Structures of *KlenTaq* DNA Polymerase Caught While Incorporating C5-Modified Pyrimidine and C7-Modified 7-Deazapurine Nucleoside Triphosphates

Konrad Bergen,[†] Anna-Lena Steck,[†] Stefan Strütt, Anna Baccaro, Wolfram Welte, Kay Diederichs, and Andreas Marx*

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

S *Supporting Information*

ABSTRACT: The capability of DNA polymerases to accept chemically modified nucleotides is of paramount importance for many biotechnological applications. Although these analogues are widely used, the structural basis for the acceptance of the unnatural nucleotide surrogates has been only sparsely explored. Here we present in total six crystal structures of modified 2′-deoxynucleoside-5′-O-triphosphates (dNTPs) carrying modifications at the C5 positions of pyrimidines or C7 positions of 7-deazapurines in complex with a DNA polymerase and a primer/template complex. The modified dNTPs are in positions poised for catalysis leading to incorporation. These structural data provide insight into the mechanism of incorporation and acceptance of modified dNTPs. Our results open the door for rational design of modified nucleotides, which should offer great opportunities for future applications.

The ability of DNA polymerases to process nucleobase-modified 2′-deoxynucleoside-5′-O-triphosphates (dNTPs) is often the essential step in biotechnological applications.[1] Modified nucleotides are used for structural characterization, immobilization, DNA conjugation, or selection of aptamers by systematic enrichment of ligands by exponential amplification (SELEX).[1b,2] For instance, dye-labeled nucleotides are of outstanding importance in DNA sequencing approaches.[1c,e,3] In most cases, the modification is linked to the nucleobase moiety.[4] Thereby the C5 positions of pyrimidines and the C7 positions of 7-deazapurines were identified as best suited for the introduction of modifications without compromising DNA polymerase activity.[4] Although modified nucleotides are widely employed, the mechanisms by which they are accepted as substrates and incorporated by DNA polymerases are still unclear, and until now, the acceptance of modified nucleotides by a DNA polymerase often has not been predictable.[5] Hitherto, structural data of DNA polymerases in complex with modified nucleotides have been limited to a single report on thymidine analogues.[4b] Here we present several crystal structures of a set of modified 2′-deoxynucleoside-5′-O-triphosphates (dN*TPs) carrying the same aminopentinyl modification[4a,6] bound to DNA polymerase and a primer/template complex. The modifications are linked either to the C5 position of a pyrimidine or the C7
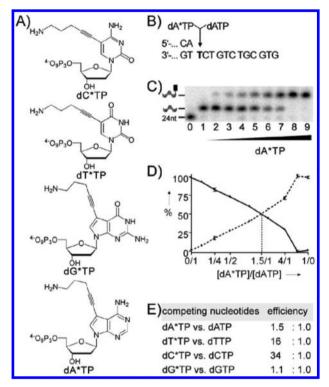

Figure 1. (A) Structures of the aminopentinyl-modified nucleoside triphosphates dN*TP. (B) Exemplary partial DNA sequences of the primer and template for the reactions employing dA*TP and dATP. (C) PAGE analysis of an exemplary competition experiment employing *KlenTaq* DNA polymerase in which the dA*TP/dATP ratio was varied: lane 0, 5′-^{32}P-labeled primer; lane 1, 0/1 ratio; lane 2, 1/10 ratio; lane 3, 1/4 ratio; lane 4, 1/2 ratio; lane 5, 1/1 ratio; lane 6, 2/1 ratio; lane 7, 4/1 ratio; lane 8, 10/1 ratio; lane 9, 1/0 ratio. (D) Evaluation of the incorporation efficiency using mixtures with varied compositions of dA*TP (■, dashed line) and dATP (●, solid line) and *KlenTaq* DNA polymerase. The % conversion is plotted vs the dA*TP/dATP ratio. The vertical dotted line marks the approximate ratio where the two nucleotides are equally incorporated. (E) Overview of the efficiencies of the presented modified nucleotides in competition with their natural counterparts (see Figure S1).
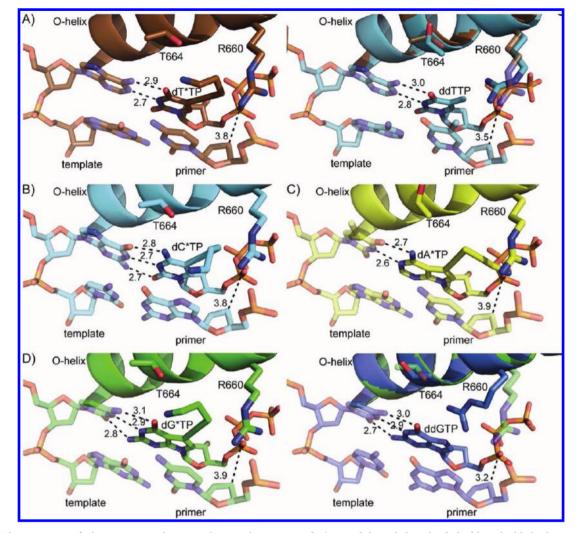
Figure 2. Close-up views of *KlenTaq* DNA polymerase showing the incoming dN*TP and the O-helix. The dashed lines highlight the Watson−Crick base-pairing interactions and the distance of the α-phosphate to the primer 3′ terminus. All distances are in Å. (A) Left panel: interaction distances and orientation of dT*TP and the position of Arg660. Right panel: interaction distances of the natural ddTTP (PDB ID 1QTM) and an overlay of the Arg660 residues. (B) Same as (A) left panel for the structure containing dC*TP. (C) Same as (A) left panel for the structure containing dA*TP. (D) Left panel: same as (A) left panel for the structure containing dG*TP. Right panel: structure containing ddGTP (PDB ID 1QSS) and an overlay of Arg660 as in (A).

