Charge Engineering of the Nucleic Acid Binding Cleft of a Thermostable HIV-1 Reverse Transcriptase Reveals Key Interactions and a Novel Mechanism of RNase H Inactivation

Javier Martínez del Río [1], Nerea López-Carrobles [1], Jesús I. Mendieta-Moreno [2], Óscar Herrera-Chacón [1], Adrián Sánchez-Ibáñez [1], Jesús Mendieta [3] and Luis Menéndez-Arias [1]*

*1 - **Centro de Biología Molecular** "**Severo Ochoa**" **(CSIC-UAM),** c/ Nicolás Cabrera 1, Campus de Cantoblanco-UAM, Madrid, Spain*

*2 - **Department of Theoretical Condensed Matter Physics,** Universidad Autónoma de Madrid, Spain*

*3 - **Department of Biotechnology,** Universidad Francisco de Vitoria, Pozuelo de Alarcón, Madrid, Spain*

***Correspondence to Luis Menéndez-Arias:** lmenendez@cbm.csic.es (L. Menéndez-Arias)*

https://doi.org/10.1016/j.jmb.2023.168219

Edited by Eric O. Freed

Abstract

Coupled with PCR, reverse transcriptases (RTs) have been widely used for RNA detection and gene expression analysis. Increased thermostability and nucleic acid binding affinity are desirable RT properties to improve yields and sensitivity of these applications. The effects of amino acid substitutions in the RT RNase H domain were tested in an engineered HIV-1 group O RT, containing mutations K358R/A359G/S360A and devoid of RNase H activity due to the presence of E478Q (O3MQ RT). Twenty mutant RTs with Lys or Arg at positions interacting with the template-primer (i.e., at positions 473–477, 499–502 and 505) were obtained and characterized. Most of them produced significant amounts of cDNA at 37, 50 and 65 °C, as determined in RT-PCR reactions. However, a big loss of activity was observed with mutants A477K/R, S499K/R, V502K/R and Y505K/R, particularly at 65 °C. Binding affinity experiments confirmed that residues 477, 502 and 505 were less tolerant to mutations. Amino acid substitutions Q500K and Q500R produced a slight increase of cDNA synthesis efficiency at 50 and 65 °C, without altering the K_D for model DNA/DNA and RNA/DNA heteroduplexes. Interestingly, molecular dynamics simulations predicted that those mutations inactivate the RNase H activity by altering the geometry of the catalytic site. Proof of this unexpected effect was obtained after introducing Q500K or Q500R in the wild-type HIV-1_{BH10} RT and mutant K358R/A359G/S360A RT. Our results reveal a novel mechanism of RNase H inactivation that preserves RT DNA binding and polymerization efficiency without substituting RNase H active site residues.

Introduction

Reverse transcriptases (RTs) synthesize complementary DNA (cDNA) from an RNA template. This activity plays a central role in the replication of hepadnaviruses, retroviruses, and retrotransposons.[1,2] RTs possess two distinct enzymatic functions: an RNA- and DNA-dependent DNA polymerase activity that synthesizes the proviral double-stranded DNA and a ribonuclease (RNase H) activity that degrades the RNA in RNA/DNA hybrids. RTs are widely used for research and biotechnological applications that require cDNA synthesis, including RT-quantitative real-time PCR (qPCR) and high-throughput RNA sequencing (RNA-seq).[3] Currently, recombinant variants of

murine leukemia virus (MLV) RT are the most popular enzymes in many research applications, although RTs of avian myeloblastosis virus (AMV), human immunodeficiency virus type 1 (HIV-1) and bacterial group II introns have been extensively characterized as well.[4–9] A main goal in generating RT mutants suitable for transcriptomics and gene expression analysis is to achieve higher optimal temperature and thermal stability, improved fidelity, higher processivity and increased catalytic efficiency.

Wild-type (WT) HIV-1 RTs and the *Geobacillus stearothermophilus* GsI-IIC RT (a group II intron RT sold commercially as TGIRT) show increased thermostability in comparison with WT AMV and MLV RTs.[5,10] However, several amino acid substitutions that eliminate RNase H activity in MLV and AMV RTs rendered enzymes with increased stability at high temperatures.[11–13] In addition, there are many examples of commercially available thermostable MLV RTs (e.g. ProtoScript®, SuperScript™, AccuScript™, AffinityScript™, etc.) (recently reviewed by Martín-Alonso *et al.*[3], and Oscorbin & Filipenko[14]). Many of them have been obtained through random mutagenesis and selection based on compartmentalized ribosome display and other techniques.[15–18]

Previous studies in our laboratory demonstrated that WT HIV-1 group O RT (ESP49 strain) had increased thermal stability and DNA/DNA template-primer binding affinity in comparison with prototypic WT HIV-1 group M/subtype B (BH10 strain) RT.[10,19] The nucleic acid binding affinity is important for a robust catalytic activity, particularly in the presence of limiting amounts of RNA. HIV-1 RTs are heterodimeric enzymes composed of subunits of 560 and 440 amino acids (designated as p66 and p51, respectively). The large subunit contains an RNase H domain (residues 441–560) responsible for the degradation of the viral RNA template during reverse transcription. The HIV-1 RT nucleic acid binding cleft accommodates 16–18 nucleotides between the active sites of the DNA polymerase and the RNase H.[20–22] Major interactions between the enzyme and the template-primer include the template grip (residues 73–83, 86–90, and 141–174), the primer grip (residues 227–235) and the helix clamp (residues 255–268 and 278–286) in the 66-kDa subunit, and the RNase H primer grip. This structure contains residues 358–361 in the connection subdomain of p66 and 395–396 in p51, in addition to RNase H domain residues 473–476, 501 and 505.[20,23] Specific mutations in the RNase H primer grip of a prototypic HIV-1 group M/subtype B RT (e.g. Y501A or N474A/Q475A) have a deleterious effect on the initiation of DNA synthesis and RNase H cleavage specificity *in vitro* and *ex vivo*.[24,25]

Theoretical studies based on molecular modeling and free energy calculations showed that in HIV-1_{BH10} RT, the stronger interactions with the template-primer occur at Arg284, Arg358, Gly359, Ala360 and His361, within the 66-kDa subunit of the enzyme.[19] Moreover, the introduction of Arg358-Gly359-Ala360 at the equivalent positions of HIV-1_{ESP49} RT (i.e., Lys358-Ala359-Ser360) produced a significant increase in DNA/DNA binding affinity. The resulting RT (mutant K358R/A359G/S360A, abbreviated as O3M) also showed enhanced cDNA synthesis efficiency above 68 °C.[19] Unlike in the case of RNase H-inactivating mutants D524A, D524G, E562Q or D583N of MLV RT,[11–13] the addition of the RNase H active site mutation E478Q to the O3M RT (i.e., mutant K358R/A359G/S360A/E478Q, abbreviated as O3MQ RT) had little effect on nucleic acid binding affinity, while both enzymes showed remarkable cDNA synthesis efficiency at 68 °C, in the presence of high amounts of template RNA.[19]

Based on previous studies showing that substituting alanine at positions 473–476 and 501 of HIV-1_{BH10} RT had minimal consequences for DNA synthesis on duplex and hybrid DNA and RNA substrates,[25] we have performed an extensive mutagenesis study by substituting Lys or Arg at residues forming part of the RNase H primer grip in the O3MQ RT. Introducing positively charged residues at selected positions unveils the fundamental role of different amino acids in the primer grip, while assessing the contribution to DNA synthesis and template-primer binding affinity of the RNase H domain. Interestingly, our study unveils a novel mechanism of RNase H inactivation that preserves nucleic acid binding and DNA polymerization efficiency, without making any replacements in the conserved active-site DEDD motif found in many nucleases, including RNase H.[21,22,26]

Results

Molecular modeling and identification of template-primer binding residues in the RNase H domain of HIV-1 RTs

Currently, there are nearly one hundred crystal structures of HIV-1 RTs complexed with nucleic acids (with or without dNTPs) and deposited in the Protein Data Bank. All these enzymes correspond to HIV-1 group M/subtype B strains (HXB2, BH10, IIIB, etc.). Available structures include binary complexes of HIV-1 RT and double-stranded DNA in pre- and post-translocation states (e.g. 1N5Y, 1N6Q and 5XN0), HIV-1 RT in complex with an RNA/DNA polypurine tract (1HYS) or with an RNA/DNA hybrid with or without nevirapine, showing structures compatible with RNA degradation (e.g. 4B3O, 4B3P, 4B3Q, 4PUO, 4PWD and 4Q0B) or ternary complexes like 1RTD containing HIV-1_{HXB2} RT bound to double-stranded DNA and dTTP. The free-energy decomposition matrix obtained with the crystal structure of the ternary complex (1RTD) identified residues 473–476, 478, 499–501 and 505 as part

of the RNase H primer grip that interacts with the double-stranded DNA in the RNase H domain of the RT.[19,27] The selected residues were also found at interacting sites with the RNA/DNA in the crystal structure of HIV-1 RT complexed with a polypurine tract.[20] As shown in Figure 1, residues interacting with the primer include Thr473, Gln475, Lys476, Tyr501 and Ile505, while Asn474 and Gln500 interact with the phosphate backbone of the RNA template.

In contrast to HIV-1 group M/subtype B strains, there are no crystal structures available for HIV-1 group O RTs. The molecular model of a ternary complex of HIV-1_{ESP49} RT based on the 1RTD structure was built and its reliability to study RT-nucleic acid interactions was proved after the successful characterization of mutants containing substitutions K358R/A359G/S360A (i.e., O3M RT).[19] In comparison with the structure of the ternary complex of HIV-1_{HXB2} RT/double-stranded DNA/dTTP, the analysis of free-energies of interaction between the HIV-1 group O enzyme and the nucleic acid showed stronger interactions with Ser499 and additional contacts with Val502, not found in the HIV-1_{HXB2} RT complex.

