

Molecular Insights into the Substrate-Assisted Mechanism of Viral DNA 3'-End Processing in Intasome of Prototype Foamy Virus Integrase from Molecular Dynamic and QM/MM Studies.

Agnieszka Krzemińska,¹ and Katarzyna Świderek^{2,*}

¹ Institute of Applied Radiation Chemistry, Lodz University of Technology, 90-924, Lodz, Poland

² Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain.

corresponding author:

Katarzyna Świderek,

e-mail: swiderek@uji.es

KEYWORDS: prototype foamy virus, integrase, QM/MM, AMBER, DNA 3'-end processing, Free energy perturbation, Density functional theory

ABSTRACT Integrases participates in two important steps for virus replication such as 3'-end processing of viral DNA ($vDNA$) and nuclear entry of host DNA ($hDNA$). In this work insight into the structural changes in intasome of prototype foamy virus integrase (PFV-IN) complexed with $vDNA$ from classical molecular dynamic (MD) simulations are done. Analysis of the results reveal the existence of alternative conformations of the enzyme active site indicating that 3'-end processing reaction can occur according to three different pathways, and taking place with the possible participation of aspartate 185, of a neighbouring phosphate group, or involving internal phosphate group of the substrate. In this work one of them, the so-called '*substrate-assisted*' mechanism was explored, by QM/MM methods. The free energy barriers of $34.4 \text{ kcal}\cdot\text{mol}^{-1}$ for the first and $35.3 \text{ kcal}\cdot\text{mol}^{-1}$ for the second step of reaction computed with free energy perturbation (FEP) methods at M06-2X/AMBER level, show that 3'-end processing has to proceed *via* a different mechanism than studied herein. Nevertheless, the obtained results are in good agreement with the experimental observations, that the substitution of the key atom for this mechanism, oxygen by sulphur, did not influence the catalysis. Additionally, the obtained mechanism reveals significant similarities to the previously studied '*substrate-assisted*' mechanism in twister ribozyme. The possible role of Mg^{2+} in the active site is discussed.

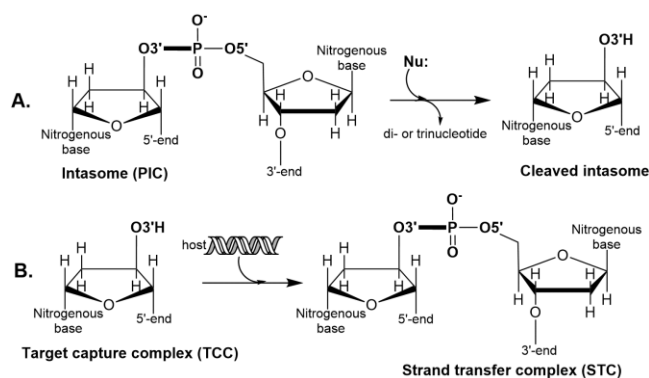
Retroviruses belong to a large and diverse family of RNA viruses defined by common taxonomic denominators that include structure, composition, and replicative properties.¹ Thus, one of the common features of this family is its replicative strategy which includes, as essential steps, the reverse transcription of the viral RNA (*v*RNA) into linear double-stranded DNA (*v*DNA) and the subsequent integration of this *v*DNA into the genome of the host cell (*h*DNA).^{2,3} Importantly, *v*DNA:*h*DNA integration is in fact the process which distinguishes retroviruses from all other viral families.

Retroviral integrases (INs) catalyze the insertion of the viral genome into the host's DNA, where the virus is able to persist in the cell indefinitely. Their role was identified as the proteins participating in two important steps for virus replication, such as 3'-end processing of *v*DNA and nuclear entry of *h*DNA. During 3'-end processing IN hydrolases a phosphodiester bond at *v*DNA end, and a di- or trinucleotide is removed, releasing 3'-hydroxyl group attached originally in the preintegration complex (PIC) to 5'-CA-3' dinucleotide.

Consequently, such cleaved intasome binds host chromosomal *h*DNA, and the target capture complex (TCC) is formed. Within the TCC, the enzyme utilizes *v*DNA 3'-hydroxyl group as nucleophile to cut *h*DNA, and joining 3' *v*DNA end to the strand of *h*DNA resulting in the strand transfer complex (STC).^{4,5} Representation of this process is presented on Scheme 1.

The first functional retroviral IN-DNA complex was characterized for intasome from the prototype foamy virus (PFV), a member of *Spumaretrovirinae* subfamily.⁵ The PFV intasome contains a tetramer of IN with a dimer-of-dimers structure, composed of two structurally and functionally distinct IN subunits. The inner one is responsible for binding and interacting with *v*DNA, as shown in Figure 1. IN binds a pair of divalent metal cations (Mg^{2+} or Mn^{2+}) which are coordinated by three carboxylates of the invariant D,DX₃5E motif within the catalytic core

domain (CCD). In order to function, IN also requires its amino-terminal domain (NTD), a three-helical bundle stabilized by binding a Zn atom, and a carboxy-terminal domain (CTD).^{6,7}



Scheme 1. Reaction of **A.** 3'-processing reaction and **B.** strand transfer occurring in PFV-IN.

The interest in understanding the replication mechanism of retroviruses increased meaningfully after the discovery of human immunodeficiency virus type 1 (HIV1).^{8,9} Despite the great progress in HIV therapy, still more than 36.9 million people worldwide are infected.¹⁰ The enormous effort has been put into the discovery of new efficient antivirals but, so far, only three such as raltegravir, dolutegravir, and elvitegravir IN strand transfer inhibitors (INSTIs) have been approved by FDA.¹¹ Thus, understanding the process of the viral replication at molecular level is crucial for describing its chemical functions and, in consequence, to reduce or completely eliminate the problem of HIV infection.

Due to the high degree of amino acid conservation within the IN active site, it is believed that knowledge obtained from the studies based on PFV-IN should be generally useful in understanding HIV1-IN, and could be expanded to the other retroviral integrases. In fact, as already demonstrated, the approach in which PFV served as a model to study inhibitors for HIV1 succeeded.^{12,13} Nevertheless, although a few clinically useful inhibitors of HIV-IN have been developed, its mechanism of action at molecular level remains unknown.

In contrast to PFV, crystallization process of HIV1-IN seem to be difficult and just recently first crystal structure of STC form was obtained.¹⁴ Unfortunately, this structure is still not sufficient (does not provide enough crystallographic data) to serve as the proper model for studies on the 3'-end-processing mechanism. Herein, we focus on conformational analysis based on MD simulations on the intasome PFV-IN, including 3'-end processing reaction; the first step of overall ν DNA: h DNA integration process. Although, previous computational studies of this step provided important insights into the reaction mechanism,^{15,16,17} additional studies are required to fully understand the mechanism of proton transfer in 3'-processing step.

The first theoretical approach to understand the reaction mechanism catalysed by PFV-IN was done by Bernardi *et al.*¹⁵ in which a cluster model of the IN active site composed of three catalytic residues (deprotonated Asp, Asp, and Glu) terminated by methyl groups, one Mg^{2+} cation and two water molecules were used. According to obtained then results, the reaction mechanism occurs via a S_N2 nucleophilic substitution, where the nucleophilic water activated by aspartate residue attacks on phosphorous atom and proton from other water molecule is simultaneously transferred to the 3' oxygen atom of scissile bond. This mechanism assumes 'single-shot' product formation with a potential energy barrier equal to $33.5 \text{ kcal}\cdot\text{mol}^{-1}$ obtained at B3LYP/6-31+G(d,p) level.

A few years later, Ruiz-Pernia *et al.* used core domain IN model composed of a protein fragment and the active site with a reconstruction of three missing residues.¹⁶ Using a QM/MM approach, a free energy barrier for the rate-limiting step of $22.6 \text{ kcal}\cdot\text{mol}^{-1}$ was computed at PM3:MP2/OPLS-AA level of theory. Two Asp residues, two water molecules, one Mg^{2+} cation and the substrate fragment of 3',5'-phosphodiester alkyl terminated by two methyl groups were treated quantum mechanically. In this case, the step-wise mechanism was explored, where in the

first step the nucleophilic water is activated by Asp residue and it attacks on the phosphorous atom. In the second step, the proton is transferred from another water molecule to 3' oxygen and phosphodiester bond breaks.

