was then eluted using a gradient of 0-80% B1 over 30 column volumes (CV) (B1 buffer: 300 mM NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 400 mM imidazole pH 8.0, 3 mM MgCl₂ and 5 mM β-mercaptoethanol). The protein sample was dialysed in buffer D (200 mM NaCl, 50 mM Na-HEPES pH 7.4, 10% (v/v) glycerol and 5 mM β-mercaptoethanol) in the presence of His-tagged TEV protease at 4 °C. The dialysed sample was subsequently applied to a HisTrap HP column (GE Healthcare), pre-equilibrated in A1 buffer. The flow-through that contained the untagged protein of interest was applied to an ion exchange chromatography HiTrap Q column (GE Healthcare). The unbound protein sample, containing nsp8 or nsp7, was concentrated using a MWCO 10,000 Amicon Ultra Centrifugal Filter (Merck) and applied to a HiLoad S200 16/600 (GE Healthcare) equilibrated in buffer A2 (300 mM NaCl, 20 mM Na-HEPES pH 7.4, 5% (v/v) glycerol, 1 mM TCEP). Peak fractions containing proteins were pooled, concentrated to 430 μM (nsp7) and 400 μM (nsp8), aliquoted, flash-frozen in liquid nitrogen and stored at -80 °C.

RNA extension assays

All unmodified and 5' FAM labelled RNA oligonucleotides (**Supplementary Table 1**) were purchased from Integrated DNA Technologies (IDT). MTP was from MedChemExpress and NTPs from ThermoScientific. The assay was performed as

described²⁰, except for the following changes. The final concentrations of nsp12, nsp8, nsp7 and RNA were 3 μ M, 9 μ M, 9 μ M and 3 μ M, respectively, except that 1 μ M of RNA was used in the assays in figure panels 1f,g and 2d. The concentration of NTPs was 37.5 μ M except for assays shown in figure panels 1f, 1g and 2d, where 4 μ M was used. RNA in annealing buffer (50 mM NaCl, 10 mM Na-HEPES pH 7.5) was annealed by heating it to 75 °C for 1 min and gradually cooling to 4 °C. Annealed RNA and pre-mixed RdRp were incubated in reaction buffer (100 mM NaCl, 20 mM Na-HEPES pH 7.5, 5% (v/v) glycerol, 10 mM MgCl₂, 5 mM β -mercaptoethanol) for 10 min at 30 °C and the reactions were started by addition of NTPs. After 20 min incubation at 30 °C the reactions were stopped with 2x stop buffer (7 M urea, 50 mM EDTA, 1x TBE buffer). RNA products were resolved on 20% denaturing polyacrylamide-urea gels in 0.5x TBE running buffer and visualized with a Typhoon 95000 FLA Imager (GE Healthcare Life Sciences).

Preparation and analysis of NHC-containing RNA oligonucleotides

5'-O-DMT-2'-O-TOM-O⁴-chlorophenyluridine was prepared as described⁵², and converted to the corresponding N⁴-hydroxy-N⁴-benzoylcytidine 3'-(2-cyanoethyl) diisopropyl phosphoramidite (M-PA) in three steps. Details of the synthetic procedures are provided below and NMR spectra of isolated compounds are given in the Supplementary Information. RNA oligonucleotides were then prepared by solid-phase synthesis on CPG support (0.6 μ mol scale) using 2'-O-TOM-protected ribonucleoside phosphoramidites (70 mM in CH₃CN) and ethylthiotetrazol (ETT, 250 mM in CH₃CN) as activator, with 4 min coupling time, as previously described^{28,52}. The oligonucleotides were deprotected with 25% NH₄OH/EtOH 3/1 at 55 °C for 6 h, followed by 1 M TBAF in THF for 12 h, and carefully purified by denaturing polyacrylamide gel electrophoresis to remove a minor fraction, in which NHC was converted to C during deprotection.

The purity and identity of the RNA oligonucleotides was analyzed by anion-exchange HPLC (Dionex DNAPac PA200, 2x250 mm, at 60 °C. Solvent A: 25 mM Tris-HCl (pH 8.0), 6 M Urea. Solvent B: 25 mM Tris-HCl (pH 8.0), 6 M Urea, 0.5 M NaClO₄. Gradient: linear, 0–48% solvent B, 4% solvent B per 1 CV), and HR-ESI-MS (Bruker micrOTOF-Q III, negative ion mode, direct injection).

Thermal melting experiments were performed in 10 mM sodium phosphate buffer pH 7.0, 100 mM NaCl, at a RNA duplex concentration of 20 μ M. Absorbance versus temperature profiles were recorded at 260 nm on a Varian Cary 100 spectrometer equipped with a Peltier temperature controller, at a heating rate of 0.5°C/min, for two heating and two cooling ramps between 10 and 90°C. Melting curves were normalized to the absorbance at 95°C, fitted to a two-state transition model with linearly sloping lower and upper baselines, and the melting temperatures were determined at the inflection point of the curves.

Synthesis and characterization of M-PA

All reactions were performed under inert nitrogen atmosphere with dry solvents (CH₂Cl₂, CH₃CN). For workup and purification distilled solvents (technical quality) were used. Column chromatography was performed on silica gel (Kieselgel 60, Merck) with a particle size of 0.040-0.063. TLC was performed on Alugram[®] aluminium sheets (Machery-Nagel, UV visualization, 254 nm). NMR spectra were recorded using Bruker Avance III (400 MHz) spectrometers. Chemical shifts (δ) are given in ppm, relative to the residual solvent signals as internal standards (CDCl₃: ¹H = 7.26, ¹³C = 77.16). Data are reported as: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad; Coupling constants (*J*) are given in Hz. High-resolution (HR) electrospray ionization (ESI) mass spectra (MS) were recorded on a Bruker micrOTOF-Q III spectrometer. The detected mass-to-charge ratio (*m/z*) is given, as well as the calculated monoisotopic mass (**Supplementary Information, Extended Data Set 1; Extended Data Fig. 1**).