Identified interacting residues were systematically replaced by Lys and Arg to evaluate whether positively charged amino acids at the RNase H primer grip would stabilize template-primer binding and improve RT's catalytic efficiency. Target sequences were those extending positions 473–477 and 499–502, plus residue 505, in the background of the K358R/A359G/S360A/E478Q RT (abbreviated as O3MQ RT). In the presence of limited amounts of RNA, this enzyme showed increased cDNA synthesis efficiency in comparison with mutant K358R/A359G/S360A RT (O3M RT) and the WT HIV-1_{BH10} RT (BH10 RT), as determined in RT-PCR reactions carried out at 65 °C (Figure 2). These results were consistent with previously published cDNA yield estimates obtained in reverse transcription reactions carried out at 68 °C, as determined in qPCR assays.[19]

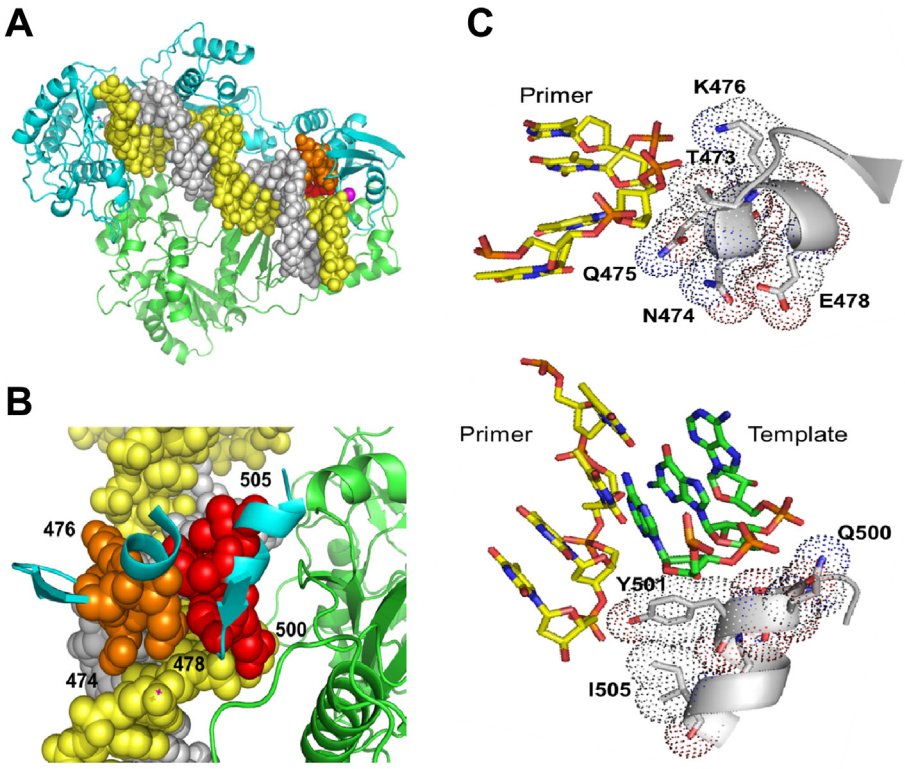

Figure 1. HIV-1 RT structure and nucleic acid interaction sites in the RNase H domain of the enzyme. (A) Crystal structure of the HIV-1 RT bound to a DNA/DNA template primer (PDB file 1RTD) showing the location of interacting residues in the vicinity of the RNase H primer grip (residues 473–478 represented with orange spheres, and 499–502 and 505 with red spheres). The template is shown in yellow and the primer in light grey. Cyan and green cartoons are used to represent the backbones of the 66- and 51-kDa subunits of the enzyme, respectively. **(B)** Close-up view showing the location of interacting residues in the RNase H domain of the RT. The cartoon representing the p66 structure has been removed for a better view of the nucleic acid binding cleft. **(C)** Views of key interactions at positions 473–478 (up) and 499–505 (down) taken from PDB file 1HYS corresponding to the HIV-1 RT bound to an RNA/DNA template-primer.

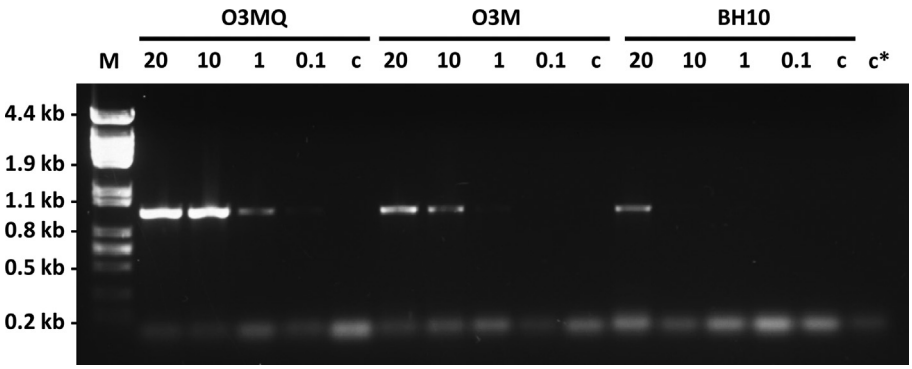

Figure 2. Efficiency of cDNA synthesis and PCR amplification in reactions catalyzed by WT BH10 and HIV-1 group O RT mutants O3M and O3MQ with different amounts of RNA. All cDNA synthesis reactions were carried out for 60 min at 65 °C with different amounts of mouse liver total RNA. Reactions were stopped when mixtures were heated at 90 °C for 10 min. The product of the cDNA synthesis reaction was amplified with Taq DNA polymerase. Amplification of actin DNA fragments of 0.95 kb are shown. Lane M shows molecular mass size markers (HindIII digest of phage Φ29 DNA). Control reactions carried out without RNA or without cDNA are shown in lanes c and c*, respectively. Numbers above the corresponding lanes indicate the amount of total RNA used in the reaction (in ng). Results are representative of at least three independent experiments.

Lysine and arginine scanning mutagenesis of the RNase H primer grip of O3MQ RT and effects on cDNA synthesis efficiency and template-primer binding affinity

Nineteen variants of the O3MQ RT, containing Lys or Arg at positions defined as interacting sites of the RNase H primer grip and the template-primer were purified and tested for cDNA synthesis efficiency. In addition to the 19 mutants containing single amino acid substitutions, we also characterized an active RT containing an insertion of 9 amino acids where K476 was replaced by the sequence GSINGDHQSK, and obtained as an unexpected variant while preparing mutant K476R. All mutant enzymes were purified as stable heterodimers. None of the residues involved in the RNase H primer grip participates in interactions between p66 and p51.[28] However, RTs containing A477K or A477R were hardly active in titration experiments and contained a small proportion of p51 relative to p66, in agreement with a previous report suggesting that substitutions at this position could have an impact on heterodimer formation due to abnormal cleavage of p66.[29]

RT efficiencies were determined using a two-step reverse transcription PCR (RT-PCR) assay that included an initial cDNA synthesis reaction at 37 °C in the presence of different amounts of mouse liver total RNA. An initial screening carried out with 200 ng of total RNA showed positive results for 85% of the tested enzymes (Figure 3). In contrast, there was no amplification in assays carried out with V502K and V502R, while yields with A477K were very low. Further experiments carried out with reduced quantities of RNA confirmed the lack of activity of these mutants. In addition, A477R, I505K, I505R and in a lesser extent S499K and S499R also showed largely diminished activity in comparison with the parental O3MQ RT. No amplification products were observed with those mutants in the assays carried out with 50 and 20 ng of total RNA. Positively charged residues were relatively well tolerated at six positions (473, 474, 475, 476, 500 and 501), although in all cases we observed a reduction in cDNA synthesis efficiency that was more pronounced for mutants N474R, Y501K and Y501R.

The lack of activity observed for mutants V502K and V502R was consistent with their weak DNA/DNA binding affinity (Table 1), while A477K and A477R mutants were unable to elongate DNA primers in single-nucleotide incorporation assays. Although the nucleic acid binding affinity of the tested mutants was broadly consistent with the RT-PCR data, in most cases, DNA/DNA dissociation equilibrium constants (K_D) were not largely affected by the introduced mutations. Most of the tested mutants showed less than 2-fold variations in the K_D, as compared with the reference RT (e.g. T473K, T473R, N474K, N474R, Q475K, S499K, S499R, Q500K and Q500R). Although most of these enzymes were able to synthesize cDNA when the amount of total RNA was as low as 20 ng, mutants S499K and S499R showed very good binding affinity but very poor RT-PCR efficiency at 50 ng. Interestingly, those amino acid substitutions had a more significant effect in the K_D when the template-primer was an RNA/DNA heteroduplex (Table 1).

Temperature effect of amino acid substitutions in cDNA synthesis efficiencies determined in RT-PCR assays

RT-PCR assays carried out at 50 °C and 65 °C highlighted the deleterious effects of mutations at

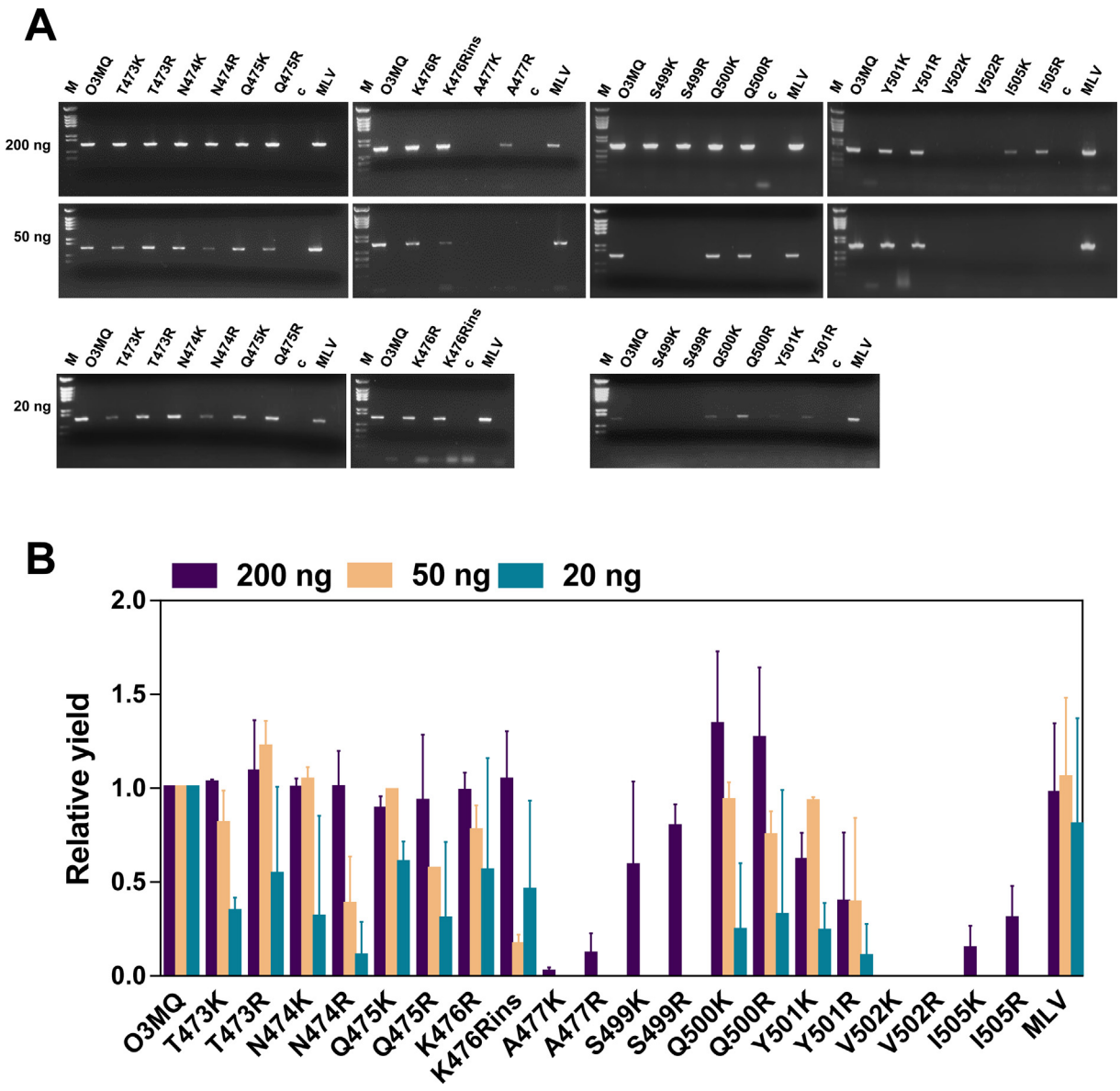

Figure 3. Efficiency of cDNA synthesis at 37 °C in reactions catalyzed by O3MQ RT variants with amino acid substitutions in the RNase H domain. All cDNA synthesis reactions were carried out for 60 min at 37 °C with different amounts of mouse liver total RNA. Reactions were stopped when mixtures were heated at 90 °C for 10 min, and products were amplified with Taq DNA polymerase, using actin primers ACT1 and ACT3. **(A)** Agarose gels showing the amplification products. Lanes M and c indicate molecular mass size markers (HindIII digest of phage Φ29 DNA) and a control reaction mixture (conducted without cDNA). Murine leukemia virus RT was used in control reactions carried out with 200 units of the enzyme. Results are representative of at least three independent experiments. **(B)** Bar diagrams indicate the relative yields of the amplification reactions carried out with different amounts of RNA relative to those obtained with the O3MQ RT. Reported values are averages ± standard deviations from at least three independent experiments.

