nature communications

Article

https://doi.org/10.1038/s41467-024-54618-y

Structural basis of deoxynucleotide addition by HIV-1 RT during reverse transcription

Received: 7 March 2024

Accepted: 15 November 2024

Sandra Vergara[1], Xiaohong Zhou[1], Ulises Santiago[1], Mounia Alaoui-El-Azher[2], James F. Conway[1], Nicolas Sluis-Cremer[2] & Guillermo Calero[1]

Reverse transcription of the retroviral RNA genome into DNA is an integral step during HIV-1 replication. Despite a wealth of structural information on reverse transcriptase (RT), we lack insight into the intermediate states of DNA synthesis. Using catalytically active substrates, and a blot/diffusion cryo-electron microscopy approach, we capture 11 structures encompassing reactant, intermediate and product states of dATP addition by RT at 2.2 to 3.0 Å resolution. In the reactant state, dATP binding to RT-template/primer involves a single Mg^{2+} (site B) inducing formation of a negatively charged pocket where a second floating Mg^{2+} can bind (site A). During the intermediate state, the α-phosphate oxygen from a previously unobserved dATP conformer aligns with site A Mg^{2+} and the primer 3′-OH for nucleophilic attack. The product state, comprises two substrate conformations including an incorporated dAMP with the pyrophosphate leaving group coordinated by metal B and stabilized through H-bonds. Moreover, K220 mutants significantly impact the rate of dNTP incorporation by RT and HIV-1 replication capacity. This work sheds light into the dynamic components of a reaction that is central to HIV-1 replication.

Reverse transcription of the retroviral single-stranded RNA genome into double-stranded DNA is an essential step in HIV-1 replication. Reverse transcriptase (RT), a multifunctional enzyme that encodes for RNA- and DNA-dependent DNA polymerase and ribonuclease H activities, catalyzes the entire reverse transcription reaction, and drugs that target this enzyme form the cornerstone of antiretroviral therapy[1–3]. RT exhibits an ordered sequential mechanism of DNA synthesis: template/primer (T/P) nucleic acid substrate binding to RT precedes dNTP binding and incorporation[4]. The rate-limiting step of the nucleotidyl transfer reaction is a conformational change that precisely aligns the dNTP substrate in the active site to facilitate phosphodiester bond formation[5,6]. dNTP addition is metal-dependent and is driven by two Mg^{2+} ions, which align substrates in the active site[7]. The Site A metal, which has only been observed in a few crystal structures of an RT-T/P-dNTP ternary complex[8–10], likely facilitates phosphodiester bond formation by lowering the pKa of the primer 3′-OH, allowing proton release to activate the nucleophile[11]. The Site B metal is thought to stabilize the incoming dNTP by coordinating with non-bridging oxygens of the triphosphate, and may participate in pyrophosphate (PPi) release following chemistry[12,13].

Multiple structures of HIV-1 RT are available for RT-T/P[14–16], RT-T/P-dNTP[8–10,17,18], and post-catalytic RT-T/P_{+1} complexes[10,18,19]. Structures for the pre-catalytic RT-T/P-dNTP complexes, however, trapped the reaction utilizing inactive substrates (e.g., a dideoxy-terminated primer[8–10,17,18] or Ca^{2+}[20,21]). These structures represent catalytically inert complexes that do not support nucleotide addition. As such, we lack insight into the kinetic intermediate states that: (i) facilitate alignment of the dNTP substrate in the RT active site and define the rate-limiting step of the catalytic reaction; (ii) underscore the contribution that Mg^{2+} plays in positioning the dNTP substrate and in phosphodiester bond formation; and (iii) illustrate the conformational changes and intermediate states in the active site that drive catalysis. To address these knowledge gaps, we devised on-grid diffusion/blot experiments for time-resolved cryo-electron microscopy (cryo-EM)

---

[1]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [2]Department of Medicine, Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. e-mail: nps2@pitt.edu; guc9@pitt.edu

where different combinations of enzyme and substrates are applied on opposite sides of a grid allowing mixing of both reaction components by diffusion through the matrix of holes in the support film, prior to blotting and the subsequent plunge-freeze procedure (Supplementary Fig. 1, see "Methods" section). Increasing the temperature of the reaction during the diffusion/blotting step provided additional means to control the reaction.

Here, we use catalytically active substrates to capture five unique intermediate states during the process of nucleotide addition by RT at near-atomic resolution. Initial dATP binding to RT-template/primer involves a single Mg^{2+} (site B) creating a negatively charged pocket where a second mobile Mg^{2+} can bind (site A). In the intermediate state, the floating site A metal, the 3′-OH of the primer, and the α-phosphate oxygen align for nucleophilic attack. Two substrate conformations were detected in the product state including an incorporated dAMP with the pyrophosphate leaving group coordinated by metal B and stabilized through H-bonds. The structural data suggest that active site residue K220 is in a position to donate a proton to the leaving pyrophosphate acting as the general acid in the reaction. Moreover, K220 mutants significantly impact the rate of dNTP incorporation by RT and HIV-1 replication capacity. These structures elucidate conformational changes in the active site of RT that facilitate dNTP incorporation and shed light on a fundamental chemical reaction that drives HIV-1 replication.

Results

Structure of HIV-1 RT in complex with a DNA aptamer T/P substrate

We first solved the structure of HIV-1 RT bound to a previously reported 38-mer hairpin DNA aptamer T/P substrate (RT-T/P)[22] at 2.7 Å, with local resolution around the active site of 2.3 Å (Fig. 1a and Supplementary Fig. 2a. See "Methods" section for the definition of local resolution around the active site). The 38-mer hairpin DNA aptamer T/P has a high affinity for RT and allows for nucleotide incorporation[20,22,23]. The aptamer contains two 2′-O-methylated nucleotides that were used as markers to monitor the positioning of the T/P. The high affinity of the DNA aptamer for RT decreased scaffold mobility allowing collection of higher resolution data both in X-ray and cryo-EM experiments[22–25]. The density of the DNA aptamer was well-resolved (Fig. 1b), including the two methyl groups that were unequivocally observed in the density map (Fig. 1c). Our cryo-EM structure of the RT-T/P complex is comparable to previously reported structures[14,19,22], however, crucial differences are detected in the fingers subdomain, especially in the β3–β4 loop (Supplementary Fig. 3a). These differences between the cryo-EM and X-ray structures arise from crystal contacts with symmetry-related molecules that alter significantly the position of the subdomain; moreover, density for these regions is well-defined in the cryo-EM map but not in the X-ray electron density maps (Supplementary Fig. 3b). The position of the β3-β4 loop observed in the solution structure shows a unique rigid conformation that allows correct placing of critical residues. In the dNTP-free structure, β3-β4 loop residues are located ~11 Å away from the active site creating an open pocket for dNTP binding (Supplementary Fig. 2a and Fig. 2e, see below). The rigid conformation of the β3-β4 loop allowed calculation of its surface potential showing a net positive charge (Supplementary Fig. 3c, left panel) that could potentially interact with the negatively charged phosphate tail. In addition to the well-defined fingers subdomain, all key structural elements of the polymerase active site were visualized in the cryo-EM maps (see below).

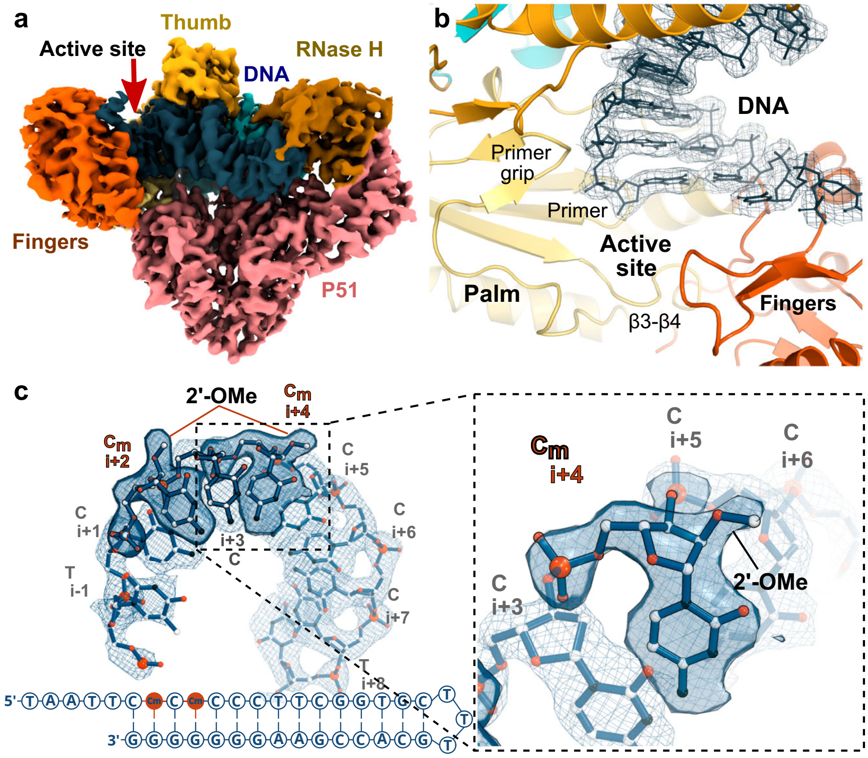
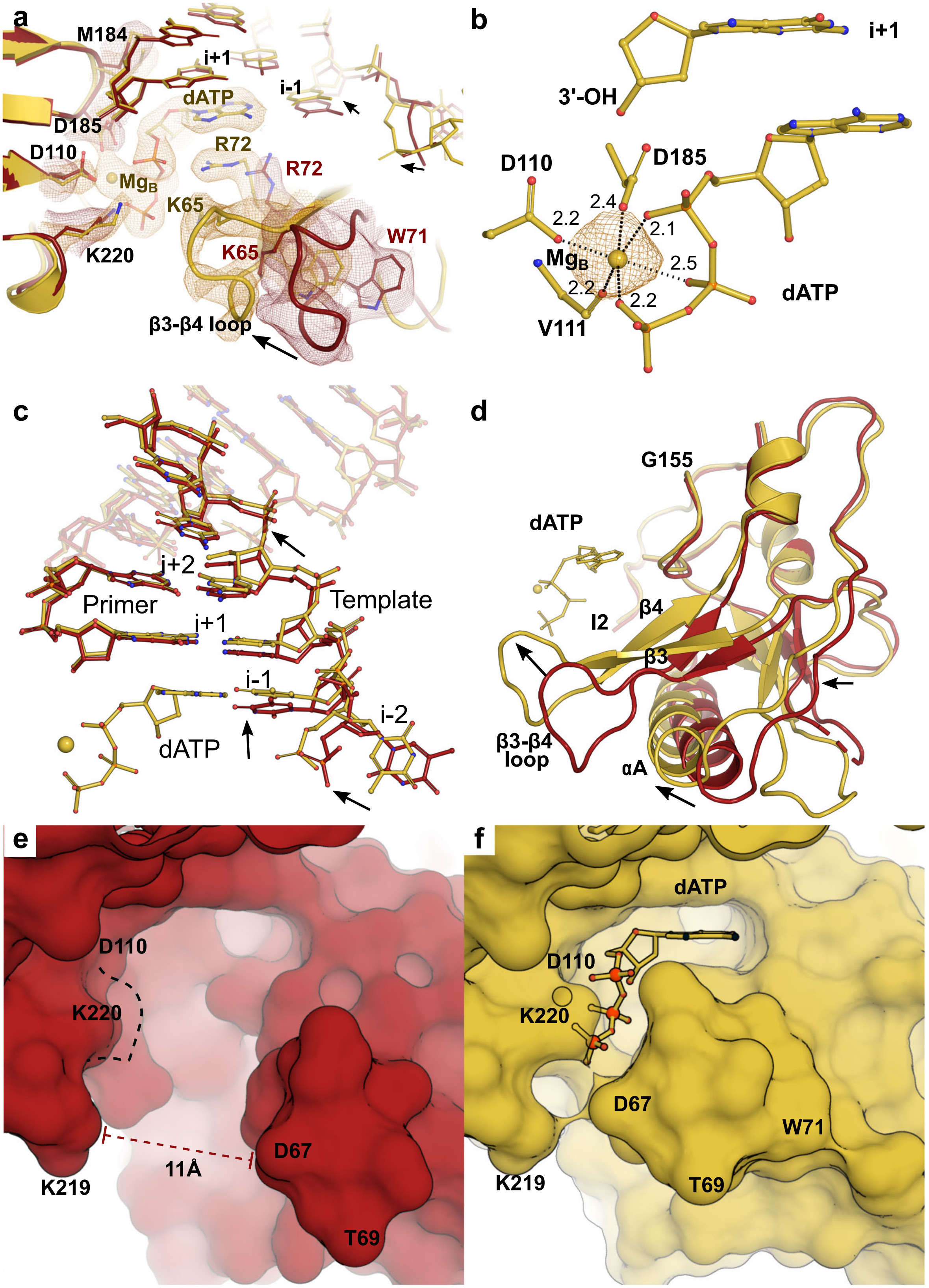