To synthesize compound **1** (5'-O-(4,4'-Dimethoxytrityl)-N⁴-hydroxy-2'-O-(triisopropylsilyloxymethyl)cytidine), 5'-O-DMT-2'-O-TOM-O⁴-chlorophenyluridine (400 mg, 474 μ mol, 1.0 eq.) was dissolved in anhydrous CH₃CN (4 mL) under nitrogen atmosphere. DMAP (174 mg, 1.42 mmol, 3.0 eq.) and NEt₃ (331 μ L, 2.37 mmol, 5.0 eq.) were added, followed by hydroxyl amine hydrochloride (165 mg, 2.37 mmol, 5.0 eq.). After stirring for 21 h at ambient temperature, the reaction mixture was diluted with CH₂Cl₂ and washed with saturated aq. NaHCO₃ (2x). The organic phase was dried over Na₂SO₄ and the solvent was removed under reduced pressure. The crude product was purified by column chromatography (*n*-hexane:EtOAc + 1% NEt₃ 1:2 to 2:1) to yield the product (compound **1**, 230 mg, 308 μ mol, 64%) as a colorless foam.

To synthesize compound **2** (N⁴-Benzoyl-5'-O-(4,4'-dimethoxytrityl)-N⁴-hydroxy-2'-O-(triisopropylsilyloxy)methylcytidine), a solution of compound **1** (200 mg, 267 μ mol, 1.0 eq.) in anhydrous CH₂Cl₂ (4 mL) was treated with DMAP (65.3 mg, 535 μ mol, 2.0 eq.) and NEt₃ (149 μ L, 1.07 mmol, 4.0 eq.) under nitrogen atmosphere. Benzoic anhydride (59.9 mg, 265 μ mol, 0.99 eq.) was added in three portions within 3 h and the resulting reaction mixture was stirred for one more hour at ambient temperature. Volatiles were removed under reduced pressure. The crude residue was purified by column chromatography (*n*-hexane:EtOAc + 1% NEt₃ 1:2) to yield compound **2** (183 mg, 215 μ mol, 80%) as a colorless foam.

To synthesize compound **3**, (N⁴-Benzoyl-5'-O-(4,4'-dimethoxytrityl)-N⁴-hydroxy-2'-O-(triisopropylsilyloxy)methyl cytidine 3'-cyanoethyl-N,N-diisopropyl phosphoramidite), compound **2** (150 mg, 176 μ mol, 1.0 eq.) was dissolved in anhydrous CH₂Cl₂ (2 mL) and cooled to 4°C. 2-Cyanoethyl N,N,N',N'-tetraisopropyl phosphoramidite (67.3 μ L, 212 μ mol, 1.2 eq.) and 4,5-dicyanoimidazol (23 mg, 194 μ mol, 1.1 eq.) were added in two portions within 1 hour. After one

additional hour at 4°C, the reaction mixture was allowed to warm up to ambient temperature, and stirred for one more hour. Then, the solvent was evaporated and the crude residue was purified by column chromatography (*n*-hexane:EtOAc + 1% NEt₃ 2:1 to 1:1) to yield compound 3 (75.0 mg, 71.3 μmol, 40%) as a colorless foam.

Cryo-EM sample preparation and data collection

RNA scaffolds for structural studies were prepared by annealing of two RNA oligonucleotides because the length of the NHC-containing RNA was limited for technical reasons. The first RNA oligo was designed to form a template-product hybrid with a hairpin leaving a 10 nt product overhang at the 3' end. A second oligo containing 8 nt complementary to the overhang was annealed to obtain an RNA template-product scaffold with a single nick at position -10 in the template RNA (**Extended Data Fig. 3, Table 1**). The second RNA oligo had NHC incorporated at position -1 and a short 5' overhang of three G nucleotides. RNA scaffolds for RdRp-RNA complex formation were prepared by mixing equimolar amounts of two RNA strands in annealing buffer (10 mM Na-HEPES pH 7.4, 50 mM NaCl) and heating to 75 °C, followed by step-wise cooling to 4 °C. RNA sequences RdRp-RNA complexes were formed by mixing purified nsp12 (1.25 nmol) with an equimolar amount of annealed RNA scaffold and 3-fold molar excess of each nsp8 and nsp7. After 10 minutes of incubation at 30 °C, the mixture was applied to a Superdex 200 Increase 3.2/300 size exclusion chromatography column (GE Healthcare), equilibrated in complex buffer (20 mM Na-HEPES pH 7.4, 100 mM NaCl, 1 mM MgCl₂, 1 mM TCEP) at 4 °C. Peak fractions corresponding to RdRp-RNA complex were pooled and diluted to 1.5 mg/ml. An additional 0.2 nmol of the annealed RNA scaffold were spiked into each sample prior to grid preparation. 3 μL of the concentrated RdRp-RNA complex were mixed with 0.5 μL of octyl β-D-glucopyranoside (0.003% final concentration) and applied to freshly glow discharged R 2/1 holey carbon grids (Quantifoil). The grids were blotted for 7 seconds with blot force 5 using a Vitrobot MarkIV (Thermo Fischer Scientific) at 4 °C and 95 % humidity and plunge frozen in liquid ethane.