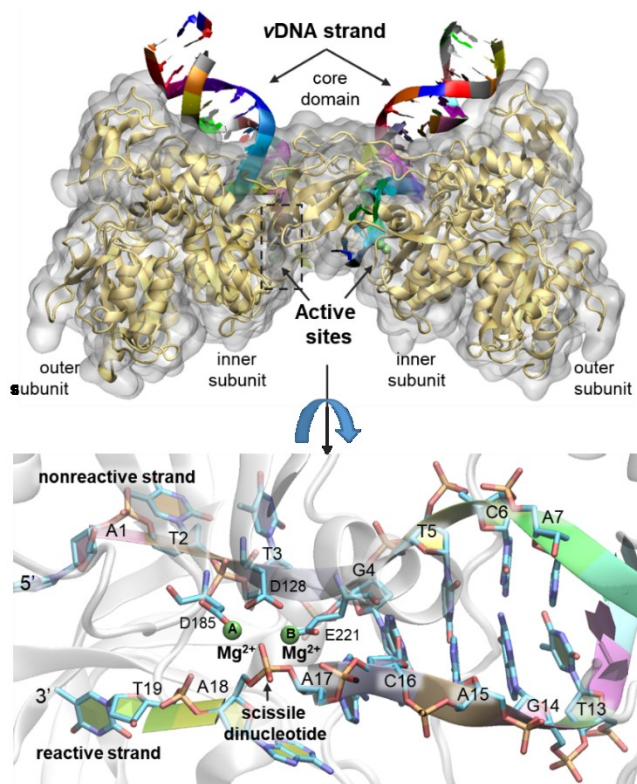


Figure 1. Crystallographic structure of biological assembly of the PFV-IN intasome (build based on structure from Protein Data Bank with PDB ID: 4E7I)²⁰ (top) together with the structure of bound vDNA and the active site (bottom).

Finally, Ribeiro *et al.* explored three different mechanisms of 3'-end processing.¹⁷ Due to the lack of crystallographic structure, an IN model was build based on fragments of PFV-IN with docked vDNA substrate into it using GOLD¹⁸ software. Based on ONIOM calculations it was concluded that the concerted mechanism activated by the presence of hydroxyl ion in the active site was the most favourable one with a potential energy barrier of 15.4 kcal·mol⁻¹ computed at MPWB1K/6-311++G(2d,2p)/AMBER level. In the hydroxyl mechanism, the water molecule is

activated by a neighbouring hydroxide group and it attacks the phosphorous atom with simultaneous phosphodiester bond breaking.

The reaction mechanism catalysed by PFV-IN has been also explored by indirect comparisons with the reaction catalysed by the retroviral ribonuclease (RNase H) of HIV-1 Reverse Transcriptase (HIV1-RT). The IN catalytic site and RNase H RT active site share similar structural features, both containing a D,DX₃5E motif (two Asp and one Glu residues) with bounded two divalent metal cofactors (Mg²⁺ or Mn²⁺).¹⁹ However, there is no scientific evidence that this comparison can be significant and conclusive about the reaction mechanism. Despite close similarities, the RNase H and IN systems likely differ in some details of their catalytic mechanism.²⁰

As it can be noticed, all previous theoretical studies were done based on the theoretical models with some serious limitations originating in the lack of full structural information. In the present work, the first theoretical studies on the complete structure of PFV-IN intasome were carried out. Firstly, the insight into the structural changes in the active site from classical molecular dynamic (MD) simulations were done. Subsequently, the proposed mechanism of 3'-end processing reaction catalysed by IN has been explored using QM/MM approach.

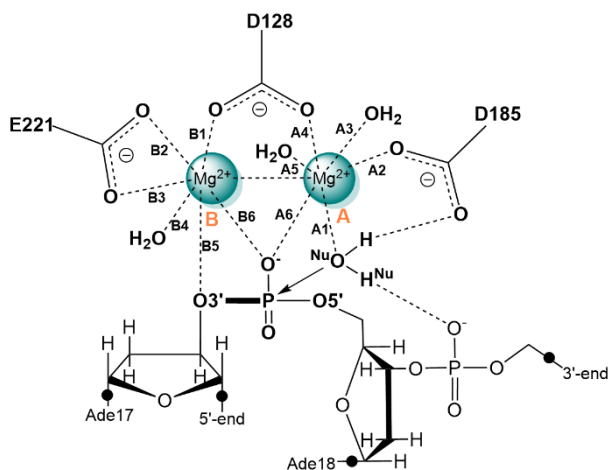
Computational methods

Setup of the system. The model of the PFV-IN was prepared based on the crystal structure obtained as a trapped Michaelis complex formed prior to 3'-processing and deposited in the Protein Data Bank with PDB ID 4E7I.²⁰ The initial crystallographic structure was a dimer with bound vDNA double helix. Two Mn²⁺ cations present in the active site were replaced by two Mg²⁺ and four water molecules from their coordination shell were preserved. The correctness of

the side-chain atoms was confirmed using SCIt tool.²¹ Due to the fact that only tetramer form of the IN is biologically active, the full tetramer structure was build using the Protein Interfaces, Surfaces and Assemblies' (PISA)²² software utilizing rotation and translation matrices available in crystallographic data. Such prepared model was composed of four polypeptide chains: two of them consisted of 367 and two of 183 amino acids with bounded two double helices of vDNA, each made of 19 base-pairs, and four Mg²⁺ cations with eight coordinated water molecules. The pK_a values for titratable amino acids were calculated using PROPKA ver.3.0.^{23,24} Considering a pH=7, and based on the obtained results all residues involved in coordination of divalent metals (two Asp and one Glu) were kept deprotonated, and the Asp13 from chain C and chain D were protonated. Sixteen of overall 30 histidines residues present in the model were protonated at ϵ position, while the remaining residues at δ position, based on the geometrical analysis of their closest environment. All computed pK_a values are available in Supporting Information (Table S1). Moreover, two cysteine bridges were detected, both between Cys88-Cys91 in Chain A and Chain B. Hydrogen atoms were added using the tLEAP module of AMBER program.²⁵ Neutralization of the system was completed in two steps, first by neutralizing the charge of DNA chains by adding 64 Na⁺, and subsequently neutralizing protein by adding 40 Cl⁻ couterions. Finally, such prepared complex was placed in the orthorhombic box of TIP3P²⁶ explicit water molecules of size 14.4x14.4x17.6 nm³. The final model contained 190 251 atoms.

Molecular dynamic (MD) simulations. Several optimizations and dynamics simulations (MDs) were performed at MM level applying AMBER ff03²⁷ force field (FF) to describe amino acids, the refined FF for DNA, Parmbsc1,²⁸ and newly developed parameters to improve the kinetic description of the Mg²⁺ and phosphate ions and their applicability in nucleic acid simulation,²⁹ as implemented in NAMD software.³⁰ Periodic boundary conditions (PBC) using the particle mesh

Ewald method were applied. The cut-off for nonbonding interactions was used including a smooth switching function with a radius range from 14.5 to 16 Å. First, the energy optimization was carried out by means of the conjugate gradient algorithm with gradient tolerance equal to 0.1 kcal·mol⁻¹·Å⁻¹. Subsequently, the system was heated from 0 to 300 K with temperature increment of 0.001 K. After 1 ns of the equilibration, 60 ns of NVT MD simulation was done with the time step of 1 fs at 300 K. The constant temperature during the MD simulation was controlled using the Langevin thermostat.³¹ The MD results were analysed using the cpptraj^{32,33} and VMD³⁴ software. The main goal of the MD simulations was to generate a relaxed system that will be used as a starting point for the subsequent QM/MM calculations and analysis.



Scheme 2. The PFV-IN active site with the non-processed 3'-viral DNA end with labelled important distances. All atoms presented on the scheme are treated at QM level of theory during the QM/MM calculations. Seven link atoms were used in order to saturate cut bond, three of them were places between C α and C β atom from the side chain of D128, D185 and E221 and the remaining 4 are indicated as thick black circles on the scheme.

