Snapshots of a modified nucleotide moving through the confines of a DNA polymerase

Heike Maria Kropp[a,b], Simon Leonard Dürr[a], Christine Peter[a,b], Kay Diederichs[b,c], and Andreas Marx[a,b,1]

[a]Department of Chemistry, Universität Konstanz, 78464 Konstanz, Germany; [b]Konstanz Research School Chemical Biology, Universität Konstanz, 78464 Konstanz, Germany; and [c]Department of Biology, Universität Konstanz, 78464 Konstanz, Germany

Edited by David Baker, University of Washington, Seattle, WA, and approved August 16, 2018 (received for review July 4, 2018)

DNA polymerases have evolved to process the four canonical nucleotides accurately. Nevertheless, these enzymes are also known to process modified nucleotides, which is the key to numerous core biotechnology applications. Processing of modified nucleotides includes incorporation of the modified nucleotide and postincorporation elongation to proceed with the synthesis of the nascent DNA strand. The structural basis for postincorporation elongation is currently unknown. We addressed this issue and successfully crystallized KlenTaq DNA polymerase in six closed ternary complexes containing the enzyme, the modified DNA substrate, and the incoming nucleotide. Each structure shows a high-resolution snapshot of the elongation of a modified primer, where the modification "moves" from the 3′-primer terminus upstream to the sixth nucleotide in the primer strand. Combining these data with quantum mechanics/molecular mechanics calculations and biochemical studies elucidates how the enzyme and the modified substrate mutually modulate their conformations without compromising the enzyme's activity significantly. The study highlights the plasticity of the system as origin of the broad substrate properties of DNA polymerases and facilitates the design of improved systems.

DNA polymerase | modified nucleotide | crystallography | alkyne | click chemistry

Numerous key biotechnological processes rely on the ability of thermostable DNA polymerases (DNA pols) to accept chemically modified nucleotides. These include DNA and RNA sequencing (1–6), systematic evolution of ligands by exponential enrichment (7, 8), and DNA labeling, e.g., for single-nucleotide polymorphism detection (9) or microarray analysis (10–12). At first glance, it is puzzling to find that DNA pols that evolved by nature to appropriately recognize the canonical nucleobases are able to process nucleotides that bear even large modifications (9, 13).

Alkyne-modified nucleotides have proven to be of high potential and utility for obtaining highly functionalized DNA exploiting DNA pol-mediated DNA synthesis in primer extension (14–16) and PCR reactions (14, 17). Interestingly, alkyne-modified nucleotides are even processed during DNA synthesis in living cells (18, 19). The alkyne modifications are particularly interesting due to their ability to react in the bioorthogonal copper-catalyzed alkyne–azide cycloaddition (CuAAC), building a versatile toolbox for postsynthetic modification of DNA that reaches from fluorophores (20) and carbohydrates (20) to the immobilization of DNA on solid support (21) and nanobiotechnology (22). Recently, we gained structural insights into the incorporation process of one and two adjacent alkyne-modified nucleotides within a thermostable DNA pol (23, 24). We found that the seemingly rigid, apolar 1,4-diethynylbenzene modification is well accepted (23). Like other alkyne modifications, it is accessible to CuAAC after incorporation by the DNA pol into DNA (23).

As the acceptance of modified nucleotides by DNA pols is dependent upon interactions between the modification and the enzyme, recent structural studies investigate the incorporation step of modified nucleotides bound in the active site of DNA pols (23–27). To perform these studies, the Klenow fragment of *Thermus aquaticus* DNA pol I (in short, KlenTaq DNA pol) was employed (23–36). These studies elucidate conformational changes of the amino acids within the DNA pol that interact with the modification of the incoming dNTP.

For several applications like sequencing, PCR reactions, and in vivo DNA synthesis, it is crucial that the DNA pol is able to perform postincorporation elongation from the modified nucleotide. As DNA pols interact with the primer template duplex from the 3′ terminus up to the sixth nucleotide upstream (37), each incorporated modified nucleotide has to migrate through the confines of the enzyme during elongation. However, the elongation process has not been investigated on a structural level so far, although insights into this process might impact future nucleotide designs for advanced applications.

Here, we addressed this issue and successfully crystallized KlenTaq DNA pol in six closed ternary complexes containing the enzyme, the modified DNA substrate, and the incoming nucleotide. We chose the 1,4-diethynylbenzene modification in this study since it is the only modified alkyne studied on a structural level so far and it's well accepted by the enzyme. Each of the six structures shows a high-resolution snapshot of the elongation of a modified primer where the 1,4-diethynylbenzene modified nucleotide "moves" from the 3′ terminus upstream to the sixth nucleotide in the primer strand. Combining the insights obtained from X-ray crystallography with quantum mechanics/molecular mechanics (QM/MM) calculations

Significance

Despite being evolved to process the four canonical nucleotides, DNA polymerases are known to incorporate and extend from modified nucleotides, which is the key to numerous core biotechnology applications. The structural basis for postincorporation elongation remained elusive. We successfully crystallized KlenTaq DNA polymerase in six complexes, providing high-resolution snapshots of the modification "moving" from the 3′ terminus upstream to the sixth nucleotide in the primer strand. Combining these data with quantum mechanics/molecular mechanics calculations and biochemical studies elucidates how the enzyme and the modified substrate mutually modulate their conformations without compromising the enzyme's activity. This highlights the unexpected plasticity of the system as origin of the broad substrate properties of the DNA polymerase and guide for the design of improved systems.

