



HAL
open science

Structural basis for the role of LYS220 as proton donor for nucleotidyl transfer in HIV-1 Reverse transcriptase

Servaas Michielssens, Samuel L.C. Moors, Mathy Froeyen, Piet Herdewijn, Arnout Ceulemans

► To cite this version:

Servaas Michielssens, Samuel L.C. Moors, Mathy Froeyen, Piet Herdewijn, Arnout Ceulemans. Structural basis for the role of LYS220 as proton donor for nucleotidyl transfer in HIV-1 Reverse transcriptase. *Biophysical Chemistry*, 2011, 10.1016/j.bpc.2011.03.009 . hal-00759072

HAL Id: hal-00759072

<https://hal.science/hal-00759072>

Submitted on 30 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

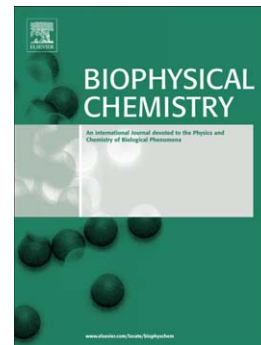
Structural basis for the role of LYS220 as proton donor for nucleotidyl transfer in HIV-1 Reverse transcriptase

Servaas Michielssens, Samuel L.C. Moors, Mathy Froeyen, Piet Herdewijn, Arnout Ceulemans

PII: S0301-4622(11)00066-4
DOI: doi: [10.1016/j.bpc.2011.03.009](https://doi.org/10.1016/j.bpc.2011.03.009)
Reference: BIOCHE 5493

To appear in: *Biophysical Chemistry*

Received date: 22 February 2011
Revised date: 25 March 2011
Accepted date: 26 March 2011



Please cite this article as: Servaas Michielssens, Samuel L.C. Moors, Mathy Froeyen, Piet Herdewijn, Arnout Ceulemans, Structural basis for the role of LYS220 as proton donor for nucleotidyl transfer in HIV-1 Reverse transcriptase, *Biophysical Chemistry* (2011), doi: [10.1016/j.bpc.2011.03.009](https://doi.org/10.1016/j.bpc.2011.03.009)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Structural basis for the role of LYS220 as proton donor for nucleotidyl transfer in HIV-1 Reverse transcriptase

Servaas Michielssens^{a,*}, Samuel L.C. Moors^a, Mathy Froeyen^b, Piet Herdewijn^b, Arnout Ceulemans^a

^a*Departement of Chemistry and INPAC Institute of Nanoscale Physics & Chemistry, Katholieke Universiteit Leuven, Leuven, Belgium*

^b*Rega Institute for Medicinal Research, Katholieke Universiteit Leuven, Leuven, Belgium*

Abstract

Biochemical studies by Castro *et al.* have recently revealed a crucial role for a general acid in the catalysis of nucleic acid transfer in distinct classes of polymerases. For HIV-RT LYS220 was identified as proton donor. This was unanticipated from a structural point of view, since in all ternary crystal structures of HIV-RT LYS220 is too distant from the active site to fulfill this role. In this work molecular dynamics simulations were used to reveal the dynamics of HIV-RT and to provide structural evidence for the role of LYS220. During a 1 μ s molecular dynamics simulation LYS220 migrates toward the active site and occupies several positions enabling direct and water mediated proton transfer towards pyrophosphate. A combination of quantum mechanical and molecular mechanics methods was used to validate the different modes of interaction.

Keywords: HIV reverse transcriptase, QM/MM, Molecular Dynamics, Protein dynamics, Enzymatic catalysis

1. Introduction

HIV reverse transcriptase is a critical enzyme in the HIV life cycle, catalyzing the transcription of the viral RNA to DNA. It is also a target for numerous drugs. Detailed knowledge of the catalytic mechanism is vital

*Celestijnenlaan 200F, 3001 Heverlee, Belgium, tel. +32 16 327384 fax. +32 16 327992
Email address: servaas.michielssens@chem.kuleuven.be (Servaas Michielssens)

to rationally develop nucleoside like drugs that cause chain termination [1]. The active site and reaction are schematically represented in Figure 1. Nucleophilic attack on the phosphorus atom of the ribonucleoside triphosphate by the primer 3'-hydroxyl leads to the formation of a phosphodiester bond and release of pyrophosphate. Two Mg^{2+} ions are present in the active site. Mg^{2+} A lowers the pKa of the primer's 3'-hydroxyl to facilitate the initial proton transfer [2, 3]. Mg^{2+} B has a structural role, it stabilizes the negative charge in the transition state and might help the release of pyrophosphate [4, 5]. A second proton transfer involves the protonation of triphosphate[6].

