UC San Diego
UC San Diego Previously Published Works

Title
Double-headed nucleotides as xeno nucleic acids: information storage and polymerase recognition

Permalink
https://escholarship.org/uc/item/3dd5v8j0

Journal
Organic & Biomolecular Chemistry, 18(36)

ISSN
1477-0520

Authors
Beck, Kasper M
Krogh, Marie B
Hornum, Mick
et al.

Publication Date
2020-09-23

DOI
10.1039/d0ob01426e

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

HHS Public Access

Author manuscript

Org Biomol Chem. Author manuscript; available in PMC 2021 September 23.

Published in final edited form as:

Org Biomol Chem. 2020 September 23; 18(36): 7213–7223. doi:10.1039/d0ob01426e.

Double-headed nucleotides as xeno nucleic acids: information storage and polymerase recognition

Kasper M. Beck[a], **Marie B. Krogh**[a], **Mick Hornum**[a], **Paul T. Ludford III**[b], **Yitzhak Tor**[b], **Poul Nielsen**[a]

[a]Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Campusvej 55, DK-5230, Odense, Denmark

[b]Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0358, USA

Abstract

Xeno nucleic acids (XNAs) are artificial genetic systems based on sugar-modified nucleotides. Herein, we investigate double-headed nucleotides as a new XNA. A new monomer, $\mathbf{A_T}$, is presented, and together with previous double-headed nucleotide monomers, new nucleic acid motifs consisting of up to five consecutive A·T base pairs have been obtained. Sections composed entirely of double-headed nucleotides are well-tolerated within a DNA duplex and can condense the genetic information. For instance, a 13-mer duplex is condensed to an 11-mer modified duplex containing four double-headed nucleotides while simultaneously improving duplex thermal stability with +14.0 °C. Also, the transfer of information from double-headed to natural nucleotides by DNA polymerases has been examined. The first double-headed nucleoside triphosphate was prepared but could not be recognized and incorporated by the tested DNA polymerases. On the other hand, it proved possible for Therminator DNA polymerase to transfer the information of a double-headed nucleotide in a template sequence to natural DNA under controlled conditions.

Table of contents entry:

Double-headed nucleotides increase the thermal stability of duplexes and are recognized by Therminator DNA polymerase

Graphical Abstract

Corresponding Author: pouln@sdu.dk.

†Electronic Supplementary Information (ESI) available: MALDI-TOF of oligonucleotides, PAGE gels, NMR spectra. See DOI: 10.1039/d0ob01426e

Conflicts of interest
There are no conflicts to declare.

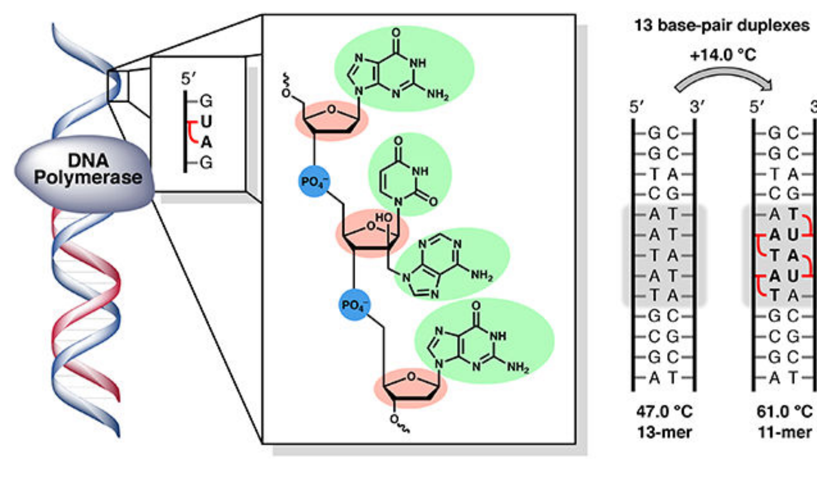
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Introduction

Xeno nucleic acids (XNA) are artificial genetic systems that consist solely of synthetic nucleotide monomers, in which the chemical composition of the sugar moiety has been changed. The nucleobases and phosphodiester linkages of natural nucleic acids (DNA and RNA) are retained.[1] Examples include the use of cyclohexene (CeNA),[2] threose (TNA),[3] acyclic L-threoninol (L-*a*TNA),[4] 2′-*O*,4′-*C*-methylene linked ribose (LNA)[5] or a glycol (GNA) instead of the natural (2′-deoxy)ribose.[6] Typically, XNAs form stable homoduplexes that are often more stable than heteroduplexes with complementary DNA and RNA counterparts. Also other properties, such as nuclease resistance and toxicity, are affected by altering the sugar moiety. In this regard, many XNAs have useful biological applications as e.g. aptamers and antisense oligonucleotides.[7] Importantly, XNAs provide orthogonal systems for the storage of genetic information. Translation from XNA to DNA and back might not be possible by standard DNA polymerases. However, for many XNAs, appropriate polymerases have been found in nature or genetically engineered for this specific purpose.[8]

We and others have focused on developing artificial nucleic acids where each sugar moiety caries two nucleobases, the so-called double-headed nucleotides.[9] Recently, we have presented a design based on arabinouridine with a second nucleobase connected to the 2′-position through a methylene-linker, $\mathbf{U_X}$ (Fig. 1a).[10–12] We envision these double-headed nucleic acids to have large potential as XNA, and in this paper we explore how they meet the criteria for an artificial genetic system.

Four different analogues of $\mathbf{U_X}$ ($\mathbf{U_T}$, $\mathbf{U_A}$, $\mathbf{U_C}$, $\mathbf{U_G}$) with each of the four natural nucleobases attached to the 2′-position have already been synthesized and studied.[11, 12] Also the 1′-base can be varied, increasing the theoretical number of combinations to 16 (considering uracil and thymine to be equal), of which we have currently prepared five including our recent cytidine-based $\mathbf{C_C}$ monomer.[13] Both nucleobases of the double-headed nucleotides have shown to participate in Watson–Crick base-pairing when introduced into a DNA duplex in exchange for the corresponding dinucleotide (Fig. 1b). The thermal stability of the duplex is not significantly affected by the incorporation of either $\mathbf{U_C}$ (+1.0 °C) or $\mathbf{U_G}$ (0.0 °C) instead of the corresponding dinucleotides UC or UG. In the case of $\mathbf{C_C}$, the stability is even

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

increased by +4.0 °C. On the other hand, incorporation of $\mathbf{U_T}$ (−3.5 °C) or $\mathbf{U_A}$ (−2.5 °C) causes some degree of destabilization compared to the unmodified reference.[10–13] When double-headed nucleotides are properly positioned in each strand of a duplex, an additional base-pair can form between the two 2′-bases (Fig. 1c). The full set of matched 2′–2′-pairs using $\mathbf{U_T}$, $\mathbf{U_A}$, $\mathbf{U_C}$ and $\mathbf{U_G}$ has been found to give a general increase in duplex stability of +6.0–6.5 °C and +5.5–9.0 °C for the A·T and C·G base pairs, respectively, compared to unmodified duplexes with the same number of nucleobases.[11] Also, acceptable mismatch discrimination was found for the 2′–2′-pairs. Finally, an internal double base-pair where two 1′–2′ A·U base-pairs form within two $\mathbf{U_A}$ monomers is tolerated (Fig. 1d) giving rise to an increase in the thermal stability of +4.5 °C.[10] Evidently, double-headed nucleotides can recognize and form stable Watson-Crick base pairs with both native dinucleotides and complementary double-headed nucleotides within a DNA duplex, thereby condensing the genetic information to fewer nucleotides.

In this regard, we want to expand our understanding of the double-headed nucleotide as an XNA by studying two new aspects: (1) The behavior of double-headed nucleotides in more extended, ultimately fully modified double-headed nucleic acids, and (2) the possibility of information transfer between double-headed nucleotides and natural DNA.

On the first aspect, we need to realize new combinations of nucleobases in our double-headed nucleotide design (Fig. 1a) beyond the five pyrimidine-based monomers ($\mathbf{U_X}$ and $\mathbf{C_C}$) already synthesized. The first purine-based double-headed nucleotide is needed, and herein we present the synthesis and evaluation of the sixth analogue in the form of $\mathbf{A_T}$. With this in hand, we study a window that contains up to five consecutive double-headed nucleotide A·T base pairs placed centrally in a DNA duplex.

For the second aspect, we continue our focus on the A·T base pair of double-headed nucleotides in relation to its recognition by several DNA polymerases. The potential transfer of information from a template containing $\mathbf{U_A}$ is studied, and we also present the synthesis of the first double-headed nucleoside triphosphate $\mathbf{U_ATP}$ and test its ability to act as a substrate for DNA polymerases.

Results and discussion

Chemical synthesis

The modified oligonucleotides were prepared from phosphoramidite building blocks using standard solid-phase DNA synthesis. The synthesis of phosphoramidite **6** (Scheme 1) started from adenosine, which was converted to the known protected ketone **1** through *N*-6-benzoylation,[14] then 3′,5′-*O*-TIPDS protection[15] and finally 2′-oxidation.[16] Ketone **1** was stereoselectively converted to the 2′(*S*)-spiroepoxide **2** in a Corey–Chaykovsky epoxidation by adapting the procedure for the equivalent reaction with the 2′-keto uridine derivative.[10, 17] The epoxide was then opened by treatment with deprotonated thymine to form the double-headed nucleoside **3** in a yield of 46%. *N*1-alkylation was confirmed by the presence of a $^{3}J_{CH}$ coupling between the 2′-methylene protons and both C2′′ and C6′′ of thymine in the HMBC spectrum of **3**, and the 2′-stereocenter was assigned using through-space couplings between H3′ and 2′-OH as well as H1′ and H6′ in the NOESY spectrum.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Removal of the silyl protecting group was achieved to give **4**, which was then tritylated to give **5** and finally phosphitylated using standard protocols to give the fully protected phosphoramidite **6**. The corresponding phosphoramidites of the double-headed nucleotides **U$_\mathbf{T}$** and **U$_\mathbf{A}$** were prepared using known procedures.[10] The modified nucleotides were introduced into oligonucleotides by standard solid-phase DNA synthesis using 4,5-dicyanoimidazole as activator and extended coupling times (15 min). Finally, the oligonucleotides were purified by reversed-phase HPLC and characterized by MALDI-TOF MS.

The double-headed nucleoside triphosphate, **U$_\mathbf{A}$TP**, was prepared in two steps from **7** (Scheme 2), which is an intermediate in the reported[10] synthesis of **U$_\mathbf{A}$**. First, **7** was treated with methanolic ammonia to remove the benzoyl protecting group, which gave the naked nucleoside **8**. In a one-pot procedure, **8** was then converted to the triphosphate **U$_\mathbf{A}$TP** by treatment with freshly distilled $POCl_3$ and tributylammonium pyrophosphate following standard procedures.[18] Purification consisting of ion-exchange flash chromatography and reversed-phase HPLC followed by precipitation with 3% sodium perchlorate in acetone gave the analytically pure triphosphate as a sodium salt, which was characterized by ^{1}H and ^{31}P NMR spectroscopy and mass spectrometry.

Hybridization studies

A 13-mer DNA sequence (Fig. 2, entry 1) containing a central five base pair wide A·T-window (marked in grey) flanked by C·G rich regions was designed for the study. With this sequence we were able to study the effect on duplex stability systematically by gradually exchanging dinucleotides with double-headed nucleotides, and thereby proceed towards a fully modified five base pair window. The thermal stability of each duplex was determined as the melting temperature (T_m), which was obtained from the maximum of the first derivative of the UV (A_{260}) melting curve.

First, a single **A$_\mathbf{T}$** monomer was introduced (Fig. 2, duplex 2) in exchange for the central 5′-AT dinucleotide and this modified 12-mer was then hybridized with the complementary unmodified 13-mer sequence. The modification was well-accommodated in the duplex structure, however, with a decrease in melting temperature of –3.0 °C compared to the unmodified reference (Fig. 2, duplex 1). This result is comparable to that of the reversed analogue, **U$_\mathbf{A}$**, that caused a T_m of –2.5 °C when introduced into a 12-mer duplex.[12]

Since the **A$_\mathbf{T}$** monomer is previously unstudied, a mismatch study was performed using this duplex to explore the base-pairing abilities of both the 1′-adenine and the 2′-thymine of **A$_\mathbf{T}$** (Fig. 3). At the 1′-position (Fig. 3, second row), mismatches were clearly discriminated with T_m-values ranging from –10.0 °C (A·C mismatch) to –6.0 °C (A·G mismatch) and there was a strong correlation with the mismatch discrimination for the unmodified duplexes (first row). For the 2′-thymine (Fig. 3, fourth row), a decent discrimination of both thymine (–6.5 °C) and cytosine (–9.0 °C) was found, whereas, the discrimination against guanine was lower (–3.0 °C). Nevertheless, this discrimination of guanine by the 2′-thymine is an improvement compared to the 2′-thymine of **U$_\mathbf{T}$** that discriminates an opposite guanine with –1.5 °C.[10] Altogether, the mismatch study showed that **A$_\mathbf{T}$** is capable of forming specific

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Watson–Crick base pairs using both its nucleobases, and so, function as a condensed dinucleotide.

With the introduction and evaluation of the new $\mathbf{A_T}$ monomer, the study proceeded towards duplexes containing multiple double-headed nucleotides (Fig. 2). Additional oligonucleotides with one or two adjacent double-headed nucleotides were prepared. To give full insight, these were first hybridized to their complementary unmodified DNA strands (Fig. 2, duplexes 3–6), and the melting temperatures were determined. For duplexes containing single incorporations of either $\mathbf{A_T}$ or $\mathbf{U_A}$, the changes in T_m were the expected −2.5 °C to −3.0 °C compared to reference duplex 1. For duplexes that have two modifications within one of its strands, the destabilization was additive with T_m-values of −5.5 °C and −6.0 °C, indicating that there is no extra penalty for multiple incorporations, even when they are placed consecutively. In this study, we compare uracils and thymines directly, although duplex stability is expected to be affected by the additional methyl group.

The modified oligonucleotides were then hybridized to each other to form duplexes with modifications in both strands (Fig. 2, duplexes 7–12). In duplex 7, where two single-modified 12-mers were combined to give the 2′–2′-pair, an increase in T_m of +4.5 °C was observed. This is slightly lower, but comparable to the corresponding 2′–2′-pairs with $\mathbf{U_T}$ and $\mathbf{U_A}$ showing T_m's of +6.0–6.5 °C in another sequence.[12] Based on our previous study of these 2′–2′-pairs, this increase in stability confirms the formation of an actual 2′–2′ A·T base pair.[12] When an additional double-headed nucleotide monomer was inserted in one or the other strand, the duplex stability was increased further (Fig. 2, entries 8 and 9). This result demonstrates the profound increase in duplex stability associated with positioning double-headed nucleotides opposite each other. In fact, interchanging one dinucleotide in duplexes 5 and 6 with the corresponding double-headed nucleotide (duplexes 8 and 9) causes a remarkable rise in T_m of +13.0 °C and +12.0 °C, respectively. In duplex 10, having the highest degree of modification, an increase in the thermal stability of +14.0 °C was found compared to the unmodified duplex 1. Hence, the use of double-headed nucleotides here enables the storage of 13 base-pairs within a very stable 11-mer duplex, when these are placed consecutively in both strands of the duplex.

Another stabilizing motif was the internal double-pair within two $\mathbf{A_T}$ monomers in duplex 11. Here, the T_m of the duplex was increased with +10.0 °C as a consequence, which is notably more than the corresponding double-pair using two $\mathbf{U_A}$ monomers, which caused a T_m of +4.5 °C in a different sequence.[10] Interestingly, the stabilizing effect of this internal double-pair could also compensate for the addition of a neighboring incorporation, as it can be seen in duplex 12. Here, the melting temperature of the duplex remained unaffected at 57.0 °C, when an additional double-headed nucleotide was placed adjacent to the double-pairing unit (duplexes 11 and 12). In practice, the incorporation of a single $\mathbf{A_T}$ monomer in the otherwise unmodified strand of duplex 5 increased the melting temperature with +15.5 °C (41.5 °C compared to 57.0 °C).

Altogether, the results show that strong base pairs are formed when double-headed nucleotides are present in both strands of the A·T window, and that duplex stability increases as the number of phosphates are reduced. Generally, condensed DNA in the form of double-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

headed nucleotides is a highly stable concept. Recognition of native DNA by double-headed nucleotides is possible with a slight thermal destabilization for A·T base pairs. From previous studies,[12, 13] we know that the C·G base pairs of double-headed nucleotides are notably stronger. Therefore, we expect even more stable duplexes as we go beyond the A·T base pair window in future studies (work in progress).

For all duplexes (1–12, Fig. 2), circular dichroism (CD) spectra were recorded in order to gain information on possible structural deviations (Fig. S1). A clear general consistency with the B-type duplex was found for all duplexes, and only small fluctuations in the CD spectra were observed. No correlation to number or position of double-headed nucleotides was observed. The CD spectra are in accordance with the results of the previous modelling study that found a remarkable structural similarity between duplexes containing double-headed nucleotides and natural DNA on several parameters including helix diameter and groove widths.[12]

Polymerase studies with double-headed nucleotides

The second part of the study explores the compatibility of double-headed nucleotides with polymerases and the transfer of their information. In practice, we set out to determine how double-headed nucleotides are interpreted by a selection of four DNA polymerases both as a nucleoside triphosphate ($\mathbf{U_A TP}$) for template-specific primer extensions, but also in the form of a modified template strand containing the double-headed nucleotide $\mathbf{U_A}$. For the sake of clarity, a complete list of primer and template sequences is presented in Fig. 4. The DNA polymerases used in this study collectively lack the 3′→5′ exonuclease activity, and consist of the *Thermus aquaticus* (*Taq*) DNA polymerase I and the Klenow fragment exo– (Kf–) of the *E. coli* DNA polymerase I from the A-family, together with the genetically engineered B-family polymerases Deep Vent exo– (DV–) and Therminator, which are derived from the *Pyrococcus* sp. GB-D and the *Thermococcus* sp. 9°N DNA polymerases, respectively. The latter was chosen for its ability to use several modified nucleoside triphosphates in the synthesis of nucleic acids.[19, 20]

The double-headed nucleoside triphosphate as a substrate for DNA polymerases.

The study was conducted on a primer–template system consisting of a previously studied[21, 22] 18 nt primer (**P4**) that was fluorescently labelled with fluorescein (FAM) at the 5′-position and a 26 nt template: the unmodified **T1** or the $\mathbf{U_A}$-modified **T2**, in which a 5′-TA dinucleotide have been exchanged with the $\mathbf{U_A}$ double-headed nucleotide. The primer-template duplex (0.1 μM) was incubated in the presence of the particular DNA polymerase and $\mathbf{U_A TP}$ (1000 μM), and the products of the reaction were separated by denaturing 20% polyacrylamide gel electrophoresis (PAGE) and visualized by fluoroimaging (Fig. 5). For each enzyme, the first lane shows a negative control without triphosphates; the second lane shows a positive control using TTP; and the third and fourth lanes show the experiment with $\mathbf{U_A TP}$ in the presence of templates **T1** and **T2**, respectively. In the preliminary experiments, primer degradation due to pyrophosphorolysis was observed with Therminator. This is a well-known weakness of Therminator,[23,24] and as a consequence all the following experiments with Therminator were conducted in the presence of thermostable inorganic pyrophosphatase (TIPP) to overcome this problem. Primer extension was observed in all the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

positive controls and the +1 nt 'overreaction' in the case of Therminator is not uncommon.[19] When **U_ATP** was used instead, no reaction happened with either of the DNA polymerases.

A series of experiments were performed with each polymerase to probe the reaction conditions by varying both enzyme and trisphosphate concentration as well as incubation temperature and time (not shown). However, these also proved unsuccessful, and so it was concluded that **U_ATP** was not a substrate for any of the four tested DNA polymerases.

Transferring information from double-headed nucleotides to natural DNA using DNA polymerases.

The other main point of interest regarding enzymatic recognition of double-headed nucleotides concerns the transfer of information from double-headed nucleotides to natural DNA. A series of single nucleotide primer extension experiments were conducted using the modified **U_A**-template **T2**, the **P4** primer and each of the natural dNTPs (Fig. 6). The first and last lane correspond for each enzyme to the negative controls using the 18 nt **P4** primer and the 19 nt **P5** primer, respectively. From the results, it was clear that Kf− specifically incorporated pA opposite of the **U_A** nucleotide, apparently without noticing the 2′-adenine (2′-A).[12] Likewise, pA was incorporated in the cases of both *Taq* and DV−, however, primer extension appeared to be more favorable with pC, suggesting a complete skip of the entire **U_A** nucleotide. With Therminator, a +1 nt product was observed in all but the dCTP case. It was decided to further examine Therminator for its unique ability to incorporate pT opposite of **U_A** and also Kf− as it specifically incorporates pA.

For the Therminator polymerase, it was essential to confirm that incorporation of pT was in fact due to recognition of the 2′-A. Therefore, the experiments with TTP were extended to also include incorporation of the second nucleotide in the **T2:P4** duplex (Fig. 7). In these experiments, TTP and the second dNTP were added sequentially with 10 min incubation after each addition to ensure initial incorporation of pT. A specific incorporation of pA was observed, which supports the belief that Therminator is able to recognize both the 1′-uracil and the 2′-adenine of **U_A**. The same specificity for pA incorporation was observed when treating the **T2**:**P5** duplex, which contains an extra T at the 3′-end of the primer, with each dNTP under the same conditions (see sup. information, Fig. S2). A control experiment excluded the possibility of non-templated primer extensions. Hence the primer was not extended by any of the dNTPs by Therminator in the absense of a template strand (Fig. S3).

For Kf−, it was of interest to verify the immediate suggestion that it recognizes the **U_A** nucleotide as a natural U without taking the additional 2′-A base into account. Therefore, experiments were performed where the **T2:P4** duplex was treated with a mixture of dATP and one other dNTP in the presence of Kf− (Fig. 8). As anticipated, the primer was extended in all four experiments due to the presence dATP. However, primer extension only proceeded further than 1 nt in the experiment that also included dCTP. This observation is in agreement with G being the next nucleotide following **U_A** in the template sequence. Thus, Kf− identifies the **U_A** double-headed nucleotide as a U and is capable of continuing the template-specific primer extension reaction thereafter.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Next, a full-length primer extension study was performed to test if the primer length had any effect on the way Kf– recognized $\mathbf{U_A}$. By hybridizing the **T2** template with each of the primers **P1**–**P7**, which varies in length from 15 to 22 nt, we probed the proximity of the modified nucleotide to the initiation site of the polymerase reaction (Fig. 9). When the 3′-terminal of the primer was situated upstream relative to $\mathbf{U_A}$, the formation of what appeared to be a 25 nt product was observed. This indicates that $\mathbf{U_A}$ in each case is reversely transcribed to an A, thereby making the product 1 nt shorter than the 26 nt reference that contains a 5′-TA dinucleotide at this spot. No reaction was observed when the primer stopped at the position of the additional 2′-nucleobase or the natural 1′-nucleobase. Yet, if the primer extends past the $\mathbf{U_A}$ (**P7**), Kf– is capable of extending it to give the full-length 26 nt product, and so, the presence of $\mathbf{U_A}$ in the template is apparently not interfering with Kf–'s ability to recognize and extend the primer-template duplex, unless the recognition site coincides with the position of double-headed nucleotide.

A similar full-length primer extension study was performed with Therminator in order to clarify whether this polymerase was able to recognize and extend the **T2**:**P5** and **T2**:**P6** duplexes (see sup. information, Fig. S4). Unfortunately, only a single nt extension was observed for the **P5** primer, while no extension was observed in the case of the **P6** primer. This indicates that Therminator is unable to recognize and extend the motif in the **T2**:**P6** duplex, as it was also the case with Kf–.

Discussion

This study of double-headed nucleotides showed that the genetic information of A·T base pairs can in fact be concentrated to fewer nucleotides with the use of double-headed nucleotides. Moreover, a profound increase in duplex stability accompanied this modification. In practice, the genetic information of 13 base pairs was contained within two double-modified 11-mers with a +14.0 °C increase in the T_m (Fig. 2, duplex 10). With these results and our former demonstration that single incorporations of C·G base pairs gave even more stable duplexes,[12, 13] we are convinced that an expansion of the concept to a higher number of modifications or even a fully modified double-headed nucleic acid duplex is achievable. The bottleneck is the synthesis of the remaining monomers, which are expected to be achievable based on the presented synthesis of $\mathbf{A_T}$. Currently, the development of replacements for 2′-adenines and 2′-thymines with improved affinity and specificity is ongoing, and for the 2′-adenine it has already been realized with the introduction of 2,6-diaminopurine (D) in the $\mathbf{U_D}$ monomer.[12]

Similarly, in the study of information transfer, the A·T base pairs of the double-headed nucleotides displayed interesting features. The primer extension experiments with the $\mathbf{U_A}$-modified template are summarized in Fig. 10, and it is shown that Therminator can recognize and transfer $\mathbf{U_A}$ and then Kf– can further extend the primer to give the full-length product. Nevertheless, recognition and extension of the motif that follows immediately after $\mathbf{U_A}$ remains a challenge, despite the fact that duplexes containing single incorporations of double-headed nucleotides generally display properties similar to those of unmodified duplexes. In a previous modelling study it was demonstrated that the presence of double-headed nucleotides in DNA duplexes leads to some adjustment of certain backbone torsion

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

angles to retain a similar global duplex structure.[12] It is therefore likely to be the absent section of backbone that limits the binding opportunities of the DNA polymerase.

Under controlled conditions, it was possible for Therminator to recognize both nucleobases of $\mathbf{U_A}$. The initial recognition of the 2′-A was unspecific, however, as the primer was extended when treated with each of TTP, dATP and dGTP. Yet, if the 'correct' pT was incorporated, a specific incorporation of pA opposite of the 1′-U followed (Fig. 7). The incorporation of pT is thought to originate from an actual recognition of the 2′-A as adenosine. In the case of pA incorporation on the other hand, the 2′-A is undetected as in the case of Kf–, and so only the 1′-U is recognized. Finally, the incorporation of pG can be justified by the reduced mismatch discrimination of guanine of −1.0 °C exhibited by the 2′-A of $\mathbf{U_A}$.[11] So, it is reasonable that pG can be incorporated opposite of the 2′-A, if pT can. A replacement of the 2′-A with 2,6-diaminopurine might solve this specificity problem.[12]

The fact that the double-headed nucleoside triphosphate was not a substrate for the tested polymerases is not unexpected due to its high degree of modification. Recognition of 2′-modified nucleoside triphosphates are generally challenging, due to the so-called 'steric gate' of DNA polymerases that is used to select dNTPs over the natural excess of rNTPs, however, they are not unheard of.[25–27] In this case the introduction of an entire nucleobase in this position is apparently too challenging, despite its complementarity to the subsequent nucleotide in the template sequence. A vast variety of genetically engineered DNA polymerases exist, and it is possible that some of the more heavily modified ones, or instead RNA polymerases, might be able to recognize and incorporate $\mathbf{U_A TP}$. Alternatively, a polymerase could be specifically engineered for this function.[28]

Overall, the double-headed nucleotides in this study have been shown to possess key features of XNAs; a structural retainment of the native nucleobases and the phosphodiester linkages, but with an altered sugar moiety, and the ability to form specific and stable duplexes. In future studies, we hope to be able to fully classify the double-headed nucleotides as an XNA with the ability to form a stable artificial genetic system independent of native nucleic acids.

Conclusions

A new double-headed nucleotide, $\mathbf{A_T}$, was conveniently synthesized and incorporated into DNA, and it was shown that sections consisting entirely of double-headed nucleotides are well-accommodated in a DNA duplex. They improve the duplex thermal stability when incorporated in both strands (up to +10 °C for a pair), and thus effectively condense the information of natural DNA. In parallel, the first double-headed nucleoside triphosphate, $\mathbf{U_A TP}$, was prepared and studied for its use as a substrate for DNA polymerases, however, without recognition by any of the four studied polymerases. Finally, the recognition of the double-headed nucleotide, $\mathbf{U_A}$, in a template for primer extension reactions by DNA polymerases was investigated and it was found that Therminator DNA polymerase, under controlled conditions, was able to recognize both nucleobases of the double-headed nucleotide in the template, albeit not with full specificity for the 2′-nucleobase. These results encourage an extension of the study, which is currently ongoing towards the inclusion

of C·G base pairs in fully modified double-headed nucleic acid sequences. Based on our preliminary results,[11] the C·G base pair is significantly more stable and specific than the A·T base pair in the double-headed nucleic acid design and thereby expected to give higher specificity in the recognition of a polymerase template. Hereby, we expect to establish the double-headed nucleic acids as an important xeno nucleic acid system with the potential as an artificial genetic system with condensed information as compared to natural DNA.

Experimentals

All commercial reagents were used as supplied except CH_2Cl_2, which was distilled prior to use. Klenow fragment 3′ 5′ exo– DNA polymerase (Kf–) and the natural dNTPs were purchased from Promega. *Taq*, Deep Vent exo– (DV–) and Therminator DNA polymerases and thermostable inorganic pyrophosphatase (TIPP) were purchased from New England Biolabs. Anhydrous solvents were dried over 3 Å (CH_3CN) or 4 Å (CH_2Cl_2, DCE, DMSO, DMF, Et_3N, petroleum ether, pyridine, THF, toluene) activated molecular sieves. Reactions were carried out under argon or nitrogen whenever anhydrous solvents were used. All reactions were monitored using TLC analysis with Merck silica gel plates (60 F_{254}). For visualisation, the plates were exposed to UV light (254 nm) and/or immersed in a solution of 5% H_2SO_4 in methanol (v/v) followed by charring. Column chromatography was performed with Silica Gel 60 (particle size 0.040–0.063 μm, Merck). ^{1}H NMR, ^{13}C NMR and ^{31}P NMR spectra were recorded at ambient temperature on a Bruker Avance III 400 instrument or a Varian VX 500 instrument. Chemical shift values (δ) are reported in ppm relative to either tetramethylsilane ($\delta_{H,C}$: 0.0) or the deuterated solvents ($CDCl_3$: δ_H 7.26, δ_C 77.16; DMSO-d_6: δ_H 2.50, δ_C 39.52) and relative to 85% H_3PO_4 as external standard for ^{31}P NMR. Assignment of NMR signals are based on 2D spectra and follow standard nucleoside convention, and the 2′-CH_2 group is assigned as C6′. Signals assigned with a double prime belong to the nucleobase in the 2′-position. High resolution mass spectra were recorded on a Bruker micrOTOF-Q II (ESI) quadrupole time of flight instrument in positive ion mode.

6-*N*-Benzoyl-3′,5′-*O*-(tetraisopropyldisiloxane-1,3-diyl)adenosine-2′(*S*)-spiroepoxide (2)

A mixture of trimethylsulfoxonium iodide (207 mg, 0.938 mmol) and NaH (60% suspension in mineral oil, 30 mg, 0.75 mmol) in anhydrous DMSO (4 mL) was stirred under N_2 atmosphere for 1 h. Anhydrous THF (4 mL) was added and the mixture was cooled to 0 °C. A solution of **1** (0.384 g, 0.628 mmol) in anhydrous THF (4 mL) was slowly added and this mixture was stirred for 30 min at 0 °C. H_2O (10 mL) was slowly added, followed by addition of a saturated aq. solution of NH_4Cl (7 mL) and EtOAc (30 mL). After stirring for 15 min, the phases were separated, and the aq. phase was extracted with EtOAc (3 × 10 mL). The combined organic phase was dried (Na_2SO_4) and concentrated under reduced pressure, and the residue was purified by silica gel column chromatography (0–100% EtOAc in petroleum ether, v/v) to give epoxide **2** (226 mg, 0.36 mmol, 58%) as a white foam. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.05 (s, 1H, NH), 8.76 (s, 1H, H2), 8.16 (s, 1H, H8), 8.05–7.98 (m, 2H, Bz), 7.65–7.56 (m, 1H, Bz), 7.55–7.48 (m, 2H, Bz), 6.33 (s, 1H, H1′), 4.96 (d, J = 8.5 Hz, 1H, H3′), 4.23–4.10 (m, 2H, H5′), 4.04 (ddd, J = 8.6, 4.2, 3.1 Hz, 1H, H4′), 3.33 (d, J = 5.3 Hz, 1H, H6$_a$′), 2.99 (d, J = 5.3 Hz, 1H, H6$_a$′), 1.19–0.90 (m, 28H, 4 × *i*-Pr); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.6 (Bz C=O), 153.0 (C2), 152.0 (C4), 149.7 (C6), 141.9

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(C8), 133.9 (Bz), 132.9 (Bz), 129.0 (Bz), 128.0 (Bz), 122.9 (C5), 81.6 (C4′), 80.9 (C1′), 69.8 (C3′), 65.9 (C2′), 61.8 (C5′), 49.4 (C6′), 17.6, 17.5, 17.4, 17.1, 17.1, 17.0, 17.0, 13.4, 13.2, 12.9, 12.7 (4 × *i*-Pr); HRMS-ESI *m*/*z* 626.2796 $[M+H]^+$; calcd ($C_{30}H_{44}N_5O_6Si_2^+$) 626.2825.

6-*N*-Benzoyl-9-(2′-*C*-(thymin-1-yl)methyl-3′,5′-*O*-(1,1,3,3-tetraisopropyldisiloxane-1,3-diyl)-β-D-arabinofuranosyl)adenine (3)

A suspension of thymine (0.697 g, 5.53 mmol) and NaH (60% suspension in mineral oil, 140 mg, 3.50 mmol) in anhydrous DMF (6 mL) was stirred under N_2 atmosphere for 2 h. A solution of **2** (0.867 g, 1.39 mmol) in anhydrous DMF (4 mL) was added, and the mixture was stirred at 65 °C for 21 h. After cooling to rt, a saturated aq. solution of NH_4Cl (5 mL) was added followed by H_2O (20 mL) and EtOAc (20 mL). The phases were separated, and the aq. phase was extracted with EtOAc (3 × 10 mL). The combined organic phase was dried (Na_2SO_4) and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (0–80% EtOAc in petroleum ether, v/v) to give the double-headed nucleoside **3** (0.49 g, 0.64 mmol, 46%) as a white foam containing 0.24 eq. DMF. ^1H NMR (400 MHz, DMSO-d_6) δ 11.35 (s, 1H, NH), 11.21 (bs, 1H, NH), 8.59 (s, 1H, H2), 8.33 (s, 1H, H8), 8.04 (d, J = 7.2 Hz, 2H, Bz), 7.67–7.61 (m, 1H, Bz), 7.55 (t, J = 7.5 Hz, 2H, Bz), 7.44 (d, J = 1.2 Hz, 1H, H6′′), 6.30 (s, 1H, H1′), 6.04 (s, 1H, 2′OH), 4.59 (d, J = 8.1 Hz, 1H, H3′), 4.35 (d, J = 15.0 Hz, 1H, $H6_a$′), 4.19–3.96 (m, 4H, H4′, $H5_a$′, $H5_b$′, $H6_b$′), 1.65 (d, J = 1.2 Hz, 3H, 5′′-CH_3), 1.26–0.89 (m, 28H, 4 × *i*-Pr); ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.5 (Bz C=O), 164.0 (C4′′), 152.5 (C2′′), 152.1 (C2), 151.4 (C4), 150.1 (C6), 142.4 (C8), 142.1 (C6′′), 133.3 (Bz), 132.3 (Bz), 128.4 (Bz), 128.4 (Bz), 124.9 (C5), 108.3 (C5′′), 84.3 (C1′), 79.8 (C2′), 79.7 (C4′), 76.8 (C3′), 61.2 (C5′), 49.6 (C6′), 17.3, 17.2, 17.1, 17.1, 16.9, 16.8, 16.7, 16.7, 12.6, 12.4, 12.3, (4 × *i*-Pr), 11.9 (5′′-CH_3); HRMS-ESI *m*/*z* 752.3221 $[M+H]^+$; calcd ($C_{35}H_{50}N_7O_8Si_2^+$) 752.3254.

6-*N*-Benzoyl-9-(2′-*C*-(thymin-1-yl)methyl-β-D-arabinofuranosyl)adenine (4)

A solution of **3** (137 mg, 0.18 mmol) in anhydrous THF (3 mL) was degassed with N_2 (15 min), $Et_3N{\cdot}3HF$ (0.10 mL, 0.61 mmol) was added and the reaction mixture was stirred at rt for 24 h. TMS-acetylene (0.23 mL, 1.62 mmol) was added and the mixture was stirred at rt for 4 h. After the addition of methanol (25 mL), the mixture was concentrated and the residue was purified by silica gel column chromatography (0–7% CH_3OH in CH_2Cl_2, v/v) to give nucleoside **4** (82 mg, 0.16 mmol, 88%) as a white solid. ^1H NMR (400 MHz, DMSO-d_6) δ 11.32 (s, 1H, NH), 11.18 (s, 1H, NH), 8.64 (s, 1H, H2), 8.53 (s, 1H, H8), 8.04 (d, J = 7.2 Hz, 2H, Bz), 7.65 (t, J = 7.4 Hz, 1H, Bz), 7.55 (t, J = 7.5 Hz, 2H, Bz), 7.28 (d, J = 0.9 Hz, 1H, H6′′), 6.32 (s, 1H, H1′), 6.17 (s, 1H, 2′OH), 5.97 (d, J = 4.7 Hz, 1H, 3′OH), 5.47 (t, J = 5.2 Hz, 1H, 5′OH), 4.13 (t, J = 4.1 Hz, 1H, H3′), 4.09–4.02 (m, 1H, $H6_a$′), 4.00–3.93 (m, 2H, H4′, $H6_b$′), 3.81–3.65 (m, 2H, H5′), 1.62 (d, J = 0.9 Hz, 3H, 5′′-CH_3); ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.5 (Bz C=O), 164.0 (C4′′), 152.3 (C2′′), 152.3 (C4), 151.5 (C2), 150.0 (C6), 143.4 (C8), 142.1 (C6′′), 133.4 (Bz), 132.4 (Bz), 128.4 (Bz), 124.7 (C5), 108.4 (C5′′), 85.0 (C1′), 84.8 (C4′), 80.2 (C2′), 76.0 (C3′), 61.0 (C5′), 47.7 (C6′), 11.8 (5′′-CH_3); HRMS-ESI *m*/*z* 510.1713 $[M+H]^+$; calcd ($C_{23}H_{24}N_7O_7^+$) 510.1732.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

6-*N*-Benzoyl-9-(5′-*O*-(4,4′-dimethoxytrityl)-2′-*C*-(thymin-1-yl)methyl-β-D-arabinofuranosyl)adenine (5)

After co-evaporating with anhydrous pyridine (2 × 5 mL), nucleoside **4** (217 mg, 0.43 mmol) was redissolved in anhydrous pyridine (3 mL). DMTrCl (0.29 g, 0.84 mmol) and 4-dimethylaminopyridine (11 mg, 89 μmol) were added, and the mixture was stirred for 22 h at rt. EtOH (96%, 6 drops) was added and the reaction was concentrated under reduced pressure. The residue was purified by silica gel column chromatography (0–3% CH_3OH in CH_2Cl_2, v/v) to give **5** (158 mg, 0.19 mmol, 44%) as a white foam containing 0.29 eq. pyridine. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.98 (bs, 1H, NH), 9.39 (s, 1H, NH), 8.78 (s, 1H, H2), 8.55 (s, 1H, H8), 8.08 (dt, J = 7.1, 1.4 Hz, 2H, Bz), 7.63 (t, J = 7.4 Hz, 1H, Bz), 7.54 (t, J = 7.5 Hz, 2H, Bz), 7.39–7.32 (m, 2H, DMTr), 7.33–7.22 (m, 7H, DMTr), 6.97 (s, 1H, H6′′), 6.83 (dd, J = 9.0, 2.6 Hz, 4H, DMTr), 6.45 (s, 1H, H1′), 5.68 (s, 1H, 2′OH), 5.53 (d, J = 2.4 Hz, 1H, 3′OH), 4.65 (d, J = 14.7 Hz, 1H, $H6_a'$), 4.28 (s, 1H, H4′), 3.85 (br s, 1H, H3′), 3.83–3.74 (m, 7H, $H5_a'$, 2 × DMTr OCH_3), 3.45 (dd, J = 11.3, 2.4 Hz, 1H, $H5_b'$), 3.06 (d, J = 14.9 Hz, 1H, $H6_b'$), 1.84 (s, 3H, 5′′-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 164.9 (Bz C=O), 164.0 (C4′′), 159.0 (DMTr), 152.9 (C2′′), 152.6 (C2), 152.4 (C4), 149.6 (C6), 143.1 (C6′′), 142.9 (C8), 134.4, 134.2 (DMTr), 133.3 (Bz), 132.9 (Bz), 130.2 (DMTr), 130.2, 129.2 (DMTr), 128.9 (Bz), 128.2, 128.2, 128.1, 127.5 (DMTr, Bz), 123.1 (C5), 113.5 (DMTr), 113.5 (DMTr), 110.8 (C5′′), 89.0 (DMTr), 85.3 (C1′), 84.1 (C4′), 81.0 (C2′), 75.9 (C3′), 63.8 (C5′), 55.3 (DMTr OCH_3), 47.1 (C6′), 11.9 (5′′-CH_3); HRMS-ESI m/z 812.3025 $[M+H]^+$; calcd ($C_{44}H_{42}N_7O_9^+$) 812.3039.

6-*N*-Benzoyl-9-(3′-*O*-(*P*-(2-cyanoethoxy)-*N*,*N*-diisopropylaminophosphinyl-5′-*O*-(4,4′-dimethoxytrityl)-2′-*C*-(thymin-1-yl)methyl-β-D-arabinofuranosyl)adenine (6)

Nucleoside **5** (88 mg, 0.11 mmol) was co-evaporated with DCE (2 × 5 mL) and dissolved in anhydrous CH_2Cl_2 (4 mL). 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.075 mL, 0.34 mmol) and DIPEA (0.10 mL, 0.57 mmol) were added, and the mixture was stirred for 21 h under N_2 atmosphere at rt. Further PCl-reagent (0.025 mL, 0.11 mmol) and DIPEA (0.07 ml, 0.40 mmol) were added, and the mixture was stirred for further 4 h at rt. EtOH (96%, 5 drops) was added and the mixture was concentrated under reduced pressure. The residue was purified by silica gel column chromatography (0–2.5% CH_3OH in CH_2Cl_2, v/v) to give phosphoramidite **6** (73 mg, 72 μmol, 67%) as a pale-yellow foam. ^{31}P NMR (162 MHz, $CDCl_3$) δ 151.6, 150.7; HRMS-ESI m/z 1012.4049 $[M+H]^+$; calcd ($C_{53}H_{59}N_9O_{10}P^+$) 1012.4117.

1-((2′-*C*-(Adenine-9-yl)methyl)-β-D-arabinofuranosyl)uracil (8)

Nucleoside **7** (0.50 g, 1.01 mmol) was dissolved in ~7 N methanolic ammonia (24 mL) and the solution was stirred at rt for 18 h. The solution was concentrated to half volume under reduced pressure at rt, and then filtrated. The residue was washed with acetone (2 × 10 mL) and dried under reduced pressure to give nucleoside **8** (0.26 g, 0.65 mmol, 65%) as a white powder. ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.38 (s, 1H, NH), 8.17 (s, 1H, H2′′), 8.05 (s, 1H, H8′′), 7.66 (d, J = 8.2 Hz, 1H, H6), 7.43 (bs, 2H, NH_2), 7.11 (d, J = 4.2 Hz, 1H, 3′OH), 6.03 (s, 1H, H1′), 5.93 (s, 1H, 2′OH), 5.65 (d, J = 8.2 Hz, 1H, H5), 5.31 (t, J = 5.1 Hz, 1H, 5′OH), 4.45 (d, J = 14.8 Hz, 1H, $H6_a'$), 4.18 (d, J = 14.8 Hz, 1H, $H6_b'$), 3.95 (dt, J = 5.1,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

1.8 Hz, 1H, H4′), 3.65 (dd, *J* = 4.2, 1.8 Hz, 1H, H3′), 3.60 (t, *J* = 5.1 Hz, 2H, H5′); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.2 (C4), 156.2 (C6′′), 151.8 (C2′′), 150.7 (C2), 149.2 (C4′′), 142.4 (C6), 142.1 (C8′′), 118.5 (C5′′), 100.6 (C5), 85.6 (C4′), 85.4 (C1′), 80.9 (C2′), 75.3 (C3′), 61.2 (C5′), 44.4 (C6′); HRMS-ESI *m*/*z* 392.1327 $[M+H]^+$; calcd ($C_{15}H_{18}N_7O_6^+$) 392.1313.

2′-(Adenin-9-yl)methyl-1′-(cytosin-1-yl)-1′-deoxy-β-D-arabinofuranoside-5′-triphosphate (U_ATP)

A suspension of **8** (76 mg, 0.19 mmol) in anhydrous trimethyl phosphate (2 mL) was stirred at 0 °C, and freshly distilled phosphoryl trichloride (46 μL, 0.49 mmol) was added. The mixture was stirred for 1.5 h at 0 °C, and then another portion of phosphoryl trichloride (46 μL, 0.49 mmol) was added. The mixture was stirred for further 30 min at 0 °C, and a 0.5 M solution of tris(tributylammonium) hydrogen pyrophosphate in anhydrous DMF (1.94 mL, 0.97 mmol) was added immediately followed by the addition of anhydrous tributylamine (0.23 mL, 0.97 mmol). The mixture was then stirred for 1.5 h at 0 °C. A 1.0 M aq. solution of triethylammonium bicarbonate (TEAB, pH 7.3) buffer (10 mL) was added, and the reaction mixture was stirred for 10 min. The resulting mixture was washed with EtOAc (10 mL), and the organic phase was extracted with 1.0 M TEAB buffer (10 mL). The combined aq. phase was concentrated under reduced pressure at rt. The residue was redissolved in a minimum of 0.05 M TEAB buffer and purified by flash column chromatography (0.05–1.0 M TEAB buffer, pH 7.3) at 4 °C using a DEAE Sephadex A25 anion-exchange column. Appropriate fractions were lyophilized twice to give crude **U_ATP** as a triethylammonium salt (44 mg, 47 μmol, 24%). A small portion of the product was further purified by reversed-phase HPLC (Sepax Bio-C18 column (PN: 106185–10025), 5 μm, 300 Å, 250 × 10 nm, 2–7% acetonitrile in 50 mM TEAA buffer, pH 6.7, 23 min) and appropriate fractions were lyophilized. The residue was dissolved in water (50 μL) and added dropwise to a solution of sodium perchlorate in acetone (3% w/w, 800 μL) upon which a precipitate formed. After centrifuging the suspension, the supernatant was removed and the solid was washed with cold acetone (2 × 600 μL) and dried. The resulting white pellet was dissolved in water (1000 μL) and treated with Chelex® 100 (sodium form) for 60 min at 0 °C. The Chelex® was filtered off and the filtrate was dried under reduced pressure to give analytically pure **U_ATP** as a white solid. ^{1}H NMR (500 MHz, D_2O) δ 8.31, 8.27 (2 × s, 2H, H2′′, H8′′), 7.94 (d, *J* = 8.1 Hz, 1H, H6), 6.18 (s, 1H, H1′), 5.83 (d, *J* = 8.1 Hz, 1H, H5), 4.68 (d, *J* = 15.1 Hz, 1H, H6′$_a$), 4.59 (d, *J* = 15.1 Hz, 1H, H6′$_b$), 4.34–4.22 (m, 4H, H3′, H4′, 2 × H5′); ^{31}P NMR (202 MHz, D_2O) δ –5.44 (d, *J* = 20.7 Hz, P_γ), –10.49 (d, *J* = 19.7 Hz, P_α), –21.43 (t, *J* = 20.1 Hz, P_β). HRMS-ESI *m*/*z* 630.0153 $[M–H]^–$; calcd ($C_{15}H_{19}N_7O_{15}P_3$) 630.0157.

Oligonucleotide synthesis: Oligonucleotide synthesis was carried out on an automated DNA synthesizer following the phosphoramidite approach. The synthesis of oligonucleotides was performed on a 0.2 μmol scale by using the amidite **6**, the previously published[10] phosphoramidites of **U_T** and **U_A** as well as the commercial 2-cyanoethyl phosphoramidites of the natural deoxynucleosides and the 6-FAM amidite. The synthesis followed the regular protocol for the DNA synthesizer. However, for modified amidites a prolonged coupling time (15 min) was used. 4,5-Dicyanoimidazole was used as the activator. The 5′-*O*-DMTr oligonucleotides were removed from the solid support by treatment with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

conc. aq. ammonia at rt (24 h) and purified by reversed-phase HPLC. All oligonucleotides were detritylated by treatment with 80% aq. acetic acid for 30 min. Aq. solutions of sodium acetate (3 M, 15 μL) and sodium perchlorate (5 M, 15 μL) were added, followed by acetone (1 mL). The pure oligonucleotides precipitated over night at −20 °C. After centrifugation at 12,000 rpm (10 min) at 4 °C, the supernatant was removed, and the pellet was washed with cold acetone (2 × 1 mL), dried (30 min) at 35 °C and dissolved in pure water (1000 μL). The concentration was determined by the optical density at 260 nm, the purity was confirmed by IC and the identity was confirmed by MALDI-TOF MS.

Thermal denaturation experiments: UV melting experiments were carried out on a PerkinElmer 35 UV/Vis spectrophotometer. The samples were comprised of a medium salt buffer containing 2.5 mM Na_2HPO_4, 5 mM NaH_2PO_4, 100 mM NaCl, and 0.1 mM EDTA at pH 7.0 and 1.5 μM of each strand. Before measuring, the oligonucleotides were annealed by heating the samples to 75 °C followed by controlled cooling to 10 °C. The absorbance at 260 nm was recorded as a function of the temperature (10–75 °C, 1 °C/min) by means of a PTP-6 Peltier Temperature Programmer. The melting temperatures (T_m) were determined as the maximum of the first derivatives of the absorbance vs. temperature curves. All determinations are averages of duplicates within 0.5 °C. For all duplexes, a clear melting temperature was found, and the process of denaturation was reversible upon cooling.

Primer extension reactions: A mixture of the template (0.5 μM, 2 μL) and the FAM-labelled primer (0.5 μM, 2 μL) in 1 μL of 10× reaction buffer containing 500 mM Tris–HCl (pH 7.2 at 25 °C), 100 mM $MgSO_4$ and 1 mM DTT (Kf−) or 200 mM Tris–HCl (pH 8.8 at 25 °C), 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$ and 1% Triton X-100 (DV−, *Taq*, Therminator) was heated to 90 °C for 2 min and then slowly cooled to rt. DNA polymerase (0.2–1.0 U·μL^{-1}, 1 μL) and TIPP (2 U·μL^{-1}, 1 μL, only for Therminator) were added to the mixture and it was heated to 37 °C (Kf−) or 75 °C (DV−, *Taq*, Therminator) for 2 min, after which a solution containing one or more of the natural dNTPs (100 μM each, 4 μL) or **U_ATP** (2.5 mM, 4 μL) was added. The mixture was then incubated at this temperature for 10 to 60 min and quenched by the addition of 20 μL stop solution (10 M urea, 50 mM EDTA, 0.1% BPB) followed by vigorous stirring. Finally, 3 μL of the resulting mixture was separated by 20% denaturing polyacrylamide gel electrophoresis (45 W, 60 °C, 120 min) in 1× TBE buffer and visualized using a fluoroimage analyzer (Amersham Typhoon).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by a grant from the Independent Research Fund Denmark | Natural Sciences (DFF – 7014-00052 to PN), and a grant from the National Institutes of Health (GM 069773 to YT).

Notes and references

1. Chaput JC and Herdewijn P, Angew. Chem. Int. Edit, 2019, 58, 11570–11572.

Author Manuscript

2. Wang J, Verbeure B, Luyten I, Lescrinier E, Froeyen M, Hendrix C, Rosemeyer H, Seela F, Van Aerschot A and Herdewijn P, J. Am. Chem. Soc, 2000, 122, 8595–8602.
3. Schöning KU, Scholz P, Guntha S, Wu X, Krishnamurthy R and Eschenmoser A, Science, 2000, 290, 1347–1351. [PubMed: 11082060]
4. Murayama K, Kashida H and Asanuma H, Chem. Commun, 2015, 51, 6500–6503.
5. Singh SK, Nielsen P, Koshkin AA and Wengel J, Chem. Commun, 1998, DOI: 10.1039/a708608c, 455–456.
6. Zhang L, Peritz A and Meggers E, J. Am. Chem. Soc, 2005, 127, 4174–4175. [PubMed: 15783191]
7. Morihiro K, Kasahara Y and Obika S, Mol. BioSyst, 2017, 13, 235–245. [PubMed: 27827481]
8. Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew S-Y, McLaughlin SH, Herdewijn P and Holliger P, Science, 2012, 336, 341–344. [PubMed: 22517858]
9. Sharma PK, Kumar P and Nielsen P, Aust. J. Chem, 2016, 69, 1094–1101.
10. Kumar P, Sharma PK and Nielsen P, J. Org. Chem, 2014, 79, 11534–11540. [PubMed: 25375974]
11. Hornum M, Sharma PK, Reslow-Jacobsen C, Kumar P, Petersen M and Nielsen P, Chem. Commun, 2017, 53, 9717–9720.
12. Hornum M, Stendevad J, Sharma PK, Kumar P, Nielsen RB, Petersen M and Nielsen P, Chem. Eur. J, 2019, 25, 7387–7395. [PubMed: 30942502]
13. Beck K, Reslow-Jacobsen C, Hornum M, Henriksen C and Nielsen P, Bioorg. Med. Chem. Lett, 2019, 29, 740–743. [PubMed: 30655212]
14. Zhu X-F, Williams HJ and Ian Scott A, Synthetic Commun., 2003, 33, 1233–1243.
15. Ohtsuka E, Ohkubo M, Yamane A and Ikehara M, Chem. Pharm. Bull, 1983, 31, 1910–1916. [PubMed: 6196133]
16. Buff R and Hunziker J, Bioorg. Med. Chem. Lett, 1998, 8, 521–524. [PubMed: 9871610]
17. Sano T, Shuto S, Inoue H and Ueda T, Chem. Pharm. Bull, 1985, 33, 3617–3622.
18. Ludwig J, Acta. Biochim. Biophys, 1981, 16, 131–133.
19. Diafa S, Evéquoz D, Leumann CJ and Hollenstein M, Chem. Asian J, 2017, 12, 1347–1352. [PubMed: 28371464]
20. Gardner AF, Jackson KM, Boyle MM, Buss JA, Potapov V, Gehring AM, Zatopek KM, Corrêa IR Jr., Ong JL and Jack WE, Front. Mol. Biosci, 2019, 6, Article 28.
21. Tomori T, Nagaoka K, Takeshita L, Shiozawa T, Miyatake Y, Masaki Y, Sekine M and Seio K, J. Org. Chem, 2018, 83, 8353–8363. [PubMed: 29952565]
22. Tsunoda H, Kudo T, Masaki Y, Ohkubo A, Seio K and Sekine M, Nucleic Acids Res, 2011, 39, 2995–3004. [PubMed: 21300642]
23. Jang M-Y, Song X-P, Froeyen M, Marlière P, Lescrinier E, Rozenski J and Herdewijn P, Chem. Biol, 2013, 20, 416–423. [PubMed: 23521798]
24. Kachare D, Song X-P and Herdewijn P, Bioorg. Med. Chem. Lett, 2014, 24, 2720–2723. [PubMed: 24813733]
25. Brown JA and Suo Z, Biochemistry, 2011, 50, 1135–1142. [PubMed: 21226515]
26. Lin Y, Qiu Q, Gill SC and Jayasena SD, Nucleic Acids Res, 1994, 22, 5229–5234. [PubMed: 7529404]
27. Lauridsen LH, Rothnagel JA and Veedu RN, ChemBioChem, 2012, 13, 19–25. [PubMed: 22162282]
28. Houlihan G, Arangundy-Franklin S, Porebski BT, Subramanian N, Taylor AI and Holliger P, Nat. Chem, 2020, 12, 683–690. [PubMed: 32690899]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

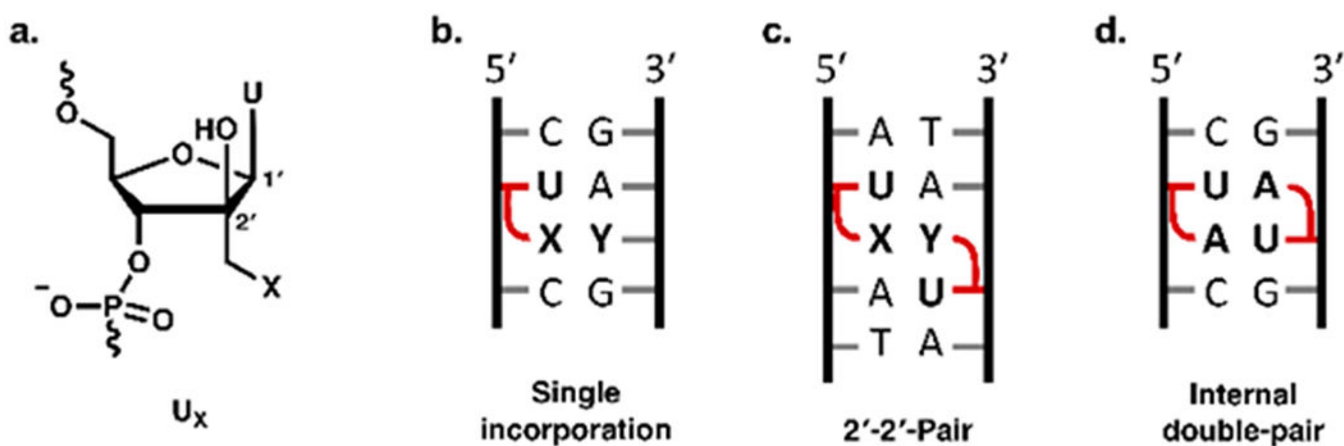


Fig. 1.

(a) General structure of the double-headed nucleotide $\mathbf{U_X}$ and (b-d) selected motifs. **U** = uracil-1-yl; **X**,**Y** = thymin-1-yl (T); adenin-9-yl (A); cytosin-1-yl (C) or guanin-9-yl (G); **Y** is complementary to **X** following Watson–Crick base-pairing.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

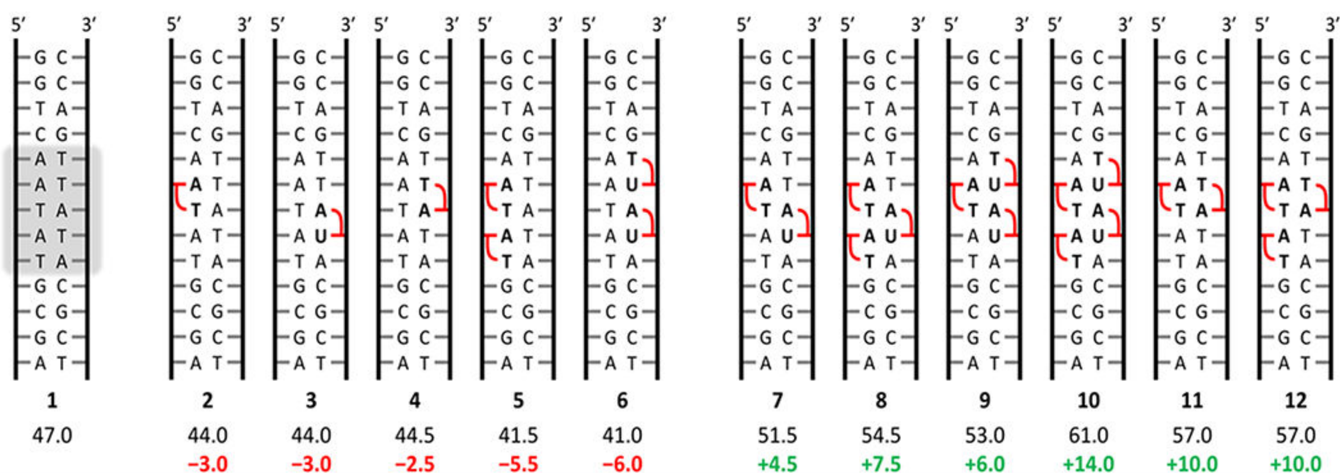


Fig. 2.

Structures of studied DNA duplexes, their melting temperatures (T_m, °C) and the difference in melting temperatures (T_m, °C) relative to the unmodified reference duplex 1.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

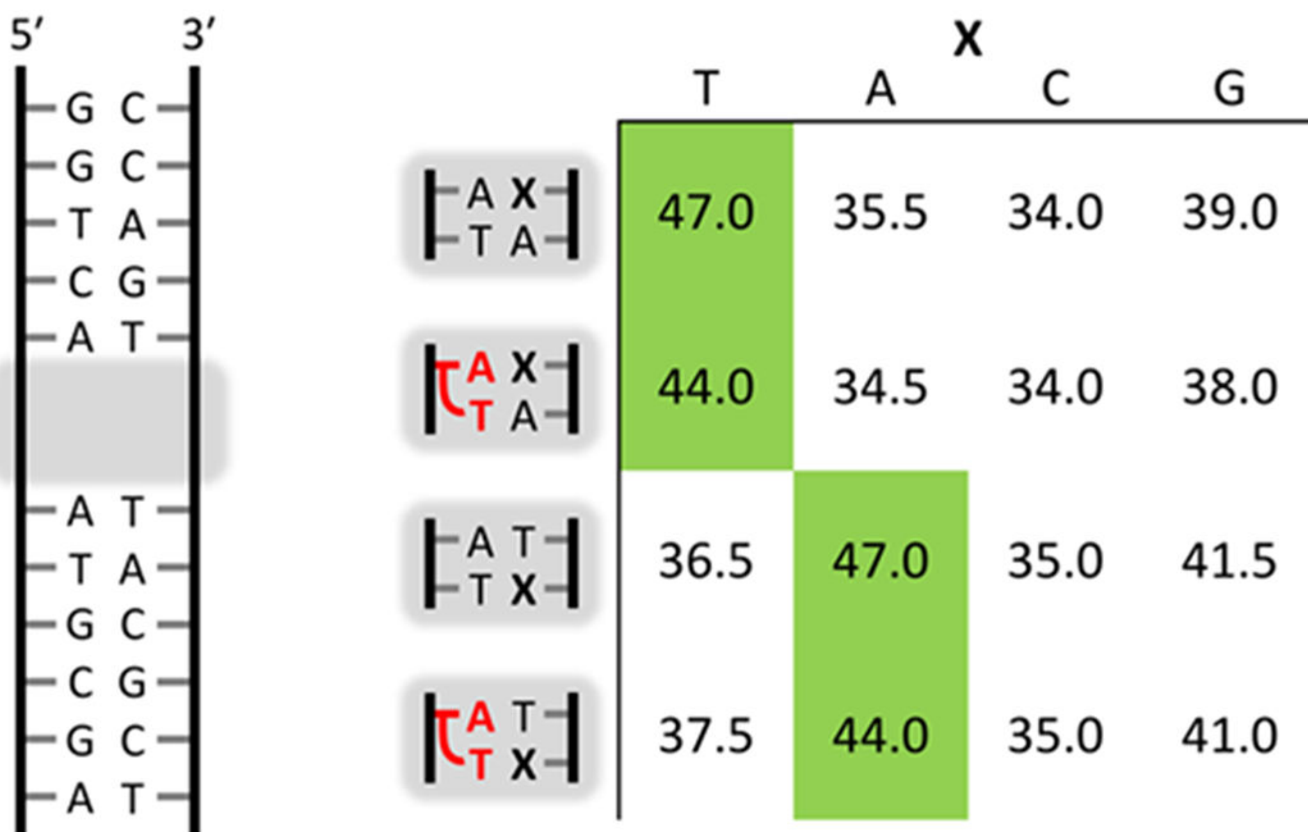


Fig. 3.

Structures and melting temperatures (°C) of duplexes used to study match and mismatch properties of the 1′-adenine and 2′-thymine of $\mathbf{A_T}$ (marked in red).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Templates

T1	$^{3'}$GCGCGCTTCTGGCCAATG ATGTCGCT$^{5'}$	26 nt
T2	$^{3'}$GCGCGCTTCTGGCCAATG $_{\mathbf{A}}$**U**GTCGCT$^{5'}$	26 nt

Primers

P1	FAM-$^{5'}$CGCGCGAAGACCGGT$^{3'}$	15 nt
P2	FAM-$^{5'}$CGCGCGAAGACCGGTT$^{3'}$	16 nt
P3	FAM-$^{5'}$CGCGCGAAGACCGGTTA$^{3'}$	17 nt
P4	FAM-$^{5'}$CGCGCGAAGACCGGTTAC$^{3'}$	18 nt
P5	FAM-$^{5'}$CGCGCGAAGACCGGTTAC T$^{3'}$	19 nt
P6	FAM-$^{5'}$CGCGCGAAGACCGGTTAC TA$^{3'}$	20 nt
P7	FAM-$^{5'}$CGCGCGAAGACCGGTTAC TACA$^{3'}$	22 nt

Expected full-length product

P8	FAM-$^{5'}$CGCGCGAAGACCGGTTAC TACAGCGA$^{3'}$	26 nt

Fig. 4.
Sequences of the used templates, primers and the expected full-length product.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

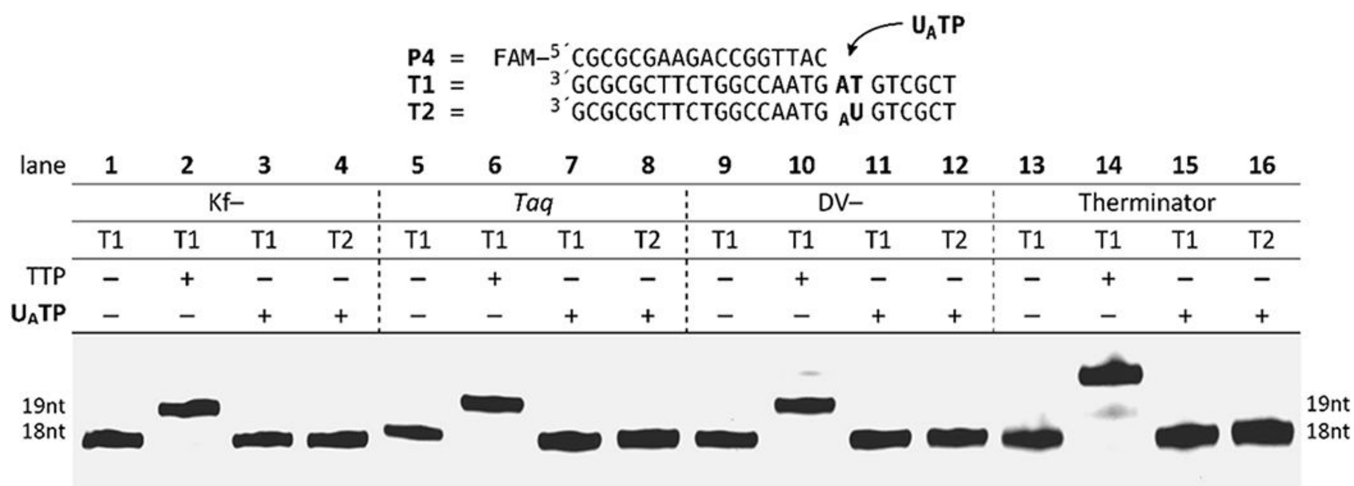


Fig. 5.

PAGE analysis of a single-nucleotide incorporation study using **U_ATP** as substrate for a DNA polymerase with primer **P4** and either template **T1** or **T2**. Conditions: [**P4**] = 0.1 μM; [template] = 0.1 μM; [TTP] = 40 μM; [**U_ATP**] = 1000 μM; [polymerase] = 0.05 U/μL (Kf−), 0.10 U/μL (*Taq*) or 0.02 U/μL (DV−, Therminator); for lanes 13–16 [TIPP] = 0.2 U/μL; temperature = 37 °C (Kf−) or 75 °C (*Taq*, DV−, Therminator); reaction time = 15 min.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

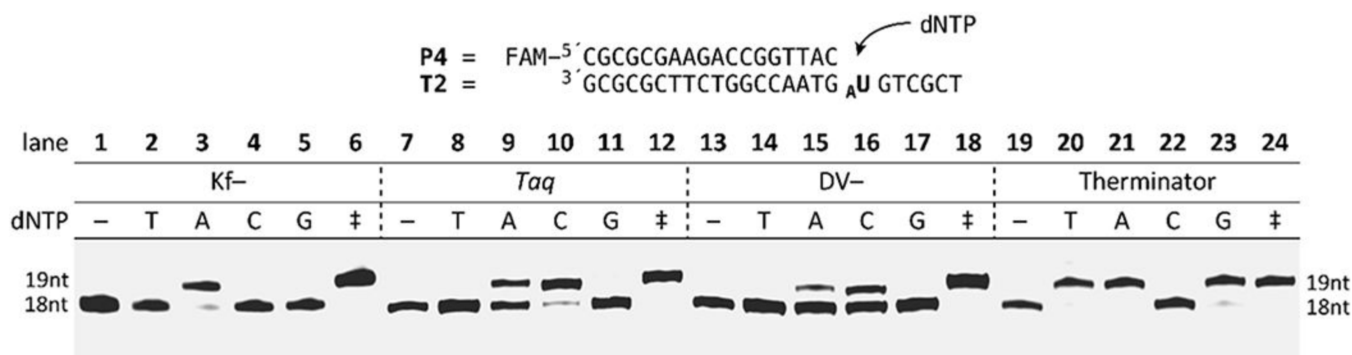


Fig. 6.

PAGE analysis of a single nucleotide incorporation study using the natural dNTPs with primer **P4** in the presence of template **T2** and a DNA polymerase. First and last lane for each enzyme correspond to the negative controls using the primers **P4** (–) and **P5** (‡), respectively. Conditions: [**P4**] = 0.1 μM; [**T2**] = 0.1 μM; [dNTP] = 40 μM; [polymerase] = 0.05 U/μL (Kf−), 0.10 U/μL (*Taq*) or 0.02 U/μL (DV−, Therminator); temperature = 37 °C (Kf−) or 75 °C (*Taq*, DV−, Therminator); reaction time = 15 min.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

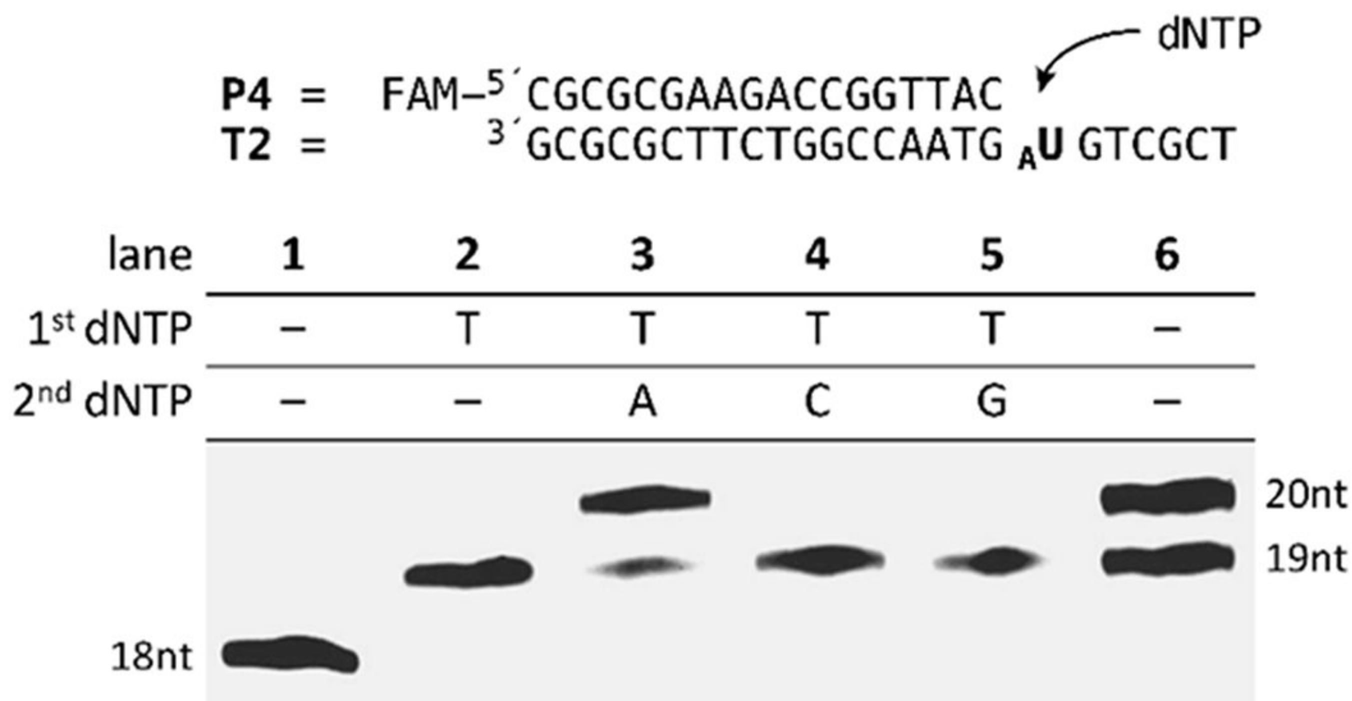


Fig. 7.

PAGE analysis of a double nucleotide incorporation study using the natural dNTPs with primer **P4** in the presence of template **T2** and Therminator. Lane 1 corresponds to the negative control and lane 6 corresponds to the 19 nt and 20 nt synthetic controls, **P5** and **P6**. Conditions: [**P4**] = 0.1 μM; [**T2**] = 0.1 μM; [dNTP] = 40 μM; [Therminator] = 0.02 U/μL; temperature = 75 °C; reaction time = 10 min for each addition.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

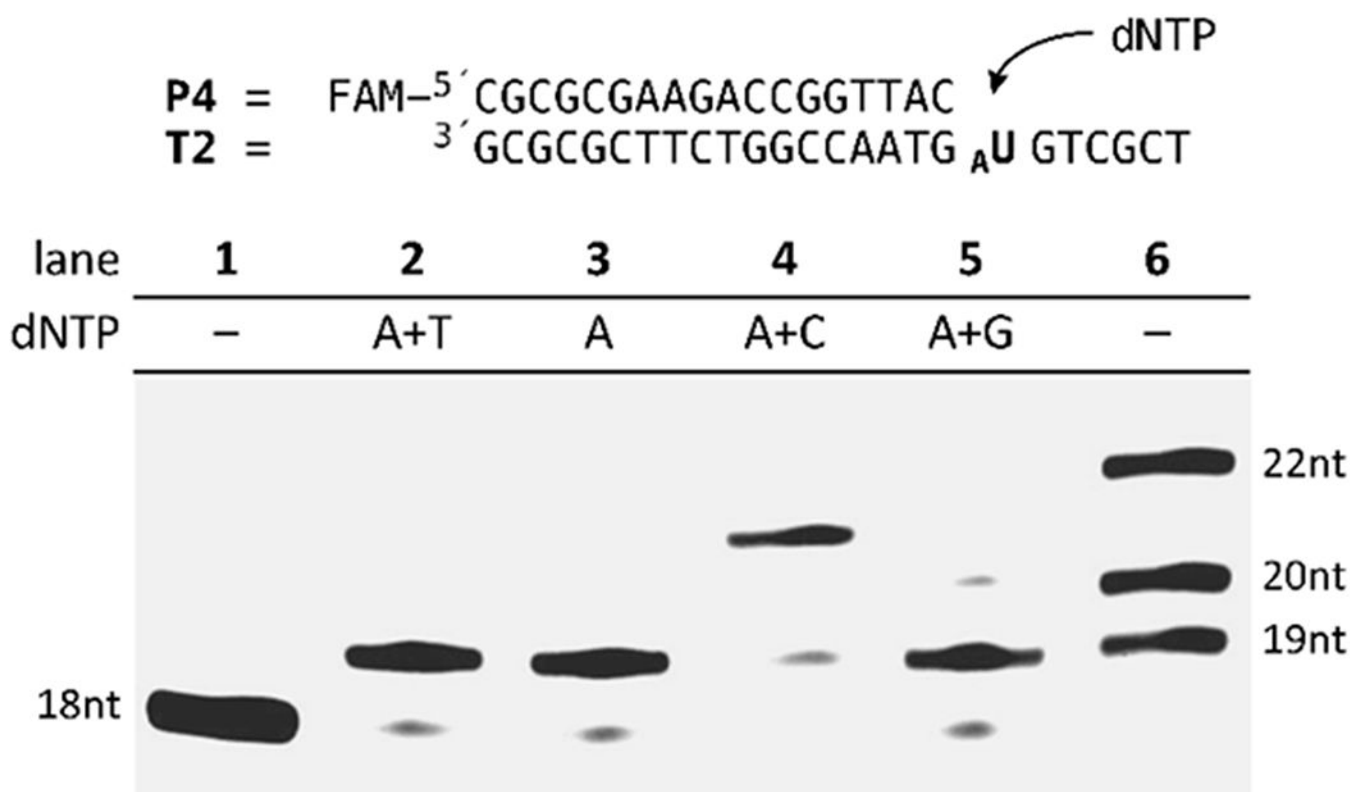


Fig. 8.

PAGE analysis of a multiple nucleotide incorporation study using the natural dNTPs with primer **P4** in the presence of template **T2** and Kf−. Lane 1 corresponds to the negative control and lane 6 corresponds to the 19 nt, 20 nt and 22 nt synthetic controls **P5**, **P6** and **P7**. Conditions: [**P4**] = 0.1 μM; [**T2**] = 0.1 μM; [dNTP] = 40 μM; [Kf−] = 0.05 U/μL; temperature = 37 °C; reaction time = 15 min.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

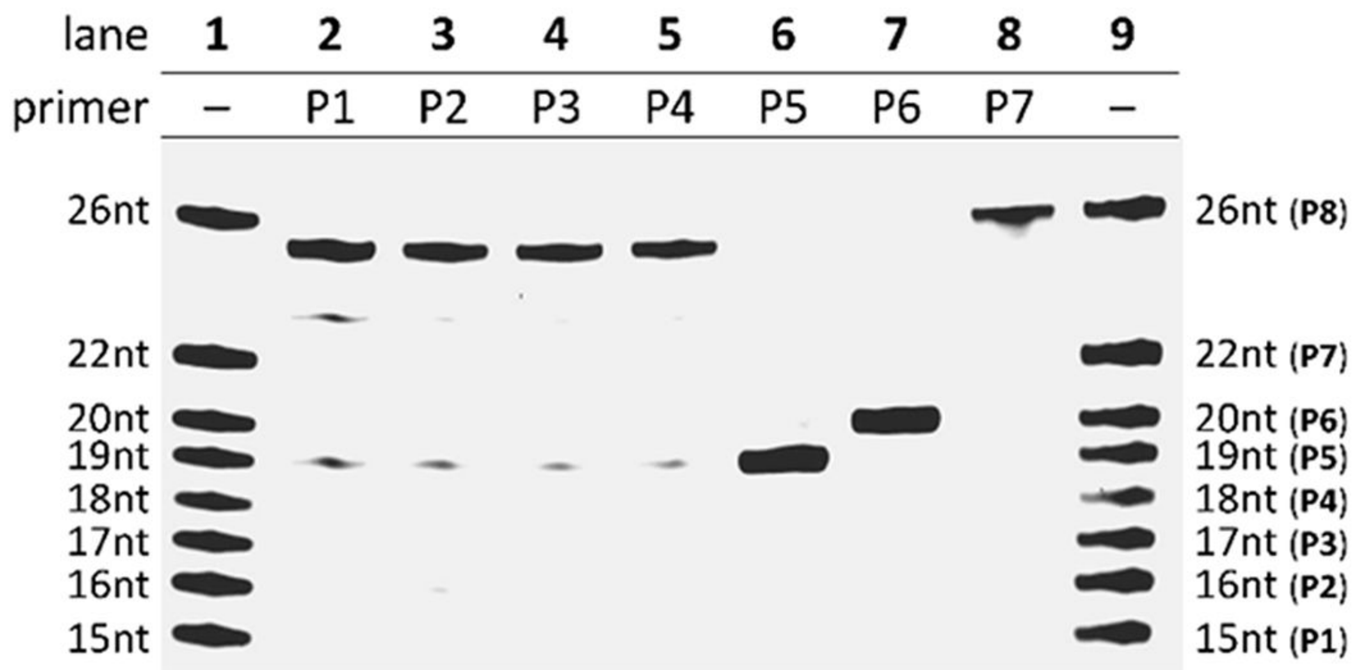


Fig. 9.

PAGE analysis of full-length primer extension reactions of **P1–P7** with natural dNTPs in the presence of **T2** and Kf−. Both lane 1 and 9 corresponds to the synthetic primers **P1–P7** and the 26 nt synthetic control **P8**. Conditions: [primer] = 0.1 μM; [**T2**] = 0.1 μM; [dNTP] = 40 μM; [Kf−] = 0.05 U/μL; temperature = 37 °C; reaction time = 60 min.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

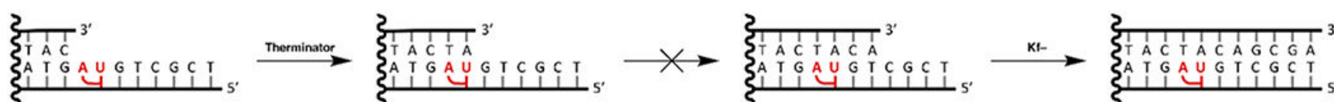
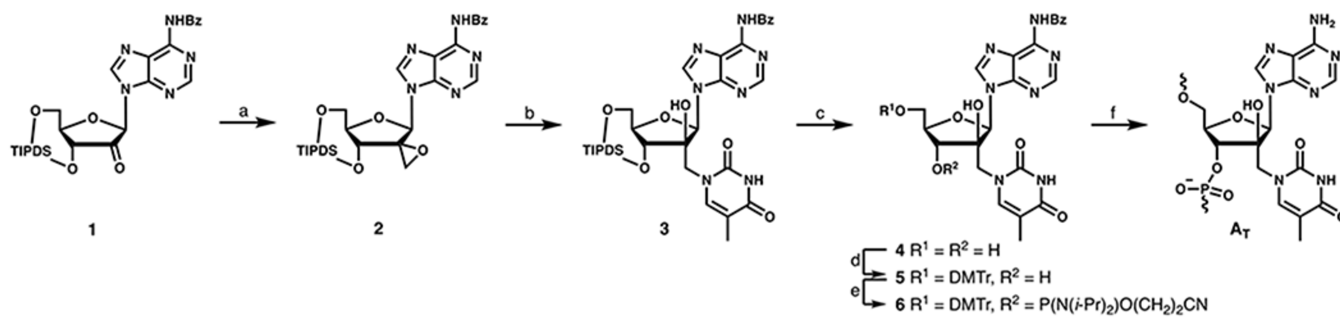


Fig. 10.

Overview of the primer extension reactions, demonstrating the ability of Therminator to correctly recognize $\mathbf{U_A}$ and extend the primer with the corresponding pT and pA, the currently lacking primer extension directly after the 5′-$\mathbf{U_A}$:3′-AT motif, and finally the ability of Kf– to fully extend the primer once it has been extended past the challenging motif.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Scheme 1.

Reagents and conditions: (a) NaH, $(CH_3)_3SOI$, THF, DMSO, 0 °C, 57%; (b) NaH, thymine, DMF, 65 °C, 46%; (c) $Et_3N·3HF$, THF, RT, 81%; (d) 4,4′-dimethoxytrityl chloride (DMTrCl), DMAP, pyridine, RT, 44%; (e) 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, $(i\text{-Pr})_2NEt$, DCE, RT, 62%; (f) incorporation into DNA oligonucleotides.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Scheme 2.

Reagents and conditions: (a) methanolic ammonia, RT, 65%; (b) i. $POCl_3$, $(MeO)_3PO$, 0–4 °C, ii. tributylammonium pyrophosphate, Bu_3N, DMF, 0–4 °C, 24%.