Author contributions: H.M.K., S.L.D., C.P., K.D., and A.M. designed research; H.M.K. and S.L.D. performed research; H.M.K., S.L.D., K.D., and A.M. analyzed data; and H.M.K., S.L.D., C.P., K.D., and A.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID codes 6FBC–6FBI).

[1]To whom correspondence should be addressed. Email: andreas.marx@uni-konstanz.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1811518115/-/DCSupplemental.

Published online September 17, 2018.

and biochemical data provides insights into how enzyme conformations are modulated by the nucleotide ligand and, in turn, the protein environment modulates the conformation of the modified nucleotide. Interestingly, these aberrations occur without compromising the enzyme's activity significantly. This highlights the unexpected plasticity of the system that may account for the substrate properties in KlenTaq DNA pol.

Results

Structures of KlenTaq DNA Pol in Complex with Modified DNA and Nonhydrolyzable Triphosphate. Former structures of KlenTaq DNA pol show protein–DNA interactions within the first six nucleotides of the primer strand (37) (Fig. 1). To investigate how these interactions at the primer strand interfere with chemical modifications, we aimed at obtaining crystal structures of KlenTaq DNA pol in ternary, closed complexes bound to primers bearing the modifications at the respective sites (*SI Appendix*, Table S2), template, and an incoming triphosphate (here, the nonhydrolyzable dGTP analog 2′-deoxyguanosine-5′-[(α,β)-imido]triphosphate, dGpNHpp; *SI Appendix*, Fig. S1) at the active site. For the crystallization experiments, we synthesized the required six modified primer strands by automated solid phase synthesis (24). The nucleotides bear a C5 modification (1,4-diethynylbenzene) at positions 1 to 6 upstream of the primer 3′ terminus (Fig. 1). For comparison, KlenTaq DNA pol was also crystallized employing an unmodified primer.

The structures were solved using difference Fourier techniques, providing structures at resolutions of 1.5 Å to 2.1 Å. Electron density maps for the modified nucleotides in the six structures are shown in *SI Appendix*, Fig. S2. The overall structures were very similar, showing only minor changes compared with the KlenTaq DNA pol structure with natural primer/template, leading to rmsd values of 0.147 to 0.226 for Cα atoms (*SI Appendix*, Fig. S3).

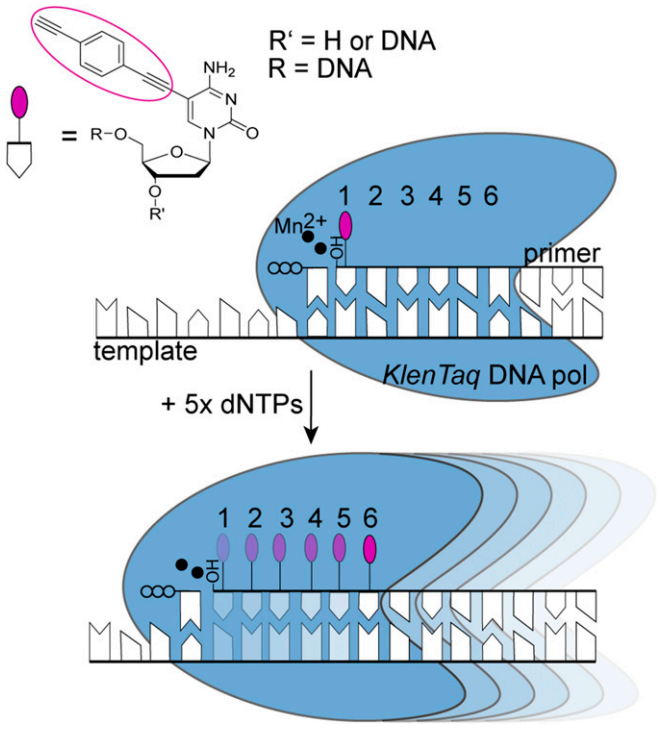

Fig. 1. Depiction of the modified nucleotide "moving" through KlenTaq DNA pol. KlenTaq DNA pol (blue) binds to the primer/template complex bearing the modified nucleotide (pink) at the 3′-primer terminus. Upon incorporation of 5 dNTPs, the primer is elongated, and the modified nucleotide moves through the enzyme from the first to the sixth primer position.

All crystals show complexes that are poised for catalysis, with the active site metal ions coordinating the triphosphate and the terminal 3′-OH of the primer. The metal ions are coordinated by the active site residues (Y611, D610, and D785) and two water molecules, with the closed O helix packing against the nascent dGpNHpp–dCMP pair. Due to the anomalous signal of the active site metal ions, we modeled Mn^{2+} at these positions, as the crystallization conditions contained 20 mM to 25 mM $MnCl_2$ (*SI Appendix*, Table S1).

The active sites shown here align to other KlenTaq DNA pol crystal structures bound to C5-modified nucleotides and Mg^{2+} in the active site (23, 25), as well as to KlenTaq DNA pol in complex with natural substrate and Mg^{2+} (30, 34).