The acceptor for the first proton transfer is still unknown. Potential acceptors are ASP186 or water. The donor of the second proton transfer was recently identified by Castro *et al.* by sequence alignment and confirmed by biochemical studies[7]. It was found that LYS220 fulfills this role in HIV-RT. Moreover, lysine was revealed to function in several classes of polymerases as general acid. However, structural evidence for the role of LYS220 is missing since it is not situated near the active site in any of the ternary crystal structures of HIV-RT[8, 9, 10]. In Figure 2 the position of LYS220 in the crystal structure is shown (the position of LYS220 in the crystal structure is indicated by the blue color, the code of the crystal structure used is 1RTD). The side chain nitrogen of LYS220 is located more than 1.5 nm from the active site. This crystal structure cannot give a structural justification for its role in the catalysis. Molecular dynamics simulations in this work reveal the dynamics of this residue and provide the necessary structural evidence.

2. Methods

2.1. Molecular dynamics

2.1.1. Model building

The simulation was started from the 1RTD crystal structure[10]. The p66 unit, including DNA was used for the simulations. The base of the entering thymidine triphosphate was modified to an adenine and the complementary adenine base (E5) was changed into a thymine by an inverse fitting procedure using Quatfit (Quatfit program in CCL software archives).

The AMBER 99SB force field was used for DNA and the protein[11]. It was verified that the DNA backbone parameters α and γ stayed close to their canonical value, since transition to unrealistic values is a known problem for free DNA simulations in the AMBER force field[12]. The triphosphate

parameters were taken from Meagher *et al.*[13] Protonation states were calculated using the H++-server[14] and confirmed by the PROPKA server[15]. The 3'-hydroxyl group of the primer strand is missing in the crystal structure and was added using tleap[16]. The AMBER topology and structure file were converted to GROMACS using amb2gmx[17, 18].

2.1.2. Simulation setup

A dodecahedral simulation box was added. The dimensions were set to the diameter of the system (largest distance between 2 atoms) plus twice 0.8 nm. The system was solvated with 38349 TIP3P water molecules[19]. To neutralize the system 32 sodium ions were added. The total system consists of 125661 atoms. The GROMACS simulation package[20] was used to simulate the neutralized solvated system. Long-range electrostatics were computed with the particle mesh Ewald method[21]. The non-bonded cutoff was set to 1.1 nm. A fourier spacing of 0.12 nm was used. All covalent bonds were constrained using P-LINCS[22]. A timestep of 2 fs was used.

The system was first minimized using the steepest descent method, followed by a slow heating of the system from 0 to 300K during 500ps. Next the system was subjected to 500ps of equilibration at 300K, to equilibrate the energy. The system temperature was controlled by the velocity rescaling thermostat[23] and the pressure was controlled using the Parinello-Rhman barostat[24]. Finally the system was simulated in a production run of 1 μ s.

2.2. QM/MM optimizations

2.2.1. Methods

The QM/MM calculations were done with a two-layer ONIOM [25, 26, 27] approach. The selected model system (Figure 3) was treated with density functional theory (DFT) with the B3LYP[28] functional and DGDZVP basis set[29, 30], the rest of the system was treated with the AMBER 99SB force field. The QM model was mechanically embedded[31] in the MM system. The boundary between the 2 regions was treated using link hydrogen atoms. This combination of the AMBER force field and B3LYP has proven its use in many studies[32, 33]. For the optimization the macro/microiterative scheme with quadratic coupling was used[34]. The calculations were done using Gaussian 09[35].

2.2.2. Model system

Three clusters of interactions between the active site were determined by molecular dynamics simulations. Representative snapshots were chosen and optimized. First, the fully solvated structure was optimized using the AMBER99SB force field followed by a QM/MM optimization of the system with a water shell. The water shell consists of all water molecules within 0.5 nm of the system or within 1.2 nm of the deoxyadenosine triphosphate substrate. Figure 3 shows the basic QM-region applied in the present work. The triphosphate moiety of dATP was included. Both Mg^{2+} ions are coordinated by 6 ligands, indicated by the broken line, and the entire coordination shell is treated on the QM-level. The full final sugar of the primer strand was included to avoid cutting the ring system. Additionally LYS220 was added since MD simulations demonstrated the interaction with the substrate. Additional elements were added in the structure where LYS220 has water mediated hydrogen bonds with the triphosphate moiety. In this structure the water mediating the bond is included, but also ASP67, which forms a salt bridge with LYS220 in this structure.

