JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

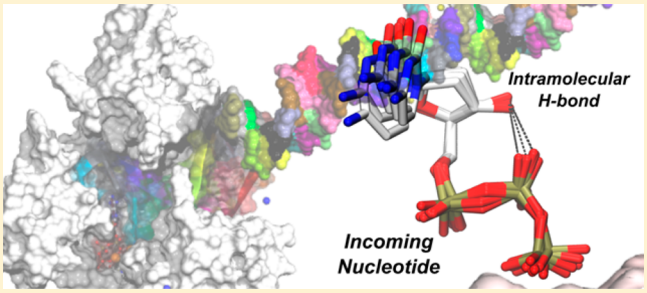
A Self-Activated Mechanism for Nucleic Acid Polymerization Catalyzed by DNA/RNA Polymerases

Vito Genna,[†,‡] Pietro Vidossich,[‡] Emiliano Ippoliti,[‡] Paolo Carloni,*[,‡] and Marco De Vivo*[,†,‡]

[†]Laboratory of Molecular Modeling & Drug Discovery, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genoa, Italy
[‡]IAS-5/INM-9 Computational Biomedicine and JARA-HPC, Forschungszentrum Jülich, Wilhelm-Johnen-Strasse, 52428 Jülich, Germany

Supporting Information

ABSTRACT: The enzymatic polymerization of DNA and RNA is the basis for genetic inheritance for all living organisms. It is catalyzed by the DNA/RNA polymerase (Pol) superfamily. Here, bioinformatics analysis reveals that the incoming nucleotide substrate always forms an H-bond between its 3′-OH and β-phosphate moieties upon formation of the Michaelis complex. This previously unrecognized H-bond implies a novel self-activated mechanism (SAM), which synergistically connects the in situ nucleophile formation with subsequent nucleotide addition and, importantly, nucleic acid translocation. Thus, SAM allows an elegant and efficient closed-loop sequence of chemical and physical steps for Pol catalysis. This is markedly different from previous mechanistic hypotheses. Our proposed mechanism is corroborated via ab initio QM/MM simulations on a specific Pol, the human DNA polymerase-η, an enzyme involved in repairing damaged DNA. The structural conservation of DNA and RNA Pols supports the possible extension of SAM to Pol enzymes from the three domains of life.



INTRODUCTION

Nucleic acid polymerization is a key process for genetic inheritance across the three domains of life.[1] This is performed by a set of DNA/RNA polymerases (Pols) that are often effective drug targets for treating cancer, viral and bacterial infections, and neurodegenerative diseases.[2−4] Pols operate via the two-metal (Mg^{2+})-ion mechanism for incorporating an incoming nucleotide [(d)NTP] into the growing nucleic acid strand, via the typical S_N2-like phosphoryl-transfer reaction, with liberation of a pyrophosphate (PP_i) leaving group (Figure 1).[5,6]

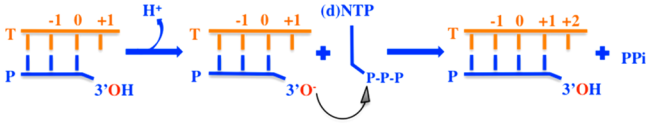


Figure 1. Diagram of nucleic acid synthesis catalyzed by RNA/DNA polymerases. Nucleophile activation, nucleotide addition, and DNA translocation for nucleic acid polymerization, with liberation of a pyrophosphate (PP_i) leaving group. Orange indicates the template strand (T) while blue indicates the primer strand (P).

The established two-metal-aided phosphoryl transfer reaction for nucleotide addition in Pols[7−10] is preceded by deprotonation of the 3′-hydroxyl (3′-OH) of the 3′-end deoxyribose. This generates the activated nucleophilic 3′-hydroxide ion. Importantly, the mechanism for nucleophile formation in Pols is yet unclear and debated.[8−10] In Pol's catalysis, the formation of the 3′-hydroxide ion is the very first chemical step to trigger a nucleophilic attack on the incoming nucleotide (Figure 1), which is bound to the enzyme thanks to a large conformational change for Watson−Crick nascent base pairs, as explained well for DNA polymerase-β catalysis.[11]

A first mechanism for nucleophile formation is via an Asp residue, which is part of the conserved DED motif that coordinates the two catalytic metal ions in Pols.[8,12] This residue can act as a general base for 3′-OH deprotonation, as shown by Warshel and collaborators for DNA polymerase of bacteriophage T7 (protein-activated mechanism).[10,13] Alternatively, the 3′-OH may be deprotonated via a transient bulk water molecule, which can then shuttle the migratory proton on the α-phosphate of the nucleotide, as first reported for catalysis in the lesion-bypass Dpo4 and Pol-κ enzymes [water-mediated and substrate-assisted (WMSA) mechanism].[9,14] Both of these mechanisms imply a stepwise catalytic process made by two formally independent chemical steps, i.e., nucleophile formation and subsequent NTP addition.

Here, bioinformatics analysis of all structures of ternary DNA/RNA Pols complexes (from all domains of life) reveals a previously unrecognized structural determinant that could play a key role in Pol catalysis and that, remarkably, is missing from all previous mechanistic proposals.[9,10] This crucial element is the intramolecular H-bond formed by the nucleophilic 3′-OH and the β-phosphate of the incoming nucleotide (distance d-PT in Figure 2), which is consistently present across all the