KlenTaq DNA Pol Bound to dGpNHpp and Natural DNA. The KlenTaq DNA pol structure bound to unmodified DNA and dGpNHpp shows a nearly identical conformation to the ddGTP-bound KlenTaq DNA pol structure from Li and Waksman (35). K663, which is crucial for the nucleotidyl transfer (38), forms hydrogen bonds to the α- and γ-phosphate of dGpNHpp (3.0 Å, *SI Appendix*, Fig. S4*B*). R660 forms hydrogen bonds to O6 (3.0 Å) and N7 (3.1 Å) of dGpNHpp, as well as to T664 (2.7 Å) (*SI Appendix*, Fig. S4*A*). These interactions were also observed by Li and Waksman (35), who showed that the R660 interaction with the nucleobase of the triphosphate is specific for guanine.

The Modified Nucleotide at the 3′-Primer Terminus. At the 3′-primer terminus, the modified dCMP coordinates the metal ion A with its 3′-hydroxyl group (2.3 Å). The modification is positioned directly underneath the O helix, displacing R660 from its interaction with dGpNHpp and T664 (Fig. 2 *A*–*C*) and opening a pocket for the modification itself (Fig. 2 *D* and *E*). R660 was modeled using phenix polder maps (39) in two conformations, both pointing upward in the direction of the primer backbone and showing an elevated B-factor of its side chain [58.2 (A conformation), 60.00 (B conformation)]. Due to the displacement of R660, K663 can occupy this position, forming hydrogen bonds to N7 of dGpNHpp and water-mediated hydrogen bonds to O6 and the α-phosphate of dGpNHpp, T664, and the primer backbone (Fig. 2 *F* and *G*). Additionally, K663 can form cation −π interactions to the π-system of the 1,4-diethynylbenzene (3.2 Å to the alkene bond, 4.4 Å to the benzene ring, Fig. 2*A*), which are in accordance with cation−π interactions found for NH_4^+, reaching from 3.1 Å (above the aromatic plane) to 4.6 Å (side-on interaction) (40) as well as for cation−π interactions found in biological systems (41, 42). Such cation−π interactions can also be found for R587 (4.7 Å, Fig. 2*A*) that also forms hydrogen bonds to the phosphate backbone. Thus, the aromatic modification is sandwiched between two positively charged amino acid side chains, showing a curved conformation toward the O helix. Furthermore, a considerable 31° twist around the diethynyl axis that connects the nucleobase and the benzene ring of the modification is observed. The terminal atom of the 1,4-diethynylbenzene is in close proximity to the protein surface (*SI Appendix*, Fig. S5), being closest to R587 (2.7 Å).

To account for the origin of the twisted conformation of both aromatic systems and the curvature of the 1,4-diethynylbenzene, we initially performed geometry optimizations of the modified nucleoside with density functional theory (DFT). These showed that the preferred geometry in vacuum is nonbent and slightly twisted (*SI Appendix*, Fig. S6*A*). The twist, which is observed in the DFT calculations, can be attributed to interactions of the N6 amino group of the cytosine and the modification (the distance of the amine hydrogen to the C6 carbon of the proximal alkyne is 2.51 Å).

As these initial theoretical investigations failed to explain the bending of the conjugated 1,4-diethynylbenzene system, found in the crystal structure, we conducted QM/MM calculations using the KlenTaq-p1 structure. We performed three energy minimizations with varying size of the QM region. We started with the