All residues which have an atom within 0.6 nm of the QM-region were treated as being flexible during the optimization, residues further away were kept fixed. The input for Gaussian09 was prepared using a toolkit developed by Tao and Schlegel[36] to assist ONIOM calculations.

3. Results and Discussion

To understand the role of LYS220, molecular dynamics (MD) simulations were made with the AMBER 99SB force field using the GROMACS simulation package. The complex was simulated for $1\mu\text{s}$. The evolution of the position of LYS220 toward the triphosphate moiety is shown in Figure 4A. Initially LYS220 stays in the crystallographic position where it is part of a beta sheet. The strand containing LYS220 has only 2 hydrogen bonds connecting it to the rest of the sheet. After 30ns of simulation the LYS220 detaches from the beta sheet. During the next 40ns the loop explores the conformational space, in which it rotates towards the active site with the head group directed toward the protein. A snapshot at 50ns is shown in green in Figure 2. After 90ns LYS220 enters a regime in which it is tightly bound to the triphosphate. A snapshot at 100ns is shown in red in Figure 2. LYS220 is now part of a 3_{10} -helix formed by LYS220 and PRO217. During the remaining 300ns LYS220 stays within 1 nm of the triphosphate. Studying

the distance, hydrogen bonds and water mediated hydrogen bonds we can discriminate between three states of interaction.

To validate the different states, representative structures of each cluster were optimized using QM/MM. The QM/MM minimized structures are shown in Figure 5. In structure I, between 92 and 112ns, LYS220 is within 0.37nm of triphosphate and in tight interaction via multiple direct hydrogen bonds with $P\alpha$ and $P\gamma$. The N-H bond length of the lysine NH_3 head group forming a hydrogen bond with the terminal phosphate is elongated from about 0.105nm to 0.118nm, indicating that this proton is more acidic in this state. Proton transfer to the terminal phosphate could be followed by transfer to the oxygen between $P\alpha$ and $P\beta$. This would be a substrate assisted proton transfer initiated by lysine. In structure II, between 112ns and 150ns, the LYS220-triphosphate distance is about 0.7nm. Here hydrogen bond interaction between LYS220 and triphosphate is mediated by one or more water molecules (see Figure 4 and 5). Notably this second state can also explain the experiments of Castro *et al.* since a water mediated proton transfer can also be catalyzed by LYS220. This has been shown for T7 DNA polymerase[37]. Structure II in Figure 5 shows an example of a water mediated hydrogen bond. In this structure LYS220 could also catalyze water mediated proton transfer. In this state ASP67 forms a salt bridge with LYS220, which could tune the pKa of lysine. In structure III, between 150 and 180ns, LYS220 is mostly bound to the γ -phosphate via direct hydrogen bonds and is at a distance of 0.41nm from LYS220. During the rest of the simulation the three states appear repeatedly.

To date, only three HIV-RT crystal structures which contain both the natural substrate and DNA are publicly available. In all three crystal structures, LYS220 is located in a beta-sheet and oriented away from the active site. To provide an extra assessment of the flexibility we studied 134 available HIV-RT crystal structures which were crystallized with different experimental conditions; namely, without substrate, without nucleic acid, with non-nucleoside like drugs, and so forth. Here, the interaction between LYS220 and the substrate cannot be studied, but both the flexibility of the structural element, of which LYS220 is part, and its secondary structure can be examined. As shown in Figure 6, the loop containing LYS220 is very flexible considering the ensemble of available crystal structures. This is on par with the results of the MD simulation and is also confirmed by hydrogen deuterium exchange rates measured by mass spectrometry[38]. Moreover in 24% of all crystal structures LYS220 is part of a 3_{10} -helix, for 7% it belongs to a beta sheet

and in 69% to a loop. The flexibility observed in the ensemble of non ternary crystal structures together with the existence of the 3_{10} -helix reinforces the validity of the simulation results.

Two other lysine residues, LYS70 and LYS219, which are initially much closer to the active site, also show a very dynamic behavior. Castro *et al.*[7] state that LYS219 can be ruled out as proton donor by comparison with HIV-2, where LYS219 is absent but the DNA polymerase activity is similar [39]. On the same basis one can rule out LYS70. In addition, the fact that the mutation of LYS220 done by Castro *et al.*[7], decreases the turnover rate shows that LYS70 and LYS219 cannot fully function as alternative for LYS220.

