

Structures of DNA Polymerases Caught Processing Size-Augmented Nucleotide Probes**

Karin Betz, Frank Streckenbach, Andreas Schnur, Thomas Exner, Wolfram Welte, Kay Diederichs, and Andreas Marx*

Dedicated to Professor Richard R. Schmidt on the occasion of his 75th birthday

The integrity of the genome relies primarily on the ability of DNA polymerases to efficiently catalyze selective DNA synthesis according to the Watson–Crick rule in a template-directed manner during DNA replication, repair, and recombination. Remarkably, some DNA polymerases achieve selective information transfer to the offspring in line with the Watson–Crick rule with intrinsic error rates as low as one mistake per one million synthesized nucleotides.^[1] This is far below the value that would be expected based on the energetic differences between canonical (i.e. Watson–Crick) and noncanonical nucleobase pairing.^[1] Geometric factors are widely cited to govern DNA polymerase selectivity.^[1] Thus, high-fidelity DNA polymerases are believed to mostly select the canonical nucleotide based on close steric complementarity of the nascent base pair to the geometry of the active site of the enzyme.

Synthetic nucleotide analogues with tailored new properties are widely used to investigate the mechanisms that govern DNA polymerase selectivity.^[2] In this context nucleotides with hydrophobic nucleobase isosteres were developed to study the contribution of hydrogen bonding to DNA polymerase selectivity.^[1d,e,3] To investigate the steric effects in interactions between DNA polymerase and its substrate, we have developed 4'-alkyl-modified nucleotides with increasing steric demand.^[4] The modifications were introduced in the sugar residues since structural and functional data of DNA polymerases show that the sugar residues of nucleotides are involved in substrate recognition. These interactions may provide the enzymes with additional paths for achieving selectivity besides editing nucleobase geometry. Alkyl groups were employed since their effects on the hydrogen-bonding patterns and conformations of the nucleotides would be minimized. By employing these size-augmented nucleotides in functional DNA polymerase studies we could evaluate varied steric effects on DNA polymerase selectivity. Herein

we report the first crystal structures of size-augmented 4'-methylated and 4'-ethylated thymidine triphosphates in complex with a DNA polymerase.

Significant mechanistic insights into nucleotide incorporation during DNA polymerization were derived from the high-resolution crystal structures reported by Waksman and colleagues^[5] of KlenTaq, an N-terminally truncated form of DNA polymerase from *Thermus aquaticus*.^[6] KlenTaq is a member of the family A DNA polymerases that play a role in prokaryotic and eukaryotic DNA replication and repair.^[7,8] Since data on the action of KlenTaq on 4'-alkylated thymidine-5'-triphosphates (dT^RTPs) is lacking, we first studied nucleotide incorporation by employing transient kinetic analysis with quench–flow technology (Table 1). Our data

Table 1: Transient kinetic parameters for nucleotide incorporation by the KlenTaq wild-type and I614 mutant.

| dT ^R TP | K_D [μM] | k_{pol} [s^{-1}] | k_{pol}/K_D [$\text{M}^{-1}\cdot\text{s}^{-1}$] | Rel. eff. ^[a] |
|---------------------|-------------------------|--------------------------------------|--|--------------------------|
| <i>wild-type</i> | | | | |
| dT ^H TP | 25.0 ± 3.1 | 8.80 ± 0.34 | 352000 | 1 |
| dT ^{Me} TP | 119 ± 20 | 0.17 ± 0.01 | 1430 | 0.0041 |
| dT ^{Et} TP | 115 ± 16 | 0.13 ± 0.01 | 1130 | 0.0032 |
| <i>I614A</i> | | | | |
| dT ^H TP | 6.83 ± 0.55 | 13.3 ± 0.3 | 1950000 | 1 |
| dT ^{Me} TP | 10.4 ± 0.8 | 3.62 ± 0.08 | 350000 | 0.18 |
| dT ^{Et} TP | 4.44 ± 0.37 | 2.01 ± 0.04 | 450000 | 0.23 |

[a] Rel. eff.: relative efficiency; efficiency: k_{pol}/K_D .