positions 477, 499, 502 and 505, while the catalytic efficiency of mutants T473K, T473R, Q475R, Y501K and Y501R was severely impaired, particularly at 65 °C (Figure 4). The loss of activity of those mutants was further confirmed by qPCR, after measuring actin cDNA yields obtained in reverse transcription reactions carried out at 65 °C, using 200 ng of mouse liver total RNA (Supplementary Figure S2). The introduction of positively charged residues at position 500 (i.e., mutants Q500K or Q500R in O3MQ RT) had a relatively minor effect on the amplification yields as compared with the reference enzyme O3MQ RT. In agreement with these results, we did not find significant differences among the three enzymes when RNA/DNA binding affinities were

Table 1 Relative dissociation equilibrium constants (K_D) for O3MQ RT variants and DNA/DNA and RNA/DNA template-primers, obtained at 37 °C.

RT	K_D (mutant) / K_D (O3MQ RT)	
	DNA/DNA	RNA/DNA
O3MQ	1	1
T473K	0.93 ± 0.57	1.59 ± 0.68
T473R	1.88 ± 0.48	ND
N474K	1.29 ± 0.37	1.95 ± 1.1
N474R	1.06 ± 0.24	2.28 ± 1.1
Q475K	1.24 ± 0.31	1.04 ± 0.39
Q475R	2.85 ± 0.71	ND
K476R	3.48 ± 1.07	ND
K476R-ins	4.17 ± 1.20	ND
A477K	No activity	ND
A477R	No activity	ND
S499K	1.16 ± 0.57	2.66 ± 0.78
S499R	1.14 ± 0.41	2.48 ± 0.69
Q500K	0.80 ± 0.38	0.98 ± 0.41
Q500R	1.08 ± 0.37	0.91 ± 0.32
Y501K	2.32 ± 0.84	3.32 ± 0.81
Y501R	1.46 ± 0.6	1.77 ± 0.72
V502K	>3.5	ND
V502R	>3.5	ND
I505K	2.48 ± 1.12	ND
I505R	2.41 ± 0.92	ND

The K_D values were determined at 37 °C using template-primer D38/25PGA (for DNA/DNA) and D38rna/25PGA (for RNA/DNA). Reported values are averages ± standard deviations from at least three independent experiments. Reference K_D values for O3MQ and DNA/DNA and RNA/DNA duplexes were 1.69 ± 0.38 and 1.02 ± 0.41 nM, respectively. ND, not determined.

determined at 50 °C. The corresponding K_D values for O3MQ, and mutants Q500K and Q500R were 2.99 ± 0.56 nM, 3.73 ± 0.23 nM and 4.22 ± 1.43 nM, respectively (Supplementary Figure S3).

Despite the modest increase in cDNA synthesis efficiency observed with mutants Q500K and Q500R of O3MQ RT at 65 °C (in the range of 20–200 ng of RNA), additional experiments carried out at 80 °C using the GADPH gene products as targets for amplification revealed that neither Q500K nor Q500R increased the sensitivity of the RT-PCR tests (Figure 5(A)). Besides, the parental O3MQ RT and both mutants showed significant activity at high temperatures. In RT-PCR assays carried out with 10 ng of total RNA from HeLa cells, a GADPH product of 0.3 kb was detected in all cases even when the cDNA reaction was carried out at 85 °C (Figure 5(B)), underlining the high thermostability of these enzymes. In these conditions, the results of qPCR assays carried out with cDNA obtained at 65 and 80 °C did not reveal significant differences in reverse transcription efficiencies among the tested enzymes, although cDNA synthesis was somewhat less efficient in the case of mutant Q500K (Supplementary Figure S4).

Effects of Q500K and Q500R in the interaction between the RT and the template-primer

Molecular dynamics studies carried out with O3MQ RT and mutants Q500K and Q500R at different temperatures showed relatively small changes in the interaction energies with protein and nucleic acid, although some remarkable differences were observed at position 500 (Figure 6(A)). These differences were also detected with RNA/DNA heteroduplexes (Supplementary Figure S5) and the analysis of molecular dynamics trajectories revealed that all complexes were stable with major interaction sites at positions 448, 473–478 and around position 500 (Supplementary Figure S6), as inferred from crystal structures (Figure 1) and molecular modeling studies.

The substitution of Arg for Gln500 produced a large conformational change that affected its side chain and neighboring interactions. In the parental O3MQ RT the side chain of Gln500 pointed towards the template. However, in the mutant containing Q500R, the Arg residue retained template binding, but changed its conformation in such a way that its side-chain interacts with Ser499 and provokes a movement of catalytic residues Asp498 and Gln478 that pushes them away from the catalytic Mg^{2+} (Figure 6(B)). In the O3MQ Q500R RT, the interatomic distances between the Cδ (Gln478) and Cγ (Asp498) and their coordinating Mg^{2+} are 3.95 and 3.03 Å, respectively (Supplementary Figure S7). However, in the parental enzyme O3MQ RT, the distances are smaller, with values of 3.63 Å for the distance Mg^{2+} - Cδ (Gln478), and 2.34 Å for Mg^{2+} - Cγ (Asp498). Reference values obtained from molecular dynamics simulations for the catalytically competent WT HIV-1 RT are 3.0 and 2.43 Å, for the distances between Cδ (Glu478) and Cγ (Asp498) and their coordinating cations.

RNase H inactivation by Q500K and Q500R

The analysis of RT-nucleic acid interactions in HIV-1 RT binary and ternary complexes showed that the side chain of Gln500 contacts the template strand in RNA/DNA hybrids (Figure 1). Moreover, the molecular dynamics simulations with O3MQ RTs containing Lys or Arg at position 500 reveal a conformational change of the side chain that leads to the loss of its interaction with the template strand, pointing away from the double-stranded nucleic acid heteroduplex (Figure 6(B)). These observations suggested that Q500K and Q500R could have an impact on the RNase H activity of the RT, although O3MQ RT is devoid of endonuclease activity due to the presence of the RNase H-inactivating mutation E478Q. Consequently, we tested the effect of substitutions Q500K and Q500R in O3M and BH10 RTs.

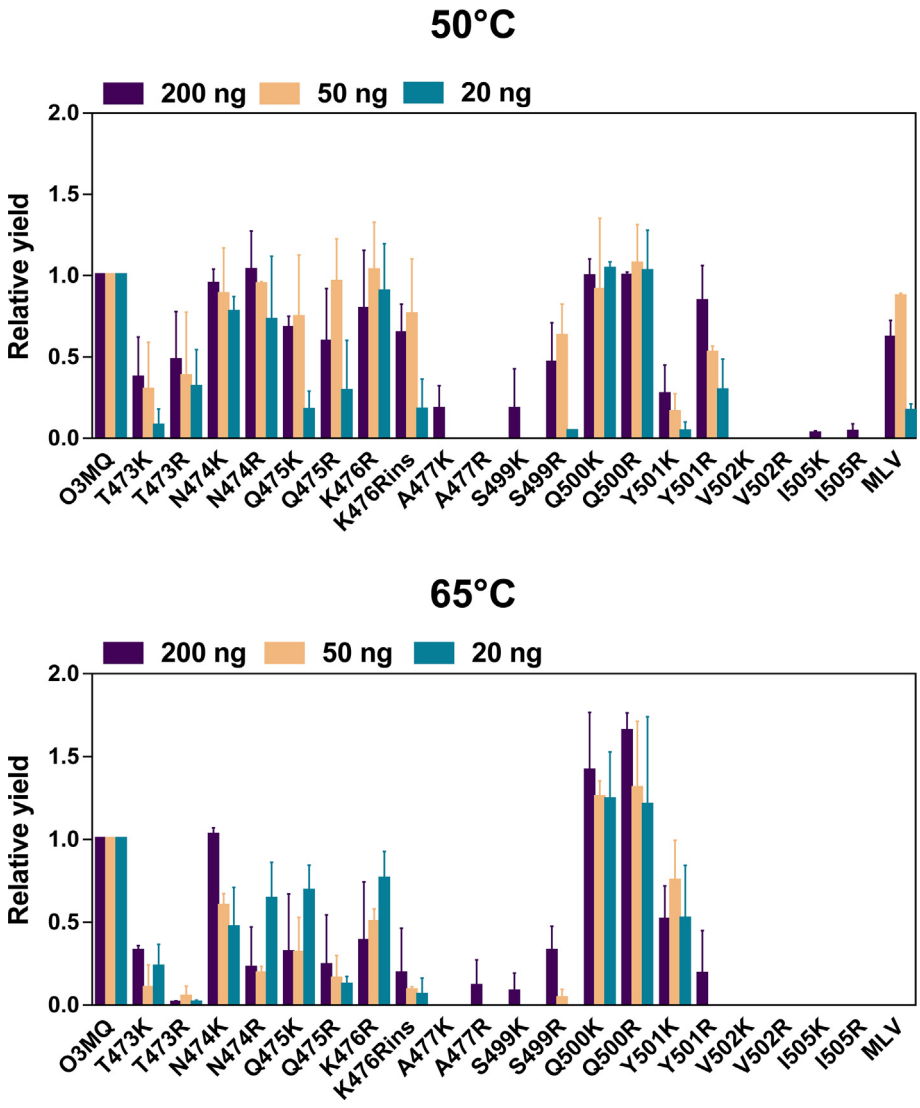

Figure 4. Efficiency of cDNA synthesis at 50 °C and 65 °C in reactions catalyzed by O3MQ RT variants with amino acid substitutions in the RNase H domain. The cDNA synthesis reactions were carried out for 60 min at the indicated temperatures with different amounts of mouse liver total RNA. Products amplified with Taq DNA polymerase, using actin primers ACT1 and ACT3, as indicated in the Materials and Methods section. Bar diagrams indicate the relative yields of the amplification reactions carried out with different amounts of RNA relative to those obtained with the O3MQ RT. Reported values are averages ± standard deviations from at least three independent experiments. Representative gels used to calculate the relative yields are given in the Supplementary Figure S1.

As shown in Figure 7(A), both amino acid changes produced a large reduction in the RT's RNase H activity, particularly in the context of BH10. The most deleterious effects were observed with Q500K. Interestingly, the loss of RNase H catalytic activity produced by Q500K and Q500R in the context of O3M RT caused a significant increase in RNA sensitivity in cDNA synthesis reactions carried out with total RNA from HeLa cells. Yields in the range of 0.1–10 ng of RNA were similar for O3M RT mutants Q500K and Q500R than for the O3MQ RT, but much higher than those obtained with O3M that retains WT RNase H activity (Figure 7(B)).