Received: May 27, 2016
Published: August 16, 2016

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

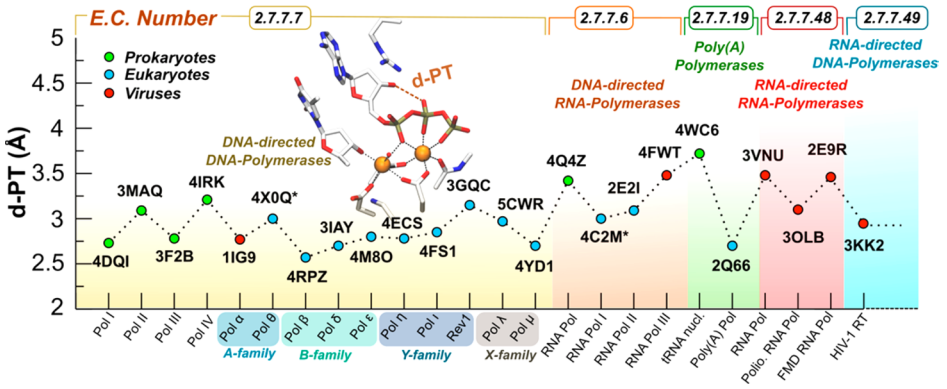


Figure 2. Graph reporting the intramolecular H-bond d-PT in different polymerases. The length of d-PT is reported for structures of Pol families from each domain of life. The *X*-axis reports the protein name. The *Y*-axis reports d-PT (Å). Green dots identify X-ray structures (PDB ID) of Pol from prokaryotes, cyan from eukaryotes, and red from viruses. The background color indicates the enzyme commission number (EC number provided above).

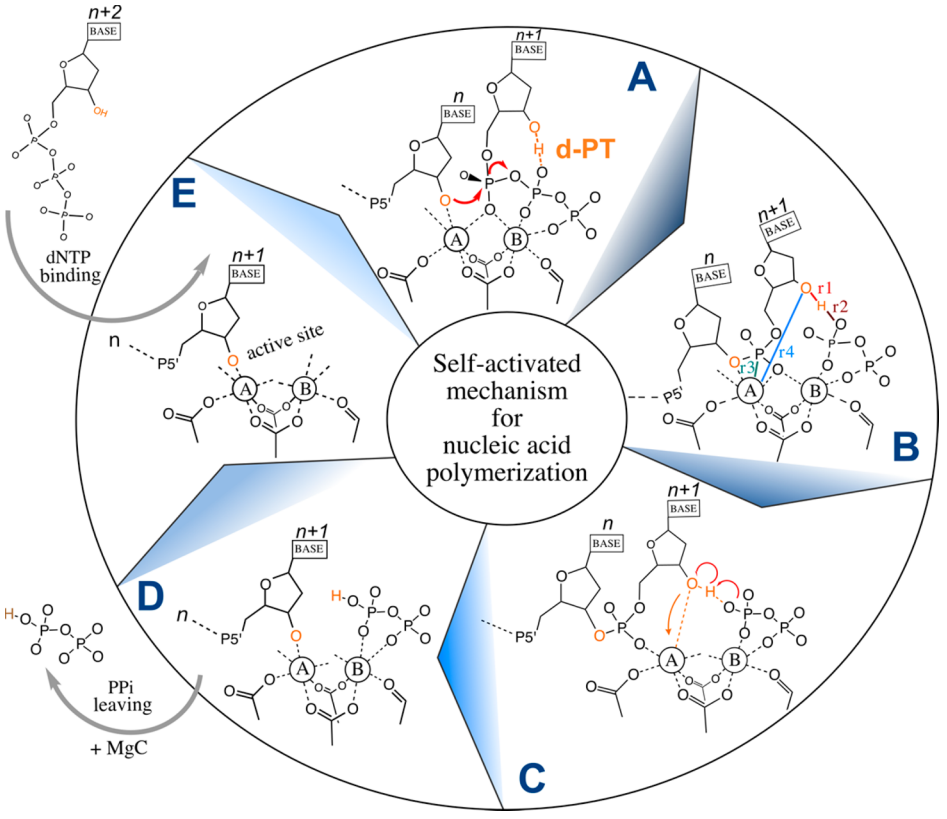


Figure 3. Reaction scheme for the proposed self-activated mechanism (SAM) for nucleic acid polymerization. (A) Michaelis−Menten complex: This state leads to the two-metal-aided S_N2-type phosphoryl transfer with liberation of pyrophosphate (PP_i) leaving group. Notably, the nucleophilic oxygen is here already activated (deprotonated). (B) Products for nucleotide addition: Here, the incoming nucleotide was added to the primer strand. Colored lines indicate selected distances taken as collective variables (CV1 = r1 − r2 and CV2 = r3 − r4 for QM/MM metadynamics) to investigate SAM. (C) Nucleophile formation and nucleic acid translocation: the nucleophile 3′-OH is activated through its deprotonation in favor of the leaving PP_i (PT_1), while r4 is progressively shortened, indicating initial nucleic acid translocation. (D) PP_i exit: at this point, the newly formed 3′-hydroxide group of the incoming nucleotide is coordinated on top of metal A, while the leaving PP_i departs from the catalytic site, helped by the transient third metal ion. (E) dNTP binding and catalytic site closure: the enzyme is ready for the subsequent polymerization cycle upon binding of a new nucleotide, with closure of the catalytic cycle.

currently available structures of Pols adducts that include the (d)NTP (see Results and Discussion). Importantly, we also found that such a short H-bond is favored only when the sugar pucker of the incoming nucleotide adopts its reactive C3′-endo conformation in the Michaelis complex, characterized by the intramolecular H-bond d-PT. Indeed, as reported by Schulten and co-workers,[15] NTP dispersed in solution adopts a more relaxed conformation that does not favor the formation of this H-bond, which therefore defines a productive state of DNA/RNA Pols when complexed with their substrates.[16,17] On the basis of these observations, we propose the following novel catalytic mechanism for nucleic acid polymerization in Pols. First, the key intramolecular H-bond in the incoming nucleotide prompts the in situ 3′-OH activation via its deprotonation in favor of the leaving PP_i (points B, C, Figure 3). Then, the newly formed 3′-hydroxide ion in the incoming nucleotide slowly moves on

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

top of MgA (points C, D, Figure 3) during DNA translocation, assuming the typical coordination required for in-line nucleophilic attack and nucleotide addition, according to the two-metal-ion mechanism.[8,12] In this way, the catalytic cycle is closed and the enzyme is ready for the subsequent round of nucleic acid polymerization (points E, A, Figure 3).