shows that KlenTaq incorporates the size-augmented nucleotides with reduced efficiency (k_{pol}/K_D) compared to the natural counterpart. While K_D was reduced by a factor of 4–5, 51- and 67-fold reduction was observed for k_{pol} . Interestingly, the two nucleotides yielded similar kinetic constants. Thus, the effect on incorporation efficiency upon introduction of one additional methylene residue in the natural nucleotide is far larger than the difference between the methyl and ethyl modifications. Similar effects were observed in studies of enzymes from the same sequence family.^[4a,f,g]

To determine the structural basis for the processing of dT^RTP by KlenTaq we attempted to obtain crystals of the enzyme in complex with primer, template, and incoming dT^RTP. Indeed, using a crystallization strategy reminiscent to one described,^[5] we were able to obtain several crystals and could solve their structures. These structures of the ternary complexes of KlenTaq with bound dT^RTP were obtained by

[*] K. Betz,^[†] F. Streckenbach,^[†] Dr. A. Schnur, Prof. Dr. T. Exner, Prof. Dr. W. Welte, Prof. Dr. K. Diederichs, Prof. Dr. A. Marx
Departments of Chemistry and Biology
Konstanz Research School Chemical Biology, Universität Konstanz
Universitätsstrasse 10, 78457 Konstanz (Germany)
Fax: (+49) 7531-88-5140
E-mail: andreas.marx@uni-konstanz.de

[†] These authors contributed equally.

[**] We gratefully acknowledge funding by the DFG.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.200905724>.

combining the enzyme with an 11 nucleotide (nt) primer strand and a complementary 16nt template, which contained a 3'-GA-5' sequence next to the 3'-end of the primer strand, in the appropriate buffer. Addition of 2',3'-dideoxycytidine-5'-triphosphate (ddCTP) and dT^RTP results in the extension of the primer strand and incorporation of 2',3'-dideoxycytidine-5'-monophosphate. Since the resulting primer terminus is devoid of any 3'-OH group of the sugar, further extension is prevented and the dT^RTP was expected to be bound in the dNTP binding site.

The obtained crystals diffracted to a resolution of 2.0 and 2.2 Å, respectively, and the structures were solved by difference Fourier techniques (for details see Table S1 in the Supporting Information). The structures provide snapshots of dT^MTP and dT^ETP incorporation opposite template adenine.

The overall structures of KlenTaq in complex containing dT^MTP and dT^ETP are very similar to the KlenTaq structure (PDB code: 1QTM) described earlier^[5e] complexed with ddT^HTP (0.34 Å and 0.31 Å root mean square deviation for C_α atoms) indicating that the modifications have only little effect on the overall enzyme conformation (Figure 1 a). Especially in the active site the amino acid backbone shows only minimal deviations. The same holds true for the primer and template DNA strand. The DNA is mostly in the B form, except for the three nucleotide pairs at the end of the duplex DNA adjacent to the active site, which are in the A form (see

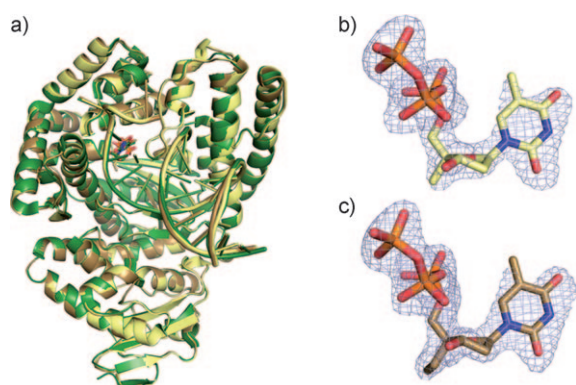


Figure 1. Structures of KlenTaq ternary complexes. a) Overall structures of KlenTaq (ribbon representation) containing primer, template, dT^MTP (yellow), and dT^ETP (bronze), respectively, superimposed on the known structure containing ddT^HTP (green). The 4'-alkylated nucleotides are shown as stick models. Close-up views of b) dT^MTP (C yellow, O red, N blue, P orange) and c) dT^ETP (C bronze) in the complexes. The structures are superimposed with the final refined mFo-DFc simulated annealing omit maps at 4σ. All figures were produced using PyMOL.^[9]