Fig. 1 | Cryo-EM structure of the dNTP-free RT-T/P complex. a Cryo-EM density map for the RT-T/P complex colored by structural subdomains. **b** Cartoon representation of the DNA polymerase active site of RT illustrating key protein motifs. The density for the T/P aptamer substrate is indicated in blue mesh with the corresponding model in the sticks. **c** Sequence of the 38-mer hairpin DNA aptamer T/P substrate used throughout this study. Nucleotides in positions +2 and +4 of the template strand labeled as C_m are 2′-O-methyl modified (2′-OMe). Inset, detail of the density for the 2′-O-methyl modified nucleotide in position +4. Map sharpened at B-factor −80 Å² and contoured at 6σ.



Structural basis of dNTP binding - reactant state 1 of the RT-T/P-dATP complex

Using on-grid diffusion/blot experiments, we successfully trapped an RT-T/P-dATP reactant state (R-1) by applying a solution of RT-T/P on one side of the grid and a solution of dATP-MgCl$_2$ on the opposite side and immediately initiating the blotting/freezing procedure (see "Methods" section). The overall structure of state R-1 of the RT-T/P-dATP ternary complex, at 2.1 Å local resolution around the active site (Fig. 2a and Supplementary Fig. 2b), includes a dATP molecule in the binding pocket in the presence of a catalytically active 3′-OH primer; and a site B Mg^{2+} ion, coordinated by the α, β, and γ-phosphates, the backbone carbonyl of V111, and the carboxylic oxygens of D110 and

Fig. 2 | Cryo-EM structure of reactant state R-1 of the RT-T/P-dATP complex. **a** Overlay of the structures of the RT-T/P complex (red) and state R-1 of the RT-T/P-dATP complex (yellow) highlighting conformational changes in the RT active site that occur upon dATP binding (black arrows). The protein backbone displayed in the cartoon and key residues displayed as ball-and-stick. Map for state R-1 is displayed in yellow mesh, with sharpening at B-factor -60 Å^2 and contoured at 10σ. The map for the RT-T/P complex is displayed in red mesh, unsharpened, and contoured at 7σ. **b** Site B Mg^{2+} coordination in state R-1 structure, yellow mesh contoured at 10σ. Distances (in Å) are indicated next to dashed lines. **c** Watson–Crick pairing of dATP with the templating (i-1) base, induces repositioning of the template strand (black arrows). **d** Large-scale conformational changes of the P66 fingers subdomain (residues 1–85 and 118–155) are observed (black arrows) upon dATP binding. **e** Surface representation (shown in red) of the Open active site in the dNTP-free RT-T/P complex and **f** the Closed active site of RT-T/P-dATP complex (shown in yellow) as a result of dATP binding (shown in ball-and-stick). Reposition of key residues D110 and K220 induces a tighter pocket in state R-1 evidenced by the extra bump region absent in RT-T/P (see **e** dashed line).

D185 (Fig. 2b). Overlay of the RT-T/P and state R-1 structures shows an overall RMSD on Cαs of 0.52 Å. However, conformational changes can be observed in the aptamer templating base (Fig. 2c); the fingers subdomain (RMSD 2.29 Å RT-T/P vs RT-T/P-dATP fingers subdomain); the β3-β4 loop, which swings toward the substrate and closes the active site pocket (Fig. 2d); and active site residues D110 and K220 (Fig. 2e, f). In this way, the position of the β3-β4 loop away from the nucleotide-binding site in the cryo-EM RT-T/P structure defines the (unbiased) open active site (Fig. 2e), while interactions of the β3-β4 loop with the incoming dATP define a closed active site (Fig. 2f). In the latter, β3-β4 residue R72 stacks against the dATP base forming salt bridges with the closest α-phosphate oxygen while K65 forms a salt bridge with the γ-phosphate oxygen (Fig. 2a and Supplementary Fig. 3d). The positively charged β3-β4 loop (Supplementary Fig. 3c) together with conformational changes in active site residues (loop D110 to Y115 and K220) create a positive surface potential to bind the negatively charged phosphate tail (Supplementary Fig. 3e, f). Other residues, including D185, and those involved in dNTP binding, stabilization and substrate selection, M184, Q151, Y115, and F116, also appear repositioned (Supplementary Fig. 3d). Overall, the protein conformational changes in the β3-β4 loop and RT active site are similar to those previously described for RT-T/P-dNTP ternary complexes[8,9] with small differences that reflect the presence of a 3′-OH terminated primer in the cryo-EM structures.

Reactant state 2 of the RT-T/P-dATP complex

A second reactant state (R-2) of the RT-T/P-dATP complex was trapped by applying an RT-T/P solution preincubated with dATP on one side of the grid and $MgCl_2$ on the other side, and then immediately initiating the blotting/freezing procedure. The structure features 2.4 Å local resolution around the active site (Supplementary Fig. 2c) with an RMSD on Cαs of 0.34 Å with respect to the state R-1 structure. The overall structure of state R-2 is almost identical to R-1, including site B Mg^{2+} ion. However, diffusion of Mg^{2+} to a pre-bound dATP enzyme allowed observation of strong, previously unreported density close to the side chain of D110, which we modeled as a Mg^{2+} ion (Fig. 3a–c). This putative Mg^{2+} does not correspond to the typical location of metal A reported for HIV-1 RT (Supplementary Fig. 4a)[8,9]; or the site C position observed in some DNA polymerases[26–28], thus we designated it as site A* Mg^{2+}. Resolving waters from ions in cryo-EM is an active area of analysis[29–31]. In this regard, the density at site A* was modeled as an ion rather than water based on the criteria described in the UnDowser procedure[29] and the segmentation-guided water and ion modeling method[32]. These criteria consider that the distance of a water atom to a nearby polar atom is 2.8 ± ~0.3 Å, whereas the distance of an ion to a charged/polar atom is 2.2 ± ~0.3 Å. In the Coulomb potential map, site A* Mg^{2+} is located ~2.3 Å from the carboxylic oxygen of D110. A prior study reported that water molecules can display high variability in positions and contour shapes at different resolutions[30]. In contrast, ions appear in a similar position and shape in all maps, with ions still being observed at 3.1 Å resolution whereas waters are no longer detectable. The density at site A* was consistent with these criteria. First, the density consistently appeared at the same position in similar experimental conditions (Supplementary Fig. 4b) and the shape and position of the density were consistent in both half maps and the full map (Supplementary Fig. 4c); moreover, there was no remarkable change when the map was sharpened at different B-factors, contrary to the noise amplification effect of sharpening (Supplementary Fig. 4d). Second, a control experiment where RT was incubated with dATP in the absence of $MgCl_2$ but with EDTA, to remove any trace Mg^{2+} contaminant (see "Methods" section), showed no density (that could correspond to site A* Mg^{2+}) near the side chain of D110 (Fig. 3d). Third, calculation of the surface electrostatic potential around site A* metal shows the presence of a negative potential pocket that would favor recruitment of positive charges (Fig. 3e)[33].

Intermediate state of the RT-T/P-dATP complex

With the aim of broadening our experimental conditions, we performed two additional simultaneous perturbations to state R-2 conditions: (1) we allowed sample diffusion for 8 s before blotting/freezing, and (2) we changed the Vitrobot chamber temperature from 22 °C to 37 °C during vitrification (Supplementary Fig. 2d). Under these conditions we captured a structure (T_{8sec}) in which we observed the primer 3′-OH group moving closer to the α-phosphate (Fig. 4a, black arrow); and the density for site A* Mg^{2+} shifting toward the 3′-OH and the α-phosphate (Fig. 4b). Intriguingly, two conformations of the dATP's phosphate tail could be observed. The first one is similar to that of dATP in states R-1 and R-2, and we suggest that this is a ground state (dATP-g). However, we observed additional continuous density around the phosphate moiety that cannot be accounted for by only the dATP-g molecule (Fig. 4c and Supplementary Fig. 5a, b). Indeed, we modeled a second conformer of dATP, an intermediate state of dATP (dATP-i). In this conformation, the phosphate tail is rotated about the γ-dihedral bond, which determines the orientation of the phosphate tail in relation to the sugar moiety. The dATP-i is within coordination distance of the site A* Mg^{2+} ion and repositions the α-phosphate within 4.1 Å from the primer 3′-OH (Fig. 4b). Of note, these ATP conformers were observed in two individual experiments, highlighting that this alternative conformation is unlikely to be an artifact (Supplementary Fig. 5c).

