Structural basis of error-prone replication and stalling at a thymine base by human DNA polymerase

Kevin Kirouac

Corresponding author: Hong Ling, University of Western Ontario

Review timeline:

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submission date</td>
<td>06 January 2009</td>
</tr>
<tr>
<td>Editorial Decision</td>
<td>05 February 2009</td>
</tr>
<tr>
<td>Revision received</td>
<td>11 March 2009</td>
</tr>
<tr>
<td>Additional correspondence to author</td>
<td>23 March 2009</td>
</tr>
<tr>
<td>Editorial Decision</td>
<td>26 March 2009</td>
</tr>
<tr>
<td>Revision received</td>
<td>31 March 2009</td>
</tr>
<tr>
<td>Accepted</td>
<td>03 April 2009</td>
</tr>
</tbody>
</table>

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 05 February 2009

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are enclosed. As you will see, all three referees show interest in your work and are broadly in favour of publication. However, all three raise criticisms that would need to be addressed in a revised version of the manuscript before we could consider publication. I will not go into the details of all the comments of the reviewers, but in particular, referee 1 has some concerns as to the domain swapping experiments, and referee 2 highlights issues with regard to the effects of divalent ions on pol iota activity.

In the light of the referees' positive recommendations, I would therefore like to invite you to submit a revised version of the manuscript, addressing all the comments of all three reviewers. I should add that at we generally allow only a single round of revision. Therefore, acceptance of your paper will depend on your ability to fully answer the points raised by the referees.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE REPORTS:
Referee #1 (Remarks to the Author):

In this paper the authors elucidate the unusual properties of human polymerase (Pol) in replicating a template thymine. Much structural work has been done on this TLS polymerase previously, showing that it replicates template purines by inducing a Hoogsteen base pair. In this way, Pol is able to replicate purines accurately, while bypassing certain base lesions. In replicating template pyrimidines, however, Pol is highly error-prone. Most notably, Pol preferentially incorporates a G opposite a template T or U and stalls at template Ts. In this study, Kirouac and Ling investigate the basis for T-G misincorporation, as well as T-stalling by Pol. The authors solved the crystal structures of Pol in complex with three substrate analogs, which mimic incorporation of A, T or G opposite a template T. The structures reveal an unusual U-turn in the single stranded portion of the template after the template T, which is proposed to be the basis for the signature T-stall of Pol.

Next, the interactions of the "finger" domain with the "little finger" (LF) domain and DNA are shown to be responsible for determining substrate specificity. The authors explain the preference for the T-G mismatch by the differences in stability of the 3 base pairs studied (TG, TA and TT). Finally, domain swapping experiments are used to test whether the "finger" domain controls replication specificity and T-stalling. This domain is found to exert a major influence on specificity, but only a minor effect on T-stalling. The crystallographic portion of this manuscript is extremely interesting and elegant, offering detailed mechanistic insights into this enigmatic activity of Pol.

However, I found the domain swapping experiment difficult to follow. Finally, while Pol is a functionally versatile enzyme, currently the biological relevance of the phenomenon studied in this work is currently unknown.

Specific comments:
1. The LF domain has been referred to in related papers as PAD. Perhaps that could be mentioned here.
2. In the introduction, the authors may want to expand on biological roles of Pol (tumor suppression, protection against oxidative & UV damage).
3. p. 8 Numbering of nucleotides in single-stranded portion of template is confusing. At one point it is referred to as "preceding" the template T, but the nucleotides are numbered +1, +2, +3. Perhaps they can be renumbered as -1, -2, -3.
4. Figure 2B, the label for Y61 is light colored.
5. In discussing the role of the T methyl group in the template U-turn, I would like to see contact distances (particularly between the methyl group and the backbone O atoms). Also, the authors don’t mention if the stall is observed with template U.
6. On p. 9, line 10: "-sheets 2 ..." should read "-strands 2 ..."
7. In the domain swapping experiments, the gels show ladders of bands in many lanes. The authors do not explain this. Is this because of the low processivity of Pol?
8. The authors do not address the appearance of multiple bands in the A/T/G/C lanes. It appears that some nucleotides are being misincorporated by all 3 enzymes, which is not always internally consistent. For instance, Fig. 5A, Dpo4 (Tepml T/A lane) -- why does lane A have 2 bands? If we look at panel D, dATP is not incorporated opposite template A.
9. pg 16 - the authors suggest there is significant C:C misincorporation, but I did not see this in the data in Fig. 5D.

Referee #2 (Remarks to the Author):

© European Molecular Biology Organization
This article reports 3 new structures of a pol iota as a complex with a primer containing a thymine in the 3’ position, paired with correct or incorrect incoming nucleotides.

The article is well written and contains new results and an interesting discussion.

However, I feel that the story is not completely finished.

1. The authors state that the typical structure found, a "U-turn" is stabilized by the methyl group of the thymine template. This should be verified by looking at the activity in the presence of the same oligonucleotide containing a dU instead of the dT.

2. In the introduction (p. 4), the authors mention the interesting study of Frank and Woodgate, 2007) on the important effect of Mn++ onto pol iota activity.

Curiously, the present article contains no discussion of this effect at all. Furthermore, it appears from Fig. 4 that the divalent ions (green dots? The legend of the Figure is incomplete in this respect) occupy different positions in the three different complexes. This should be discussed in the text.

Also, Table I should contain an entry describing how many divalent ions were assigned in the density, and if they were Mg++ or Ca++ (and why). Indeed the T:ddADP crystal was obtained in the presence of 150mM Ca++ (p. 21), whereas T:dGTP and T:dTTP crystals were not (they also crystallize in a different space group). Mg++ is present in all cases.