Table S2 in the Supporting Information). The same was reported in the structures containing nonmodified nucleotides.^[5] Interestingly, in the earlier reported structures of KlenTaq in ternary complexes with a 16nt template the electron density map allowed modeling of 14 nucleotides only.^[5] In the structures solved in this work the entire template was resolved. The two 5'-terminal nucleotides of the template strand point towards the thumb domain of the

polymerase and adopt slightly different orientations in the two structures (see Figure S1 in the Supporting Information). The difference electron density maps at the nucleotide binding site indicated that dT^MTP and dT^ETP are indeed trapped in the structure. The best fit of the model to the electron density indicates that the 2'-deoxyribose of the nucleotides are puckered in the N conformation (Figure 1 b,c) similar to the structure obtained for ddT^HTP (see Figure S2 in the Supporting Information). Noteworthy, the orientation of the thymine nucleobase towards the coding adenine and the triphosphate moiety are similar to data obtained with the natural nucleotide. Binding of the nucleotide triphosphate by the enzyme is accompanied with structural reorganizations. In particular the movement of the O helix was discussed to be crucial for guiding the incoming nucleotide to the active site and its closure to promote subsequent catalysis.^[10] In the structures reported here the conformation of the O helices is similar to that of the reported structure comprising ddT^HTP (see Figure S3 in the Supporting Information).

The data obtained for dT^MTP and dT^ETP indicate that the closing motion proceed to completion and all components of the active site are correctly aligned to allow the enzyme to catalyze formation of the phosphodiester bond (Figure 2).

Similar conformations are found for the structure containing ddT^HTP reported earlier.^[5] Only subtle differences in terms of alignment of the incoming nucleotides, magnesium ions, and water molecules are observed (Figure 2). The dT^MTP-, dT^ETP-, and ddT^HTP-containing structures superimpose well, suggesting that the enzyme employs similar mechanistic routes for nucleotide incorporation. However, our kinetic studies show that the phosphodiester bond to the

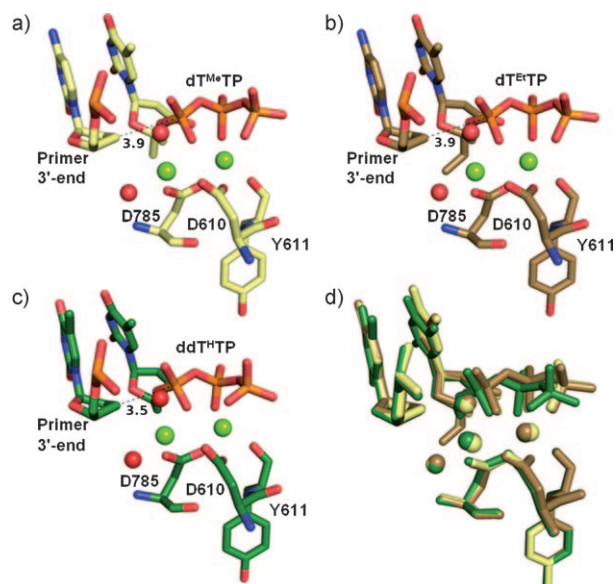


Figure 2. Close-up views of the primer 3'-end, incoming nucleotide, and amino acids complexing divalent cations; color code as in Figure 1 b,c. a) Stick model from the structure derived from the ternary complex of KlenTaq, primer, template, and dT^MTP. Water molecules (red) and magnesium ions (green) are shown as spheres. b) Same as in (a) for the ternary complex containing dT^ETP. c) Same as in (a) using the reported structure containing ddT^HTP. d) Superposition of the structures depicted in (a)–(c) using the color code of Figure 1 a.

incoming dT^RTP analogues is formed with significantly lower efficiency (Table 1). Distances of 3.9 Å between C3' and the phosphorus atom of the α -phosphate are observed in the structures that contain dT^{Me}TP and dT^{Et}TP, respectively. A shorter distance of 3.5 Å was found in the ternary structure of KlenTaq complexed with ddT^HTP reported previously.^[5e] Even if this increase in distance correlates with the decrease in reaction rates, one has to note that the structural variations are within the error of model refinement at this resolution.

