

---

# Non-Covalent Interactions between dUTP C5-Substituents and DNA Polymerase Result in a Decrease in PCR Efficiency

---

[Olga A. Zasedateleva](#)<sup>\*</sup>, [Sergey A. Surzhikov](#), Viktoriya E. Kuznetsova, Valeriy E. Shershov, Victor E. Barsky, Alexander S. Zasedatelev, [Alexander V. Chudinov](#)

Posted Date: 2 August 2023

doi: 10.20944/preprints202308.0184.v1

Keywords: C5-modified dUTPs; A and B family DNA polymerases; PCR amplification; X-ray structure; molecular modeling; non-covalent interactions



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

# Non-Covalent Interactions between dUTP C5-Substituents and DNA Polymerase Result in a Decrease in PCR Efficiency

Olga A. Zasedateleva \*, Sergey A. Surzhikov, Viktoriya E. Kuznetsova, Valeriy E. Shershov, Victor E. Barsky, Alexander S. Zasedatelev and Alexander V. Chudinov

<sup>1</sup> Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov Street, 119991 Moscow, Russia; ssergey77@mail.ru (S.A.S); kuzneimb@gmail.com (V.E.K.); shershov@list.ru (V.E.S.); victorbarsky@gmail.com (V.E.B); zas@biochip.ru (A.S.Z.); chud@eimb.ru (A.V.C.)

\* Correspondence: for\_lana3@mail.ru; Tel.: 7-499-135-99-80

**Abstract:** The approach based on molecular modeling was developed to study dNTP derivatives characterized by new polymerase-specific properties. For this purpose, the relative efficiency of PCR amplification with modified dUTPs was studied using Taq, Tth, Pfu, Vent, Deep Vent, Vent (exo-), and Deep Vent (exo-) DNA polymerases. The efficiency of PCR amplification with modified dUTPs was compared with the results of molecular modeling using the known 3D structures of KlenTaq polymerase–DNA–dNTP complexes. The dUTPs were C5-modified with bulky functional groups (the Cy5 dye analogs) or lighter aromatic groups. Comparing the experimental data and the results of molecular modeling revealed the decrease of PCR efficiency in the presence of modified dUTPs with an increase in the number of non-covalent bonds between the substituents and the DNA polymerase (about 15% decrease per one extra non-covalent bond). Generalization of the revealed patterns to all the studied polymerases of the A and B families is discussed herein. The number of non-covalent bonds between the substituents and polymerase amino acid residues is proposed to be a potentially variable parameter for regulating enzyme activity.

**Keywords:** C5-modified dUTPs; A and B family DNA polymerases; PCR amplification; X-ray structure; molecular modeling; non-covalent interactions

## 1. Introduction

The regularities and mechanisms of the enzymatic incorporation of natural and modified nucleoside triphosphates by various DNA polymerases are intensively investigated [1–3]. The relevant studies allow the creation of modified dNTPs to inhibit the polymerase's enzymatic activity in pathogenic bacteria and viruses and not affect the polymerases of higher organisms [4–7]. Other aspects of this field include the creation of aptamers—enzymatically synthesized modified oligonucleotides capable of binding to target proteins with high specificity. This approach can be used to diagnose and treat infectious, cancer, and other diseases [8–20]. The key to this technology is the ability of various polymerases to synthesize DNA strands using nucleoside triphosphates with specific modifications in the structure of bases [21–29] and, in particular, to synthesize oligonucleotides with modifications to all bases [21,25,29].

3D-structural models of the enzymatic incorporation of natural and modified nucleotides are also investigated [1,30–40]. Based on X-ray diffraction studies, Steitz T.A [34] and Rothwell P.J. and Waksman G. [35] found that, notwithstanding the significantly different amino acid composition, DNA polymerases of various organisms have a common shape resembling that of a half-open right hand and consisting of several subdomains—the so-called fingers, thumb, and palm [34,35]. These subdomains cover template fragment and synthesized DNA strand and form an active center where the complementary dNTP is localized. The dNTP is fixed in the necessary conformation by a set of non-covalent bonds with amino acid residues of the polymerase. The process is followed by the subsequent attachment of nucleotide to the growing DNA strand due to the formation of a

phosphodiester bond and by the translocation of the substrate for a new round of incorporation [32,35].