QM/MM calculations. Subsequently, the structure from the 60 ns of the MM MD was selected to build up the systems for QM/MM calculations. The QM region was composed of 78 atoms (as presented in Scheme 2) and treated quantum mechanically at semiempirical AM1³⁵ level as implemented in Mopac2009³⁶ and at density functional theory (DFT) using B3LYP³⁷, MPW1PW91³⁸ and M06-2X^{39,40} functionals with 6-31+G(d,p) basis set as available in Gaussian09.⁴¹ The remaining part of the protein, vDNA chains, counterions and water molecules were described by AMBER and TIP3P FFs as implemented in fDynamo package.^{42,43,44} All atoms beyond 20 Å from the QM part were kept frozen.

Potential energy surfaces (PESs). In order to explore the proposed mechanisms, potential energy surfaces (PESs) at AM1, semiempirical level combined with MM FF were computed. Structures generated in this process were then used to compute potential energy at higher level of theory (M06-2X/MM). Subsequently, based on chosen structures, a micro-macro iteration optimization algorithm^{45,46} at DFT/6-31+G(d,p)/MM level was used to localize, optimize, and characterize the TS structures using a Hessian matrix containing all the coordinates of the QM subsystem, whereas the gradient norm of the remaining movable atoms was maintained less than 0.01 kcal·mol⁻¹·Å⁻¹. By following the gradient that was specified by the transition vector in the TSs, the paths of the chemical reaction were traced by using intrinsic reaction coordinates (IRCs)⁴⁷ to the valleys of the reactants and products in mass-weighted Cartesian coordinates. Subsequently, last structures from IRC were used to localize, optimize, and characterize the minima energy structures i.e. reactant, intermediate and product complexes.

Free energy calculations. In order to describe the mechanism of the reaction in condensed media, a free energy perturbation (FEP)⁴⁸ was used employing the M06-2X functional to

describe the QM sub-set of atoms. In this studies FEP employment originates in the limitations of semiempirical methods (SM), as it was shown by Otyepka and co-workers.⁴⁹

Since FEP required the sampling of the environment along a previously traced IRC from TS located at QM/MM level, the free energy profile is obtained along a realistic reaction coordinate. FEP method opens the possibility of exploring the reaction path directly at high level of theory. Therefore, the QM wave function is polarized by the charges of the MM subset of atoms.

The structures for which the free energy differences were estimated using the FEP methodology correspond to those obtained along the IRC and are characterized by a single s coordinate:

$$s_i = s_{i-1} + \sqrt{\sum_{j \in \text{QM}} m_j \left((x_{j,i} - x_{j,i-1})^2 + (y_{j,i} - y_{j,i-1})^2 + (z_{j,i} - z_{j,i-1})^2 \right)}$$

where $x_{j,i}$, $y_{j,i}$ and $z_{j,i}$ are the coordinates of the i th structure for the j th QM atom belonging to the IRC traced from the transition state structure (x_{j0} , y_{j0} and z_{j0} coordinates) and m_j are the corresponding masses of the atoms. Within this treatment, the free energy relative to the reactant can be expressed as a function of the s coordinate as:

$$\Delta G_{\text{FEP}}(s^R - s^j) = \left(E_{\text{QM}}^0(s^j) - E_{\text{QM}}^0(s^R) \right) + \left(\text{ZPE}(s^j) - \text{ZPE}(s^R) \right) - k_B T \sum_{i=R}^j \ln \left\langle \exp \left(- \frac{E_{\text{QM/MM}}(s^{i+1}) - E_{\text{QM/MM}}(s^i)}{k_B T} \right) \right\rangle_{\text{MM},i}$$

where E_{QM}^0 is the gas-phase energy of the QM subsystem computed at M06-2X level, k_B is the Boltzmann constant, and T is the temperature. The contribution of the QM/MM interaction to the free energy difference between two different values of s is obtained by averaging the QM/MM interaction energy (including the polarization energy) over all the MM coordinates of the system obtained during the MD simulation carried out for the i th window. Then, the coordinates of the

QM atoms are those corresponding to the s_i and s_{i+1} structures along the IRC. Thus, it can be assumed that the evolution of the chemical system on the FES follows a path that is very close to the minimum energy path on the PES. The MDs for the FEP calculation were performed at 300 K, using the NVT ensemble. 20 ps of production, with a time step of 1 fs, were done in each window. The total amount of windows required to generate the full free energy path was 51 and 43 for the first and second step of the reaction, respectively.

Results and discussion

The MD simulations. In order to understand the key interactions between the enzyme and ν DNA in the PFV-IN intasome, and especially the interactions established in the active site, a long classical MD simulation was done. The total energy of the system and its temperature were monitored. As presented in Figure S1 of Supporting Information the temperature was remaining constant at 300 K with a standard deviation of ± 1 K along MD simulations. The stability of the protein backbone (C-C α -N atoms) was analysed by measuring the root-mean-square deviations (RMSD) with respect to the optimized crystal structure, as presented in Figure S1. The RMSD values changed from 0 to around 5 Å. According to the evolution of RMSD, the system can be considered as equilibrated after 20 ns of MD simulations. Observed RMSD changes are small enough to ensure the stability of the structure, and thus such structure can be used in further studies. Moreover, the total energy also becomes stable after 20 ns of MD confirming the equilibration of the system.

Active site interactions. MD simulations have revealed a very conserved structure of the active site, which is mainly dictated by the presence of two Mg²⁺ metal ions. As shown in the Table S2 of Supporting Information, distances which correspond to the first coordination shell of both Mg²⁺ ions did not change meaningfully along MD simulations. In fact, the highest value of the standard deviation was 0.1 Å.

Interestingly, as it can be observed in Figure 2 the distances between the divalent metals and their ligands were slightly elongated in case of optimized structures with QM/MM method compared to the MM calculations. The average distance between each metal ion and its ligands obtained at MM level of 1.97 and 1.94 Å for Mg²⁺(A) and (B), respectively, is shorter than the

value obtained from QM/MM optimizations where they vary, depending on the QM method, between 2.06 and 2.15 Å for $\text{Mg}^{2+}(\text{A})$, and 2.03 and 2.12 Å in case of $\text{Mg}^{2+}(\text{B})$. This is because the accumulated charge on these two metals, as shown in Table S2 is not exactly the same as the values assumed in the classical force field ($+2e^-$). In fact calculation of atomic charges at DFT/MM level reveals an increment up to $2.111 e^-$ and a substantial reduction up to $1.559 e^-$ on ion (A) and ion (B), respectively.

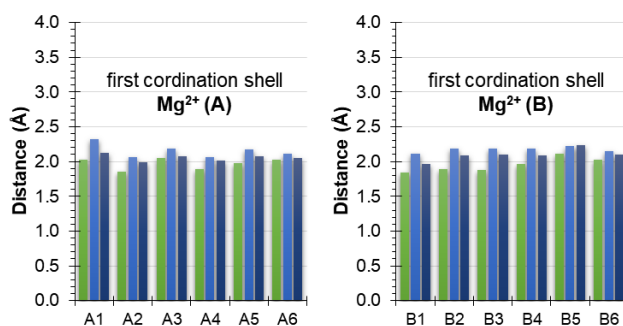


Figure 2. Distances of first coordination shell of Mg^{2+} ions bound in the active site of PFV-IN obtained for optimized structures after 60 ns of MD simulations at different levels of theory, including MM shown in green (AMBER FF), and QM/MM (AM1/MM in light blue, and M06-2X/MM in dark blue) methods.

During the MD simulation an important observation was done related to the position of the active site water molecule. It must bear in mind that it is believed that this water plays an essential role in the catalysis by acting as a nucleophile (see Scheme 2). The position of this water is preserved with respect to the phosphorus atom with a distance oscillating around 3.23 ± 0.13 Å as shown in Figure 3.A. The reason for the unchangeable position of this water can be explained by the fact that it is coordinated to the Mg^{2+} cation and any change in its position would be energetically unfavorable. In fact, Figure 3.B. shows very conserved ‘in-line’ position

of this water, dictated by the coordination to the Mg^{2+} (A) that results in a very high population (47%) of structures fulfilling the condition of the so-called reactive ‘in-line’ conformation.⁵⁰ This is a very interesting behavior by comparison to other systems, (that also require ‘in-line’ conformation between nucleophile, phosphate atom and leaving group) such as hairpin ribozyme previously studied by York and co-workers[50] or twister ribozyme,⁵¹ and ANT4,⁵² studied in our group. In the previous models reactive conformations appear much less often along MD simulations with the calculated probabilities not higher than 7.7%.