Thus, we describe a new mechanism characterized by a concerted closed-loop catalytic sequence of steps for nucleophile formation, nucleotide addition, and, importantly, nucleic acid translocation. These are synergistically interconnected chemical and physical steps that form a novel enzymatic mechanism for Pol catalysis. Hereafter, we refer to this mechanism as the "self-activated mechanism" (SAM) because it is initiated by a proton transfer for nucleophile formation that occurs within the incoming nucleotide for nucleic acid elongation.

RESULTS AND DISCUSSION

Self-Activated Mechanism for Human DNA Pol-η Catalysis. First and most importantly, we identified a previously unrecognized and conserved H-bond formed by the nucleophilic 3′-OH and the β-phosphate of the incoming nucleotide (distance d-PT in Figures 2 and 4) in all the currently available structures of Pols ternary complexes, with values from ∼2.50 to ∼3.75 Å [Figure 4 and Supplementary Table 1 of the Supporting Information (SI)]. On the basis of this experimental evidence, we propose a new catalytic mechanism for nucleic acid polymerization, which is characterized by a d-PT-prompted proton transfer for in situ 3′-OH activation (SAM, Figure 3). Here, we define SAM in human DNA polymerase-η (Pol-η) catalysis, aided by the wealth of structural and kinetics data on this important enzyme.[18−20] Pol-η is a trans-lesion Pol that catalyzes elongation of DNA affected by UV-induced cyclobutane−pyrimidine dimers (CPDs),[21,22] which are related to skin cancer onset.[23,24]

Recent high-resolution time-resolved X-ray structures of the ternary Pol-η/DNA/dNTP complex have shown the incoming dNTP assuming its reactive C3′-endo sugar pucker conformation, which allows a short (2.78 Å) intramolecular H-bond formed by the nucleophilic 3′-OH and the β-phosphate of dNTP[16] (distance d-PT in Figure 2). According to SAM, this H-bond d-PT, together with the initial DNA translocation, facilitates the deprotonation of the 3′-OH in favor of the β-phosphate (r1 and r2, Figures 3 and 5) of the incoming dNTP

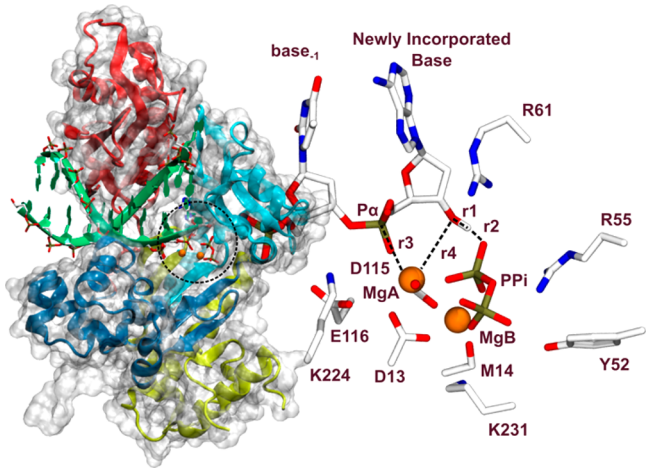


Figure 5. Human DNA Pol-η structure after incorporation of the incoming base. (Left) Overview of the ternary Pol-η/DNA/(d)NTP complex. Each domain of Pol-η is a different color: palm, yellow; thumb, blue; fingers, cyan; and little finger, red. (Right) Close view of the catalytic site of Pol-η. The two Mg^{2+} ions are in orange, nitrogen is in blue, carbon is in white, oxygen is in red, and phosphorus is in maroon.

(points B−C, Figure 3). At this point of the catalytic cycle, the forming interaction between MgA and the approaching 3′-OH group is known to facilitate 3′-hydroxide formation by lowering the pK_a of the 3′-OH within the protein environment (typically ∼7.5−10.5 instead of ∼10.5−12.5).[11,25,26] Thus, the progressive decrease of the 3′-O^--MgA distance during SAM (r4, Figures 3 and 5) implies a significant electrostatic influence of the metal ion on the ionization state of the 3′-OH and nearby residues/groups,[27−30] as comprehensively explained by Warshel and collaborators for other nucleotidyltransferases undergoing significant

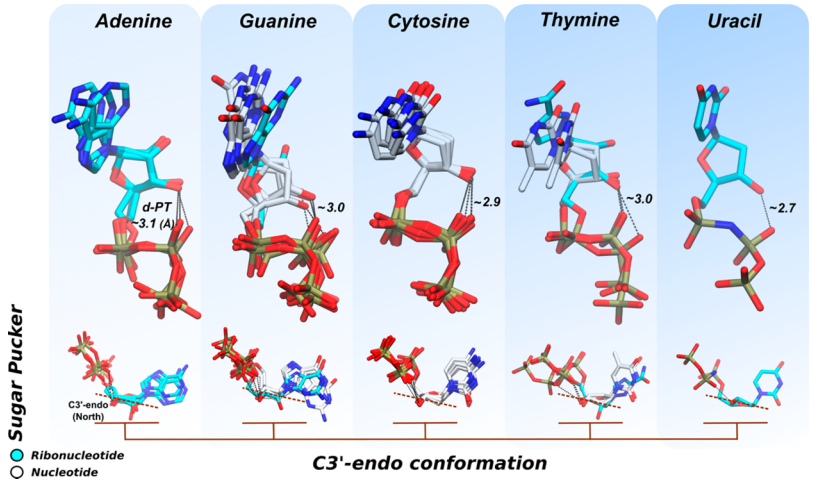


Figure 4. Superimposition of (ribo)nucleotides cocrystallized in Pol's reactive ternary complexes. Structures extracted from different crystals (in Figure 2 and Supplementary Table 1, SI) are superimposed following their species (A, C, G, T, U). The upper part indicates the conserved presence of the intramolecular H-bond (d-PT) in those (ribo)nucleotides complexed with Pol/DNA(RNA) binary complexes. The lower part shows the C3′-endo sugar pucker conformation always detected in those structures. Ribonucleotides (RNA) are cyan. Nucleotides (DNA) are white. Value reported for d-PT is the average value obtained for each type of (ribo)nucleotide.