Altogether, one would very much wish the authors had collected additional data sets in the presence of 10 mM Mn++, just to see if the general conclusions are maintained. At the very least, one would like to see a primer extension test (as in Fig. 5a,b,c,d) in the presence of 150 mM Ca++.

Referee #3 (Remarks to the Author):

Objective: The authors sought to explain why human DNA polymerase exhibits such high levels of mis-insertion opposite thymidine template residues using x-ray crystallography and domain swapping approaches. The highly-error prone nature of pol nucleotide incorporation opposite template pyrimidines is unusual, even compared with other Y-family DNA polymerases. Previous structural studies have shown that pol uses Hoogsteen base pairing during dNTP insertion opposite guanosine or N2-adducted guanosine. In those studies a narrowed C1'-C1' distance was proposed to influence the syn orientation observed for the purine template, but it remains unclear from a mechanistic perspective why pol loses its fidelity when inserting opposite template pyrimidines.

Comments: The structures reported in this work are of high interest to the polymerase field and are of excellent quality (based on the reported statistics). They serve to explain the mechanism of pol misincorporation/stalling at pyrimidine template bases. The uniqueness of the mechanism apparently lies in the ability of pol to stabilize the "U-turn" template DNA conformation only when a pyrimidine (or at least thymine) is present in the enzyme active site. The reported structures are quite similar to ones previously published by the Aggarwal and Perrino groups. I have essentially one major comment, which should not prohibit publication of the article.

One of the major conclusions drawn from the current work is that the "U-turn" is unique to pol AND a template thymidine. However, unless I am mistaken, none of the other pol structures utilize a template with the six base ssDNA overhang, such as the one observed here. Comparisons with Dpo4 structures (where the ssDNA overhang is usually four or five bases) and sequence-based alignments are supportive of the uniqueness of the "U-turn" to pol, but I do not think the authors can state definitively that the "U-turn" does not occur when a template purine (or cytosine) is present in the same sequence. Saying that purines "are too large" is one thing, but I'm not convinced that this is true. The purines would be expected to adopt a syn orientation, placing the bulk of the ring-system into the major groove side of the active site, but this may not prohibit the "U-turn". Hydrogen bonding between the exocyclic amino group of dG and the phosphate backbone at positions +1 and
+2 might serve to stabilize the "U-turn". In fact, Fig. 5C shows that pol stalls at similar points whether T, G, or C is present at the 0 position (albeit to a lesser extent for G and C), which is qualitatively indicative of similar mechanisms for these three template bases. In either case, there is no conclusive evidence to support the view that the "U-turn" is exclusive to thymidine.

It is also unclear how the contacts between the template T methyl group and the phosphate backbone are important for stabilizing the "U-turn". The authors draw some dotted lines between O5' and the C6 atom of template T, as well as between several oxygen atoms and the C5-Me group of T, but I'm not sure what kind of "hydrophobic interactions" occur between oxygen and carbon. Additionally, the absence of the methyl group might open a space to allow greater bending at the "U-turn". I would place more emphasis upon the conserved interactions between the ssDNA and the different domains.

The authors should also comment specifically on how the "U-turn" might result in the reduced catalytic efficiency observed for dNTP insertion opposite pyrimidines. For example, is there evidence in the literature for decreased dNTP binding affinity when template T or C is present? Decreased kpol values? I believe that there is at least some evidence for the latter (see Washington, M.T. 2004 Mol. Cell. Biol. & compare template dT with dA and with Choi, J-Y. 2006 JBC, template dG).

Overall, I think that this is an excellent study and worthy of publication in The EMBO Journal because it is of general interest to the DNA replication field and it provides much needed insight into why pol is highly error-prone during insertion opposite template pyrimidines. However, I also think that the authors should state their conclusions in a more cautious manner and reconsider some of their assertions. Specifically, they might consider that the "U-turn" does occur when cytosine is present b/c of the reduced catalytic efficiency and highly-error prone nature of pol insertion opposite template, as well as the similar extension properties observed from dA:dT and dG:dC primer:template pairs (see Vaisman, A. 2001 JBC).

Minor Comments:

1. The term "substrate recognition site" is vague. I believe that the authors are trying to indicate the region contacting the nascent base pair, but it could mean other things (e.g. Tyr39 check upon deoxyribose moiety of incoming dNTP) and should be clarified.

2. Calling pol "exemplary" is not a good choice of words. It is unusual in its mode of dNTP selection but it is also the most error-prone of the Y-family members, so I wouldn't call it "exemplary".

3. I would not say that pol incorporation opposite purines is "quite accurate". Pol only favors dCTP over dTTP by ~50-fold during insertion opposite G. Just say it's "more accurate" during insertion opposite purines than it is opposite pyrimidines.

4. Throughout: a space needs to be added between last word in the sentence and reference; pg. 10, near bottom purine is misspelled; pg. 10 near bottom, don't use the possessive form when describing an enzyme (i.e. "enzyme active site" not "enzyme's active site").

5. pg. 11-12: ".which has not been observed in other pol structures THAT contain purine bases in the active site (Fig. 4C)." Also, there should be a reference to the "other structures", and I don't think any of those other structures have an incoming dGTP so it's misleading to say purines. Just say, adenine or dATP.