Cryo-EM data were collected with SerialEM⁵³ on a Titan Krios transmission electron microscope (Thermo Fischer Scientific) operated at 300 keV. Inelastically scattered electrons were filtered out with a GIF Quantum energy filter (Gatan) using a slit width of 20 eV. Images were acquired using a K3 direct electron detector in counting mode (non-super resolution) at a nominal magnification of 105,000x, resulting in a calibrated pixel size of 0.834 Å/pixel. Images were exposed for a total of 2.0 seconds with a dose rate of 20.5 e-/px/s, resulting in a total dose of 59.6 e-/Å² that was fractionated into 80 frames. Our previous cryo-EM analysis of SARS-CoV-2 RdRp-RNA complexes showed strong preferred particle orientation in ice²⁰. To obtain more particle orientations, all data were collected with 30° stage tilt. Motion correction,

CTF-estimation, and particle picking and extraction were performed using Warp⁵⁴. A total of 11,060 and 10,230 movies were collected for M-A and M-G containing structures, respectively.

Cryo-EM data processing and structural modeling

For the RdRp-RNA complex containing the M-A base pair, 2.5 million particles were extracted using Warp⁵⁴ 1.0.9. Particles were imported to cryoSPARC⁵⁵ 2.15 and subjected to 2D classification. 2D classes representing contamination or broken particles were selected and used for calculating two *ab initio* structures. All particles were then used for supervised 3D classification against five references, where four were originating from the *ab initio* reconstruction of contamination and broken particles and one was a previous RdRp-RNA complex structure (EMD-11995²⁸). The class containing RNA-bound RdRp yielded ~800,000 particles that were subjected to homogenous 3D refinement. The refined particles were then exported to RELION 3.1⁵⁶ and focus-refined in 3D with an initial local angular sampling of 3.7° and a mask around RdRp that omitted the nsp8 sliding poles and the upstream, second RNA turn. To improve the quality of the density in the active site, particles were 3D classified without image alignment (T=4, four classes) using the same mask. The best class was focus-refined to an overall resolution of 3.3 Å. Local resolution was estimated with Relion 3.1 using a kernel size of 10 Å. For the RdRp-RNA complex containing the M-G base pair, 2.2 million particles were extracted using Warp⁵⁴ 1.0.9. Further processing was as for the first complex, except that the class containing RNA-bound RdRp yielded 850,000 particles and the refinement resulted in a reconstruction at an overall resolution of 3.2 Å.

Atomic models were built using our previously published SARS-CoV-2 RdRp-RNA complex structure (PDB 7B3D²⁸) as starting model. The model was first rigid-body fit into the density and then manually adjusted in Coot⁵⁷. The protruding second RNA turn and the nsp8 extensions were removed from the model due to low quality of density for these regions. Restraints for molnupiravir monophosphate (M) were generated in phenix.elbow⁵⁸ and the structures were refined using phenix.real_space_refine⁵⁹ with secondary structure restraints. Model quality was assessed using MolProbity within Phenix⁶⁰ which revealed excellent stereochemistry for both structural models (**Table 1**). Figures were prepared with PyMol (Schrödinger) and ChimeraX⁶¹.

REFERENCES

1. Dolgin, E. The race for antiviral drugs to beat COVID - and the next pandemic. *Nature* **592**, 340-343, doi:10.1038/d41586-021-00958-4 (2021).