4. Conclusions

Extensive MD simulation demonstrated the dynamics of LYS220 and revealed for the first time several states that can explain its catalytic role in nucleic acid transfer. Simulations at higher level of theory indicated that the states found using force field methods are also valid minima on a quantum mechanical potential energy surface. This dynamic behavior of LYS220 is also observed in the ensemble of crystal structures which are crystallized without nucleic acid and substrate. Moreover, the presence of the 3_{10} -helix is also observed for some of these structures. The structural basis for the interaction between triphosphate and LYS220, proposed in this work, is of major importance for further study of the detailed mechanism for nucleotidyl transfer in HIV-RT.

- [1] O. Adelfinskaya, M. Terrazas, M. Froeyen, P. Marliere, K. Nauwelaerts, P. Herdewijn, Polymerase-catalyzed synthesis of DNA from phosphoramidate conjugates of deoxynucleotides and amino acids, *NUCLEIC ACIDS RESEARCH* 35 (2007) 5060–5072.
- [2] J. Florian, M. Goodman, A. Warshel, Computer simulation of the chemical catalysis of DNA polymerases: Discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase, *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* 125 (2003) 8163–8177.
- [3] S. Michielssens, N. T. Trung, M. Froeyen, P. Herdewijn, M. T. Nguyen, A. Ceulemans, Hydrolysis of aspartic acid phosphoramidate nucleotides:

- a comparative quantum chemical study, *PHYSICAL CHEMISTRY CHEMICAL PHYSICS* 11 (2009) 7274–7285.
- [4] M. Fothergill, M. Goodman, J. Petruska, A. Warshel, Structure-Energy analysis of the role of metal ions in phosphodiester bond hydrolysis by DNA-polymerase-I, *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* 117 (1995) 11619–11627.
- [5] T. Steitz, Structural biology - A mechanism for all polymerases, *NATURE* 391 (1998) 231–232.
- [6] C. Castro, E. Smidansky, K. R. Maksimchuk, J. J. Arnold, V. S. Korneeva, M. Gotte, W. Konigsberg, C. E. Cameron, Two proton transfers in the transition state for nucleotidyl transfer catalyzed by RNA- and DNA-dependent RNA and DNA polymerases, *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA* 104 (2007) 4267–4272.
- [7] C. Castro, E. D. Smidansky, J. J. Arnold, K. R. Maksimchuk, I. Moustafa, A. Uchida, M. Gotte, W. Konigsberg, C. E. Cameron, Nucleic acid polymerases use a general acid for nucleotidyl transfer, *NATURE STRUCTURAL & MOLECULAR BIOLOGY* 16 (2009) 212–218.
- [8] E. B. Lansdon, D. Samuel, L. Lagpacan, K. M. Brendza, K. L. White, M. Hung, X. Liu, C. G. Booramra, R. L. Mackman, T. Cihlar, A. S. Ray, M. E. McGrath, S. Swaminathan, Visualizing the Molecular Interactions of a Nucleotide Analog, GS-9148, with HIV-1 Reverse Transcriptase-DNA Complex, *JOURNAL OF MOLECULAR BIOLOGY* 397 (2010) 967–978.
- [9] K. Das, R. P. Bandwar, K. L. White, J. Y. Feng, S. G. Sarafianos, S. Tuske, X. Tu, J. A. D. Clark, P. L. Boyer, X. Hou, B. L. Gaffney, R. A. Jones, M. D. Miller, S. H. Hughes, E. Arnold, Structural Basis for the Role of the K65R Mutation in HIV-1 Reverse Transcriptase Polymerization, Excision Antagonism, and Tenofovir Resistance, *JOURNAL OF BIOLOGICAL CHEMISTRY* 284 (2009) 35092–35100.
- [10] H. Huang, R. Chopra, G. Verdine, S. Harrison, Structure of a covalently trapped catalytic complex of HIV-I reverse transcriptase: Implications for drug resistance, *SCIENCE* 282 (1998) 1669–1675.