6. I believe that it is correct to italicize the atom if you are using superscript representation.

7. Check the methods section: spaces need to be added throughout (e.g. 0.2mM needs to be 0.2 mM etc...).
Thank you for the opportunity to revise our manuscript for your consideration for acceptance by The EMBO Journal and for the enclosed comments pertaining to our manuscript. The referees comments and suggestions were greatly appreciated and we thank them for their time and effort to improve this manuscript.

Referee #1

In this paper the authors elucidate the unusual properties of human polymerase (Pol) in replicating a template thymine. Much structural work has been done on this TLS polymerase previously, showing that it replicates template purines by inducing a Hoogsteen base pair. In this way, Pol is able to replicate purines accurately, while bypassing certain base lesions. In replicating template pyrimidines, however, Pol is highly error-prone. Most notably, Pol preferentially incorporates a G opposite a template T or U and stalls at template Ts. In this study, Kirouac and Ling investigate the basis for T-G misincorporation, as well as T-stalling by Pol. The authors solved the crystal structures of Pol in complex with three substrate analogs, which mimic incorporation of A, T or G opposite a template T. The structures reveal an unusual U-turn in the single stranded portion of the template after the template T, which is proposed to be the basis for the signature T-stall of Pol. Next, the interactions of the "finger" domain with the "little finger" (LF) domain and DNA are shown to be responsible for determining substrate specificity. The authors explain the preference for the T-G mismatch by the differences in stability of the 3 base pairs studied (TG, TA and TT). Finally, domain swapping experiments are used to test whether the "finger" domain controls replication specificity and T-stalling. This domain is found to exert a major influence on specificity, but only a minor effect on T-stalling. The crystallographic portion of this manuscript is extremely interesting and elegant, offering detailed mechanistic insights into this enigmatic activity of Pol. However, I found the domain swapping experiment difficult to follow. Finally, while Pol is a functionally versatile enzyme, currently the biological relevance of the phenomenon studied in this work is currently unknown.

Specific comments:

1. The LF domain has been referred to in related papers as PAD. Perhaps that could be mentioned here.

Yes, we have added a sentence in the introduction referring to this nomenclature:
"Y-family polymerases possess a unique C-terminal domain, called the "little finger" or polymerase associated domain (PAD)" (p.3)

2. In the introduction, the authors may want to expand on biological roles of Pol (tumor suppression, protection against oxidative & UV damage).

Yes, we agree that more background information on the biological roles of pol would be beneficial to the readers. Therefore, we have expanded this section in the introduction (p.5), specifically on the roles of suppressing UV-light induced tumors and protection against oxidative damage:
"Mice deficient in both pol and pol have an earlier onset on UV-light induced tumors than pol deficiency alone (Dumstorf et al., 2006; Ohkumo et al., 2006), indicating a role for pol in UV-induced lesion bypass. In addition, it has recently been observed that pol plays a significant role in cellular protection from oxidative damage (Petta et al., 2008). Although pol likely facilitates the repair of oxidative DNA lesions, the specificity and mechanism of this repair is unknown."

3. p. 8 Numbering of nucleotides in single-stranded portion of template is confusing. At one point it is referred to as "preceding" the template T, but the nucleotides are numbered +1, +2, +3. Perhaps
they can be renumbered as -1, -2, -3.

The standard nomenclature for numbering template nucleotides within a DNA strand is; 0 for the replicating template, -1, -2, -3, etc. for bases are replicated before the replicating base and +1, +2, +3, etc. for bases are to be replicated after the replicating base. We understand how the word "preceding" could be misinterpreted and therefore have changed it to "downstream" (p. 10).

4. Figure 2B, the label for Y61 is light colored.

Yes, the figure has been relabeled in black.

5. In discussing the role of the T methyl group in the template U-turn, I would like to see contact distances (particularly between the methyl group and the backbone O atoms). Also, the authors don't mention if the stall is observed with template U.

Yes, contact distances have been added for the methyl group and backbone O atoms in figure 2D. We thank the referee for bringing up the question of template U stalling. We have added in the introduction that template U stalling has been observed for pol similar to template T (p. 4):

"A similar pattern of misincorporation and replication stalling by pol is observed opposite template uracil (U) (Vaisman and Woodgate, 2001)."

We have also included this point within the results section (p. 9):

"Although the extensive contacts between this unique methyl group and the +1 nucleotide reinforce the unusual bending template, the methyl group may not be an absolute requirement, due to a similar stalling effect opposite template U (Vaisman and Woodgate, 2001)"

6. On p. 9, line 10: "-sheets 2 ..." should read "-strands 2 ...

Yes, this has been changed as noted (p.10)

7. In the domain swapping experiments, the gels show ladders of bands in many lanes. The authors do not explain this. Is this because of the low processivity of Pol?

Yes, this is due to the low processivity of the enzymes. This has been mentioned within the results section for clarity (p. 15):

"Multiple bands are observed due to the low processivity of these enzymes."

8. The authors do not address the appearance of multiple bands in the A/T/G/C lanes. It appears that some nucleotides are being misincorporated by all 3 enzymes, which is not always internally consistent. For instance, Fig. 5A, Dpo4 (Tepml T/A lane) -- why does lane A have 2 bands? If we look at panel D, dATP is not incorporated opposite template A.

