

# Learning from Directed Evolution: *Thermus aquaticus* DNA Polymerase Mutants with Translesion Synthesis Activity

Samra Obeid,<sup>[a]</sup> Andreas Schnur,<sup>[a, b]</sup> Christian Gloeckner,<sup>[a]</sup> Nina Blatter,<sup>[a]</sup> Wolfram Welte,<sup>[b]</sup> Kay Diederichs,<sup>[b]</sup> and Andreas Marx<sup>\*[a]</sup>

DNA is being constantly damaged by endo- and exogenous agents such as reactive oxygen species, chemicals, radioactivity, and ultraviolet radiation. Additionally, DNA is inherently labile, and this can result in, for example, the spontaneous hydrolysis of the glycosidic bond that connects the sugar and the nucleobase moieties in DNA; this results in abasic sites. It has long been obscure how cells achieve DNA synthesis past these lesions, and only recently has it been discovered that several specialized DNA polymerases are involved in translesion synthesis. The underlying mechanisms that render one DNA polymerase competent in translesion synthesis while another DNA polymerase fails are still indistinct. Recently two variants of *Taq* DNA polymerase that exhibited higher lesion bypass ability than the wild-type enzyme were identified by directed-evolution approaches. Strikingly, in both approaches it was independently found that substitution of a single nonpolar amino acid side chain by a cationic side chain increases the

capability of translesion synthesis. Here, we combined both mutations in a single enzyme. We found that the *KlenTaq* DNA polymerase that bore both mutations superseded the wild-type as well as the respective single mutants in translesion-bypass proficiency. Further insights in the molecular basis of the detected gain of translesion-synthesis function were obtained by structural studies of DNA polymerase variants caught in processing canonical and damaged substrates. We found that increased positive charge of the surface potential in the area proximal to the negatively charged substrates promotes translesion synthesis by *KlenTaq* DNA polymerase, an enzyme that has very limited naturally evolved capability to perform translesion synthesis. Since expanded positively charged surface potential areas are also found in naturally evolved translesion DNA polymerases, our results underscore the impact of charge on the proficiency of naturally evolved translesion DNA polymerases.

## Introduction

DNA polymerases are responsible for all DNA synthesis in nature: in DNA replication, repair, and recombination. Efficient DNA synthesis relies on recognition of the parental DNA strand that serves as the template for the selection and incorporation of canonical nucleotides into the nascent DNA strand. However, DNA is constantly being damaged by endogenous and exogenous agents such as reactive oxygen species, chemicals, radioactivity, and ultraviolet radiation.<sup>[1,2]</sup> Additionally, DNA is inherently labile, and can experience, for example, spontaneous hydrolysis of the glycosidic bond that connects the sugar and the nucleobase moieties; this results in abasic sites.<sup>[3–5]</sup> These lesions are the most frequent DNA lesions observed under physiological conditions, with an estimated 10 000 abasic sites per human cell per day; these arise mainly from hydrolysis of the glycosidic bonds that connect purines to their ribose units. Since these lesions lack genetic information, they are potentially mutagenic. Most DNA damage is restored by DNA repair systems that use the sister strand to guide incorporation of the right nucleotide in places of the lesion. However, undetected lesions, and those formed during S-phase, pose a challenge to DNA polymerases. Indeed, abasic sites as well as other lesions are strong blockers of DNA synthesis catalyzed by replicative DNA polymerases.<sup>[6–7]</sup>

Although abasic sites are devoid of any base-coding information and are thus considered “noninstructive”, in vitro and

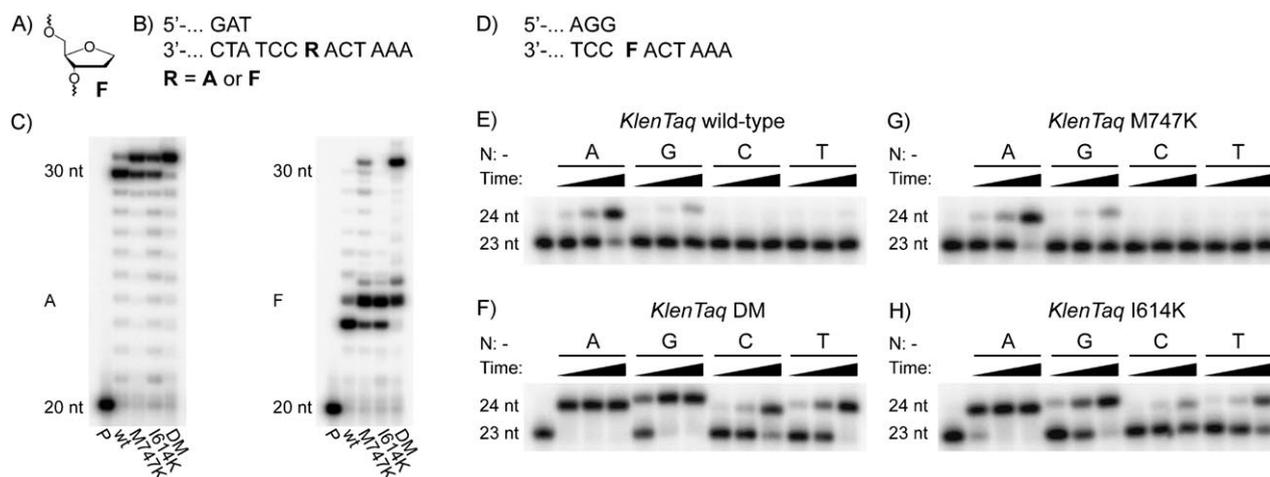
in vivo studies of abasic sites and of the stabilized tetrahydrofuran analogue F (Figure 1A) have shown that adenine and, to a lesser extent, guanine are most frequently incorporated opposite the lesion.<sup>[8–15]</sup> This preference of DNA polymerases for adenine incorporation has been termed the “A-rule”. Intrinsic properties of purines, such as the superior stacking ability of an incoming purine nucleotide to the nucleobase  $\pi$ -system at the primer end, are proposed as the molecular basis for the A-rule.<sup>[16–28]</sup> However it was recently shown that a DNA polymerase that follows the A-rule uses an evolutionary conserved tyrosine side chain of the protein to mimic the absent nucleobase for selective incorporation of purines when bypassing

