Point-by-point response to reviewer comments for "molecular events"

Dataset · June 2016

1 author:

Zhong Ren

Renz Research, Inc.

49 PUBLICATIONS **2,341** CITATIONS

SEE PROFILE

Molecular Events during Translocation and Proofreading Extracted from 200 Static Structures of DNA Polymerase

Zhong Ren

Department of Chemistry, The University of Illinois at Chicago, Chicago, IL 60607, USA
Renz Research, Inc., Westmont, IL 60559, USA
zren@uic.edu

Point-by-point responses to referee comments

I appreciate the constructive comments and detailed suggestions from the editor and referees. I have carefully reworked and revised parts of the manuscript to address all the points raised. Here I list specifically the changes made throughout the text and figures together with my point-by-point responses to the comments. The topics are arranged by relevance to one another. The referee comments are copied verbatim first in *italics*.

Referee 1: Overview: The author presents a detailed analysis of the large data base of RB69 DNA polymerase structures in the protein data bank with the goal to deduce (1) how family B DNA polymerases translocate along the template strand during DNA replication and (2) how translocation is avoided after the primer strand is excised by the exonuclease and returned back to the polymerase active site. These are worthy, but ambitious projects. While new information is provided by the SVD analysis (singular value decomposition), which is useful to the field, there are several shortcomings. It is necessary for the author to consider these shortcomings and to point out to readers the limitations of structural studies.

Referee 2: The manuscript by Ren entitled, "Molecular Events during Translocation and Proofreading Extracted from 200 Static Structures of DNA Polymerase" presents a meta-analysis of a large number of family B DNA polymerase structures in an effort to understand the conformational space traversed by family B polymerases during translocation and proofreading. The approach utilizes a singular value decomposition (SVD) analysis which has been applied previously by the author to an examination of hemoglobin structures. The result is a detailed molecular description of polymerase conformations, including an ordered distribution of conformational states. These data are then used to develop a model for how protein dynamics may drive translocation and processing events. This work provides a significant distillation of 200 crystal structures and is an important contribution to understanding molecular mechanisms of family B polymerases. The general approach will be of interest to the broader structural biology community.

Referee 3: This paper reports very detailed studies of DNA polymerases and their conformations with an emphasis on the possibilities to change their forms in a drastic way between a primary role as translocation enzyme for selecting nucleotides in the replication of DNA and later roles

with a controlling, proofreading effect when checking that an actual nucleotide was correctly put in its proper place. This is described by a very careful and detailed analysis complemented by many illustrating pictures. Clearly, with all its details, this provides a valuable presentation of important events in DNA replication. As I see it, these are very well illustrated and clearly presented. It is very difficult to point out parts which are less well presented and I cannot find any parts which should be improved. This is definitely a valuable contribution to our knowledge of protein actions.

I fully agree that a more complete mechanistic understanding can only come from a unified view of functional and structural studies. This work presents such an attempt to explain the dynamics involved in two basic functions of this polymerase – translocation and proofreading based on a joint analyzes of a large number of static structures. Each static structure observed by cryocrystallography potentially carries its limitation. This manuscript provides a case study that explores and exploits the nature of such limitation ("The reconstructed movie" paragraph on P. 11). It is the understanding of the limitation in structural studies that motivates a decomposition of observed motions and an extraction of structural information relevant to function. Identification and removal of artifacts due to static structures that are isolated from a transient process are facilitated by the numerical decomposition.

Referee 1: Potential shortcomings to consider: 1. An inherent problem with structure studies is that they are snapshots of the most stable DNA polymerase conformations; thus, the dynamic nature of DNA polymerases in action must be inferred from stable structures that are formed before and after rapid events. Solution studies can help fill in the blanks.

Agreed. Crystallographic structures are snapshots. This work provides one more example to demonstrate that the dynamic properties of a protein structure can be extracted from a collection of abundant snapshots. It is entirely possible that a stable structure chemically trapped in the crystal lattice may differ from those during function. However, such possibility is not a valid ground to dismiss a structure. We can still learn valid structural information if potential artifacts are understood. For example, the downward bending hook appeared in the trajectory of translocation (Fig. 5g; also reprinted below) indicates that the polymerase structure moves back towards the "near" conformation momentarily while the thumb disengages from the product duplex. Although this feature is likely due to artifacts of static structures (circled in red below), these static structures clearly carry crucial information about the disengagement between the protein and DNA. However, another aspect of these structures needs to be amended for a better understanding of the mechanism.