2. Merck and Ridgeback Biotherapeutics Provide Update on Progress of Clinical Development Program for Molnupiravir, an Investigational Oral Therapeutic for the Treatment of Mild-to-Moderate COVID-19. Press release, Merck & RidgebackBio, April 15, 2021. (2021).
3. Toots, M. *et al.* Quantitative efficacy paradigms of the influenza clinical drug candidate EIDD-2801 in the ferret model. *Transl Res* **218**, 16-28, doi:10.1016/j.trsl.2019.12.002 (2020).
4. Sheahan, T. P. *et al.* An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci Transl Med* **12**, doi:10.1126/scitranslmed.abb5883 (2020).
5. Wahl, A. *et al.* SARS-CoV-2 infection is effectively treated and prevented by EIDD-2801. *Nature* **591**, 451-457, doi:10.1038/s41586-021-03312-w (2021).
6. Cox, R. M., Wolf, J. D. & Plemper, R. K. Therapeutically administered ribonucleoside analogue MK-4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. *Nat Microbiol* **6**, 11-18, doi:10.1038/s41564-020-00835-2 (2021).
7. Hilgenfeld, R. & Peiris, M. From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. *Antiviral Res* **100**, 286-295, doi:10.1016/j.antiviral.2013.08.015 (2013).
8. Snijder, E. J., Decroly, E. & Ziebuhr, J. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. *Adv Virus Res* **96**, 59-126, doi:10.1016/bs.aivir.2016.08.008 (2016).
9. Posthuma, C. C., Te Velthuis, A. J. W. & Snijder, E. J. Nidovirus RNA polymerases: Complex enzymes handling exceptional RNA genomes. *Virus Res* **234**, 58-73, doi:10.1016/j.virusres.2017.01.023 (2017).
10. Romano, M., Ruggiero, A., Squeglia, F., Maga, G. & Berisio, R. A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells* **9**, doi:10.3390/cells9051267 (2020).
11. Jiang, Y., Yin, W. & Xu, H. E. RNA-dependent RNA polymerase: Structure, mechanism, and drug discovery for COVID-19. *Biochem Biophys Res Commun*, doi:10.1016/j.bbrc.2020.08.116 (2020).
12. Cannalire, R., Cerchia, C., Beccari, A. R., Di Leva, F. S. & Summa, V. Targeting SARS-CoV-2 Proteases and Polymerase for COVID-19 Treatment: State of the Art and Future Opportunities. *J Med Chem*, doi:10.1021/acs.jmedchem.0c01140 (2020).
13. Vicenti, I., Zazzi, M. & Saladini, F. SARS-CoV-2 RNA-dependent RNA polymerase as a therapeutic target for COVID-19. *Expert Opin Ther Pat*, 1-13, doi:10.1080/13543776.2021.1880568 (2021).
14. Tian, L. *et al.* RNA-dependent RNA polymerase (RdRp) inhibitors: The current landscape and repurposing for the COVID-19 pandemic. *Eur J Med Chem* **213**, 113201, doi:10.1016/j.ejmech.2021.113201 (2021).
15. Subissi, L. *et al.* One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci U S A* **111**, E3900-3909, doi:10.1073/pnas.1323705111 (2014).
16. Kirchdoerfer, R. N. & Ward, A. B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat Commun* **10**, 2342, doi:10.1038/s41467-019-10280-3 (2019).
17. Gao, Y. *et al.* Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*, eabb7498, doi:10.1126/science.abb7498 (2020).
18. Yin, W. *et al.* Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*, doi:10.1126/science.abc1560 (2020).
19. Wang, Q. *et al.* Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell*, doi:10.1016/j.cell.2020.05.034 (2020).
20. Hillen, H. S. *et al.* Structure of replicating SARS-CoV-2 polymerase. *Nature* **584**, 154-156, doi:10.1038/s41586-020-2368-8 (2020).
21. Jochheim, F. A. *et al.* Dimeric form of SARS-CoV-2 polymerase. *bioRxiv*, 2021.2003.2023.436644, doi:10.1101/2021.03.23.436644 (2021).
22. Hillen, H. S. Structure and function of SARS-CoV-2 polymerase. *Curr Opin Virol* **48**, 82-90, doi:10.1016/j.coviro.2021.03.010 (2021).
23. Bouvet, M. *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc Natl Acad Sci U S A* **109**, 9372-9377, doi:10.1073/pnas.1201130109 (2012).
24. Robson, F. *et al.* Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. *Mol Cell* **79**, 710-727, doi:10.1016/j.molcel.2020.07.027 (2020).
25. Ferron, F. *et al.* Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc Natl Acad Sci U S A* **115**, E162-E171, doi:10.1073/pnas.1718806115 (2018).
26. Gordon, C. J., Tchesnokov, E. P., Feng, J. Y., Porter, D. P. & Gotte, M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem*, doi:10.1074/jbc.AC120.013056 (2020).
27. Gordon, C. J. *et al.* 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* **295**, 6785-6797, doi:10.1074/jbc.RA120.013679 (2020).
28. Kocic, G. *et al.* Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat Commun* **12**, 279, doi:10.1038/s41467-020-20542-0 (2021).
29. Bravo, J. P. K., Dangerfield, T. L., Taylor, D. W. & Johnson, K. A. Remdesivir is a delayed translocation inhibitor of SARS-CoV-2 replication. *Mol Cell* **81**, 1548-1552 e1544, doi:10.1016/j.molcel.2021.01.035 (2021).
30. Beigel, J. H. *et al.* Remdesivir for the Treatment of Covid-19 - Preliminary Report. *N Engl J Med*, doi:10.1056/NEJMoa2007764 (2020).
31. Grein, J. *et al.* Compassionate Use of Remdesivir for Patients with Severe Covid-19. *N Engl J Med* **382**, 2327-2336, doi:10.1056/NEJMoa2007016 (2020).
32. Teoh, S. L., Lim, Y. H., Lai, N. M. & Lee, S. W. H. Directly Acting Antivirals for COVID-19: Where Do We Stand? *Front Microbiol* **11**, 1857, doi:10.3389/fmicb.2020.01857 (2020).