Same as above; the multiple bands are a result of the low processivity of these enzymes, and also the results of misincorporation. The misincorporated bands are not internally consistent because the sequence context of the template DNA changes. Misincorporation by Y-polymerases is sequence dependent. Therefore, different sequence contexts at the primer-template junction result in different incorporation efficiencies. For Dpo4, the sequence dependence is commonly achieved by strand misalignment, due to +1/-1 frameshift mutations observed in our previous structure work (Ling et al, 2004a, Bauer et al, 2007). In strand misaligning, the bases 3i or 5i to the replicating base can function as templates for the misincorporation. In Fig. 5A, Dpo4 has two bands for lane A. The first band is Dpo4 incorporating A opposite template T and the second band is Dpo4 misincorporating A opposite template A. It seems the second misincorporation is achieved by misaligning the template DNA and utilizing the 3i T as the replicating base. Strand misalignment fits well for the misincorporation of A in 5A. In Fig. 5D, Dpo4 has very little misincorporation of A bases opposite template C. This may be due to the lack of cognate template bases around the template C. More likely, other factors, such as the stability of local base pairs (matched/mismatched) in the active site of Dpo4 are contributing to the misincorporation rate at this position, as all the misincorporations are restricted with this sequence by Dpo4 compared to
poli.

9. pg 16 - the authors suggest there is significant C:C misincorporation, but I did not see this in the data in Fig. 5D.
We agree that the word "significant", could mislead the readers as to the extent of misincorporation opposite template C. Therefore, we have changed this description to: "Higher misincorporations of T:T, T:C, C:T and C:C were observed in our primer extension assays for both poli and Dpo4-i relative to misincorporations opposite template purines" (p.18).

10. There are some minor typographical errors in the text.

We have gone through the manuscript and corrected the typographical errors.

Referee #2

This article reports 3 new structures of a pol iota as a complex with a primer containing a thymine in the 3' position, paired with correct or incorrect incoming nucleotides. The article is well written and contains new results and an interesting discussion. However, I feel that the story is not completely finished.

1. The authors state that the typical structure found, a "U-turn" is stabilized by the methyl group of the thymine template. This should be verified by looking at the activity in the presence of the same oligonucleotide containing a dU instead of the dT.

We thank the referee for bringing up the point of template dU stalling. After reviewing the literature, it was found that template dU stalling has been reported for pol (Vaisman and Woodgate, 2001). Therefore, it will not be necessary to repeat these experiments. We have included this observation in the introduction and incorporated these results into our structural interpretation. (p. 4):
"A similar pattern of misincorporation and replication stalling by pol is observed opposite template uracil (U) (Vaisman and Woodgate, 2001)."
(p. 9):
"Although the extensive contacts between this unique methyl group and the +1 nucleotide reinforce the unusual bending template, the methyl group may not be an absolute requirement, due to a similar stalling effect opposite template U (Vaisman and Woodgate, 2001)"

2. In the introduction (p. 4), the authors mention the interesting study of Frank and Woodgate, 2007) on the important effect of Mn2+ onto pol iota activity. Curiously, the present article contains no discussion of this effect at all.

We agree that the effect of Mn2+ on pol activity is an important observation that is relevant to this manuscript and thus deserves attention. Therefore, we have included in the discussion, how Mn2+ may affect pol activity by comparing this observation with our current three structures. (p.18):
"Interestingly, in the presence of Mn2+, pol has increased fidelity opposite template thymine with a preference of incorporating the correct A nucleotide instead of incorrect G (Frank and Woodgate, 2007). The Mn2+ ion has a more relaxed and mobile coordination than Mg2+ within the active site of polymerases. This effect likely allows the incoming nucleotide to adopt a variety of conformations that would not be possible with Mg2+. In this manner, Mn2+ ion coordination by pol may render a favourable interaction that selects A over G opposite template T."

Furthermore, it appears from Fig. 4 that the divalent ions (green dots? The legend of the Figure is incomplete in this respect) occupy different positions in the three different complexes. This should be discussed in the text.

Yes, the figure legend has been changed to specify the divalent ions (green dots). We thank the referee for bringing up the point of divalent metal ion positions. We agree that this point should be addressed and we have therefore mentioned it within the results section (p.6-7). A supplemental figure has been made to show metal positions clearly with structural superimposition (Figure S2).
The metal ion positions in the site ëBí are identical in all three structures (T:ddADP, T:dTTP, T:dGTP), previous pol structures containing template purines (Nair et al, 2006) and Dpo4 ternary structures (Ling et al, 2001; Vaisman et al, 2005), but metal positions in the site A are variable, which is a general structural feature of Y family polymerases.

Also, Table I should contain an entry describing how many divalent ions were assigned in the density, and if they were Mg2+ or Ca2+ (and why). Indeed the T:ddADP crystal was obtained in the presence of 150mM Ca2+ (p. 21), whereas T:dGTP and T:dTTP crystals were not (they also crystallize in a different space group). Mg2+ is present in all cases.

Yes, the number and identity of the divalent metal ions were included in Table I and mentioned within the results section (p.6):

"In the T:ddADP structure the active site metal ions have been refined as Ca2+, due to the presence of 150mM CaCl2 within the crystalization buffer and the high electron density. The active site Ca2+ ions were verified by generating an anomalous signal map for calcium. This is analogous to Dpo4 structures crystallized with 100mM Ca(AC)2 (Ling et al., 2001; Wong et al., 2008). Both T:dTTP and T:dGTP structures contain two active site Mg2+ ions as no calcium salt was added in the crystalization buffer."

 Altogether, one would very much wish the authors had collected additional data sets in the presence of 10 mM Mn++, just to see if the general conclusions are maintained.