Product states of the RT-T/P-dATP complex

We trapped two product states of the RT-T/P-dATP complex by incubating a solution with all reactants (RT-T/P, dATP, and $MgCl_2$) for 2 and 6 min (P_{2min} and P_{6min}, respectively) at 37 °C prior to grid preparation (local resolution of 2.2 and 2.3 Å around the active site, respectively; Supplementary Fig. 2e). Inspection of Coulomb potential maps for both states showed additional density for the 2′-O-methyl groups in the DNA aptamer at positions +3 and +5. For states R-1, R-2 and T_{8sec}, we observed density at positions +2 and +4 as expected (Fig. 5a); however, the state P_{2min} showed a mixture of partial occupancies at positions +2 and +3 and at +4 and +5 (Fig. 5b). Conversely, density in the state P_{6min} had entirely shifted to the +3 and +5, indicating that one nucleotide was fully added and translocated (Fig. 5c and Supplementary Fig. 6). Further analysis of P_{2min} and P_{6min} maps highlighted further difference for these states. The P_{2min} map shows that the position of the 3′-OH and site A* metal are similar to T_{8sec} (with additional coordination from residue D110) hence featuring a mixture of intermediate and product states (Supplementary Fig. 7a, b). The Coulomb potential map of P_{6min} shows two

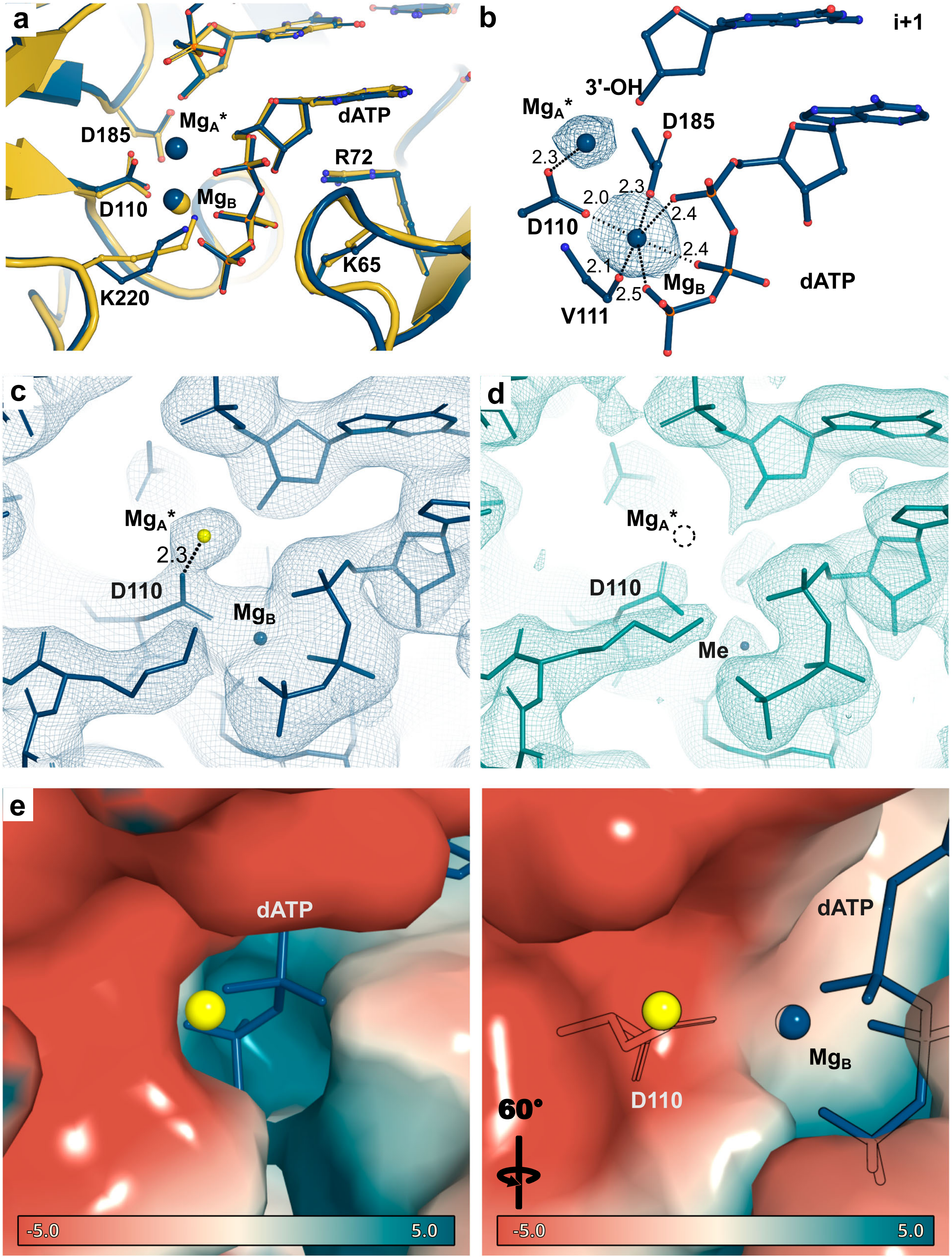


Fig. 3 | Cryo-EM structure of reactant state R-2 of the RT-T/P-dATP complex. **a** Overlay of the active sites of states R-1 (yellow) and R-2 (blue) of the RT-T/P-dATP complex. The protein backbone is shown in the cartoon representation; selected key residues are shown in the ball-and-stick representation. **b** Two Mg^{2+} sites, B and A*, are observed at state R-2. The map displayed as blue mesh (sharpening at B-factor = −80 Å^2 and contoured at 11σ). Coordination distances (in Å) are indicated next to dashed lines. **c** Unsharpened density map for state R-2, contoured at 10σ (blue mesh) showing density around site A* Mg^{2+} (yellow sphere); and **d** unsharpened density map, contoured at 10σ (teal mesh) for RT-T/P-dATP in the absence of MgCl$_2$ showing no density at the A* position (dashed circle). A mobile small density was observed at site B (labeled as Me: metal), but its identity could not be determined. **e** Two views of the surface representation of RT-T/P-dATP in state R-2 colored according to electrostatic potential, from red (−5 kT/e, negatively charged) to blue (+5 kT/e, positively charged), showing Mg^{2+} A* positioned in a negatively charged pocket. Mg^{2+} ion site A* is colored as a yellow sphere.

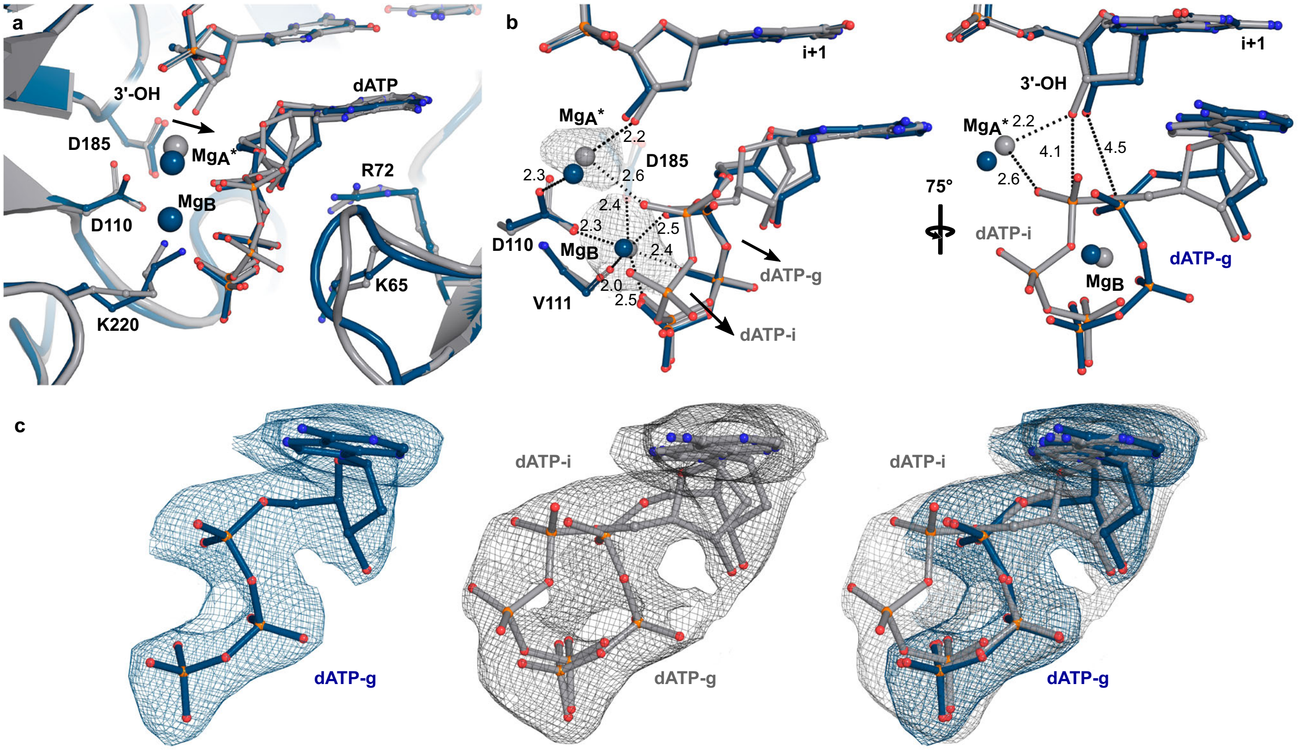


Fig. 4 | Cryo-EM structure of the intermediate state of the RT-T/P-dATP complex. **a** Overlay of the active sites of state R-2 (blue) and state T_{8sec} (gray) of the RT-T/P-dATP complex. The protein backbone is shown in the cartoon representation; selected key residues are shown in the ball-and-stick representation. **b** Coordination of Mg^{2+} at sites B and A* in the active site of RT by the two different dATP conformers (dATP-g and dATP-i) that are observed in the intermediate state of the RT-T/P-dATP complex. The map around the Mg^{2+} ions at sites B and A* is shown as a grey mesh and is contoured at 8.5σ (Map sharpened at B-factor −60 Å^2). Coordination distances (in Å) are indicated next to dashed lines. **c** Density around dATP-g in state R-2 of the RT-T/P-dATP complex (blue mesh, left panel) and around dATP-g and dATP-i in state T_{8sec} of the RT-T/P-dATP complex (gray mesh, middle panel). Overlay of the R-2 and T_{8sec} state maps contoured at 13σ and 8.5σ, respectively, is shown in the right panel. Maps sharpened at B-factors −60 Å^2 and −100 Å^2 for states R-2 and T_{8sec}, respectively.

conformations of the dATP in the binding pocket, which could be explained by a mixture of reactants and products with two chemical states of the substrate: one where the density of the α-phosphate stretches toward the 3′-OH of the primer to form a phosphodiester bond and a scissile PPi (Fig. 5d); and a second one corresponding to an intact dATP, with density for the α-β-phosphate bond (Fig. 5e). To further corroborate the presence of the two dATP states in the active site pocket we overlayed the P_{6min} map with that of the earliest state, R-1. Four major structural changes could be identified during addition (Fig. 5f): first, the primer 3′-OH swings 1.8 Å toward the incoming dATP; second, density for a phosphodiester bond emerges between the α-phosphate and the 3′-OH (black arrows); third, density weakens for the α- and β-phosphate bond (red arrows); and fourth, density at the phosphate tail is extended, indicative of PPi separation (green arrows). Indeed, increasing the density map threshold (to reveal only the highest scattering atoms) (Fig. 5g) shows that the α-phosphate center of density in state P_{6min} shifts within bonding distance of the 3′-OH (1.9 Å vs 4.3 Å for state R-1), and that the distance between the α- and β-phosphate has increased beyond bonding length (3.7 Å vs 3.0 Å in state R-1). The latter suggests that a high proportion of dATP reactant molecules have been incorporated. The β-phosphate of the leaving PPi is coordinated by the site B metal while the γ-phosphate is stabilized through hydrogen bonding with K220, the amide nitrogen of D113 and β3-β4 loop K65 (Fig. 5h). Overlay of P_{6min} with the crystal structure of an RT/AZT-terminated DNA and Foscarnet (PFA, a pyrophosphate mimic[20], PDB ID: 5HP1) shows major differences in the active site. The position of PFA and pyrophosphate do not overlap since the N_3 moiety in the AZT-MP forces a different PFA binding site. In addition, the freed pyrophosphate in our structure would clash with the non-natural N_3 moiety; and PFA would clash with K220. Overall, the two structures have significantly different chemistries, including atom type, charge, and bond distances (Supplementary Fig. 8).

