

Snapshot of a DNA Polymerase while Incorporating Two Consecutive C5-Modified Nucleotides

Samra Obeid,[‡] Holger Bußkamp,[‡] Wolfram Welte, Kay Diederichs, and Andreas Marx*

Departments of Chemistry and Biology, Konstanz Research School Chemical Biology, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany

Supporting Information

ABSTRACT: Functional nucleotides are important in many cutting edge biomolecular techniques. Often several modified nucleotides have to be incorporated consecutively. This structural study of KlenTaq DNA polymerase, a truncated form of *Thermus aquaticus* DNA polymerase, gives first insights how multiple modifications are processed by a DNA polymerase and, therefore, contribute to the understanding of these enzymes in their interplay with artificial substrates.

Chemically modified 2' deoxynucleoside triphosphates (dNTPs) play a paramount role in many cutting edge biomolecular technologies, such as next generation sequencing,¹ single molecule sequencing,² microarray analysis based on labeled DNA amplicates,³ DNA conjugation,⁴ or the in vitro selection of ligands like aptamers by systematic enrichment of ligands by exponential amplification.⁵ Therefore, the rational design of modified dNTPs is getting increasingly important to ensure efficient processing by DNA polymerases. When pyrimidines are used for these endeavors the modifications are almost exclusively attached to the C5 position.⁶ For rational design of the modification the knowledge of structure–function relationship of the enzyme is of outstanding importance. Recently obtained structural insights of KlenTaq (large fragment of *Thermus aquaticus* DNA polymerase I) highlight the processing of several modified dNTPs.⁷ These studies indicate that flexible and polar functionalities allow beneficial substrate–protein interactions and thereby might improve the substrate properties. In contrast, the first structures of KlenTaq harboring a rigid and nonpolar modification of the dNTP suggested that the modification results in distortion of substrate–enzyme interactions leading to low incorporation efficiency.^{7a} However, recently published structures of KlenTaq processing another nonpolar, ridged but aromatic nucleotide modification indicated that the introduction of an aromatic ring in the rigid linkage enables additional interactions to the protein and thereby may contribute to good substrate properties.^{7c}

Until now, only structural data harboring a modified nucleotide bound in the dNTP binding site have been reported. However, incorporation of multiple modified nucleotides is crucial for many applications.^{1,2,6} In order to get first structural insight into incorporation of multiple modified nucleotides we next studied the extension process of a modified 3' primer terminus by a modified nucleotide. We

investigated the recently reported alkyne modified nucleotide dC* (Figure 1A) since it has very good substrate properties,

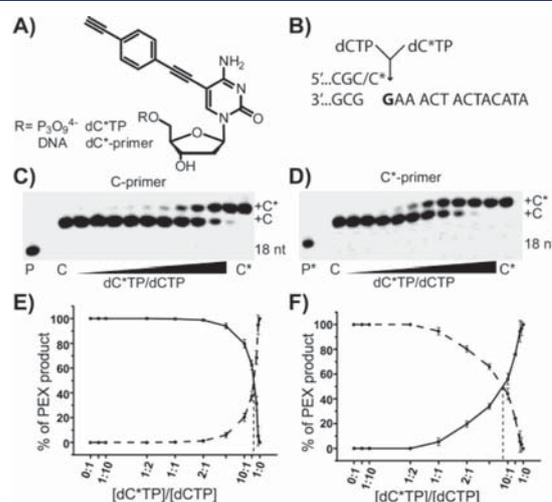


Figure 1. (A) Structure of the dC*TP. (B) Partial sequence of the primer/template complex. Primers are used with either dC or dC* at the 3' end. (C) PAGE analysis of the competition experiment using primer with dC* at the 3' end. (D) PAGE analysis of the competition experiment using primer with dC at the 3'-end. (E) Graphical readout of PAGE analysis of the primer ending with dC. (F) Same as (E) for the primer ending with dC.

which can be functionalized by Cu(I) catalyzed azide–alkyne cycloaddition (“Click” chemistry).^{7c} First we aimed at investigating the extension of the modified 3' primer terminus on a functional level. Therefore we synthesized a DNA primer strand that bears a dC* residue at the 3' terminus (see Scheme S1). We next functionally investigated extension from modified vs unmodified primer strands, respectively, by incorporation of the modified nucleotide in comparison to the unmodified counterpart (Figure 1C–F). Single nucleotide incorporation experiments were performed in which dC*TP directly competes for incorporation with dCTP. This experimental setup was previously used for the same purpose⁷ as well as to study DNA polymerase selectivity.⁸ The ratio of unmodified versus modified nucleotide incorporation is easily accessible via denaturing polyacrylamide gel electrophoresis (PAGE)

analysis and phosphorimaging through the significantly different retention times resulting from the modification of the dC*TPs (Figure 1C).