Although, a crystal structure of pol containing Mn2+ would be of great interest and may further our understanding of pol function, such an experiment would be extremely difficult to attempt for a number of reasons. First, both crystalization conditions contain high amounts of either CaCl2 (150mM) or MgSO4 (200mM). This would cause the Mn2+ ion to be replaced with either Ca2+ or Mg2+, as was the case for the T:ddADP structure. This would prevent the desired metal ion from crystallizing in the complex. Second, Mn2+ ion binding within Y family polymerases is notoriously mobile or ëslipperyí, resulting in crystal instability. This would prevent crystalization with the Mn2+ ions. We indeed tried to add Mn2+ ions in the crystalization trials without Ca2+ and Mg2+ but failed to generate stable crystals so far.

At the very least, one would like to see a primer extension test (as in Fig. 5a,b,c,d) in the presence of 150 mM Ca2+.

We understand the referees concern regarding Ca2+ within the crystalization buffer, and Ca2+ ion coordination within the active site of the T:ddADP structure. Therefore, we preformed primer extension tests with pol in the presence of 150 mM CaCl2 to observe the effects of Ca2+ ions. As expected, the presence of 150 mM CaCl2 only reduces the catalytic efficiency, while the nucleotide incorporation specificity is identical (Figure S1, a supplementary figure for the assay results). Particularly, opposite template T, the most preferred incoming nucleotide is incorrect G, for both 5 mM MgCl2 and 150 mM CaCl2 conditions (Figure S1-B). The Ca2+ ion coordination with the T:ddADP structure is the same as Mg2+ ion coordination in the other two structures (T:dTTP and T:dGTP, crystallized without Ca2+) and does not change the conformation of the active site. Therefore, the T:ddADP structure should represent the same replication mechanism as the other two structures in this study. We have added the following sentence in page 6: 

"Primer extension assays have been preformed on pol in the presence of 150 mM CaCl2, which demonstrates that Ca2+ ions do not change the nucleotide incorporation specificity of pol (Figure S1)."

Referee #3

Objective: The authors sought to explain why human DNA polymerase exhibits such high levels of mis-insertion opposite thymidine template residues using x-ray crystallography and domain swapping approaches. The highly-prone nature of pol nucleotide incorporation opposite template pyrimidines is unusual, even compared with other Y-family DNA polymerases. Previous structural studies have shown that pol uses Hoogsteen base pairing during dNTP insertion opposite guanosine or N2-adducted guanosine. In those studies a narrowed C1'-C1' distance was proposed to
influence the syn orientation observed for the purine template, but it remains unclear from a mechanistic perspective why pol loses its fidelity when inserting opposite template pyrimidines.

Comments: The structures reported in this work are of high interest to the polymerase field and are of excellent quality (based on the reported statistics). They serve to explain the mechanism of pol misincorporation/stalling at pyrimidine template bases. The uniqueness of the mechanism apparently lies in the ability of pol to stabilize the "U-turn" template DNA conformation only when a pyrimidine (or at least thymine) is present in the enzyme active site. The reported structures are quite similar to ones previously published by the Aggarwal and Perrino groups. I have essentially one major comment, which should not prohibit publication of the article.

One of the major conclusions drawn from the current work is that the "U-turn" is unique to pol AND a template thymidine. However, unless I am mistaken, none of the other pol structures utilize a template with the six base ssDNA overhang, such as the one observed here. Comparisons with Dpo4 structures (where the ssDNA overhang is usually four or five bases) and sequence-based alignments are supportive of the uniqueness of the "U-turn" to pol, but I do not think the authors can state definitively that the "U-turn" does not occur when a template purine (or cytosine) is present in the same sequence. Saying that purines "are too large" is one thing, but I'm not convinced that this is true. The purines would be expected to adopt a syn orientation, placing the bulk of the ring system into the major groove side of the active site, but this may not prohibit the "U-turn". Hydrogen bonding between the exocyclic amino group of dG and the phosphate backbone at positions +1 and +2 might serve to stabilize the "U-turn".

The statement that the "U-turn" does not occur with template purines is not a postulation, but rather an experimental observation. We used two types of DNA substrates in the co-crystallization, both show eU-turn at the T base. One of them (in the structures T:dTTP and T:dGTP) is identical to that of the previous purine template structures in size. Obviously we had not made it clear to readers. Therefore, we have added the following statements in the results section to clarify this point (p.9): "The unique eU-turn conformation is not observed in the presence of template purine bases (Nair et al., 2005; Nair et al., 2006). Two of our template thymine structures (T:dTTP and T:dGTP) have the same sized DNA substrate with an identical three base ssDNA overhang as that in the previous purine template structures and isomorphous to the previous purine template structures. Yet, our template thymine structures adopt the eU-turn conformation, similar to the T:ddADP structure which has a longer ssDNA overhang (six-base ssDNA overhang) and a distinct crystal form (Figure 1A). This excludes any structural variations caused by differences in the ssDNA overhang size and packing environments of the complexes in the crystals."

Therefore, we believe that the "U-turn" conformation does not occur with a purine template is a structural observation.

In the case of a "U-turn" conformation in the presence of template C, we agree with the referee that we cannot say definitively that this conformation does not occur, due to lack of structural information. Therefore, we have added a postulation to what might occur to DNA with a template C base (p.10):

"Although template C has a similar size to T and U, which may lead itself to an eU-turn structure, such a conformation may not be stable when template C is in the active site, as no significant replication stalling has been observed with this template base."