Recent studies show that the closing of the O helix is too fast to be rate limiting.^[11] Thus, other steps occurring after the closing of the O helix and the correct placement of the nascent nucleobase are rate limiting. Since it is not possible to experimentally trap additional intermediates of the chemical reaction, we can only speculate on the reasons for the different kinetics. Previously, Warshel et al. performed empirical valence bond calculations on T7 DNA polymerase; they found that the nucleophilic attack, and not the formation of the ternary complex, is rate limiting.^[12] The simulations also suggested that very minor conformational differences introduced by the mismatched base pairs could result in increased reaction barriers. Studies on DNA polymerase β came to the same conclusion: subtle changes in the binding patterns of the substrate can lead to profound changes in the reaction mechanism.^[13] In a similar way, the steric bulk introduced by the additional alkyl groups of the probes could force the reaction to follow a route with a higher reaction barrier. Finally, also the opening step could be influenced by the modified steric demands.^[14]

To gain insight into potentially crucial interactions of the modified nucleotides with proximate amino acid residues we inspected the protein environment in more detail. The 4'-alkyl modifications are buried in a cavity within the nucleotide-binding pocket (Figure 3). The shortest distances between the nucleotides and the proximal amino acid residues were measured between the 4'-methyl group and the methylene moiety of the 4'-ethyl group to amino acids I614 (3.5 Å for dT^{Me}TP and dT^{Et}TP, respectively) and E615 (3.4 and 3.3 Å, respectively) rendering these residues prime targets for mutational analysis. Recent results highlight the necessity of a carboxylate at position 615 for enzyme catalysis.^[15] Thus, we focused our investigations on I614, which had been previously discussed to be involved in selectivity and sugar recognition.^[15,16] We converted isoleucine into alanine by site-directed mutagenesis. The I614A mutation should remove steric constraints within the active site of KlenTaq because of the absence of a β -alkyl side chain. The purified mutant was investigated by transient kinetic analysis as described above (Table 1). Strikingly, the I614A mutant processes the 4'-alkylated nucleotides dT^{Me}TP and dT^{Et}TP with the same activity as the wild-type enzyme processes the unmodified dT^HTP. This results in more than 240- and 400-fold increased incorporation efficiencies for the modified nucleotides, respectively. Interestingly, this is paralleled by a sixfold increase for the efficiency of processing dT^HTP by the I614A mutant in comparison to KlenTaq wild-type. Thus, decrease of steric demand in the active site of the enzyme by mutating the amino acid I614 proximate to the 4'-alkyl modification results in steep increases of incorporation