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

conformational changes.[10,31,32] Thus, within SAM, the electrostatic attraction of the forming hydroxide ion with MgA helps DNA translocation. This was also demonstrated qualitatively by ab initio steered MD simulations and Car−Parrinello (CP) quantum mechanics/molecular mechanics (QM/MM) metadynamics, which consistently indicated that DNA translocation (i.e., shortening of the distance r4) is favored when in the presence of the activated 3′-O^- group, compared to the case with the nucleophile 3′-OH still protonated (see Supplementary Figure 1, SI). Indeed, the X-ray structure of Pol-η, in a state preceding nucleotide addition and DNA translocation (PDB ID 4ECS),[16] has P^{α}−MgA (r3, Figure 3 and 5) and r4 distances equal to 3.42 and 7.05 Å, respectively. Then, the postreactive structure of Pol-η (PDB ID 4ECW)[16] shows r3 increased to 6.14 Å and r4 diminished to 2.29 Å, which reflect initial DNA translocation, with the complete translocation of the 3′-end after the breakage of the P^{α}−MgA interaction. In this way, SAM leads to the (re)formation of an optimal 3′-O^-−MgA coordination, with the newly formed nucleophilic 3′-O^- properly placed to perform the subsequent nucleophilic attack at the incoming nucleotide.[13] Thus, SAM infers a closed-loop catalytic cycle, in which the S_N2-type phosphoryl transfer for nucleotide incorporation in Pols ends by originating a new 3′-hydroxide group that, in turn, initiates the following catalytic addition of the next incoming nucleotide, after DNA translocation and PP_i departure (points A−E in Figure 3).

Remarkably, similar values and variation of r3 and r4 are found in X-ray structures of several other Pol reactive complexes, further suggesting a closed-loop catalytic sequence of both chemical and physical steps formed by nucleophile formation, nucleotide addition, and DNA translocation, as proposed in SAM. For example, bacteriophage N4 RNA-Pol is an enzyme recently studied by means of time-resolved X-ray crystallography to capture real-time intermediates in the pathway of transcription.[33] The series of crystallographic structures for bacteriophage N4 RNA-Pol shows RNA extension, from prereactive to postreactive states. In this case, the prereactive complex (PDB ID 4FF3) has r3 and r4 equal to 3.79 and 7.34 Å, respectively. These two distances correspond to 4.31 and 6.08 Å in the postreactive structure (PDB ID 4FF4), indicating initial nucleic acid translocation and formation of nucleophile−MgA coordination. These data further support the key role for Pol's catalysis of an intimate interconnection between the physical step for nucleic acid translocation and the chemical steps for nucleophile formation and nucleotide addition, as proposed in SAM.

Taken together, this structural evidence and extensive conservation between DNA and RNA Pols suggest an evolutionary convergence to preserve those specific structural features that are key to nucleic acid binding and processing in Pols. There are the conserved DED motif,[34,35] multiple catalytic Mg^{2+} ions,[8,36,37] a positively charged residue in the active site,[38] and, ultimately, a short d-PT, which (according to SAM) is needed to trigger the 3′-OH deprotonation for nucleophile activation. Hence, SAM is remarkably different from previous mechanistic hypotheses of Pols catalysis. This is because SAM is characterized by a synergistic interplay between chemical (i.e., nucleophile formation and nucleotide addition) and physical (i.e., nucleic acid translocation) steps to form a closed-loop cycle for efficient Pols catalysis.[39,40]

QM/MM Simulations of Nucleophile Activation in Pol-η Catalysis. To further corroborate SAM, we next performed ab initio CP QM/MM simulations[41,42] coupled with metadynamics-based free-energy calculations[43] of Pol-η's catalysis. This allowed us to determine the dynamics and semiquantitative energetics of SAM for nucleic acid extension in Pol-η. Here, we analyzed only the coupling between the chemical and physical steps for nucleophile formation and nucleic acid translocation (points B−D in Figure 3), which precede the already well-characterized S_N2-like phosphoryl-transfer reaction for nucleotide addition[7−10] (point A in Figure 3). Thus, we first investigated the proton-transfer along d-PT for in situ formation of the catalytically active 3′-hydroxide ion, using two selected collective variables (CV1 and CV2). CV1 is defined as the difference between the lengths of the breaking 3′-O−H (r1, Figure 3 and 5) and forming $H-O_{PP_i}$ (r2) bonds; CV2 is the difference between the lengths of the P^{α}−MgA (r3) and the 3′-O^-−MgA (r4) coordination bonds. The free-energy surface (FES, Figure 6), projected on those CVs, shows that