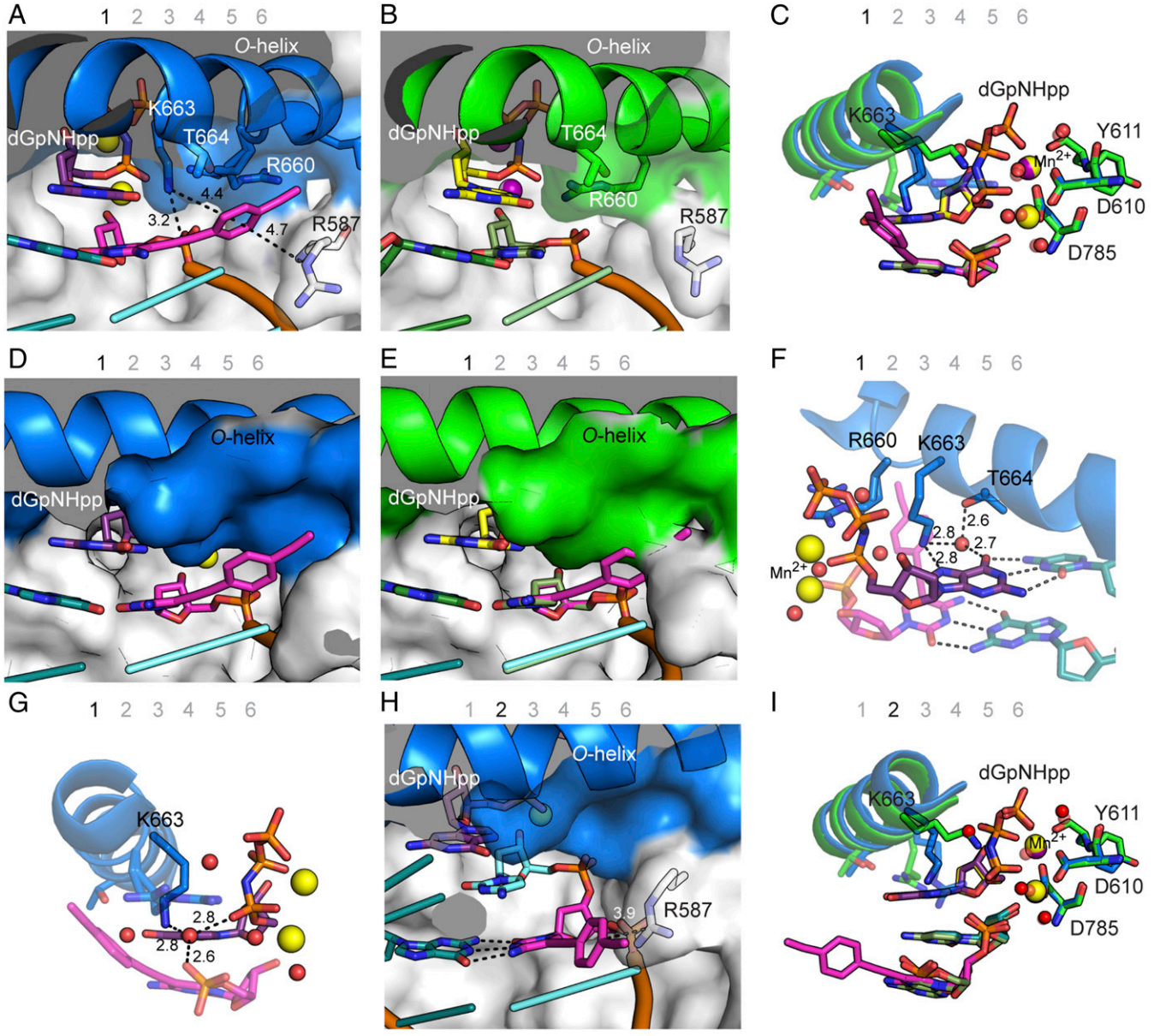

Fig. 2. Interactions between the modified nucleotide and the protein at positions 1 and 2 upstream from the 3′-primer terminus. The position of the nucleotide is indicated by a black number on top of each depiction. Distances are shown in black dashes and are given in angstroms. The protein surface is shown in gray. (*A*) The modification (magenta) is positioned underneath the O helix (blue) and forms cation–π interactions with R587 (4.7 Å) and K663 (3.2 Å and 4.4 Å); R660 points toward the primer backbone. (*B*) In KlenTaq-p^{natural}, R660 points toward O6 and N7 of dGpNHpp, forming a pocket at the nucleotide binding site. (*C*) Superimposition of the active site of KlenTaq-p1 and KlenTaq-p^{natural}. (*D*) Due to the displacement of R660, a pocket is formed in KlenTaq-p1 in which the modification is embedded. (*E*) Superimposition of KlenTaq-p1 with KlenTaq-p^{natural}, showing the protein surface and the O helix (green) of KlenTaq-p^{natural}, highlighting the clash of the modification with the protein in case of a nondisplaced R660. (*F*) K663 (blue) interacts with N7 of dGpNHpp and, via a water molecule, with O6 of dGpNHpp and T664. (*G*) View from the primer backbone to K663, showing water-mediated interactions with the α-phosphate of dGpNHpp and the primer backbone. (*H*) At position 2, the benzene ring of the modified dCMP is twisted by 57° with regard to the plane of the nucleobase and forms cation–π interactions with R587. (*I*) Superimposition of the active site of KlenTaq-p2 and KlenTaq-p^{natural}.

side chain of the putative interaction partner K663 only and the modified cytosine (dCMP^{alkyne}, small QM region, *SI Appendix*, Fig. S7 *A* and *B*) in the QM region, followed by an extended simulation with the side chains of V586, R587, R660, K663, and T664 and dCMP^{alkyne} in the QM region (extended QM region, *SI Appendix*, Fig. S7*C*). We observed that the modification gets less bent compared with the crystal structure (*SI Appendix*, Fig. S7), but stays twisted, when changing the size of the QM region. Additionally, K663 moves out of its electron density toward the α-phosphate of dGpNHpp during the optimization of the small QM region (*SI Appendix*, Fig. S7 *A* and *B*), as it is displaced from its position by a water molecule. Thus, the bending and twisting of the modification is not only caused by interaction with K663 but, moreover, by the protein environment, as K663 shows different conformations depending on the size of the QM region. To further investigate how the protein modulates the conformation of the modification, we generated an in silico mutant, where the amino acids most likely interacting with the 1,4-diethynylbenzene (V586A, R587A, R660A, K663A, T664A) were mutated to alanine. Using this mutant for the calculation, we obtained a propeller twist of about 87° of the benzene ring in relation to the nucleobase (*SI Appendix*,