Taq DNA polymerase is currently one of the most studied DNA polymerases. As shown by Waksman G. et al. [32], a large fragment of the Taq polymerase, KlenTaq, has open and closed forms, the so-called "relaxed" and "tight" structures, depending on the stage of nucleotide incorporation. The amino acid residues belonging to the active center of the Taq DNA polymerase (Tyr611, Ser612, Gln613, Ile614, Glu615, His639, the amino acid residues of O alpha helix—Arg659, Lys663, and Phe667, as well as Tyr671) form non-covalent interactions with potentially active chemical groups of the nucleotide being incorporated [32]. Arg660 residue also plays an essential role in incorporating a nucleotide stabilizing the complex's structure due to forming a hydrogen bond with a phosphate group at the 3' end of the primer DNA strand [32,33,36].

A significant study has been done by Marx A. et al. [36–41], who analyzed the structures of the "closed" DNA–KlenTaq polymerase complexes with modified dNTPs non-covalently bound to the active center of the enzyme. dNTPs bases were modified with substituents consisting of a linker and aliphatic or aromatic hydrocarbon functional groups. The enzymatic incorporation of these nucleotides occurs due to the presence of free spaces in the structure of the Taq polymerase adjacent to the active center. These free spaces, namely cavity A and cavity B, are located between the finger and the thumb or between the palm and the finger domains, respectively. More specifically, cavity A is "bordered by the Arg587 side chain and residues from the O-helix", while cavity B "runs parallel to the O-helix" [40]. When incorporating a modified nucleotide, the enzyme flexibly adapts its structure to various base modifications, redistributing interactions between amino acid residues and the 3' end of the primer DNA strand. Polymerase forms additional non-covalent bonds with the substituent of the modified nucleotide depending on the structure of its linker and functional group. Arg587, Arg660, Ala661, Lys663, and Thr664 amino acid residues were shown to form non-covalent interactions with the substituents.

As noted by Marx A. et al. [36,38,39], efficient enzymatic incorporation of modified nucleotides is facilitated by the chemical groups in the structure of the substituents capable of forming hydrogen or  $\pi$ -cationic bonds with amino acid residues of the DNA polymerase's active center.

Hocek M. et al. [42] used a docking procedure to localize modified dNTPs at the previously known Bst polymerase–DNA complex structure.  $\pi$ -cation interaction was shown to be between Arg629 of Bst polymerase and the phenyl group of dG<sup>Ph</sup>TP.

Our recent results [28] are somewhat consistent with the conclusions of Marx A. et al. [36,38,39]. Using the Taq polymerase, as well as the Tth, Vent (exo-), and Deep Vent (exo-) polymerases, we have shown that the efficiency of PCR incorporation of deoxyuridines modified with aromatic hydrocarbon groups correlated with the hydrophilicity of these groups [28].

The authors of the above-cited articles used dNTPs with bases modified by 39-577 Da molecular weight substituents. Meanwhile, according to recent studies [43–46], enzymatic DNA synthesis may occur in the presence of dNTPs carrying bulkier substituents (molecular weights up to 692 Da). For example Holliger P. et al. [43] showed the possibility of "colored" DNA synthesis by PCR with genetically modified DNA polymerase in the presence of dCTPs labeled with Cy3 or Cy5 fluorescent dyes.

We found earlier [46] that the incorporation efficiency of dUMPs modified via linker with analogs of Cy3 and Cy5 cyanine dyes into the DNA chain by Taq polymerase significantly depends on the total charge of the fluorophores. In particular, incorporation efficiency is approximately ten times higher for dUMPs fluorescently labeled with electroneutral zwitterionic analogs of cyanine dyes than dUMPs labeled with negatively charged analogs of cyanine dyes.