In reference to our statement that purine nucleotides are too large to accommodate the "U-turn" conformation, we acknowledge that this explanation is likely oversimplified. There are likely other factors, which are also involved in preventing the "U-turn" conformation. Therefore, we have toned down this statement in our results section (p. 9):

"It appears that the purine bases A and G are too large to be accommodated in the "U-turn" conformation observed in our structures, which would disrupt bending in the major groove. However, there may be other unidentified factors also involved in preventing this conformation."

In fact, Fig. 5C shows that pol stalls at similar points whether T, G, or C is present at the 0 position (albeit to a lesser extent for G and C), which is qualitatively indicative of similar mechanisms for these three template bases.
We are not sure what 'stalls' the referee is referring to in Fig.5C. The stalling effects were only shown for poli in the reactions at the 8th position in the second lanes of every panel (with horizontal arrows pointing to) where all four dNTPs are present. The limited extensions of the T/G/C lanes in Fig5C are due to low extension efficiencies of misincorporations. The stalling effects of poli are well documented. We only wanted to demonstrate that non-poli polymerases (e.g. Dpo4) do not have this thymine stalling feature.

In either case, there is no conclusive evidence to support the view that the "U-turn" is exclusive to thymidine.

Yes, we agree that there is no conclusive evidence to support that a "U-turn" does not occur opposite template C. We have discussed this point in the previous paragraph. Since the base at the 0 position is the only base component of the εU-turn structure and involved in εU-turn interactions, the εU-turn structure should be, at least, template base identity dependent.

It is also unclear how the contacts between the template T methyl group and the phosphate backbone are important for stabilizing the "U-turn". The authors draw some dotted lines between O5' and the C6 atom of template T, as well as between several oxygen atoms and the C5-Me group of T, but I'm not sure what kind of "hydrophobic interactions" occur between oxygen and carbon.

The referee is right. As shown in Figure 2D, the interactions are a mixture of hydrophobic contacts and some backbone O atoms in the vicinity of the methyl group. We deleted the words "hydrophobic interactions" in page 9, as the interactions are not purely hydrophobic. "All the contact distances are between 3.40 and 4.45 ≈ with the distances of C5-Me group to the backbone oxygen atoms labeled" was added to the figure legend for Fig. 2D.

Additionally, the absence of the methyl group might open a space to allow greater bending at the "U-turn". I would place more emphasis upon the conserved interactions between the ssDNA and the different domains.

Yes, we agree that the methyl group is likely not essential for the 'U-turn' conformation and we are focusing more upon the conserved pol interactions with the ssDNA. We have included this into our results section (p.9)
"Although the extensive contacts between this unique methyl group and the +1 nucleotide reinforce the unusual bending template, the methyl group may not be an absolute requirement, due to a similar stalling effect opposite template U (Vaisman and Woodgate, 2001)."
We have also removed comments within the discussion that emphasized the methyl group interaction (p. 19) to tune down the importance of the methyl group:
"Without the methyl group, other DNA bases will not maintain the same stable εU-turn template as the T base located at the template-primer junction at the active site." n REMOVED In addition, the subheading: "Replication stalling is T specific and stabilized by conserved residues over three domains" (p.18) within the discussion, has been changed to: "Replication stalling is stabilized by conserved residues over three domains"

The authors should also comment specifically on how the "U-turn" might result in the reduced catalytic efficiency observed for dNTP insertion opposite pyrimidines. For example, is there evidence in the literature for decreased dNTP binding affinity when template T or C is present? Decreased kpol values? I believe that there is at least some evidence for the latter (see Washington, M.T. 2004 Mol. Cell. Biol. & compare template dT with dA and with Choi, J-Y. 2006 JBC, template dG).

We thank the referee for their insightful comments on the possible link between the "U-turn" conformation and reduced catalytic efficiency. The first paper mentioned by the referee (Washington et al, 2004) states that dGTP binding affinity opposite template T is greater than dTTP or dATP and is the same for dTTP binding affinity opposite template A. This reaffirms our structural observation that incoming dGTP opposite template T is bound in a highly stable conformation. This has been included in the results section (p.14)
"The structural observation that incoming dGTP is the most stable incoming nucleotide opposite T
is supported by the observation that dGTP binding affinity opposite template T by pol is greater than dTTP or dATP and is the same for dTTP binding affinity opposite template A."

Although there is no decreased dNTP binding affinity reported, there is a decreased dNTP incorporation rate opposite template T compared to template A. As the referee stated, this could be the result of the "U-turn" conformation, and we have included this with the discussion section (p.19):

"In addition, the highly bent DNA may also reduce the catalytic efficiency of pol, as the observation that dNTP incorporation opposite template T is much slower than opposite template A (Washington et al., 2004)."

The second paper mentioned (Choi et al, 2006) has dNTP binding affinities opposite template G, which are compared to lesion G templates. Although this study is interesting, it provides no information on dNTP efficiencies for template pyrimidines, and is therefore of little help in solidifying or re-evaluating our conclusions.

Overall, I think that this is an excellent study and worthy of publication in The EMBO Journal because it is of general interest to the DNA replication field and it provides much needed insight into why pol is highly error-prone during insertion opposite template pyrimidines. However, I also think that the authors should state their conclusions in a more cautious manner and reconsider some of their assertions. Specifically, they might consider that the "U-turn" does occur when cytosine is present b/c of the reduced catalytic efficiency and highly-error prone nature of pol insertion opposite template, as well as the similar extension properties observed from dA:dT and dG:dC primer:template pairs (see Vaisman, A. 2001 JBC).