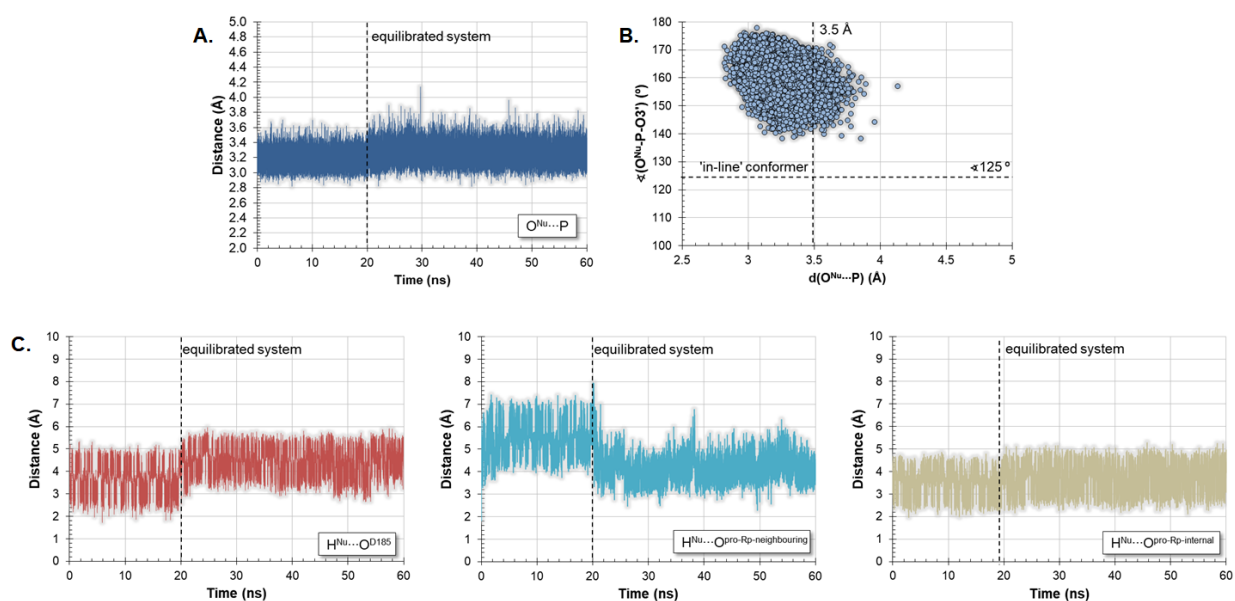


Figure 3. **A.** Evolution of $O^{Nu}\dots P$ distance, **B.** Reactive ‘in-line’ conformations as a function of a distance between nucleophilic water and phosphorus atom $d(O^{Nu}\dots P)$ and an angle between the same nucleophilic water, phosphorus atom and oxygen of the leaving group $\alpha(O^{Nu}\dots P-O3')$ (to specify amount of the reactive conformations the following boundary conditions were assigned: distance not longer than 3.5 Å, and angle not smaller than 125°). and **C.** Evolution of distances established between hydrogen atom of water molecule, H^{Nu} and its possible acceptors O^{pro-Rp} , $OD2^{D185}$ and O^{pro-Rp} of neighbouring phosphate group, along 60ns of MD simulations.

The first step of the full enzymatic process catalysed by PFV-IN entails the P-O3' breaking between A17 and A18 nucleotides in the reactive strand of the ν DNA. This reaction requires at least one proton transfer onto the 3'-aloxide group, in order to generate the 3'-hydroxyl leaving group. Structural analysis of the active site presented above indicated, similarly as in previous MD studies done by De Vivo *et al.*,⁵³ and Rosta *et al.*¹⁹ for RNase H system, that the water coordinated to Mg²⁺ (metal A) oriented in the 'in-line' position to the scissile phosphodiester is the best candidate to play the role of nucleophile. Thus, in such case additional proton transfer from the water that will generate active hydroxide anion has to be included in the mechanism studies. Unfortunately, the transfer of these two protons, and more precisely the determination of their donor and acceptor, is the source of many speculations. Thus, based on the results from the MD simulations we performed a search of possible candidates which could accept the proton from the water. Their identification can provide new ideas about possible scenarios of the 3'-end processing mechanism.

Three important interactions between the nucleophilic water and the active site residues were observed, as shown in Scheme 2 and Figure 3.C. First, the existence of an important interaction between the hydrogen atom of water and the oxygen atom of D185 was recognized, with an original value of distance established at 3.5 Å. However, after 20 ns of MD simulations, as shown in Figure 3.C, this interaction became weaker since a meaningful increase in the distance between proton and D185 was observed. In the last 40 ns, the distance between H^{Nu} and OD2^{D185} oscillates around 4.43 ± 0.64 Å. Hence, it is difficult to determine the significance of D185 for the catalysis. It could be assumed that the transfer of the proton to the oxygen atom of D185 could be mediated by an additional water molecule, as it was assumed in previous

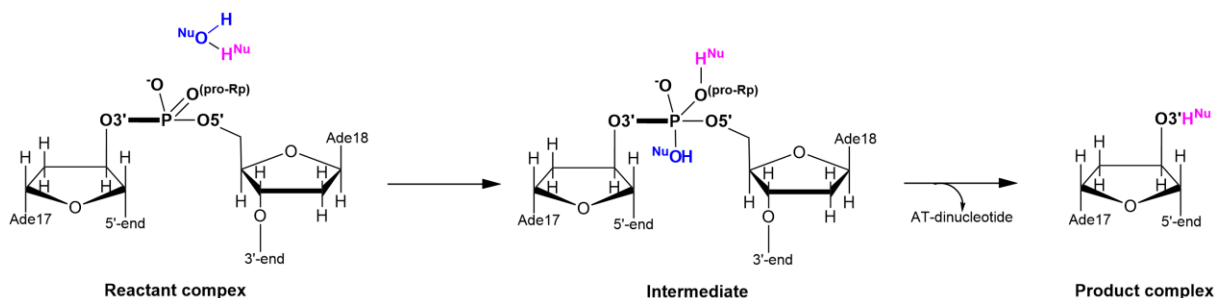
studies.¹⁷ Nevertheless, as shown, the energetic barrier obtained using ONIOM approach for such mechanism was too high (35.8 kcal·mol⁻¹).

Alternative option for a proton acceptor is the pro-*R_P* oxygen atom of neighbouring phosphate group located between A18 and T19 nucleotide. The neighbouring phosphate group is definitely beyond the reach of the nucleophile with a distance of 5.5 Å, at the beginning of the MD simulation and cannot be responsible for the water activation. Interestingly, after 20 ns of the classical simulations, this distance is reduced to 4.18 ± 0.54 Å, as shown in Figure 3.C. Thus, this change should be taken into consideration and the possible role of phosphate group cannot be excluded. Similarly, as in case of D185, the transfer of the proton should involve participation of an additional water molecule. Direct transfer of the proton to this phosphate group was, nevertheless previously studied, and the computed high barrier of 36.3 kcal·mol⁻¹ precluded this hypothesis.¹⁷

Finally, the last identified candidate to serve as a proton acceptor is an internal phosphodiester of the scissile dinucleotide. However, analysis of the crystal structure at molecular level, made elsewhere,²⁰ suggested that phosphate group located between A17 and A18 is rather less promising candidate to act as a general base, since its non-bridging pro-*R_P* oxygen is pointing away from the attacking water. This was not confirmed during MD simulations. In fact, the shortest distance between hydrogen from the water, H^{Nu} and this oxygen (O^{pro-*R_P*}) was found. This distance was oscillating around 3.70 ± 0.60 Å. Thus, in our contribution to the mechanistic studies of 3'-end processing reaction we will focus on this last scenario.

The substrate-assisted 3'-end processing of νDNA. As discussed in the previous section, it can be assumed that mechanism of 3'-end processing reaction proceeds in such way that pro-*R_P*

oxygen atom of the phosphodiester group serves as a general base that abstracts a proton from the water molecule. Subsequently, the same pro- R_P oxygen serves as a proton donor to the 3'-alkoxine group and it is closing the catalytic cycle, as presented in Scheme 3. By assuming such course of the reaction pathway, the main role of the PFV-IN in the 3'-end processing is reduced to: firstly provide an appropriate environment for binding two divalent metals, which role is to activate a water molecule to attack the phosphorus atom, and additionally to destabilize the scissile bond; and secondly, to partially unfold the double helix of ν DNA and bind it into the proper position in the active site. Thus, in such case, direct involvement of the PFV-IN in the catalytic process has been eliminated, consequently reducing role of the enzyme to a passive spectator of the chemical changes. In such situation this chemical process will be named herein as '*substrate-assisted*' mechanism.