In addition to Mg^{2+} ions at sites B and A*, density corresponding to the canonical site A metal[8,9] could be observed; however, density was weak (around the noise level for the map) and could only be observed after map sharpening (Supplementary Fig. 7c). This was not surprising as Mg^{2+} at this canonical A site has been detected only in few crystal structures of an RT-T/P-dNTP ternary complex[8–10]; yet all of these structures had nucleic acid scaffolds featuring 3′ chain terminated primers. We next tested whether we could observe a Mg^{2+} ion at site A mimicking the conditions that allowed its observation in the original reports, i.e., employing a 3′ chain terminated primer (Supplementary Fig. 7d, e). Indeed, the cryo-EM density map shows the presence of density at the site A position, and no additional density at site A* (Supplementary Fig. 7f). Thus, the absence/presence of the 3′-OH primer could cause fluctuations in the location of the Mg^{2+} ion, suggesting that site A and site A* observed positions could relate to a single, oscillating Mg^{2+} ion. A control experiment with a non-hydrolyzable dNTP (dApCpp) showed a bound, yet frozen (not-added) substrate (Supplementary Fig. 7g). The structure exhibits well-defined density for both sites B and A*. The latter metal is within coordination distance of D110, α-phosphate oxygen, and the primer's 3′-OH, highlighting the high mobility of this Mg^{+2}. Coordination of metal B showed small variations with the state R-2 structure, mainly due to differences in bond lengths (Supplementary Fig. 7h).

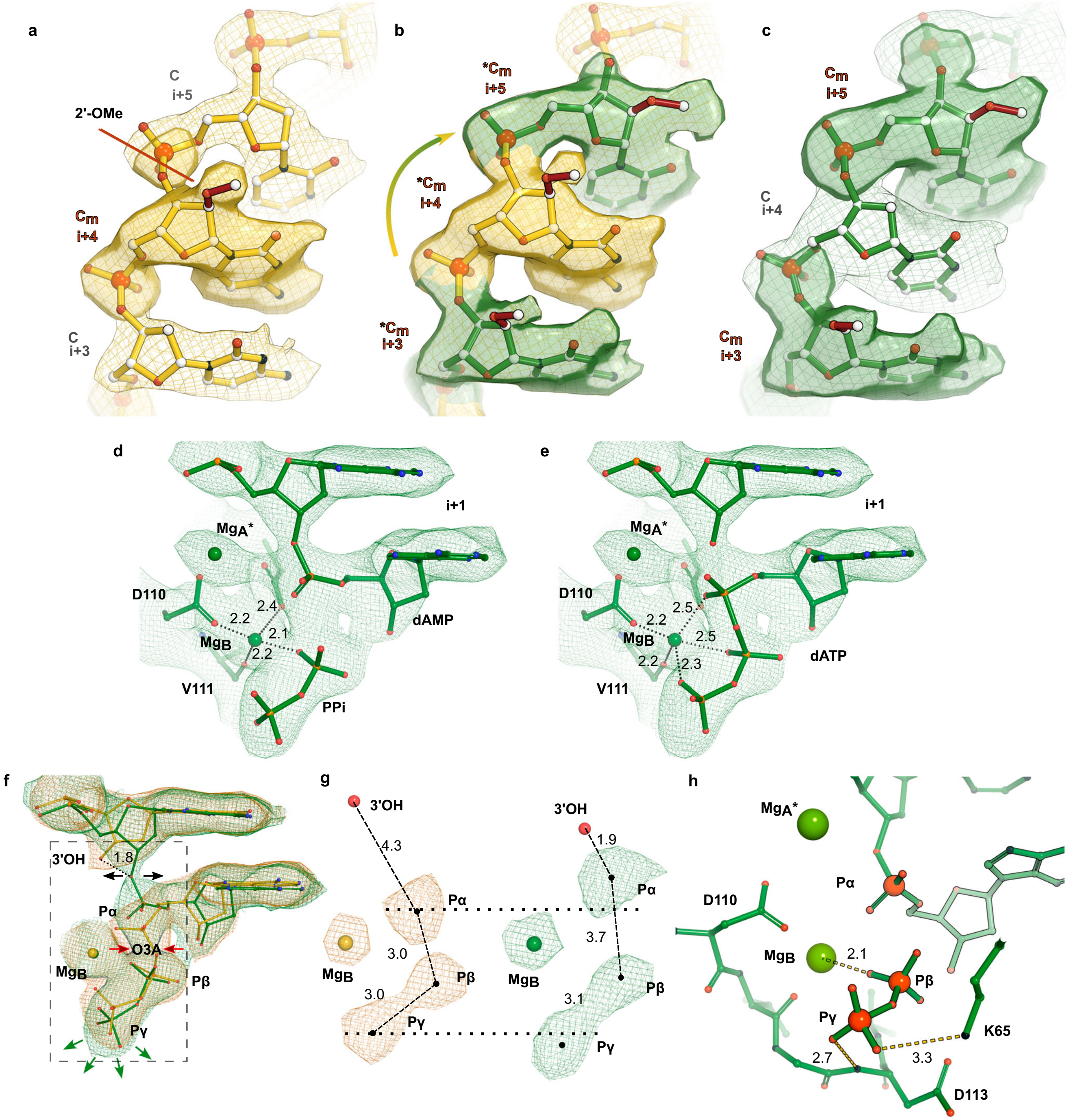
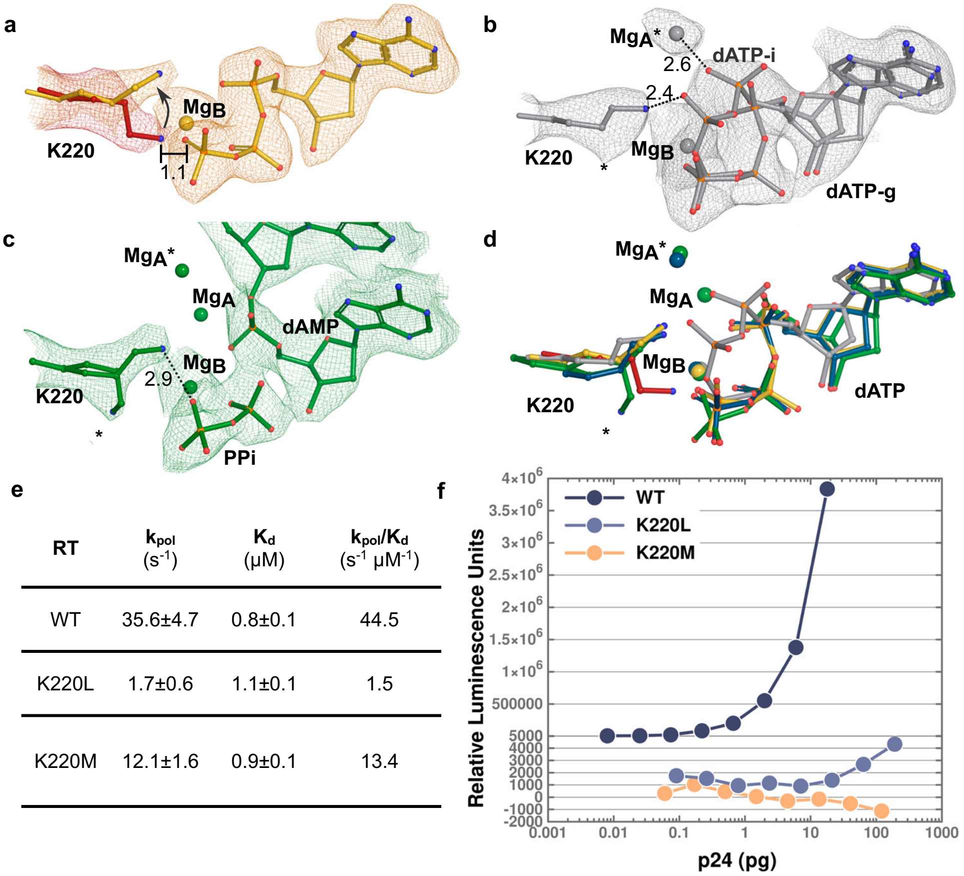


Fig. 5 | Cryo-EM structure of product states of the RT-T/P-dATP complex. Location of the 2′-O-methyl moiety (2′-OMe) in states **a** R-1, **b** P_{2min}, and **c** P_{6min} of the RT-T/P-dATP complex. **a** In states R-1, R-2, and T_{8sec} (R-2 and T_{8sec} not displayed for clarity purposes) density (yellow mesh) for the 2′-OMe (Cm, red sticks) is observed in position +2 and +4 and is clearly absent in positions +3 and +5 of the template strand (map displayed with the sharpening of B-factor = −60 Å^2 contoured at 12σ). **b** In-state P_{2min} density for 2′-OMe is observed in positions +2, +3, +4, and +5 as a mixture of translocated and non-translocated molecules. *Cm denotes partial occupancy of the 2′-OMe template scaffold. Map density is colored with yellow mesh at position +4, and green mesh at positions +3 and +5 (contoured at 7.5σ and sharpened at B-factor of −40 Å^2). **c** In P_{6min} density for 2′-OMe (green mesh) is observed only in positions +3 and +5, indicating all molecules have translocated one position (map displayed with B-factor sharpening = −40 Å^2 contoured at 8.5σ). **d**, **e** Ball-and-stick representation of the active site in state P_{6min} with a mixture of **d** products (added dAMP and pyrophosphate (PPi)) and **e** reactants (i.e., dATP). **f** Overlay of density maps of states R-1 and P_{6min} of the RT-T/P-dATP complex. Black arrows point to the growth of density for the newly formed phosphodiester bond; red arrows indicate thinning of density around O3A; green arrows point to stretched density for the leaving pyrophosphate (PPi) group (densities contoured at σ = 12.5 and σ = 9.5, for R-1 and P_{6min}, respectively. Maps are sharpened at B-factor of −60 Å^2). **g** Detail of density maps inside box in (**f**), displayed at high threshold (σ = 35 and σ = 14, for state R-1 and state P_{6min}, respectively. B-factor sharpening = −100 Å^2). **h** The PPi is stabilized by metal B and surrounding amino acids K65, D113, and K220 (Fig. 6c). Distances are indicated next to dashed lines (in angstroms).



RT	k_{pol} (s^{-1})	K_d (μM)	k_{pol}/K_d (s^{-1} $μM^{-1}$)
WT	35.6±4.7	0.8±0.1	44.5
K220L	1.7±0.6	1.1±0.1	1.5
K220M	12.1±1.6	0.9±0.1	13.4

Fig. 6 | Interaction of K220 in the RT active site. Structural changes of K220 during HIV-1 RT reaction. **a** The arrival of dATP to the active site forces a repositioning (Black arrow) in K220 to avoid clashes (red, dATP-free state, map in red mesh with B-factor −40 Å^2 and contoured at 6.5σ; yellow, state R-1, map in yellow mesh with B-factor −60 Å^2 and contoured at 12σ). **b** K220 directly interacts with the phosphate tail of dATP-i prior to nucleotide addition (gray, state T_{8sec}; map in gray mesh with B-factor −60 Å^2 and contoured at 8σ). **c** Following translocation (green, state P_{6min}, map in green mesh with B-factor −60 Å^2 and contoured at 8σ), a second conformation for K220 is detected (marked with *). **d** Overlay of all stages showing the movement of K220 during the dATP addition cycle (including state R-2, in blue). Coordination distances (in Å) are indicated next to dashed lines. **e** Transient kinetic analysis of nucleotidyl transfer by wild-type (WT), K220L, and K220M HIV-1 RT. Data represent mean ± SD ($n = 3$). **f** Replication capacity of wild-type, K220L and K220M HIV-1 in TZM-bl cells. Source data are provided as a Source Data file.