Yes, we have changed our tunes and stated our conclusions in a more cautious manner and have reconsidered some assertions; specifically that the "U-turn" conformation is the only factor for stalling. We tuned down the statements regarding the 'U-turn' as described above. We have also stated that the ëU-turní may occur on cytosines as described above. These points are noted in the results section (p. 9-10).

Minor Comments:

1. The term "substrate recognition site" is vague. I believe that the authors are trying to indicate the region contacting the nascent base pair, but it could mean other things (e.g. Tyr39 check upon deoxyribose moiety of incoming dNTP) and should be clarified.

Yes, we were referring to the region contacting the nascent base pair. In order to clarify this point we have changed the sentence to: "The finger domain contacts the replicating base pair and is therefore the substrate recognition site". (p.3)

2. Calling pol "exemplary" is not a good choice of words. It is unusual in its mode of dNTP selection but it is also the most error-prone of the Y-family members, so I wouldn't call it "exemplary".

Yes, "exemplary" has been removed (p.3)

3. I would not say that pol incorporation opposite purines is "quite accurate". Pol only favors dCTP over dTTP by ~50-fold during insertion opposite G. Just say it's "more accurate" during insertion opposite purines than it is opposite pyrimidines.

Yes, we agree that "quite accurate" may mislead readers and therefore have changed it to "more accurate". (p.4)

4. Throughout: a space needs to be added between last word in the sentence and reference; pg. 10, near bottom purine is misspelled; pg. 10 near bottom, don't use the possessive form when describing an enzyme (i.e. "enzyme active site" not "enzyme's active site").

Yes, spaces have been added, purine spelling has been corrected, and "enzymeís" has been changed to "enzyme".
5. pg. 11-12: "which has not been observed in other pol structures THAT contain purine bases in the active site (Fig. 4C)." Also, there should be a reference to the "other structures", and I don't think any of those other structures have an incoming dGTP so it's misleading to say purines. Just say, adenine or dATP.

Yes, "which" has been changed to "that" and references have been added. We agree that saying "purines" alone could be misleading, so we have changed it to "template purines", because we were not referring to incoming purine nucleotides but rather the conformation of template purines within the active site.

6. I believe that it is correct to italicize the atom if you are using superscript representation.

We believe the referee is referring to the description of hydrogen bonding in the results section (p.14) where exocyclic base atoms are denoted as N2, N6, O2, and O4. We used to have all the base atoms italic. It seems now that the base atoms are in normal letter in many literatures. Italicizing the atom is only used when describing modifications to specific atoms, for instance, a methyl group added to the O2 atom would be represented as: 2O$_2$-methyl. We could not find the consensus by checking the literatures, thus italicized all the base atom names in the manuscript following the referee's suggestion.

7. Check the methods section: spaces need to be added throughout (e.g. 0.2mM needs to be 0.2 mM etc...).

Yes, changed as noted. Thanks a lot for taking care of the details.

We feel the additional materials added and the modifications made to the manuscript have strengthened the paper and thank the reviewers for their helpful suggestions. We hope our work is now suitable for publication in The EMBO Journal and look forward to hearing from you.

Additional correspondence to author

23 March 2009

Thank you for sending the revised version of your manuscript. It has now been seen again by two of the original referees (#s 1 and 2). Referee 1 is now fully happy with the paper and supports publication without further revision. However, referee 2, while supportive of publication, still has concerns regarding the assignment of calcium in the active site (see comments below), and has requested to see the original data for this. Before we can make a final decision on your manuscript, I would therefore ask if you could send the original data: the anomalous map to which you refer. It is not clear to me that you will need to modify the manuscript and I hope that, provided referee 2 is happy with the data, we will be able to accept the manuscript. Therefore, I suggest the easiest solution would be for you to send me the data by e-mail, rather than uploading it through the system. Should you need to make any changes at a later stage, you can then do this via a formal revision.

I look forward to receiving the data from you, and please do not hesitate to contact me should you have any further questions.

With best wishes,

Editor
The EMBO Journal
REFEREE REPORTS:

Referee #1 (Remarks to the Author):

I am satisfied with the authors' revisions. This article will be a highly informative addition to the field.

Referee #2 (Remarks to the Author):

I am satisfied by all corrections except for one concerning the assignment of Ca++ in the active site (p. 5)

At what wavelength were the data collected for the anomalous map used by the authors to assign Ca++ ions?
I would like to see this (anomalous) map contoured at 3 sigmas.

2nd Editorial Decision 26 March 2009

Many thanks for sending the anomalous map, which has now been seen by referee 2, whose comments are enclosed below. He/she remains unconvinced that this anomalous map conclusively demonstrates the presence of Ca++ at the active site. As this is not a central issue in the paper, neither the referee nor we feel it should preclude publication of your manuscript. However, we would ask that you revise the text to tone down your claim as to the assignment of Ca++ at the active site - as suggested in point 5 below.

I would therefore like to invite you to submit a revised version of your manuscript incorporating these minor changes. Please can you also make sure to include the PDB accession codes for the structures? Once we have the final revision, we should be ready to accept with no further delay.

I look forward to receiving your revision.