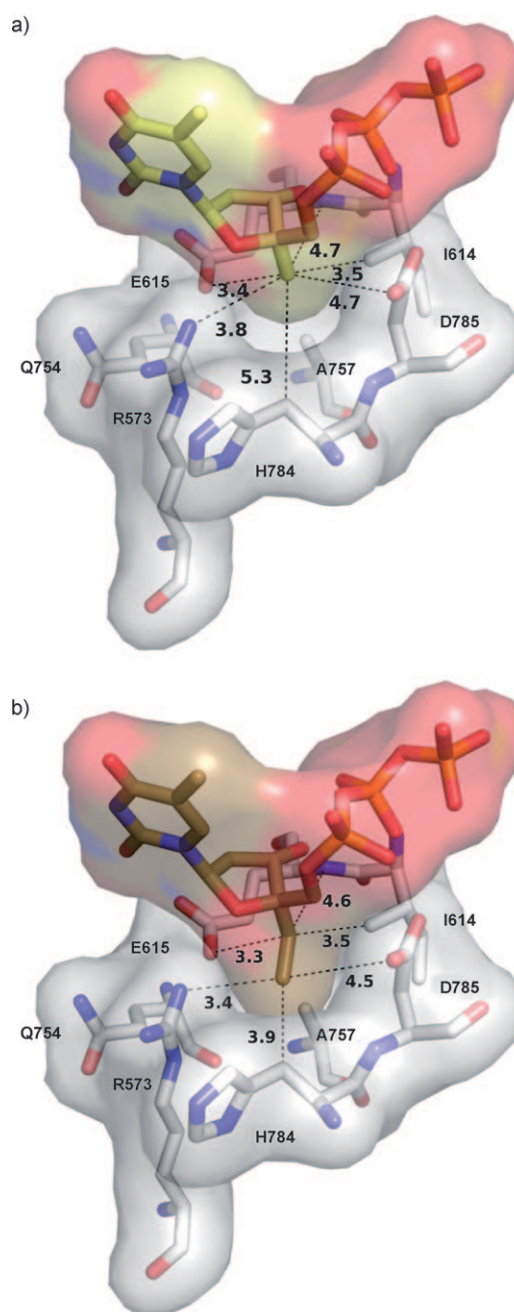


Figure 3. Close-up views of the amino acid residues closest to the 4'-modifications of dT^{Me}TP (a) and dT^{Et}TP (b); color code as in Figure 1 b,c. The stick models are shown inside transparent Connolly surfaces.

efficiencies of the size augmented nucleotides. This finding shows that I614, which is highly conserved from bacteria to humans among DNA polymerases from this sequence family (see Table S3 in the Supporting Information), plays an important role for sugar recognition by KlenTaq.

In summary, we have described the structures of size-augmented nucleotide probes bound to the active site of a DNA polymerase. The nucleotide probes adopt conformations nearly identical to those observed for unmodified substrates. The same holds true for the protein side chains.

Since the kinetics of nucleotide incorporation differ significantly between the natural and size-augmented nucleotides, the finding of nearly identical enzyme and substrate conformations suggests that mechanistic steps not resolved by the crystal structures are rate limiting. The interplay with distinct enzyme side chains may account for this observation. Indeed, mutation of a single side chain most proximate to the 4'-modification increases nucleotide incorporation efficiencies for the size-augmented nucleotides to the level found for the unmodified substrate. Future efforts on elucidation of the interplay of size-augmented nucleotides with DNA polymerases in atomic detail will focus on gaining insight into altered selectivity that derive from enzyme mutation.

Received: October 12, 2009

Revised: January 5, 2010

Keywords: DNA · DNA polymerase · DNA replication · nucleotides · structure elucidation