[a] S. Obeid,<sup>+</sup> Dr. A. Schnur,<sup>+</sup> Dr. C. Gloeckner,<sup>+</sup> N. Blatter, Prof. Dr. A. Marx  
Department of Chemistry, Konstanz Research School Chemical Biology  
University of Konstanz  
Universitätsstrasse 10, 78457 Konstanz (Germany)  
Fax: (+49) 7531-885140  
E-mail: andreas.marx@uni-konstanz.de

[b] Dr. A. Schnur,<sup>+</sup> Prof. Dr. W. Welte, Prof. Dr. K. Diederichs  
Department of Biology, Konstanz Research School Chemical Biology  
University of Konstanz  
Universitätsstrasse 10, 78457 Konstanz (Germany)

[<sup>+</sup>] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201000783>.



**Figure 1.** Comparison of *KlenTaq* wild-type, *KlenTaq* single mutants and *KlenTaq* DM in primer extension and single incorporation experiments opposite an abasic site F. A) Structure of the abasic site analogue F. B) Partial primer (upper) and template (lower) sequences used in primer extension experiments. C) Primer extension with wild-type *KlenTaq* (wt), the single mutants (M747K, I614K), or *KlenTaq* DM. P: primer. D) Partial primer and template sequences used in single nucleotide incorporation experiments. E)–H) Single nucleotide incorporation of dNTP opposite F after 0.5, 2, and 10 min, with wild-type *KlenTaq*, *KlenTaq* DM, *KlenTaq* M747K, and *KlenTaq* I614K, respectively. The specific dNTPs are indicated.

abasic sites.<sup>[29]</sup> This observation is consistent with the geometric model for generic DNA polymerase substrate selection.<sup>[21]</sup>

It has long been obscure how cells perform DNA synthesis past these lesions, and only recently has it been discovered that several specialized DNA polymerases belonging to the new “Y-family” are involved in translesion synthesis.<sup>[6,7,30,31]</sup> One prominent functional characteristic is their high error-propensity when dealing with undamaged DNA; this distinguishes them from the known high-fidelity DNA polymerases. Translesion DNA polymerases are found in a variety of organisms ranging from *E. coli* to human. It has been shown that these enzymes have important cellular functions including, for example, the suppression of cancer.<sup>[6,7,30,31]</sup>

What are the determinants for a DNA polymerase to be proficient in translesion synthesis? Structural and functional studies have added significantly to our understanding of the basic mechanisms of translesion synthesis by DNA polymerases.<sup>[6,7,30,31]</sup> It has been suggested that, because of their larger, sterically more “open” active sites, Y-family translesion DNA polymerases are competent to process aberrant substrates.<sup>[30–33]</sup> Other models suggest that these enzymes surround the nascent base pair closely, but are more flexible, and can thus accommodate altered nucleobase pairs.<sup>[34]</sup> However, data from crystal structures of ternary complexes of translesion and non-translesion DNA polymerases show that the nature of electrostatic interactions between these enzymes and their bound substrates is different; this suggests another cause of their altered properties.<sup>[33,35]</sup> Recently, Loeb and co-workers identified a variant of *Thermus aquaticus* (*Taq*) DNA polymerase, I614K, that exhibited superior lesion-bypass ability than that of the wild-type enzyme.<sup>[36]</sup> The enzyme was found by a directed-evolution approach that combined randomization of the gene, genetic complementation (to identify active mutants), and their subsequent screening. Following a similar approach, we found a methionine to lysine mutant (M747K) of

*KlenTaq* DNA polymerase (an N-terminally truncated form of *Taq* DNA polymerase similar to the Klenow fragment of *E. coli* DNA polymerase I) that also showed a higher activity in bypassing damaged DNA sites.<sup>[37]</sup> Strikingly, in both approaches it was independently found that substitution of a nonpolar amino acid side chain with a cationic side chain increased the capability for translesion synthesis. Motivated by this coincidence, we set out to combine both mutations in one enzyme. Here we describe the functional and structural properties of the resulting double mutant. We found that the DNA polymerase that bore both mutations was superior to the wild-type, as well as the single mutants, in translesion synthesis proficiency. Further insights in the molecular basis of this detected gain of function were obtained from structural studies of the DNA polymerase variants when processing canonical and damaged substrates.

## Results and Discussion

### Mutation of *KlenTaq* DNA polymerase and functional characterization

First, we generated a double mutant of *KlenTaq* DNA polymerase that carried lysines at positions 614 and 747. The resulting enzyme, *KlenTaq* DM, was compared to the wild-type enzyme and to the single-substituted variants (I614K, M747K). All enzymes were Ni-NTA purified and diluted to equal protein concentrations. Next, we investigated the effect of the combination of the two mutations on the ability of *KlenTaq* DNA polymerase to bypass an abasic site. We employed a qualitative primer extension assay by using templates that bore either a 2'-deoxyadenosine (dA) moiety or the abasic site analogue F (Figure 1). Figure 1C shows the autoradiogram of the polyacrylamide gel analysis of the extension products, and it can be seen that all variants were able to extend the primer to

full length in the control reaction (non-damaged DNA template). Additionally, they show an elevated template-independent addition of a nucleotide to the primer strand compared to the wild-type enzyme.