A role for K220 in RT during nucleotide addition

K220 in RT is conserved across all HIV-1 isolates and has been implicated in a general acid catalysis mechanism of nucleotidyl transfer[34]. In all of our cryo-EM structures, K220 formed key interactions with the incoming dATP molecule that were not previously described in X-ray crystal structures. In state R-1, the incoming dATP promotes repositioning of the K220 side chain (Fig. 6a) such that the Nζ moiety forms a salt bridge with the γ-phosphate of dATP. These interactions were maintained in state R-2. In state T_{8sec}, K220 interacts with the β-phosphate of the dATP-i conformation (Fig. 6b). We hypothesize that K220 together with the site A* Mg^{2+} induces and/or stabilizes the dATP-i conformation. In state P_{6min}, K220 displays two alternative conformations: in the first one, with higher occupancy, K220 forms a H-bond with the scissile pyrophosphate; while pointing away in the second one (Fig. 6c, d). To probe the role of K220 in nucleotide incorporation, we introduced either leucine or methionine at residue 220 by side-directed mutagenesis (Fig. 6e). Using transient kinetic analyses, and consistent with a prior study[34], we found that the K220L and K220M mutations in RT significantly diminished the rate of dATP incorporation (i.e., k_{pol}) without substantially impacting the affinity of the nucleotide for the RT active site (i.e., K_d), suggesting that K220 is involved in the catalytic event. We also introduced K220L and K220M into the RT gene of a subtype B HIV-1^{LAI} infectious virus clone. Both recombinant viruses displayed significantly attenuated replication capacity in TZM-bl cells (Fig. 6f), further highlighting the key role that K220 plays in viral reverse transcription.

Discussion

In this study, we used an on-grid diffusion/blot cryo-EM approach to trap intermediate states of nucleotide incorporation by HIV-1 RT with catalytically active substrates. This approach is easier to implement compared to other more complex cryo-EM grid preparation techniques[35,36]. In vitro at 37 °C, the rate of dNTP incorporation by HIV-1 RT, defined by the constant k_{pol}, is relatively fast, ranging anywhere from 0.1 to 35 s^{-1}, depending on multiple factors including the RT sequence as well as the T/P length and sequence. The methylated DNA aptamer used in this study is unique in that it binds to RT with high affinity (14 pM) compared to a double-stranded DNA with a similar configuration (~100 nM)[22]. DNA aptamers have been used in the past as HIV-1 RT inhibitors[37–39]. Given the high rate of RT's nucleotide incorporation (millisecond time scale), the use of a high affinity DNA aptamer[39] slowed down (significantly) translocation allowing capture of RT-T/P-dATP-(Mg^{2+}) structures in the seconds time scale. These structures reveal aspects of the chemistry of nucleotide addition.

During the transition from substrate binding to state R-1, the fingers subdomain and the active site pocket undergo significant rearrangements upon dNTP binding (Fig. 2e, f and Supplementary Fig. 3e, f). These changes include a rigid body rotation that propels the β3-β4 loop toward the dNTP resulting in partial closure of the active site pocket[8]. This swing motion could be driven by a combination of two factors: (1) dNTP Watson–Crick base-pairing of the templating nucleotide pulling the template strand which lies right above the loop (Fig. 2a, c), releasing the β3-β4 (stacking) residue L74 (see below); and (2) electrostatic interactions between the positively charged β3-β4 loop and the dNTP (state R-1) (Supplementary Fig. 3c). Notably, our data show that in state R-1, site B Mg^{2+} (but not site A* metal) arrives first, and coordinates all three oxygen atoms of the phosphate tail; it is possible that this metal could be carried into the binding site pre-bound to the dNTP as the K_d for Mg^{2+} binding to nucleotide is 29 μM[7].

Progression of state R-1 into R-2 involves binding of site A metal to a previously undescribed metal ion binding site in RT, site A* (Figs. 3 and 4). Structures for many classes of DNA polymerases have clearly defined density for Mg^{2+} sites A and B but, notably, a third Mg^{2+}ion has been observed for human DNA Pol η[26,40], Pol β[27] and Pol μ[28]. Binding of the third Mg^{2+} ion in the DNA Pol η, Pol β, and Pol μ structures was transient and occurred during phosphodiester bond formation where it has been shown to stabilize the leaving pyrophosphate group[26,41]. Our cryo-EM structure shows fundamental differences between the active sites of DNA polymerases and HIV-1 RT. Structural overlay of DNA Pol η with RT shows a third Mg^{2+} ion across site A and B metals (Supplementary Fig. 7i). On the other hand, the tightly closed pocket (Fig. 2f and Supplementary Fig. 3f) observed in RT sterically hinders potential binding of a third Mg^{2+} (metal C) at a homologous position. However, a small negatively charged pocket forms in the neighborhood of the aspartic acid triad, Asp110, Asp185, Asp186, and the 3′-OH of the primer template (Fig. 7a). This pocket can accommodate the binding of site A* metal at different positions.

Evolution of state R-2 into T_{8sec} required a boost of activation energy; this intermediate state was achieved with samples exposed to temperatures between 30 and 37° C immediately before plunge-freezing. The structure for the state T_{8sec} shows density in the Coulomb potential map that can accommodate two dATP conformers (Fig. 4 and Supplementary Fig. 5). The α-phosphate of dATP-i moves within ~4 Å of the 3′-OH primer and merges into the coordination sphere of the Mg^{2+} site A*. This observation is unique among all previous structural (and time-resolved) work on DNA/RNA polymerases since high-temperature data collection could not be achieved for X-ray crystallographic experiments.

Employing the nucleoside analog 4′-ethynyl-2-fluoro-2′-deoxyadenosine (EFdA)[10] Salie et al. obtained intermediates of the RT inhibition reaction including a (pre-catalytic) EFdA-MP complex bound to a 3′-OH chain terminated RT-T/P binary complex; and two post-incorporation complexes where EFdA-MP was extended by deoxythymidine-MP or an additional EFdA-MP. However, the former included a 3′-OH terminated primer (inert substrate), and the latter was thermodynamically stable. Conversely, our cryo-EM structures feature primer extensions of a catalytically active RT-T/P binary complex using natural substrates. In particular, it was possible to capture snapshots of a primer extension where the structure of P_{2min} evolved into P_{6min} (Fig. 5a–c). The former features a mixture of transient

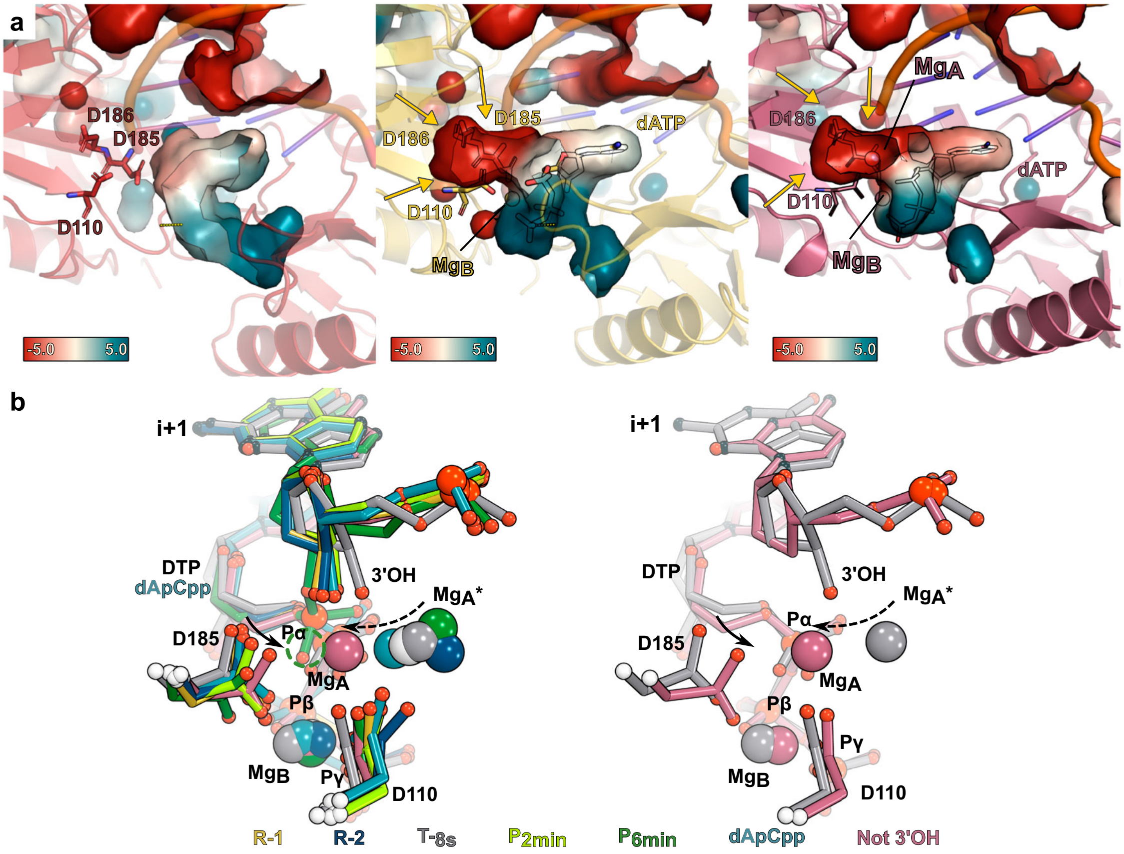


Fig. 7 | High mobility of Mg^{+2} at site A*. **a** Electrostatic surface potential of the dNTP binding site. Graphical surface representation of the charged cavities colored by the electrostatic potential for RT-T/P binary complex (left), state R-1 of RT-T/P-dATP (center), and RT-T/P-dATP 3′ chain terminated (right). A negatively charged pocket is formed upon dATP binding (yellow arrows). Pockets were detected using a 4 Å cavity detection radius and a 4 Å solvent radius cavity detection cutoff in Pymol. Protein backbone is shown in cartoon representation with key residues displayed in ball-and-stick representation. **b** Superposition of all states (left panel) reported in this study showing that site A* Mg^{2+} has high mobility, while the position of site B Mg^{2+} had minor changes in all experiments. Dash green circle corresponds to site A in P_{6min}. Right panel, overlay of the 3′OH chain terminated structure (violet) with the T_{8sec} structure shows that the Mg^{2+} ion shifts from the A* to the canonical A position (dash arrow) in the absence of a 3′OH group, with concomitant D185 residue rearrangement (solid arrow).