Thus in light of the abovementioned results the aims of the current study are:

- (i) To investigate the efficiency of PCR amplification on the charge of dNTP-attached bulky aromatic hydrocarbon substituents for different types of DNA polymerases;
- (ii) To find the potentially reactive chemical groups of dNTP-attached small and bulky substituents that can participate in non-covalent interactions with Taq polymerase amino acid residues;

(iii) To identify which Taq DNA polymerase amino acid residue groups could be involved in non-covalent interactions with dNTP-attached substituents.

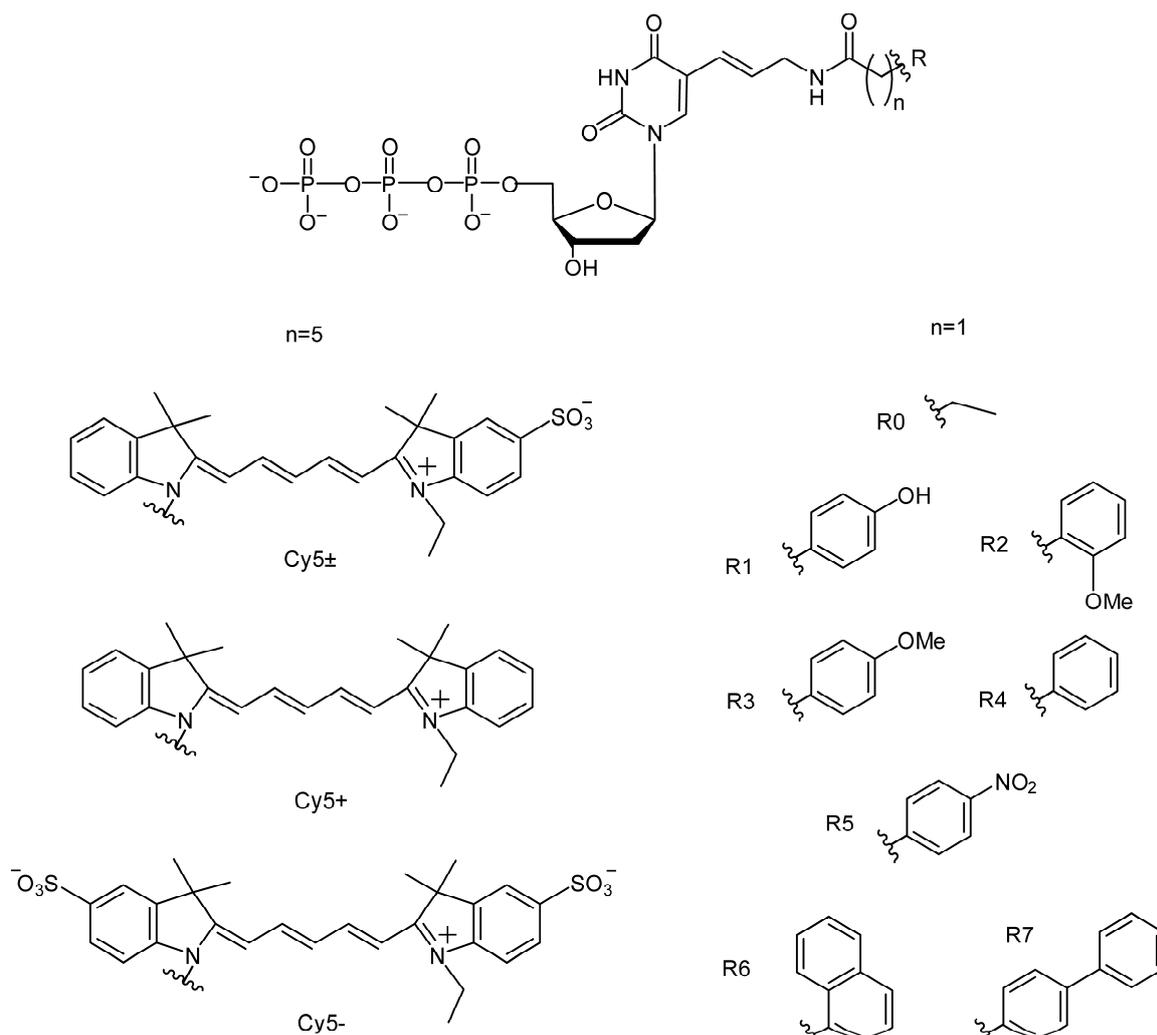
In this work, we used seven DNA polymerases from families A and B (Taq, Tth, Pfu, Vent (exo-), Deep Vent (exo-), Vent, and Deep Vent) to study the efficiency of PCR amplification in the presence of dUTPs C5-modified with bulky substituents (536 - 694 Da). These substituents carry Cy5 dye analogs (Cy5 $\pm$ , Cy5<sup>+</sup> or Cy5<sup>-</sup>) with a neutral (zwitterionic), positive, or negative charge, respectively. The experimental results obtained in this study for bulky substituents as well as those for lighter substituents (126-251 Da) obtained previously [28] were further compared with the results of molecular modeling using the known 3D structures of KlenTaq polymerase–DNA–dNTP complexes [33,40]. It was obtained that the increase in the quantity of non-covalent bonds between various C5-conjugated dUTP substituents and amino acid residues of KlenTaq polymerase resulted in a decrease in PCR amplification efficiency.