As described above for O3MQ RT mutants, the effects of amino acid substitutions Q500K and Q500R on RNA/DNA and DNA/DNA binding affinity were not significant when introduced in the O3M RT (Table 2). However, we observed a modest but significant loss of affinity when the substitution Q500K was introduced in BH10 RT, particularly with DNA/DNA hybrids. This finding is consistent with results of RT-PCR assays that showed the higher efficiency of Q500R in the detection of minute amounts of RNA as compared with the WT enzyme (Figure 7(B)). However, the mutant Q500K was less efficient than the WT enzyme in detecting the GAPDH gene product

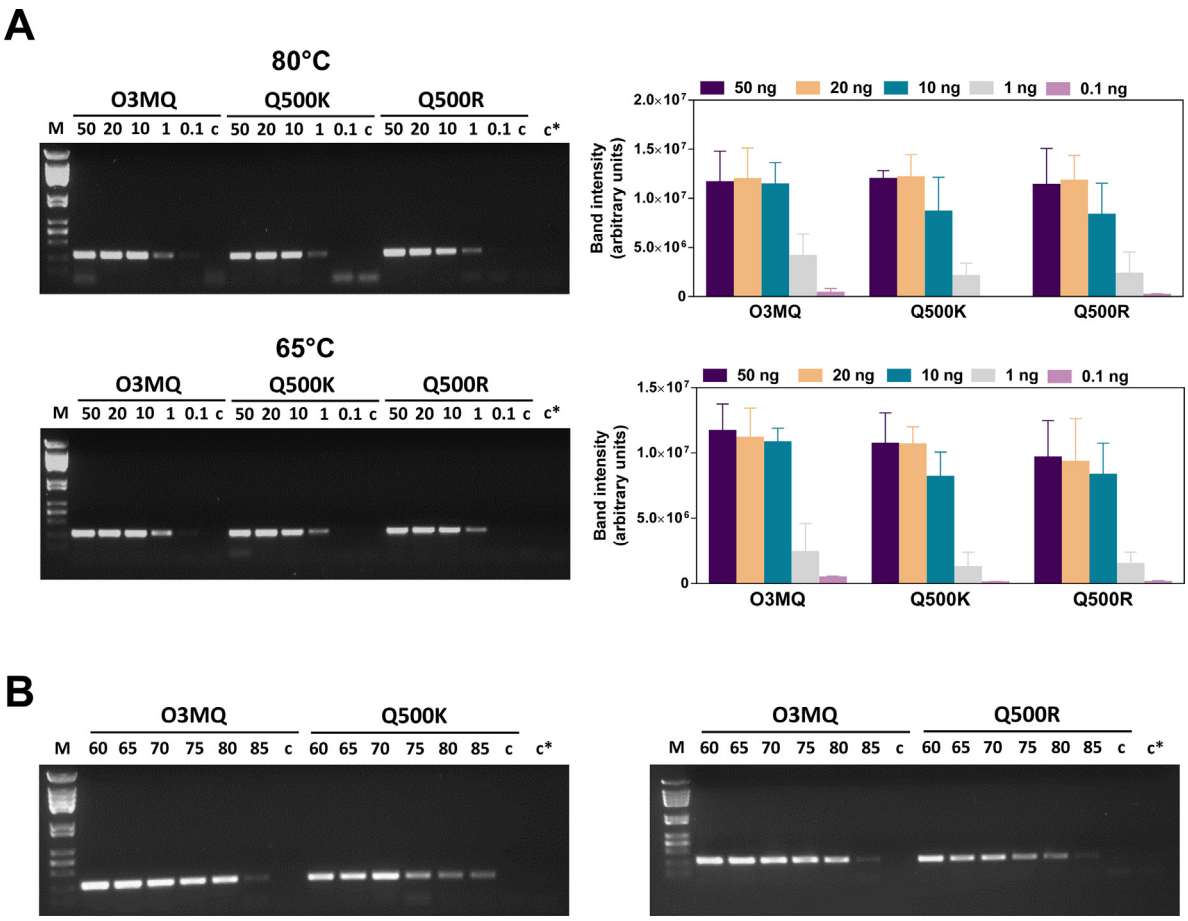

Figure 5. Temperature effects on the reverse transcription efficiency of mutants Q500K and Q500R in the context of O3MQ RT. The cDNA synthesis reactions were carried out for 60 min at the indicated temperatures with different amounts of total RNA from HeLa cells. Products were amplified with Taq DNA polymerase, using GAPDH primers G1 and G2, as indicated in the Materials and Methods section. **(A)** The gel shows the amplification results obtained with different amounts of RNA in the range of 0.1 – 50 ng, when cDNA synthesis reactions were carried out at 65 °C and 80 °C. Lane M shows molecular mass size markers (HindIII digest of phage Φ29 DNA). Control reactions carried out without RNA or without cDNA are shown in lanes c and c*, respectively. Bar diagrams on the right indicate yields of the RT-PCR reactions carried out at 65 °C and 80 °C with different amounts of RNA. **(B)** Effect of the temperature (range 60–85 °C) in the RT-PCR yields obtained in reactions catalyzed by O3MQ RT and mutants Q500K and Q500R, in the presence of 10 ng of HeLa cells total RNA. Lane M shows molecular mass size markers (HindIII digest of phage Φ29 DNA). Control reactions carried out without RNA or without cDNA are shown in lanes c and c*, respectively. Results are representative of at least three independent experiments.

when the cDNA synthesis reaction was carried out at 65 °C in the presence of less than 10 ng of total RNA from HeLa cells.

Molecular dynamics and mechanism of RNase H inactivation

Experimental results were consistent with predictions rendered by molecular dynamics simulations carried out with O3M and O3M Q500R RTs. As shown in Figure 8, the observed conformational changes for these two enzymes were similar to those obtained with O3MQ and O3MQ Q500R RTs (Figure 6(B)). According to the molecular dynamics predictions, the observed conformational change affecting the side chain at position 500 produced a shift that affected the RT protein backbone, while moving the side chain of Glu478 away from the catalytic site. In the O3M Q500R RT, the interatomic distance between Cγ of Asp498 and its coordinating Mg^{2+} was 3.07 Å, similar to that observed in the mutant O3MQ Q500R (Supplementary Figure S7). The presence of Glu or Gln at position 478 had no significant effect on the positioning of Asp498. However, the distance between Cδ of Glu478 and the coordinating Mg^{2+} increased up to 5.56 Å, thereby facilitating the introduction of a water molecule between the side chain of Glu478 and the catalytic Mg^{2+}, and abrogating the catalytic activity of the RNase H.

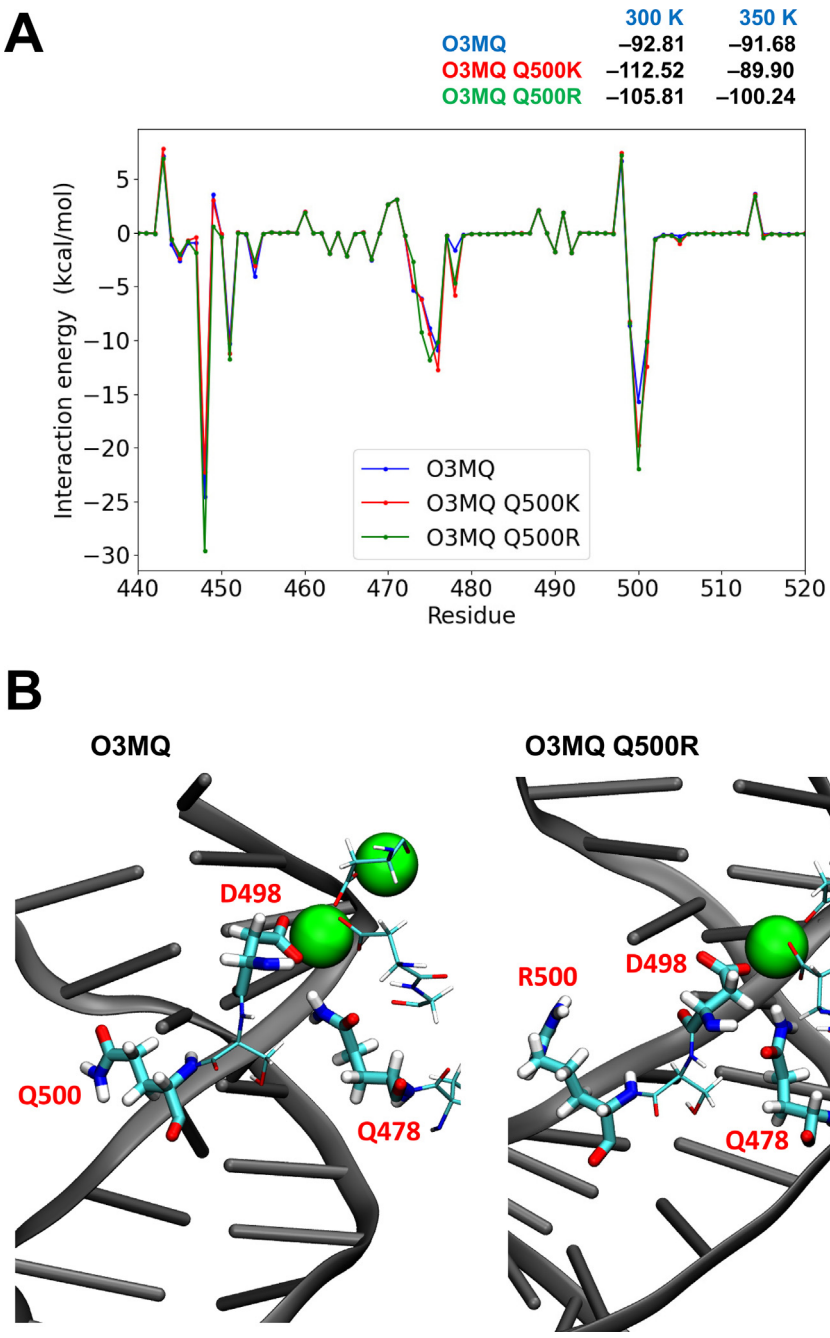

Figure 6. Molecular dynamics of O3MQ RT and effects of substituting Lys or Arg for Gln500. (A) Interaction energies for the three RTs and double-stranded DNA, at the RNase H domain (residues 441–560 of p66). Values shown in the table correspond to interaction energies calculated for residues 380–520 of the RT and the template-primer. **(B)** Structures showing the side-chain conformations of key RT residues (stick representations) at positions 478, 498 and 500 in relation to the nucleic acid, for the O3MQ RT, and the mutant containing Q500R. Green spheres show the Mg^{2+} ions located at the RNase H catalytic site. Shown simulations were obtained with DNA/DNA hybrids at 300 K.