Referee 2: The following issues should be addressed before accepting for publication: Pg. 3 ln 18. The author states, "Each experimentally determined static structure is a snapshot of the protein structure during its function." It seems likely that the database of polymerase structures will contain trapped states that are off the normal pathway. How does the author account for or identify off-pathway structures/conformations in the database? Does the inclusion of these structures in the SVD analysis significantly alter or affect the results? Please address.

This sentence has been revised (P. 3, Material & Methods). No assumption is made entering a structural meta-analysis that each snapshot of protein structure must follow a normal reaction trajectory. However, each structure represents an accessible conformation in the vast conformational space as evidenced by the experimentally observed structure. Trapped structures off the normal pathway are identified during the process of meta-analysis itself. First, a single or a small number of structures lying off the trajectory that is defined by the majority of observed structures raises the suspicion of outliers. Their functional relevance must be carefully examined. For example, a so-called R3 structure of hemoglobin may represent a dead-end reaction pathway that occurs abnormally, which has been clearly identified as an outlying structure off the main trajectory (Ren PLoS One 8, e77141, 2013). Second, a subset of the observed structures may collectively skew the reaction trajectory such as those identified in this work (Fig. 5g reprinted below and "The reconstructed movie" paragraph on P. 11). See above.

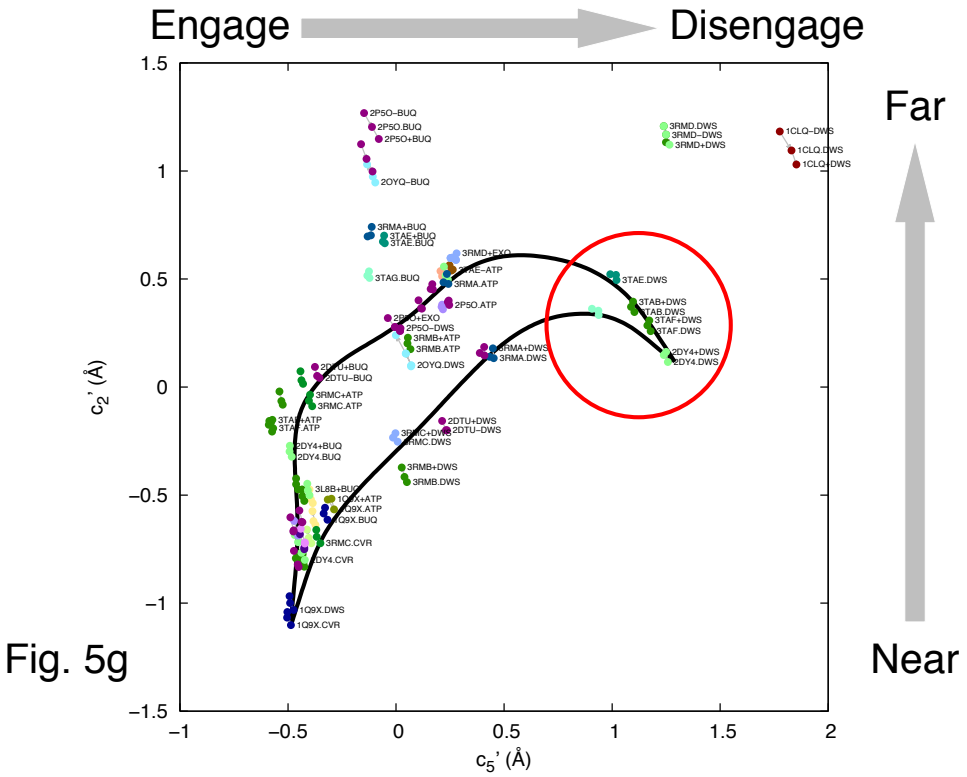

Fig. 5g

Referee 2: Pg.7 ln 23. "Apparently, such discreteness of states is distinct from the behavior of polymerases in family A." The author cites one example (ref 25) where an "ajar" conformation was observed in the fingers domain upon formation of a mismatch. Aside from that observation (to this reviewer's knowledge), the majority, if not all other family A crystal structures adopt distinct open and closed states. Does the author mean to suggest that the proposed mechanism for family B polymerases does not apply to family A polymerases on the basis of one unique set of crystal structures in the database? (Again, how does the author distinguish between on-pathway

and off-pathway events?) Obviously, it would be very informative to perform a similar meta-analysis with family A structures, but the scope of that work is beyond what would be expected for this manuscript. However, this manuscript would be strengthened by some discussion of how the proposed mechanisms relate to other polymerase families that contain similar domains, including family A polymerases.