position of a 7-deazapurine (see Figure 1A for the structures of dC*TP, dT*TP, dG*TP, dA*TP). The amine-modified nucleotides are well-suited for further functionalization via amide bond formation and are therefore of great interest.[4a,5] Additionally, these nucleotides are compared with their natural counterparts by structural and functional means. For our studies, we chose the N-terminally truncated form of the DNA polymerase I from *Thermus aquaticus* (*KlenTaq* DNA polymerase) because of its well-known characteristics at the structural and functional level and its relevance in biotechnological applications.[4b,7]

We first investigated the efficiency of nucleotide incorporation of the nucleobase-modified nucleotides dN*TP in comparison to their natural counterparts. We performed single-nucleotide incorporation experiments in which the modified nucleotides directly competed with their natural counterparts for incorporation. This experimental setup was previously used for the same purpose[4a,b] as well as to study DNA polymerase selectivity.[8] In Figure 1B−D, an exemplary study employing dA*TP and dATP is depicted [see Figure S1 in the Supporting Information (SI) for results for the other dN*TPs].

We used a 24 nucleotide (nt) primer with a ^{32}P label at the 5′ end and four different 36 nt templates that code for the extension of the primer by a single complementary nucleotide (the DNA sequences are listed in the SI). The ratio of unmodified versus modified nucleotide incorporation is easily accessible via denaturing polyacrylamide gel electrophoresis (PAGE) analysis and phoshorimaging because of the significantly different retention times resulting from the modification of the dN*TPs (Figure 1C). Similar observations of lower mobility for modified DNA have been reported previously.[9] Interestingly, we found that *KlenTaq* DNA polymerase incorporated the purine analogues with approximately the same efficiency as the natural counterparts, whereas the pyrimidine analogues were incorporated with 16−34-fold lower efficiency than their natural counterparts (Figure 1E).

Knowing that *KlenTaq* DNA polymerase accepts the dN*TPs, we aimed at solving crystal structures of *KlenTaq* DNA polymerase in complex with the modified nucleotides. To obtain crystals suitable for structure elucidation of the *KlenTaq* DNA polymerase in complex with DNA primer/template and the dN*TPs, we employed different
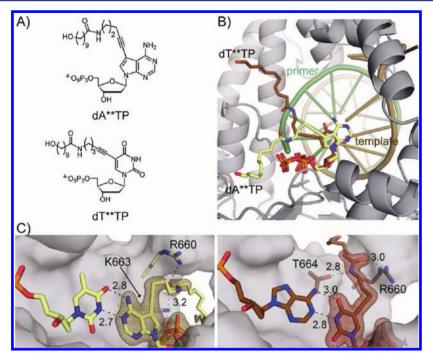
Figure 3. (A) Structures of the dN**TPs used. (B) Close-up view of the active center. The orientations of the modifications are shown in sand for dA**TP and in brown for dT**TP. Parts of the enzyme, including the O-helix, have been removed for better visibility. (C) Close-up views of the nascent base pairs and the orientation of the attached modifications. Depicted is the model density at 1σ. The distances in Å are indicated as dashed lines. The left panel shows dA**TP with the stabilizing residues Lys663 (K663, in the background) and Arg660 (R660), and the active site is shown as the Connolly surface. The right panel shows dT**TP with Arg660 and Thr664 forming hydrogen bonds to the amide of the modification. Both cavities are lined by Arg660.