Discussion

Recombinant HIV-1 group O RTs O3M and O3MQ show remarkable catalytic efficiency at temperatures above 75 °C, as determined in RT-PCR assays.[19] Free energy calculations based on molecular modeling of binary complexes of HIV-1 RTs bound to nucleic acid showed the key role of template-primer binding interactions in conferring thermal stability to those enzymes. The introduction of the triad K358R/A359G/S360A enhanced the cDNA synthesis efficiency of WT HIV-1 group O

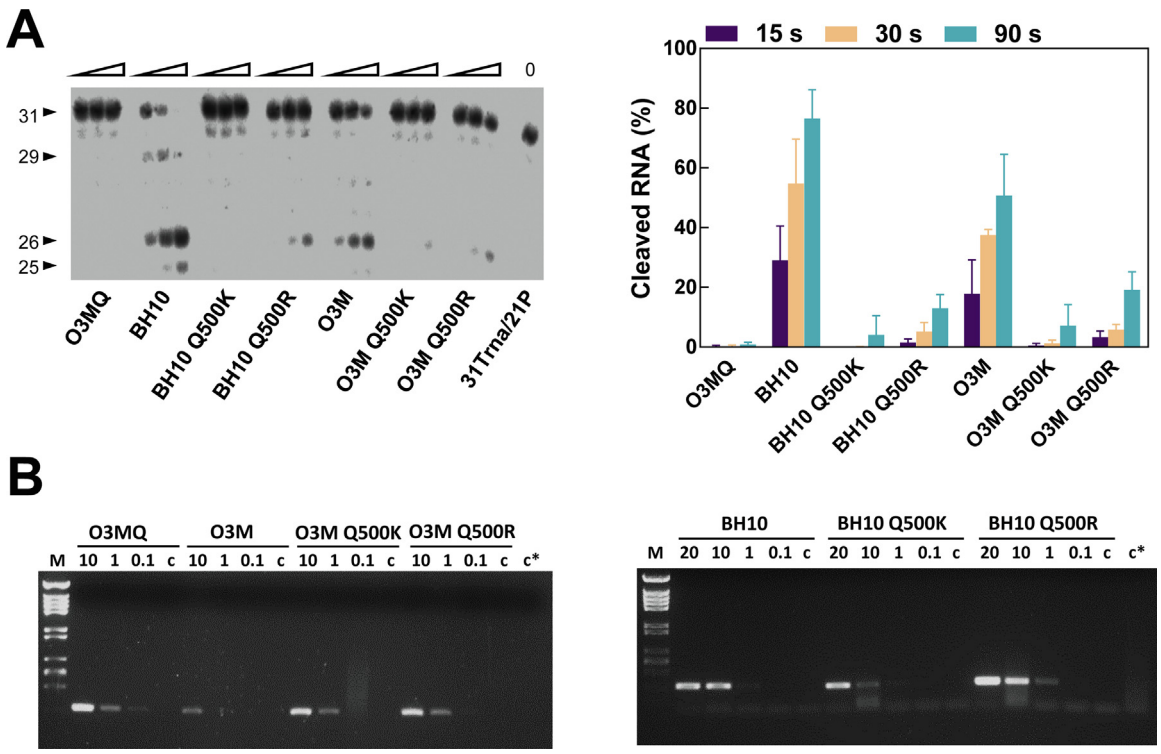

Figure 7. Effects of Q500K and Q500R in different sequence backgrounds on RNase H activity and cDNA synthesis efficiency. (A) Cleavage of a [^{32}P]RNA/DNA substrate (25 nM) was carried out at 37 °C in the presence of the corresponding RT at 25–50 nM (active enzyme concentration). The gel shows for each enzyme the uncleaved substrate and aliquots taken after 15, 30 and 90 s of incubation in reaction buffer. The bar diagram shows relative amounts of cleaved RNA template in RNase H activity assays carried out with O3MQ, O3M and BH10 RTs and mutants Q500K and Q500R for 15, 30 and 90 s. Reported values are averages ± standard deviations from at least three independent experiments. **(B)** RT-PCR amplification reactions with O3MQ, O3M and BH10 RTs and mutants Q500K and Q500R in the context of the O3M and BH10 enzymes. The cDNA synthesis reactions were carried out for 60 min at 65 °C with 0.1, 1, 10 and 20 ng of total RNA from HeLa cells. Products were amplified with Taq DNA polymerase, using GAPDH primers G1 and G2, as indicated in the Materials and Methods section. Lane M shows molecular mass size markers (HindIII digest of phage Φ29 DNA). Control reactions carried out without RNA or without cDNA are shown in lanes c and c*, respectively. Results are representative of at least three independent experiments.

Table 2 Effects of Q500K and Q500R in the dissociation equilibrium constants (K_D) of WT BH10 and O3M RTs and double-stranded DNA/DNA and RNA/DNA complexes.

substitution	K_D (mutant) / K_D (reference RT)			
	WT BH10 RT		O3M RT	
	DNA/DNA	RNA/DNA	DNA/DNA	RNA/DNA
Q500K	3.03 ± 1.14	1.59 ± 0.36	1.30 ± 0.62	1.54 ± 0.51
Q500R	0.53 ± 0.22	0.94 ± 0.28	1.41 ± 0.89	1.17 ± 0.59

The K_D values were determined at 37 °C using template-primer D38/25PGA (for DNA/DNA) and D38rna/25PGA (for RNA/DNA). Reported values are averages ± standard deviations from at least three independent experiments. In these experiments, the reference K_D values for WT BH10 RT and DNA/DNA and RNA/DNA duplexes were 2.75 ± 0.61 and 1.69 ± 0.16 nM, respectively. Reference K_D values for O3M RT and DNA/DNA and RNA/DNA duplexes were 1.35 ± 0.58 and 1.02 ± 0.24 nM, respectively.

RT, while increasing its nucleotide binding affinity.[19] In binary complexes of HIV-1_{HXB2} RT and RNA/DNA hybrids, Gly359 and Ala360 together with several amino acids of the RNase H domain (e.g. residues 473–476, 478, 499–501 and 505) participate in a network of contacts, involving the sugar-phosphate backbone of the DNA primer, located 4–9 nucleotides away from the RNase H catalytic site.[20] These interactions are important for the trajectory and proper orientation of the template-primer within the nucleic acid binding cleft.

Our mutagenesis study shows restrictions in the variability of the RNase H primer grip. Unexpectedly, in most cases, the introduction of

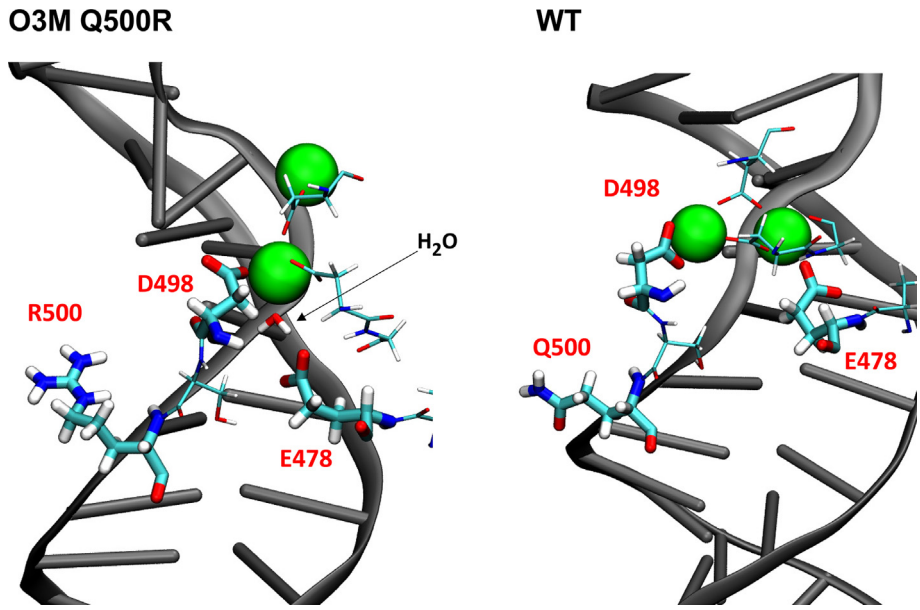

Figure 8. Molecular dynamics of O3M RT and effects of substituting Arg for Gln500. Detailed view of the structural models obtained for the O3M RT bearing Arg at position 500 (left) compared with the wild-type enzyme. Changes in the side-chain conformation of residues at positions 498–500 in relation to the template nucleic acid destabilize the catalytic site moving away from the Mg^{2+} cations the side chains of Asp498 and Glu478, facilitating the insertion of a water molecule between this residue and the coordinating Mg^{2+}. Shown simulations were obtained with DNA/DNA hybrids at 300 K.

positive charges at interaction sites between the RNase H domain and the template-primer had little effect in nucleic binding affinity, while being detrimental for cDNA synthesis efficiency at high temperatures (e.g. 65 °C, Figure 4). The most deleterious effects were observed with mutants A477K/R, V502K/R and I505K/R and can be attributed to the disruption of α-helices (A' for Ala477, and B' for Val502 and Ile505), resulting in total or partial destabilization of the RNase H domain. We also observed large reductions in reverse transcription efficiency (particularly at high temperatures) with mutants T473K/R, S499K/R, and in a lesser extent with Y501K/R. Tyr501 locates at α-helix B' and points towards the phosphate group in the DNA primer, five nucleotides away from the RNase H catalytic site.[20] Interestingly, Thr473 and Tyr501 are also key conserved targets of antiviral intervention. Vinylogous ureas such as NSC727447,[30] β-thujaplicinol,[31] diketo acid derivatives derived from 6-[1-(4-fluorophenyl)methyl-1*H*-pyrrol-2-yl)]-2,4-dioxo-5-hexenoic acid ethyl ester (RDS1643)[32] and the thiazolidine-2,4-dione AA[33] are examples of allosteric and active site RNase H inhibitors that interfere with nucleic acid binding at the RNase H primer grip.

Despite being non-conservative substitutions, mutants Y501K and Y501R retained significant reverse transcription activity at 37 °C and rendered RT-PCR products in reactions carried out at 50 and 65 °C. These results are consistent with early studies showing that in the HIV-1 group M/subtype B RT, amino acid substitutions Y501F, Y501W and Y501R had a minor effect on RNA-dependent DNA polymerization and RNase H activity, while other residues at this position (e.g. Gly, Leu, Ser, Glu, Ala or His) were detrimental for both activities.[34] These results were further confirmed by alanine scanning mutagenesis of the RNase H primer grip. In enzymatic assays the mutant Y501A RT showed 4-fold reduced double-stranded binding affinity, altered polypurine tract (PPT) processing activity and about 5-fold reduced viral titer in single round of replication assays.[24,25,35]

Other alanine substitutions at the RNase H primer grip such as T473A abolished viral replication, while Q475A and the double-mutant N474A/Q475A reduced the titer 5- to 10-fold.[24] In contrast, the viral replication capacity of mutants N474A and K476A was reduced less than two-fold relative to the WT.[24,36] In the case of I505A, we found conflicting results in the literature. While Delviks-Frankenberry *et al.*[36] reported a > 10-fold decrease in viral replication capacity, associated to a 10-fold reduction in RT activity, Julias *et al.*[24] found that the mutation did not affect the viral titer in single round of replication assays. Although both results are plausible from a structural point of view, the largely detrimental effects of I505K and I505R in the sequence background of O3MQ RT could be explained by the potential destabilization of α-helix B'.