(i.e., position and coordination of metal A*) and product states with a population of non-translocated and translocated particles (i.e., partial occupancies of the 2′-O-methylated dCMPs at positions +2, +3, +4 and +5). Conversely, in $P_{6\text{min}}$ most or all particles have translocated one position. Moreover, state $P_{6\text{min}}$ captures two conformers of the dATP in the binding pocket: (1) a pre-translocated dATP, or a reacting molecule, and (2) an inserted dAMP forming a phosphodiester bond with the primer and the released PPi, or reaction products. PPi is coordinated by metal B and stabilized through H-bonds with the surrounding residues. PPi release is likely concomitant with metals release as previous structures of RT with incorporated nucleotides on the N-site do not retain Mg^{+2} in either sites A or B[10]; however, our structures do not allow us to confirm these observations.

The role of D186 has long been reported as part of the catalytic triad in HIV-1 RT[42,43] and has been associated with site A Mg^{2+} coordination. Intriguingly, we did not observe interactions of D186 with Mg^{2+} in our structures. However, analysis of the positions adopted by D186 in six previously reported structures indicates that D186 can be observed either pointing toward site A metal[9,20,44] or away[8,10,45] (Supplementary Fig. 9). The proposed architecture of the conserved mechanism of nucleotide addition by DNA polymerases[13], where a catalytic triad comprising three aspartates coordinate metals A and B was not fully observed in our structures. However, this was not completely unexpected given that the classic 3′-OH terminated structure of RT by Harrison and coworkers[8], shows that site A Mg^{2+} is only coordinated by D110 and D185, with D186 lying 3.9 Å away from site A (Supplementary Fig. 9a). Therefore, while it is likely that the conserved mechanism applies to RT, no structural evidence is yet available to support it. Thus, it is possible that for RT, the canonical architecture could just be achieved during the transition state.

A crucial event during nucleotide addition is the transient deprotonation of the 3′-OH for nucleophilic attack of the α-phosphate[46–48]; this event could be aspartate- or water-as-base-mediated[47]. In the former, RT's D185 could act as the general base to deprotonate the 3′-OH; however, molecular simulations of aspartic acid residues (in DNA polymerase η) suggested that aspartate-deprotonation could be energetically unfavorable. In the water-as-base mechanism, which is energetically favored, a water molecule, originating from Mg^{2+}-coordinated water, could deprotonate de 3′-OH and thus trigger phosphoryl transfer. A water molecule coordinated by site A* metal may well be responsible for such a nucleophilic attack. Our data shows the presence of two metals in HIV-1 RT's active site, metals B and A*. Metal B is unlikely to be involved since it is fully coordinated by three dNTP phosphate oxygens and three neighboring residues (Fig. 2b); conversely, given that site A* Mg^{2+} located next to the α-phosphate and has a low (residue) coordination number, it thus might provide the water molecule involved in 3′-OH deprotonation. Moreover, this event may be possibly triggered during site A* coordination by the dATP-i conformer (Fig. 4b, c).

Intriguingly, no residue acting as a general acid (i.e., donating a proton) to the leaving pyrophosphate has been identified for HIV-1 RT; the structural and biochemical data presented here suggest that K220 may perhaps be this residue. Hence, it is possible that a water molecule coordinated by site A* metal could function as the general base, and K220 could function as the general acid[34].

RT-mediated DNA synthesis has been proposed to follow the two-metal-ion mechanism similar to other nucleic acid polymerases[7,13,49]. In this model, metal ion A binds near the primer terminus and lowers the pK_a of 3′-OH group, thereby activating the nucleophile and facilitating its attack on the α-phosphate of the incoming dNTP. Metal ion B forms a tight complex with the incoming dNTP via coordination with the non-bridging oxygens of the phosphates and neutralizes the developing negative charge during chemistry. This model was supported by RT-T/P-dNTP structures obtained with catalytic inactive primers to trap the complexes[8,9]. The structural data presented here suggest the possibility that metal A is not fixed, hence varying experimental conditions such as the presence or absence of the 3′-OH cause fluctuations in its position. This metal arrives at a previously undescribed position A*, drifting around this position during the reaction, and is observed at position A during phosphodiester bond formation (Fig. 7b). The structures presented here shed light on the chemistry of nucleotide addition, namely, the role of metal A* to drive the chemical reaction.

A nucleotide addition cycle

The described intermediate states suggest a structural pathway to nucleotide addition, i.e., a structural dNTP addition cycle (Fig. 8 and Supplementary Movie 1). (1) In the RT-T/P binary complex, the rigid positively charged β3-β4 loop is in the open conformation exposing the templating nucleotide (i-1). (2) The incoming dNTP binds in the active site pocket and recruits site B metal. (3) Watson–Crick pairing of the dNTP triggers sliding of the DNA/RNA hybrid template releasing the β3-β4 loop stacking residue L74, promoting interactions between the positively charged β3-β4 loop and the negatively charged phosphate tail closing the binding pocket. (4) dNTP in the active site pocket attracts site A* metal. (5) Coordination of site A* metal elicits a precatalytic conformation. (6) Concurrence of the primer 3′-OH, the α-phosphate and site A* metal trigger phosphodiester bond formation, and PPi release. (7) The DNA/RNA hybrid translocation swings back the β3-β4 loop toward its open conformation.

The reactant, intermediate, and product states presented here help to define the conformational changes that facilitate nucleotide incorporation. Collectively, this study provides insights into a fundamental chemical reaction that is central to HIV-1 replication.

Methods

RT expression and purification

The expression plasmid RT152A was obtained as a kind gift from Professor Eddy Arnold (Rutgers University)[22]. The protein was expressed in BL21(DE3) with 0.5 mM isopropyl β-d-1-thiogalactopyranoside at 16 °C overnight. Cells were lysed by sonication in 20 mM Tris pH 8.0, 150 mM NaCl, 5% Glycerol. After centrifugation at 18,000 rpm for 40 min, the supernatant was loaded to Nickel beads (HisTrap FF, Cytiva) and eluted with an imidazole gradient from 0 mM to 500 mM. Collected fractions were analyzed with SDS-PAGE, and selected peak fractions were concentrated and loaded onto a heparin column (HiTrap heparin HP, Cytiva). Eluted fraction was concentrated and loaded to Superdex 200 10/300 GL column equilibrated with 20 mM Hepes pH 7.5, 100 mM NaCl, 5% Glycerol. The peak fractions were concentrated to 10 mg/mL, aliquoted, and stored at −80 °C.

Preparation of the RT-T/P binary complex

The 38-mer DNA-hairpin T/P aptamer used in this study (5′-TAAT $TCC_mCC_mCCCTTCGGTGCTTTGCACCGAAGGGGGGG$-3′) was purchased from Integrated DNA Technologies (Coralville, IA) (Fig. 1c). The RT-T/P was prepared by mixing RT with the T/P at a molar ratio of 1:1.3. The mixture was set on ice for 30 min before loading onto a Superdex 200 5/150 GL column equilibrated with 25 mM Tris-HCl pH 7.5, 100 mM NaCl, to remove the excess of DNA. The RT-T/P peak fraction was used for cryo-EM grid preparation.

Cryo-EM grid preparation

We prepared cryo-EM grids with slightly different protocols for the eleven structures reported here (Supplementary Table 1). UltrAuFoil R1.2/1.3 300-mesh grids (Quantifoil Micro Tools GmbH, Großlöbichau, Germany) were freshly glow-discharged for 60 s and used immediately. If a solution was applied on both sides of the grid, each side of the grid was glow-discharged. Grids were vitrified with liquid ethane using a Vitrobot Mark IV (Thermo Fischer Scientific, USA) set at 22 °C (except for intermediate state $T_{8\text{sec}}$), 100% humidity, and a blotting time of 3 s. The final concentration of the liquid drop on the grid was 2.1 μM RT-T/P, 0.4 mM dATP, and 1 mM $MgCl_2$. Solutions were kept on

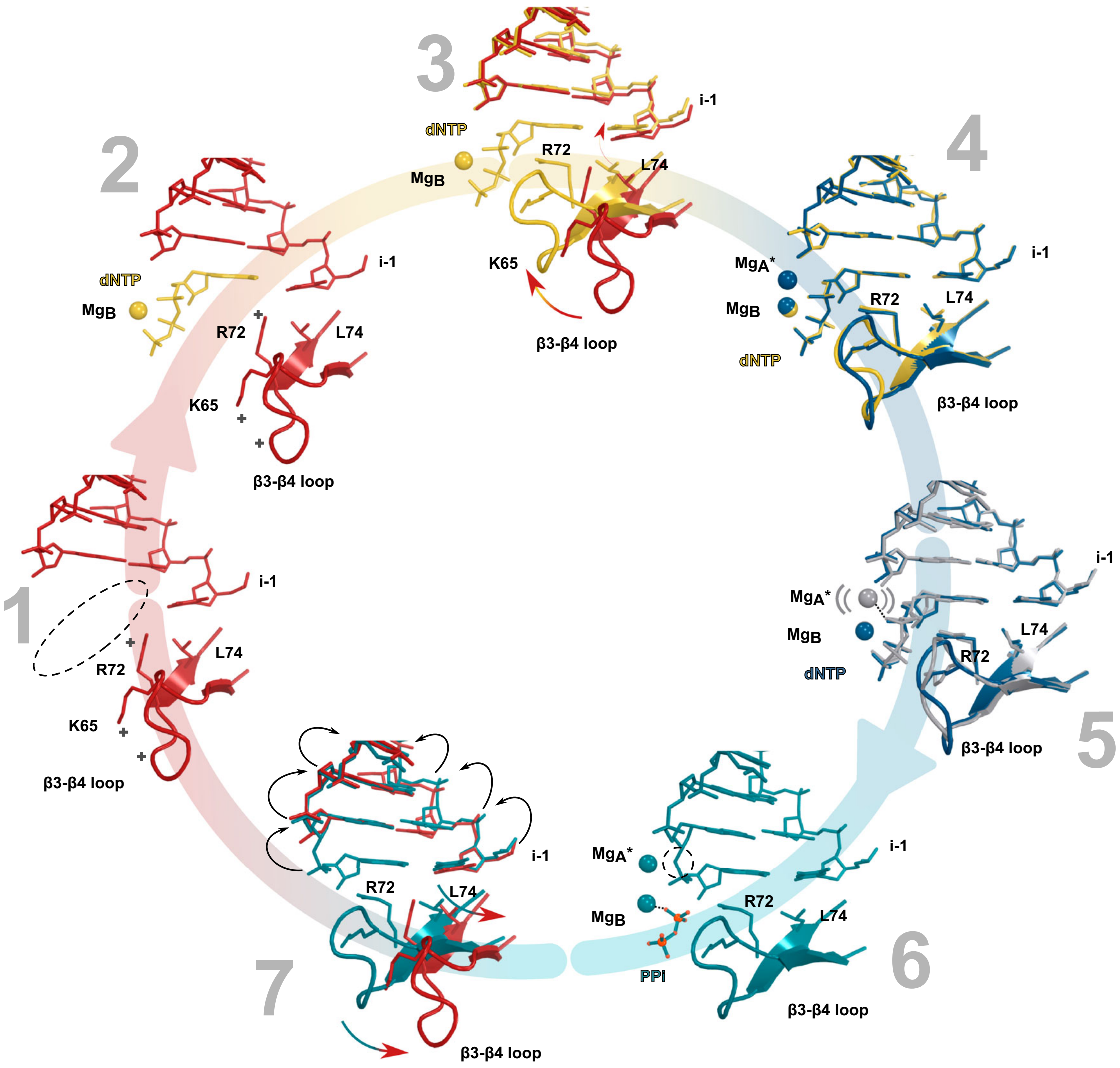


Fig. 8 | Proposed nucleotide addition cycle. We observed five unique intermediate states (R-1, R-2, T_{8sec}, P_{2min}, P_{6min}) that suggest a possible cycle for nucleotide addition. See the text for a detailed description. (**1**) dNTP-free (dash ellipse) RT-T/P binary complex (red). + symbol indicates positive charges. (**2**) In state R-1 (yellow), a dNTP binds in the active site pocket and recruits site B metal. (**3**) Binding of dNTP triggers active site closure. (**4**) dNTP in the active site pocket attracts site A* metal (state R-2, blue). (**5**) Coordination of site A* metal elicits a precatalytic conformation (state T_{8sec}, gray). Curved lines around site A* metal indicate that its position is mobile. (**6**) Concurrence of the primer 3′-OH, the α-phosphate, and site A* metal (state P_{2min} and P_{6min}, teal) trigger phosphodiester bond formation (dash circle) and PPi release. (**7**) Upon translocation (black arrows), the β3-β4 loop returns to its open conformation (teal/red arrow) to restart the cycle.