When studying extension from a primer terminating with dC, we found that KlenTaq DNA polymerase incorporates the modified analogue equally as the unmodified nucleotide when an ~ 27 fold excess of dC*TP was used (Figure 1C,E). Interestingly, only a 7 fold excess of dC*TP was required when the dC* modified primer was extended (Figure 1D,F). To gain insight into how the modification is processed when extending a modified primer by the modified nucleotide, we crystallized KlenTaq in presence of DNA primer/template complex and ddC*TP (Schemes S2 and S3). The template sequence was designed in a way that it codes for the consecutive incorporation of two modified nucleotide analogs. At first, one modified ddC*MP is incorporated by the DNA polymerase, terminating the polymerization reaction, due to the missing 3' OH group. Thereby, the second modified ddC*TP is captured at the nucleotide binding site poised for incorporation. The crystallization trials in the presence of ddC*TP resulted in a high resolution structure of KlenTaq in ternary complex capturing two modified nucleotide analogs (henceforth termed KlenTaq(ddC*), Scheme S3). The structure was solved by difference Fourier techniques at a resolution of 1.9 Å (Table S1 and Figure S2). In KlenTaq(ddC*), the polymerase adopts a similar overall structure as it is observed for KlenTaq processing of either an unmodified ddCTP Samra: KlenTaq(3KTQ) [ohne Leerzeichen] or a modified dC*TP (KlenTaq(dC*), PDB 4ELU) capturing a dC*TP in the waiting position, resulting in RMSD values for C α atoms of 0.761 and 0.664 Å, respectively (note: ddC*TP and dC*TP bear the identical C5 modification). The close ups of the active site illustrate that in KlenTaq(ddC*) the enzyme is stalled in an active mode defined by the closure of the finger domain by the O helix comparable to KlenTaq(3KTQ) and KlenTaq(dC*) (Figure 2A–C). The modified ddC*MP was incorporated according to the Watson Crick rule (Figure 2A). However, the introduction of a modification at the nucleobase affects the protein–substrate interactions of the surrounding amino acids. One critical amino acid thereby is Arg660. In the natural case (KlenTaq(3KTQ)) Arg660 interacts with the phosphate backbone of the 3' primer terminus forming a kind of clamp between the enzyme and the DNA substrate (Figure 2B).

However, the previously published structures of KlenTaq harboring a rigid and nonpolar modified nucleotide in the active site indicate that the Arg660–primer interaction is perturbed, and instead Arg660 is flipped out to make room for the modification as seen in KlenTaq(dC*) (Figure 2C). The structure of KlenTaq(ddC*) shows that in case of the incorporation of two consecutive modified pyrimidine analogs the benzene rings of the modification interact with each other via π – π stacking (Figure 2D). Thereby the benzene ring is twisted out of the plane with respect to the nucleobase (Figure S3). This interaction might account for the higher extension proficiency of extending the dC* terminated primer by dC*TP in comparison to extension of an unmodified primer. However, this assembly results in a reorganization of the surrounding amino acid residues. For instance, Arg660 shows weak electron density indicating a high flexibility, which is manifested in an increased B factor (average B factor of Arg660 side chain atoms: 81.2 Å²). Since the modifications are slightly rotated with respect to the nucleobases, Arg660 seems to orientate