- [11] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple amber force fields and development of improved protein backbone parameters, *PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS* 65 (2006) 712–725.
- [12] A. Perez, I. Marchan, D. Svozil, J. Sponer, I. T. E. Cheatham, C. A. Laughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: Improving the description of alpha/gamma conformers, *BIOPHYSICAL JOURNAL* 92 (2007) 3817–3829.
- [13] K. Meagher, L. Redman, H. Carlson, Development of polyphosphate parameters for use with the AMBER force field, *JOURNAL OF COMPUTATIONAL CHEMISTRY* 24 (2003) 1016–1025.
- [14] J. Gordon, J. Myers, T. Folta, V. Shoja, L. Heath, A. Onufriev, H++: a server for estimating pK(a)s and adding missing hydrogens to macromolecules, *NUCLEIC ACIDS RESEARCH* 33 (2005) W368–W371.
- [15] M. H. M. Olsson, C. R. Sondergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK(a) Predictions, *JOURNAL OF CHEMICAL THEORY AND COMPUTATION* 7 (2011) 525–537.
- [16] D. A. Case, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, J. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Morgan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, P. A. Kollman, *Amber* 10, 2008.
- [17] A. J. DePaul, E. J. Thompson, S. S. Patel, K. Haldeman, E. J. Sorin, Equilibrium conformational dynamics in an RNA tetraloop from massively parallel molecular dynamics, *NUCLEIC ACIDS RESEARCH* 38 (2010) 4856–4867.
- [18] E. Sorin, V. Pande, Exploring the helix-coil transition via all-atom equilibrium ensemble simulations, *BIOPHYSICAL JOURNAL* 88 (2005) 2472–2493.
- [19] M. Mahoney, W. Jorgensen, A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential

- functions, *JOURNAL OF CHEMICAL PHYSICS* 112 (2000) 8910–8922.
- [20] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, *JOURNAL OF CHEMICAL THEORY AND COMPUTATION* 4 (2008) 435–447.
- [21] U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, L. Pedersen, A smooth particle mesh ewald method, *JOURNAL OF CHEMICAL PHYSICS* 103 (1995) 8577–8593.
- [22] B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, *JOURNAL OF CHEMICAL THEORY AND COMPUTATION* 4 (2008) 116–122.
- [23] G. Bussi, T. Zykova-Timan, M. Parrinello, Isothermal-isobaric molecular dynamics using stochastic velocity rescaling, *JOURNAL OF CHEMICAL PHYSICS* 130 (2009) 074101 (9 pp.).
- [24] M. Parrinello, A. Rahman, Polymeric transitions in single-crystals - A new molecular-dynamics method, *JOURNAL OF APPLIED PHYSICS* 52 (1981) 7182–7190.
- [25] S. Dapprich, I. Komaromi, K. Byun, K. Morokuma, M. Frisch, A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *JOURNAL OF MOLECULAR STRUCTURE-THEOCHEM* 461 (1999) 1–21.
- [26] T. Vreven, K. Morokuma, Investigation of the S-0 to S-1 excitation in bacteriorhodopsin with the ONIOM(MO : MM) hybrid method, *THEORETICAL CHEMISTRY ACCOUNTS* 109 (2003) 125–132. Symposium on Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials held at the 222nd American-Chemical-Society National Meeting, Chicago ,Illinois AUG 26-30, 2001.
- [27] M. Svensson, S. Humbel, R. Froese, T. Matsubara, S. Sieber, K. Morokuma, ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for

- Diels-Alder reactions and $\text{Pt}(\text{P}(\text{t-Bu})(3))(2)+\text{H}_2$ oxidative addition, *JOURNAL OF PHYSICAL CHEMISTRY* 100 (1996) 19357–19363.
- [28] A. D. Becke, Density-functional thermochemistry .3. The role of exact exchange, *JOURNAL OF CHEMICAL PHYSICS* 98 (1993) 5648–5652.
- [29] N. Godbout, D. R. Salahub, J. Andzelm, E. Wimmer, Optimization of gaussian-type basis-sets for local spin-density functional calculations .1. Boron through neon, optimization technique and validation, *CANADIAN JOURNAL OF CHEMISTRY-REVUE CANADIENNE DE CHIMIE* 70 (1992) 560–571.
- [30] C. Sosa, J. Andzelm, B. C. Elkin, E. Wimmer, K. D. Dobbs, D. A. Dixon, A local density functional-study of the structure and vibrational frequencies of molecular transition-metal compounds, *JOURNAL OF PHYSICAL CHEMISTRY* 96 (1992) 6630–6636.
- [31] F. R. Clemente, T. Vreven, M. J. Frisch, Chapter 2. getting the most out of oniom: Guidelines and pitfalls, *Quantum Biochemistry* (2010) 63.
- [32] M. Torrent, T. Vreven, D. Musaev, K. Morokuma, O. Farkas, H. Schlegel, Effects of the protein environment on the structure and energetics of active sites of metalloenzymes. ONIOM study of methane monooxygenase and ribonucleotide reductase, *JOURNAL OF THE AMERICAN CHEMICAL SOCIETY* 124 (2002) 192–193.
- [33] T. Vreven, K. Morokuma, On the application of the IMOMO (integrated molecular orbital plus molecular orbital) method, *JOURNAL OF COMPUTATIONAL CHEMISTRY* 21 (2000) 1419–1432.
- [34] T. Vreven, M. Frisch, K. Kudin, H. Schlegel, K. Morokuma, Geometry optimization with QM/MM methods II: Explicit quadratic coupling, *MOLECULAR PHYSICS* 104 (2006) 701–714.
- [35] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima,

- Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian 09, Revision A.1, 2009. Gaussian, Inc., Wallingford, CT.
- [36] P. Tao, H. B. Schlegel, Software News and Updates A Toolkit to Assist ONIOM Calculations, *JOURNAL OF COMPUTATIONAL CHEMISTRY* 31 (2010) 2363–2369.
- [37] L. Wang, S. Broyde, Y. Zhang, Polymerase-Tailored Variations in the Water-Mediated and Substrate-Assisted Mechanism for Nucleotidyl Transfer: Insights from a Study of T7 DNA Polymerase, *JOURNAL OF MOLECULAR BIOLOGY* 389 (2009) 787–796.
- [38] J. Seckler, K. Howard, M. Barkley, P. Wintrod, Solution Structural Dynamics of HIV-1 Reverse Transcriptase Heterodimer, *BIOCHEMISTRY* 48 (2009) 7646.
- [39] A. Hizi, R. Tal, M. Shaharabany, S. Loya, Catalytic properties of the reverse transcriptases of human immunodeficiency viruses type-1 and type-2, *JOURNAL OF BIOLOGICAL CHEMISTRY* 266 (1991) 6230–6239.

Figure Legends

Figure 1.

Schematic representation of the reaction and active site of HIV-RT. Nucleophilic attack on the-phosphorus atom of the ribonucleoside triphosphate by the primer 3'-hydroxyl leads to the formation of a phosphodiester bond and release of pyrophosphate. Catalytic Mg^{2+} cations and their coordination shell are represented. Parts not present in the active site of the crystal structure but which occur after modeling (3'-hydroxyl) and MD-simulations (water and lys220) are colored blue.

Figure 2.

Position of LYS220 at different times with respect to the active site in HIV-RT. With 0 ns the crystal structure. Adenine triphosphate (dark blue) and Mg^{2+} of the active site are shown together with the structural domain containing LYS220.

Figure 3.

The QM region used for the QM/MM optimizations. Red are the link atoms. Blue parts are elements of the active site that only appear after modeling (3' hydroxyl) or molecular dynamics simulations (H_2O and LYS220).

Figure 4.

Time evolution of the interaction between LYS220 and triphosphate and their relative position: the center of mass distance between the NH_3 head group of LYS220 and triphosphate (A), labels I,II,III correspond to the structures shown in Figure 5., the number of direct hydrogen bonds (B) and the number of water mediated hydrogen bonds (C).

Figure 5.

The different modes of interaction between triphosphate and LYS220 after QM/MM optimization. Structure I shows the tight interaction with α and γ phosphate, II shows the state with water mediated interaction, and III the state with interaction with the terminal phosphate.

Figure 6.

The polymerase domain of HIV-RT of crystal structure 1RTD, including DNA (green) and thymidine triphosphate substrate (magenta). 134 other crystal structures were superimposed via a least squares fit onto this structure

using protein backbone atoms of residues 205 to 210 and 228 to 233. The loops containing residues 210 to 228 of the 134 structures superimposed onto 1RTD are shown.

ACCEPTED MANUSCRIPT

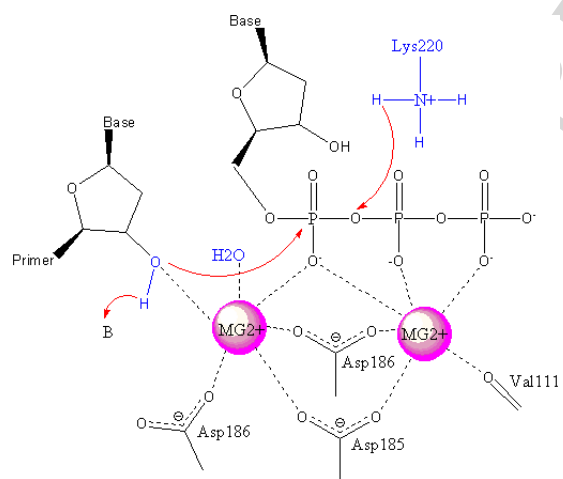


Figure 1:

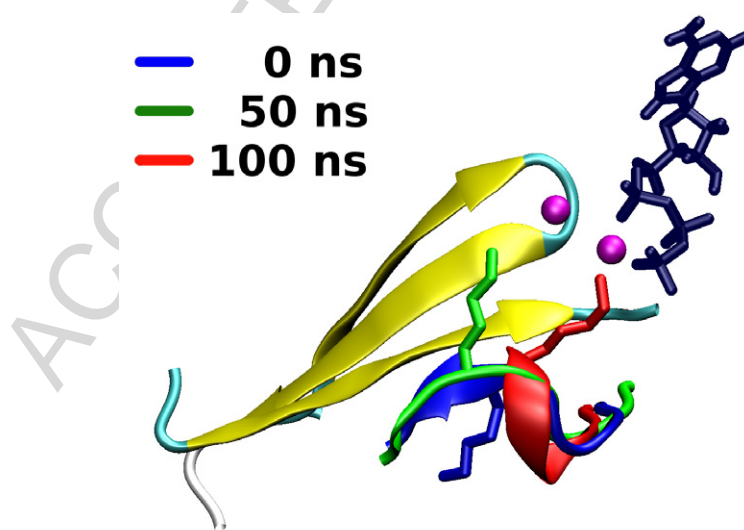


Figure 2:

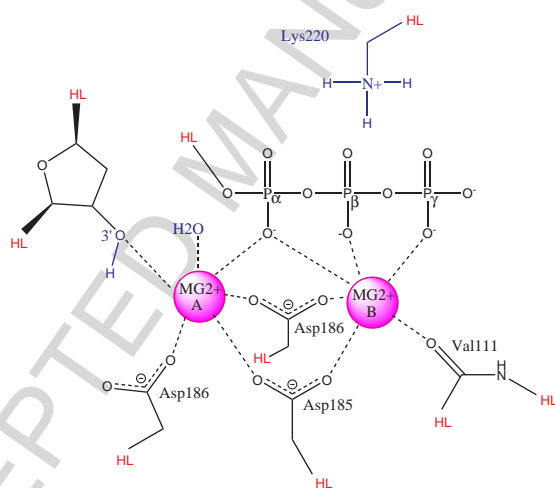


Figure 3:

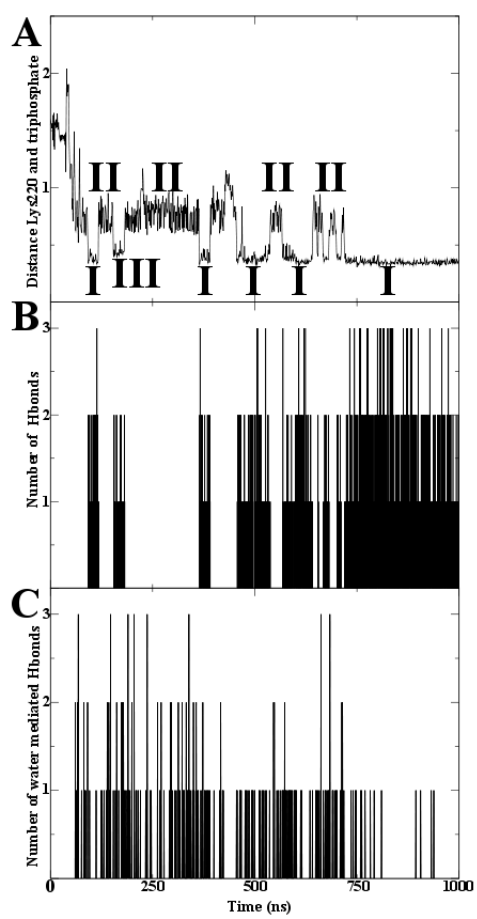


Figure 4:

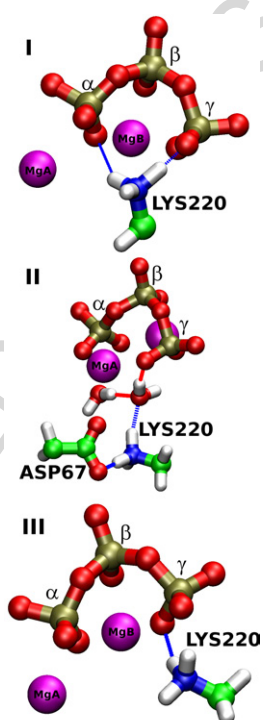


Figure 5:

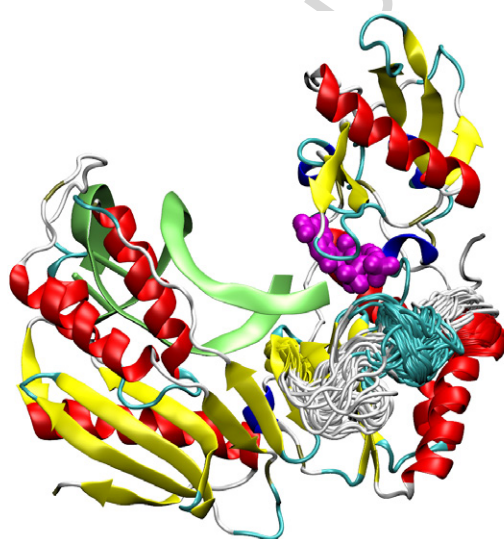
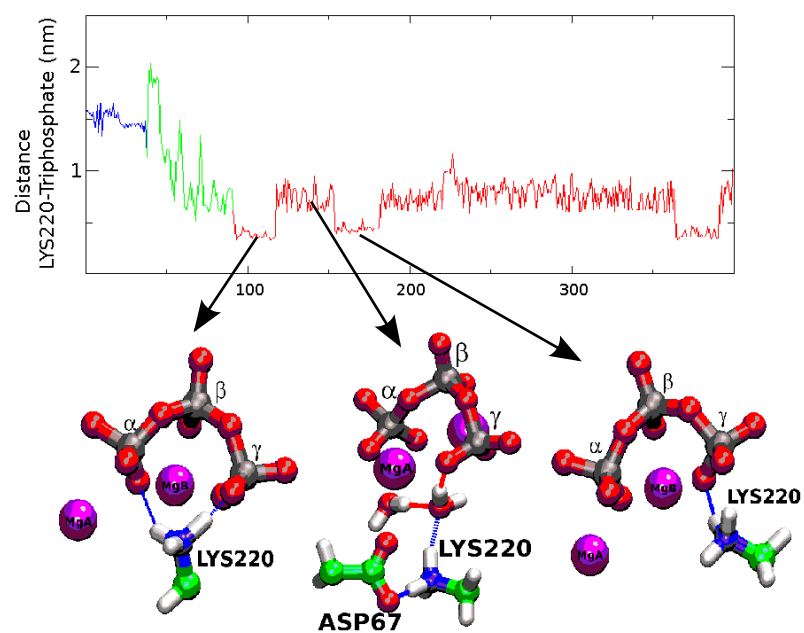
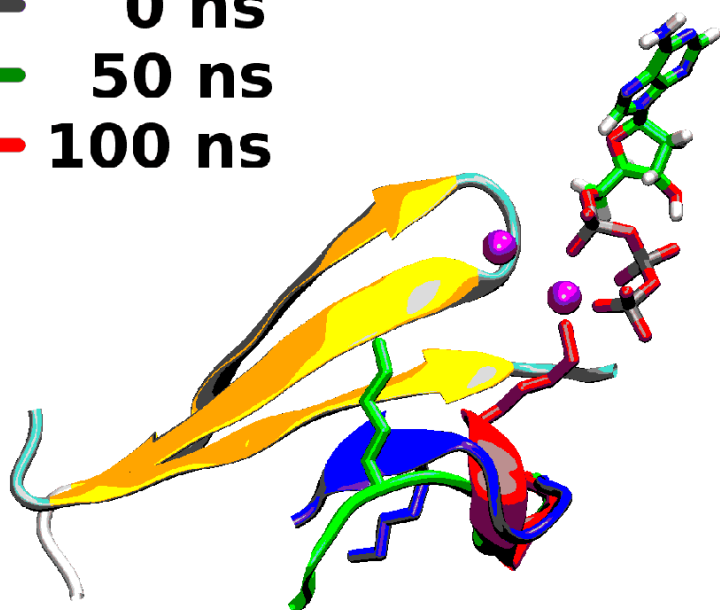


Figure 6:

*Graphical Abstract

— 0 ns
— 50 ns
— 100 ns



***Research Highlights**

A structural explanation for the catalytic role of Lys220 in HIV-RT is offered using molecular dynamics simulations. QM/MM optimizations are used to assess the quality of the structures found on a quantum mechanical potential. The results are compared to the ensemble of available crystal structures of this protein.

ACCEPTED MANUSCRIPT