## 2. Results

### 2.1. PCR Amplification Efficiency Using Different Types of DNA Polymerases and dUTPs C5-Modified with Bulky Aromatic Hydrocarbon Substituents

#### 2.1.1. The Structures of C5-Attached dUTP Substituents and Electrophoretic Separation of PCR Products

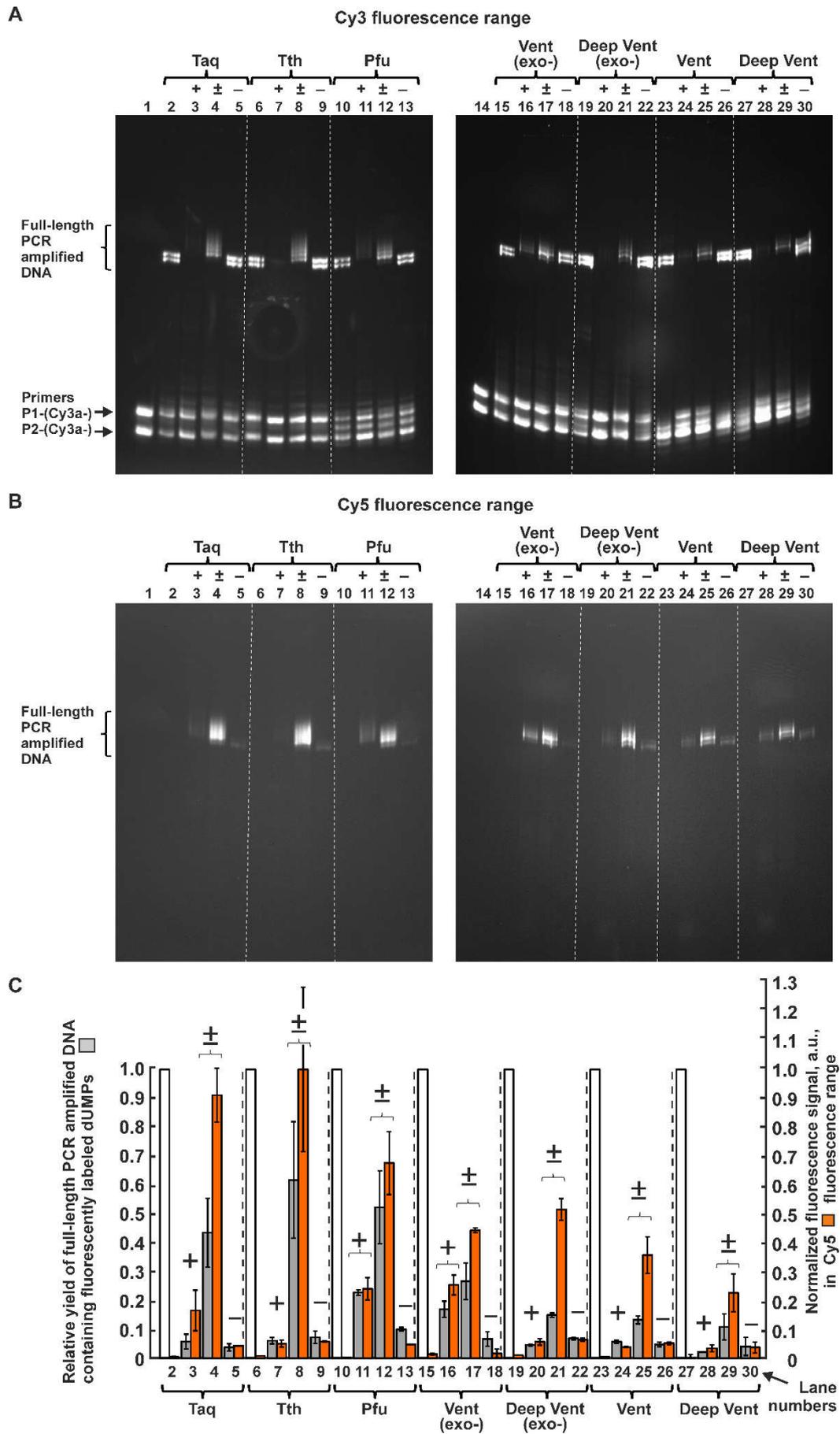
PCR efficiency was measured for seven A and B family DNA polymerases (Taq, Tth, Pfu, Vent (exo-), Deep Vent (exo-), Vent, and Deep Vent). In accordance with our previous data [51] PCR by Taq polymerase was most efficient in the presence of 5% of C5-modified dUTPs. Therefore in the present study of PCR efficiency in the presence of dUTPs C5-modified with bulky substituents, 5% dTTP in the reaction mixtures was replaced by modified dUTPs. Figure 1 shows that modifications were carried out via a linker (-CH=CH-CH<sub>2</sub>-NHCO-(CH<sub>2</sub>)<sub>5</sub>-) at the C5 position of the uridine base of dUTPs using differently charged Cy5 dye analogs and denoted as dU(Cy5<sup>+</sup>)TP, dU(Cy5 $\pm$ )TP, and dU(Cy5<sup>-</sup>)TP. The synthesized structures of Cy5 dye analogs differ from each other by the number and localization of negatively charged SO<sub>3</sub> groups.



**Figure 1.** dUTPs modified at the C5 position of the pyrimidine ring with aromatic hydrocarbon groups via linker  $-\text{CH}=\text{CH}-\text{CH}_2-\text{NHCO}-(\text{CH}_2)_n-$ . In the case of bulky substituents carrying Cy5 dye analogs, the linker part is longer ( $n=5$ ), while in the case of lighter substituents carrying  $R_0$ - $R_7$ , the linker part is shorter ( $n=1$ ).

Figure 2A,B show the electrophoregram of PCR products in polyacrylamide gel. The amount of PCR products was estimated after amplification in the presence of a 68-nucleotide template and two primers 18 and 17 nucleotides long, labeled at the 5' end with Cy3 dye (see also Materials and Methods).

Figure 2A (left and right) and Figure 2B (left and right) show the images of gels in the fluorescence ranges specifically for Cy3 and Cy5 dye, respectively.