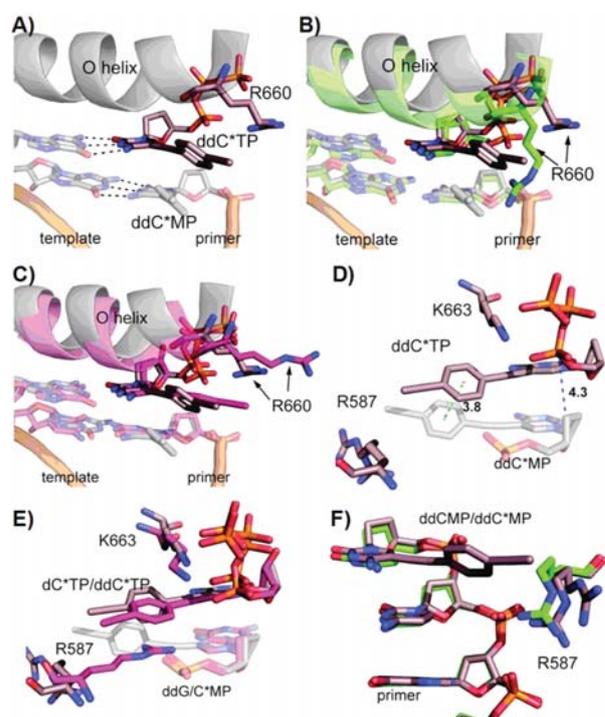


Figure 2. KlenTaq(ddC*) structure (brown). (A) Zoom into the active site of KlenTaq(ddC*). The incoming ddC*TP and the already incorporated ddC*MP are stabilized via Watson Crick base pairing. The O helix packs against the nascent base pair. R660 is indicated as stick. (B) Same orientation as in (A) superimposed structure with KlenTaq(3KTQ) (green). (C) Same orientation as in (A) superimposed structure with KlenTaq(dC*) (pink). (D) The consecutive introduction of two modified substrates result in π – π stacking interactions of the aromatic rings, which is indicated in green dashed lines. (E) Overlay with KlenTaq(dC*) (pink) showing the same orientation as in (D). (F) An overlay of KlenTaq(ddC*) and KlenTaq(3KTQ) (green) shows the primer strand and the amino acid R587. Distances are in Å.

itself in direction of the primer strand. Furthermore, Arg587, which in KlenTaq(3KTQ) is involved to position the DNA primer, is influenced by the introduction of a benzene ring in KlenTaq(dC*) forming a cation– π interaction (Figure 2E). In KlenTaq(ddC*), Arg587 is released from its cation– π interaction pattern to the incoming modified nucleotide making room for the modified 3' primer terminus (Figure 2D,E). Instead it adopts a similar conformation as in KlenTaq(3KTQ) and interacts with the phosphate backbone of the primer strand (Figure 2F). However, Arg587 in KlenTaq(ddC*) shows high flexibility, reflected by the observation of two possible conformations (Figure 2F). The stabilization of the primer strand seems to be slightly affected and reduced by the introduction of the modified pyrimidine analog. Inspecting the residues responsible for catalysis reveals that the complexation of the crucial bivalent metal ions differs slightly from the known arrangement in KlenTaq processing natural substrates (Figure 3). In KlenTaq(ddC*) metal ion B is a Ca²⁺ ion and metal ion A is octahedrally coordinated by Asp610, Glu786, and four water molecules (Figure 3B), whereas metal ion A in KlenTaq(3KTQ) has only five ligands (Figure 3C). In addition, metal ion A in KlenTaq(ddC*) shows no interaction with the α phosphate of ddC*TP, which needs to be activated for proper phosphoryl transfer, suggesting that conformational

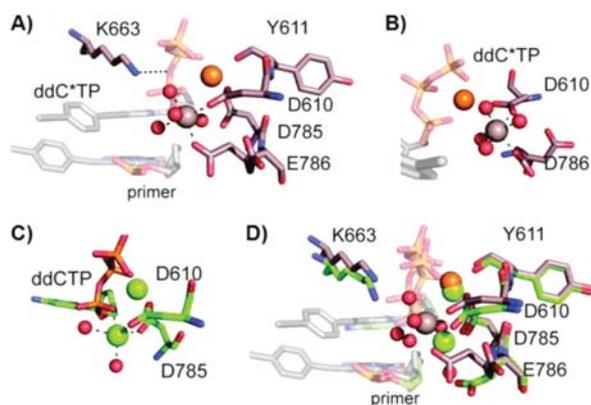


Figure 3. Zoom into the active site of KlenTaq(ddC*) (brown). (A) Close-up view of the complexation of the metal ions (Ca^{2+} , orange; Mg^{2+} , brown) by the incoming ddC*TP and the amino acids crucial for catalysis. (B) Octahedral coordination of the metal ion A (Mg^{2+}) in KlenTaq(ddC*). (C) Coordination of metal ion A by five ligands in KlenTaq(3KTQ) (green). (D) Superimposition of KlenTaq(ddC*) and KlenTaq(3KTQ) in the same orientation as in (A).

changes are required in order to facilitate the phosphodiester bond formation.

This structural study indicates the plasticity of a DNA polymerase. DNA polymerases are able to tolerate several modifications attached to pyrimidines, as numerous incorporation studies and applications demonstrate.^{1–6} The structures of KlenTaq in complex with DNA and modified substrates show that the DNA polymerase accepts the modification and can also interact with it and clearly responds to the introduction of modifications.⁷ The enzyme is equipped with several flexible and positively charged amino acids near the substrate binding sites that stabilize aberrant conformation. Especially this detail is illustrated in the KlenTaq(ddC*) structure by the flexibility of Arg660 and Arg587 (see Figure S4 for an overview). Furthermore, the introduction of a benzene ring offers them the opportunity to interact via π – π stacking as elucidated by this study that might contribute to a higher incorporation proficiency of modified nucleotides. These interactions might overrule interactions with the enzyme and, thereby, alter the substrate properties.