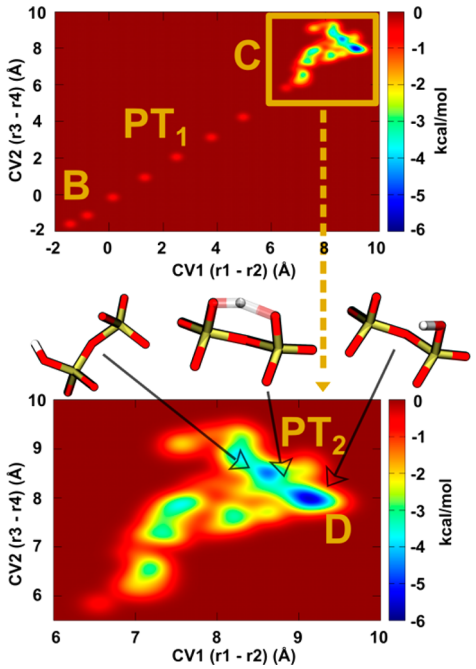


Figure 6. Free-energy surface for SAM in human DNA Pol-η. **B**, **PT_1**, **PT_2**, **C**, and **D** identify saddle points for SAM-catalyzed nucleic acid polymerization in DNA Pol-η, moving from point B of the catalytic cycle to an ensemble of global minima at point D (see reaction scheme and points B and D in Figure 3).

our starting system was initially located in a metastable state B, retrieved by previous extensive MD simulations connecting pre- and postreactive states.[21,22,38] Thus, as expected, the system quickly fell from B into a large minimum D, where the 3′-hydroxide was fully formed, while the leaving PP_i was stably protonated (see Figure 3).

Importantly, two proton transfers occurred moving from **B** to **D**. First, the proton transfer for the self-activation of the nucleophile 3′-O^- occurred at **PT_1**. Then, the transferred proton was shuttled further away on the departing PP_i through a second proton transfer **PT_2**, before the systems fell into **D** (Figure 6). In detail, in **B** (CV1 ∼ −4.0 Å and CV2 ∼ −3.0 Å), the system was only ∼1.2 kcal/mol more stable than its surrounding conformational space. However, a well-structured H-bond network centered on the catalytic Mg^{2+} ions stabilized the overall architecture of Pol-η's catalytic site. In **B**, r3 was 3.28 Å, reflecting a stable P^{α}−MgA coordination. The distance

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

r4 was 5.01 Å, close to the value detected in the X-ray structure of the postreactive state conformation (PDB ID 4ECW,[16] r4 = 7.05 Å). Also, the conserved surrounding residues R61, R55, Y52, and K231 formed a distinctive XRYK motif centered on the PP_i. From **B** to $\mathbf{PT_1}$, the system overcame a series of four small energetic barriers (~1 kcal/mol each, Figure 6). Then, we observed the in-line 3′-O−H−O_{PP_i} proton transfer $\mathbf{PT_1}$, with a barrier of ~2.0 kcal/mol, leading to the final 3′-hydroxide. Notably, the protonation of the leaving PP_i was also observed in other similar enzymatic reactions, where the leaving PP_i served as the final proton acceptor for nucleophile formation.[9,14,25,44,45] Here, the 3′-OH deprotonation event is well-captured by r1 and r2, which gradually changed from 1.02 and 2.58 Å in **B** to 1.42 and 1.07 Å in $\mathbf{PT_1}$, respectively. Interestingly, at this point, the variation of r3 and r4 (of 3.75 and 3.55 Å, respectively) reflects the shift of the newly generated 3′-O^-, which slowly moved on top of MgA, while the phosphate group of the 3′-terminal base slid away (points B, C, Figure 3). Altogether, this indicates an initial DNA translocation, which occurs concomitantly to nucleophile formation (see below). Also, during DNA translocation in SAM, the two catalytic metal ions increase their initial internuclear distance from 3.36 ± 0.14 Å in point **A** to about 4 Å in point **B**. Then, after DNA translocation, the two ions slowly return to their initial internuclear distance of ~3.5 Å, moving from **C** to **D**−**E**−**A** to stabilize the transition state along the phosphoryl transfer for nucleotide addition. Noteworthy, the cooperative motion of the two catalytic ions was reported for other nucleic acid-processing two-metal-ion enzymes.[7,8,35,36,46,47] Clearly, additional costly simulations of the overall catalytic cycle are needed to better establish the level of synchronicity and synergy of SAM's chemical and physical steps.

From $\mathbf{PT_1}$, the system evolved toward PT_2 (CV1 ~ 6.5 Å and CV2 ~ 5.8 Å). This second intramolecular proton transfer $\mathbf{PT_2}$ occurred from the β-group to the adjacent γ-group of the PP_i, with a barrier of ~2.0 kcal/mol. PT_2 is also shown by r1 and r2, which became ~10.5 and ~4.0 Å respectively, while r3 and r4 changed to ~5.8 and ~3.5 Å, further suggesting the initial DNA translocation. Precisely, the proton previously shuttled to O_{PP_i} from 3′-OH in $\mathbf{PT_1}$ was rotated by about ~270° with respect to its donor species. In this way, this proton pointed toward one of the nonbridging oxygen atoms of the γ-phosphate of PP_i. From here, it was then quickly shuttled ($\mathbf{PT_2}$) on the adjacent phosphate of the PP_i, where it stably remained for the rest of the simulations. This protonation state of the PP_i was also found for T7 DNA polymerase catalysis,[9] further confirming the likely role of the PP_i as the ultimate acceptor of the shuttled proton generated by the 3′-OH deprotonation. Immediately after PT_2, the system rapidly fell into the deepest energetic minimum D of the FES (CV1 ~ 7.5 Å and CV2 ~ 6.0 Å), which is at ca. −6.0 kcal/mol (Figure 6). This energetic minimum was confirmed by additional ~25 ps of unbiased QM/MM simulations, during which the architecture of the metal-aided catalytic site, as well as the transferred proton on the PP_i γ-group, were maintained, matching well the crystallographic prereactive state of Pol-η (PDB ID 4ECS,[16] RMSD ~ 3.0 Å; see Supplementary Figure 2, SI).