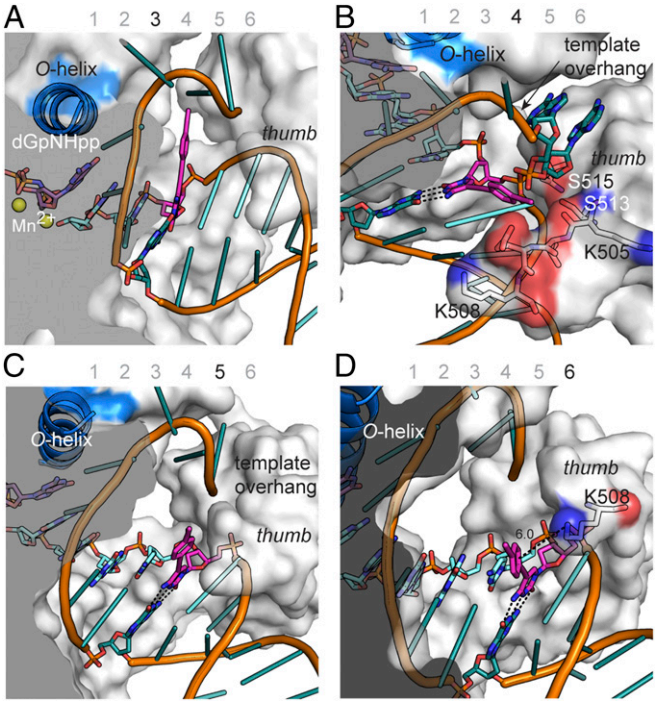

Fig. 3. Following the modified nucleotide's movement to position 6 upstream from the 3′-primer terminus. The position of the nucleotide is indicated by a black number on top of each depiction. Superimpositions with KlenTaq-p^{natural} are shown in *SI Appendix*, Fig. S8. Distances are shown as black dashes and are given in angstroms. The protein surface is shown in gray. (*A*) Position 3: The modified dCMP (magenta) points toward the thumb domain. (*B*) Position 4: The modified dCMP is in close proximity to the template overhang and the thumb domain with the residues K505, S513, and S515, showing a twist of the benzene ring in respect to the nucleobase. (*C*) Position 5: The modification points toward the tip of the thumb domain. (*D*) Position 6: The modification is located close to the tip of the thumb and comes close to K508 (6.0 Å). The benzene ring is flipped and is parallel to the amino group of K508.

Fig. S7*D*), which greatly emphasizes the role of the protein environment in modulating the conformation of the modification.

The energy barrier of rotation around the diethynyl axis of the nucleobase and the benzene ring of the modification from the coplanar to the state with a 90° twist angle (energy maximum) of both aromatic systems was determined, in an in vacuo potential energy scan, to be 3.5 kJ·mol^{-1} at the DFT B3LYP-D3/def2-TZVP level of theory (*SI Appendix*, Fig. S6*B*). Our findings are consistent with earlier calculations of the rotation around the alkyne bond in diphenylethyne of ∼4 kJ·mol^{-1} at the same level of theory (43). This low energy barrier explains the rotation of the modification, not only in the QM simulation but, moreover, in the KlenTaq DNA pol structures.

The Modified Nucleotide at Position 2 from the 3′-Primer Terminus. Next, we investigated the interaction of the modification at position 2 upstream from the 3′-primer terminus. The modified nucleotide is positioned in the crevice between the finger and thumb domain. The benzene ring of the modification is significantly twisted by 57° and forms cation–π interactions with R587, which is also located to interact with the primer phosphate backbone (Fig. 2*H*). However, bending, as described for the KlenTaq-p1 structure, is not observed at this position anymore. The same holds true when the modification is located farther upstream from the 3′-primer terminus (*vide infra*). The amino acid side chains of K663 and R660 show high flexibility (B-factors 76.0 Å^2 and 94.6 Å^2, respectively), where K663 seems to point toward the N7 of dGpNHpp and R660 seems to point toward the primer backbone (Fig. 2*I*).

The Modified Nucleotide Moving Upstream to Position 6. During further elongation, the modification will pass positions 3 to 6 upstream from the 3′-primer terminus (Fig. 3). The conformation of the modification at these four positions differs, showing twists of the benzene ring, in relation to the nucleobase, ranging from 5° to 58° (Fig. 4). The most planar conformations are observed for KlenTaq-p3 and KlenTaq-p5, with 5° and 8°, respectively. The modification in KlenTaq-p3 is located in the crevice between finger and thumb domain, pointing toward the thumb, showing no interactions with the protein (Fig. 3*A*) and being in close proximity to A517 (2.7 Å, *SI Appendix*, Fig. S5).

In KlenTaq-p4, the modification points into a pocket formed by the thumb domain and the single-stranded template (Fig. 3*B*). The phosphate backbone of the single-stranded template, the amino acid side chains of S513 and S515, and the backbone of K505 (Fig. 3*B*) are in close proximity to the 1,4-diethynylbenzene. Here, the twist of the benzene ring in relation to the nucleobase is 58°.

The modification in KlenTaq-p5 is positioned at the tip of the thumb domain, not displaying any interactions with the protein (Fig. 3*C*). Finally, the modification reaches position 6 in KlenTaq-p6, where the benzene ring is in close proximity to K508 (Fig. 3*D*). The benzene ring is positioned parallel to the amino group of K508, and twisted by 29° in relation to the plane of the nucleobase.