In summary, this study provides first insights into the mechanisms of how multiple modifications are processed by a DNA polymerase and, therefore, contribute to the understanding of these enzymes in their interplay with artificial substrates.

■ ASSOCIATED CONTENT

● Supporting Information

Synthetic procedures and crystallisation and crystallographic analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

Andreas.Marx@uni-konstanz.de

Author Contributions

[‡]These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge funding by the Konstanz Research School Chemical Biology, the Zukunftskolleg of the University of Konstanz, and the DFG for funding, and support and access to beamlines PXI and PXIII at the Swiss Light Source (SLS) at the Paul Scherrer Institute (PSI), Villigen, Switzerland.

■ REFERENCES

- (1) (a) Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R.; et al. *Nature* **2008**, *456*, 53. (b) Bowers, J.; Mitchell, J.; Beer, E.; Buzby, P. R.; Causey, M.; Efcavitch, J. W.; Jarosz, M.; Krzymanska-Olejnik, E.; Kung, L.; Lipson, D.; Lowman, G. M.; Marappan, S.; McInerney, P.; Platt, A.; Roy, A.; Siddiqi, S. M.; Steinmann, K.; Thompson, J. F. *Nat. Methods* **2009**, *6*, 593.
- (2) (a) Ruparel, H.; Bi, L.; Li, Z.; Bai, X.; Kim, D. H.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5932. (b) Harris, T. D.; Buzbyl, P. R.; Babcock, H.; Beer, E.; Bowers, J.; Braslavsky, I.; Causey, M.; Colonell, J.; DiMeo, J.; Efcavitch, J. W.; et al. *Science* **2008**, *320*, 106.
- (3) (a) Guo, J.; Xu, N.; Li, Z.; Zhang, S.; Wu, J.; Kim, D. H.; Marma, M. S.; Meng, Q.; Cao, H.; Li, X.; Shi, S.; Yu, L.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9145. (b) Pollack, J. R.; Perou, C. M.; Alizadeh, A. A.; Eisen, M. B.; Pergamenschikov, A.; Williams, C. F.; Jeffrey, S. S.; Botstein, D.; Brown, P. O. *Nat. Genet.* **1999**, *23*, 41.
- (4) (a) Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2008**, 1828. (b) Hollenstein, M. *Molecules* **2012**, *17*, 13569.
- (5) Mayer, G. *Angew. Chem., Int. Ed.* **2009**, *48*, 2672.
- (6) Examples: (a) Baccaro, A.; Steck, A.-L.; Marx, A. *Angew. Chem., Int. Ed.* **2012**, *51*, 254–257. (b) Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 6782. (c) Srivatsan, S. G.; Tor, Y. *J. Am. Chem. Soc.* **2007**, *129*, 2044. (d) Cahová, H.; Havran, L.; Brázdilová, P.; Pivoňková, H.; Pohl, R.; Fojta, M.; Hocek, M. *Angew. Chem., Int. Ed.* **2008**, *47*, 2059. (e) Gramlich, P. M.; Wirges, C. T.; Manetto, A.; Carell, T. *Angew. Chem., Int. Ed.* **2008**, *47*, 8350. (f) Hocek, M.; Fojta, M. *Org. Biomol. Chem.* **2008**, *6*, 2233. (g) Kaufmann, G. F.; Meijler, M. M.; Sun, C.; Chen, D. W.; Kujawa, D. P.; Mee, J. M.; Hoffman, T. Z.; Wirsching, P.; Lerner, R. A.; Janda, K. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 2144. (h) Thoresen, L. H.; Jiao, G.-S.; Haaland, W. C.; Metzker, M. L.; Burgess, K. *Chem. Eur. J.* **2003**, *9*, 4603. (i) Thum, O.; Jager, S.; Famulok, M. *Angew. Chem., Int. Ed.* **2001**, *40*, 3990. (j) Seela, F.; Zulauf, M. *Chem. Eur. J.* **1998**, *4*, 1781. (k) Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2007**, 1828.
- (7) (a) Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21327. (b) Bergen, K.; Steck, A. L.; Strutt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. *J. Am. Chem. Soc.* **2012**, *134*, 11840. (c) Obeid, S.; Bußkamp, H.; Welte, W.; Diederichs, K.; Marx, A. *Chem. Commun.* **2012**, 8320.
- (8) Bertram, J. G.; Oertell, K.; Petruska, J.; Goodman, M. F. *Biochemistry* **2010**, *49*, 20.