Notably, our calculations provide only a thermodynamics description of the process under investigation, while the overall relaxation step of the whole ternary complex, after PP_i release, is suggested to be the rate-limiting step of the polymerization process catalyzed by human Pol-η, as already proposed for the structurally similar Y-family members Dpo4 and Pol-k.[20,48,49] The overall relaxation step of the whole ternary complex is therefore likely to remain the rate-limiting step of SAM, although this point remains to be clarified by further investigations. In addition, the recent time-resolved crystallographic structures of Pol-η have revealed a transient third ion bound at the catalytic site after nucleotide insertion.[50−52] This third ion is suggested to facilitate product formation during nucleotide addition and, as also proposed by our previous MD simulations,[38] to serve as an exit shuttle for the leaving PP_i. In this respect, we preliminarily evaluated the effect of the third ion in SAM. First, additional QM/MM simulations demonstrated that this transient third ion hampers nucleophile formation and DNA translocation, if bound to the pretranslocation complex, point B in Figure 3 (see Supplementary Figure 3, SI). This explains the structural evidence that a third metal ion cannot be placed in the reactant enzyme−substrate complex, mainly because of steric clashes.[50] On the other hand, further QM/MM simulations revealed also that a third ion bound at the catalytic site of Pol-η in the product state, i.e. after nucleophile formation and nucleotide addition, facilitates the exit of the PP_i leaving group, while preventing the reverse reaction of pyrophosphorolysis (see Supplementary Figure 4, SI). These results further corroborate the evidence that the third metal can be transiently bound only at the product state during catalysis.[50,52] Therefore, the key initial steps in SAM (i.e., nucleophile 3′-O^- formation and DNA translocation) do not require a transient metal ion that, again, was in fact experimentally found only in the products. This puzzling and nascent concept of a functional and cooperative dynamics of multiple catalytic metal ions for DNA and RNA processing undoubtedly merits further studies.[8,50]

Often, DNA polymerases contain a highly flexible positively charged residue, like an arginine or a lysine, which is conserved and located near the catalytic site.[38] This residue is R61 in Pol-η.[18,19,38,53] We analyzed the role of this residue in SAM, and found that R61 stabilizes the negatively charged 3′-hydroxyl nucleophile, when it adopts what is referred to as the "A" conformation (A-conf). This conformation is characterized by bifurcated hydrogen bonds established with the leaving group PP_i. This likely prevents a back-proton migration from the protonated PP_i to the active nucleophile. R61 in the "C" conformation (C-conf), where it forms two H-bond interactions with the incoming base, generates an approximately 6.0 kcal/mol higher barrier for nucleophile formation and initial DNA translocation, compared to the system with R61 in A-conf (see Supplementary Figure 5, SI). Thus, A-conf favors the nucleotide incorporation, while C-conf guides the incoming base into the catalytic site and assists PP_i departure toward the solvent-exposed part of the cavity, as previously reported.[38]

Overall, the present work does not rule out other possible mechanisms for the 3′-OH deprotonation in Pol-η[16] (see Supplementary Figure 6, SI) and other previously reported mechanisms for Pol's catalysis.[9,10] Indeed, the WMSA mechanism remains a valid hypothesis for Pol-η's catalysis given the persistent presence, in the recent crystals, of a bulk water molecule properly located to act as a general base for nucleophile deprotonation.[9,14,16,50] However, we found that nucleophile formation via this bulk water molecule is energetically unfavored compared to SAM (see Supplementary Figure 6, SI). Indeed, other transient bulk waters, as well as surrounding residues, could in principle accept the proton from the nucleophile 3′-OH group.[10,13,54] However, when compared to these previously

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

proposed mechanisms, we underline that only SAM does (i) account for the absolutely conserved intramolecular H-bond d-PT at the active site of DNA/RNA Pols, formed within the incoming nucleotide, and (ii) imply a highly efficient coupling of DNA translocation with nucleophile formation for enzymatic nucleic acid polymerization.

CONCLUSION

We propose a novel self-activated mechanism for efficient polymerase catalysis, which is based on the identification of an evolutionary convergence to preserve a key enzymatic structural element in all the available X-ray structures of DNA/RNA polymerases from all domains of life. This is a structurally conserved H-bond formed by the nucleophilic 3′-OH and the nonbridging oxygen of the β-phosphate in the incoming nucleotide, in the Michaelis complex only. SAM is characterized by the synergistic interplay of in situ nucleophile formation (via 3′-OH deprotonation), nucleotide addition, and, importantly, DNA translocation. Thus, SAM allows formation of a closed-loop catalytic cycle characterized by a concerted sequence of steps of an elegant and efficient nucleic acid polymerization, as shown here by our analyses of polymerase structures and by our simulations of DNA elongation catalyzed by Pol-η. Importantly, on the basis of the extensive structural conservation of RNA and DNA polymerases, we propose SAM to be transferable to a broad range of other nucleic acid-processing enzymes.

METHODS

Structural Model and Car−Parrinello QM/MM Simulations. Our ternary Pol-η/DNA/dNTP model system is based on the crystallographic structure of the enzyme structure after completion of the nucleotidyl-transfer reaction and consequent formation of products (PDB ID 4ECW).[16,38] This structural model was used here to verify the coupling between nucleophile formation and DNA translocation, as proposed in SAM. Toward this end, we performed ab initio CP simulations, in the QM/MM implementation,[41] coupled with metadynamics-based free-energy calculations[43] of Pol-η catalysis. As these are enhanced sampling simulations, they cannot provide information on the time scale of the events. The reactive region of the ternary complex was treated at the DFT/BLYP level and includes the Mg^{2+} coordination sphere (DED motif: D13, E115, D118, M14), part of the DNA dA:dT, dT_{-1} nucleotides, R61, pyrophosphate, and solvation water molecules (for a total of 183 QM atoms, Figure 5). The remaining part of the complex (~70 000 atoms) was treated using the Amber force field. The valence electrons are described by a plane wave basis set up to a cutoff of 70 Ry. A 20 × 20 × 18 $Å^3$ cell includes the QM part of the system. The interactions between valence electrons and ionic cores are described with norm-conserving Martins−Troullier pseudopotentials. CP QM/MM dynamics is carried out with a time step of 0.12 fs (for a total simulation time of ~250 ps, including plain, steered, and metadynamics QM/MM simulations) and a fictitious electron mass of 500 au; constant temperature simulations are achieved by coupling the system with a Nose′−Hoover thermostat at 500 cm^{-1} frequency. The interactions between the MM and QM regions are coupled in a Hamiltonian scheme as discussed by Laio et al.[41] Notably, a rigorous Hamiltonian treatment of the electrostatic interaction between QM and MM regions is used, as in ref 41. The approach has been shown to accurately describe a variety of metal-dependent enzymes[55−61] and, specifically, protein−DNA complexes.[34−36]