Scheme 3. 3'-end processing *via* '*substrate-assisted*' mechanism catalysed by PFV-IN.

At molecular level the '*substrate-assisted*' mechanism is the reaction in which the proton of the water molecule is transferred to the oxygen (pro- R_P) of phosphate group participating directly in the chemical reaction with a simultaneous water attack on the phosphorus atom (P). In the second step, the proton is transferred from the same oxygen (pro- R_P) to O3' of scissile bond and

phosphodiester bond (P-O3') simultaneously breaks, as shown in Scheme 3. In order to explore the proposed mechanism on PESs two-dimensional scans were done. The first proposed step in this mechanism was explored by controlling the distance between the nucleophilic water oxygen atom and phosphorus atom $d(\text{O}^{\text{Nu}}-\text{P})$, together with antisymmetric combination of distances between oxygen and hydrogen atoms of the nucleophilic water, and the same hydrogen and pro- R_P oxygen of phosphate group: $d(\text{O}^{\text{Nu}}-\text{H}^{\text{Nu}})-d(\text{H}^{\text{Nu}}-\text{O}^{\text{pro-}R_P})$. Subsequently, the second step of this mechanism was described using the distance between the phosphorus atom and 3'-oxygen of alkoxide group $d(\text{P}-\text{O}3')$, and the antisymmetric combination of the two distances, first defined for the same pro- R_P oxygen and hydrogen H^{Nu} of nucleophile water molecule, and second assigned between hydrogen H^{Nu} and O3' oxygen: $d(\text{O}^{\text{pro-}R_P}-\text{H}^{\text{Nu}})-d(\text{H}^{\text{Nu}}-\text{O}3')$. AM1/MM and M06-2X:AM1/MM PESs computed as explained in Computational methods section are shown in Figure 4. The TS structures for both steps were localized at AM1/MM, B3LYP/MM, MPW1PW91/MM and M06-2X/MM levels of theory. For clarity purpose, only key interatomic distances for structures optimized at AM1/MM and M06-2X/MM are provided in Table 1, while the results obtained using other DFT method are collected in Table S3 of Supporting Information. The representation of the stationary structures computed at M06-2X/MM level of theory are presented in Figure 5. Obtained MEP together with localized stationary points using different methods are projected on PESs.

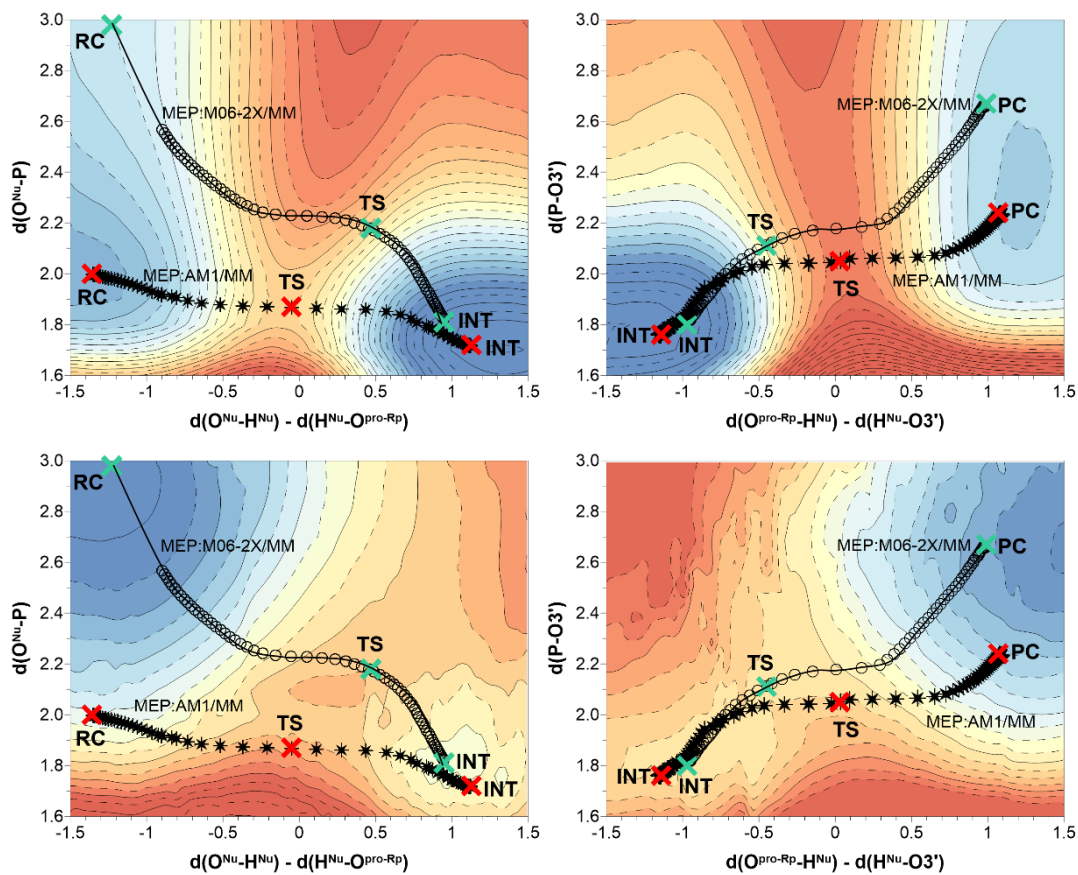


Figure 4. Potential energy surfaces for the first (left) and second steps (right) of ‘*substrate-assisted*’ mechanism catalysed by PFV-IN computed at AM1/MM (top) and M06-2X:AM1/MM (bottom) level of theory. Position of localized stationary points such as reactant (RC), intermediate (INT) and product (PC) complexes and transition states (TS) structure are indicated by red (AM1/MM) and green (M06-2X/MM) crosses, together with traced minimum energy path (MEP) obtained in IRC calculations. Distances of the axes are given in Å.

As can be noticed the PESs and the MEPs computed at AM1/MM level are significantly different from those calculated using M06-2X/MM method. Despite that both QM levels describe ‘*substrate-assisted*’ mechanism as a two-step process, meaningful differences between them can be indicated. First of all, as can be seen on all PESs (Figure 4), and confirmed by distance analysis (Table 1), the position of minima for reactant and product complexes obtained

at AM1/MM level do not overlap with those delivered from M06-2X method. Thus, it can be concluded that the main problem of the semiempirical method used in this work originates in the erroneous description of non-bonding interactions between oxygen and phosphorus atom ($O\cdots P$). Both distances for not formed bonds obtained for optimized structures of reactant and product complexes are too short, 2.00 and 2.24 Å for $O^{Nu}-P$, and $P-O3'$, respectively, compared with the values obtained at M06-2X level, i.e. 2.98 and 2.67 Å, and with those delivered from calculation employing others DFT methods. (see Table S4). Additionally, the quadratic regions on which TS structures can be localized are shifted. Thus, in TS1 process of the bond formation between oxygen from water and phosphorus atom ($O^{Nu}-P$) is much more advanced at AM1 (1.87 Å), than at M06-2X level. (2.18 Å). On the contrary, proton transfer seems to be almost accomplished at DFT level, since proton-acceptor distance in TS1 is 1.04 Å, while this is not the case at AM1 (1.33 Å). The same problem with proton transfer is observed in the case of TS2. In the second step, the transfer of the proton is delayed with respect to the $P-O3'$ bond breaking at DFT level, while at AM1 this proton is already partially transferred to 3'-alkoxide group.

Furthermore, computed potential energy barriers (ΔE^\ddagger), of 27.1 and 25.2 kcal·mol⁻¹ at AM1/MM level for first and second step, respectively, as shown in Table 2, are underestimated with respect to those delivered from calculations done at M06-2X of 35.1 and 32.5 kcal·mol⁻¹. Interestingly, AM1 method significantly over stabilizes intermediate complex, while all DFT methods indicate its metastable character.