Additionally, the amino acid side chains of the catalytically important residues K663 and R660 show different conformations in KlenTaq-p4 and KlenTaq-p5 and high flexibility in KlenTaq-p3 and KlenTaq-p6. In positions 4 and 5, K663 interacts with the bound triphosphate, but in different modes. In KlenTaq-p4, K663 forms hydrogen bonds to the α- and β-phosphate, whereas, in KlenTaq-p5, K663 forms hydrogen bonds to N7 of the nucleobase of dGpNHpp. Also, R660 adopts different conformations, forming

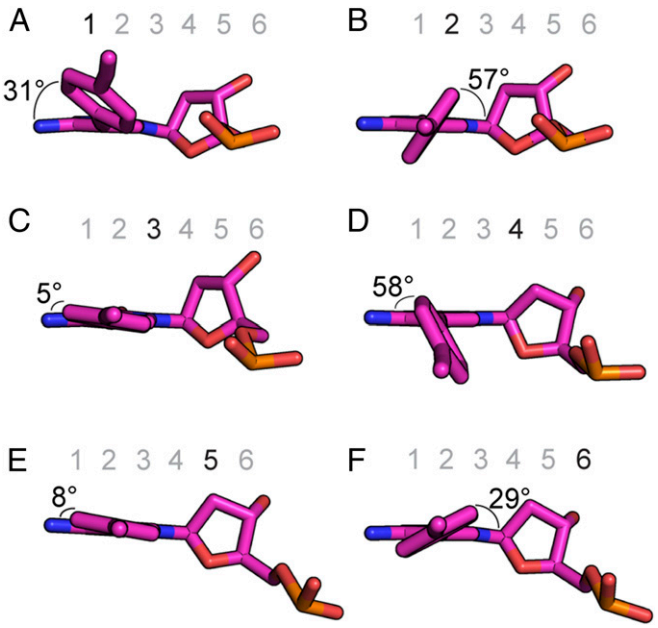

Fig. 4. Depiction highlighting the conformational flexibility of the modification at the six different positions (*A*–*F*), respectively. Shown is the twist of the benzene ring in respect to the nucleobase plane. The position of the nucleotide is indicated by a black number above the image.

hydrogen bonds to N7 of dGpNHpp in KlenTaq-p4 and pointing toward the primer backbone in KlenTaq-p5.

Also, the terminal atoms of the modification in KlenTaq-p4 and KlenTaq-p5 are in close proximity to the protein surface (2.9 Å and 2.3 Å, respectively), whereas KlenTaq-p6 shows a larger distance to the protein surface (4.0 Å) (*SI Appendix*, Fig. S5).

Primer Extensions. The structural studies show most effects of the modification on the system at position 1. To access the impact of the modification on the enzyme function at this position, we measured single-nucleotide incorporation kinetics under steady-state single completed hit conditions as described previously (44–47) (*SI Appendix*, Fig. S9 and Table S3). The K_M and k_{cat} values derived from the modified substrate are very similar to the ones derived from the unmodified system, indicating that the structural alterations that are caused by the modification have little impact on the enzyme function (*SI Appendix*, Table S3).

To investigate the acceptance of the modification during primer elongation to the full-length product by the incorporation of multiple nucleotides, we performed primer extension experiments employing the p1 primer and all four natural dNTPs in a time-dependent manner, showing the full elongation of the p1 primer (*SI Appendix*, Fig. S10). This finding reveals that the conformational changes as described above are well accommodated within the enzyme substrate complexes without significantly compromising activity.

Discussion

Structural investigations of DNA pols in complex with modified nucleotides have so far been reported on modified triphosphates bound in the active site (23–27). First, these studies indicate that modifications, capable of forming hydrogen-bonding interactions, are well suited for incorporation by the enzyme (25–27). However, also the apolar 1,4-diethynylbenzene C5-modification in pyrimidine triphosphates is also well accepted as substrate for KlenTaq DNA pol (23). The formation of cation−π interactions with positively charged amino acid side chains might be the mechanistic basis for this property. Here, we investigated the interactions of the modification in postincorporation elongation events by KlenTaq DNA pol. We first synthesized six different chemically modified primer strands that bear the modified nucleotide at positions 1 to 6 upstream from the 3′ terminus of the primer (Fig. 1). Employing these oligonucleotides in crystallization trials, we obtained the desired six structures of closed ternary KlenTaq DNA pol complexes, mimicking the postincorporation movement of a nucleobase modification through the confines of a DNA pol.

These structures elucidate that the 1,4-diethynylbenzene modification forms cation−π interactions to positively charged amino acid side chains (K663, R587), if it is located at the 3′ terminus or the second upstream position. These cation−π interactions may contribute to an additional stabilization of the ternary closed complex and the primer. However, more surprisingly, the structures revealed that the modification adopts several distinct conformations, depending on its positioning in the primer (Fig. 4). Among them, the bent conformation of the 1,4-diethynylbenzene in the KlenTaq-p1 structure at the 3′ terminus is the most pronounced. QM/MM simulations showed that the bending of the modification depends on the protein environment which modulates the conformation of the aromatic system.

The QM/MM calculations further indicate that the preference for K663 to be in the position observed in the X-ray KlenTaq-p1 structure is rather small, as a change of the QM region alters the conformation of K663 (*SI Appendix*, Fig. S7). Even though K663 is important for the nucleotidyl transfer reaction (38) and R660 may specifically interact with an incoming guanine (35), both amino acid side chains are found to be very flexible.