33. Pan, H. *et al.* Repurposed antiviral drugs for COVID-19; interim WHO SOLIDARITY trial results. *medRxiv*, 2020.2010.2015.20209817, doi:10.1101/2020.10.15.20209817 (2020).
34. Consortium, W. H. O. S. T. *et al.* Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results. *N Engl J Med* **384**, 497-511, doi:10.1056/NEJMoa2023184 (2021).
35. Painter, G. R. *et al.* The prophylactic and therapeutic activity of a broadly active ribonucleoside analog in a murine model of intranasal venezuelan equine encephalitis virus infection. *Antiviral Res* **171**, 104597, doi:10.1016/j.antiviral.2019.104597 (2019).
36. Yoon, J. J. *et al.* Orally Efficacious Broad-Spectrum Ribonucleoside Analog Inhibitor of Influenza and Respiratory Syncytial Viruses. *Antimicrob Agents Chemother* **62**, doi:10.1128/AAC.00766-18 (2018).
37. Ehteshami, M. *et al.* Characterization of beta-d-N(4)-Hydroxycytidine as a Novel Inhibitor of Chikungunya Virus. *Antimicrob Agents Chemother* **61**, doi:10.1128/AAC.02395-16 (2017).
38. Reynard, O. *et al.* Identification of a New Ribonucleoside Inhibitor of Ebola Virus Replication. *Viruses* **7**, 6233-6240, doi:10.3390/v7122934 (2015).
39. Costantini, V. P. *et al.* Antiviral activity of nucleoside analogues against norovirus. *Antivir Ther* **17**, 981-991, doi:10.3851/IMP2229 (2012).
40. Stuyver, L. J. *et al.* Ribonucleoside analogue that blocks replication of bovine viral diarrhea and hepatitis C viruses in culture. *Antimicrob Agents Chemother* **47**, 244-254, doi:10.1128/aac.47.1.244-254.2003 (2003).
41. Agostini, M. L. *et al.* Small-Molecule Antiviral beta-d-N (4)-Hydroxycytidine Inhibits a Proofreading-Intact Coronavirus with a High Genetic Barrier to Resistance. *J Virol* **93**, doi:10.1128/JVI.01348-19 (2019).
42. Painter, W. P. *et al.* Human Safety, Tolerability, and Pharmacokinetics of Molnupiravir, a Novel Broad-Spectrum Oral Antiviral Agent with Activity Against SARS-CoV-2. *Antimicrob Agents Chemother*, doi:10.1128/AAC.02428-20 (2021).
43. Crotty, S., Cameron, C. E. & Andino, R. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A* **98**, 6895-6900, doi:10.1073/pnas.111085598 (2001).
44. Urakova, N. *et al.* beta-d-N (4)-Hydroxycytidine Is a Potent Anti-alphavirus Compound That Induces a High Level of Mutations in the Viral Genome. *J Virol* **92**, doi:10.1128/JVI.01965-17 (2018).
45. Les, A., Adamowicz, L. & Rode, W. Structure and conformation of N4-hydroxycytosine and N4-hydroxy-5-fluorocytosine. A theoretical ab initio study. *Biochim Biophys Acta* **1173**, 39-48, doi:10.1016/0167-4781(93)90240-e (1993).
46. Jena, N. R. Role of different tautomers in the base-pairing abilities of some of the vital antiviral drugs used against COVID-19. *Phys Chem Chem Phys* **22**, 28115-28122, doi:10.1039/d0cp05297c (2020).
47. Shannon, A. *et al.* Rapid incorporation of Favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. *Nat Commun* **11**, 4682, doi:10.1038/s41467-020-18463-z (2020).
48. Naydenova, K. *et al.* Structure of the SARS-CoV-2 RNA-dependent RNA polymerase in the presence of favipiravir-RTP. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2021946118 (2021).
49. Chen, J. *et al.* Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell* **182**, 1560-1573 e1513, doi:10.1016/j.cell.2020.07.033 (2020).
50. Malone, B. *et al.* Structural basis for backtracking by the SARS-CoV-2 replication-transcription complex. *bioRxiv*, 2021.2003.2013.435256, doi:10.1101/2021.03.13.435256 (2021).
51. Suzuki, T., Moriyama, K., Otsuka, C., Loakes, D. & Negishi, K. Template properties of mutagenic cytosine analogues in reverse transcription. *Nucleic Acids Res* **34**, 6438-6449, doi:10.1093/nar/gkl761 (2006).
52. Buttner, L., Seikowski, J., Wawrzyniak, K., Ochmann, A. & Hobartner, C. Synthesis of spin-labeled riboswitch RNAs using convertible nucleosides and DNA-catalyzed RNA ligation. *Bioorg Med Chem* **21**, 6171-6180, doi:10.1016/j.bmc.2013.04.007 (2013).
53. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).
54. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146-1152, doi:10.1038/s41592-019-0580-y (2019).
55. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
56. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).
57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
58. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074-1080, doi:10.1107/S0907444909029436 (2009).
59. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/S2059798318006551 (2018).
60. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci* **27**, 293-315, doi:10.1002/pro.3330 (2018).
61. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci*, doi:10.1002/pro.3943 (2020).

ACKNOWLEDGEMENTS

We thank U. Steuerwald for maintenance of the EM infrastructure. H.S.H. was supported by the Deutsche Forschungsgemeinschaft (SFB1190, FOR2848, EXC 2067/1-390729940). C.H. was supported by the DFG (SPP1784) and the ERC Consolidator Grant illumizymes (grant agreement No 682586). P.C. was supported by the Deutsche Forschungsgemeinschaft (SFB860, SPP2191, EXC 2067/1-390729940) and the ERC Advanced Investigator Grant CHROMATRANS (grant agreement No 693023).

AUTHOR CONTRIBUTIONS

F.K. and J.S. designed and carried out biochemical experiments and analyzed biochemical data. C.S. synthesized and analyzed M-containing RNA oligonucleotides. F.K. and C.D. carried out cryo-EM data collection and processing. H.H. assisted with model building. C.H. designed and supervised RNA synthesis and analysis. P.C. designed and supervised research and wrote the manuscript, with input from all authors.

COMPETING INTERESTS

The authors declare no competing interests.