- [1] a) H. Echols, M. F. Goodman, *Annu. Rev. Biochem.* **1991**, *60*, 477; b) M. F. Goodman, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10493; c) T. A. Kunkel, K. Bebenek, *Annu. Rev. Biochem.* **2000**, *69*, 497; d) E. T. Kool, J. C. Morales, K. M. Guckian, *Angew. Chem.* **2000**, *112*, 1046; *Angew. Chem. Int. Ed.* **2000**, *39*, 990; e) E. T. Kool, *Annu. Rev. Biochem.* **2002**, *71*, 191; f) A. J. Berdis, *Chem. Rev.* **2009**, *109*, 2862.
- [2] K.-H. Jung, A. Marx, *Cell. Mol. Life Sci.* **2005**, *62*, 2080.
- [3] a) T. W. Kim, J. C. Delaney, J. M. Essigmann, E. T. Kool, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15803; b) S. Mizukami, T. W. Kim, S. A. Helquist, E. T. Kool, *Biochemistry* **2006**, *45*, 2772; c) T. W. Kim, L. G. Briebe, T. Ellenberger, E. T. Kool, *J. Biol. Chem.* **2006**, *281*, 2289; d) H. O. Sintim, E. T. Kool, *J. Am. Chem. Soc.* **2006**, *128*, 396; e) H. O. Sintim, E. T. Kool, *Angew. Chem. Int. Ed.* **2006**, *45*, 1974; *Angew. Chem.* **2006**, *118*, 2008; f) A. P. Silverman, Q. Jiang, M. F. Goodman, E. T. Kool, *Biochemistry* **2007**, *46*, 13874.
- [4] a) D. Summerer, A. Marx, *Angew. Chem. Int. Ed.* **2001**, *40*, 3693; *Angew. Chem.* **2001**, *113*, 3806; b) M. Strerath, D. Summerer, A. Marx, *ChemBioChem* **2002**, *3*, 578; c) M. Strerath, J. Cramer, T. Restle, A. Marx, *J. Am. Chem. Soc.* **2002**, *124*, 11230; d) J. Cramer, M. Strerath, A. Marx, T. Restle, *J. Biol. Chem.* **2002**, *277*, 43593; e) F. Di Pasquale, D. Fischer, D. Grohmann, T. Restle, A. Geyer, A. Marx, *J. Am. Chem. Soc.* **2008**, *130*, 10748; f) J. Cramer, G. Rangam, A. Marx, T. Restle, *ChemBioChem* **2008**, *9*, 1243; g) F. Streckenbach, G. Rangam, H. M. Möller, A. Marx, *ChemBioChem* **2009**, *10*, 1630.
- [5] a) S. Korolev, M. Nayal, W. M. Barnes, E. Di Cera, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9264; b) Y. Li, S. Korolev, G. Waksman, *EMBO J.* **1998**, *17*, 7514; c) Y. Li, V. Mitaxov, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9491; d) Y. Li, G. Waksman, *Protein Sci.* **2001**, *10*, 1225.
- [6] W. M. Barnes, *Gene* **1992**, *112*, 29.
- [7] a) U. Hübscher, G. Maga, S. Spadari, *Annu. Rev. Biochem.* **2002**, *71*, 133; b) M. F. Goodman, *Annu. Rev. Biochem.* **2002**, *71*, 17; c) W. Yang, R. Woodgate, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15591.
- [8] M. Seki, C. Masutani, L. W. Yang, A. Schuffert, S. Iwai, I. Bahar, R. D. Wood, *EMBO J.* **2004**, *23*, 4484.
- [9] W. L. DeLano, DeLano Scientific, Palo Alto, CA, USA, **2002**.
- [10] a) H. Pelletier, M. R. Sawaya, A. Kumar, S. H. Wilson, J. Kraut, *Science* **1994**, *264*, 1891; b) S. Doublié, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251; c) H. F. Huang, R. Chopra, G. L. Verdine, S. C. Harrison, *Science* **1998**, *282*, 1669; d) M. C. Franklin, J. Wang, T. A. Steitz, *Cell* **2001**, *105*, 657; e) H. Ling, F. Boudsocq, R. Woodgate, W. Yang, *Cell* **2001**, *107*, 91; f) H. Ling, F. Boudsocq, B. S. Plosky, R. Woodgate, W. Yang, *Nature* **2003**, *424*, 1083.
- [11] P. J. Rothwell, V. Mitaksov, G. Waksman, *Mol. Cell* **2005**, *19*, 345.
- [12] a) J. Florián, A. Warshel, M. F. Goodman, *J. Phys. Chem. B* **2002**, *106*, 5739; b) J. Florián, M. F. Goodman, A. Warshel, *J. Am. Chem. Soc.* **2003**, *125*, 8163; c) J. Florián, M. F. Goodman, A. Warshel, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6819.
- [13] a) K. Arora, W. A. Beard, S. H. Willson, T. Schlick, *Biochemistry* **2005**, *44*, 13328; b) R. Radhakrishnan, T. Schlick, *J. Am. Chem. Soc.* **2005**, *127*, 13245; c) R. Radhakrishnan, T. Schlick, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5970.
- [14] L. Yang, W. A. Beard, S. H. Wilson, B. Roux, S. Broyde, T. Schlick, *J. Mol. Biol.* **2002**, *321*, 459.
- [15] P. H. Patel, L. A. Loeb, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5095.
- [16] P. H. Patel, H. Kawate, E. Adman, M. Ashbach, L. A. Loeb, *J. Biol. Chem.* **2001**, *276*, 5044.