Since the structural studies revealed most pronounced conformational effects caused by the modification at position 1, enzyme kinetics were measured. The obtained kinetic data indicate that the observed conformational changes do not impair the enzymes' capability to process the modified substrates. Moreover, KlenTaq DNA pol's activity is a little higher in the case of the p1 primer compared with $p^{natural}$, even though K663 is displaced from its usual position and interacts with dGpNHpp via hydrogen bonds and the modification via cation−π interaction. For KlenTaq-p2, Klentaq-p3, and KlenTaq-p5, K663 was also observed to be in this displaced position. On first glance, this observation is puzzling, since it was shown that K663 is important for the nucleotidyl transfer reaction (38) and one would expect diminished activity. Considering that crystal structures represent snapshots of the average conformation that were here measured at 100 K, we assume that residues are more flexible in solution at ambient temperature and may change their conformation to promote catalysis. Additionally, cation−π interactions were shown to be significantly attenuated by solvent (40). Moreover, the modification is positioned underneath the O helix of the finger domain, where these interactions may stabilize the closed complex, contributing to the enhanced activity of KlenTaq DNA pol with the p1 primer.

The low energy barrier of rotation along the dietynyl axis that connects the nucleobase and the benzene ring of the modification (3.50 $kJ \cdot mol^{-1}$) allows comparatively easy rotation of the modification and results in twisted conformations in KlenTaq-p2, KlenTaq-p4, and KlenTaq-p6 (Fig. 4). Moreover, the twisted conformation of the modification at the fourth and sixth positions may also be due to the protein environment, even though no direct interaction with the protein is observed. However, the modifications in KlenTaq-p4 and KlenTaq-p6 point into a pocket formed by the thumb domain and the template overhang or are in close proximity to a lysine residue, respectively.

Overall, the good acceptance of the modification by KlenTaq DNA pol seems not only to rely on the modification's flexibility but also on its size. The terminal atom of the modification is always in close proximity to the protein surface (*SI Appendix*, Fig. S5), with distances ranging from 2.1 Å to 4.2 Å, but still being small enough to be well positioned in the crevice between finger and thumb domain.

In summary, we have gained insight into the processing of a modified nucleotide within the confines of KlenTaq DNA pol. As the modification "moves" from the 3′ terminus upstream to position 6, different conformations of the substrate nucleotide were observed in the confines of the enzyme. X-ray crystallography, QM/MM calculations, and biochemical data show how KlenTaq DNA pol's conformations are modulated by the nucleotide ligand and, in turn, the protein environment modulates the conformation of the modified nucleotide. Interestingly, these aberrations come without compromising the enzyme's activity. This highlights the unexpected plasticity of the system that may account for the substrate properties in KlenTaq DNA pol.

Materials and Methods

Detailed description of protein purification, crystallization, acquisition of X-ray diffraction data, refinement and model building, primer extension experiments, and DFT calculations are given in *SI Appendix, SI Materials and Methods*.

Atomic coordinates and structure factors for the KlenTaq DNA pol structures have been deposited in the Protein Data Bank (PDB) (www.rcsb.org). The accession codes for each structure are given in *SI Appendix*, Table S1.

ACKNOWLEDGMENTS. We acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG), the European Research Council (AdG Grant 339834 EvoEPIGEN), and the Swiss Light Source of the Paul Scherrer Institute for access to beamlines PXI and PXIII and great support. Theoretical calculations have been carried out on the JUSTUS cluster, funded by the state of Baden-Württemberg and the DFG (Grant INST 40/467-1 FUGG).