As suggested, a short paragraph is added to the end of Discussion (P. 20). The commonality among different families could be limited, not because a single ajar conformation was captured, rather because the overall architectures of different polymerases families differ significantly although they share some domains. It would require a similar meta-analysis to show whether the ajar conformation is indeed off-pathway. However, such analysis is technically more challenging, although highly desirable.

Referee 1: The results of the analysis have the potential to provide new insights, but only if there is further study. For example, the author should suggest engineering specific DNA polymerase mutants to test his proposals.

As requested, a sentence has been added to suggest potential experiments for testing (P. 8 above the summary paragraph). The structural mechanism of translocation and proofreading derived from this joint analysis also leads to predictions of structural events that have not been observed directly. For example, the split-end interface consisting of only main chain and rigid side chains such as Pro, Ala, and Gly is important to the function. Mutants that change the property of the interface will severely affect the function. Second, the model predicts that synchronized translocations of the single-stranded template and product duplex during normal replication would not cause the β-hairpin to lift. The passage behind the β-hairpin would only open when the single-stranded template and product duplex translocate differentially (P. 15-16). Third, the model also predicts that a short deletion opposite an abasic site in the template tends to occur when consecutive dTs are incorporated to the primer [point 5) on P. 18]. All these are experimentally testable hypotheses. However, I felt that specific experimental designs such as mutagenesis to validate or disprove these predictions are independent of this work.

Referee 1: Some mutants in the fingers and thumb domains that affect translocation may be relevant to the author's proposals (JMB (2010) 400:295).

Some mutants that affect the stability of the protein-DNA complex and its processivity do not seem to be related to translocation and active site switching. Certain aspects of this polymerase are still beyond our comprehension.

Referee 1: For example, fluorescence studies with the base analog 2-aminopurine demonstrate that there is a rapid equilibrium between open pre- and post-translocation states and that amino acid changes for conserved residues in the polymerase active site affect the equilibrium.

Rapid, transient structural changes cannot be captured by static crystallographic structures in the current PDB. I cannot correlate fluorescence studies with the reaction

trajectory. In my opinion, what is lacking in current structural biology is the coupling between functional and structural studies and between different techniques of studies. The polymerase field is only a specific case illustrating such urgency for cross-study coupling. Many independent studies are difficult to be unified to illustrate a more complete view because of large gaps in spatial and temporal resolutions between studies. Crystallographic structures archived in PDB at the atomic resolution completely lack time stamps in a reaction pathway. This work, together with my previous attempts, tries to arrange a large number of structures into a logical order. But I agree that it remains difficult to confidently assign a structural event without any temporal resolution to dynamic fluorescence measurements that have only limited spatial resolution.

Referee 1: The L415M substitution favors the pre-translocation state, which increases processivity and pyrophosphorolysis.

This mutant is located on one side of the base pair binding pocket. One structure of this mutant is available in the closed state 4J2E. However, I cannot explain its effect to processivity based on the information we know.

Referee 1: Other RB69 DNA polymerase mutants form exceptionally stable closed pre-T complexes; these complexes have been observed with the pyrophosphate-like drug, PFA, bound in the PPi site (JBC (2011)286:25246). This structure does not appear to have been included in the SVD analyses.

Two related structures 3KD1 and 3KD5 of a chimeric polymerase are available, in which some portions of the fingers adopt the sequence of human cytomegalovirus. Both structures were included in the analysis. They turned out to be the only exceptions in the closed state without triphosphate binding. Discussion regarding these structures has been expanded and moved into the main text (first paragraph on P. 8). A figure is added to show that the exception is caused by the extensive substitution in the fingers that may not be completely compatible to the other parts of the structure (Fig. 5f). The zoom-in plot of the closed structures shows that 3KD1 and 3KD5 are located towards the extreme of the closed conformation when compared to the other closed structures. It is puzzling to me why the original paper claims that phosphonoformic acid (PFA) traps the polymerase in the closed form, which in turn inhibits viral replication. The structure 3KD1 without PFA reported by the same paper is clearly in a very closed conformation too. PFA may have stabilized the closed conformation even further. But the closed conformation is not resulted from PFA binding, rather from the extensive substitution in the fingers.

Referee 1: Summary: The author raises several interesting possibilities re translocation and proofreading, but the picture is incomplete. Some predictions from the SVD analysis are not explained with respect to the current literature. For example, it would be useful to include the closed pre-T complex with PFA bound in the PPi binding site along with open structures. The

SVD analysis opens the door to potentially new insights, but these must be followed up with experimentation.