The CP QM/MM protocol includes an initial equilibration phase, followed by a short run where only the MM part is free to move, while the QM part is kept frozen. Notably, the starting configurations were retrieved from our recent microsecond-long classical MD study.[38] Then, the whole system is allowed to move and heat up to 300 K (~2 ps). Trajectories are then collected for analysis. Configurations from the equilibrated CP QM/MM simulations are used for free-energy calculations. Specifically, we used the extended Lagrangian metadynamics techniques in the context of first-principle simulations to reconstruct the free-energy landscape associated with nucleophile activation and DNA translocation. The free energy was determined as a function of two selected collective variables (CVs; see Figure 5) that identify the main motions taken into consideration. CV1 is defined as the difference between the length of the breaking 3′-O−H bond (r1) and that of the forming $H-O_{PP_i}$ bond (r2). CV2 is the difference between the length of the breaking P^{α}−MgA (r3) and that of the 3′-O^{-}−MgA interaction (r4). The Gaussian function deposition rate was set to 24 fs. The initial hills height and width were set to 0.05 kcal mol^{-1} and 0.01 Å, respectively. A total of ~600 Gaussians were deposited from A to D, in two replica systems (~120 000 steps). The Lagrangian simulations were carried out until their convergence (~60 000 steps per replica), i.e. the progressive stabilization of the energetic minima on the free-energy surface (see Supplementary Figure 7, SI). All other parameters correspond to those used for the plain QM/MM MD simulations described above. See the Supporting Information for further details on the computational setup and calculations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b05475.

An explanation of Car−Parrinello molecular dynamics and of the role of metal ions in SAM, Supplementary Figures 1−7, and Supplementary Table 1 (PDF)

AUTHOR INFORMATION

Corresponding Authors

*p.carloni@fz-juelich.de

*marco.devivo@iit.it

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

M.D.V. thanks the Italian Association for Cancer Research (AIRC) for financial support (MFAG no. 14140). The authors gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and the JARA-HPC Vergabegremium, provided on the supercomputer JUROPA (Project id HGR25) and on the JARA-HPC Partition part of JURECA (Project id JIAS52) at Jülich Supercomputing Centre (JSC). We thank Grace Fox for proofreading and copy-editing the manuscript.

REFERENCES

(1) Loeb, L. A.; Monnat, R. J. *Nat. Rev. Genet.* **2008**, *9*, 594.

(2) Lange, S. S.; Takata, K.; Wood, R. D. *Nat. Rev. Cancer* **2011**, *11*, 96.

(3) Bae, B.; Nayak, D.; Ray, A.; Mustaev, A.; Landick, R.; Darst, S. A. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E4178.

(4) Warren, T. K.; Wells, J.; Panchal, R. G.; Stuthman, K. S.; Garza, N. L.; Van Tongeren, S. A.; Dong, L.; Retterer, C. J.; Eaton, B. P.; Pegoraro, G.; Honnold, S.; Bantia, S.; Kotian, P.; Chen, X.; Taubenheim, B. R.; Welch, L. S.; Minning, D. M.; Babu, Y. S.; Sheridan, W. P.; Bavari, S. *Nature* **2014**, *508*, 402.

(5) Steitz, T. A. *Curr. Opin. Struct. Biol.* **1993**, *3*, 31.

(6) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22*, 5.

(7) Hou, G. H.; Cui, Q. *J. Am. Chem. Soc.* **2013**, *135*, 10457.

(8) Palermo, G.; Cavalli, A.; Klein, M. L.; Alfonso-Prieto, M.; Dal Peraro, M.; De Vivo, M. *Acc. Chem. Res.* **2015**, *48*, 220.