ice unless stated otherwise. For reactant states R-1 (data set 2) and R-2 (data sets 3 and 4) and intermediate states T_{8sec} (data set 5) and T'_{8sec} (data set 6) we performed on-grid diffusion/blot experiments where two separate solutions are applied on the opposite sides of the cryo-EM grid and mixing occurs on the grid by diffusion through the 1.2 µm diameter holes in the UltrAuFoil support (Supplementary Fig. 1). Application of the second solution starts the RT reaction. For state R-1, 2 µl of RT-T/P (4.2 µM) was applied on one side of the grid, and 2 µl of the second solution (0.8 mM dATP, 2 mM $MgCl_2$) was applied to the opposite side of the grid, and the blotting/freezing cycle was immediately started (~4 s of diffusion during blotting/freezing). For state R-2, 2 µl of RT-T/P preincubated with dATP (5 min, no $MgCl_2$) were applied on the first side of the grid, and 2 µl of solution containing $MgCl_2$ were applied on the opposite side (blotting/freezing immediately). For state T_{8sec} (data set 5) the Vitrobot and reactant solutions were pre-equilibrated at 37 °C (or 30 °C, for T'_{8sec} data set 6) before solutions were applied to the grid as described for state R-2, but with a delay of 8 s before the blotting/freezing cycle was started. For the remaining cases (data sets 1, and 7 through 11), regular grid preparation protocols were followed, with 3 µl of sample applied to one side of the grid and the blotting/freezing cycle performed immediately. In the case of states P_{2min} (data set 7) and P_{6min} (data set 8), all components of the reaction were incubated in solution at 37 °C on a heat block for 2 min or 6 min, respectively, and then used immediately for grid preparation. For data set 9, the non-hydrolyzable dATP analog dApCpp (2′-Deoxyadenosine-5′-[(α,β)-methyleno]triphosphate, Sodium salt, Jena Bioscience, Germany) was incubated with RT-T/P and $MgCl_2$ at 37 °C for 15 min. To obtain the structure of RT-T/P-dATP in the absence of $MgCl_2$ (data set 10), RT-T/P purification buffer was supplemented with 1 mM EDTA; dATP was preincubated for 1 h with EDTA (4 mM dATP, 40 mM EDTA) and subsequently all components (RT-T/P-dATP) were incubated 10 min at room temperature (final concentration 0.4 mM dATP, 4 mM EDTA). For the structure of the 3′ chain terminated primer (data set 11), RT-T/P was prepared as described above with a ddCMP-terminated DNA aptamer (5′-TAATTGC$_m$CC$_m$CCCTTCGGTGCTTTGCACCGAAGGGGGG-ddC-3′) (Integrated DNA Technologies). In this case, the final concentration of $MgCl_2$ was increased to 13 mM.

Data collection

Cryo-EM data collection was performed using EPU software on a Titan Krios 3Gi transmission electron microscope operated at 300 kV (Thermo Fisher Scientific, USA). Movies were collected on a Falcon 3 direct electron detecting CMOS camera at a nominal magnification of 96,000× corresponding to a calibrated pixel size of 0.825 Å (Data sets 1–6 and 8). Data set 7 was collected on a Falcon 4i camera (96,000× magnification, 0.82 Å pixel size). For data sets 9, 10, and 11, the microscope had been upgraded to include a Selectris energy filter in addition to the Falcon 4i camera; movies were recorded at 165,000× corresponding to a calibrated pixel size of 0.72 Å/pixel, using a slit width of 10 eV. Defocus values ranged from −0.75 μm to −2.25 μm. Movies were recorded in counting mode with a total dose of ~50–60 e^- per $Å^2$. Further details of data collection are summarized in Supplementary Table 1.

Cryo-EM data processing

All cryo-EM data processing was performed using the CryoSPARC v3.2 pipeline[50]. Movie stacks were aligned using patch motion correction followed by defocus estimation with patch CTF estimation. For the RT-T/P complex (data set 1), a subset of 200 images was used for automatic particle picking using a blob picker with 60 Å and 130 Å as minimum and maximum particle diameters, respectively. After extensive 2D classification, a selection of the best 2D class averages was used as templates to pick particles from the full data set. Particles were subjected to reference-free 2D classification. A subset of 20% of particles selected after 2D classification were used for heterogeneous ab initio reconstruction with 2 classes. One of the classes resembles the expected 3D map for RT while the other class seems to deviate (Supplementary Fig. 10). We used these two classes to represent a good class and a junk class for 3D heterogeneous refinement. After this clean-up, the final good particle set was used for non-uniform 3D Refinement[51]. For the remaining data sets, the 3D reconstruction of data set 1 was used to generate projection-based templates for automatic particle picking. After reference-free 2D classification, particles were submitted to 3D heterogeneous refinement using the two classes generated ab initio for the RT-T/P complex (data set 1). This strategy allowed us to remove junk particles still present in the stack after 2D classification by sinking them into the junk 3D class. Additional 3D classification with a single reference did not detect underlying heterogeneity or improve the final reconstruction. Cleaned particle stacks were used for non-uniform 3D Refinement (Supplementary Fig. 11). Details of data processing are shown in Supplementary Fig. 10 and Supplementary Table 1. Resolutions were estimated by Fourier shell correlation using the 0.143-criterion according to the gold-standard method. Local resolutions were also calculated in cryoSPARC using BlocRes algorithm[52] (Supplementary Fig. 2). The reported local resolution around the active site was calculated based on the local resolution maps. Fragments of the cryo-EM maps were extracted in Coot v0.9[53] with a 10 Å cutoff using the α-phosphate of the dATP molecule as the center of the extraction box. The extracted local resolution maps were used to calculate the average local resolution value around the active site using ChimeraX v.1.5[54] (Supplementary Fig. 2 and Supplementary Table 1). Directional resolution was evaluated by running Orientation Diagnostics jobs in CryoSPARC v.4.4.1 with a half-angle cone of 20° and 3072 total directions. Sphericity was calculated in the 3DFSC server with a volume threshold at 0.5[55].

Model building and refinement

The X-ray structure of the RT-T/P complex (PDB ID: 5D3G) was used as a starting model and docked into the cryo-EM maps by MOLREP[56] inside the CCP-EM interface. A series of sharpened maps (B-factors −20, −40, −60, −80, −100, −120 $Å^2$) were obtained with the sharpening tool of cryoSPARC and used to manually adjust the model with Coot v0.9[53]. Mg^{2+} ions and dATP were placed into the densities in unsharpened maps. Sharpened maps were used to model side chain density and the models were then refined with PHENIX v1.18.2[57] using real-space refinement. After this refinement, each residue of the model, in particular those of the active site, was manually checked and refined in Coot. Structural models were validated with MolProbity[58]. The final refinement statistics were obtained against unsharpened maps (Supplementary Fig. 12) and are summarized in Supplementary Table 1. Electrostatic potentials were calculated using the APBS plugin inside Pymol v.2.5.0 (Schrödinger, LLC).

Pre-steady-state kinetic analyses

Rapid quench experiments were carried out using a KinTek RQF-3 instrument (KinTek Corporation, Clarence, PA, USA), using a 19-nucleotide DNA primer (5′-GTCCCTGTTCGGGCGCCAC-3′) annealed to a 45-nucleotide DNA template (5′-TAGTCAGAATGGAAAATCTCTAGC AGTGGCGCCCGAACAGGGACA-3′)[59] purchased from Integrated DNA Technologies (Coralville, IA). Pre-steady-state data were fitted by nonlinear regression using Sigma Plot software (Jandel Scientific). For the burst experiments, the following equation was used: $[T/P_{+1}] = A[1 - \exp(-k_1 t) + mt)$, where A represents the burst amplitude, k_1 the burst rate, and m the slope. For single turnover experiments, the following equation was used: $[T/P_{+1}] = A(1\text{-}e^{-kobst})$. K_d and k_{pol} values were calculated by fitting the observed single rate constants (k_{obs}) obtained at different concentrations of dNTP to the hyperbolic expression as follows: $k_{obs} = k_{pol}[dNTP]/(K_d + [dNTP])$, where K_d is the equilibrium dissociation constant for the interaction of dNTP with the RT-T/P complex and k_{pol} is the maximum first-order rate constant for dNTP incorporation.

HIV-1 replication capacity assays

The genes for K220L and K220M RT were cloned into a subtype B HIV-1_{LAI} clone. HEK293 cells (ATCC) were used to produce the virus which was quantified by HIV-1 p24 ELISA (ZeptoMetrix). Replication capacity was determined by adding different amounts of virus (pg of p24) to TZM-bl cells, and the relative luminescence units (RLU) were quantified. RLU for each experiment was measured by duplicate. The experiment was done twice for Wild-Type (WT) RT and K220L RT and once for K220M RT. TZM-bl cells were obtained from the HIV AIDS repository (www.beiresources.org, ARP-8129).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Cryo-EM maps and coordinates have been deposited in the Electron Microscopy Data Bank (EMDB) and the Protein Data Bank (PDB), respectively, with accession numbers EMD-43114 and 8VB6 (RT-T/P); EMD-43115 and 8VB7 (RT-T/P-DNA state R-1); EMD-43116 and 8VB8 (RT-T/P-DNA state R-2); EMD-43117 and 8VB9 (RT-T/P-DNA state R-2′); EMD-43120 and 8VBC (RT-T/P-DNA state T_{8sec}); EMD-43121 and 8VBD (RT-T/P-DNA T'_{8sec}); EMD-43122 and 8VBE (RT-T/P-DNA P_{2min}); EMD-43123 and 8VBF (RT-T/P-DNA P_{6min}); EMD-43124 and 8VBG (RT-T/P-dApCpp); EMD-43125 and 8VBH (RT-T/P-DNA No $MgCl_2$); and EMD-43126 and 8VBI (RT-T/P-DNA 3′-terminated primer). Source data are provided with this paper.

References

1. Sarafianos, S. G. et al. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J. Mol. Biol.* **385**, 693–713 (2009).
2. Coffin, J. M. & Fan, H. The discovery of reverse transcriptase. *Annu. Rev. Virol.* **3**, 29–51 (2016).
3. Xavier Ruiz, F. & Arnold, E. Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Curr. Opin. Struct. Biol.* **61**, 113–123 (2020).