crystallization strategies[7a,b,f,10] and obtained the best results using a method similar to the one reported by Beard et al.[11] We crystallized binary complexes of *KlenTaq* DNA polymerase in complex with primer and template first and then soaked the crystals with the respective dN*TP (Figure 1A; also see the SI). Details concerning structure solution and refinement can be found in the SI. All of the structures of *KlenTaq* DNA polymerase in complex with the four modified dN*TPs were obtained at high resolution (1.9−2.0 Å) and found to adopt conformations similar to the unmodified cases (rmsd for Cα atoms = 0.33−0.38 Å). The respective modified dN*TPs are bound in positions poised for catalysis and undergo canonical Watson−Crick nucleobase pairing to the templating nucleobases (Figure 2). As well, the O-helix of the finger domain is packed tightly to the nascent base pair, thereby forming a closed, active complex, comparable to the structures observed when natural substrates were used. The distances of the 3′ end of the primer to the α-phosphate of the modified nucleotide are slightly higher than in the natural case, ranging between 3.8 and 3.9 Å (Figure 2). In previously reported structures containing C5-modified dTTPs, Arg660,[7d,g] which interacts with the primer strand when an unmodified nucleotide is bound, is displaced substantially as a result of the steric hindrance of the bulky C5 modification.[4b] The structures presented here show a smaller reorientation of Arg660 relative to the previously reported structure (PDB ID 3OJU; see Figure S2), except in the case when dG*TP is bound (Figure 2).[7d] On the basis of the amino acid alignment in several A-family DNA polymerases, it is known that this Arg660 in particular is located within the B motif[4b,12] and is conserved in bacteria.[4b] Hence, it is likely that the effects on Arg660 apply to other DNA polymerases in this sequence family as well.

Direct comparison of the modified dN*TPs bound in the active site of the enzyme revealed an unexpected feature concerning the orientation of the aminopentinyl modifications. Interestingly, whereas the modifications of dC*TP, dG*TP, and dT*TP point toward the base, with only variation of the plane for electrostatic reasons (Figure S2), the modification of dA*TP points toward the phosphate of the primer terminus (Figure 2C). To verify whether this is a single observation specific for this modification, we studied extended modifications using 7-(*N*-(10-hydroxydecanoyl)aminopentinyl)-7-deaza-dATP (dA**TP) and 5-(*N*-(10-hydroxydecanoyl)aminopentinyl)-dUTP (dT**TP) (Figure 3A). The compounds were synthesized and crystallized in the same manner as described above (for details, see the SI). Notably, both analogues were accepted by *KlenTaq* DNA polymerase. In competition experiments, dA**TP was incorporated with approximately the same efficiency and dT**TP with ∼12-fold lower efficiency than their natural counterparts (Figure S3). The structures derived from these data showed overall properties similar to those of the unmodified cases (Table S1 in the SI; rmsd for Cα atoms = 0.30−0.54 Å). Intriguingly, the orientations as well as the interacting amino acid side chains differ for dA**TP and dT**TP, and consequently, the modifications extend from the enzyme active site through different cavities (Figure 3B,C). In detail, dA**TP is stabilized mainly by Lys663 and most likely by Arg660 (Figure 3C, left panel; Lys663 is in the background). Lys663 is within the interaction distance to the amide bond. In regard to dT**TP, hydrogen bonds to the amide are formed by residues Thr664 and Arg660, leading the modification toward the cavity mainly formed by residues of the O-helix.

In comparison to the previous structures of nucleobase-modified nucleotides,[4b] the modifications used in this study are more flexible and lead to only small disorder in the active site.

The relatively low disturbing effects on the enzyme conformations can be explained by the interaction patterns of the modifications with either the base (in the cases of dT*TP and dG*TP), the phosphate of the primer terminus (dA*TP), or the residues of the O-helix (in the cases of dA**TP and dT**TP). The amide bond in the modification leads to further hydrogen bonding of the modification to residues in the O-helix. These properties might also explain the high incorporation efficiencies of the modified nucleoside triphosphates proven by the functional studies (Figure 1and Figure S3). In direct comparison with the previously solved structure of modified dTTP with a spin-label modification,[4b] the smaller displacement of R660 as well as the additional hydrogen-bonding capacity of the modified triphosphate can explain the higher incorporation rates and acceptance of dT*TP and dT**TP (Figure S2A−C). As a consequence of these observations, the combination of the modification length and the positioning of the amide bond between the subsets of the modification seems to be an important factor for efficiency of the incorporation of the modified building block.

These structural data provide insight into the mechanism of acceptance and incorporation of modified dNTPs by a DNA polymerase that is widely used in biotechnological applications and should open the door for rational design of modified nucleotides. The beneficial combination of rationally designed modified nucleotides and directed evolution of DNA polymerases[13] offers great opportunities for future applications.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Synthetic procedures for dA**TP and dT**TP; the reactions in solution; preparations of the crystals; additional figures of the crystal structures and tables of crystallographic data; and complete refs 1a, 1c, and 3a. 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

[†]K.B. and A.-L.S. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) (a) Bentley, D. R.; et al. *Nature* **2008**, *456*, 53. (b) Obeid, S.; Yulikov, M.; Jeschke, G.; Marx, A. *Angew. Chem., Int. Ed.* **2008**, *47*, 6782. (c) Harris, T. D.; et al. *Science* **2008**, *320*, 106. (d) Ohbayashi, T.; Kuwahara, M.; Hasegawa, M.; Kasamatsu, T.; Tamura, T.; Sawai, H. *Org. Biomol. Chem.* **2005**, *3*, 2463. (e) Seo, T. S.; Bai, X.; Kim, D. H.; Meng, Q.; Shi, S.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5926.