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598

(9) Wang, L. H.; Yu, X. Y.; Hu, P.; Broyde, S.; Zhang, Y. K. *J. Am. Chem. Soc.* **2007**, *129*, 4731.
(10) Florian, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163.
(11) Sucato, C. A.; Upton, T. G.; Kashemirov, B. A.; Batra, V. K.; Martinek, V.; Xiang, Y.; Beard, W. A.; Pedersen, L. C.; Wilson, S. H.; McKenna, C. E.; Florian, J.; Warshel, A.; Goodman, M. F. *Biochemistry* **2007**, *46*, 461.
(12) Steitz, T. A.; Steitz, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6498.
(13) Florian, J.; Goodman, M. F.; Warshel, A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6819.
(14) Lior-Hoffmann, L.; Wang, L. H.; Wang, S. L.; Geacintov, N. E.; Broyde, S.; Zhang, Y. K. *Nucleic Acids Res.* **2012**, *40*, 9193.
(15) Harrison, C. B.; Schulten, K. *J. Chem. Theory Comput.* **2012**, *8*, 2328.
(16) Nakamura, T.; Zhao, Y.; Yamagata, Y.; Hua, Y. J.; Yang, W. *Nature* **2012**, *487*, 196.
(17) McGinnis, J. L.; Dunkle, J. A.; Cate, J. H. D.; Weeks, K. M. *J. Am. Chem. Soc.* **2012**, *134*, 6617.
(18) Biertumpfel, C.; Zhao, Y.; Kondo, Y.; Ramon-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. *Nature* **2010**, *465*, 1044.
(19) Zhao, Y.; Gregory, M. T.; Biertumpfel, C.; Hua, Y. J.; Hanaoka, F.; Yang, W. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8146.
(20) Patra, A.; Nagy, L. D.; Zhang, Q.; Su, Y.; Muller, L.; Guengerich, F. P.; Egli, M. *J. Biol. Chem.* **2014**, *289*, 16867.
(21) Ucisik, M. N.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2015**, *137*, 13240.
(22) Ucisik, M. N.; Hammes-Schiffer, S. *J. Chem. Inf. Model.* **2015**, *55*, 2672.
(23) Inui, H.; Oh, K. S.; Nadem, C.; Ueda, T.; Khan, S. G.; Metin, A.; Gozukara, E.; Emmert, S.; Slor, H.; Busch, D. B.; Baker, C. C.; DiGiovanna, J. J.; Tamura, D.; Seitz, C. S.; Gratchev, A.; Wu, W. H.; Chung, K. Y.; Chung, H. J.; Azizi, E.; Woodgate, R.; Schneider, T. D.; Kraemer, K. H. *J. Invest. Dermatol.* **2008**, *128*, 2055.
(24) Zhao, Y.; Biertumpfel, C.; Gregory, M. T.; Hua, Y. J.; Hanaoka, F.; Yang, W. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7269.
(25) Castro, C.; Smidansky, E.; Maksimchuk, K. R.; Arnold, J. J.; Korneeva, V. S.; Gotte, M.; Konigsberg, W.; Cameron, C. E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 4267.
(26) Harris, T. K.; Turner, G. J. *IUBMB Life* **2002**, *53*, 85.
(27) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210.
(28) Hammes, G. G.; Benkovic, S. J.; Hammes-Schiffer, S. *Biochemistry* **2011**, *50*, 10422.
(29) Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E. *Science* **2011**, *332*, 234.
(30) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *J. Phys. Chem. B* **2009**, *113*, 1253.
(31) Xiang, Y.; Goodman, M. F.; Beard, W. A.; Wilson, S. H.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 231.
(32) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283.
(33) Basu, R. S.; Murakami, K. S. *J. Biol. Chem.* **2013**, *288*, 3305.
(34) Palermo, G.; Stenta, M.; Cavalli, A.; Dal Peraro, M.; De Vivo, M. *J. Chem. Theory Comput.* **2013**, *9*, 857.
(35) De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 10955.
(36) Ho, M. H.; De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2010**, *132*, 13702.
(37) Rosta, E.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2014**, *136*, 3137.
(38) Genna, V.; Gaspari, R.; Dal Peraro, M.; De Vivo, M. *Nucleic Acids Res.* **2016**, *44*, 2827.
(39) Golosov, A. A.; Warren, J. J.; Beese, L. S.; Karplus, M. *Structure* **2010**, *18*, 83.
(40) Silva, D. A.; Weiss, D. R.; Pardo Avila, F. P.; Da, L. T.; Levitt, M.; Wang, D.; Huang, X. H. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 7665.
(41) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941.
(42) Carloni, P.; Rothlisberger, U.; Parrinello, M. *Acc. Chem. Res.* **2002**, *35*, 455.
(43) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302.
(44) Wang, Y.; Schlick, T. *J. Am. Chem. Soc.* **2008**, *130*, 13240.
(45) Da, L. T.; Wang, D.; Huang, X. H. *J. Am. Chem. Soc.* **2012**, *134*, 2399.
(46) Lopez-Canut, V.; Roca, M.; Bertran, J.; Moliner, V.; Tunon, I. *J. Am. Chem. Soc.* **2011**, *133*, 12050.
(47) Andrews, L. D.; Fenn, T. D.; Herschlag, D. *PLoS Biol.* **2013**, *11*, e1001599.
(48) Beckman, J. W.; Wang, Q.; Guengerich, F. P. *J. Biol. Chem.* **2008**, *283*, 36711.
(49) Lone, S.; Townson, S. A.; Uljon, S. N.; Johnson, R. E.; Brahma, A.; Nair, D. T.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Mol. Cell* **2007**, *25*, 601.
(50) Gao, Y.; Yang, W. *Science* **2016**, *352*, 1334.
(51) Vyas, R.; Reed, A. J.; Tokarsky, E. J.; Suo, Z. C. *J. Am. Chem. Soc.* **2015**, *137*, 5225.
(52) Perera, L.; Freudenthal, B. D.; Beard, W. A.; Shock, D. D.; Pedersen, L. G.; Wilson, S. H. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E5228.
(53) Su, Y.; Patra, A.; Harp, J. M.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2015**, *290*, 15921.
(54) Miller, B. R.; Beese, L. S.; Parish, C. A.; Wu, E. Y. *Structure* **2015**, *23*, 1609.
(55) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198.
(56) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2010**, *46*, 2354.
(57) Monard, G.; Merz, K. M. *Acc. Chem. Res.* **1999**, *32*, 904.
(58) Campomanes, P.; Rothlisberger, U.; Alfonso-Prieto, M.; Rovira, C. *J. Am. Chem. Soc.* **2015**, *137*, 11170.
(59) Dal Peraro, M.; Ruggerone, P.; Raugei, S.; Gervasio, F. L.; Carloni, P. *Curr. Opin. Struct. Biol.* **2007**, *17*, 149.
(60) Vidossich, P.; Magistrato, A. *Biomolecules* **2014**, *4*, 616.
(61) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.

DOI: 10.1021/jacs.6b05475
J. Am. Chem. Soc. 2016, 138, 14592−14598