Table 1. Key interatomic distances measured for optimized stationary structures at AM1/MM and M06-2X/MM levels of theory. Values are given in Å.

	RC	TS1	INT	TS2	PC
AM1/MM					
H ^{Nu} -O ^{Nu}	0.99	1.28	2.10	3.57	4.12
H ^{Nu} -O ^{pro-Rp}	2.35	1.33	0.97	1.27	2.06
O ^{Nu} -P	2.00	1.87	1.72	1.67	1.67
P-O3'	1.70	1.70	1.72	2.05	2.24
H ^{Nu} -O3'	4.24	3.60	3.37	1.24	0.99
M06-2X/MM					
H ^{Nu} -O ^{Nu}	0.97	1.51	1.93	3.25	3.97
H ^{Nu} -O ^{pro-Rp}	2.20	1.04	0.98	1.04	1.96
O ^{Nu} -P	2.98	2.18	1.81	1.66	1.64
P-O3'	1.62	1.64	1.72	2.11	2.67
H ^{Nu} -O3'	4.31	3.32	3.31	1.49	0.97

Calculation of Gibbs free energy (ΔG^\ddagger) barriers with FEP method at M06-2X/MM level, as shown in Figure 6, did not change significantly the values obtained at PESs, and just a slight decrease of 0.5 and increase of 2.8 kcal·mol⁻¹ was observed in first and second step of reaction, respectively. This could indicate, that in case of PFV-IN, environmental fluctuations around active site are somehow limited along the reaction pathway.

situation in which the replacement for position pro- S_P is done. Substitution of position pro- S_P oxygen by sulphur resulted in a strong inhibition. Our results confirm that pro- R_P cannot be involved in the catalytic process, and thus its modification to the less electronegative atom than oxygen would not affect the reaction mechanism.

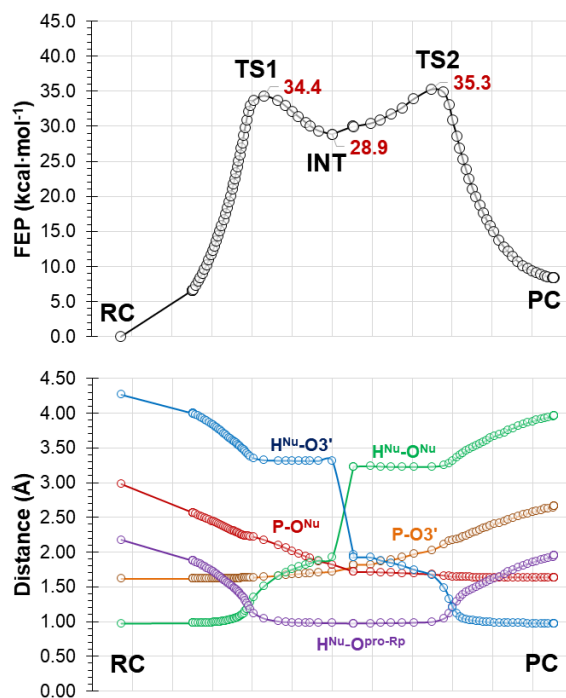


Figure 6. Free energy profile (top) and key distances evolution (bottom) along the two-step mechanism of ‘substrate-assisted’ 3’-end processing of vDNA catalysed by PFV-IN obtained with free energy perturbation method at M06-2X/MM level of theory.

The ‘substrate-assisted’ mechanism of P-O bond cleavage was previously studied in our group in the other systems such as twister ribozyme,⁵¹ or 4’-O-nucleotidyltransferase (ANT4’).⁵² In the first one, the reaction of self-cleavage RNA phosphodiester bond proceeds *via* a very similar two-steps mechanism to that described in this work, while in the ANT4’ the transfer of adenylyl group from adenosine 5’-triphosphate (ATP) to a OH-4’ hydroxyl group of the

kanamycin A proceeds *via* bimolecular nucleophilic substitution (S_N2), and the mechanism cannot be directly compared to one studied in PFV-IN. It is worthy to point out that in both previously studied cases, the same as in this one, '*substrate-assisted*' mechanism revealed existence of high energy barriers, and could not be considered as the most probable reaction mechanism. Nevertheless, comparative analysis of the results obtained herein and for '*substrate-assisted*' mechanism studied in ribozyme reveals some interesting observations. For example, the barriers obtained for the similar reaction, but catalysed by ribozyme without the presence of a metal ion in the active site, are both lower than those obtained for PFV-IN system i.e. for nucleophilic attack the free energy barrier, $\Delta G^\ddagger(TS1)$ of 20.9 kcal·mol⁻¹ (13.5 kcal·mol⁻¹ lower) and in subsequent step, where P-O5' bond is breaking, the $\Delta G^\ddagger(TS2)$ of 31.9 kcal·mol⁻¹ (3.4 kcal·mol⁻¹ lower) were computed at M06-2X/MM level. This meaningful difference in energies observed especially for the first step of reaction could be explained by the presence of divalent Mg²⁺ ions in the active site of PFV-IN, which could be responsible for reducing the base character of the phosphate group positioned directly in their first coordination shell. However, the negative charge on proton acceptor (O^{Pro-Rp}) of -0.930 e⁻, does not differ meaningfully from that observed on acceptor in case of twister ribozyme of -0.923 e⁻ (as presented in Table 3). Thus, the higher reaction barrier obtained in PFV-IN active site can be only explained based on the geometrical comparison of reactant complexes. While the distance of proton and its acceptor in ribozyme in Michaelis complex was 1.57 Å, in the case of PFV-IN the equivalent distance is much longer and equal to 2.20 Å.

Moreover, the analysis of the atomic charge distribution on key atoms along the reaction pathway in PFV-IN reveals more similarities with the '*substrate-assisted*' mechanism in twister ribozyme. For instance, the most negative charge on leaving group (O5' in case of ribozyme and

O3' in PFV-IN) was reached in TS2. As in the case of the oxygen atom of the nucleophile (O2' and O^{Nu} in case of ribozyme and PFV-IN, respectively) the highest negative charge was found in TS1.

Table 2. Relative potential (ΔE_{pot}) and Gibbs free (ΔG_{pot}) energies computed at semiempirical (AM1/MM) and DFT (M06-2X/MM) level for the stationary structures along the reaction pathway. Values are given in kcal·mol⁻¹.

	ΔE_{pot}		ΔG
	AM1	M06-2X	M06-2X
RC	0.0	0.0	0.0
TS1	27.1	35.1	34.4
INT	-5.4	28.3	28.9
TS2	25.2	32.5	35.3
PC	4.5	5.3	8.4

Table 3. Atomic charges computed using ChelpG method at M06-2X/MM level of theory.

Atoms	RC	TS1	INT	TS2	PC
O ^{Nu} :	-1.091	-1.119	-1.011	-0.674	-0.675
O ^{pro-Rp}	-0.930	-0.966	-1.003	-0.795	-0.930
O ^{pro-Sp}	-0.867	-0.783	-0.901	-0.760	-0.901
O3'	-0.507	-0.598	-0.667	-0.830	-0.625
P	1.126	1.417	1.584	1.362	1.445
H ^{Nu} :	0.561	0.623	0.568	0.397	0.344
Mg ²⁺ (A)	2.127	1.827	1.899	1.719	1.806
Mg ²⁺ (B)	1.607	1.604	1.669	1.635	1.621

Changes observed for the values of atomic charges computed on Mg^{2+} ions along the reaction path indicate that metal ions are involved in the charge transfer related to chemical changes occurring in their close surrounding. In particular, a strong effect on ion(A) was observed. In its case, the most positive charge of $2.127 e^-$, and the lowest value of $1.719 e^-$ were found in Michaelis complex and TS2, respectively. Interestingly, no significant change in the charge accumulated on the second metal (B) was perceived.