CORRESPONDENCE AND REQUESTS

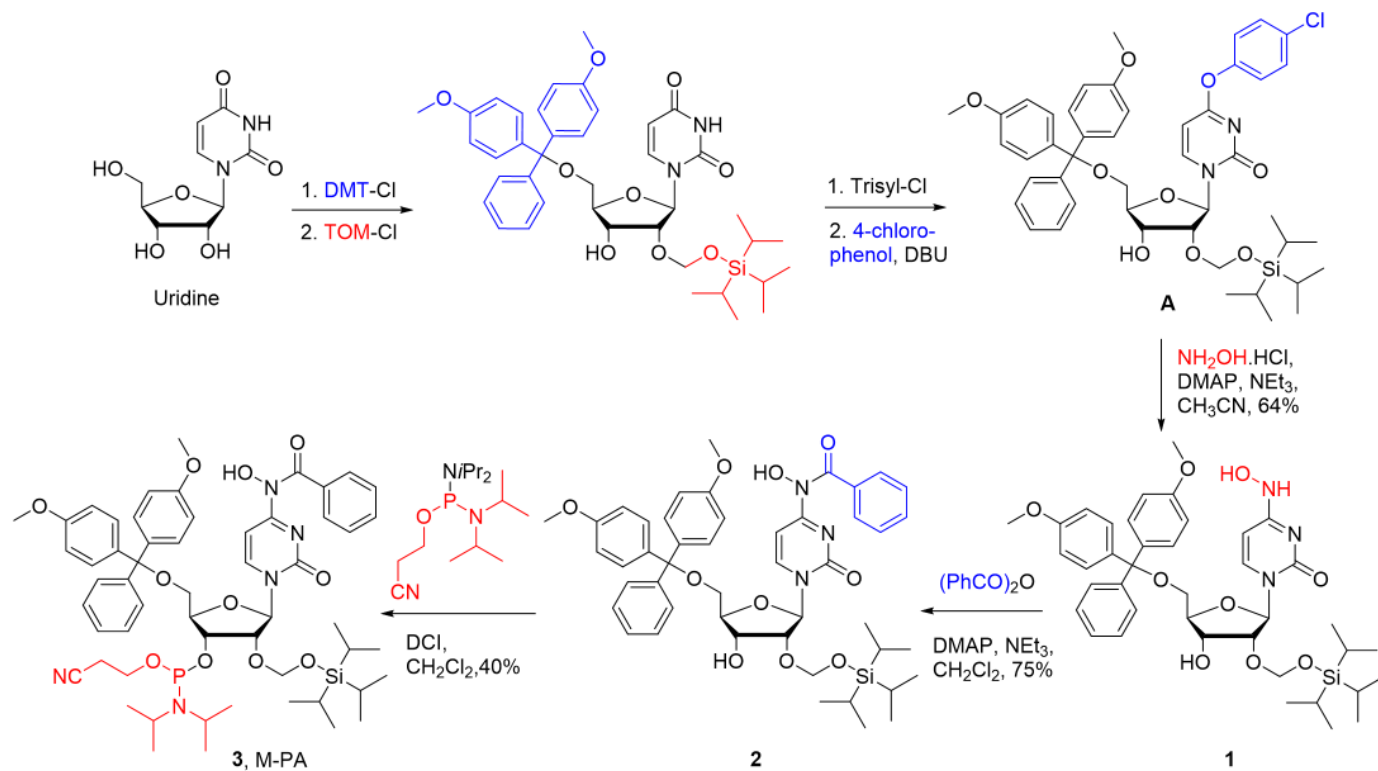
Correspondence and requests for materials should be addressed to C.H. (claudia.hoebartner@uni-wuerzburg.de) and P.C. (patrick.cramer@mpibpc.mpg.de).

DATA AVAILABILITY

The cryo-EM reconstructions and structure coordinates for the RdRp-RNA structures containing M-G or M-A base pairs were deposited with the Electron Microscopy Database (EMDB) under accession codes EMD-XXXXX and EMD-YYYYY and with the Protein Data Bank (PDB) under accession codes XXXX and YYYYY, respectively. Source data are provided with this paper. Other data are available from corresponding authors upon reasonable request.

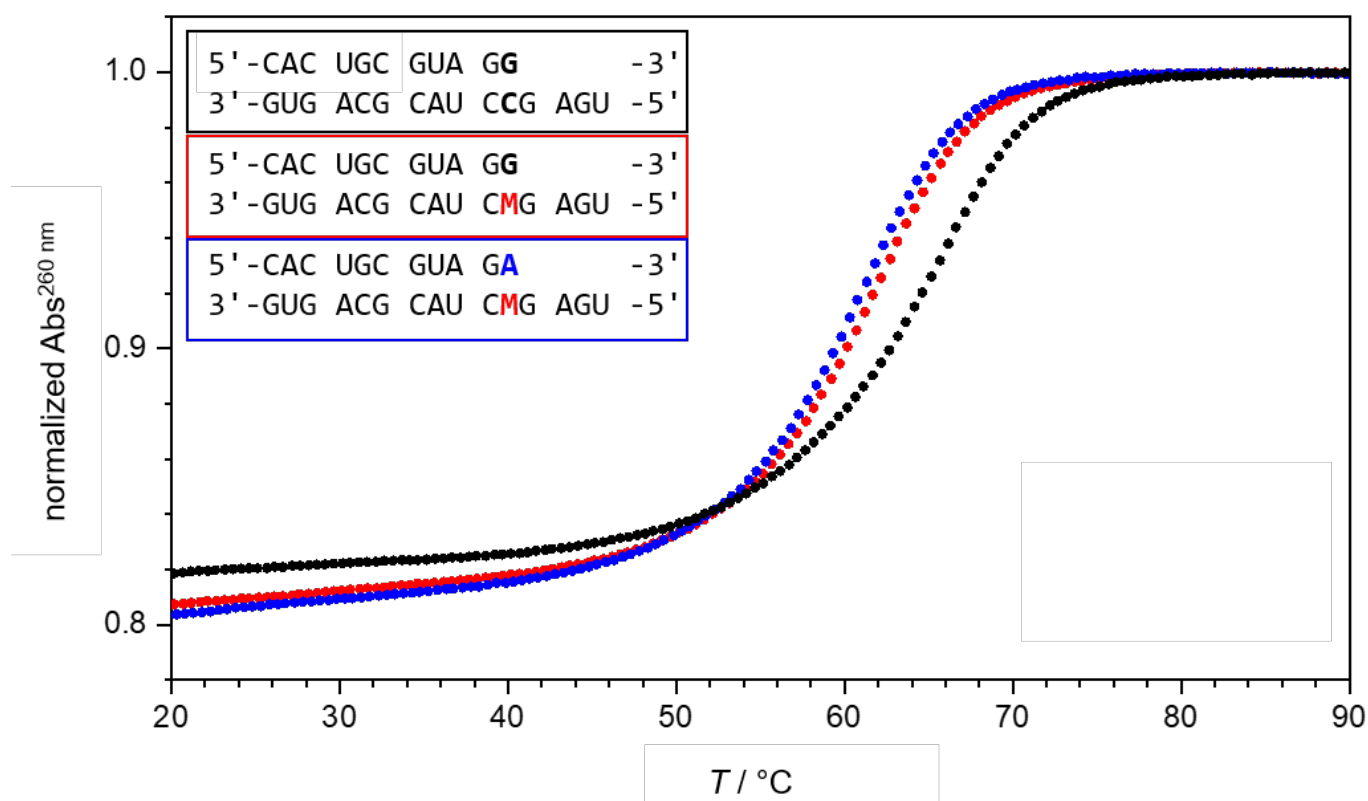
Table 1: Cryo-EM data collection, refinement and validation statistics