Next, we investigated lesion bypass. The wild-type shows significant pausing prior to incorporation of a nucleotide opposite the abasic site, and minimal formation of the longer reaction product (Figure 1 C). Both single mutants incorporated a nucleotide opposite the lesion with increased efficiency compared to the wild-type enzyme, but only little extension beyond the lesion. *KlenTaq* DM showed the most elevated ability to incorporate a nucleotide opposite the lesion, and promote extension beyond the lesion that resulted in the formation of significant amounts of full-length product. In order to gain insights into which nucleotide is preferentially incorporated opposite the lesion we performed single-nucleotide incorporation studies (Figure 1 D–H). The data obtained suggests that *KlenTaq* DM, as the wild-type, preferentially incorporated a 2'-deoxyadenosine-5'-triphosphate (dATP) opposite the lesion, and thus followed the "A-rule". In order to quantify the data we conducted kinetic experiments by using rapid-quench flow techniques in a single-nucleotide incorporation assay.<sup>[29,38,39]</sup>

We investigated dAMP incorporation opposite abasic and dT sites in a <sup>32</sup>P-radiolabeled primer–template complex. Incorporation of dAMP was found to be most efficient when using *KlenTaq* DM (56-fold more efficient than wild-type) or *KlenTaq* I614K (53-fold more efficient), while the effect of the M747K mutation was only moderate (Table 1). The increase in incorporation efficiencies opposite the abasic site stems from increases in  $k_{\text{pol}}$  and decreases of  $K_{\text{D}}$  compared to the wild-type enzyme. To quantify the selectivity of the mutants in comparison to that of the wild-type, we determined the error spectrum of all

enzymes by employing a reported PCR-based assay.<sup>[36]</sup> The wild-type enzyme exhibited an error rate of  $8.8 \times 10^{-5}$ , which is of the same order as that reported before (Table 2). Interestingly, *KlenTaq* M747K exhibited a similar error rate ( $11.5 \times 10^{-5}$ ), while *KlenTaq* I614K showed an approximately threefold increased error rate. The highest error rate was detected for *KlenTaq* DM (14 times higher than that of the wild-type enzyme).

### Structural characterization

We determined several crystal structures of *KlenTaq* DM. One structure, *KlenTaq* DM<sub>ddCTP</sub> was obtained with the enzyme bound to an undamaged primer–template complex with an incoming 2',3'-dideoxycytidine-5'-triphosphate (ddCTP). A second structure, *KlenTaq* DM<sub>dATP</sub> was solved of a complex of *KlenTaq* DM bound to a primer–template complex containing an abasic site opposite an incoming 2',3'-dideoxyadenosine-5'-triphosphate (ddATP). The crystals were obtained by a strategy previously used for wild-type *KlenTaq*,<sup>[29,39–43]</sup> and the structures were solved by difference Fourier techniques. The structures, determined at 1.7–2.4 Å resolution (see Table S1 in the Supporting Information), provide snapshots of nucleotide incorporation opposite undamaged template (dG), and nucleobase-deficient template (F). The structure of *KlenTaq* DM in complex with ddCTP and undamaged template was very similar to that of the wild-type *KlenTaq* structure (3KTQ) described earlier<sup>[43]</sup> (0.30 Å root mean squared deviation for C $\alpha$  atoms); this indicates that the mutations have little effect on enzyme structure. Subtle changes around position 747 were detected (Figure 2 A and B); however, the substitution of the less-polar amino acid methionine with the positively charged lysine resulted in increased positively charged surface potential near the negatively charged DNA backbone of the template strand (Figure 2 E and F). Close to position 614 different conformations were observed (Figure 2 C and D). Position 614 is located in proximity to motif C, which harbors catalytically essential carboxylates. In *KlenTaq* DM<sub>ddCTP</sub> the I614K mutation introduces water-mediated hydrogen bond network between K614 and the backbone of motif C.

**Table 1.** Kinetic analysis of nucleotide incorporation opposite abasic site F by *KlenTaq* and mutants.

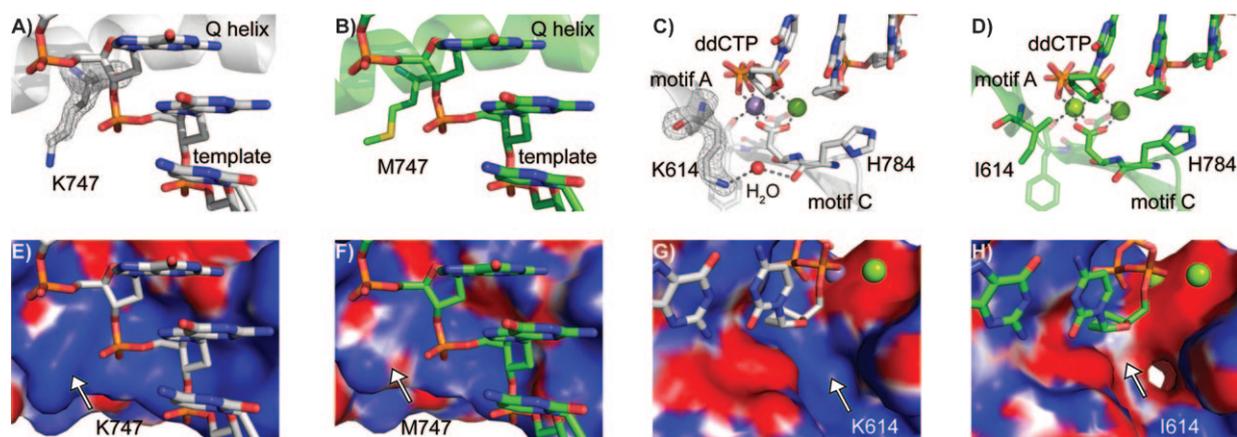
Enzyme	Template	dNTP	$k_{\text{pol}}$ [ $\text{s}^{-1} \times 10^{-2}$ ]	$K_{\text{D}}$ [ $\mu\text{M}$ ]	$k_{\text{pol}}/K_{\text{D}}$ [ $\mu\text{M}^{-1} \text{s}^{-1} \times 10^{-4}$ ]	Rel. eff. <sup>[a]</sup>
<i>KlenTaq</i>	dT	A	516 ± 44	15.3 ± 4.7	3373	1843
<i>KlenTaq</i>	F	A	2.73 ± 0.23	149 ± 35	1.83	1
<i>KlenTaq</i> (M747K)	dT	A	557 ± 47	11.3 ± 3.7	4929	2693
<i>KlenTaq</i> (M747K)	F	A	3.02 ± 0.13	108 ± 15	2.80	1.53
<i>KlenTaq</i> (I614K)	dT	A	351 ± 14	2.23 ± 0.34	15740	8601
<i>KlenTaq</i> (I614K)	F	A	41.9 ± 3.3	43.3 ± 14.4	96.8	53
<i>KlenTaq</i> DM	dT	A	462 ± 36	1.46 ± 0.35	31 644	17 292
<i>KlenTaq</i> DM	F	A	52.7 ± 2.2	51.4 ± 8.6	103	56

[a] Relative efficiency.

**Table 2.** PCR error rate and error spectrum of *KlenTaq* and mutants.

Enzyme	No. of clones	Mutations per clone (average)	Error rate <sup>[a]</sup>	No. of individual substitutions						No. of deletions	No. of insertions
				Transitions		Transversions					
				AT → GC	GC → AT	AT → TA	AT → CG	GC → TA	GC → CG		
<i>KlenTaq</i>	32	0.4	$8.8 \times 10^{-5}$	5	4	0	0	1	0	3	0
<i>KlenTaq</i> (M747K)	29	0.5	$11.5 \times 10^{-5}$	4	5	0	0	1	0	3	1
<i>KlenTaq</i> (I614K)	24	0.9	$26 \times 10^{-5}$	10	7	1	1	0	0	3	0
<i>KlenTaq</i> DM	28	5.5	$125 \times 10^{-5}$	34	43	41	1	13	2	19	2

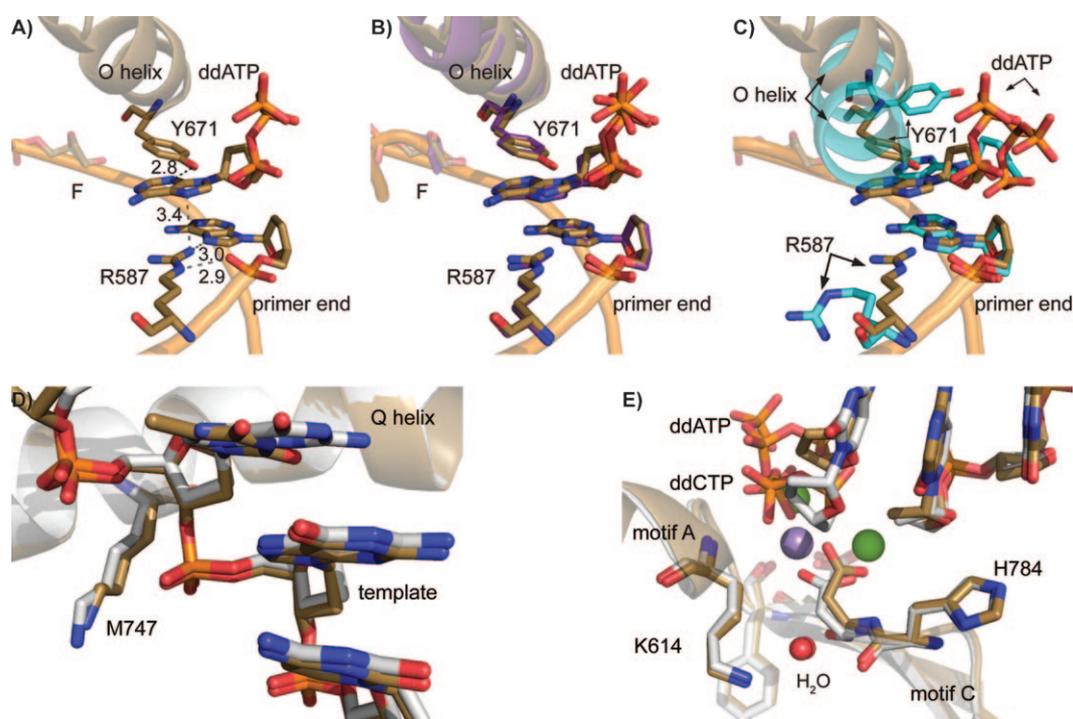
[a] Number of mutations per base per replication.



**Figure 2.** Impact of the mutations on enzyme substrate complex. A) Close-up view of *KlenTaq* DM<sub>ddCTP</sub> at the mutation site M747K. The K747 side chain and the template are shown as sticks, and the Q helix as a ribbon. The final refined simulated annealing omit map mFo-DFc encountered at  $3\sigma$  is shown for K747. B) *KlenTaq* wild-type (PDB ID: 3KTQ) with M747. C) Close-up view of *KlenTaq* DM<sub>ddCTP</sub> at I614K. The K614 side chain in motif A interacts via a water molecule with the backbone of motif C (----). The incoming ddCTP located at the 3' terminus of the primer is coordinated via  $Mn^{2+}$  and  $Mg^{2+}$  ions with the catalytically essential carboxylates (----). The final refined simulated annealing omit map mFo-DFc encountered at  $3\sigma$  is shown for K614. D) *KlenTaq* wild-type (PDB ID: 3KTQ) with I614. E–H) Electrostatic surface potentials of A–D), respectively. Blue and red represent the positive ( $+10\text{ kTe}^{-1}$ ) and negative ( $-10\text{ kTe}^{-1}$ ) charge potentials, respectively. Arrows indicate the amino acid side chains.

*KlenTaq* DM<sub>AP</sub> (the complex with ddATP opposite an abasic template site) adopted an overall conformation very similar to that of 3LWL,<sup>[29]</sup> a recently described wild-type structure complexed with the same substrates (0.27 Å root mean squared

deviation for  $C\alpha$  atoms; Figure 3 A–C). Again, compared to the structures of *KlenTaq* complexed to undamaged substrates, there are several structural changes to the O helix near the primer terminus and the dNTP binding site. In the wild-type



**Figure 3.** Interaction network of incoming ddATP opposite abasic site F. A) Hydrogen bonds stabilize ddATP opposite an abasic site in *KlenTaq* DM<sub>AP</sub> (sand), showing side chains of R587 and Y671. B) Superimposed structures of *KlenTaq* DM<sub>AP</sub> (sand) and *KlenTaq* wild-type (purple, PDB ID: 3LWL) in complex with F-containing template. C) Superimposed structures of *KlenTaq* DM<sub>AP</sub> (sand) and *KlenTaq* wild-type (cyan, PDB ID: 1QSY) in the presence of an undamaged template show the difference in orientation of residues R587 and Y671. D) Overlay of *KlenTaq* DM<sub>ddCTP</sub> (gray) with *KlenTaq* DM<sub>AP</sub> (sand) at M747K. Highlighted is the amino acid side chain K747 and the template strand (stick model), Q helix (cartoon). E) Overlay of *KlenTaq* DM<sub>ddCTP</sub> (gray) with *KlenTaq* DM<sub>AP</sub> (sand) at the mutation site I614K. Highlighted is the amino acid side chain K614 in motif A as stick. K614 interacts via a water molecule with the backbone of motif C. The incoming ddCTP located at the 3' terminus of the primer is coordinated via  $Mn^{2+}$  (violet) and  $Mg^{2+}$  (green) ions with the catalytically essential carboxylates.

structure (as well as the *KlenTaq* DM<sub>ddCTP</sub> with undamaged substrates; see Figure S1) the O helix sits against the template nucleobase and the incoming ddCTP, and thereby “closes” the active site. However, in *KlenTaq* structures that incorporate abasic sites, the O helix is in a conformation that leaves the active site more solvent-exposed (“opened”; Figure 3A–C). These differences in the conformation of the O helix result in variations in the protein–substrate assembly at the active site. In the structures obtained with the abasic site opposite an incoming ddATP the phenol ring of Tyr671 is in a different orientation (Figures 3 and S2). It directly interacts with the nucleobase of the incoming ddATP. The vacant space (missing nucleobase) is filled by the aromatic side chain of the amino acid; this mimics a pyrimidine nucleobase and thus preferentially directs purine incorporation because of the superior geometric fit to the active site. Notably, the amino acid conformations at positions 614 and 747 are similar to those found when the enzyme is processing undamaged template (Figure 3D–E).

An additional feature of many DNA polymerases is their ability to catalyze non-template-directed nucleotide addition to the 3′-termini of blunt-ended DNA, with a strong preference for the incorporation of adenine.<sup>[29,44–47]</sup> The same holds true for *KlenTaq* and *KlenTaq* DM as our investigations have revealed. Our recently published data<sup>[29]</sup> strongly suggest an amino acid side-chain templating mechanism that involves Tyr671 for blunt-end extension—reminiscent of the mechanism for incorporation opposite abasic sites. To obtain structural insights into the template-independent nucleotidyl transferase activity of *KlenTaq* DM, we crystallized the enzyme in the presence of a blunt-end primer–template complex and ddATP (*KlenTaq* DM<sub>BE</sub>) by following our recently employed crystallization strategy.<sup>[29]</sup> By the incorporation of ddAMP at the primer terminus, a blunt-end primer–template duplex lacking a 3′-terminal hydroxyl group is formed, and a ddATP molecule is captured at the active site for template-independent elongation of the primer end. This structure was determined with a resolution of 2.2 Å (Table S1). Here again, when processing a blunt-end primer template substrate, the *KlenTaq* wild-type (3LWM) and *KlenTaq* DM<sub>BE</sub> structures showed high similarity, and the interaction patterns of the Tyr671s are very much alike (Figures S3 and S4).

## Conclusions

As demonstrated, the substitutions of hydrophobic amino acids with cationic residues, I614 and M747 (located near the nucleotide and DNA substrates for *KlenTaq* wild-type), rendered the resulting *KlenTaq* DM more competent at bypassing abasic sites. Both mutations led to an expanded positively charged surface potential area that is in close proximity to the negatively charged substrates (DNA template and incoming dNTP). The positive charge might foster attractive interactions with the anionic substrates and/or stabilize productive enzyme conformations that are required for nucleotide incorporation opposite aberrant template sites such as abasic sites. The relevance of positively charged amino acids on lesion bypass has been recently discussed. Giesecking et al. identified a variant of

human DNA polymerase  $\beta$  from a library of more than 10000 mutants (generated by error-prone PCR) that exhibited superior lesion-bypass ability than the wild-type enzyme.<sup>[48]</sup> It was found that a mutation that has a pronounced effect was the substitution of a negatively charged glutamic acid with a positively charged lysine. Furthermore, the presence of a large positively charged DNA binding surface has been shown to promote DNA-lesion bypass by human DNA polymerase  $\eta$ , an enzyme that is able to perform translesion synthesis, as well as *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4), an intensively studied translesion polymerase.<sup>[33,35]</sup>

Our results show that the increase in translesion-synthesis proficiency comes at the cost of selectivity. Similar effects were reported earlier.<sup>[36]</sup> In addition, natural translesion DNA polymerases, such as the above mentioned human DNA polymerase  $\eta$  and Dpo4, also have elevated error rates.<sup>[6,7,31]</sup> This effect was ascribed to more open solvent-accessibility and to a loose active site compared to those found in high-fidelity enzymes. Our results show that mutations that do not significantly affect active site solvent accessibility but do affect protein surface charge decrease DNA polymerase selectivity. Recent reports indicate that aberrant DNA polymerase conformations are required when the enzyme is processing mismatched substrates.<sup>[30,35]</sup> Again positive charge might foster attractive interactions with the anionic substrates and/or stabilize conformations, including those that are required for mismatch formation.

The model of adaptive evolution concludes that the inherent plasticity of an enzyme allows mutations to change the substrate scope of the enzyme without compromising the general activity of the enzyme.<sup>[49]</sup> Thereby, new and beneficial factors could be provided that allow adaptation to changes in the environment. Our findings that *KlenTaq* DNA polymerase can accommodate mutations that increase the activity of the enzyme and foster translesion synthesis proficiency support this model.

In summary, we found that an increase in the positively charged surface potential of the area in proximity to the negatively charged substrates promotes the lesion-bypass ability of *KlenTaq* DNA polymerase, an enzyme that has a very limited naturally evolved capability to achieve lesion-bypass synthesis. Expanded positively charged surface potential areas are found in naturally evolved translesion DNA polymerases as well. Thus, the results from directed evolution underscore the contribution of charge for the proficiency of naturally evolved translesion DNA polymerases.

In addition, our evolved DNA polymerase is interesting as it exhibits high activity in combination with a lesion-bypass propensity in a single enzyme. Particularly, the propensity to process highly damaged DNA templates in PCR amplification may be useful for forensic applications and the amplification of “ancient” DNA.<sup>[37,50]</sup>

## Experimental Section

**Proteins, nucleotides and oligonucleotides:** *KlenTaq* DNA polymerase mutation, expression, and purification were conducted as

described.<sup>[29]</sup> dNTPs were commercially available (Roche). DNA oligonucleotides were purchased from Purimex, Germany, and double HPLC purified.

**Primer extension assay:** The reaction mixture (20  $\mu$ L) contained 150 nM primer (5'-d(CGT TGG TCC TGA AGG AGG AT)-3'), 225 nM of either conventional template (5'-d(AAA TCA ACC TAT CCT CCT TCA GGA CCA ACG TAC)-3') or F-containing template (5'-d(AAA TCA FCC TAT CCT CCT TCA GGA CCA ACG TAC)-3'), dNTPs (100  $\mu$ M), buffer (Tris-HCl (50 mM, pH 9.2), (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (16 mM), MgCl<sub>2</sub> (2.5 mM), and Tween20 (0.1%)) and the relevant *KlenTaq* polymerase (150 nM). Reaction mixtures were incubated at 72 °C for 30 min. Primer was labeled by using [ $\gamma$ -<sup>32</sup>P]-ATP according to standard techniques. Reactions were stopped by addition of 45  $\mu$ L stop solution (80% (v/v) formamide, EDTA (20 mM), 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol), and analyzed by 12% denaturing PAGE. Visualization was performed by phosphorimaging.

**Single nucleotide incorporation assay:** The reaction mixture (20  $\mu$ L) contained 100 nM primer (5'-d(CGT TGG TCC TGA AGG AGG ATA GG)-3'), 130 nM template (5'-d(AAA TCA FCC TAT CCT CCT TCA GGA CCA ACG TAC)-3'), dNTPs (100  $\mu$ M), buffer (Tris HCl (20 mM, pH 7.5), NaCl (50 mM), and MgCl<sub>2</sub> (2 mM)) and of the relevant *KlenTaq* polymerase (200 nM). Reaction mixtures were incubated at 37 °C; incubation times are provided in the legend of Figure 1. Primer was labeled by using [ $\gamma$ -<sup>32</sup>P]-ATP according to standard techniques. Reactions were stopped by addition of stop solution (45  $\mu$ L, as above) and analyzed by 12% denaturing PAGE. Visualization was performed by phosphorimaging. Blunt-end extension experiments were performed as described above but with DNA primer (5'-d(CGT TGG TCC TGA AGG AGG AT)-3') and template (5'-d(ATC CTC CTT CAG GAC CAA CGA AA)-3').

**Enzyme kinetics:** The rate of single nucleotide incorporation was determined by rapid-quench flow-kinetics in chemical quench-flow apparatus (RQF-3, KinTek Corp., University Park, PA) as described earlier.<sup>[29,40]</sup> For reaction times longer than 5 s a manual quench was performed. In brief, radiolabelled primer-template complex (15  $\mu$ L, 200 nM) and DNA polymerase (2  $\mu$ M) in 15  $\mu$ L reaction buffer (Tris-HCl (20 mM, pH 7.5), NaCl (50 mM), and MgCl<sub>2</sub> (2 mM)) were rapidly mixed with a dNTP solution at 37 °C. Quenching was achieved by adding EDTA (0.3 M) at defined time intervals before mixing with the previously described stop solution. For the analysis of dNTP incorporation opposite dT or F, primer (5'-d(CGT TGG TCC TGA AGG AGG ATA GG)-3') and either template (5'-d(AAA TCA TCC TAT CCT CCT TCA GGA CCA ACG TAC)-3'), or F-containing template (5'-d(AAA TCA FCC TAT CCT CCT TCA GGA CCA ACG TAC)-3') were used. Quenched samples were analyzed by 12% denaturing PAGE followed by phosphorimaging. For kinetic analysis, experimental data were fitted by nonlinear regression by using the program GraphPad Prism 4. The data were fit to a single-exponential Equation (1):

$$[\text{conversion}] = A(1 - \exp(-k_{\text{obs}}t)) \quad (1)$$

The observed catalytic rates ( $k_{\text{obs}}$ ) were then plotted against the dNTP concentrations, and the data were fitted to a hyperbolic equation to determine the  $K_D$  of the incoming nucleotide. The incorporation efficiency is given by  $k_{\text{pol}}/K_D$ .

**Error spectrum:** The error spectrum was determined as described earlier.<sup>[36,37]</sup> Briefly, a ~1.8 kb fragment was PCR amplified followed by subsequent cloning of the product and DNA sequencing. PCR reaction mixture (400  $\mu$ L) contained plasmid harboring the target gene (31.5 fmol), primer (400 nM each), dNTPs (100  $\mu$ M), DNA polymerase (80 nM) in buffer (50 mM Tris-HCl, pH 9.2, 16 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>,

2.5 mM MgCl<sub>2</sub>, and 0.1% Tween20). Thirty PCR cycles were performed for 60 s at 94 °C, 60 s at 65 °C, and 120 s at 72 °C. PCR products were purified by agarose gel electrophoresis and cloned into pASK-IBA37plus (IBA) vector. The cloned vector was transformed into *E. coli*, single colonies were randomly picked, and a stretch of 600 bp was sequenced for each isolated plasmid. The mutation frequency equals the number of mutations per clone and per bp sequenced (600 bp). The error rate was calculated as mutation frequency per number of replications. Replications were calculated from the PCR product yield determined by agarose gel electrophoresis and the amount of starting material (31.5 fmol).

**Crystallization and structure determination:** The closed ternary complexes of *KlenTaq* DM were obtained by incubating *KlenTaq* DM in the presence of DNA primer (5'-d(GAC CAC GGC GC)-3'), various templates (template1: 5'-d(AAA GGG CGC CGT GGT C)-3', template2: 5'-d(AAA FTG CGC CGT GGT C)-3', template3: 5'-d(TG CGC CGT GGT C)-3'), and the appropriate ddNTP to terminate the polymerase reaction.

*KlenTaq* DM<sub>ddCTP</sub>: Crystallization was performed by using purified *KlenTaq* DM (11 mg mL<sup>-1</sup>) in buffer (Tris HCl (20 mM pH 7.5), NaCl (150 mM), EDTA (1 mM), and  $\beta$ -mercaptoethanol (1 mM)), with DNA template1/primer duplex, and ddCTP in a molar ratio of 1:8.3:16.7 and in presence of MgCl<sub>2</sub> (20 mM). The crystallization solution was mixed in 1:1 ratio with a reservoir solution that contained HEPES (0.1 M, pH 7.1), MnCl<sub>2</sub> (20 mM), NaOAc (0.1 M), and PEG 4000 (12% (w/v)). Crystals were produced by the hanging drop vapor diffusion method by equilibrating against 1 mL of the reservoir solution for one week at 18 °C. The crystals were frozen in liquid nitrogen. Datasets were collected at the beamline ID 23-1 at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) at a wavelength of 1.072 Å with an ADSC Quantum Detector. Data reduction was performed with the XDS package.<sup>[51,52]</sup> Datasets from three different crystals were merged by using XSCALE<sup>[51,52]</sup> for higher redundancy. The structure was solved by molecular replacement using PHASER<sup>[53]</sup> and refined by using PHENIX<sup>[54]</sup> at a resolution of 1.7 Å.

*KlenTaq* DM<sub>AP</sub>: The crystallization was set up using purified *KlenTaq* DM (8 mg mL<sup>-1</sup>) in buffer (Tris HCl (20 mM, pH 7.5), NaCl (150 mM), EDTA (1 mM), and  $\beta$ -mercaptoethanol (1 mM)), DNA template2/primer duplex, and ddATP, in a molar ratio of 1:3:40 and in presence of MgCl<sub>2</sub> (20 mM). The crystallization solution was mixed in 1:1 ratio with the reservoir solution (Tris HCl (0.05 M, pH 8.8), NH<sub>4</sub>Cl (0.2 M), CaCl<sub>2</sub> (0.01 M), and 28% PEG 4000). *KlenTaq* DM<sub>BE</sub>: The crystallization was set up using purified *KlenTaq* DM (14 mg mL<sup>-1</sup>) in buffer (as above), DNA template3/primer duplex, and ddATP in a molar ratio of 1:3:20 and in presence of MgCl<sub>2</sub> (20 mM). The crystallization solution was mixed in 1:1 ratio with the reservoir solution (sodium cacodylate (0.05 M, pH 7.0), KCl (0.2 M), Mg(OAc)<sub>2</sub> (0.1 M), and 20% PEG 8000). The crystals of the respective DNA polymerase were produced by the hanging drop vapor-diffusion method by equilibrating against 1 mL of the reservoir solution for five days at 18 °C. They were frozen in liquid nitrogen and kept at 100 K during data collection. Data were measured at the beamline PXI (X06SA) at the Swiss Light Source of the Paul Scherrer Institute (PSI, Villigen, Switzerland), at a wavelength of 1.000 Å and with a PILATUS detector. Data reduction was performed with the XDS package.<sup>[51]</sup> The structures were solved by difference Fourier techniques with *KlenTaq* wild type (PDB ID: 3KTQ) as model. Refinement was performed with PHENIX<sup>[54]</sup> and model rebuilding was done with COOT.<sup>[55]</sup> The structures were solved and refined to a resolution of 2.4 Å for *KlenTaq* DM processing an abasic site, and 1.8 Å for *KlenTaq* DM processing a blunt-end primer template.

## Acknowledgements

We would like to thank the DFG for funding within SPP 1170 (MA 2288/5–3) and the beamline staff of the Swiss Light Source (SLS) and European Synchrotron Radiation Facility (ESRF) for support. S.O. acknowledges funding by the Zukunftskolleg, University of Konstanz.