Conclusions

In this work theoretical studies of the reaction mechanism of 3'-end processing step catalysed by PFV-IN were done. A conserved water of the first coordination shell of Mg^{2+} was identified as a nucleophile for which 'in-line' conformation was found in almost half of the structures populated along 60 ns of MD simulations. Moreover, results for classical MD revealed the existence of three possible candidates to play the role of a proton acceptor from the nucleophile, i.e. aspartate 185, the neighbouring phosphate group or the internal phosphate group of the ν DNA substrate. Since first two mechanisms have been previously theoretically studied, and in both cases too high energy barriers were obtained, in this work, we have focused on the '*substrate-assisted*' mechanism, in which internal phosphodiester of the scissile dinucleotide plays the key role. Our calculations applying QM/MM methods with semiempirical (AM1), and DFT (M06-2X, B3LYP and MPW1PW91) level of theory combined with AMBER force field, showed that this mechanism can exist and proceed through a two-step process. In the first step the proton of the water molecule is transferred to the oxygen of phosphate group participating directly in the chemical reaction with a simultaneous water attack on the phosphor atom, while in the second

step, the proton is transferred from the same oxygen to O3' of scissile bond and phosphodiester bond (P-O3') simultaneously breaks.

Unfortunately, the obtained potential and free energy barriers excluded the possibility that 3'-end processing of ν DNA takes place according to this scenario. Nevertheless, the studies of this mechanism confirmed the experimental observations in which an exchange of oxygen atom ($O^{\text{pro-Rp}}$) by sulphur did not affect the catalysis. The experimental evidence together with results of this work gives unwavering certainty that this atom is not playing any significant catalytic role. Moreover, it can be assumed that at this moment the most possible reaction mechanism of 3'-end processing catalysed by IN is the one involving the hydroxide (as described in the Introduction) proposed by Fernandes and co-workers,¹⁷ with computed potential energy barrier of $10.6 \text{ kcal}\cdot\text{mol}^{-1}$ and improved by adding activation energy, ZPE, entropic effects, solvation, and water/hydroxide exchange terms to $15.4 \text{ kcal}\cdot\text{mol}^{-1}$, which is within the upper limit of experimentally measured rate constant for this process.

The results presented in this work for '*substrate-assisted*' mechanism together with previously computed for ANT4' or twister ribozyme models indicate that regardless of the number of Mg^{2+} ions in the active site or the type of the substrate containing phosphorus atoms (ATP, RNA or DNA), the internal phosphate group (involved directly in the reaction of P-O bond breaking) cannot play a role of proton acceptor in any of these systems.

Additionally, our observations indicated the importance of the presence of Mg^{2+} ions in the active site of PVF-IN. In particular, one of them, which coordinates the nucleophilic water molecule, seems to be especially vital since it participates in the charge transfer along the reaction path. On the contrary, no significant changes on Mg^{2+} ion responsible for the destabilization of P-O bond were observed. Anyway, until the proper reaction mechanism is

known, it is difficult to determine the actual role of both magnesium cations present in the active site of PVF-IN. However, from the technical point of view, as it was shown herein, it is crucial to describe both Mg^{2+} ions together with the atoms of their first coordination shell at QM level in calculations using QM/MM approaches.

Finally, as expected, significant differences between the results obtained using semiempirical methods, as AM1, and DFT method were found. These differences are mostly related to the description of too short distances for non-bonding interaction between O and P atoms, and over stabilization of the intermediate complex obtained at AM1 level by comparison with DFT methods.

ACKNOWLEDGMENT We would like especially to thank to Prof. Vicent Moliner from Universitat Jaume I (Spain) for his helpful comments done on this work and congratulate him on the occasion of receiving the award for research excellence from the Spanish Royal Society of Chemistry. The authors would like to acknowledge the Servei d'Informàtica, Universitat Jaume I, and Informatic Center on the “Blueocean” supercomputer of Lodz University of Technology for generous allotment of computer time.

Supporting Information. The Supporting Information is available free of charge on the ACS Publications website at DOI:

Evolution of the total energy, the temperature changes monitored along MM MD simulations with RMSD of the protein backbone; pK_a values computed for titratable residues; Distances of first coordination shell of A and B Mg^{2+} ions; key interatomic distances measured for optimized stationary structures at different AM1/MM, B3LYP/MM, MPW1PW91/MM and M06-2X/MM

level of theory; relative potential energies computed at semiempirical (AM1/MM) and DFT (B3LYP/MM, MPW1PW91/MM and M06-2X/MM) level; Coordinates of localized TS at B3LYP/MM, MPW1PW91/MM and M06-2X/MM levels of theory.

Corresponding Author

*Dr. Katarzyna Świderek, e-mail: swiderek@uji.es

ORCID

Katarzyna Świderek: 0000-0002-7528-1551

Agnieszka Krzemińska: 0000-0003-3250-4193

Present Addresses

† present address of A. K. is Institute of Physics, Lodz University of Technology, Lodz, Poland

Author Contributions

A.K. performed MD calculations, and QM/MM studies, K.Ś. designed research, performed FEP calculations, analysed the data, prepared the figures and wrote the manuscript.

Funding Sources

This work was supported by the Polish Ministry of Science and Higher Education (“Iuventus Plus” program project no. 0478/ IP3/2015/73, 2015-2016). K.Ś. would like especially to thank the Spanish Ministerio de Economía y Competitividad for a Juan de la Cierva–Incorporación (ref. IJCI-2016-27503) contract.

Notes

The authors declare no competing financial interest.

REFERENCES

- (1) Coffin, J. M. Retroviridae and Their Replication In Virology Raven Press, New York **1996**, 1767–1848. (ed. Fields, B. N. et al.)
- (2) Craigie, R. Retroviral DNA integration. In Mobile DNA II , Craig, N. L.; Craigie, R.; Gellert, M.; Lambowitz, A. M. (eds) **2002**, 613–630. Washington, DC: ASM Press.
- (3) Lewinski, M. K.; Bushman, F. D. Retroviral DNA Integration—Mechanism and Consequences. *Adv. Genet.* **2005**, *55*, 147–181.
- (4) Lesbats, P.; Engelman, A. N.; Cherepanov, P. Retroviral DNA Integration. *Chem. Rev.* **2016**, *116*, 12730–12757.
- (5) Hare, S.; Gupta, S.S.; Valkov, E.; Engelman, A. N.; Cherepanov, P. Retroviral Intasome Assembly and Inhibition of DNA Strand Transfer. *Nature* **2010**, *464*, 232–236.
- (6) Cai, M.; Zheng, R.; Caffrey, M.; Craigie, R.; Clore, G. M.; Gronenborn, A. M. Chronic 'jet lag' Produces Temporal Lobe Atrophy and Spatial Cognitive Deficits. *Nat. Struct. Biol.* **1997**, *4*, 567–577.
- (7) Eijkelenboom, A. P.; Lutzke, R. A. P.; Boelens, R.; Plasterk, R. H. A.; Kaptein, R.; Hard, K. Cancer Genetics: It Doesn't Always Take Two. *Nat.. Struct. Biol.* **1995**, *2*, 807–810.
- (8) Barre-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. Isolation of a T-lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science* **1983**, *220*, 868–871.

(9) Popovic, M.; Sarngadharan, M. G.; Read, E.; Gallo, R. C. Detection, Isolation, and Continuous Production of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and Pre-AIDS. *Science* **1984**, *224*, 497–500.

(10) World Health Organization (<https://www.who.int/en>) Last updated: 2018-07-11

(11) Antiretroviral Drugs Used in the Treatment of HIV Infection. Source: U.S. FOOD & DRUG ADMINISTRATION (FDA) www.fda.gov - last updated on April 12, 2018

(12) Krishnan, L.; Li, X.; Naraharisetty, H. L.; Hare, S.; Cherepanov, P.; Engelman, A. N. Structure-Based Modeling of the Functional HIV-1 Intasome and Its Inhibition. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 15910–15915

(13) Hare, S.; Vos, A. M.; Clayton, R. F.; Thuring, J. W.; Cummings, M. D.; Cherepanov, P. Molecular Mechanisms of Retroviral Integrase Inhibition and the Evolution of Viral Resistance. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20057–20062.

(14) Passos, D. O.; Li, M.; Yang, R.; Rebersburg, S. V.; Ghirlando, R.; Jeon, Y.; Shkriabai, N.; Kvaratskhelia, M.; Craigie, R.; Lyumkis, D. Cryo-EM Structures and Atomic Model of the HIV-1 Strand Transfer Complex Intasome. *Science* **2017**, *355*, 89–92.