**Figure 2. The efficiency of PCR amplification using seven A and B family DNA polymerases in the presence of dU(Cy5+)TP, dU(Cy5±)TP, or dU(Cy5-)TP in the reaction mixtures.** PCR mixtures with modified dUTPs contained  $10^{-4}$  M dATP, dTTP, dCTP,  $0.95 \times 10^{-4}$  M of dTTP, and  $0.05 \times 10^{-4}$  M of dU(Cy5+)TP, dU(Cy5±)TP, or dU(Cy5-)TP as correspondingly noted by the symbols (+, ± or -) in the plots. A. Electrophoretic separations of PCR products are shown in the fluorescence range of the Cy3 dye (the labeling dye for both primers). Lanes 1 and 14 are the PCR mixtures without the polymerases. Lanes 2, 6, 10, 15, 19, 23, and 27 are the PCR mixtures containing dNTPs and Taq, Tth, Pfu, Vent (exo-), Deep Vent (exo-), Vent, or Deep Vent polymerases, respectively. Lanes 3, 7, 11, 16, 20, 24, and 28 are the PCR mixtures containing, respectively, the same polymerases and dATP, dCTP, dGTP, as well as 95% dTTP, and 5% dU(Cy5+)TP. Lanes 4, 8, 12, 17, 21, 25, and 29 are the PCR mixtures containing, respectively, the same polymerases and dNTPs, as well as 5% dU(Cy5±)TP. Lanes 5, 9, 13, 18, 22, 26, and 30 show the PCR mixtures containing, respectively, the same polymerases and dNTPs, as well as 5% dU(Cy5-)TP. B. Electrophoretic separations of PCR products are shown in the fluorescence range of the Cy5 dye (the bulky substituents of dUTPs contained its derivatives (Cy5+, Cy5±, or Cy5-)) (see also Materials and Methods section). C. Histogram of the relative amounts of the full-size PCR-amplified DNA fragments. Grey columns represent the relative PCR efficiency obtained by normalizing the fluorescence intensities in the Cy3 dye fluorescence range of the DNA products containing incorporated modified dUTPs to the fluorescence intensities of the products obtained in the presence of only natural dNTPs. Orange columns show the fluorescence intensities of PCR products normalized to the maximum value in the Cy5 fluorescence range (for Tth polymerase). White columns reflect the relative fluorescent intensity of amplified DNA with only natural dNTPs in mixtures in the Cy3 fluorescence range (see the Materials and Methods section). The experiments shown in A and B were performed twice. The average values from the two experiments are listed in Table S1 and presented in C. The bars indicate the absolute deviations.

The natural reaction products are visible in lanes 2, 6, 10, 15, 19, 23, and 27 (Figure 2A). The reaction products amplified in the presence of C5-modified dUTPs are in lanes 3-5 (for Taq polymerase), 7-9 (for Tth polymerase), 11-13 (for Pfu polymerase), 16-18 (for Vent (exo-)) 20-22 (for Deep Vent (exo-)), 24-26 (for Vent polymerase) and 28-30 (for Deep Vent polymerase).

Unspent primers are in the lower half of the gel images in Figure 2A. Polymerases with exonuclease activity, Pfu, Vent, and Deep Vent, cut off one nucleotide at the 3' end of the primers, and thus the primers have a couple of lighter fragments (lanes 10-13, 23-30).

### 2.1.2. Inhibiting Effect of Modified dUTPs in the Reaction Mixture on the Yield of PCR Product Consisting Only of Natural Nucleotides

As can be seen in Figure 2, even low concentration ( $5 \times 10^{-6}$  M) of dU(Cy5+)TP or dU(Cy5±)TP in the PCR mixture leads to significant synthesis inhibition of PCR products consisting of only natural nucleotides (lanes 3, 4, 7, 8, 11, 12, 16, 17, 20, 21, 24, 25, 28, 29). However, similar inhibition does not occur when the reaction mixture contains dU(Cy5-)TP at the same concentration (lanes 5, 9, 13, 18, 22, 26, 30). See the corresponding paragraph "Part A: Inhibiting effect of modified dUTPs in the reaction mixture on the yield of PCR product consisting only of natural nucleotides" in Supplementary data for details.