Findings of this work certainly do not present a complete picture. Functional behaviors of many mutants could be very complicated. However, the PFA bound structure is included in the analysis. Relevant discussion is expanded. Some potential experiments to validate the findings have been suggested. See above.

Referee 1: 2. Another problem with DNA polymerase ternary structures is that there is little information about the DNA polymerase binding to the template strand. Typically, the template overhang is short and what is observed in structures is often distorted. Yet, the DNA polymerase surely interacts with the template strand and these interactions are important for translocation.

The first section of Discussion is on this topic (P. 14-16; Fig. 9a; Movie S4). Without direct observation of the single-stranded template, the motions of the N-terminal and exo domains are the best indication on how the single-stranded template would move.

Referee 1: For example, archaeal family B DNA polymerases can detect uracil in the template that is 4 nucleotides ahead of the primer end.

This is a property unique to archaeal polymerases B as reported. A surface cavity of the N-terminal domain is responsible for uracil recognition (Fogg et al. Nat. Struct. Biol. 9, 922, 2002). This finding is consistent with the hypothesis that the single-stranded template binds in the Arg-Lys-rich groove (P. 15). However, without direct observation of a single-stranded template with a significant length, it is difficult to conclude any further.

Referee 1: In studies with 2-aminopurine, 2 aminopurine in the templating position is either flipped out of the template which produces fluorescence or is sandwiched between neighboring bases, which quenches fluorescence. A structure of the templating base in a flipped out position has not yet been observed (JBC (2002) 286:25246).

Indeed, a flipped out templating base has not been observed. However, either the fluorescence time courses may be explained differently, or the flipped out position is not stable enough to be captured by crystallography.

Referee 1: 3. With respect to proofreading, the author states that when the corrected primer returns back to the polymerase active center, the translocation step is bypassed or undone. This statement does not include observations that proofreading likely requires DNA polymerase dissociation, which means that pol-to-exo switching is not processive; however, return of the trimmed primer end to the polymerase active site is rapid and processive (JBC (1992) 267: 14157. However, maybe 2 wrong nucleotides can be removed by a processive mechanism.

It has been clarified at several places in the Introduction and the overview of Results that only processive proofreading is discussed here (P. 2, 6, and 7). Removal of multiple nucleotides is discussed in the section of A-rule (P. 18).

Referee 1: 4. Also with respect to proofreading, a beta hairpin structure in the RB69 DNA polymerase has been observed as a wedge that separates the primer strand from the template in exonuclease complexes (JBC(2006) 282:1432). However, the bet hairpin structure is largely missing from the family B yeast DNA polymerase epsilon; yet, both DNA polymerases proofread.

The β-hairpin in polymerase ε is relatively short, but not completely missing. Residues 403-415 can be considered as the β-hairpin. However, polymerase δ features an extra long β-hairpin consisting residues 433-455. The relevant sentence is modified for clarity (P. 15).

Referee 1: Most readers will need more background information about the SVD analysis method. Referee 3: That it is long and, maybe, not easily penetrated but that is probably necessary for its main goal. I have nothing against it and rather propose acceptance without any need for any revision.

Agreed. However, this is not the first paper on application of structural meta-analysis or an introduction to SVD. Background information on SVD and previous applications are cited where the readers can find information. I hesitate to lengthen the Methods even further.

Referee 2: Several grammar issues were noted in the text. Specific examples are listed below: Pg 2 ln 8-10. "In bacteriophage T7…" These sentence construction is somewhat confusing. Please reword.

Done.

Referee 2: Pg. 2 ln 33-36. "It is remarkable…" Confusing sentence. Please reword.

Done.

Referee 2: Pg. 3 last sentence. "…and reveals those embedded structural information." Grammar issue. Need to reword sentence.

Done.

Referee 2: The frequent use of first person throughout the text is somewhat unusual.

The use of first person has been reduced.

Referee 2: Pg. 11 ln 4. "The rest 60% rotation…" Consider replacing the word "rest" with "remaining" or something similar.

Done.

Referee 2: Pg 15 last sentence. "…four orders of magnitude less capable." Does the author mean "less active"?

Changed to active.

Referee 2: Pg 19 ln 40. "…nucleotides by about 10 folds." Should be "fold" (singular).

Done.

Referee 2: "The increased error rate of Y567A could be resulted from first that…" Grammar issues should be corrected.

Done.

Referee 2: Figure legend 2, pg 26 ln 8. "The motion of the fingers is isolate from those of…" Isolated?

Revised.