1. Bentley DR, et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456:53–59.
2. Metzker ML (2010) Sequencing technologies–The next generation. *Nat Rev Genet* 11: 31–46.
3. Fuller CW, et al. (2016) Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array. *Proc Natl Acad Sci USA* 113:5233–5238.
4. Ju J, et al. (2006) Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators. *Proc Natl Acad Sci USA* 103:19635–19640.
5. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: Ten years of next-generation sequencing technologies. *Nat Rev Genet* 17:333–351.
6. Guo J, et al. (2008) Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc Natl Acad Sci USA* 105:9145–9150.
7. Sefah K, et al. (2014) In vitro selection with artificial expanded genetic information systems. *Proc Natl Acad Sci USA* 111:1449–1454.
8. Mayer G (2009) The chemical biology of aptamers. *Angew Chem Int Ed Engl* 48: 2672–2689.
9. Welter M, Verga D, Marx A (2016) Sequence-specific incorporation of enzyme-nucleotide Chimera by DNA polymerases. *Angew Chem Int Ed Engl* 55:10131–10135.
10. Mischel PS, Cloughesy TF, Nelson SF (2004) DNA-microarray analysis of brain cancer: Molecular classification for therapy. *Nat Rev Neurosci* 5:782–792.
11. Perou CM, et al. (2000) Molecular portraits of human breast tumours. *Nature* 406: 747–752.
12. Pollack JR, et al. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23:41–46.
13. Hottin A, Marx A (2016) Structural insights into the processing of nucleobase-modified nucleotides by DNA polymerases. *Acc Chem Res* 49:418–427.
14. Gierlich J, et al. (2007) Synthesis of highly modified DNA by a combination of PCR with alkyne-bearing triphosphates and click chemistry. *Chemistry* 13:9486–9494.
15. Slavíčková M, et al. (2018) Turning off transcription with bacterial RNA polymerase through CuAAC click reactions of DNA containing 5-ethynyluracil. *Chemistry* 24: 8311–8314.
16. Panattoni A, Pohl R, Hocek M (2018) Flexible alkyne-linked thymidine phosphoramidites and triphosphates for chemical or polymerase synthesis and fast post-synthetic DNA functionalization through copper-catalyzed alkyne-azide 1,3-dipolar cycloaddition. *Org Lett* 20:3962–3965.
17. Gramlich PME, Wirges CT, Gierlich J, Carell T (2008) Synthesis of modified DNA by PCR with alkyne-bearing purines followed by a click reaction. *Org Lett* 10:249–251.
18. Salic A, Mitchison TJ (2008) A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc Natl Acad Sci USA* 105:2415–2420.
19. Neef AB, Samain F, Luedtke NW (2012) Metabolic labeling of DNA by purine analogues in vivo. *ChemBioChem* 13:1750–1753.
20. Gramlich PME, Warncke S, Gierlich J, Carell T (2008) Click-click-click: Single to triple modification of DNA. *Angew Chem Int Ed Engl* 47:3442–3444.
21. Timper J, et al. (2012) Surface "click" reaction of DNA followed by directed metalization for the construction of contactable conducting nanostructures. *Angew Chem Int Ed Engl* 51:7586–7588.
22. Fischler M, et al. (2008) Chain-like assembly of gold nanoparticles on artificial DNA templates via 'click chemistry'. *Chem Commun (Camb)*:169–171.
23. Obeid S, Busskamp H, Welte W, Diederichs K, Marx A (2012) Interactions of non-polar and "Click-able" nucleotides in the confines of a DNA polymerase active site. *Chem Commun (Camb)* 48:8320–8322.
24. Obeid S, Bußkamp H, Welte W, Diederichs K, Marx A (2013) Snapshot of a DNA polymerase while incorporating two consecutive C5-modified nucleotides. *J Am Chem Soc* 135:15667–15669.
25. Obeid S, Baccaro A, Welte W, Diederichs K, Marx A (2010) Structural basis for the synthesis of nucleobase modified DNA by Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci USA* 107:21327–21331.
26. Bergen K, et al. (2012) Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J Am Chem Soc* 134:11840–11843.
27. Hottin A, Betz K, Diederichs K, Marx A (2017) Structural basis for the KlenTaq DNA polymerase catalysed incorporation of alkene- versus alkyne-modified nucleotides. *Chemistry* 23:2109–2118.
28. Obeid S, et al. (2010) Replication through an abasic DNA lesion: Structural basis for adenine selectivity. *EMBO J* 29:1738–1747.
29. Obeid S, Welte W, Diederichs K, Marx A (2012) Amino acid templating mechanisms in selection of nucleotides opposite abasic sites by a family a DNA polymerase. *J Biol Chem* 287:14099–14108.
30. Betz K, et al. (2012) KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nat Chem Biol* 8:612–614.
31. Betz K, et al. (2013) Structural insights into DNA replication without hydrogen bonds. *J Am Chem Soc* 135:18637–18643.
32. Betz K, et al. (2010) Structures of DNA polymerases caught processing size-augmented nucleotide probes. *Angew Chem Int Ed Engl* 49:5181–5184.
33. Li Y, Kong Y, Korolev S, Waksman G (1998) Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates. *Protein Sci* 7:1116–1123.
34. Li Y, Korolev S, Waksman G (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: Structural basis for nucleotide incorporation. *EMBO J* 17:7514–7525.
35. Li Y, Waksman G (2001) Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP-trapped ternary complex of Klentaq1: Insights into nucleotide incorporation and selectivity. *Protein Sci* 10:1225–1233.
36. Betz K, et al. (2017) Structural basis for the selective incorporation of an artificial nucleotide opposite a DNA adduct by a DNA polymerase. *Chem Commun (Camb)* 53: 12704–12707.
37. Bergen K, Betz K, Welte W, Diederichs K, Marx A (2013) Structures of KOD and 9°N DNA polymerases complexed with primer template duplex. *ChemBioChem* 14: 1058–1062.
38. Castro C, et al. (2009) Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat Struct Mol Biol* 16:212–218.
39. Liebschner D, et al. (2017) Polder maps: Improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol* 73:148–157.
40. Marshall MS, Steele RP, Thanthiriwatte KS, Sherrill CD (2009) Potential energy curves for cation-π interactions: Off-axis configurations are also attractive. *J Phys Chem A* 113:13628–13632.
41. Gallivan JP, Dougherty DA (1999) Cation-π interactions in structural biology. *Proc Natl Acad Sci USA* 96:9459–9464.
42. Schärer K, et al. (2005) Quantification of cation-π interactions in protein-ligand complexes: Crystal-structure analysis of factor Xa bound to a quaternary ammonium ion ligand. *Angew Chem Int Ed Engl* 44:4400–4404.
43. Bagheri B, Baumeier B, Karttunen M (2016) Getting excited: Challenges in quantum-classical studies of excitons in polymeric systems. *Phys Chem Chem Phys* 18: 30297–30304.
44. Boosalis MS, Petruska J, Goodman MF (1987) DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J Biol Chem* 262:14689–14696.
45. Petruska J, et al. (1988) Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc Natl Acad Sci USA* 85:6252–6256.
46. Goodman MF, Creighton S, Bloom LB, Petruska J, Kunkel TA (1993) Biochemical basis of DNA replication fidelity. *Crit Rev Biochem Mol Biol* 28:83–126.
47. von Watzdorf J, Leitner K, Marx A (2016) Modified nucleotides for discrimination between cytosine and the epigenetic marker 5-methylcytosine. *Angew Chem Int Ed Engl* 55:3229–3232.