### 2.1.3. PCR Efficiency of Seven A and B Family Polymerases for dUTPs C5-Modified with Bulky Aromatic Hydrocarbon Substituents

As seen in Figure 2A, for each polymerase, the PCR amplification in the mixture with 5% dU(Cy5±)TP leads to the appearance of heavier electrophoretic bands (lanes 4, 8, 12, 17, 21, 25, and 29) in comparison with the bands containing natural DNA product.

Up to three such bands are visible in lanes 4, 8, 12, 17, 21, 25 and 29. These additional bands appear due to the weighting of the synthesized DNA strands when they contain, respectively, from 1 to 3 nucleotides C5-modified with a bulky substituent carrying Cy5± (see also Figure 1). These additional bands can also be observed for each DNA polymerase in the Cy5 fluorescence range (Figure 2B), though we could not resolve them in that range.

The values of relative fluorescence intensities in the Cy3 fluorescence range, characterizing the yields of the full-size PCR products containing nucleotide-modified DNA fragments, are summarized in Table S1 and plotted for each of the DNA polymerases in the histogram in Figure 2C (grey columns). The histogram also shows the relative fluorescence intensities in the Cy5 fluorescence range for the full-size PCR products containing Cy5-modified nucleotides (Figure 2C, orange columns).

The data presented in Figure 2C enables us to rate the ability of DNA polymerases of the A and B families to use the dUTPs C5-modified with bulky Cy5 dye analogs in PCR amplification. For dU(Cy5±)TP, the rating can be approximately estimated as follows:

$$\text{Tth} \geq \text{Pfu} \geq \text{Taq} > \text{Vent(exo-)} > \text{Deep Vent (exo-)} \geq \text{Vent} \geq \text{Deep Vent} \quad (1)$$

At the same time, the rating for dU(Cy5+)TP is

$$\text{Pfu} \geq \text{Vent(exo-)} > \text{Taq} = \text{Tth} = \text{Vent} \geq \text{Deep Vent (exo-)} > \text{Deep Vent} \quad (2)$$

Notably, the ability of the tested DNA polymerases to use dU(Cy5-)TP in PCR amplification is neglectable.

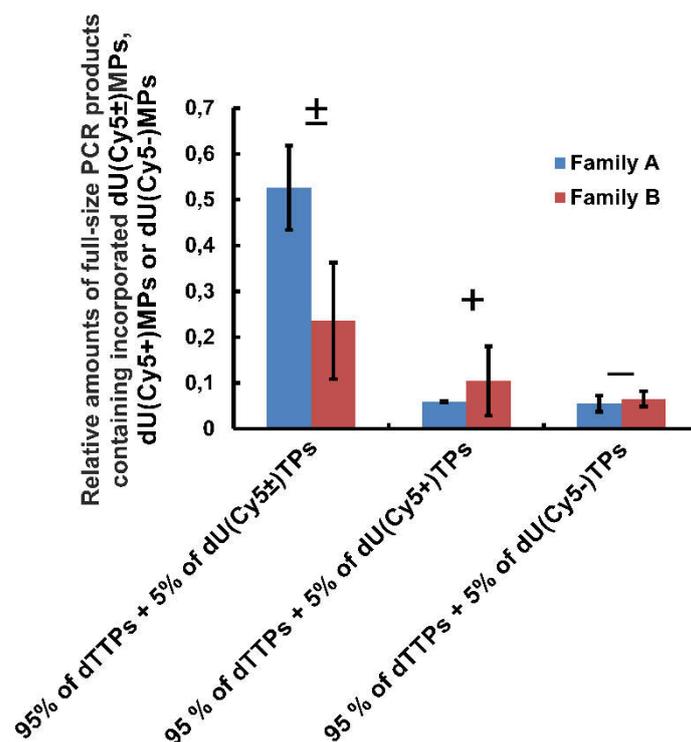
#### 2.1.4. Efficiency of PCR Amplification by DNA Polymerases of the A and B Families in the Presence of dUTPs C5-Modified with Cy5 Dye Analogs

Each polymerase synthesized the full-size PCR products containing incorporated dU(Cy5±)MPs less efficiently than entirely natural DNA products (Figure 2C and Figure 3), but more efficiently than the full-size DNA products modified by Cy5+ or Cy5- dye analogs.

