CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Structural basis for the KlenTaq DNA Polymerase catalyzed Incorporation of Alkene- versus Alkyne-modified Nucleotides

Authors: Andreas Marx, Audrey Hottin, Karin Betz, and Kay Diederichs

To be cited as: *Chem. Eur. J.* 10.1002/chem.201604515

Link to VoR: http://dx.doi.org/10.1002/chem.201604515

Supported by
ACES

WILEY-VCH

10.1002/chem.201604515

WILEY-VCH

FULL PAPER

Structural basis for the KlenTaq DNA Polymerase catalyzed Incorporation of Alkene- versus Alkyne-modified Nucleotides

Audrey Hottin[a], Karin Betz[a], Kay Diederichs[a] and Andreas Marx*[a]

Abstract: Efficient incorporation of modified nucleotides by DNA polymerases is essential for many cutting-edge biomolecular technologies. The present study compares the acceptance of either alkene- or alkyne-modified nucleotides by KlenTaq DNA polymerase and provides structural insights into how 7-deaza-adenosine and deoxyuridine with attached alkene-modifications are incorporated into the growing DNA strand. Thereby, we identified modified nucleotides, which prove to be superior substrates for KlenTaq DNA polymerase compared to their natural analogues. The knowledge can be used to guide future design of functionalized nucleotide building blocks.

Introduction

Nucleobase-functionalized DNA is exploited for many biotechnological applications such as next-generation sequencing with dye-labeled nucleotides[1, 2], the in vitro selection of ligands, like aptamers by SELEX where numerous modifications improve chemical and structural diversity of DNA,[3] post-synthetic DNA-conjugation,[4] and DNA-microarrays.[5, 6] Efficient approaches for the preparation of functionalized oligonucleotides comprise the enzymatic incorporation of modified 2'-deoxyribonucleoside monophosphates (dNMPs) by DNA polymerases via primer extension (PEX) or polymerase chain reaction (PCR).[7, 8] Therefore, it is essential that the modification attached to the nucleotide does not significantly compromise enzyme activity. During the last years, constant efforts have been made to identify modified triphosphates (dNTPs) that are well processed by DNA polymerases.[9-13] Thereby nucleobase modifications are mainly attached at the C5 position of pyrimidines and the C7 position of 7-deazapurines (Figure 1A) to direct the modifications pointing into the developing major groove of the DNA.[14-16] Larger functionalities are often connected to the nucleobase via linkers which vary in composition, length and flexibility. With this strategy a variety of probes, even very bulky ones could successfully be incorporated into DNA oligonucleotides.[17-22] Functional studies revealed that the success of efficient incorporation of a specific modified substrate by a DNA polymerase depends amongst others on the nucleobase and the attachment position,[23, 24] the linker type,[25]

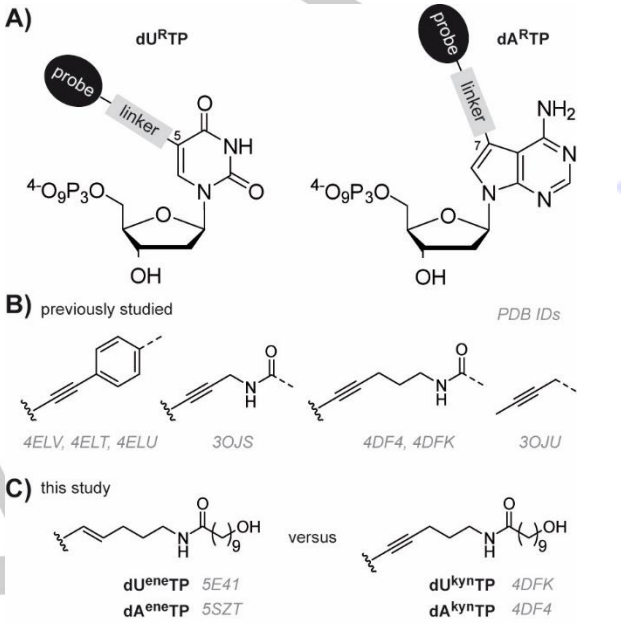


Figure 1. A) General design of modified nucleotide substrates based on dUTP and 7-deaza-dATP. B) Structure of the linkers (or parts of them) that were structurally characterized in the active site of KlenTaq DNA polymerase in our group with the respective PDB codes of the complexes. C) Structures of the linkers investigated in this study attached either to dUTP or 7-deaza-dATP.

the modification coupled and the polymerase used.[26] Recently, there were even some dNTP analogues identified that are better substrates for DNA polymerases than their natural counterparts.[27-29]

To better understand the acceptance of nucleobase-modified substrates by DNA polymerases structural studies were performed in our group (reviewed in [30]). We solved crystal structures of KlenTaq DNA polymerase (large fragment of *Thermus aquaticus* DNA polymerase I) in complex with nucleotides that carry different functionalities attached to the nucleobase via alkyne linkers (Figure 1B). Thereby, interaction patterns were identified that can help the enzyme to tolerate the modifications. In detail, one study suggests that the introduction of an aromatic ring in a rigid, non-polar alkyne linker can improve the processing by newly formed cation-π interactions with positively charged amino acid side chains.[31, 32] Other structures with derivatives harbouring more flexible modifications such as branched polyamide or aminopentynyl moieties highlight the beneficial effects of including functionalities with hydrogen bonding capability that can stabilize the enzyme-substrate complex.[28, 33] Rigid and bulky modifications with little capability to perform stabilizing interactions with the protein, however, lead to low incorporation efficiency, probably due to enzyme/substrate complex perturbation.[33] These results already provide first

[a] Dr. A. Hottin, Dr. K. Betz, Prof. Dr. Kay Diederichs, Prof. Dr. A. Marx
Department of Chemistry and Department of Biology
University of Konstanz
Universitätsstrasse 10, 78457 Konstanz, Germany
E-mail: andreas.marx@uni-konstanz.de

Supporting information for this article is given via a link at the end of the document.

10.1002/chem.201604515

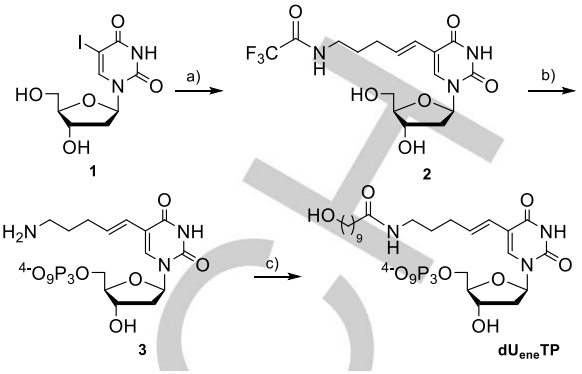
guidelines to tweak incorporation efficiency when designing nucleotides for specific applications. As the described structural studies were solely performed with alkyne-modified nucleotides the expansion to other linkage types is desirable.

Besides alkyne linkers also alkene linkage to the nucleobase is frequently used and the modified nucleotides have proven to be good substrates in PEX and PCR.[12, 25, 34, 35] In addition, alkene linkers present spectroscopic properties that can be of interest for electrochemical studies as redox label.[36] Direct comparison of these two linkage types with the same modification are only sparsely reported. In a study of Lee *et al* different linker types attached to 2'-deoxyuridine triphosphates were evaluated and more rigid *E*-alkene and alkyne linkers turned out to be preferable compared to *Z*-alkenes or highly flexible alkane linkers.[25] Moreover, modified 2'-deoxyuridine triphosphates with *E*-alkenyl linkers performed better in generation of longer DNA products than substrates with the same modifications attached via alkyne linkers. Though there was no significant difference in incorporation efficiency using alkynyl-, *Z*-alkenyl- or alkyl-linkers attached to 7-deaza-2'-deoxyadenosine triphosphates.[35]

As alkene linkers were not yet explored on a structural basis we aim to expand our previous studies with KlenTaq DNA polymerase to get further insights into the interplay of enzyme and substrate. Here, we chose to investigate the performance of alkene-modified nucleotides with the triphosphates **dU^{ene}TP** and **dA^{ene}TP** bearing a 5-N-(10-hydroxydecanoyl)-aminopentenyl-modification that is tailored e.g. for dye labeling of DNA (Figure 1C). To probe the impact of the modification chemistry on DNA polymerase substrate scope, the alkene derivatives were compared to their alkyne counterparts (**dU^{kyn}TP, dA^{kyn}TP**) by competitive incorporation experiments with the respective natural substrates as well as by structural means.

Results and Discussion

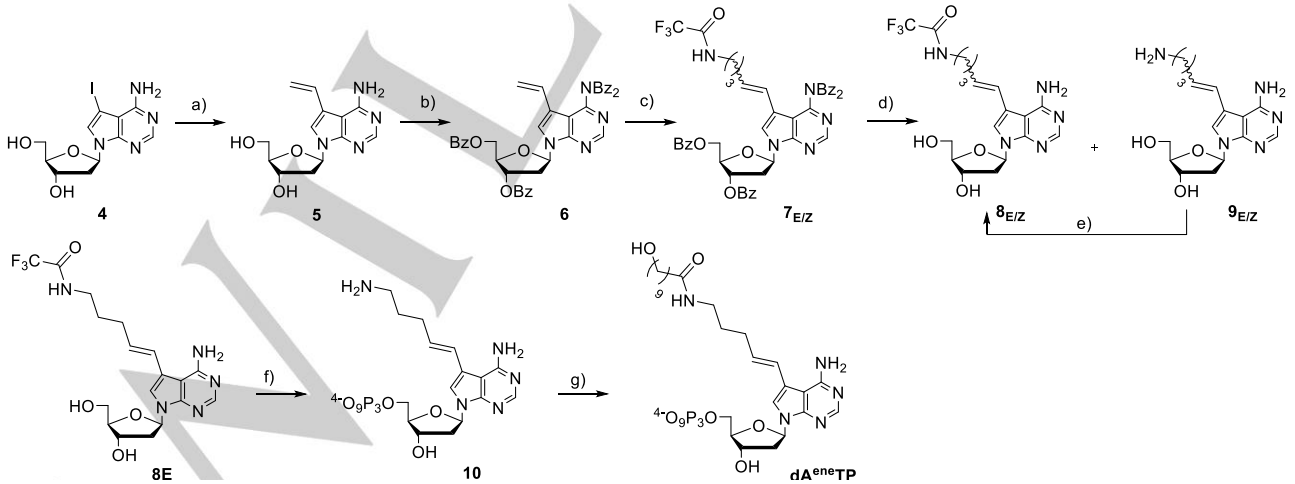
Synthesis of the alkene-modified triphosphates



Scheme 1. Synthesis of modified triphosphate **dU^{ene}TP**: a) $Pd(OAc)_2$, PPh_3, Et_3N, 2,2,2-trifluoro-N-(pent-4-enyl)acetamide, DMF, 100°C, 1 h ; b) proton sponge, $POCl_3$, $PO(OMe)_3$, 0°C, then $(Bu_3NH)_2H_2P_2O_7$, nBu_3N, 0.1 M TEAB buffer, ammonium hydroxide. c) succinimidyl 10-hydroxydecanoate, 0.1 M sodium carbonate buffer (pH = 8.7), DMF, 2 h, rt**.**

dU^{ene}TP: The alkene linker was introduced by palladium catalyzed Heck cross coupling from the unprotected halogenated nucleoside 5-iodo 2´-deoxyuridine **1** with trifluoropentene-acetamide (Scheme 1).[37] The reaction catalyzed by palladium (II) acetate in the presence of triphenylphosphine as ligand and triethylamine as base, leads to the desired trans nucleoside **2** in 14% yield after purification. The low reactivity can be explained by the use of a non-activated alkene for this reaction. Thereupon, nucleoside **2** was converted into uridine triphosphate analogue **3**, using an optimized Yoshikawa phosphorylation procedure.[38] By subsequent coupling of **3**, bearing an amino-linker, with an NHS-activated acid to form the amide bond, triphosphate **dU^{ene}TP** was obtained.

dA^{ene}TP: For the adenosine derivative, the palladium catalyzed Heck cross coupling from 7-deaza 2´-deoxy-7-iodoadenosine only lead to degraded material. Thus, we developed a two-stage strategy for the formation of the alkene modification by first introducing a vinyl group at position C7 of the purine by Stille cross-coupling[39] and then by undergoing a cross-metathesis of



Scheme 2. Synthesis of 2´-deoxyadenosine analogue **dA^{ene}TP**: a) tributyl(vinyl)tin, $Pd(PPh_3)_4$, DMF, 100 °C, 2 h ; b) BzCl, pyr, rt, 2 h, 83% ; c) pentenyltrifluoroacetamide, Grubbs II, toluene, 100 °C, 2 h ; d) 7 N NH_3 in MeOH, 50 °C, overnight, e) CF_3COOEt, DMAP, THF, 90 °C, 4 h. f) proton sponge, $POCl_3$, $PO(OMe)_3$, 0°C, then $(Bu_3NH)_2H_2P_2O_7$, nBu_3N, 0.1 M TEAB buffer, ammonium hydroxide. g) succinimidyl 10-hydroxydecanoate, 0.1 M sodium carbonate buffer (pH = 8.7), DMF, 2 h, rt.

10.1002/chem.201604515

that terminal alkene with a terminal olefin of interest.[40] In detail iodo-adenosine **4** reacts with tributyl(vinyl)tin using $Pd(PPh_3)_4$ as catalyst giving 7-vinyl substituted nucleoside **5** in high yield. Before metathesis, compound **5** is protected with benzoyl groups due to possible Ru-catalyst poisoning by the exocyclic amine. The cross metathesis between adenosine **6** and pentenyltrifluoroacetamide performed in toluene in the presence of Grubbs II catalyst[41] leads to a mixture of diastereoisomeres (E/Z:3/1). The further deprotection of benzoylated derivative 7E/Z using 7N NH_3 in methanol gives nucleoside **8E/Z** as E/Z mixture and compound **9E/Z** that was easily recycled through its protection with a trifluoroacetyl group. Combined fractions of **8E/Z** are then purified by preparative HPLC to isolate the trans-alkene nucleoside. The triphosphorylation of 8E and the subsequent coupling with an NHS-activated acid provides the desired nucleotide **dA^{ene}TP**. Triphosphates **dU^{kyn}TP** and **dA^{kyn}TP** were synthesized as previously described.[6, 28]

Primer extension studies

First, we investigated the acceptance of the four 2'-deoxyribonucleoside triphosphates **dN^{R}TP**s in DNA polymerase catalyzed primer extension reaction. Using a 32P-labeled 22-mer primer and a 36-mer template, each modified **dN^{R}TP** was incubated with KlenTaq DNA polymerase in the presence of the other three natural nucleotides. After 4 minutes, the full-length product was obtained in all cases, clearly showing the ability of the DNA polymerase to use these modified nucleotides as substrates (Figure S1). To investigate the incorporation efficiency of modified versus natural substrate we performed single nucleotide incorporation experiments in which the modified nucleotide competes directly with its natural counterpart. This setup was previously used for the same purpose[33] as well as to study DNA polymerase selectivity.[42] The relative incorporation of natural nucleotide versus modified one was monitored by denaturing polyacrylamide gel electrophoresis (PAGE) and phosphorimaging, taking advantage of the significantly different migration of naturally extended primer compared to modified extended primer due to the bulky modification (Figure 2B). Competition experiments for analogs **dU^{kyn}TP** and **dA^{kyn}TP** have been published before[28] but were repeated under the same conditions (incubation time and enzyme batch & concentration) as for **dU^{ene}TP** and **dA^{ene}TP**. The results for all four nucleotides are summarized in Figure 2A. The competition experiments identified three candidates, which proved to be superior substrates for KlenTaq DNA polymerase compared to their natural analogues, namely **dU^{ene}TP**, **dA^{ene}TP** and **dA^{kyn}TP**. The adenosine derivatives **dA^{ene}TP** and **dA^{kyn}TP** display similar incorporation efficiencies with 4-fold excess of dATP concentration required to obtain an equal ratio of primer extended by dATP and **dA^{R}TP**. KlenTaq incorporates the alkene-modified uridine nucleotide **dU^{ene}TP** equally as dTTP when a 2-fold excess of dTTP is used. However, its alkyne counterpart **dU^{kyn}TP** is a poorer substrate than dTTP for KlenTaq DNA polymerase.[28]

Structural Studies

The interesting results from the incorporation studies prompted us to solve crystal structures of KlenTaq DNA polymerase in complex with **dU^{ene}TP** and **dA^{ene}TP** bound in the active site and to compare them with a natural substrate as well as with the already available data of KlenTaq DNA polymerase with **dU^{kyn}TP** and **dA^{kyn}TP** bound (PDB IDs 4DFK and 4DF4[28]). The primer contained a ddC at the 3'-end to capture the incoming modified **dN^{ene}TP** at the insertion site, yielding a structural snapshot just

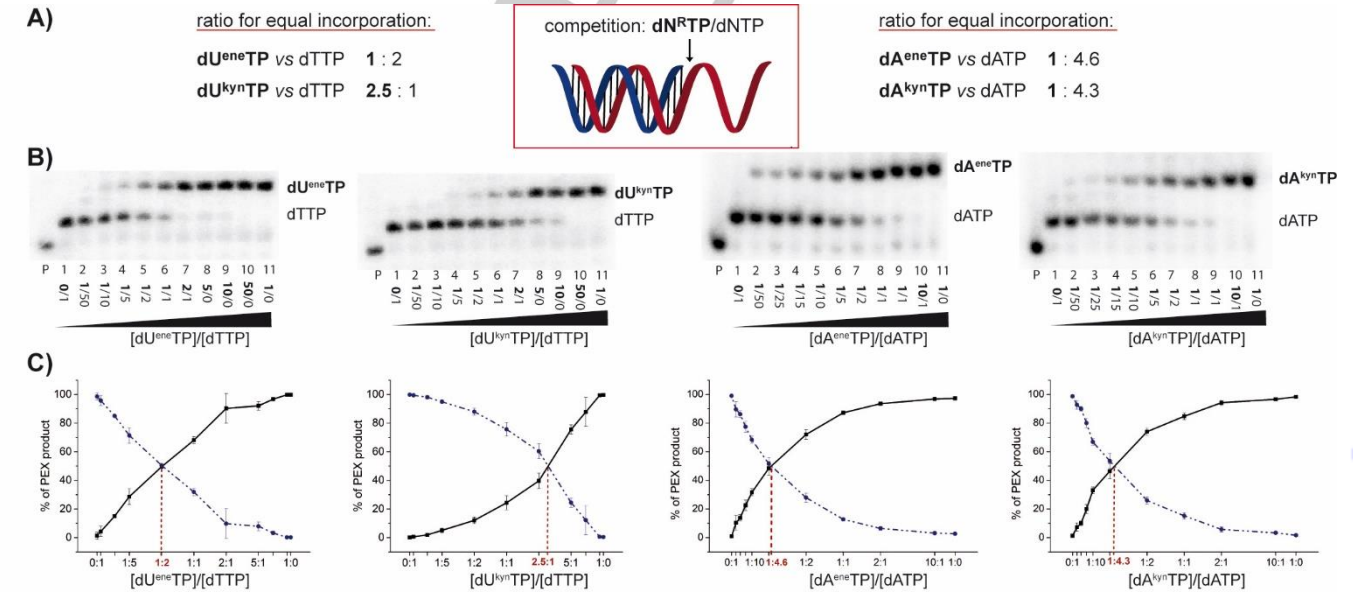


Figure 2. A) Ratio **dN^{R}TP** vs dNTP when both nucleotides are equally incorporated by KlenTaq DNA polymerase. B) PAGE analysis of the competition experiments with various ratios of **dN^{R}TP** versus dNTP . Lane P: 5'-^{32}P-labeled primer; lane 1-11 different ratios of **dN^{R}TP**/dNTP. C) Quantification of the PAGE analysis. The point of equal incorporation is indicated by a red dashed line.

10.1002/chem.201604515

before the imminent incorporation reaction. Ternary complexes, consisting of polymerase, primer/template and the **dN^{ene}TP**s, were obtained by co-crystallization. Crystals exhibit the same space group ($P3_12_12$) and similar cell dimensions as the previously solved KlenTaq DNA polymerase structures. The structures were refined, using data to a resolution of 1.8Å (for data collection and refinement statistics see Table S1).

Ternary complex with dU^{ene}TP

The overall structure of KlenTaq(**dU^{ene}TP**) is similar to KlenTaq DNA polymerase in complex with a ddTTP substrate (PDB ID: 1QTM, core root-mean-square deviation (rmsd): 0.36Å)[43] or the more recently published structure with the natural substrate dCTP bound (PDB ID: 3RTV,[44] core rmsd: 0.26Å Figure 2B) that is referred to as KlenTaq(dCTP) in this paper. The polymerase adopts a closed conformation with the O-helix (part of the finger domain, residues 657-671) covering the nascent dA-**dU^{ene}TP** base pair. The linker does not disturb the complex and usual binding interactions of the substrate, such as canonical Watson-Crick base-pairing to the template or π-stacking with the primer strand, as well as H-bonds of the α and γ phosphate with Lys663 are maintained (Figure 3A,D). Residues in the catalytic site are oriented as observed in KlenTaq(dCTP). The complex is trapped in a state prior to the incorporation reaction where the two catalytic magnesium ions are coordinated by the side chains of Asp610, Asp785, the backbone of Tyr611, the phosphate groups of **dU^{ene}TP** and two water molecules (Figure 2D, E). The distance of the α-phosphate of the substrate to the nucleophilic 3' primer end (for distance measurement the ribose C3 is used) is virtually identical to the natural complex (3.8 Å vs 3.9 Å). The **dU^{ene}TP** sugar, base and triphosphate moieties are well resolved and electron density for the linker is unambiguous up to the amide functionality (Figure 4A and S2A, B). The remaining linker could not be traced and was refined with zero occupancy. The linker extends from the C5 position of **dU^{ene}TP** through a cavity between the finger and thumb domain (denoted cavity A), mainly bordered by the Arg587 side chain and residues from the O-helix as was already observed in KlenTaq(**dU^{kyn}TP**) (Figure 4A, C). The two amino acids Arg660 and Arg587 show different conformations compared to the natural case KlenTaq(dCTP) (Figure 3B, E). Although Arg660 is slightly displaced to make space for the modification it can still maintain its interaction with the primer 3'-end in the closed ternary complex (Figure 3A). The side chain of Arg587 is now oriented along the linker and thereby loses its interaction with the primer backbone. The interaction of the Arg587 main chain with the backbone, however, is maintained. (Figure 3D). Via its amide functionality, the **dU^{ene}TP** linker itself is capable to engage two specific interactions to adjacent O-helix residues: one with the guanidinium group of the already mentioned Arg660 and the other with the side chain of Thr664 (Figure 3A). We assume that these interactions may well stabilize the linker-bearing substrate and contribute to the proficient processing of this analogue by KlenTaq DNA polymerase. Identical enzyme/substrate interactions are observed in the previously solved structure of alkyne modified **dU^{kyn}TP** in which the linker protrudes from the enzyme through the same cavity (Figure 4A,C). However, this nucleotide is less well processed by KlenTaq DNA polymerase than dTTP in our assay. Therefore, other differences between unmodified, dT-alkene- or dT-alkyne-modified nucleotides must occur in the process of substrate

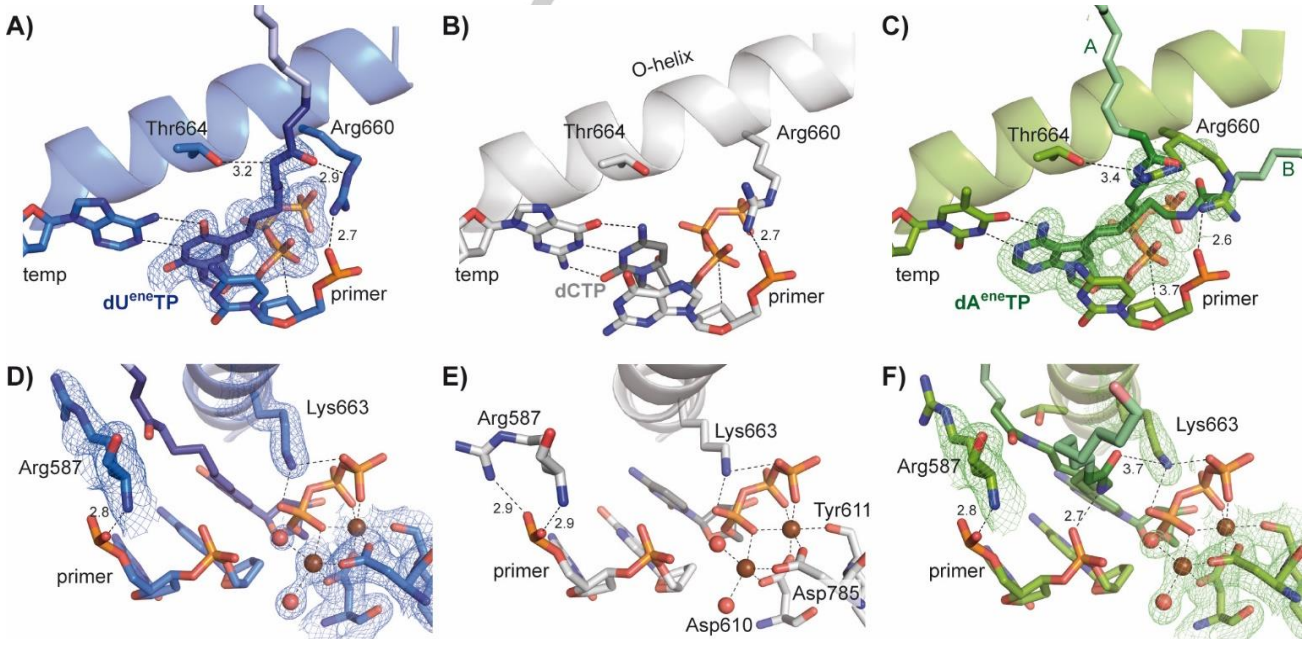


Figure 3. Structure of KlenTaq DNA polymerase in complex with dN^{R}TPs. The dashed lines highlight the interaction pattern described in the text. All distances are given in Å A-C) Detailed view on the insertion site of KlenTaq(**dU^{ene}TP**), blue; KlenTaq(dCTP), grey and KlenTaq(**dA^{ene}TP**), green. The substrate and near protein and primer/template residues are shown as sticks and the O-helix is shown as cartoon. The linker portion refined with zero occupancy is displayed in lighter color. **dU^{ene}TP** and **dA^{ene}TP** are surrounded by omit maps contoured at 2.5σ. The modification of **dA^{ene}TP** leaves the enzyme in two different conformations, which are labelled with A and B. D-F) View on the respective catalytic sites of the three structures. Magnesium ions and water molecules are shown as brown and red spheres, respectively. 2mFo-DFc map (1σ) is shown for interacting protein residues.

10.1002/chem.201604515

binding, active complex formation, catalytic reaction, and/or substrate release that result in the difference of processing.

Ternary complex with dA^{ene}TP

KlenTaq(**dA^{ene}TP**) was also obtained as a closed active complex and the overall structure is again almost identical to the unmodified case KlenTaq(dCTP) (Figure 3C, F; core rmsd: 0.40Å). The distance of the α-phosphate of the substrate to the nucleophilic 3' primer end is 3.7 Å (3.9 Å in the natural complex). In contrast to KlenTaq(**dU^{ene}TP**), it appears that the modification of **dA^{ene}TP** can take two different orientations to exit the enzyme. The relatively weak electron density of the linker portion and the near Arg660 side chain indicated flexibility and was initially difficult to interpret. The best solution was to model alternative conformations of the ligand with the linker pointing in two different directions as well as two conformations of the Arg660 side chain (Figure 3C, F and 4 B, D). For modeling we made use of polder and omit maps generated with the program Phenix[45] (Figure S3 C,D). The alternative conformations of Arg660 and the ligand including triphosphate, sugar, nucleobase and linker up to the amide functionality were refined with an occupancy of 0.5. The rest of the linker that was not visible in the electron density was modelled as extension to the more rigid part and refined with zero occupancy. In conformation one the modification extends to the surface of the enzyme through the cavity A already mentioned with **dU^{ene}TP** and **dU^{kyn}TP**. Here again, the enzyme substrate complex may be stabilized by Thr664 through interaction with the amide present on the linker (Figure 3C). In contrast to the natural case Arg587 is again oriented along the linker but maintains its backbone interaction with the primer (Figure 3F). In the second conformation the modification runs parallel to the O-helix and passes through another cavity (cavity B) mainly formed by the palm and finger domains (Figure 4D). This orientation has already been observed in the previously solved structure of KlenTaq(**dA^{kyn}TP**).[28] In this conformation Lys663 can interact with the amide bond (as it is the case in KlenTaq(**dA^{kyn}TP**)), however the distance of 3.7Å is too long for a H-bond. Due to the two conformations of Arg660 and the resulting weakly defined electron density it is hard to judge if the modelled conformations display the actual situation of that residue and therefore we can only speculate on interactions it may engage. However, it is well possible that Arg660 interacts with the primer backbone (as in KlenTaq(**dU^{ene}TP**), KlenTaq(**dU^{kyn}TP**) and KlenTaq(dCTP)) when the modification points to cavity A (Figure 3C). When the modification occupies cavity B, Arg660 is most probably too far displaced from the primer backbone. In this situation, interaction of the substrates amide moiety with the phosphate group of the primer terminus is possible which can compensate the "lost" interaction of Arg660 with the primer terminus. This interaction with the primer is also possible in KlenTaq(**dA^{kyn}TP**) (however not mentioned in the publication[28]).

In contrast to the alkene or alkyne modification attached to uridine we do not observe a difference in incorporation efficiency when attached to 7-deaza-adenosine. Both adenosine-modified

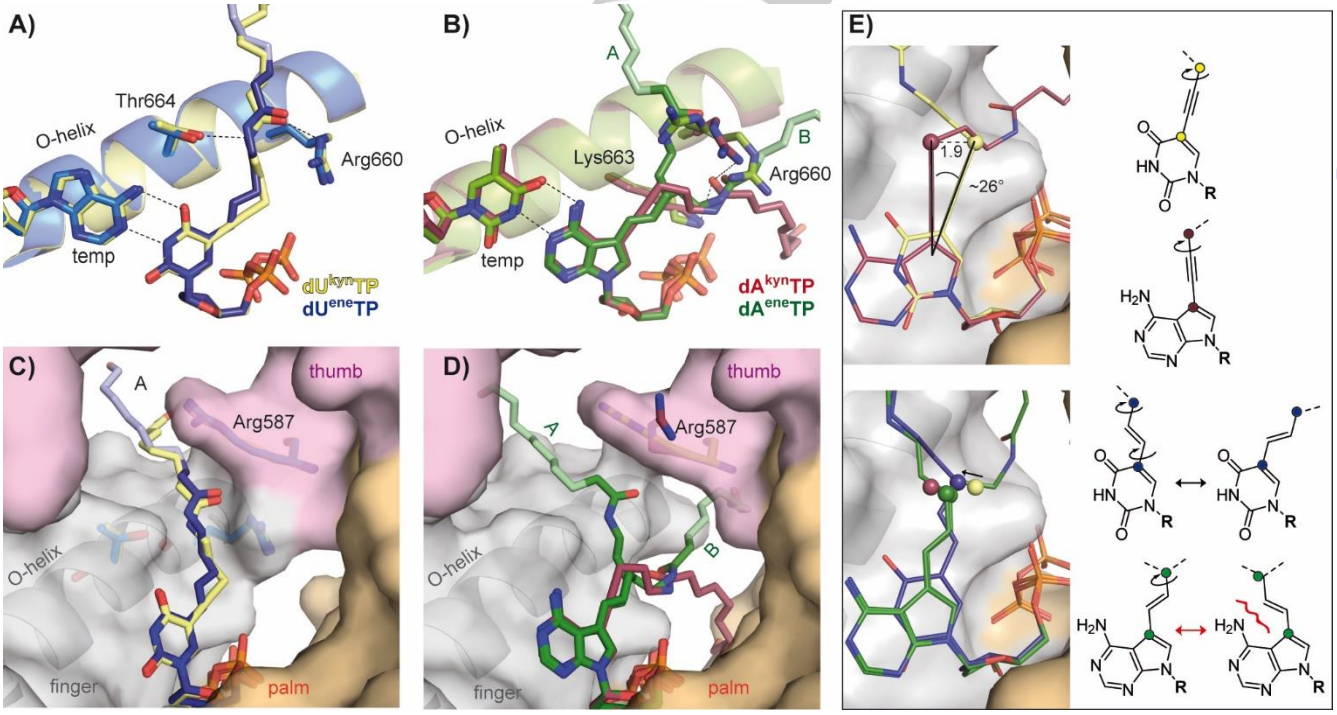


Figure 4. Comparison of alkyne- or alkene-modified nucleotides bound in the active site of KlenTaq DNA polymerase. A) Same arrangement as in Figure 3A but superimposition of KlenTaq(**dU^{ene}TP**), blue with KlenTaq(**dU^{kyn}TP**), yellow. Dashed lines show interactions for KlenTaq(**dU^{kyn}TP**). B) Same arrangement as in Figure 3C but superimposition of KlenTaq(**dA^{ene}TP**), green with KlenTaq(**dA^{kyn}TP**), red. C, D) Cavities A and B through which the modifications pass the enzyme. The surface of KlenTaq(**dA^{ene}TP**) is shown (Arg660 omitted) with the thumb, finger and palm domain coloured pink, grey and orange, respectively. E) Superimposition of **dT^{R}TPs** and **dA^{R}TPs** in the crystal structures and schematic view on the respective rotation points (displayed as coloured spheres) for alkyne or alkene modifications. The potential clash for **dA^{ene}TP** when rotating around the C7-linker bond is indicated in red.

nucleotides are to the same rate better incorporated than dAMP although they show differences in the structure. Whereas in KlenTaq(**dA^{kyn}TP**) the linker follows only one direction (direction B) it can adopt two different conformations in **dA^{ene}TP**. Both conformations show additional interactions with the enzyme or primer backbone and the substrate amide moiety, emphasizing the importance of this moiety within the linker portion as already discussed previously.[28] The different linkage type (alkene *vs* alkyne) does not make a difference with the linker investigated here but may influence incorporation efficiency when sterically more demanding groups are attached. The knowledge that both modification types are equally well accepted when linked to 7-deaza-adenine enlarges the repertoire from which to choose in future when designing modified nucleotides for a specific application.

Comparison of dT^{R}TPs and dA^{R}TPs

As just discussed, the linkage type does not change incorporation efficiency whether the 5-N-(10-hydroxydecanoyl)-aminopentenyl- or the 5-N-(10-hydroxydecanoyl)-aminopentynyl-modification is attached to 7-deaza-dATP. When attached to dTTP, however, different linkage type results in different incorporation behavior for **dU^{ene}TP** and **dU^{kyn}TP** by KlenTaq DNA polymerase. The same tendency has already been observed before by Williams and coworkers.[25, 35] The crystal structures solved in this and previous work are not able to explain the difference between **dT^{R}TPs** and **dA^{R}TPs** as both modifications do not disturb the enzyme to form an active closed complex and in both cases potentially stabilizing interactions of the linker with the enzyme or primer are possible. However, the presented crystal structures only represent a single snapshot of the whole insertion reaction, namely the step after binding of the substrate and closure of the finger domain. Therefore, we assume that differences arise in a different step in the insertion process, e.g. during substrate binding, enzyme closure, the actual chemical reaction or substrate release. What can nicely be seen in the structures is that the angle by which the rigid part of the alkyne linker protrudes from the nucleobases when base-pairing in the active site differs by approximately 26° (Figure 4E). This results in different positions of the points from where the C-C bonds can rotate and the modification gets flexible (distance 1.9Å). With alkene linkers this difference is reduced due to the additional kink resulting from the double bond. For **dU^{ene}TP** compared to **dU^{kyn}TP** the point of rotation shifts towards the templating nucleotide and away from the triphosphate moiety (arrow in Figure 4E). This different position might be more favorable somewhere in the reaction cascade. Another point that can be noted is that alkene linkers attached to dTTP have additional flexibility compared to the alkyne linkers, as rotation around the bond at C5 leads to different positioning of the double bond and the rest of the linker. For 7-deaza-dATP this rotation seems less favorable as a clash between the free amine of 7-deaza-adenine and one hydrogen of the alkene moiety may occur. In some way these differences may lead to the fact that both **dA^{R}TP**s are better substrates than dATP but only **dU^{ene}TP** and not **dU^{kyn}TP** is preferably incorporated than dTTP.

Conclusions

In this work two modified nucleotides were successfully synthesized and evaluated towards their acceptance by KlenTaq DNA polymerase. We show by a competitive incorporation assay that the two alkene derivatives **dA^{ene}TP** and **dU^{ene}TP** are excellent substrates for KlenTaq DNA polymerase displaying even better substrate properties than the respective natural analogue. In addition the proficient incorporation of **dA^{kyn}MP** over dAMP from a previous study was confirmed. The presented ternary structures KlenTaq(**dA^{ene}TP**) and KlenTaq(**dU^{ene}TP**) support this observation, as in analogy to the previously solved structures KlenTaq(**dU^{kyn}TP**) and KlenTaq(**dA^{kyn}TP**) the modifications do not disturb formation of a catalytic complex. We observe a full closure of the enzymes finger domain and proper interaction of the substrate triphosphate within the active site. The modifications point into directions where they can leave the enzyme through two different cavities. Interactions of the protein or the primer backbone with the amide moiety of the linker may stabilize the ternary complex in both possible directions and support the proficient incorporation behaviour of **dA^{kyn}TP**, **dA^{ene}TP** and **dU^{ene}TP**. An obvious reason why **dU^{kyn}TP** is less well processed than **dU^{ene}TP**, however, could not be deduced from the obtained structure. The obtained results represent new insights for the design of modified nucleotides as substrates for KlenTaq DNA polymerase. When designing modified nucleotides based on 7-deaza-dATP both linkers (alkene and alkyne) should be considered and it might well be that for larger modifications the linkage type makes a difference here. If the modification should be attached to dTTP the alkene linkage may be preferred.

Experimental Section

Chemical Synthesis of modified nucleotides

General experimental details: Succinimidyl 10-hydroxydecanoate,[46] 2,2,2-trifluoro-N-(pent-4-enyl)acetamide,[47] 5-(N-(10-hydroxydecanoyl)-aminopentynyl)-2´-deoxyuridine-triphosphate,[28] and 7-(N-(10-hydroxydecanoyl)-aminopentynyl)-7-deaza-2´-deoxyadenosine-triphosphate[28] were prepared according to literature. 1H, 13C and 31P NMR spectra were recorded on a Bruker DRX 300 MHz or Avance II+ 600 MHz instrument. The chemical shifts are referenced to the residual proton signal of the deuterated solvents: $CDCl_3$ (7.26 ppm), MeOD (3.31) or d6-DMSO (2.49 ppm) for 1H NMR spectra; $CDCl_3$ (77.0 ppm), MeOD (49.9) or d6-DMSO (39.5 ppm) for 13C NMR spectra. 31P-shifts are relative to external 85% phosphoric acid. 1H- and 13C-assignments were based on COSY and HSQC experiments. MS experiments were performed on a Finnigan LCQ Advantage MAX ion trap instrument. Analytical thin-layer chromatography (TLC) was carried out on Marcherey-Nagel Polygram SIL G/UV254 plates. Flash column chromatography was carried out on silica gel 60 (70-230 mesh). All reactions were carried out under argon atmosphere. Chemical reagents and solvents were purchased from commercial suppliers and used without further purification. Organic solvents for reactions were dried overnight over freshly activated molecular sieves (4 Å).

Synthesis of 2´-deoxyuridine analogue dU^{ene}TP

5-(trifluoroacetamidopentenyl)-2´-deoxyuridine 2

To a solution of 5-iodo-2´-deoxyuridine **1** (400 mg, 1.1 mmol) in DMF under nitrogen atmosphere, were added 2,2,2-trifluoro-N-(pent-4-enyl)acetamide (1.2 g, 6.7 mmol, 6 equiv), $Pd(OAc)_2$ (24 mg, 0,11 mmol, 0.1 equiv), and PPh_3 (58 mg, 0.22 mmol, 0.2 equiv.) followed by addition of triethylamine (308 µl, 2.2 mmol, 2 equiv). The reaction mixture was stirred at 100 °C for 1 h and then evaporated in vacuo. Complete conversion of the starting material was detected by RP-TLC (MeCN/water: 1/3). The crude mixture was first purified by column chromatography on silica gel (DCM/MeOH: 0 to 10 % MeOH). Then the resulted fractions were further purified by RP-HPLC (Nucleosil 100-5 C18 PPN, 5-100% acetonitrile / 0.05 M aqueous triethylammonium acetate buffer (TEAA buffer, pH 7.0) to give modified nucleoside **2** as a white solid (64 mg, 14%).
RP-TLC (MeCN/water: 1/3): Rf = 0.29;
1H NMR (400 MHz, CD3OD): δ 8.14 (s, 1H, H-6), 6.46 (dt, J = 15.7, 7.1 Hz, 1H, CH=CH-CH2), 6.33 (t, J = 6.6 Hz, 1H, H-1´), 6.20 (d, J = 15.7 Hz, 1H, CH=CH-CH2), 4.46 (dt, J = 7.1, 3.3 Hz, 1H, H-3´), 3.97 (q, J = 3.3 Hz, 1H, H-4´), 3.88 (dd, J = 12.0, 3.3 Hz, 1H, Ha-5´), 3.80 (dd, J = 12.0, 3.3 Hz, 1H, Hb-5´), 3.27-3.39 (m, 2H, CH2-NH), 2.24-2.43 (m, 2H, H-2´), 2.22 (q, J = 7.4, 7.1 Hz, 2H, CH=CH-CH2), 1.74 (p, J = 7.4 Hz, 2H, CH2-CH2-CH2) ; 13C NMR (101 MHz, CD3OD) δ 164.6 (CO), 151.6 (CO), 137.8 (C-6), 131.6 (CH=CH-CH2), 122.9 (CH=CH-CH2), 113.4 (C-5), 80.0 (C-4´), 86.6 (C-1´), 72.0 (C-3´), 62.6 (C-5´), 41.6 (C-2´), 40.3 (CH2-NH), 31.6 (CH=CH-CH2), 29.4 (CH2-CH2-CH2).
HRMS (m/z, positive mode) calcd for [C16H20F3N3O6Na]+: 430,1202; found: 430,1169.

5-(aminopentenyl)-2´-deoxyuridinetriphosphate 3

Nucleoside **2** (40 mg, 100 µmol) and proton sponge (1,8-bis(dimethylaminonaphthalene)) (32 mg, 149 µmol) were dried overnight in vacuo and dissolved in trimethyl phosphate (1 mL) and cooled down to 0 °C. $POCl_3$ (20 µL, 218 µmol) was added to the mixture and stirred for 1 h at RT. A 0.5 M solution of $(Bu_3NH)_2H_2P_2O_7$ in anhydrous DMF (1.2 mL, 493 µmol) and nBu_3N (261 µL, 987 µmol) were added simultaneously to the mixture. After 30 min, 0.1 M aqueous triethylammonium bicarbonate (TEAB buffer, 1 M triethylamine, saturated with CO_2, pH 7.5, 3 mL) was added and the aqueous layer was washed with EtOAc (3 × 4 mL). The aqueous layer was evaporated under vacuum. The residue was dissolved in 40 ml of a solution of ammonium hydroxide (25%) and the reaction stirred slowly at RT for 4 h. The mixture was concentrated in vacuo and the resulting residue was purified by ion-exchange chromatography (DEAE-Sephadex A-25, flow 2 mL/min, linear gradient of 0.1 M to 1 M TEAB buffer) and further purified by RP-HPLC (VP 250/16 NUCLEODUR C18 HTec column, linear gradient from 5 to 100 % MeCN in 0.05 M TEAA buffer (0.05 M acetic acid, 0.05 M triethylamine, pH 7.0)) to give the triphosphate 3 (5,5 mg, 10 µmol, 10 %).
1H NMR (400 MHz, D2O): δ 7.95 (s, 1H, H-6), 6.43 (dt, J = 16.1, 6.5 Hz, 1H, CH=CH-CH2), 6.36 (t, J = 6.3 Hz, 1H, H-1´), 6.27 (d, J = 16.1 Hz, 1H, CH=CH-CH2), 4.70 (q, J = 4.8 Hz, 1H, H-3´), 4.33-4.24 (m, 1H, H-4´), 4.26-4.16 (m, 2H, H-5´), 3.04 (t, J = 6.5 Hz, 1H, CH2-NH2), 2.41 (t, J = 6.3 Hz, 2H, H-2´), 2.31 (q, J = 6.5 Hz, 2H, CH=CH-CH2), 1.88 (p, J = 6.5 Hz, 2H, CH2-CH2-CH2-NH2) ; 13C NMR (101 MHz, D2O) δ 164.6, 151.0, 136.8, 131.4, 121.2, 113.0, 85.7, 85.2, 70.5, 65.2, 39.1, 38.6, 28.9, 25.7 ; 31P NMR (162 Hz, D2O) δ -7.11 (d, J = 20.2 Hz, 1P, Pγ), -11.24 (d, J = 19.1 Hz, 1P, Pα), 21.91 (t, J = 20.2 Hz, 1P, Pβ) ;
HRMS (negative mode): m/z: calcd for [C14H23N3O14P3]-: 550.0415; found: 550.0403.

5-(N-(10-hydroxydecanoyl)-aminopentenyl)-2´-deoxyuridine-triphosphate $dU^{ene}TP$

To a solution of 5-(aminopentenyl)-2´-deoxyuridinetriphosphate **3** (10 µmol) in 400 µl of sodium carbonate buffer (0.1 M, pH = 8.7) was added a solution of succinimidyl 10-hydroxydecanoate (20 mg, 70 µmol, 7 equiv) in DMF (400 µl). The reaction was stirred slowly at RT for 2 h. The solution was filtered before purification by RP-HPLC (VP 250/8 NUCLEODUR C18 HTec column, 5-100 % acetonitrile/ 0.1 M TEAA buffer (pH 7.0) to give triphosphate **$dU^{ene}TP$** in quantitative yield.
1H NMR (400 MHz, D2O) δ 7.90 (s, 1H, H-6), 6.49 (dt, J = 16.0, 6.9 Hz, 1H, CH=CH-CH2), 6.38 (t, J = 7.0 Hz, 1H, H-1´), 6.23 (d, J = 16.0 Hz, 1H, CH=CH-CH2), 4.74-4.66 (m, 1H, H-3´), 4.36-4.12 (m, 3H, H-4´ and H-5´), 3.60 (t, J = 6.7 Hz, 2H, CH2-OH), 3.32-3.18 (m, 19H, CH2-NH-, superimposed by TEA), 2.51-2.37 (m, 2H, H-2´), 2.33-2.24 (m, 2H, CH=CH-CH2), 2.23 (t, J = 7.4 Hz, 2H, NH-CO-CH2), 1.72 (p, J = 6.9 Hz, 2H, CH2-CH2-CH2-NH), 1.61-1.48 (m, 4H, 2*CH2), 1.38-1.18 (m, 36H, 5*CH2, superimposed by TEA) ; 31P NMR (162 MHz, D2O) δ -10.75 (d, J = 14.2 Hz, 1P, Pγ), -11.59 (d, J = 19.6 Hz, 1P, Pα), 22.62 – -23.74 (m, 1P, Pβ) ;
HRMS (negative mode): m/z: calcd for [C24H41N3O16P3]-: 720.1722; found: 720.1736.

Synthesis of 2´-deoxyadenosine analogue dAeneTP

7-deaza-2´-deoxy-7-vinyladenosine 5

To 7-deaza-2´-deoxy-7-iodoadenosine **4** (500 mg, 1,329 mmol) in DMF (10 ml) were added subsequently tetrakis(triphenylphosphine)palladium (150 mg, 0.133 mmol, 0.1 equiv) and tributyl(vinyl)tin (1,555 ml, 5.317 mmol, 4 equiv). The mixture was stirred at 100°C for 2 h. After cooling, the solvent was removed under vacuum. The crude mixture was purified by silica gel column chromatography (DCM/MeOH: 9/1) to yield compound **5** (314 mg, 86%) as yellow solid.
1H NMR (400 MHz, CD3OD) δ 8.07 (s, 1H, H-2), 7.51 (s, 1H, H-8), 7.00 (dd, J = 17.3, 11.0 Hz, 1H, -CH=CH2), 6.51 (dd, J = 8.3, 6.0 Hz, 1H, H-1´), 5.59 (dd, J = 17.3, 1.6 Hz, 1H, -CH=CH2), 5.26 (dd, J = 11.0, 1.6 Hz, 1H, -CH=CH2), 4.53 (dt, J = 6.0, 2.6 Hz, 1H, H-3´), 4.01 (q, J = 3.3 Hz, 1H, H-4´), 3.86-3.69 (m, 2H, H-5´), 2.67 (ddd, J = 13.7, 8.3, 6.0 Hz, 1H, H-2´), 2.31 (ddd, J = 13.7, 6.0, 2.6 Hz, 1H, H-2´) ; 13C NMR (101 MHz, CD3OD) : δ 159.3 (Cq), 152.1 (C2), 151.1 (Cq), 129.6 (-CH=CH2), 121.3 (C8), 116.6, 115.5 (-CH=CH2), 103.0 (Cq), 100.2 (Cq), 89.0 (C4´), 86.5 (C1´), 73.0 (C3´), 63.7 (C5´), 41.4 (C2´) ;
HRMS (positive mode): m/z: calcd for [C13H17N4O3]+: 277.1301; found: 277.1303.

7-deaza-2´-deoxy-3´,5´-dibenzoate-6N,6N-dibenzoyl-7-vinyladenosine 6

To 7-deaza-2´-deoxy-7-vinyladenosine **5** (600 mg, 2.173 mmol) in dry pyridine (10 ml) at 0°C was added benzoylchlorid (1.514 ml, 13.038 mmol) over 30 min. Then the mixture was stirred at room temperature for 2 hours. Water was slowly added and the solution was extracted with dichloromethane. The combined organic layers were dried ($MgSO_4$), filtered, evaporated and the residue was purified by silica gel column chromatography (Petroleum Ether/EtOAc: 8/2 to 4/6) to yield protected adenosine **6** (1.253 g, 83%) as a yellow solid. Mp = 62°C.

1H NMR (400 MHz, CDCl3): δ 8.59 (s, 1H, H-2), 8.18-8.04 (m, 4H, Bz), 7.90-7.81 (m, 4H, Bz), 7.74-7.29 (m, 13H, Bz and H-8), 6.93 (dd, J = 8.7, 6.0 Hz, 1H, H-1´), 6.70 (dd, J = 17.4, 10.8 Hz, 1H, , -CH=CH2), 5.83 (dt, J = 6.0, 2.2 Hz, 1H, H-3´), 5.29 (dd, J = 17.4, 1.3 Hz, 1H, -CH=CH2), 5.00 (dd, J = 10.8, 1.3 Hz, 1H, -CH=CH2), 4.84 (dd, J = 12.0, 3.6 Hz, 1H, H-5´), 4.73 (dd, J = 12.0, 3.6 Hz, 1H, H-5´), 4.67-4.62 (m, 1H, H-4´), 2.98-2.67 (m, 2H, H-2´). 13C NMR (101 MHz, CDCl3) δ 172.8, 166.3, 166.2 (3xCO), 153.3, 153.2 (C4 and C6), 151.4 (C2), 133.8, 133.65, 133.61, 133.0, 130.2, 130.0, 129.8, 129.6, 129.4, 129.4, 129.3, 129.0, 128.9, 128.7, 128.6, 126.9 (-CH=CH2), 122.4 (C8), 116.4 (-CH=CH2), 115.4 (C7), 113.1 (C5), 84.0 (C1´), 82.6 (C4´), 75.5 (C3´), 64.5 (C5´), 38.5 (C2´).

7-deaza-2´-deoxy-3´,5´-dibenzoate-6N,6N-dibenzoyl-7-(trifluoroacetamidopentenyl)adenosine 7E/Z

A solution of adenosine 6 (200 mg, 0.289 mmol) in dry toluene (5 ml) was degassed 5 min. The pentenyltrifluoroacetamide (520 µl, 2.89 mmol, 10 equiv) was added and the solution was degassed 5 more minutes. The solution was heated up to 100°C and Grubbs II catalyst (48 mg, 20%) was added by portions (16 mg every 30 min). After 2 h, the solution was cool down and the solvent was evaporated. The residue was purified by silica gel chromatography (Petroleum Ether/EtOAc: 7/3) affording **7E/Z** (108 mg, 45%) as a mixture of diastereoisomers (E/Z : 3/1).

7-deaza-2´-deoxy-7-(trifluoroacetamidopentenyl)adenosine 8

The mixture of diastereoisomers **7** (217 mg, 0.257 mmol) in a 7 N ammonia solution in methanol (5 ml) was stirred in a closed vial at 50 °C overnight. The mixture was concentrated in vacuo and the resulting residue was first purified by column chromatography on silica gel (DCM/MeOH: 9/1) giving compound **8E/Z** as E/Z mixture (80 mg, 72%) and entire deprotected compound **9E/Z** as E/Z mixture (20 mg, 23%).

To compound **9** (20 mg, 0.060 mmol) in THF (0.5 ml) was added Ethyltrifluoroacetate (8 µl, 0.066 mmol, 1.1 equiv) and DMAP (0.7 mg, 0.006 mmol, 0.1 equiv). The solution was stirred at 90 °C for 4 h and the the resulting residue was first purified by column chromatography on silica gel (DCM/MeOH: 9/1) affording **8** (12 mg, 47%) as a mixture of diastereoisomers (E/Z : 3/1).

Then the combined fractions of **8E/Z** (92 mg) were purified by RP-HPLC (VP 250/16 NUCLEODUR C18 HTec column, 5-100 % acetonitrile/ water) to give pure trans-alkene nucleoside **8** as a white solid (28 mg, 27%).

1H NMR (400 MHz, CD3OD): δ 8.08 (s, 1H, H-2), 7.42 (s, 1H, H-8), 6.69 (dd, J = 15.5, 1.1 Hz, 1H, CH=CH-CH2), 6.53 (dd, J = 8.3, 5.9 Hz, 1H, H-1´), 6.06 (dt, J = 15.6, 7.0 Hz, 1H, , -CH=CH-CH2), 4.55 (dt, J = 5.6, 2.6 Hz, 1H, H-3´), 4.03 (q, J = 3.3 Hz, 1H, H-4´), 3.88-3.66 (m, 2H, H-5´), 3.40 (t, J = 7.0 Hz, 2H, -CH2-NH), 2.69 (ddd, J = 13.4, 8.3, 6.0 Hz, 1H, H-2´), 2.42-2.21 (m, 3H, H-2´ and -CH=CH-CH2-), 1.79 (p, J = 7.1 Hz, 2H, -CH2-CH2-CH2-) ; 13C NMR (101 MHz, CD3OD) δ 159.2 (Cq), 151.9 (C2), 150.8 (Cq), 132.2 (-CH=CH-CH2), 123.2 (-CH=CH-CH2), 120.6 (C8), 116.2 (Cq), 103.1 (Cq), 89.0 (C-4´), 86.4 (C-1´), 73.0 (C-3´), 63.7 (C-5´), 41.3 (C2´), 39.9 (-CH2-NH), 30.9 (-CH=CH-CH2-), 29.5 (-CH2-CH2-CH2-) ;

HRMS (positive mode): m/z: calcd for [C18H23F3N5O4]+: 230.1702; found: 230.1693.

7-(aminopentenyl)- 7-deaza-2´-deoxyadenosinetriphosphate 10

Nucleoside **8** (65 mg, 152 µmol) and proton sponge (1,8-bis(dimethylaminonaphthalene)) (49 mg, 229 µmol) were dried overnight in vacuum, dissolved in trimethyl phosphate (3.25 mL), and cooled down to 0 °C. $POCl_3$ (28 µL, 218 µmol) was added to the mixture and stirred for 1 h at RT. A 0.5 M solution of $(Bu_3NH)_2H_2P_2O_7$ in anhydrous DMF (1.5 mL, 760 µmol) and nBu_3N (361 µL, 1,52 mmol) were added simultaneously to the mixture. After 30 min, 0.1 M aqueous triethylammonium bicarbonate (TEAB buffer, pH 7.5, 6 mL) was added and the aqueous layer was washed with EtOAc (3 × 6 mL). The aqueous layer was evaporated under vacuum. The residue was dissolved in 45 ml of a solution of ammonium hydroxide (25%) and the reaction stirred slowly at RT for 4 h. The mixture was concentrated in vacuo and the resulting residue was purified by ion-exchange chromatography (DEAE-Sephadex A-25, linear gradient of TEAB buffer (0.1 M to 1 M, 1000 mL), flow 2 mL/min) and further purified by RP-HPLC (VP 250/16 NUCLEODUR C18 HTec column, linear gradient from 5 to 100% MeCN in 0.05 M TEAA buffer (pH 7.0)) to give the triphosphate 10 (19.8 µmol, 14 %).

1H NMR (400 MHz, CD3OD) δ 8.05 (s, 1H, H-6), 7.85 (s, 1H, H-2), 6.71-6.59 (m, 2H, H-1´ and CH=CH-CH2), 6.28 (dt, J = 14.7, 6.9 Hz, 1H, CH=CH-CH2), 4.78-4.70 (m, 1H, H-3´), 4.31-4.22 (m, 2H, H-5´), 4.07-4.01 (m, 1H, H-4´), 3.26-3.01 (m, 15H, CH2-NH2 superimposed by TEA), 2.58-2.48 (m, 1H, H-2´), 2.44-2.24 (m, 3H, H-2´ and CH=CH-CH2), 1.93-1.83 (m, 2H, CH2-CH2-CH2-NH2); 31P NMR (162 MHz, CD3OD) : δ -10.97 (d, J = 20.5 Hz, Pγ), -11.93 (d, J = 19.9 Hz, Pα), -23.72 (t, J = 19.9 Hz, Pβ) ; HRMS (negative mode): m/z: calcd for [C16H25N5O12P3]-: 572.0735; found: 572.0759.

7-(N-(10-hydroxydecanoyl)-aminopentenyl)- 7-deaza-2´-deoxyadenosinetriphosphate dAeneTP

To a solution of 7-(aminopentenyl)- 7-deaza-2´-deoxyadenosinetriphosphate **10** (6.6 µmol) in 300 µl of sodium carbonate buffer (0.1 M, pH = 8.7) was added a solution of succinimidyl 10-hydroxydecanoate (9.4 mg, 33 µmol, 5 equiv) in DMF (300 µl). The reaction was stirred slowly at RT for 2 h. The solution was filtered before purification by RP-HPLC (VP 250/8 NUCLEODUR C18 HTec column, 5-100 % acetonitrile/ 0.05 M TEAA buffer (pH 7.0)) to give triphosphate **dA^{ene}TP** in quantitative yield.

1H NMR (400 MHz, CD3OD) δ 8.23 (s, 1H, H-2), 7.68 (s, 1H, H-8), 6.63 (t, J = 6.7 Hz, 1H, H-1´), 6.27 (d, J = 15.5 Hz, 1H, CH=CH-CH2), 6.16 (dt, J = 15.5, 6.6 Hz, 1H, CH=CH-CH2), 4.75-4.62 (m, 1H, H-3´), 4.36-4.26 (m, 2H, H-5´), 4.16-4-13 (m, 1H, H-4´), 3.55 (t, J = 6.7 Hz, CH2-OH) 3.57-3.08 (m, 14H, CH2-NH- superimposed by TEA), 2.56-2.46 (m, 1H, H-2´), 2.39-2.28 (m, 1H, H-2´), 2.27-2.19 (m, 2H, CH=CH-CH2), 1.80-1.67 (m, 4H, CH2-CH2-CH2-NH- and CO-CH2), 1.67-1.58 (m, 2H, -CO-CH2-CH2-), 1.58-1.48 (m, 2H, -CH2-CH2-OH) 1.41-1.21 (m, 28H, 5xCH2 superimposed by TEA). 31P NMR (162 MHz, CD3OD) δ -9.95 (d, J = 19.6 Hz, Pγ), -10.75 (d, J = 20.8 Hz, Pα), -22.95 – -23.69 (m, Pβ) ; HRMS (negative mode): m/z: calcd for [C26H43N5O14P3]-: 742.2019, found: 742.2010.

Primer extension reactions

KlenTaq DNA polymerase was expressed and purified as described before.[48] T4 polynucleotide kinase PNK was purchased from New England Biolabs. Primer and template were purchased from Biomers. [γ-32P]ATP was purchased from Hartmann Analytics and natural dNTPs from ThermoFisher. Sequences of Oligonucleotides are found in the supporting information.

5´ Radioactive labeling of ODNs

DNA oligonucleotide primers were radioactively labeled at the 5´ terminus using T4 PNK and [γ-32P]ATP. The reaction mixture contained primer (0.4 µM), PNK reaction buffer (1 x), [γ-32P]ATP (0.4 µCi/µl) and T4 PNK (0.4 U/µl) in a total volume of 50 µl and were incubated for 1 h at 37 °C. The reaction was stopped by incubation for 2 min at 95 °C and purified by gel filtration (MicroSpin Sephadex G-25).

Full-length incorporation.

Primer extension reaction (10 µl) employing KlenTaq DNA polymerase contained 1 x KlenTaq reaction buffer (50 mM Tris HCl (pH 9.2), 16 mM $(NH4)_2SO_4$, 2.5 mM $MgCl_2$, 0.1% Tween 20), 150 nM 32P-labeled primer, 200 nM template, 50 µM of modified dN^RTP and 50 µM each of dATP/dTTP, dCTP, dGTP, and 8 nM KlenTaq DNA polymerase. First primer and template were annealed. Afterwards the primer/template complex, nucleotides and KlenTaq DNA polymerase were incubated (55°C; 30 min). The reactions were quenched by addition of 50 µL PAGE gel loading buffer/stop solution (80% formamide, 20 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanole FF) and the product mixtures were analyzed by 12% denaturing polyacrylamide gel and subjected to autoradiography.

Competition Experiments

A typical competition experiment (10 µl) employing KlenTaq DNA polymerase contained 1 x KlenTaq reaction buffer, 150 nM 32P-labeled primer, 200 nM template, 50 µM modified dNTP/dN^RTP mixture, and 8 nM KlenTaq DNA polymerase. First primer and template were annealed. Afterwards the primer/template complex, nucleotides and DNA polymerase were incubated (55°C; 30 min). The reactions were quenched by addition of 50 µL PAGE gel loading buffer/stop solution and the product mixtures were analyzed by 12% denaturing polyacrylamide gel and subjected to autoradiography. Quantification was done by using the Bio-Rad Quantity One software. The conversion in % was plotted versus the concentration using the program ImageLab. All reactions were done in triplicates.

Gel electrophoresis

Denaturing polyacrylamide gels (12 %) were prepared by polymerization of a solution of urea (8.3 M) and bisacrylamide/acrylamide (12 %) in TBE buffer using ammonium peroxodisulfate (APS, 0.08 %) and N,N,N´,N´-tetramethylethylene-diamine (TEMED, 0.04 %). Immediately after addition of APS and TEMED the solution was filled in a sequencing gel chamber (Bio-Rad) and left for polymerization for at least 45 min. After addition of TBE buffer (1 x) to the electrophoresis unit, gels were prewarmed by electrophoresis at 100 W for 30 min and samples were added and separated during electrophoresis (100 W) for approx. 1.5 h. The gel was transferred to Whatman filter paper, dried at 80 °C, in vacuo, using a gel dryer (model 583, Bio-Rad) and exposed to an imager screen. Readout was performed with a molecular imager (FX, Bio-Rad).

Crystallization, data collection and analysis

Primer and template were purchased from MWG Eurofins. The protein was overexpressed and purified as described earlier[48] and concentrated to 10 mg/ml for storage at 4°C. For crystallization, a 2´3´-dideoxy-cytidine terminated primer (5´-d(GAC CAC GGC GC)ddC) and a 16-mer template (5´-d(AAA A/TGG CGC CGT GGT C)) were used.

Primer/template construct was annealed and the ternary complex of KlenTaq was formed by mixing the protein (final concentration ~6.0 mg/ml in 20 mM Tris-HCl pH 7.5, 1mM EDTA, 0.15 M NaCl, 1 mM β-Mercaptoethanol), the primer/template DNA (final concentration ~0.120 mM) and triphosphate **$dN^{ene}TP$** in a molar ratio of 1:1.2:15 and the solution was set to a final concentration of 20 mM $MgCl_2$. The mixture was incubated for 30 min at 30 °C. Crystals were grown in hanging drop plates (Qiagen) against 800 µl reservoir (reservoir condition: 0.05 M sodium cacodylate pH 6.5, 0.01 M $Mg(OAc)_2$, 0.2 M NH_4OAc, 30% PEG 4000) in a 1:1 ratio protein/DNA to reservoir. Crystals were harvested after 6 days of growth and dipped into a cryo-solution consisting of the crystallization reservoir with additional 20% glycerol before freezing in liquid nitrogen.

Data were collected at the beamlines PXI and PXIII at the Swiss Light Source (SLS), Paul-Scherrer Institute, Villigen, Switzerland. Data reduction was performed with the XDS package.[49] The structure was solved by difference Fourier techniques using KlenTaq wild-type (PDB 3M8S)[50] as model. Refinement was performed with PHENIX[45] and model rebuilding was done with COOT.[51] In both structures the complete enzyme (residues 293-832) and the primer/template duplex was modelled. Side chains without defined electron density were not deleted but modelled in a common rotamer conformation and high B-factors demonstrate their flexibility. The substrate nucleotide **$dA^{ene}TP$** was modelled in two conformations and refined with an occupancy of 0.5 for each conformation. The terminal alkyl chain of the linker in **$dU^{ene}TP$** as well as **$dA^{ene}TP$** was not resolved in the electron density and was modelled in extension to the more rigid part (nucleotide plus linker until the amide functionality) and refined with zero occupancy. The restraints files (cif files) for the ligands were generated using the grade webserver.[52] Figures were made with PyMOL.[53] Omit maps for the triphosphates are shown in Figure S2 and data collection and refinement statistics in Table S1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (www.rcsb.org) with PDB codes 5E41 for KlenTaq(**$dU^{ene}TP$**) and 5SZT for KlenTaq(**$dA^{ene}TP$**).

Acknowledgements

We gratefully acknowledge financial support by the DFG and the support and access to beamlines PXI and PXIII at the Swiss Light Source (SLS) of the Paul Scherrer Institute (PSI), Villigen, Switzerland.

Keywords: linker modified nucleotides • DNA polymerase crystal structure • KlenTaq DNA polymerase • artificial nucleotide substrate

[1] D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, et al., *Nature* **2008**, *456*(7218), 53–59.

[2] E. R. Mardis, *Annu. Rev. Genomics Hum. Genet.* **2008**, *9*, 387–402.

[3] G. Mayer, *Angew. Chem. Int. Ed.* **2009**, *48*(15), 2672–2689.

[4] S. H. Weisbrod, A. Marx, *Chem. Commun.* **2008**, (44), 5675–5685.

[5] C. M. Perou, T. Sorlie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, D. T. Ross, H. Johnsen, L. A. Akslen, et al., *Nature* **2000**, *406*(6797), 747–752.

[6] A. Baccaro, A.-L. Steck, A. Marx, *Angew. Chem. Int. Ed.* **2012**, *51*(1), 254–257.

[7] M. Hocek, M. Fojta, *Org. Biomol. Chem.* **2008**, *6*(13), 2233–2241.
[8] M. Hollenstein, *Molecules* **2012**, *17*(12), 13569–13591.
[9] O. Thum, S. Jäger, M. Famulok, *Angew. Chem. Int. Ed.* **2001**, *40*(21), 3990–3993.
[10] T. Ohbayashi, M. Kuwahara, M. Hasegawa, T. Kasamatsu, T. Tamura, H. Sawai, *Org. Biomol. Chem.* **2005**, *3*, 2463–2468.
[11] S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum, M. Famulok, *J. Am. Chem. Soc.* **2005**, *127*(43), 15071–15082.
[12] M. Kuwahara, J.-i. Nagashima, M. Hasegawa, T. Tamura, R. Kitagata, K. Hanawa, S.-i. Hososhima, T. Kasamatsu, H. Ozaki, H. Sawai, *Nucleic Acids Res.* **2006**, *34*(19), 5383–5394.
[13] M. Hocek, *J. Org. Chem.* **2014**, *79*(21), 9914–9921.
[14] J. Sági, A. Szemzö, K. Ébinger, A. Szabolcs, G. Sági, É. Ruff, L. Ötvös, *Tetrahedron Letters* **1993**, *34*(13), 2191–2194.
[15] B. M. Znosko, T. W. Barnes, T. R. Krugh, D. H. Turner, *J. Am. Chem. Soc.* **2003**, *125*(20), 6090–6097.
[16] F. Seela, E. Feiling, J. Gross, F. Hillenkamp, N. Ramzaeva, H. Rosemeyer, M. Zulauf, *Journal of Biotechnology* **2001**, *86*(3), 269 – 279.
[17] L. H. Thoresen, G.-S. Jiao, W. C. Haaland, M. L. Metzker, K. Burgess, *Chem. Eur. J.* **2003**, *9*(19), 4603–4610.
[18] D. Verga, M. Welter, A.-L. Steck, A. Marx, *Chem. Commun.* **2015**, *51*(34), 7379–7381.
[19] D. Cherkasov, T. Biet, E. Bäuml, W. Traut, M. Lohoff, *Bioconjugate Chemistry* **2010**, *21*(1), 122–129.
[20] Z. Zhu, J. Chao, H. Yu, A. Waggoner, *Nucleic Acids Res.* **1994**, *22*(16), 3418–3422.
[21] C. J. Lacenere, N. K. Garg, B. M. Stoltz, S. R. Quake, *Nucleosides Nucleotides Nucleic Acids* **2006**, *25*(1), 9–15.
[22] M. Welter, D. Verga, A. Marx, *Angew. Chem. Int. Ed.* **2016**, *55*(34), 10131–10135.
[23] T. Tasara, *Nucleic Acids Research* **2003**, *31*(10), 2636–2646.
[24] G. Giller, T. Tasara, B. Angerer, K. Mühlegger, M. Amacker, H. Winter, *Nucleic Acids Res.* **2003**, *31*(10), 2630–2635.
[25] S. E. Lee, *Nucleic Acids Res.* **2001**, *29*(7), 1565–1573.
[26] J. P. Anderson, B. Angerer, L. A. Loeb, *Biotechniques* **2005**, *38*(2), 257–264.
[27] P. Kielkowski, J. Fanfrlík, M. Hocek, *Angew. Chem. Int. Ed.* **2014**, *53*(29), 7552–7555.
[28] K. Bergen, A.-L. Steck, S. Strütt, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *J. Am. Chem. Soc.* **2012**, *134*(29), 11840–11843.
[29] H. Cahová, A. Panattoni, P. Kielkowski, J. Fanfrlík, M. Hocek, *ACS Chemical Biology* **2016**, *11*(11), 3165–3171.
[30] A. Hottin, A. Marx, *Acc. Chem. Res.* **2016**, *49*(3), 418–427.
[31] S. Obeid, H. Bußkamp, W. Welte, K. Diederichs, A. Marx, *Chem. Commun.* **2012**, *48*(67), 8320.
[32] S. Obeid, H. Bußkamp, W. Welte, K. Diederichs, A. Marx, *J. Am. Chem. Soc.* **2013**, *135*(42), 15667–15669.
[33] S. Obeid, A. Baccaro, W. Welte, K. Diederichs, A. Marx, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*(50), 21327–21331.
[34] K. Sakthivel, C. F. Barbas III, *Angew. Chem. Int. Ed.* **1998**, *37*(20), 2872–2875.
[35] T. Gourlain, *Nucleic Acids Res.* **2001**, *29*(9), 1898–1905.
[36] J. Dadová, P. Vidláková, R. Pohl, L. Havran, M. Fojta, M. Hocek, *J. Org. Chem.* **2013**, *78*(19), 9627–9637.
[37] C. C. C. Johansson_Seechurn, M. O. Kitching, T. J. Colacot, V. Snieckus, *Angew. Chem. Int. Ed.* **2012**, *51*(21), 5062–5085.
[38] M. Yoshikawa, T. Kato, T. Takenishi, *Tetrahedron Letters* **1967**, *8*(50), 5065–5068.
[39] P. Espinet, A. M. Echavarren, *Angew. Chem. Int. Ed.* **2004**, *43*(36), 4704–4734.
[40] S. J. Connon, S. Blechert, *Angew. Chem. Int. Ed.* **2003**, *42*(17), 1900–1923.
[41] R. H. Grubbs, *Angew. Chem. Int. Ed.* **2006**, *45*(23), 3760–3765.
[42] J. G. Bertram, K. Oertell, J. Petruska, M. F. Goodman, *Biochemistry* **2010**, *49*(1), 20–28.
[43] Y. Li, V. Mitaxov, G. Waksman, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*(17), 9491–9496.
[44] K. Betz, D. A. Malyshev, T. Lavergne, W. Welte, K. Diederichs, T. J. Dwyer, P. Ordoukhanian, F. E. Romesberg, A. Marx, *Nat. Chem. Biol.* **2012**, *8*(7), 612–614.
[45] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta. Crystallogr. Sect. D: Biological Crystallography* **2010**, *66*(2), 213–221.
[46] A. V. Lebedev, D. Combs, R. I. Hogrefe, *Bioconjugate Chemistry* **2007**, *18*(5), 1530–1536.
[47] J. F. Teichert, M. Fañanás Mastral, B. L. Feringa, *Angew. Chem. Int. Ed.* **2011**, *50*(3), 688–691.
[48] K. Betz, D. A. Malyshev, T. Lavergne, W. Welte, K. Diederichs, F. E. Romesberg, A. Marx, *J. Am. Chem. Soc.* **2013**, *135*(49), 18637–18643.
[49] W. Kabsch, *Acta. Crystallogr. Sect. D: Biological Crystallography* **2010**, *66*(2), 125–132.
[50] K. Betz, F. Streckenbach, A. Schnur, T. Exner, W. Welte, K. Diederichs, A. Marx, *Angew. Chem. Int. Ed.* **2010**, *49*(30), 5181–5184.
[51] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta. Crystallogr. Sect. D: Biological Crystallography* **2010**, *66*, 486–501.
[52] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonrhein, G. Bricogne, grade, version 1.2.9, http://www.globalphasing.com, **2011**.
[53] Schrödinger, LLC, The PyMOL Molecular Graphics System, Version 1.7.6, **2015**.

10.1002/chem.201604515

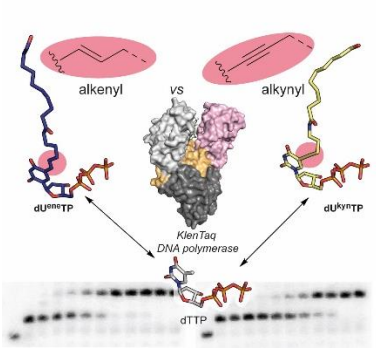
WILEY-VCH

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

Double or triple?: Two generally used linker types to attach functionalities to nucleobases are compared by functional and structural means. Incorporation studies reveal three modified substrates that are preferentially incorporated by KlenTaq than their natural counterparts. Crystal structures provide knowledge that can help in future design of functionalized nucleotide building blocks.



*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title