Yours sincerely,

Editor
The EMBO Journal

Referee 2 comments:

After looking at the anomalous map my impressions go like this:
1. The anomalous map indeed shows some peaks above 3.3 sigmas where Mg++ are expected.
2. However, the expected anomalous signal for Ca++ at this wavelength (0.9832 Angstrom) is quite low (about 0.6 electrons), so it is not certain at all that it is indeed Ca++ that is causing these peaks. In other words, the anomalous map is not a proof of the identity of the anomalous scatterer, unless one varies the wavelength and collect data at the absorption edge (3.070 Angstrom for Ca++).
However, this range of wavelengths is not easy to reach at synchrotron beamlines.
Florescence spectra of crystals might be an option but the crystals should be washed extensively with a solution without Ca++ before doing it and it is not certain they will stand it.
3. So what might cause these peaks? One possibility is that either Ni++ or Co++ were copurified during the NTA column elution (Protein Preparation, Materials and Methods).
In this case, one should also see this kind of peaks in the two other complexes crystallized in the absence of Ca++. Could the authors also have a look at the corresponding anomalous maps?
4. The issue of knowing with certainty the identity of these anomalous scatterers is interesting but not central for the paper.
My concern is simply that if these divalent ions are deposited in the
PDB without any warning, people interested in metal coordination will probably consider this PDB file as an outlier. Also, some polymerases are inhibited by Ca++, but this does not seem to be the case for this one (p. 5 of author's answers).

5. One option would be that the authors change their phrasing about the assignment of these peaks to Ca++ and state that the assignment was based on the presence of peaks in the anomalous map (give the wavelength) AND the estimated relative concentration of divalent ions in mother liquor, or something along these lines.

Referee 2

After looking at the anomalous map my impressions go like this:

1. The anomalous map indeed shows some peaks above 3.3 sigmas where Mg++ are expected. There are indeed anomalous peaks present at the active site metal ion coordination sites, indicating a metal ion other than Mg++ is present.

2. However, the expected anomalous signal for Ca++ at this wavelength (0.9832 Angstrom) is quite low (about 0.6 electrons), so it is not certain at all that it is indeed Ca++ that is causing these peaks. In other words, the anomalous map is not a proof of the identity of the anomalous scatterer, unless one varies the wavelength and collect data at the absorption edge (3.070 Angstrom for Ca++).

We agree with the Referee that simply observing this anomalous signal does not enable verification of the atom producing the anomalous scatter. We have therefore tuned down our statement of calcium verification. See response to question 5.

However, this range of wavelengths is not easy to reach at synchrotron beamlines. Florescence spectra of crystals might be an option but the crystals should be washed extensively with a solution without Ca++ before doing it and it is not certain they will stand it.

We thank the Referee for their suggestion of using Florescence spectra. Although this would be an interesting experiment to perform, these crystals are quite fragile and such extensive washing would likely disintegrate the crystal.

3. So what might cause these peaks? One possibility is that either Ni2+ or Co2+ were copurified during the NTA column elution (Protein Preparation, Materials and Methods).

In this case, one should also see this kind of peaks in the two other complexes crystallized in the absence of Ca++. Could the authors also have a look at the corresponding anomalous maps? It is possible that Ni2+ could have been copurified with the protein during the Nickel column elution. However, an ion exchange column was run after the Nickel column, thus the concentration of Ni2+ would be minute compared to the concentrations of Mg2+ (10mM) during the complex formation and Ca2+ (100mM) in the crystallization buffer. The coordination atoms surrounding the metal ions are O atoms not N atoms, which make Mg/Ca ions more likely there with O coordination. In addition, the anomalous map generated for the T:dGTP and T:dTTP structures, which had the same protein sample and no Ca2+ in the crystallization buffer, showed no anomalous peaks on the Mg2+ ions (see figure below). Thus, it is most likely that the anomalous scatterer is Ca2+. We agree that the most defined experiments, Florescence spectra analysis, is not carried out to exclude other possibilities. We thus stated that "the active site metal ions have been refined as Ca2+." in the
manuscript (p. 6), instead of "the active site metal ions are Ca2+.") to leave a room for other possibilities.

4. The issue of knowing with certainty the identity of these anomalous scatterers is interesting but not central for the paper. My concern is simply that if these divalent ions are deposited in the PDB without any warning, people interested in metal coordination will probably consider this PDB file as an outlier. Also, some polymerases are inhibited by Ca++, but this does not seem to be the case for this one (p. 5 of author's answers).

We thank the Referee for his/her cautions regarding the divalent metal ions and agree that the identity of these ions should be stated cautiously; see above and below.

5. One option would be that the authors change their phrasing about the assignment of these peaks to Ca++ and state the assignment was based on the presence of peaks in the anomalous map (give the wavelength) AND the estimated relative concentration of divalent ions in mother liquor, or something along these lines.

We have tuned down the phrasing about anomalous signals as following (p. 6): We changed the sentence "The active site Ca2+ ions were verified by generating an anomalous signal map for calcium." to "In addition, anomalous signal peaks were observed at the metal ion sites which are distinct from surrounding non-metal atoms though weak at the A sites."

Additional notes:

Locating metal ions at the A sites of Y-family polymerases sometimes is a tough task as the metal ions are mobile at this site, especially for Mg2+ ions which are usually not distinguishable to water molecules. With accurate phase information from refined structures, however, we can use detectable anomalous peaks as reference points to location Ca ions at the A sites, in combination with coordination environment inspection. We agree that cautions must be taken to avoid errors in the structures which will be used by the community. We re-checked all three structures in this work for the metal sites and found no errors in the Ca2+ bound T:ddADP structure, but the Mg2+ at the A sites of T:dGTP and T:dTTP need adjustment. After adjustment, we re-made the Figure 4 and Figure 2S (see attached figures) with the Mg2+A positions changed only. The minor changes do not affect the structural description and conclusions in the manuscript but making the figures consistent with the coordinates deposited into the PDB database. We are very grateful to the referee and the editor for their time and efforts to improve this work.