	RdRp-RNA with M–A base pair (EMD-XXXX) (PDB XXXX)	RdRp-RNA with M–G base pair (EMD-YYYYY) (PDB YYYY)
Data collection and processing		
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e ⁻ /Å ²)	59.6	59.6
Defocus range (μm)	0.4-2.2	0.4-2.7
Pixel size (Å)	0.834	0.834
Symmetry imposed	C1	C1
Initial particle images (no.)	2,528,775	2,183,996
Final particle images (no.)	373,938	851,168
Map resolution (Å)	3.3	3.2
FSC threshold	0.143	0.143
Map resolution range (Å)	2.7-4.5	2.8-3.9
Refinement		
Initial model used (PDB code)	7B3D	7B3D
Model resolution (Å)	2.9	2.9
FSC threshold	0.5	0.5
Model resolution range (Å)	2.7-4.5	2.8-3.9
Map sharpening <i>B</i> factor (Å ²)	-142	-159
Model composition		
Non-hydrogen atoms	8405	8406
Protein residues	991	991
Ligands	3	3
<i>B</i> factors (Å ²)		
Protein	66.68	61.32
Ligand	68.17	60.27
R.m.s. deviations		
Bond lengths (Å)	0.004	0.005
Bond angles (°)	0.587	0.581
Validation		
MolProbity score	1.52	1.56
Clashscore	7.14	8.23
Poor rotamers (%)	0.11	0.11
Ramachandran plot		
Favored (%)	97.34	97.45
Allowed (%)	2.66	2.55
Disallowed (%)	0.00	0.00



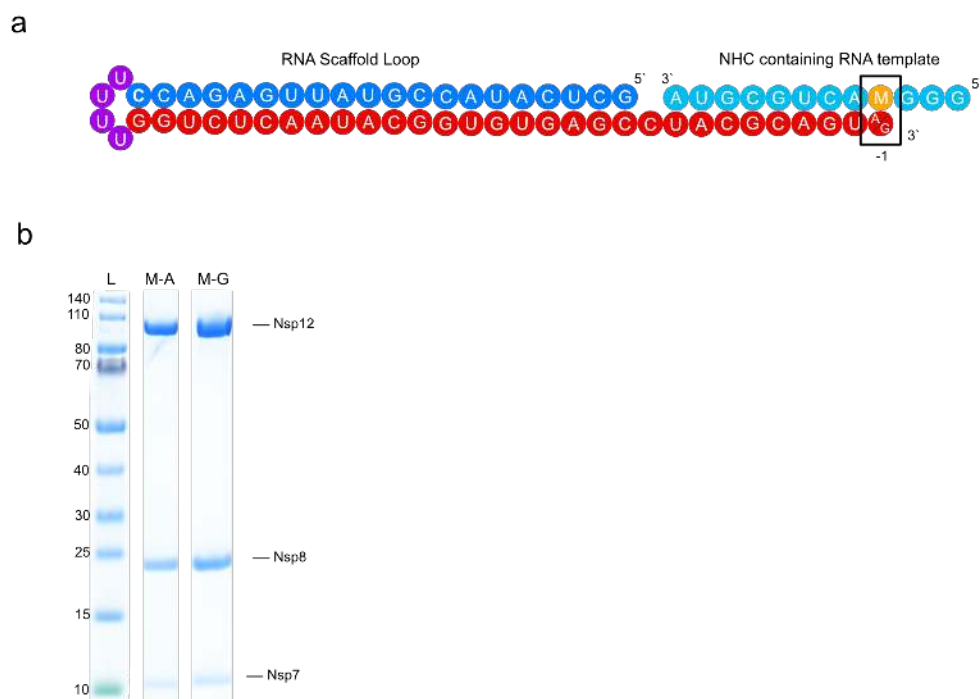
Extended Data Fig. 1 | Synthesis of NHC-phosphoramidite 3 (M-PA).

5'-O-DMT-2'-O-TOM-4-chlorophenyluridine (A) was synthesized from uridine as previously reported (L. Buttner, J. Seikowski, K. Wawrzyniak, A. Ochmann, C. Hobartner, Synthesis of spin-labeled riboswitch RNAs using convertible nucleosides and DNA-catalyzed RNA ligation. *Bioorg Med Chem* 21, 6171-6180 (2013)). The chlorophenol group was displaced by hydroxylamine to give new compound 1. After selective benzoyl protection at N4 with benzoic anhydride, compound 2 was converted to the phosphoramidite M-PA (3) using 2-cyanoethyl-N,N,N',N'-tetraisopropylphosphorodiamidite and 4,5-dicyano-imidazol (DCI) in analogy to a previous report (J. Lu, L. Nan-Sheng, J. A. Piccirilli, Efficient Synthesis of N 4-Methyl- and N 4-Hydroxycytidine Phosphoramidites. *Synthesis* 16, 2708-2712 (2010)). DMT-Cl = 4,4'-dimethoxy-trityl chloride, TOM-Cl = triisopropylsilyloxymethyl chloride, DBU = 1,8-diazabicyclo [5.4.0]undec-7-ene, DMAP = 4-(N,N-dimethylamino)-pyridine.