4. Kati, W. M., Johnson, K. A., Jerva, L. F. & Anderson, K. S. Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.* **267**, 25988–25997 (1992).
5. Reardon, J. E. Human immunodeficiency virus reverse transcriptase. A kinetic analysis of RNA-dependent and DNA-dependent DNA polymerization. *J. Biol. Chem.* **268**, 8743–8751 (1993).
6. Hsieh, J. C., Zinnen, S. & Modrich, P. Kinetic mechanism of the DNA-dependent DNA polymerase activity of human immunodeficiency virus reverse transcriptase. *J. Biol. Chem.* **268**, 24607–24613 (1993).
7. Gong, S. et al. Kinetic and thermodynamic analysis defines roles for two metal ions in DNA polymerase specificity and catalysis. *J. Biol. Chem.* **296**, 100184 (2021).
8. Huang, H., Chopra, R., Verdine, G. L. & Harrison, S. C. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669–1675 (1998).
9. Das, K., Martinez, S. E., Bandwar, R. P. & Arnold, E. Structures of HIV-1 RT-RNA/DNA ternary complexes with dATP and nevirapine reveal conformational flexibility of RNA/DNA: insights into requirements for RNase H cleavage. *Nucleic Acids Res.* **42**, 8125–8137 (2014).
10. Salie, Z. L. et al. Structural basis of HIV inhibition by translocation-defective RT inhibitor 4′-ethynyl-2-fluoro-2′-deoxyadenosine (EFdA). *Proc. Natl Acad. Sci. USA* **113**, 9274–9279 (2016).
11. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA* **90**, 6498–6502 (1993).
12. Atis, M., Johnson, K. A. & Elber, R. Pyrophosphate release in the protein HIV reverse transcriptase. *J. Phys. Chem. B* **121**, 9557–9565 (2017).
13. Steitz, T. A. A mechanism for all polymerases. *Nature* **391**, 231–232 (1998).
14. Ding, J. et al. Structure and functional implications of the polymerase active site region in a complex of HIV–1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution. *J. Mol. Biol.* **284**, 1095–1111 (1998).
15. Sarafianos, S. G. et al. Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**, 1449–1461 (2001).
16. Larsen, K. P. et al. Architecture of an HIV-1 reverse transcriptase initiation complex. *Nature* **557**, 118–122 (2018).
17. Bertoletti, N., Chan, A. H., Schinazi, R. F., Yin, Y. W. & Anderson, K. S. Structural insights into the recognition of nucleoside reverse transcriptase inhibitors by HIV-1 reverse transcriptase: First crystal structures with reverse transcriptase and the active triphosphate forms of lamivudine and emtricitabine. *Protein Sci.* **28**, 1664–1675 (2019).
18. Lansdon, E. B. et al. Visualizing the molecular interactions of a nucleotide analog, GS-9148, with HIV-1 reverse transcriptase-DNA complex. *J. Mol. Biol.* **397**, 967–978 (2010).
19. Das, K., Martinez, S. E., Bauman, J. D. & Arnold, E. HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism. *Nat. Struct. Mol. Biol.* **19**, 253–259 (2012).
20. Das, K. et al. Conformational states of HIV-1 reverse transcriptase for nucleotide incorporation vs pyrophosphorolysis-binding of foscarnet. *ACS Chem. Biol.* **11**, 2158–2164 (2016).
21. Ruiz, F. X., Hoang, A., Dilmore, C. R., DeStefano, J. J. & Arnold, E. Structural basis of HIV inhibition by L-nucleosides: Opportunities for drug development and repurposing. *Drug. Discov. Today* **27**, 1832–1846 (2022).
22. Miller, M. T., Tuske, S., Das, K., DeStefano, J. J. & Arnold, E. Structure of HIV-1 reverse transcriptase bound to a novel 38-mer hairpin template-primer DNA aptamer. *Protein Sci.* **25**, 46–55 (2016).
23. Tuske, S. et al. Integrative structural biology studies of HIV-1 reverse transcriptase binding to a high-affinity DNA aptamer. *Curr. Res. Struct. Biol.* **2**, 116–129 (2020).
24. Singh, A. K. et al. Cryo-EM structures of wild-type and E138K/M184I mutant HIV-1 RT/DNA complexed with inhibitors doravirine and rilpivirine. *Proc. Natl Acad. Sci. USA* **119**, e2203660119 (2022).
25. Singh, A. K. et al. Sliding of HIV-1 reverse transcriptase over DNA creates a transient P pocket - targeting P-pocket by fragment screening. *Nat. Commun.* **12**, 7127 (2021).
26. Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. J. & Yang, W. Watching DNA polymerase eta make a phosphodiester bond. *Nature* **487**, 196–U177 (2012).
27. Freudenthal, B. D., Beard, W. A., Shock, D. D. & Wilson, S. H. Observing a DNA polymerase choose right from wrong. *Cell* **154**, 157–168 (2013).
28. Jamsen, J. A. et al. Time-lapse crystallography snapshots of a double-strand break repair polymerase in action. *Nat. Commun.* **8**, 253 (2017).
29. Prisant, M. G., Williams, C. J., Chen, V. B., Richardson, J. S. & Richardson, D. C. New tools in MolProbity validation: CaBLAM for CryoEM backbone, UnDowser to rethink "waters," and NGL Viewer to recapture online 3D graphics. *Protein Sci.* **29**, 315–329 (2020).
30. Pintilie, G. et al. Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nat. Methods* **17**, 328–334 (2020).
31. Pintilie, G. & Chiu, W. Validation, analysis and annotation of cryo-EM structures. *Acta Crystallogr. D Struct. Biol.* **77**, 1142–1152 (2021).
32. Zhang, K., Pintilie, G. D., Li, S., Schmid, M. F. & Chiu, W. Resolving individual atoms of protein complex by cryo-electron microscopy. *Cell Res.* **30**, 1136–1139 (2020).
33. Lin, G. et al. Structural basis of transcription: RNA polymerase II substrate binding and metal coordination using a free-electron laser. *Proc. Natl Acad. Sci. USA* **121**, e2318527121 (2024).
34. Castro, C. et al. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat. Struct. Mol. Biol.* **16**, 212–218 (2009).
35. Dandey, V. P. et al. Time-resolved cryo-EM using Spotiton. *Nat. Methods* **17**, 897–900 (2020).
36. Klebl, D. P., White, H. D., Sobott, F. & Muench, S. P. On-grid and in-flow mixing for time-resolved cryo-EM. *Acta Crystallogr. D. Struct. Biol.* **77**, 1233–1240 (2021).
37. Kissel, J. D., Held, D. M., Hardy, R. W. & Burke, D. H. Single-stranded DNA aptamer RT1t49 inhibits RT polymerase and RNase H functions of HIV type 1, HIV type 2, and SIVCPZ RTs. *AIDS Res. Hum. Retrovir.* **23**, 699–708 (2007).
38. Ratanabunyong, S. et al. Biophysical characterization of novel DNA aptamers against K103N/Y181C double mutant HIV-1 reverse transcriptase. *Molecules* **27**, 285 (2022).
39. DeStefano, J. J. & Nair, G. R. Novel aptamer inhibitors of human immunodeficiency virus reverse transcriptase. *Oligonucleotides* **18**, 133–144 (2008).
40. Gao, Y. & Yang, W. Capture of a third Mg2+ is essential for catalyzing DNA synthesis. *Science* **352**, 1334–1337 (2016).
41. Stevens, D. R. & Hammes-Schiffer, S. Exploring the role of the third active site metal ion in DNA polymerase eta with QM/MM free energy simulations. *J. Am. Chem. Soc.* **140**, 8965–8969 (2018).
42. Lowe, D. M., Parmar, V., Kemp, S. D. & Larder, B. A. Mutational analysis of two conserved sequence motifs in HIV-1 reverse transcriptase. *FEBS Lett.* **282**, 231–234 (1991).
43. Kaushik, N. et al. Biochemical analysis of catalytically crucial aspartate mutants of human immunodeficiency virus type 1 reverse transcriptase. *Biochemistry* **35**, 11536–11546 (1996).
44. Das, K., Martinez, S. E. & Arnold, E. Structural insights into HIV reverse transcriptase mutations Q151M and Q151M complex that confer multinucleoside drug resistance. *Antimicrob. Agents Chemother.* **61**, e00224–17 (2017).
45. Tuske, S. et al. Structures of HIV-1 RT-DNA complexes before and after incorporation of the anti-AIDS drug tenofovir. *Nat. Struct. Mol. Biol.* **11**, 469–474 (2004).
46. Yoon, H. & Warshel, A. Simulating the fidelity and the three Mg mechanism of pol eta and clarifying the validity of transition state theory in enzyme catalysis. *Proteins* **85**, 1446–1453 (2017).
47. Gregory, M. T., Gao, Y., Cui, Q. & Yang, W. Multiple deprotonation paths of the nucleophile 3′-OH in the DNA synthesis reaction. *Proc. Natl Acad. Sci. USA* **118**, e210399011 (2021).

48. Roston, D., Demapan, D. & Cui, Q. Extensive free-energy simulations identify water as the base in nucleotide addition by DNA polymerase. *Proc. Natl Acad. Sci. USA* **116**, 25048–25056 (2019).
49. Steitz, T. A. DNA-dependent and Rna-dependent DNA-polymerases. *Curr. Opin. Struct. Biol.* **3**, 31–38 (1993).
50. 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 (2017).
51. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
52. Cardone, G., Heymann, J. B. & Steven, A. C. One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J. Struct. Biol.* **184**, 226–236 (2013).
53. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
54. Meng, E. C. et al. UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
55. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
56. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25 (2010).
57. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
58. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
59. Radzio, J. & Sluis-Cremer, N. Stereo-selectivity of HIV-1 reverse transcriptase toward isomers of thymidine-5′-O-1-thiotriphosphate. *Protein Sci.* **14**, 1929–1933 (2005).

Acknowledgements

We thank Professor Eddy Arnold and Dr. Francesc Ruiz for the kind gift of RT152A plasmid and technical assistance. We also thank Professor Craig D. Kaplan for his critical reading of the manuscript and Professor Jonathan Coleman for his suggestions on data processing. This research was supported by NIH Grants P50 GM082251 and R01 AI175067 (to G.C.). All Cryo-EM data were collected at the Cryo-Electron microscopy Facilities at the Department of Structural Biology. The University of Pittsburgh Titan Krios microscope and Falcon 3 camera were supported by the Office of the Director, National Institutes of Health, under award numbers S10 OD025009 and S10 OD019995. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

X.Z. selected the constructs and performed protein purification. M.A. performed HIV replication assays. S.V. prepared cryo-EM samples. S.V. and J.F.C. collected and processed cryo-EM data. S.V. and G.C. performed model building. U.S. contributed to structure analysis. N.S.C. performed pre-steady-state kinetic analyses. S.V., N.S.C., and G.C. wrote the manuscript with input from all co-authors. G.C. conceived the project and together with N.S.C. coordinated and supervised experimental work and data analysis.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54618-y.

Correspondence and requests for materials should be addressed to Nicolas Sluis-Cremer or Guillermo Calero.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints