Bioorganic Chemistry 99 (2020) 103829

Contents lists available at ScienceDirect

Bioorganic Chemistry

journal homepage: www.elsevier.com/locate/bioorg

PCR incorporation of dUMPs modified with aromatic hydrocarbon substituents of different hydrophilicities: Synthesis of C5-modified dUTPs and PCR studies using Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases

Olga A. Zasedateleva*, Sergey A. Surzhikov, Valeriy E. Shershov, Rinat A. Miftakhov, Dmitry A. Yurasov, Viktoriya E. Kuznetsova, Alexander V. Chudinov

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov Street, 119991 Moscow, Russia

ARTICLE INFO

Keywords:
C5-modified dUTPs
Aromatic hydrocarbon groups
Hydrophilicity
PCR incorporation
Modified DNA

ABSTRACT

Deoxyuridine triphosphate derivatives (dUTPs) modified at the C5 position of the pyrimidine ring with various aromatic hydrocarbon substituents of different hydrophilicities have been synthesized. The aromatic hydrocarbon substituents were attached to dUTPs via a CH═CH−CH_2−NHCO−CH_2 linker. The efficiency of the PCR incorporation of modified dUMPs using Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases and a model DNA template containing one, two and three adjacent adenine nucleotides at three different sites within the sequence was investigated. For all the polymerases used, the yield of the modified PCR product was significantly increased with increasing hydrophilicity of the aromatic hydrocarbon substituent. In particular, for the above polymerases, the efficiency of the incorporation of dUMPs modified with the most hydrophilic of the studied aromatic hydrocarbon substituents, a 4-hydroxyphenyl residue, was 60–85% of the efficiency of dTMP incorporation. At the same time, the relative efficiencies of the incorporation of dUMPs modified with 2-, 4-methoxyphenyl, phenyl and 4-nitrophenyl substituents ranged from 20 to 50% and were 2–18% for the 1-naphthalene and 4-biphenyl groups, which were the most hydrophobic of the studied aromatic hydrocarbon substituents.

1. Introduction

The substrate specificities of different DNA polymerases in relation to nucleoside triphosphates modified with different functional groups are currently under intense investigation. Such studies are relevant, in particular, due to the importance of technologies used for the creation of modified aptamers. These aptamers useful for future medical applications are DNA oligonucleotides containing bases modified with different functional groups providing the ability for oligonucleotides to interact with target proteins with high specificity [1,2]. To date, protein biomarker discovery and clinical diagnostics company SomaLogic founded by L. Gold has developed modified aptamers for more than 1000 human proteins [3–6], with one of them, Macugen, available for purchase on the market [7]. Also, for a number of proteins, this company has developed pairs of modified aptamers whose dissociation constants are comparable to the dissociation constants of antigen-antibody complexes [6,8].

However, details of the technology for creating aptamers, including the corresponding nucleotide sequences, have not been published, and a number of laboratories are conducting independent researches aimed at finding and developing methods for creating specific and highly affine modified DNA aptamers [9–13]. A key feature of the technology for creating a modified DNA aptamer is the effective enzymatic incorporation of nucleotides modified by various functional groups in the nucleotide chain. Therefore, molecular mechanisms of enzymatic incorporation of modified nucleotides are also intensively investigated [14–16].

In 2005, Jäger et al. demonstrated the possibility of the effective enzymatic production of fully modified DNA products by B family polymerases (both in the primer extension reaction and in PCR) using all four nucleotides modified with basic, acid or lipophilic amino acid residues [9]. It has also been shown in this work that the efficiency of A family polymerases (Taq, Tth and KF (exo-)) in primer extension reactions depends significantly on the template DNA sequences.

* Corresponding author.
E-mail address: for_lana3@mail.ru (O.A. Zasedateleva).

https://doi.org/10.1016/j.bioorg.2020.103829
Received 24 December 2019; Received in revised form 25 March 2020; Accepted 5 April 2020
Available online 08 April 2020

In the above mentioned work [9], six different types of DNA templates were used: a DNA template with an almost equal nucleotide distribution without repeats of the same nucleotide except for one AA motif; four templates containing long A-, T-, G- and C-rich regions with two to four repetitions of the same nucleotide in adjacent positions; and a template containing a complete set of random sequences of 40 nucleotides (i.e., a complete set of sequences 40 nucleotides long, which is typically used for in vitro aptamer selection experiments). It was shown that Taq polymerase cannot produce a primer elongation product from an A-rich template or a template containing a complete set of random sequences of 40 nucleotides. At the same time, B-family polymerases (Pwo, Tgo, Pfu (exo-) and Vent (exo-)) produced modified primer extension products using all of the listed templates with efficiencies of more than 40% compared to the yields of primer extension products produced using only natural dNTPs. Pwo and Vent (exo-) polymerases produced modified DNA with efficiencies of more than 70%. In addition, Pwo polymerase can produce a fully modified PCR product using any of the listed templates. Meanwhile, G- and C-rich matrices were least effectively amplified by Pwo polymerase.

In 2006, Kuwahara et al. [10] also investigated the PCR amplification of plasmid DNA sites containing AAAA and GGGG poly motifs by polymerases from A and B families in the presence of dTTP and dCTP modified with aliphatic hydrocarbon substituents. In this work, it was found that polymerases from the B family (Vent (exo-), KOD Dash, and KOD (exo-)) more effectively incorporated dTMP and dCMP modified with a wider variety of substituents than polymerases from the A family (Taq and Tth).

As in [9], in 2010, Vaught et al. used a template containing a complete set of random 40 nucleotide-long sequences flanked by regions with fixed sequences for primer annealing and studied the enzymatic incorporation of dUMPs modified with various aromatic and aliphatic hydrocarbon substituents [3]. PCR in the presence of such modified nucleoside triphosphates did not lead to the formation of full-length products; however, full-length products containing modified dUMPs were efficiently produced in primer extension reactions by B family polymerases (KOD XL and Deep Vent (exo-)). The reverse reaction of primer extension with a modified sequence was subsequently performed in the presence of natural dNTPs followed by PCR. This scheme was used initially for selection of the modified aptamers for a member of the TNF receptor super family, TNFRSF9, and tumor-associated calcium signal transducer, TACSTD2, as protein targets [3]. Further the same approach was used to create aptamers for a wide range of proteins [4–6]. Dissociation constants for the most of protein-aptamer complexes were estimated to be lower than 1 nM. In all of the above studies using pyrimidine-based nucleoside triphosphates, substituents were introduced at the C5 position of the pyrimidine ring.

In the work of Kuwahara et al., pyrimidine nucleotides modified with substituents containing an amino group that was positively charged due to protonation were inefficiently incorporated by all the tested polymerases during PCR [10]. Recently, we have shown that Taq polymerase incorporates dUMPs modified by electroneutral hydrophilic (due to the zwitterionic structure) analogues of cyanine dyes more effectively, than dUMPs containing positively or negatively charged analogues of these dyes as substituents [17]. Based on that work, it followed that when creating dNTPs conjugated with dyes for the effective enzymatic labelling of DNA, the physical and chemical properties of the fluorophores, which can be considered as rather large substituents, must be taken into account. The introduction of zwitterionic groups into dyes improves their hydrophilicity, concomitantly improving the aqueous solubility of fluorescently labelled dNTPs and the enzymatic incorporation of nucleotides in DNA.

In the present work, by using Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases, we investigated the efficiency of the PCR incorporation of dUMPs modified via a linker at the C5 position of the pyrimidine ring with aromatic hydrocarbons. These substituents differed in hydrophilicity/hydrophobicity, and among them were those that mimic tyrosine and phenylalanine residues. For all the studied polymerases, the efficiency of PCR amplification increased with the increasing hydrophilicity of the aromatic hydrocarbon substituent. In this study, all the polymerases most effectively (with efficiencies equal to 60–85% of that of dTMP incorporation efficiency) incorporated dUMP modified with 4-hydroxyphenyl group, which mimics a tyrosine residue.

2. Materials and methods

2.1. dUTPs modified by different aromatic hydrocarbon substituents using $CH{=}CH{-}CH_2{-}NHCO{-}CH_2$ as a linker

Homogeneity of synthesized deoxynucleoside triphosphate derivatives and the reactions were monitored by TLC on TLC Silica gel 60 RP-18 F254 and TLC Silica gel 60 F254 plates (Merck, Darmstadt, Germany). The column chromatography was carried out using Lichroprep RP-18 sorbent (0.040–0.063 mm, Merck, Darmstadt, Germany). The compounds on chromatograms were detected at 280 nm with a EM-1 Econo UV Monitor (Bio-Rad Laboratories Inc., Hercules, CA, USA). High resolution mass spectra (HRMS) were registered on a Bruker Daltonics microTOF-Q II instrument (Bruker, Billerica, Germany) using electrospray ionization (ESI). The measurements were done in a negative and positive ion mode. The ^{1}H, ^{13}C and ^{31}P NMR spectra were recorded using a Bruker AMX-400 spectrometer (400 MHz, Bruker, Billerica, Germany) in DMSO-d_6 and D_2O. Chemical shifts (δ) were expressed in ppm.

All reagents were obtained from Sigma-Aldrich or Fluka unless otherwise stated.

The synthesis of activated aromatic acid esters and modified deoxyuridine triphosphates was carried out according to the synthetic procedures described earlier [18]. Detailed synthetic procedures and spectroscopic characteristics, including data on the intermediates, are provided in the Supplementary Information.

2.2. The DNA template and primers were similar to those used in [17]

Template
5′ – TCTCTTGCCCTTTCGTCTCTAAATTGTCTTAATCTCTTCTATCC
TTCTCTCTCACCACTTACATCCGC – 3′
Primer P1-Cy3
Cy3 – NH – 5′ – GCGGATGTAAGTGGTGAG – 3′
Primer P2-Cy3
Cy3- NH – 5′ – TCTCTTGCCCTTTCGTC – 3′
Primer sequences and the corresponding primer binding sites are underlined. Also underlined are adenines that are subject to complementary binding with modified deoxyuridines of the enzymatically synthesized DNA strand.

2.3. PCR and electrophoresis

PCR was performed following a protocol similar to one described earlier [17]. The above 68-nucleotide template and a pair of 18- and 17-nucleotide-long primers labelled with Cy3 dye were used. To study the efficiency of the complementary incorporation of modified dUMPs the template contained one, two and three adjacent adenine nucleotides at three different sites within the sequence. We used *Thermus aquaticus* (Taq) and *Thermus thermophilus* (Tth) DNA polymerases (Sileks, Badenweiler, Germany) and forms of the native DNA polymerases from Thermococcus litoralis (Vent (exo-)) and Pyrococcus GB-D strain (Deep Vent (exo-)) that were genetically engineered to eliminate the 3′→5′ proofreading exonuclease activity (New England Biolabs, Ipswich, UK).

As in the work of Kuwahara et al. [10], the PCR conditions were selected for each polymerase so that the full-length PCR products synthesized using only natural dNTPs were efficiently produced by all polymerases. Herein, the amount of PCR products was the same for all

the considered polymerases within an accuracy of 20%, which allowed the facile observation, registration and comparison of the electrophoregrams obtained using different polymerases.

The PCR conditions for each of the polymerases were as follows. All reaction mixtures (25 μl) for PCR amplification of the 68-nucleotide template contained each primer (10^{-6} M), the template (4×10^{-9} M), each deoxynucleoside triphosphate (dATP, dCTP, dGTP, dTTP or modified dUTP, 10^{-4} M), and the manufacturer's recommended buffer (70 mM Tris HCl, 16.6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, pH 8.6, for Taq and Tth polymerases, and 20 mM Tris HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton® X-100, pH 8.8, for Vent (exo-) and Deep Vent (exo-) polymerases). In the case of PCR amplification by Vent (exo-) polymerase, the reaction mixture additionally contained 8% formamide.

Tubes containing the reaction mixture without polymerase were placed into the heating block of a Peltier thermal cycler (Dyad from Bio-Rad, Hercules, CA, USA) and heated at 95 °C for 6 min. After that, the temperature was decreased to 62 °C, which is the calculated melting temperature of the P1-Cy3 and P2-Cy3 primers. The melting temperatures of the primers were calculated using the DI-nucleic acid hybridization and melting prediction web server [19]. Immediately after heating for 6 min at 62 °C, the appropriate amounts of polymerase (3 units of Taq polymerase, 1 unit of Tth polymerase, 0.5 units of Vent (exo-) polymerase or 0.08 units of Deep Vent (exo-) polymerase) were added to the reaction mixtures, which were then stirred without removing the tubes from the block. This excluded the possibility that the activity of the polymerases decreased during the pre-denaturation stage. Then, 20 reaction cycles of heating at 94 °C for 1 min and cooling to 62 °C for 1.5 min were conducted.

The reactions were terminated with 450 μl of 2.5% (for the case of amplification using Taq and Tth polymerases) or 4% (for the case of amplification using Vent (exo-) and Deep Vent (exo-) polymerases) lithium perchlorate trihydrate in 95% ethanol to form 90% of the final ethanol concentration followed by short-term mixing and 18 h of precipitation at −20 °C. After that the samples were centrifuged for 10 min at 13,000 rpm using Biofuge pico centrifuge (Kendro Laboratory Products GmbH, Hanau, Germany), the supernatants were removed and the samples were washed with 40 μl of 95% ethanol with lithium perchlorate trihydrate at the same corresponding concentrations at 20 °C for 5 min and 450 rpm using Thermomixer comfort (Eppendorf, Hamburg, Germany). Then the samples were additionally exposed for centrifugation for 4 min at 13,000 rpm using Biofuge pico centrifuge, the supernatants were removed and the samples were dried at 37 °C for 10 min using Termo 24–15 thermostat (Biocom, Moscow, Russia). Each sample was dissolved in 6 μl of water solution of 7 M Urea and 0.01% Bromophenol blue and heated at 95 °C for 50 s before being loaded into gel wells.

The PCR products were separated by electrophoresis in 18% polyacrylamide denaturing gels ((19:1 (w/w) acrylamide/bis-acrylamide, 7 M urea; 700 V; thermostabilized 16 × 16 cm glass sandwich with 1-mm gel thickness; TBE buffer (89 mM Tris-borate and 2 mM ethylenediaminetetraacetic acid, pH 8.3)) using PROTEAN II xi cell (Bio-Rad Laboratories Inc., Hercules, CA, USA) and an Elf-8 power supply (DNA Technology, Moscow, Russia). Samples of the PCR products obtained by replacing the dTTPs with modified dUTPs were loaded in descending order of hydrophilicity (increasing hydrophobicity) of the aromatic hydrocarbon substituent.

After electrophoresis, images of the gels were obtained in the Cy3 fluorescence range (λ_{max}^{abs} = 550 nm, λ_{max}^{em} = 570 nm, as described in the manufacturer's instructions for Amersham CyDye mono-reactive NHS Esters) using a gel imager for fluorescence spectroscopy with an image field of 20 × 16 cm, as described earlier [17]. The gel imager was equipped with an RTE/CCD-1536-K/1 CCD camera (Roper Scientific, Sarasota, FL, USA), a mercury lamp, a pair of 535DF35 and 580DF27 filters (Omega Optical, Brattleboro, VT, USA) and a computer running ImaGel Research software [17,20].

The relative quantities of the PCR-amplified full-length DNA fragments containing modified dUMPs were considered proportional to the relative fluorescence intensities of the Cy3-labelled primers. The relative Cy3 fluorescence intensities of the doublet bands that resulted from PCR amplification of the template using the forward and reverse primers labelled with Cy3 dye were quantified using virtual rectangular frames surrounding the bands. The fluorescence intensities of all the pixels within the frame were summed to obtain the total fluorescence intensity within the frame. The fluorescence intensity of the blank gel region within the same frame was then subtracted from the obtained value. The obtained quantities were normalized to the fluorescence intensity of the full-length PCR amplified DNA products obtained using the forward and reverse primers and only natural nucleotides. The relative quantities of PCR-amplified DNA containing modified dUMPs obtained in two experiments were averaged.

3. Results

3.1. Properties of the aromatic hydrocarbon substituents

To study the effect of the hydrophilicity of the aromatic hydrocarbon substituents, dUTPs modified at the C5 position of the pyrimidine ring by various aromatic hydrocarbon residues (R1–R7) were synthesized. The tested substituents included the residues that mimic natural tyrosine and phenylalanine amino acid residues as well as other structures that could potentially interact with the target proteins (Table 1). Following the approach developed by us in [18,21,22], aromatic hydrocarbon residues were added at the C5 position of the pyrimidine ring of dUTP via a CH═CH−CH_2−NHCO−CH_2 linker.

The hydrophilicity/hydrophobicity of substituents R1–R7 could be approximately characterized based on the water solubility values (S values) of the aromatic hydrocarbons, as their residues formed the corresponding substituents. The S values were taken from the PubChem (USA) database [23] and are presented in Table 1. These values range from S = $4.5*10^{-5}$ M (for diphenyl) to S = 0.88 M (for phenol). In addition, certain functional groups of some residues (R1–R7) may act as hydrogen bond donors (in the case of R1) or acceptors (in the cases of R1, R2, R3 and R5), which can influence the incorporation of the corresponding modified dUMP by each of the studied polymerases.

Thus, the selected aromatic hydrocarbon substituents presented and characterized in Table 1 allowed the investigation of the substrate specificity of the DNA polymerases for a wide range of hydrophilic/hydrophobic aromatic hydrocarbons conjugated at the C5 position of the pyrimidine ring of dUTPs via a linker.

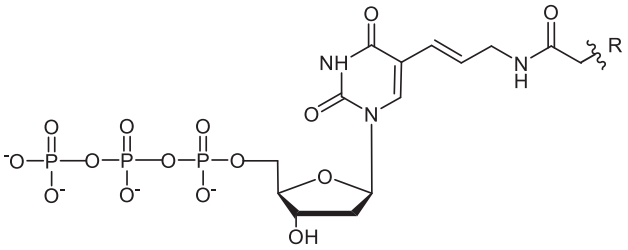
3.2. Efficiency of the PCR amplification of DNA strands by Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases using the C5-modified dUTPs

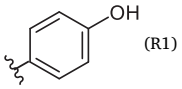
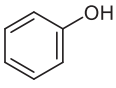
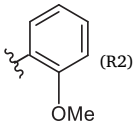
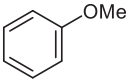
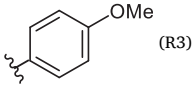
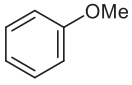
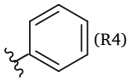
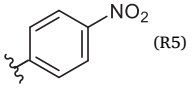
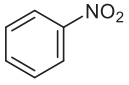
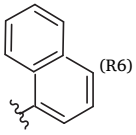
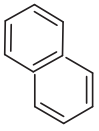
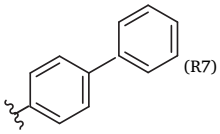
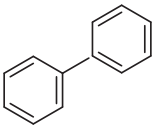
Fig. 1A and B shows the electrophoretic separation of the PCR products obtained with the Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases in the presence of the 68-nucleotide-long template and the pair of 18- and 17-nucleotide-long primers labelled with Cy3 dye (see the Materials and Methods section).

As shown in Fig. 1A and B, the full-length PCR products form two close electrophoretic fluorescent bands in the gels, since both the direct and reverse primers were labelled with Cy3 dye (see the Materials and Methods section). The exception is the full-length PCR-amplified strands in lanes 3, 12, 22 and 31, in which the PCR products of the forward and reverse primers overlap in the gel. This coelution can be attributed to the fact that in the presence of natural dNTPs, the "heavier" of the amplified DNA strands (the direct primer product) contains 10 dTMPs in its sequence and has a mass of 22,737 Da, and the "lighter" (the reverse primer product) has 29 dTMPs and a mass of 21,725 Da (lanes 2, 11, 21 and 30). Thus, the difference in the masses of these DNA strands is 1012 Da.

Since the primer regions of the amplified DNA strands contained

Table 1
The chemical structures of the modified dUMPs, the physicochemical properties of the aromatic hydrocarbon substituents, R1–R7, and the water solubilities of the aromatic hydrocarbon compounds forming the corresponding aromatic hydrocarbon substituents.



Substituent, R	Properties		
	Number of potential hydrogen bond donors	Number of potential hydrogen bond acceptors	Water solubility, S,[a] at 25 °C, of aromatic hydrocarbons the residues of which form the corresponding aromatic hydrocarbon substituents (M)
 (R1)	1	1	 Phenol 0.88
 (R2)	0	1	 Anisole 0.01
 (R3)	0	1	 Anisole 0.01
 (R4)	0	0	 Benzene 0.02
 (R5)	0	2	 Nitrobenzene 0.02
 (R6)	0	0	 Naphthalene 0.00024
 (R7)	0	0	 Biphenyl 0.000045

[a] Water solubilities of aromatic hydrocarbon compounds were taken from the PubChem (USA) database (23).

only natural nucleotides, it is easy to determine that in the case of PCR in the presence of dUTPs modified with a methylated linker (CH═CH−CH_2−NHCO−CH_2-Me) instead of dTTP, the product of the forward primer contains 6 modified dUMPs, and summarizing the other dNMPs presented in its' sequence, it has a mass of 23,319 Da. The product of the reverse primer contains 21 of the same dUMPs, has a mass of 23,762 Da and is the "heavier" of the amplified DNA strands. However, the difference in mass between these DNA strands is only 443 Da. These PCR products overlap in the gel and emit fluorescence as a single band (lanes 3, 12, 22 and 31).

Upon addition of an aromatic hydrocarbon residue (R1–R7) to the linker, the DNA strand possessing 21 modified dUMPs is significantly heavier than the DNA strand possessing 6 modified dUMPs. The mass difference for such PCR-amplified strands is 1375 Da (for the case of dUMPs modified via a linker with a phenyl residue (substituent R4, Table1), lanes 7, 16, 26 and 35 in Fig. 1A and B) or more (for other aromatic hydrocarbon residues). Therefore, in these cases, the full-length PCR products appear as two distinct fluorescent bands.

As shown in Fig. 1A and B, the PCR products that emitted the brightest fluorescence were those obtained in the presence of only natural dNTPs and those that possessed dUMPs that were substituted at the C5 position of the pyrimidine ring with a methylated linker or a linker to which the most hydrophilic of the aromatic hydrocarbon residues was attached, namely, a 4-hydroxyphenyl residue (R1) (lanes 2–4, 11–13, 21–23 and 30–32), instead of dTMPs. With increasing hydrophobicity of the aromatic hydrocarbon residue, the fluorescence, which is proportional to the amount of full-length PCR product, decreases.

The results of the quantitative analysis of the PCR products are presented in Table 2 and Fig. 1C. In the cases of Taq and Tth polymerases, the relative amounts of PCR-amplified products containing dUMPs modified with a 4-hydroxyphenyl residue (R1) were 80–85% of the amounts of product obtained in the presence of only natural dNTPs (lanes 4 and 13 in Fig. 1C). For Vent (exo-) and Deep Vent (exo-) polymerases, these amounts were 60–70% (lanes 23 and 32 in Fig. 1C). Thus, for Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases the relative efficiencies of PCR incorporation of dUMPs modified with a 4-hydroxyphenyl residue were 60–85%. The relative efficiencies of PCR

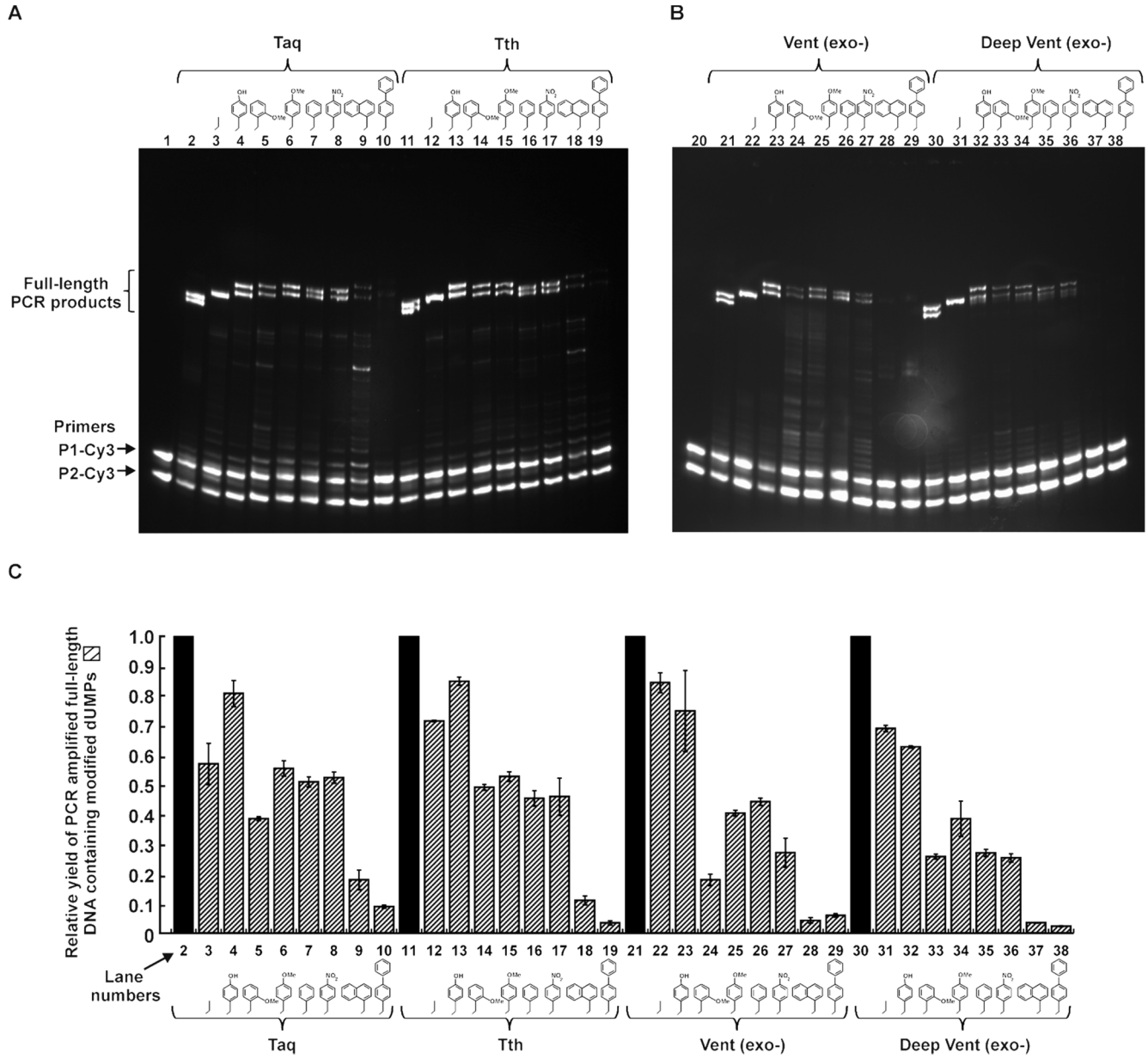


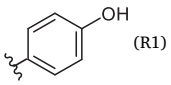
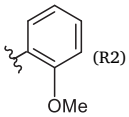
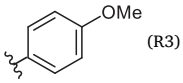
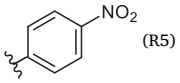
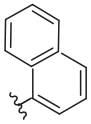
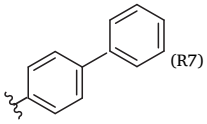
Fig. 1. Efficiency of PCR amplification by Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases when dTTP is replaced by dUTP modified via a CH═CH−CH_2−NHCO−CH_2 linker at the C5-position of the pyrimidine ring with aromatic hydrocarbon residues of different hydrophilicities. A and B. Electrophoretic separation of PCR products recorded in the fluorescence range of the Cy3 dye, which was used to label the primers (see Materials and Methods section). Lanes 1 and 20: PCR mixtures in the absence of polymerase. Lanes 2, 11, 21 and 30: PCR mixtures containing dNTPs and Taq, Tth, Vent (exo-) or Deep Vent (exo-) polymerase, respectively. Lanes 3, 12, 22 and 31: the same PCR mixtures as in lanes 2, 11, 21 and 30, respectively, but with the dTTPs replaced with dUTPs modified with the methylated linker CH═CH−CH_2−NHCO−CH_2-Me. Lanes 4–10, 13–19, 23–29 and 32–38: the same PCR mixtures as in lanes 2, 11, 21 and 30, respectively, but with dTTP replaced with dUTPs modified with residues R1-R7 using a CH═CH−CH_2−NHCO−CH_2 linker. C. Histogram of the fluorescence intensities of the full-length PCR products normalized for each polymerase to the fluorescence intensities of the PCR products produced in the presence of only natural dNTPs. The black columns are the fluorescent signals of the completely natural PCR products (taken as 1). The hatched columns are the normalized fluorescence signals of the PCR products containing modified dUMPs instead of dTMPs. The experiments shown in A and B were performed twice. The average values from the two experiments are shown. The bars indicate the absolute deviations.

incorporation of dUMPs conjugated with 2-, 4-methoxyphenyl, phenyl and 4-nitrophenyl residues (R2-R5) were 20–50%. The relative efficiencies of PCR incorporation of dUMPs modified by the most hydrophobic substituents, 1-naphthalene and 4-biphenyl residues (R6-R7), were 2–18%.

Thus, for all the polymerases studied, the amounts of PCR products that were obtained by replacing dTTP with modified dUTP decreased with decreasing hydrophilicity of the aromatic hydrocarbon substituent.

It should be mentioned that dUMPs conjugated via a linker with a 2-methoxyphenyl residue (R2) were incorporated by Taq, Vent (exo-) and Deep Vent (exo-) polymerases with an efficiency 1.5–2 times less than dUMPs conjugated with a 4-methoxyphenyl residue (R3). Therefore, the attachment of 4-methoxyphenyl residue is better for effective PCR amplification.

Table 2
The relative amounts of PCR products amplified by Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases when replacing dTTP with dUTPs modified with aromatic hydrocarbon substituents (R1–R7) via a CH═CH−CH_2−NHCO−CH_2 linker.

Substituent, R	Polymerase				
	Taq	Tth	Vent (exo-)	Deep Vent (exo-)	Polymerase averaged values
 (R1)	0.81 ± 0.04	0.85 ± 0.01	0.75 ± 0.14	0.63 ± 0.01	0.76 ± 0.13
 (R2)	0.39 ± 0.01	0.49 ± 0.01	0.18 ± 0.02	0.26 ± 0.01	0.33 ± 0.16
 (R3)	0.55 ± 0.03	0.53 ± 0.02	0.40 ± 0.01	0.39 ± 0.06	0.47 ± 0.08
 (R4)	0.51 ± 0.02	0.45 ± 0.03	0.44 ± 0.01	0.27 ± 0.01	0.42 ± 0.15
 (R5)	0.52 ± 0.02	0.46 ± 0.06	0.27 ± 0.05	0.25 ± 0.01	0.38 ± 0.14
 (R6)	0.18 ± 0.03	0.11 ± 0.01	0.04 ± 0.01	0.03 ± 0.01	0.09 ± 0.09
 (R7)	0.09 ± 0.01	0.03 ± 0.01	0.06 ± 0.01	0.02 ± 0.01	0.05 ± 0.04

3.3. Correlation between the yield of the PCR-amplified modified product and the hydrophilicity of the aromatic hydrocarbon substituent at the C5 position of dUMP

Fig. 2 shows the plot of the relative amounts of full-length PCR products, which were averaged over all four polymerases (Taq, Tth, Vent (exo-) and Deep Vent (exo-), see Table 2) against the hydrophilicity of the aromatic hydrocarbon substituents conjugated at the C5 position of the pyrimidine ring of the dUTPs. The hydrophilicities were approximated by the water solubility values, S values, of the aromatic hydrocarbons, the residues of which form the corresponding substituents (Table 1).

As shown in Fig. 2, the points in the obtained plot show a positive correlation between the relative amounts of full-length PCR products and the hydrophilicity of the aromatic hydrocarbon substituents. Thus, it can be concluded that when PCR is performed using Taq, Tth, Vent (exo-) or Deep Vent (exo-) polymerase, the observed efficiency of the incorporation of dUMPs modified at the C5 position of the pyrimidine ring increases with increasing hydrophilicity of the aromatic hydrocarbon substituent.

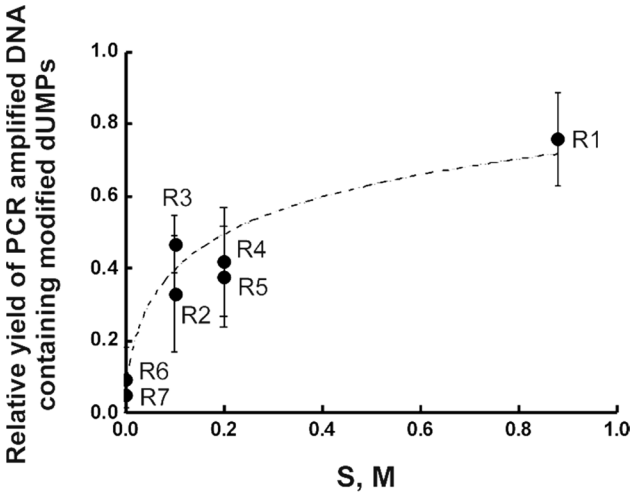


Fig. 2. Plot of the relative amounts of enzymatically produced full-length modified PCR products averaged over Taq, Tth, Vent (exo-) and Deep Vent (exo-) polymerases as a function of the hydrophilicity of the aromatic hydrocarbon substituents (R1–R7) conjugated at the C5 position of the dUMPs via a CH═CH−CH_2−NHCO−CH_2 linker. Black circles – plot of the averaged relative amounts of full-length PCR-amplified DNA strands (see Table 2) against the water solubility values, S values (M) (Table 1). The bars indicate the maximum deviations from the average. The dashed line indicates the approximation of the "black circles" using the least squares method and the following logarithmic equation: $y = A \cdot \log(B \cdot S + C) + D$, where y is the averaged relative efficiency of PCR incorporation, and A, B, C and D are the parameters. The coefficient of determination of the logarithmic approximation, R^2, was 0.91.

4. Discussion

Hydrogen bonds, electrostatic, van der Waals and hydrophobic interactions form the basis of protein-protein, protein-DNA and ligand-DNA intermolecular interactions [24–28]. The dissociation constants for the sequence-specific complexes between transcription factors and corresponding DNA sites are in the range of 10^{-6}–10^{-9} M [29–31] and for antigen-antibody complexes these values lie in the range 10^{-9}–10^{-12} M [32]. As reported by L. Gold et al. [4,5] using SELEX technology [33–35], it was possible to create more than 1000 DNA oligonucleotides (or DNA aptamers) containing bases modified by aromatic and aliphatic functional groups with a relatively high binding affinity for target proteins (with K_d values 10^{-9}–10^{-12} M for the most aptamer-protein complexes). Therefore, it is promising to obtain modified DNA aptamers exhibiting high affinity to disease-related target

proteins for the diagnostics and therapy of infectious, oncological, neurodegenerative and cardiovascular diseases [1,2].

One of the main stages of SELEX technology used in the search for aptamers is enzymatic synthesis, i.e. primer extension or PCR [33–35]. Therefore, modified DNA template and modified dNTPs should be compatible with the enzymatic synthesis. To meet this requirement, various DNA polymerases were previously tested to select the most efficient ones, depending on certain modifications of the bases of dNTPs [3,9,12]. Another approach is based on the search for structural parameters characterizing as linkers and attached functional groups, which would be more compatible with the most of DNA polymerases [10].

4.1. Efficiency of PCR in the presence of modified dUTPs depending on the hydrophilicity of aromatic hydrocarbon substituents

Previous publications have included studies of basic, acidic, or lipophilic groups [3,9,10], but their hydrophilicity has not been considered separately or systematically studied.

In the present study, using DNA polymerases from the A (Taq and Tth) and B (Vent (exo-) and Deep Vent (exo-)) families, similar efficiencies of PCR incorporation of dUMPs modified by an aromatic hydrocarbon substituent were obtained. Moreover, it was shown for all the studied polymerases, that the relative amount of full-length modified PCR product obtained was increased with increasing hydrophilicity of the aromatic hydrocarbon substituent. This correlation, common to the four polymerases, appears to be a consequence of the common basic structure of DNA polymerases consisted of subdomains referred to as fingers, palm, and thumb domains which resemble a right hand [28,36] and also a consequence of the presence of similar parts of the DNA template and the synthesized DNA chain in the active centers of DNA polymerases.

It should be mentioned, that the increase of PCR efficiency depending on hydrophilicity of substituent at the C5 position of dUTPs also correlates with the previous observation that the dUMPs which are labeled at the same position by hydrophilic zwitterionic analogs of cyanine dyes are efficiently incorporated by Taq DNA polymerase [17].

4.2. PCR-compatibility of substituents attached at the C5 position of dUTPs, depending on the linker structure

In the previous studies, Vaught et al. [3] and Gold et al. [4] using a CONH-CH_2 linker, demonstrated the ability to effectively incorporate dUTPs modified with aromatic and aliphatic hydrocarbon substituents in primer extension reactions, but not in PCR. A distinctive feature of the approach developed by Jäger et al. [9], which is the most successful approach in terms of the efficiency of the PCR incorporation of modified dUMPs, was the use of a C≡C−CH_2−NHCO−CH_2 linker containing a C≡C bond and an “inverted” peptide group. Thus, the effective PCR incorporation of modified dUMPs is strongly influenced by the linker structure and is facilitated by the presence of multiple bonds between carbon atoms (in this case, a triple bond) in the linker as well as by the presence of the “inverted” peptide group (NHCO).

Most recently, in 2019, Lapa et al. used dUTPs modified with aliphatic or aromatic hydrocarbon substituents via a CH═CH−CH_2−NHCO−CH_2 linker containing a C═C bond and an “inverted” peptide group (NHCO) [22]. It was demonstrated that the PCR amplification of full-length products by polymerases of both A and B families, Taq and Vent (exo-), using the above mentioned modified dUTPs and either a bacterial DNA or a template containing a complete set of random 40 nucleotide-long sequences, differs in efficiency by no more than a factor of two. Apparently, as follows from the above discussion of previous works, the presence of a carbon-carbon double bond and an “inverted” peptide group (NHCO) in the linker structure is crucial for the effective amplification of the modified PCR product by Taq polymerase.

In the present study, as in [22], the CH═CH−CH_2−NHCO−CH_2 linker was used to modify dUTPs by seven various aromatic hydrocarbon substutuents with rather different hydrophilicities. Using DNA polymerases from the A (Taq and Tth) and B (Vent (exo-) and Deep Vent (exo-)) families, we obtained quite similar efficiencies (with the difference of about 20–50% for the most of substituents) of PCR incorporation of dUMPs modified by one and the same aromatic hydrocarbon substituent. Consequently, this linker structure contributes to the effective incorporation of modified nucleotides by a broad range of DNA polymerases.

4.3. Complete set of random sequences of a fixed length or only those without homonucleotide repeats longer than 3 adjacent nucleotides?

In the present work, the incorporation of a relatively small number (no more than three) of adjacent modified dUMPs in the PCR-amplified strands apparently contributed to the effective PCR amplification of full-length modified products.

As was expected on initial stages of SELEX development the necessary step in aptamer production was enzymatic synthesis using a template containing a complete set of random sequences of a fixed length including those containing homonucleotide repeats [33–35]. Accordingly, the previous investigations were partly focused on finding combinations of polymerases and modified dNTPs that would contribute to enzymatic synthesis with minimal dependence on DNA sequences [3,9,10].

However, it was shown later that the relative enzymatic incorporation efficiency decreases with an increase in the number of consequently incorporated modified nucleotides of one type [10,17]. Therefore, the sequences containing homonucleotide repeats are enzymatically synthesized less efficiently than those without such sites. At the same time, Taq polymerase seems to amplify an array of random DNA sequences of a fixed length including those containing homonucleotide repeats of a modified nucleotide no longer than a triplet. Some of these oligonucleotides may with a certain probability contain motifs capable of highly specific interactions with the target proteins.

Therefore, it can be assumed that in order to obtain aptamers, the enzymatic synthesis of modified DNA fragments using dUTPs modified with certain aromatic hydrocarbon substituents can be carried out by PCR using Taq polymerase, which allows the incorporation of up to three adjacent modified nucleotides.

5. Conclusions

When using DNA templates containing no more than three adjacent adenine nucleotides:

(i) In general, Taq, Tth, Vent (exo-) and Deep Vent (exo-) DNA polymerases exhibit the same capability for PCR incorporation of dUMPs modified via a CH═CH−CH_2−NHCO−CH_2 linker at the C5 position of the pyrimidine ring with an aromatic hydrocarbon substituent;
(ii) The efficiency of incorporation by the above mentioned polymerases increases with increasing hydrophilicity of the aromatic hydrocarbon substituent.

Also, it is of practical significance that Taq polymerase can be used for the effective PCR amplification of modified DNA strands containing no more than three adjacent dUMPs modified at the C5 position of the pyrimidine ring with residues of aromatic hydrocarbons having hydrophilicities comparable to or greater than that of anisole.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Dr. Mikhailovich V.M. (EIMB RAS) for providing useful advice regarding the PCR conditions. We thank American Journal Experts (Durham, NC, USA) for providing language editing services.

Funding

This study was supported by the Russian Science Foundation [grant number 18-29-09151].

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bioorg.2020.103829.

References

[1] R.B. Ladju, D. Pascut, M.N. Massi, C. Tiribelli, C.H. Sukowati, Aptamer: A potential oligonucleotide nanomedicine in the diagnosis and treatment of hepatocellular carcinoma, Oncotarget 9 (2017) 2951–2961, https://doi.org/10.18632/oncotarget.23359.

[2] G. Zhu, X. Chen, Aptamer-based targeted therapy, Adv. Drug. Deliv. Rev. 134 (2018) 65–78, https://doi.org/10.1016/j.addr.2018.08.005.

[3] J.D. Vaught, C. Bock, J. Carter, T. Fitzwater, M. Otis, D. Schneider, J. Rolando, S. Waugh, S.K. Wilcox, B.E. Eaton, Expanding the chemistry of DNA for in vitro selection, J. Am. Chem. Soc. 132 (2010) 4141–4151, https://doi.org/10.1021/ja908035g.

[4] L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E.N. Brody, J. Carter, A.B. Dalby, B.E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T.R. Keeney, N. Kim, T.H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S.K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R.M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D.I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J.D. Vaught, M. Vrkljan, J.J. Walker, M. Watrobka, S. Waugh, A. Weiss, S.K. Wilcox, A. Wolfson, S.K. Wolk, C. Zhang, D. Zichi, Aptamer-based multiplexed proteomic technology for biomarker discovery, PLoS One 5 (2010) e15004, https://doi.org/10.1371/journal.pone.0015004.

[5] L. Gold, J.J. Walker, S.K. Wilcox, S. Williams, Advances in human proteomics at high scale with the SOMAscan proteomics platform, N. Biotechnol. 29 (2012) 543–549, https://doi.org/10.1016/j.nbt.2011.11.016.

[6] B.N. Gawande, J.C. Rohloff, J.D. Carter, I. von Carlowitz, C. Zhang, D.J. Schneider, N. Janjic, Selection of DNA aptamers with two modified bases, Proc. Natl. Acad. Sci. 114 (2017) 2898–2903, https://doi.org/10.1073/pnas.1615475114.

[7] D.W. Drolet, L.S. Green, L. Gold, N. Janjic, Fit for the eye: aptamers in ocular disorders, Nucl. Acid Ther. 26 (2016) 127–146, https://doi.org/10.1089/nat.2015.0573.

[8] U.A. Ochsner, L.S. Green, L. Gold, N. Janjic, Systematic selection of modified aptamer pairs for diagnostic sandwich assays, Biotechniques 56 (2014) 125–133, https://doi.org/10.2144/000114134.

[9] S. Jäger, G. Rasched, H. Kornreich-Leshem, M. Engeser, O. Thum, M. Famulok, A versatile toolbox for variable DNA functionalization at high density, J. Am. Chem. Soc. 127 (2005) 15071–15082, https://doi.org/10.1021/ja051725b.

[10] M. Kuwahara, J. Nagashima, M. Hasegawa, T. Tamura, R. Kitagata, K. Hanawa, S. Hososhima, T. Kasamatsu, H. Ozaki, H. Sawai, Systematic characterization of 2'-deoxynucleoside-5'-triphosphate analogs as substrates for DNA polymerases by polymerase chain reaction and kinetic studies on enzymatic production of modified DNA, Nucl. Acids Res. 34 (2006) 5383–5394, https://doi.org/10.1093/nar/gkl637.

[11] M. Famulok, G. Mayer, Aptamers and SELEX in chemistry & biology, Chem. Biol. 21 (2014) 1055–1058, https://doi.org/10.1016/j.chembiol.2014.08.003.

[12] M. Hollenstein, C.J. Leumann, Synthesis and biochemical characterization of tricyclothymidine triphosphate (tc-TTP), ChemBioChem 15 (2014) 1901–1904, https://doi.org/10.1002/cbic.201402116.

[13] M. Hocek, Synthesis of base-modified 2′-deoxyribonucleoside triphosphates and their use in enzymatic synthesis of modified DNA for applications in bioanalysis and chemical biology, J. Org. Chem. 79 (2014) 9914–9921, https://doi.org/10.1021/jo5020799.

[14] S. Obeid, A. Baccaro, W. Welte, K. Diederichs, A. Marx, Structural basis for the synthesis of nucleobase modified DNA by Thermus aquaticus DNA polymerase, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 21327–21331, https://doi.org/10.1073/pnas.1013804107.

[15] K. Bergen, A.L. Steck, S. Strütt, A. Baccaro, W. Welte, K. Diederichs, A. Marx, Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates, J. Am. Chem. Soc. 134 (2012) 11840–11843, https://doi.org/10.1021/ja3017889.

[16] H.M. Kropp, S.L. Dürr, C. Peter, K. Diederichs, A. Marx, Snapshots of a modified nucleotide moving through the confines of a DNA polymerase, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) 9992–9997, https://doi.org/10.1073/pnas.1811518115.

[17] O.A. Zasedateleva, V.A. Vasiliskov, S.A. Surzhikov, V.E. Kuznetsova, V.E. Shershov, T.O. Guseinov, I.P. Smirnov, R.A. Yurasov, M.A. Spitsyn, A.V. Chudinov, dUTPs conjugated with zwitterionic Cy3 or Cy5 fluorophore analogues are effective substrates for DNA amplification and labelling by Taq polymerase, Nucl. Acids Res. 46 (2018) e73, https://doi.org/10.1093/nar/gky247.

[18] A.V. Chudinov, Y.Y. Kiseleva, V.E. Kuznetsov, V.E. Shershov, M.A. Spitsyn, T.O. Guseinov, S.A. Lapa, E.N. Timofeev, A.I. Archakov, A.V. Lisitsa, S.P. Radko, A.S. Zasedatelev, Structural and functional analysis of biopolymers and their complexes: Enzymatic synthesis of high-modified DNA, Mol. Biol. (Mosk) 51 (2017) 474–482, https://doi.org/10.1134/S0026893317030025.

[19] M. Zuker, N. Markham, http://unafold.rna.albany.edu.

[20] O.A. Zasedateleva, A.L. Mikheikin, A.Y. Turygin, D.V. Prokopenko, A.V. Chudinov, E.E. Belobritskaya, V.R. Chechetkin, A.S. Zasedatelev, Gel-based oligonucleotide microarray approach to analyze protein–ssDNA binding specificity, Nucl. Acids Res. 36 (2008) e61, , https://doi.org/10.1093/nar/gkn246.

[21] V.A. Vasiliskov, S.A. Lapa, V.E. Kuznetsova, S.A. Surzhikov, V.E. Shershov, M.A. Spitsyn, T.O. Guseinov, R.A. Miftahov, O.A. Zasedateleva, A.V. Lisitsa, S.P. Radko, A.S. Zasedatelev, E.N. Timofeev, A.V. Chudinov, Novel 5-alkylcarboxamide-2'-deoxyuridine-5'-triphosphates for enzymatic synthesis of highly modified DNA, Russ. J. Bioorg. Chem. 45 (2019) 221–223, https://doi.org/10.1134/S1068162019030063.

[22] S.A. Lapa, A.S. Pavlov, V.E. Kuznetsova, V.E. Shershov, M.A. Spitsyn, T.O. Guseinov, S.P. Radko, A.S. Zasedatelev, A.V. Lisitsa, A.V. Chudinov, Enzymatic preparation of modified DNA: study of the kinetics by real-time PCR, Mol. Biol. (Mosk) 53 (2019) 460–469, https://doi.org/10.1134/S0026893319030099.

[23] https://pubchem.ncbi.nlm.nih.gov/.

[24] A.S. Krylov, S.L. Grokhovsky, A.S. Zasedatelev, A.L. Zhuze, G.V. Gursky, B.P. Gottikh, Quantitative estimation of the contribution of pyrrolcarboxamide groups of the antibiotic distamycin A into specificity of its binding to DNA AT pairs, Nucl. Acids Res. 6 (1979) 289–304, https://doi.org/10.1093/nar/6.1.289.

[25] G.V. Gursky, A.S. Zasedatelev, A.L. Zhuze, A.A. Khorlin, S.L. Grokhovsky, S.A. Streltsov, A.N. Surovaya, S.M. Nikitin, A.S. Krylov, V.O. Retchinsky, M.V. Mikhailov, R.S. Beabealashvili, B.P. Gottikh, Synthetic sequence-specific ligands, Cold Spring Harb. Symp. Quant. Biol. 47 (1983) 367–378, https://doi.org/10.1101/sqb.1983.047.01.043.

[26] R.S. Richard, DNA Structure and Function. Elsevier Inc., 1994. https://doi.org/10.1016/C2009-0-02451-9.

[27] A.L. Mikheikin, A.L. Zhuze, A.S. Zasedatelev, Molecular modelling of ligand-DNA minor groove binding: role of ligand-water interactions, J. Biomol. Struct. Dyn. 19 (2001) 175–178, https://doi.org/10.1080/07391102.2001.10506729.

[28] J.M. Walsh, P.J. Beuning, Synthetic nucleotides as probes of DNA polymerase specificity, J. Nucl. Acids 2012 (2012), https://doi.org/10.1155/2012/530963 ID 530963.

[29] S. Akashi, R. Osawa, Y. Nishimura, Evaluation of protein-DNA binding affinity by electrospray ionization mass spectrometry, J. Am. Soc. Mass Spectrom. 16 (2005) 116–125, https://doi.org/10.1016/j.jasms.2004.09.021.

[30] R. Joshi, J.M. Passner, R. Rohs, R. Jain, A. Sosinsky, M.A. Crickmore, V. Jacob, A.K. Aggarwal, B. Honig, R.S. Mann, Functional specificity of a Hox protein mediated by the recognition of minor groove structure, Cell 131 (2007) 530–543, https://doi.org/10.1016/j.cell.2007.09.024.

[31] O.A. Zasedateleva, V.A. Vasiliskov, S.A. Surzhikov, A.Y. Sazykin, L.V. Putlyaeva, A.M. Schwarz, D.V. Kuprash, A.Y. Rubina, V.E. Barsky, A.S. Zasedatelev, UV fluorescence of tryptophan residues effectively measures protein binding to nucleic acid fragments immobilized in gel elements of microarrays, Biotechnol. J. 9 (2014) 1074–1080, https://doi.org/10.1002/biot.201300556.

[32] A.W. Drake, D.G. Myszka, S.L. Klakamp, Characterizing high-affinity antigen/antibody complexes by kinetic- and equilibrium-based methods, Anal. Biochem. 328 (2004) 35–43, https://doi.org/10.1016/j.ab.2003.12.025.

[33] A.R. Oliphant, C.J. Brandl, K. Struhl, Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: analysis of yeast GCN4 protein, Mol. Cell Biol. 9 (1989) 2944–2949, https://doi.org/10.1128/mcb.9.7.2944.

[34] A.D. Ellington, J.W. Szostak, In vitro selection of RNA molecules that bind specific ligands, Nature 346 (1990) 818–822, https://doi.org/10.1038/346818a0.

[35] C. Tuerk, L. Gold, Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, Science 249 (1990) 505–510, https://doi.org/10.1126/science.2200121.

[36] P.J. Rothwell, G. Waksman, Structure and mechanism of DNA polymerases, in: J.M. Squire, D.A.D. Parry (Eds.), Advances in protein chemistry, 71 2005, pp. 401–440, , https://doi.org/10.1016/S0065-3233(04)71011-6.