In our study, the substitutions of Lys or Arg for Gln500 in the O3MQ RT had relatively minor effects on reverse transcription efficiency over a

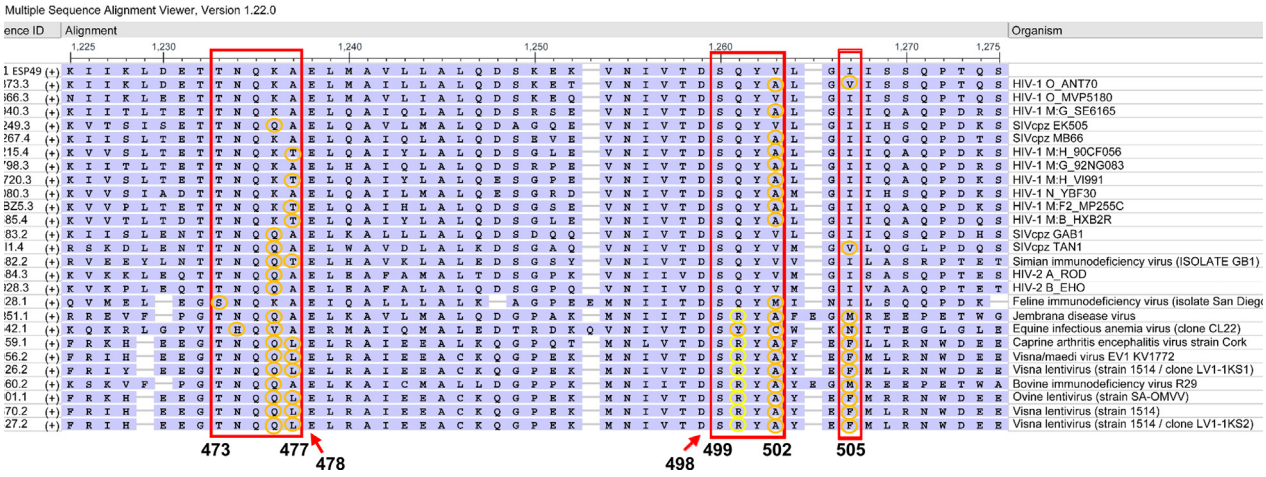

Figure 9. Representative amino acid sequence alignment of the RNase H primer grip (residues 465–513) of lentiviral RTs. The program BLASTP (NCBI) was run with the amino acid sequence of the HIV-1_{ESP49} RT RNase H domain (top sequence) *versus* the UniProt/Swiss-Prot Database, using *Retroviridae* (taxid: 11632) as the target organism. The complete alignment is provided in the Supplementary Figure S8. Representative sequences are shown.

wide range of temperatures. Furthermore, at temperatures above 65 °C, differences in RNA sensitivity between O3MQ RT and mutants Q500K and Q500R were not significant. The molecular dynamics simulations obtained with those mutants showed little differences in template-primer binding in comparison with the parental enzyme, O3MQ RT, while free energies of interaction between the RT and the RNA/DNA hybrid remain unchanged along the nucleic acid binding cleft. However, those simulations predicted a change in the conformation of the basic side chain at position 500 relative to the parental enzyme. While Gln500 interacts with the RNA template in the O3MQ RT (as observed in crystal structures of HIV-1_{HXB2} RT), the side chains of Lys500 or Arg500 in the mutant enzymes point away from the RNA, suggesting that those amino acids could have a negative effect in RNase H activity.

Our experiments carried out with RTs having a WT RNase H activity (i.e., HIV-1_{BH10} RT and mutant O3M RT) revealed a loss of endonuclease activity to the presence of Q500K or Q500R in both enzymes. Interestingly, all three mutants of O3M RT (Q500K, Q500R and E478Q) showed increased reverse transcription efficiency in comparison with the parental enzyme.

The molecular dynamics simulations revealed alterations in the geometry of the RNase H catalytic site due to the presence of Q500K or Q500R. The side-chain of Glu478 is shifted away from the closest catalytic Mg^{2+}, due to altered packaging with Ser499. These results are consistent with current knowledge on the catalytic mechanism of the RNase H. During the nucleolytic reaction, the cleaved phosphodiester bond containing an O3´ leaving oxygen remains tightly bound to the Mg^{2+}, and its protonation is critical to stabilize the cleavage product. The highly conserved Glu478 residue acts as a proton donor for the leaving product.[37]

Interestingly, Gln500 is highly conserved among HIV and simian immunodeficiency virus (SIV) strains, although sequence alignments carried out with viral proteins deposited in the UniProt/Swiss-Prot database revealed the natural occurrence of Arg at the equivalent position in the RTs of cattle-infecting lentiviruses such as Jembrana disease virus, caprine arthritis encephalitis virus, bovine immunodeficiency virus, visna/maedi virus and visna and ovine lentiviruses (Figure 9). However, in all cases there are additional mutations in the primer grip, including the loss of a basic residue at position 476 (which is Gln in the Arg500-containing RTs).

Taken together, our results underline limitations in further optimization of a catalytically efficient RT at high temperatures. The straightforward introduction of positive charges in the nucleotide binding cleft does not seem to affect template-primer binding affinity and cDNA synthesis efficiency. Major determinants of HIV-1 RT binding to duplex and hybrid nucleic substrates concentrate in the template and primer grips of the polymerase domain, including β-strand 4 and α-helix B of the p66 fingers subdomain and the β8–αE connecting loop and β-strand 5a of the palm subdomain, as well as α-helices H and I and the β12-β13 hairpin in the thumb subdomain.[38] The process of reverse transcription requires coordination between the two catalytic centers. However, in agreement with the notion that most interactions with the nucleic acid substrate are mediated through the polymerase domain, the equilibrium dissociation constants (K_D) obtained at the RNase

H active site were about 5-fold higher than those obtained at the polymerase active site.[31] During reverse transcription nucleotide incorporation rate is about 5-fold faster than the RNase H cleavage rate.[39] The inefficient RNase H reaction may, in fact, be an important feature of the enzyme activity so that the RNase H site does not bind tightly to the RNA/DNA duplex at any one site as that may impede the movement of the duplex driven by the polymerase.[39]

The discussion above is also pertinent in the case of MLV RT. Two engineered MLV RTs with increased thermal stability (E69K/E302R/W313F/L435G/N454K and L139P/D200N/T330P/L603W/E607K) also showed increased affinity for the template-primer.[15,16] However, Glu302 was the only residue located at the nucleic acid binding cleft of those enzymes, based on the available crystal structures of MLV RT. In contrast, other studies showed that introducing basic residues in the nucleic acid binding cleft of the enzyme (e.g. E286R, E302K or L435R) increased the reverse transcription efficiency of the MLV RT at high temperatures.[12,40] Moreover, further characterization of these mutants revealed that their improved performance in RT-PCR assays was due to RNase H inactivation,[41] as shown in our study for Q500K and Q500R mutants of O3M RT. This loss of RNase H activity could not be attributed to weaker nucleic acid binding interactions but to changes in the conformation of critical residues in the RNase H active site (mainly at Glu478) that play a critical role in the stabilization of the cleavage product, acting as proton donors in the nucleolytic reaction. Despite the loss of RNase H activity, the Q500K and Q500R mutants still retain remarkable cDNA synthesis efficiency and provide an alternative option to inactivate the RNase H activity of HIV/SIV RTs. To the best of our knowledge these are the first reported RNase H-inactivating mutations outside the catalytic site that preserve DNA polymerization function in HIV-1 RT and structurally related enzymes.

Materials and Methods

Nucleotides, oligonucleotides and template-primers

Stock solutions (100 mM) of individual dNTPs were obtained from GE Healthcare. DNA oligonucleotides 21P (5′-ATACTTTAACCATATG TATCC-3′), 25PGA (5′-TGGTAGGGCTATA CATTCTTGCAGG-3′), D38 (5′-GGGTCCTTTCTT ACCTGCAAGAATGTATAGCCCTACCA-3′), ACT1 (5′-CCTAGGCACCAGGGTGTGAT-3′), ACT3 (5′-CGTACTCCTGCTTGCTGATCC-3′), ACT-q1 (5′-CTAAGGCCAACCGTGAAAAG-3′), ACT-q2 (5′-ACCAGAGGCATACAGGGACA-3′), G1 (5′-CATCACTGCCACCCAGAAGA-3′) and G2 (5′-CCACCTGGTGCTCAGTGTAG-3′), and RNA oligonucleotides 31Trna (5´-UUUUUUUUUAG GAUACAUAUGGUUAAAGUAU-3′) and D38rna (5′-GGGUCCUUUCUUACCUGCAAGAAU GUAUAGCCCUACCA-3′) were obtained from Integrated DNA Technologies (Coralville, Iowa, USA). Oligonucleotides used in enzymatic assays were labeled at their 5′ termini with [γ-^{32}P]ATP (10 mCi/mL; 3000 Ci/mmol) (Perkin Elmer) and T4 polynucleotide kinase (Promega), and then annealed to their corresponding templates or primers depending on the experiment.

Mutagenesis

Site-directed mutagenesis was carried out using the standard QuikChange™ protocol (Stratagene). The templates used were derivatives of plasmid p66RTB containing either the wild-type (WT) RT-coding sequence of the HIV-1$_{BH10}$ strain[42] or the mutant HIV-1 group O RTs K358R/A359G/S360A (O3M) and K358R/A359G/S360A/E478Q (O3MQ).[19] Complementary mutagenic primers (Supplementary Table S1) were used to amplify entire plasmids in a thermocycling reaction carried out with high-fidelity Pfu DNA polymerase. After mutagenesis, RT-coding regions were entirely sequenced and, if correct, used for RT expression and purification.

Expression and purification of recombinant HIV-1 RTs

WT HIV-1$_{BH10}$ RT and mutant HIV-1 group O RTs were obtained as previously described.[10,19,42,43] Briefly, poly-histidine tagged p66 subunits were co-expressed together with the HIV-1 protease in *Escherichia coli* XL1 Blue. The obtained p66/p51 heterodimers were purified by cation exchange chromatography followed by inmobilized metal affinity chromatography on Ni^{2+}-nitroacetic acid agarose. RTs were quantified by active site titration before biochemical studies, using template-primer D38/25PGA.[42,44]

DNA binding affinity assays

Dissociation equilibrium constants (K_D) for RTs and template-primers were obtained by following a previously described procedure.[45] DNA/DNA and RNA/DNA duplexes were prepared with the ^{32}P-labeled 25PGA primer and templates D38 and D38rna, respectively. Briefly, RTs (at concentrations between 3 and 5 nM) were pre-incubated during 10 min at 37 °C (or 50 °C in some experiments) with increasing concentrations of the 5´-^{32}P-labeled template-primer (1.25 to 60 nM). Reactions were carried out in 50 mM Hepes buffer (pH 7.0), containing 15 mM NaCl, 15 mM magnesium acetate, 130 mM potassium acetate, 1 mM dithiothreitol, and 5% polyethylene glycol, and initiated by adding dTTP to a final concentration of 100 μM. The high concentration of Mg^{2+} was used to avoid premature dissociation of the template-primer in the presence of E478Q,[46] and to ensure an excess of Mg^{2+} over

dTTP in the polymerization reaction, particularly for enzymes showing low dNTP binding affinity. Aliquots were removed at different times (between 10 and 40 s) and quenched with an equal amount of sample loading buffer (10 mM EDTA in 90% formamide containing 3 mg/mL xylene cyanol FF and 3 mg/mL bromophenol blue). After heating the sample at 90 °C for 10 min, reaction products were separated on denaturing polyacrylamide gel electrophoresis and quantified by phosphorimaging in a BAS 1500 scanner (Fuji), using the program TINA version 2.09 (Raytest Isotopenmessgerate Gmbh, Staubenhardt, Germany). Burst amplitudes (i.e., RT bound to template-primer at time zero) were plotted against the template-primer concentration in the assay and the data were fitted to a quadratic equation to get the equilibrium dissociation constant (K_D) for RT binding to template-primer.[47] K_D values were determined using GraphPad Prism version 6.00 for Windows (GraphPad Software, La Jolla, CA).