(15) Bernardi, F.; Bottoni, A.; De Vivo, M.; Garavelli, M.; Keserü, G.; Náráy-Szabó, G. A Hypothetical Mechanism for HIV-1 Integrase Catalytic Action: DFT Modelling of a Bio-Mimetic Environment. *Chem. Phys. Lett.* **2002**, *362*, 1–7.

(16) Ruiz-Pernia, J. J.; Alves, C. N.; Moliner, V.; Silla, E.; Tunon, I. A QM/MM Study of the Reaction Mechanism for the 3'-Processing Step Catalyzed by HIV-1 Integrase. *J. Mol. Struct. : THEOCHEM* **2009**, *898*, 115–120.

(17) Ribeiro, A. J. M.; Ramos, M. J.; Fernandes, P. A. The Catalytic Mechanism of HIV-1 Integrase for DNA 3'-End Processing Established by QM/MM Calculations. *J. Am. Chem. Soc.* **2012**, *134*, 13436–13447

(18) Jones, G.; Willett, P.; Glen, R. C. Molecular Recognition of Receptor Sites Using a Genetic Algorithm with a Description of Desolvation. *J. Mol. Biol.* **1995**, *245*, 43-53.

(19) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **2011**, *133*, 8934–8941.

(20) Hare, S.; Maertens, G. N.; Cherepanov, P. 3'-Processing and Strand Transfer Catalysed by Retroviral Integrase in Crystallo. *EMBO J.* **2012**, *31*, 3020.

(21) Gautier, R.; Camproux, A. C.; Tufféry, P. SCit: Web Tools for Protein Side Chain Conformation Analysis. *Nucleic. Acids. Res.* **2004**, *32*, W508–W511.

(22) Krissinel, E.; Henrick K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774-797.

(23) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theor. Comput.* **2011**, *7*, 525-537.

(24) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M., Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theor. Comput.* **2011**, *7*, 2284-2295.

(25) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(27) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters *Proteins: Struct. Funct. Bioinf.* **2006**, *65*, 712–725.

(28) Ivani, I.; Dans, P.D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpí, J.L.; González, C.; Vendruscolo, M.; Laughton, C.A.; Harris, S.A; Case, D.A.; Orozco M. PARMBSC1: A Refined Force-Field for DNA Simulations. *Nat. Methods* **2016**, *13*, 55–58.

(29) Allnér, O.; Nilsson, L.; Villa, A. Magnesium Ion-Water Coordination and Exchange in Biomolecular Simulations. *J. Chem. Theory Comput.* **2012**, *8*, 1493-1502.

(30) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, W.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

-
- (31) Grest, G. S.; Kremer, K. Molecular Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys. Rev. A* **1986**, *33*, 3628-3631.
- (32) Roe, D.R.; Cheatham III, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.
- (33) Roe, D.R.; Cheatham III, T.E. Parallelization of CPPTRAJ Enables Large Scale Analysis of Molecular Dynamics Trajectory Data. *J. Comput. Chem.* **2018**, *39*, 2110-2117.
- (34) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Molec. Graphics* **1996**, *14*, 33-38.
- (35) Dewar, M. J. S.; Zoebisch, E.G.; Healy, E.F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A new General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909
- (36) Stewart, J. J. P. MOSOL, MOPAC for Solid-State Physics. *Quant. Chem. Prog. Exch.* **1985**, *5*, 62-63.
- (37) Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785-789.
- (38) Adamo, C.; Barone, V. Exchange Functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW. *J. Chem. Phys.* **1998**, *108*, 664-675.

(39) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor, Chem, Acc.* **2008**, *120*, 215–241.

(40) Zhao, Y.; Truhlar, D. G. Density Functionals with Broad Applicability in Chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.

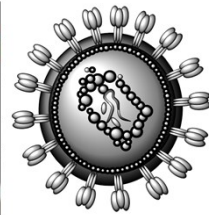
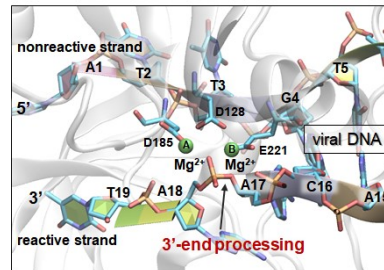
(41) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.

(42) Field, M. J. A Practical Introduction to the Simulation of Molecular Systems Cambridge University Press: Cambridge, UK, 1999.

-
- (43) Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K. Binding Isotope Effects as a Tool for Distinguishing Hydrophobic and Hydrophilic Binding Sites of HIV-1 RT. *J. Phys. Chem. B* **2015**, *119*, 917-927.
- (44) Aranda, J.; Zinovjev, K.; Świderek, K.; Roca, M.; Tuñón, I. Unraveling the Reaction Mechanism of Enzymatic C5-Cytosine Methylation of DNA. A Combined Molecular Dynamics and QM/MM Study of Wild Type and Gln119 Variant. *ACS Catal.* **2016**, *6*, 3262-3276.
- (45) Turner, A.; Moliner, V.; Williams, I. H. Transition State Structural Refinement with GRACE and CHARMM: flexible QM/MM Modelling for Lactate Dehydrogenase. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323-1331.
- (46) Martí, S.; Moliner, V.; Tuñón, I. Improving the QM/MM Description of Chemical Processes: A Dual Level Strategy To Explore the Potential Energy Surface in Very Large Systems. *J. Chem. Theory. Comput.* **2005**, *1*, 1008-1016.
- (47) Fukui, K. The Path of Chemical Reactions - the IRC Approach. *Acc. Chem. Res.* **1981**, *14*, 363.
- (48) Świderek, K.; Tuñón, I.; Martí, S.; Moliner, V.; Bertran, J. Role of Solvent on Nonenzymatic Peptide Bond Formation Mechanisms and Kinetic Isotope Effects. *J. Am. Chem. Soc.* **2013**, *135*, 8708-8719.
- (49) Mlýnský, V.; Banáš, P.; Šponer, J.; van der Kamp, M. W.; Mulholland, A. J.; Otyepka, M. Comparison of ab Initio, DFT, and Semiempirical QM/MM Approaches for Description of Catalytic Mechanism of Hairpin Ribozyme. *J. Chem. Theory. Comput.* **2014**, *10*, 1608-1622.

-
- (50) Heldenbrand, H.; Janowski, P. A.; Giambaşu, G.; Giese, T.J.; Wedekind, J. E.; York, D. M. Evidence for the Role of Active Site Residues in the Hairpin Ribozyme from Molecular Simulations along the Reaction Path. *J. Am. Chem. Soc.* **2014**, *136*, 7789–7792.
- (51) Świderek, K.; Marti, S.; Tuñón, I.; Moliner, V.; Bertran Molecular Mechanism of the Site-Specific Self-Cleavage of the RNA Phosphodiester Backbone by a Twister Ribozyme. *J. Theor. Chem. Acc.* **2017**, *136*, 31.
- (52) Marti, S.; Bastida, A.; Świderek, K. Theoretical Studies on Mechanism of Inactivation of Kanamycin A by 4'-O-Nucleotidyltransferase. *Front. Chem.* **2019**, *6*, 660.
- (53) De Vivo, M.; Dal Peraro, M.; Klein, M.L. Phosphodiester Cleavage in Ribonuclease H Occurs via an Associative Two-Metal-Aided Catalytic Mechanism. *J. Am. Chem. Soc.* **2008**, *130*, 10955–10962.
- (54) Lee, S.P.; Kim, H.G.; Censullo, M.L.; Han, M.K. Characterization of Mg²⁺-Dependent 3'-Processing Activity for Human Immunodeficiency Virus Type 1 Integrase in Vitro: Real-Time Kinetic Studies Using Fluorescence Resonance Energy Transfer. *Biochemistry* **1995**, *34*, 10205-10214.
- (55) Smolov, M.; Gottikh, M.; Tashlitskii, V.; Korolev, S.; Demidyuk, I.; Brochon, J.; Mouscadet, J. Deprez, E. Kinetic Study of the HIV-1 DNA 3'-end Processing Single-Turnover Property of Integrase. *FEBS J.* **2006**, *273*, 1137–1151.
- (56) Gerton, J. L.; Herschlag, D.; Brown, P. O. Stereospecificity of Reactions Catalyzed by HIV-1 Integrase (1999) *J. Biol. Chem.* **1999**, *274*, 33480–33487.

Table of Contents/Abstract Graphics



Prototype Foamy Virus (PFV)