Extended Data Fig. 2 | Melting curves for RNA duplexes containing M-G or M-A base pairs.

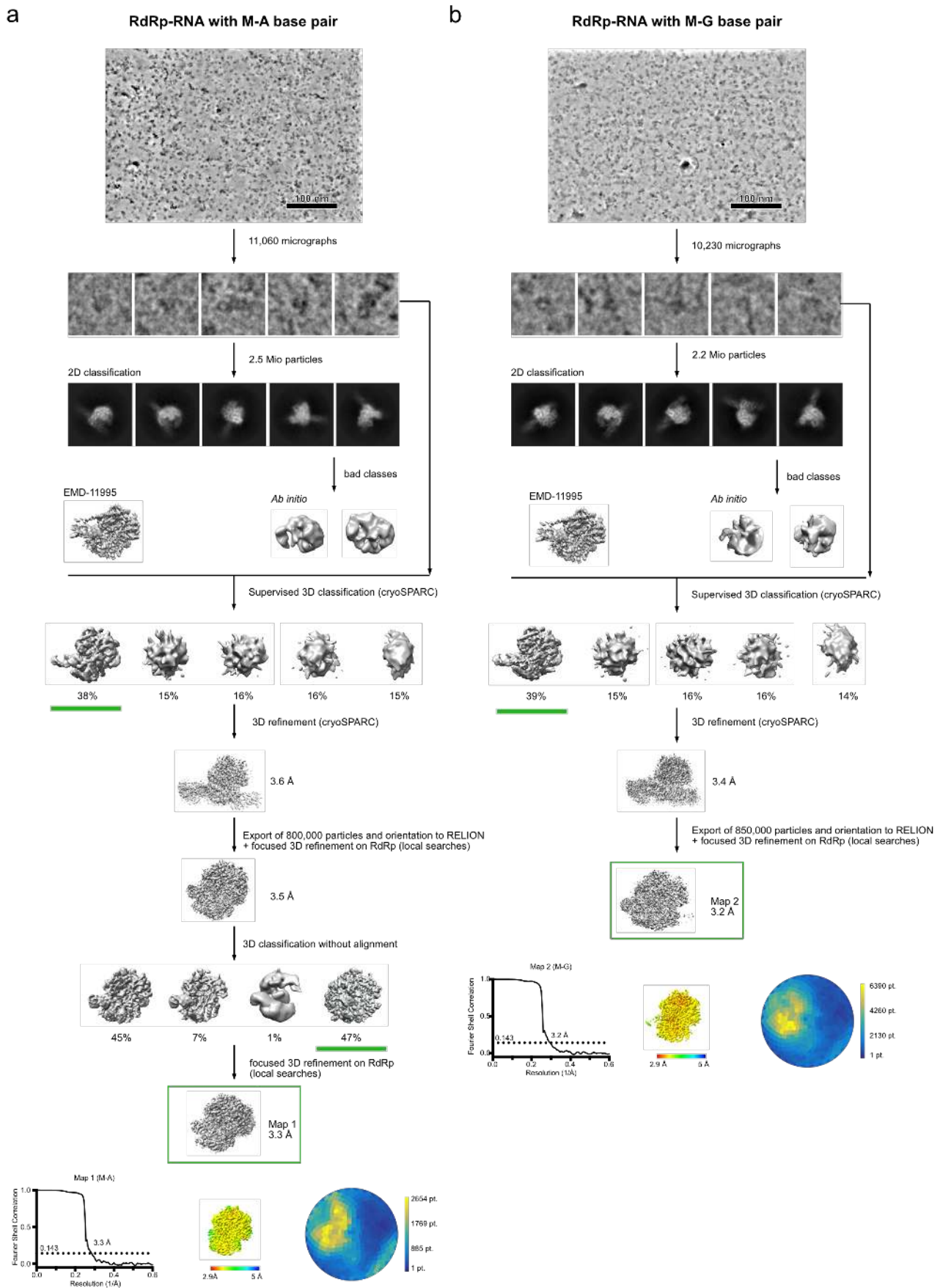
UV thermal melting monitored at 260 nm for 20 μ M duplexes (11 bp, 4 nt single-stranded overhang) in 100 mM NaCl, 10 mM Na-phosphate buffer pH 7.0. black: unmodified duplex C-G ($T_m = 64.7^\circ\text{C}$), red terminal M-G ($T_m = 61.2^\circ\text{C}$): blue: terminal M-A base pair ($T_m = 60.6^\circ\text{C}$). See also Supplementary Table 2.



Extended Data Fig. 3 | Protein preparation and RNA scaffold for structural studies.

a, RNA scaffold was obtained by annealing a short M-containing oligonucleotide to a hairpin RNA duplex.

b, SDS-PAGE of purified RdRp-RNA complexes used for cryo EM. Purified proteins were run on 4-12 % Bis-Tris SDS-PAGE gels in 1x MOPS buffer and stained with Coomassie Blue. M-A corresponds to the RdRp complex with RNA scaffold, where M in template base pairs to A. M-G corresponds to the RdRp complex with RNA scaffold, where M in template base pairs to G. L: PageRuler Prestained Protein Ladder (Thermo Scientific).



Extended Data Fig. 4 | Cryo-EM sorting trees and quality of reconstructions.

a, Cryo-EM sorting tree (top); local resolution, FSC plot and angular distribution of the final reconstruction (bottom) for M-A bp-containing RdRp-RNA structure.

b, Cryo-EM sorting tree (top); local resolution, FSC plot and angular distribution of the final reconstruction (bottom) for M-G bp-containing RdRp-RNA structure.