RNase H activity assays

RT's RNase H cleavage efficiencies were determined with the RNA/DNA template-primer 31Trna/21P.[10] The RNA oligonucleotide (31Trna) was first labeled at its 5′ end with ^{32}P, and then purified with a Quick Spin™ column (Roche Applied Science) and annealed to DNA primer 21P. Cleavage reactions were carried out in 50 mM Tris-HCl pH 8.0 buffer, containing 50 mM NaCl, and 6 mM $MgCl_2$.[45,48] The concentration of the template-primer in these assays was 25 nM, and RT concentrations varied between 25 and 50 nM. RTs and template-primers were pre-incubated for 5 min at 37 °C and then reactions were initiated by adding $MgCl_2$. Aliquots were removed at appropriate times (15, 30 and 90 s) and quenched with an equal volume of sample loading buffer. RNase H cleavage products were analyzed after heating the quenched aliquots at 90 °C for 10 min, by denaturing polyacrylamide gel electrophoresis followed by phosphorimaging analysis as described above.

Reverse Transcription-PCR assays

The efficiency of RTs in reverse transcription reactions was determined using a two-step RT-PCR assay.[10,19] RT reactions were carried out in 20 μL of 50 mM Tris-HCl (pH 8.3) buffer, containing 75 mM KCl, 3 mM $MgCl_2$, 10 mM dithiothreitol, 0.8 U/μL RNase inhibitor (RNasin® Plus, Promega), 500 μM each of dATP, dCTP, dGTP and dTTP, 5 μM oligo$(dT)_{16}$, 1–200 ng of mouse liver total RNA (Cat# 736009–41; Lot# 0006062881; Agilent Technologies, Santa Clara, CA, USA) and the corresponding RT at 150 nM. The RNA template and the oligo$(dT)_{16}$ were pre-incubated at 70 °C for 5 minutes and cooled on ice to prevent secondary structure formation. The cDNA synthesis reactions were carried out at different temperatures in the range of 37–85 °C for 60 min and then stopped after heating the sample for 10 min at 90 °C. Subsequent PCR reactions were performed in 50 μL solutions, containing 2 μL of each cDNA synthesis reaction, 200 μM of each dNTP, 1.25 U of Taq DNA Polymerase (New England Biolabs) and 200 ng each of primers ACT1 and ACT3 (for detecting β-actin cDNA). Taq DNA polymerase reaction buffer was used as indicated by the enzyme supplier. Cycling conditions for all targets were as follows: an incubation at 95 °C for 2 min; 30 cycles at 95 °C for 30 s, 55 °C for 1 min, and 72 °C for 1 min; and a final incubation at 72 °C for 10 min. RT-PCR reactions were also carried out as described above but using total RNA from human HeLa cells (Cat# 636543, Lot#2101023) (Takara Bio, Kusatsu, Japan) in the reverse transcription step, and primers G1 and G2 in the PCR amplification step to detect the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene product. PCR products were analyzed by agarose gel electrophoresis to determine PCR product yields. Band intensities were quantified with the image processing program ImageJ [49] as previously described.[50] Yields were calculated as band intensity ratios using as reference the amplification products obtained with O3MQ RT.

Quantitative real-time PCR assays

Quantitative PCR reactions were carried out in a 10 μL volume, containing 1 μL of cDNA obtained in reverse transcription reactions described above, EvaGreen® Dye Supermix and gene-specific primers ACT-q1 and ACT-q2 for detecting β-actin cDNA, or G1 and G2 for detecting GAPDH cDNA. Reactions were carried out in triplicate in a BioRad CFX-384 system. The cycling protocol included an initial denaturation at 95 °C for 30 s, 40 two-step cycles (5 s at 95 °C and 5 s at 60 °C) and a melting curve from 60 to 95 °C. Obtained cycle threshold (C_T) values were used to calculate cDNA levels, and reverse transcription efficiencies were determined as fold differences relative to data obtained with O3MQ RT using the formula $^{2\text{-}\Delta CT}$.[51]

Molecular modeling and molecular dynamics simulations

A structural model of mutant K358R/A359G/S360A/E478Q HIV-1_{ESP49} RT bound to double-stranded DNA and dTTP was constructed by standard homology modeling using as reference the one previously obtained for the WT HIV-1_{ESP49} RT, which was based on the crystal structure of the ternary complex containing WT HIV-1_{HXB2} RT (PDB file 1RTD).[19,27] The effects of different amino acid substitutions in the conformation and interaction energies were analyzed after performing molecular dynamics simulations using the AMBER software.[52] In these studies, we used the pam19sb forcefield[53] for the protein and the OL15 forcefield[54] for the nucleic acid. The systems were solvated with

a box of TIP3P waters[55] with 12 Å as the shortest distance between any atom in the protein and the periodic box. For each system we start with a minimization followed by a 500-ps heating phase in which the temperature was raised from 1.0 to 300 K. Finally, we perform a free molecular dynamics simulation with time steps of 2 fs during 100 ns, at 300 K or 350 K. The interaction energies between protein and nucleic acids were calculated with the MMPBSA software[56] for the last 50 ns of the free molecular dynamics simulation. Molecular dynamics trajectories were analyzed using intermolecular contact maps.[57]

Author Contributions

Experimental design and execution: J.M.R., N.L.-C., J.I.M.-M., O.H.-C. and A.S.-I. performed experiments and analyzed data; supervision: J.M. and L.M.-A.; writing and editing: J.M.R., N.L.-C. and L.M.-A.; project conceptualization and administration: L.M.-A.

DATA AVAILABILITY

Data will be made available on request.

DECLARATION OF COMPETING INTEREST

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: O3M and O3MQ RTs (sold under the commercial name SunScript™) and methods for their use are the subject of patents and patent applications that have been licensed by CSIC to 4Basebio AG. L. M.-A. receives royalty payments for the sales of relevant enzymes and kits.

Acknowledgements

We thank the "Servicio de genómica y secuenciación masiva del CBMSO" for technical assistance with qPCR assays. This work was supported in part by the Ministry of Science and Innovation of Spain through grants PID2019-104176RB-I00/AEI/MCI/10.13039/501100011033 awarded to L.M.-A, and PID2021-125604NB-I00/AEI/MCI/10.13039/501100011033 (supporting J. M. and J.I.M.-M.). J.M.R. is a predoctoral fellow of the Spanish Ministry of Universities (Formación de Profesorado Universitario, FPU19/01653). N.L.-C. was supported by a contract (PEJ-2020-AI/BMD-19429) of the Youth Guarantee programme of the European Union, with the participation of the Comunidad de Madrid (Consejería de Educación, Universidades, Ciencia y Portavocía). A.S.-I. received a short-term fellowship from CSIC (JAE-Intro program, JAEINT_20_00736/JAEINT_20_EX_0774). An institutional grant of Fundación Ramón Areces to the CBMSO is also acknowledged. The CBMSO has been certified since 2023 as Severo Ochoa Center of Excellence by AEI/MCI/10.13039/501100011033. The team at CBMSO is member of the Global Virus Network.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2023.168219.

Received 5 May 2023;
Accepted 26 July 2023;
Available online 1 August 2023

Keywords:
reverse transcriptase;
RNase H;
polymerase engineering;
RT-PCR;
template-primer binding

Abbreviations:
AMV, avian myeloblastosis virus; cDNA, complementary DNA; DTT, dithiothreitol; HIV-1, human immunodeficiency virus type 1; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; IN, integrase; MLV, murine leukemia virus; qPCR, quantitative real-time PCR; RNase H, ribonuclease H; RT, reverse transcriptase; SIV, simian immunodeficiency virus; WT, wild-type

References

1. Hu, W.S., Hughes, S.H., (2012). HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**, a006882
2. Menéndez-Arias, L., Sebastián-Martín, A., Álvarez, M., (2017). Viral reverse transcriptases. *Virus Res.* **234**, 153–176.
3. Martín-Alonso, S., Frutos-Beltrán, E., Menéndez-Arias, L., (2021). Reverse transcriptase: from transcriptomics to genome editing. *Trends Biotechnol.* **39**, 194–210.
4. Barrioluengo, V., Alvarez, M., Barbieri, D., Menéndez-Arias, L., (2011). Thermostable HIV-1 group O reverse transcriptase variants with the same fidelity as murine leukaemia virus reverse transcriptase. *Biochem. J* **436**, 599–607.
5. Mohr, S., Ghanem, E., Smith, W., Sheeter, D., Qin, Y., King, O., Polioudakis, D., Iyer, V.R., et al., (2013). Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA* **19**, 958–970.
6. Stamos, J.L., Lentzsch, A.M., Lambowitz, A.M., (2017). Structure of a thermostable group II intron reverse transcriptase with template-primer and its functional and evolutionary implications. *Mol. Cell* **68**, 926–939.e4.
7. Álvarez, M., Sebastián-Martín, A., García-Marquina, G., Menéndez-Arias, L., (2017). Fidelity of classwide-resistant HIV-2 reverse transcriptase and differential contribution of