**Keywords:** A-rule · DNA damage · DNA polymerases · DNA replication · translesion synthesis

- [1] O. D. Schärer, *Angew. Chem.* **2003**, *115*, 3052–3082; *Angew. Chem. Int. Ed.* **2003**, *42*, 2946–2974.
- [2] E. C. Friedberg, G. C. Walker, W. Siede, *DNA Repair and Mutagenesis*, American Society for Microbiology, Washington DC, **1995**.
- [3] T. Lindahl, *Nature* **1993**, *362*, 709–715.
- [4] T. Lindahl, B. Nyberg, *Biochemistry* **1972**, *11*, 3610–3618.
- [5] L. A. Loeb, B. D. Preston, *Annu. Rev. Genet.* **1986**, *20*, 201–230.
- [6] U. Hübscher, G. Maga, S. Spadari, *Annu. Rev. Biochem.* **2002**, *71*, 133–163.
- [7] M. F. Goodman, *Annu. Rev. Biochem.* **2002**, *71*, 17–50.
- [8] S. Avkin, S. Adar, G. Blander, Z. Livneh, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3764–3769.
- [9] M. F. Goodman, H. Cai, L. B. Bloom, R. Eritja, *Ann. N. Y. Acad. Sci.* **1994**, *726*, 132–142.
- [10] C. W. Lawrence, A. Borden, S. K. Banerjee, J. E. LeClerc, *Nucleic Acids Res.* **1990**, *18*, 2153–2157.
- [11] V. Pagès, R. E. Johnson, L. Prakash, S. Prakash, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1170–1175.
- [12] D. Sagher, B. Strauss, *Biochemistry* **1983**, *22*, 4518–4526.
- [13] R. M. Schaaper, T. A. Kunkel, L. A. Loeb, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 487–491.
- [14] S. Shibutani, M. Takeshita, A. P. Grollman, *J. Biol. Chem.* **1997**, *272*, 13916–13922.
- [15] B. S. Strauss, *DNA Repair* **2002**, *1*, 125–135.
- [16] W. A. Beard, D. D. Shock, V. K. Batra, L. C. Pedersen, S. H. Wilson, *J. Biol. Chem.* **2009**, *284*, 31680–31689.
- [17] J. Chen, F. Y. Dupradeau, D. A. Case, C. J. Turner, J. Stubbe, *Nucleic Acids Res.* **2008**, *36*, 253–262.
- [18] K. A. Fiala, C. D. Hypes, Z. Suo, *J. Biol. Chem.* **2007**, *282*, 8188–8198.
- [19] E. Freisinger, A. P. Grollman, H. Miller, C. Kisker, *EMBO J.* **2004**, *23*, 1494–1505.
- [20] M. Hogg, S. S. Wallace, S. Doublé, *EMBO J.* **2004**, *23*, 1483–1493.
- [21] E. T. Kool, *Annu. Rev. Biochem.* **2002**, *71*, 191–219.
- [22] H. Ling, F. Boudsocq, R. Woodgate, W. Yang, *Mol. Cell* **2004**, *13*, 751–762.
- [23] D. T. Nair, R. E. Johnson, L. Prakash, S. Prakash, A. K. Aggarwal, *Structure* **2009**, *17*, 530–537.
- [24] S. K. Randall, R. Eritja, B. E. Kaplan, J. Petruska, M. F. Goodman, *J. Biol. Chem.* **1987**, *262*, 6864–6870.
- [25] E. Z. Reineks, A. J. Berdis, *Biochemistry* **2004**, *43*, 393–404.
- [26] M. Seki, C. Masutani, L. W. Yang, A. Schuffert, S. Iwai, I. Bahar, R. D. Wood, *EMBO J.* **2004**, *23*, 4484–4494.
- [27] J. S. Taylor, *Mutat. Res.* **2002**, *510*, 55–70.
- [28] K. E. Zahn, H. Belrhali, S. S. Wallace, S. Doublé, *Biochemistry* **2007**, *46*, 10551–10561.
- [29] S. Obeid, N. Blatter, R. Kranaster, A. Schnur, W. Welte, K. Diederichs, A. Marx, *EMBO J.* **2010**, *29*, 1738–1747.
- [30] W. Yang, R. Woodgate, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15591–15598.
- [31] S. Prakash, R. E. Johnson, L. Prakash, *Annu. Rev. Biochem.* **2005**, *74*, 317–353.
- [32] W. Yang, *Curr. Opin. Struct. Biol.* **2003**, *13*, 23–30.
- [33] H. Ling, F. Boudsocq, R. Woodgate, W. Yang, *Cell* **2001**, *107*, 91–102.
- [34] S. Mizukami, T. W. Kim, S. A. Helquist, E. T. Kool, *Biochemistry* **2006**, *45*, 2772–2778.
- [35] C. Biertümpfel, Y. Zhao, Y. Kondo, S. Ramón-Maiques, M. Gregory, J. Y. Lee, C. Masutani, A. R. Lehmann, F. Hanaoka, W. Yang, *Nature* **2010**, *465*, 1044–1048.
- [36] P. H. Patel, H. Kawate, E. Adman, M. Ashbach, L. A. Loeb, *J. Biol. Chem.* **2001**, *276*, 5044–5051.
- [37] C. Gloeckner, K. B. M. Sauter, A. Marx, *Angew. Chem.* **2007**, *119*, 3175–3178; *Angew. Chem. Int. Ed.* **2007**, *46*, 3115–3117.
- [38] F. Di Pasquale, D. Fischer, D. Grohmann, T. Restle, A. Geyer, A. Marx, *J. Am. Chem. Soc.* **2008**, *130*, 10748–10757.
- [39] S. Obeid, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21327–21331.
- [40] K. Betz, F. Streckenbach, A. Schnur, W. Welte, K. Diederichs, A. Marx, *Angew. Chem.* **2010**, *49*, 5181–5184; *Angew. Chem. Int. Ed.* **2010**, *49*, 5181–5184.
- [41] S. Korolev, M. Nayal, W. M. Barnes, E. Di Cera, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9264–9268.
- [42] Y. Li, G. Waksman, *Protein Sci.* **2001**, *10*, 1225–1233.
- [43] Y. Li, S. Korolev, G. Waksman, *EMBO J.* **1998**, *17*, 7514–7525.
- [44] J. M. Clark, *Nucleic Acids Res.* **1988**, *16*, 9677–9686.
- [45] J. M. Clark, C. M. Joyce, G. P. Beardsley, *J. Mol. Biol.* **1987**, *198*, 123–127.
- [46] H. Hwang, J. S. Taylor, *Biochemistry* **2004**, *43*, 14612–14623.
- [47] K. A. Fiala, J. A. Brown, H. Ling, A. K. Kshetry, J. Zhang, J.-S. Taylor, W. Yang, Z. Suo, *J. Mol. Biol.* **2007**, *365*, 590–602.
- [48] S. Giesecking, K. Bergen, F. di Pasquale, K. Diederichs, W. Welte, A. Marx, *J. Biol. Chem.* **2011**, *286*, 4011–4020.
- [49] P. H. Patel, L. A. Loeb, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5095–5100.
- [50] M. d'Abbadie, M. Hofreiter, A. Vaisman, D. Loakes, D. Gasparutto, J. Cadet, R. Woodgate, S. Pääbo, P. Holliger, *Nat. Biotechnol.* **2007**, *25*, 939–943.
- [51] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.
- [52] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 133–144.
- [53] L. C. Storoní, A. J. McCoy, R. J. Read, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 432–438.
- [54] P. D. Adams, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58*, 1948–1954.
- [55] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

Received: December 23, 2010  
Published online on April 8, 2011