(2) (a) Mayer, G. *Angew. Chem., Int. Ed.* **2009**, *48*, 2672. (b) Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2007**, 1828.

(3) (a) Guo, J.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9145. (b) 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.

(4) (a) Baccaro, A.; Steck, A.-L.; Marx, A. *Angew. Chem., Int. Ed.* **2012**, *51*, 254. (b) Obeid, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21327. (c) Borsenberger, V.; Kukwikila, M.; Howorka, S. *Org. Biomol. Chem.* **2009**, *7*, 3826. (d) Kore, A. R. *Tetrahedron Lett.* **2009**, *50*, 793. (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) Anderson, J. P.; Angerer, B.; Loeb, L. A. *Biotechniques* **2005**, *38*, 257. (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.

(5) Weisbrod, S. H.; Marx, A. *Chem. Commun.* **2008**, 5675.

(6) (a) Augustin, M. A.; Ankenbauer, W.; Angerer, B. *J. Biotechnol.* **2001**, *86*, 289. (b) Seela, F.; Feiling, E.; Gross, J.; Hillenkamp, F.; Ramzaeva, N.; Rosemeyer, H.; Zulauf, M. *J. Biotechnol.* **2001**, *86*, 269.

(7) (a) Betz, K.; Streckenbach, F.; Schnur, A.; Exner, T.; Welte, W.; Diederichs, K.; Marx, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 5181. (b) Obeid, S.; Blatter, N.; Kranaster, R.; Schnur, A.; Diederichs, K.; Welte, W.; Marx, A. *EMBO J.* **2010**, *29*, 1738. (c) Rothwell, P. J.; Mitaksov, V.; Waksman, G. *Mol. Cell* **2005**, *19*, 345. (d) Li, Y.; Waksman, G. *Protein Sci.* **2001**, *10*, 1225. (e) Li, Y.; Mitaxov, V.; Waksman, G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9491. (f) Li, Y.; Kong, Y.; Korolev, S.; Waksman, G. *Protein Sci.* **1998**, *7*, 1116. (g) Li, Y.; Korolev, S.; Waksman, G. *EMBO J.* **1998**, *17*, 7514. (h) Korolev, S.; Nayal, M.; Barnes, W. M.; Di Cera, E.; Waksman, G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9264.

(8) Bertram, J. G.; Oertell, K.; Petruska, J.; Goodman, M. F. *Biochemistry* **2010**, *49*, 20.

(9) (a) Baccaro, A.; Marx, A. *Chem.—Eur. J.* **2010**, *16*, 218. (b) Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. *J. Am. Chem. Soc.* **2005**, *127*, 15071.

(10) Upton, T. G.; Kashemirov, B. A.; McKenna, C. E.; Goodman, M. F.; Prakash, G. K.; Kultyshev, R.; Batra, V. K.; Shock, D. D.; Pedersen, L. C.; Beard, W. A.; Wilson, S. H. *Org. Lett.* **2009**, *11*, 1883.

(11) Beard, W. A.; Shock, D. D.; Batra, V. K.; Pedersen, L. C.; Wilson, S. H. *J. Biol. Chem.* **2009**, *284*, 31680.

(12) Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P. *Protein Eng.* **1990**, *3*, 461.

(13) (a) Leconte, A. M.; Patel, M. P.; Sass, L. E.; McInerney, P.; Jarosz, M.; Kung, L.; Bowers, J. L.; Buzby, P. R.; Efcavitch, J. W.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2010**, *49*, 5921. (b) Staiger, N.; Marx, A. *ChemBioChem* **2010**, *11*, 1963. (c) Ramsay, N.; Jemth, A. S.; Brown, A.; Crampton, N.; Dear, P.; Holliger, P. *J. Am. Chem. Soc.* **2010**, *132*, 5096. (d) Ghadessy, F. J.; Ramsay, N.; Boudsocq, F.; Loakes, D.; Brown, A.; Iwai, S.; Vaisman, A.; Woodgate, R.; Holliger, P. *Nat. Biotechnol.* **2004**, *22*, 755. (e) Fa, M.; Radeghieri, A.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 1748.

D

dx.doi.org/10.1021/ja3017889 | *J. Am. Chem. Soc.* XXXX, XXX, XXX−XXX