K65R to the accuracy of HIV-1 and HIV-2 reverse transcriptases. *Sci. Rep.* **7**, 44834.
8. Sebastián-Martín, A., Barrioluengo, V., Menéndez-Arias, L., (2018). Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases. *Sci. Rep.* **8**, 627.
9. Oscorbin, I.P., Novikova, L.M., Filipenko, M.L., (2022). Comparison of reverse transcriptase (RT) activities of various M-MuLV RTs for RT-LAMP assays. *Biology (Basel)* **11**, 1809.
10. Alvarez, M., Matamoros, T., Menéndez-Arias, L., (2009). Increased thermostability and fidelity of DNA synthesis of wild-type and mutant HIV-1 group O reverse transcriptases. *J. Mol. Biol.* **392**, 872–884.
11. Gerard, G.F., Potter, R.J., Smith, M.D., Rosenthal, K., Dhariwal, G., Lee, J., Chatterjee, D.K., (2002). The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res.* **30**, 3118–3129.
12. Yasukawa, K., Mizuno, M., Konishi, A., Inouye, K., (2010). Increase in thermal stability of Moloney murine leukaemia virus reverse transcriptase by site-directed mutagenesis. *J. Biotechnol.* **150**, 299–306.
13. Konishi, A., Yasukawa, K., Inouye, K., (2012). Improving the thermal stability of avian myeloblastosis virus reverse transcriptase α-subunit by site-directed mutagenesis. *Biotechnol. Letter* **34**, 1209–1215.
14. Oscorbin, I.P., Filipenko, M.L., (2021). M-MuLV reverse transcriptase: selected properties and improved mutants. *Comput. Struct. Biotechnol. J.* **19**, 6315–6327.
15. Arezi, B., Hogrefe, H., (2009). Novel mutations in Moloney murine leukemia virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res.* **37**, 473–481.
16. Baranauskas, A., Paliksa, S., Alzbutas, G., Vaitkevicius, M., Lubiene, J., Letukiene, V., Burinskas, S., Sasnauskas, G., et al., (2012). Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Eng. Des. Sel.* **25**, 657–668.
17. Skirgaila, R., Pudzaitis, V., Paliksa, S., Vaitkevicius, M., Janulaitis, A., (2013). Compartmentalization of destabilized enzyme-mRNA-ribosome complexes generated by ribosome display: a novel tool for the directed evolution of enzymes. *Protein Eng. Des. Sel.* **26**, 453–461.
18. Katano, Y., Li, T., Baba, M., Nakamura, M., Ito, M., Kojima, K., Takita, T., Yasukawa, K., (2017). Generation of thermostable Moloney murine leukemia virus reverse transcriptase variants using site saturation mutagenesis library and cell-free protein expression system. *Biosci. Biotech. Bioch.* **81**, 2339–2345.
19. Matamoros, T., Barrioluengo, V., Abia, D., Menéndez-Arias, L., (2013). Major groove binding track residues of the connection subdomain of human immunodeficiency virus type 1 reverse transcriptase enhance cDNA synthesis at high temperatures. *Biochemistry* **52**, 9318–9328.
20. Sarafianos, S.G., Das, K., Tantillo, C., Clark Jr, A.D., Ding, J., Whitcomb, J.M., Boyer, P.L., Hughes, S.H., et al., (2001). Crystal structure of HIV-1 reverse transcriptase in complex with a polypurine tract RNA:DNA. *EMBO J.* **20**, 1449–1461.
21. Figiel, M., Krepl, M., Poznanski, J., Golab, A., Sponer, J., Nowotny, M., (2017). Coordination between the polymerase and RNase H activity of HIV-1 reverse transcriptase. *Nucleic Acids Res.* **45**, 3341–3352.
22. Tian, L., Kim, M.S., Li, H., Wang, J., Yang, W., (2018). Structure of HIV-1 reverse transcriptase cleaving RNA in an RNA/DNA hybrid. *PNAS* **115**, 507–512.
23. Ding, J., Das, K., Hsiou, Y., Sarafianos, S.G., Clark Jr, A. D., Jacobo-Molina, A., Tantillo, C., Hughes, S.H., et al., (1998). 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.
24. Julias, J.G., McWilliams, M.J., Sarafianos, S.G., Arnold, E., Hughes, S.H., (2002). Mutations in the RNase H domain of HIV-1 reverse transcriptase affect the initiation of DNA synthesis and the specificity of RNase H cleavage in vivo. *PNAS* **99**, 9515–9520.
25. Rausch, J.W., Lener, D., Miller, J.T., Julias, J.G., Hughes, S.H., Le Grice, S.F.J., (2002). Altering the RNase H primer grip of human immunodeficiency virus reverse transcriptase modifies cleavage specificity. *Biochemistry* **41**, 4856–4865.
26. Ilina, T.V., Brosenitsch, T., Sluis-Cremer, N., Ishima, R., (2021). Retroviral RNase H: Structure, mechanism, and inhibition. *Enzymes* **50**, 227–247.
27. Huang, H., Chopra, R., Verdine, G.L., Harrison, S.C., (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. *Science* **282**, 1669–1675.
28. Menéndez-Arias, L., Abraha, A., Quiñones-Mateu, M.E., Mas, A., Camarasa, M.J., Arts, E.J., (2001). Functional characterization of chimeric reverse transcriptases with polypeptide subunits of highly divergent HIV-1 group M and O strains. *J. Biol. Chem.* **276**, 27470–27479.
29. Abram, M.E., Sarafianos, S.G., Parniak, M.A., (2010). The mutation T477A in HIV-1 reverse transcriptase (RT) restores normal proteolytic processing of RT in virus with Gag-Pol mutated in the p51-RNH cleavage site. *Retrovirology* **7**, 6.
30. Wendeler, M., Lee, H.F., Bermingham, A., Miller, J.T., Chertov, O., Bona, M.K., Baichoo, N.S., Ehteshami, M., et al., (2008). Vinylogous ureas as a novel class of inhibitors of reverse transcriptase-associated ribonuclease H activity. *ACS Chem. Biol.* **3**, 635–644.
31. Beilhartz, G.L., Wendeler, M., Baichoo, N., Rausch, J., Le Grice, S., Götte, M., (2009). HIV-1 reverse transcriptase can simultaneously engage its DNA/RNA substrate at both DNA polymerase and RNase H active sites: implications for RNase H inhibition. *J. Mol. Biol.* **388**, 462–474.
32. Corona, A., Di Leva, F.S., Thierry, S., Pescatori, L., Cuzzucoli Crucitti, G., Subra, F., Delelis, O., Esposito, F., et al., (2014). Identification of highly conserved residues involved in inhibition of HIV-1 RNase H function by diketo acid derivatives. *Antimicrob. Agents Chemother.* **58**, 6101–6110.
33. Poongavanam, V., Corona, A., Steinmann, C., Scipione, L., Grandi, N., Pandolfi, F., Di Santo, R., Costi, R., et al., (2018). Structure-guided approach identifies a novel class of HIV-1 ribonuclease H inhibitors: binding mode insights through magnesium complexation and site-directed mutagenesis studies. *Medchemcomm.* **9**, 562–575.
34. Arion, D., Sluis-Cremer, N., Min, K.L., Abram, M.E., Fletcher, R.S., Parniak, M.A., (2002). Mutational analysis of Tyr-501 of HIV-1 reverse transcriptase. Effects on ribonuclease H activity and inhibition of this activity by N-acylhydrazones. *J. Biol. Chem.* **277**, 1370–1374.

35. Rausch, J.W., Qu, J., Yi-Brunozzi, H.Y., Kool, E.T., Le Grice, S.F., (2003). Hydrolysis of RNA/DNA hybrids containing nonpolar pyrimidine isosteres defines regions essential for HIV type 1 polypurine tract selection. *PNAS* **100**, 11279–11284.
36. Delviks-Frankenberry, K.A., Nikolenko, G.N., Barr, R., Pathak, V.K., (2007). Mutations in human immunodeficiency virus type 1 RNase H primer grip enhance 3'-azido-3'-deoxythymidine resistance. *J. Virol.* **81**, 6837–6845.
37. Dürr, S.L., Bohuszewicz, O., Berta, D., Suardiaz, R., Jambrina, P.G., Peter, C., Shao, Y., Rosta, E., (2021). The role of conserved residues in the DEDDh motif: the proton-transfer mechanism of HIV-1 RNase H. *ACS Catal.* **11**, 7915–7927.
38. Dash, C., Scarth, B.J., Badorrek, C., Götte, M., Le Grice, S. F., (2008). Examining the ribonuclease H primer grip of HIV-1 reverse transcriptase by charge neutralization of RNA/DNA hybrids. *Nucleic Acids Res.* **36**, 6363–6371.
39. Li, A., Li, J., Johnson, K.A., (2016). HIV-1 reverse transcriptase polymerase and RNase H (ribonuclease H) active sites work simultaneously and independently. *J. Biol. Chem.* **291**, 26566–26585.
40. Konishi, A., Hisayoshi, T., Yokokawa, K., Barrioluengo, V., Menéndez-Arias, L., Yasukawa, K., (2014). Amino acid substitutions away from the RNase H catalytic site increase the thermal stability of Moloney murine leukemia virus reverse transcriptase through RNase H inactivation. *Biochem. Biophys. Res. Commun.* **454**, 269–274.
41. Konishi, A., Ma, X., Yasukawa, K., (2014). Stabilization of Moloney murine leukemia virus reverse transcriptase by site-directed mutagenesis of surface residue Val433. *Biosci. Biotech. Bioch.* **78**, 75–78.
42. Matamoros, T., Deval, J., Guerreiro, C., Mulard, L., Canard, B., Menéndez-Arias, L., (2005). Suppression of multidrug-resistant HIV-1 reverse transcriptase primer unblocking activity by alpha-phosphate-modified thymidine analogues. *J. Mol. Biol.* **349**, 451–463.
43. Boretto, J., Longhi, S., Navarro, J.M., Selmi, B., Sire, J., Canard, B., (2001). An integrated system to study multiply substituted human immunodeficiency virus type 1 reverse transcriptase. *Anal. Biochem.* **292**, 139–147.
44. Kati, W.M., Johnson, K.A., Jerva, L.F., Anderson, K.S., (1992). Mechanism and fidelity of HIV reverse transcriptase. *J. Biol. Chem.* **267**, 25988–25997.
45. Betancor, G., Puertas, M.C., Nevot, M., Garriga, C., Martínez, M.A., Martinez-Picado, J., Menéndez-Arias, L., (2010). Mechanisms involved in the selection of HIV-1 reverse transcriptase thumb subdomain polymorphisms associated with nucleoside analogue therapy failure. *Antimicrob. Agents Chemother.* **54**, 4799–4811.
46. Cristofaro, J.V., Rausch, J.W., Le Grice, S.F.J., DeStefano, J.J., (2002). Mutations in the ribonuclease H active site of HIV-RT reveal a role for this site in stabilizing enzyme-primer-template binding. *Biochemistry* **41**, 10968–10975.
47. Menéndez-Arias, L., (1998). Studies on the effects of truncating α-helix E´ of p66 human immunodeficiency virus type 1 reverse transcriptase on template-primer binding and fidelity of DNA synthesis. *Biochemistry* **37**, 16636–16644.
48. Betancor, G., Álvarez, M., Marcelli, B., Andrés, C., Martínez, M.A., Menéndez-Arias, L., (2015). Effects of HIV-1 reverse transcriptase connection subdomain mutations on polypurine tract removal and initiation of (+)-strand DNA synthesis. *Nucleic Acids Res.* **43**, 2259–2270.
49. Schneider, C.A., Rasband, W.S., Eliceiri, K.W., (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.
50. Wittmeier, P., Hummel, S., (2022). Agarose gel electrophoresis to assess PCR product yield: comparison with spectrophotometry, fluorometry and qPCR. *Biotechniques* **72**, 155–158.
51. Schmittgen, T.D., Livak, K.J., (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature Protoc.* **3**, 1101–1108.
52. Case, D.A. et al, (2022). Amber 2022. University of California, San Francisco.
53. Tian, C., Kasavajhala, K., Belfon, K., Raguette, L., Huang, H., Migues, A., Bickel, J., Wang, Y., et al., (2020). ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* **16**, 528–552.
54. Zgarbova, M., Sponer, J., Otyepka, M., Cheatham, T.E., Galindo-Murillo, R., Jurecka, P., (2015). Refinement of the sugar-phosphate backbone torsion beta for the AMBER force fields improves the description of Z-DNA and B-DNA. *J. Chem. Theory Comput.* **11**, 5723–5736.
55. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
56. Miller, B.R., McGee, T.D., Swails, J.M., Homeyer, N., Gohlke, H., Roitberg, A.E., (2012). MMPBSA.py: an efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321.
57. Abdel-Azeim, S., Chermak, E., Vangone, A., Oliva, R., Cavallo, L., (2014). MDcons: intermolecular contact maps as a tool to analyze the interface of protein complexes from molecular dynamics trajectories. *BMC Bioinf.* **15**, S1.