Figure 3 shows that the efficiency of PCR amplification in the presence of modified dUTPs (as percentage from dTTP values) averaged independently for A family and B family DNA polymerases are (50±10)% and (24±13)% for dU(Cy5±)TP, (6±1)% and (11±8)% for dU(Cy5+)TP, and (6±2)% and (7±2)% for dU(Cy5-)TP, respectively. Thus, according to the efficiency of PCR amplification, the dNTPs are ranked in the following order:

$$\text{dTTP} > \text{dU(Cy5±)TP} > \text{dU(Cy5+)TP} \geq \text{dU(Cy5-)TP}. \quad (3)$$

This rating does not depend on whether the polymerase has 3'-5' exonuclease activity, as in the case of Pfu, Vent, and Deep Vent polymerases, or does not have this activity, as in the case of Taq and Tth, as well as in the cases when the polymerase is genetically modified so that the 3'-5' exonuclease function is turned off, as in the case of Vent(exo-) and Deep Vent (exo-) polymerases.



**Figure 3.** The relative amounts of full-size PCR products containing incorporated dU(Cy5±)MPs, dU(Cy5+)MPs, or dU(Cy5-)MPs averaged over the DNA polymerases of the A and B families compared to entirely natural PCR products. The relative amounts of the PCR products were taken from the experimental data presented in Figure 2 and Table S1. The bars indicate the mean absolute deviations.

## 2.2. Structural Factors Reducing the PCR Efficiency When dUTPs C5-Modified by Various Low-Molecular-Weight Substituents Are Used

The molecular modeling of possible conformations of the substituents at the C5 position of dUTPs localized in the active center of the KlenTaq polymerase was performed using the covalent docking procedure of Discovery Studio program [49] (a free version of the program was kindly provided by Dassault Systèmes software corporation). The substituents were virtually attached at the C5 position of pyrimidine rings of the known structures of the KlenTaq polymerase–DNA–(dUTP or ddTTP) "closed" triple complexes first determined by Waksman G. et al. [33] and later by Marx A. et al. [40] by X-ray structural analysis and published in the protein database [50] (PDB IDs: 1QTM and 5E41).

Before analyzing bulky aromatic hydrocarbon substituents (536 - 694 Da) attached at the C5 position of the pyrimidine ring of dUTP (characterized in Section 2.1), we analyzed the dUTPs modified using comparatively smaller and simpler substituents. For this purpose, the published experimental data [28] on the efficiency of PCR amplification using Taq, Tth, Vent (exo-), and Deep Vent (exo-) polymerases when replacing dTTP with dUTPs modified with small low-molecular-weight substituents (126-251 Da) containing R0 – R7 functional groups (see Figure 1) were used.

Accordingly, the following low molecular weight substituents were virtually constructed (see Figure 1):

(a) Small (126-251 Da) substituents consisting of -CH=CH-CH<sub>2</sub>-NHCO-CH<sub>2</sub>- linker and R0 – R7 functional groups.

(b) Bulky (536 - 694 Da) substituents consisting of -CH=CH-CH<sub>2</sub>-NHCO-(CH<sub>2</sub>)<sub>5</sub>- linker and the analogs of Cy5 dye.

### 2.2.1. Identifying the Non-Covalent Interactions between the Substituents Containing Small Functional Groups R0-R7 and the KlenTaq Polymerase Amino Acid Residues

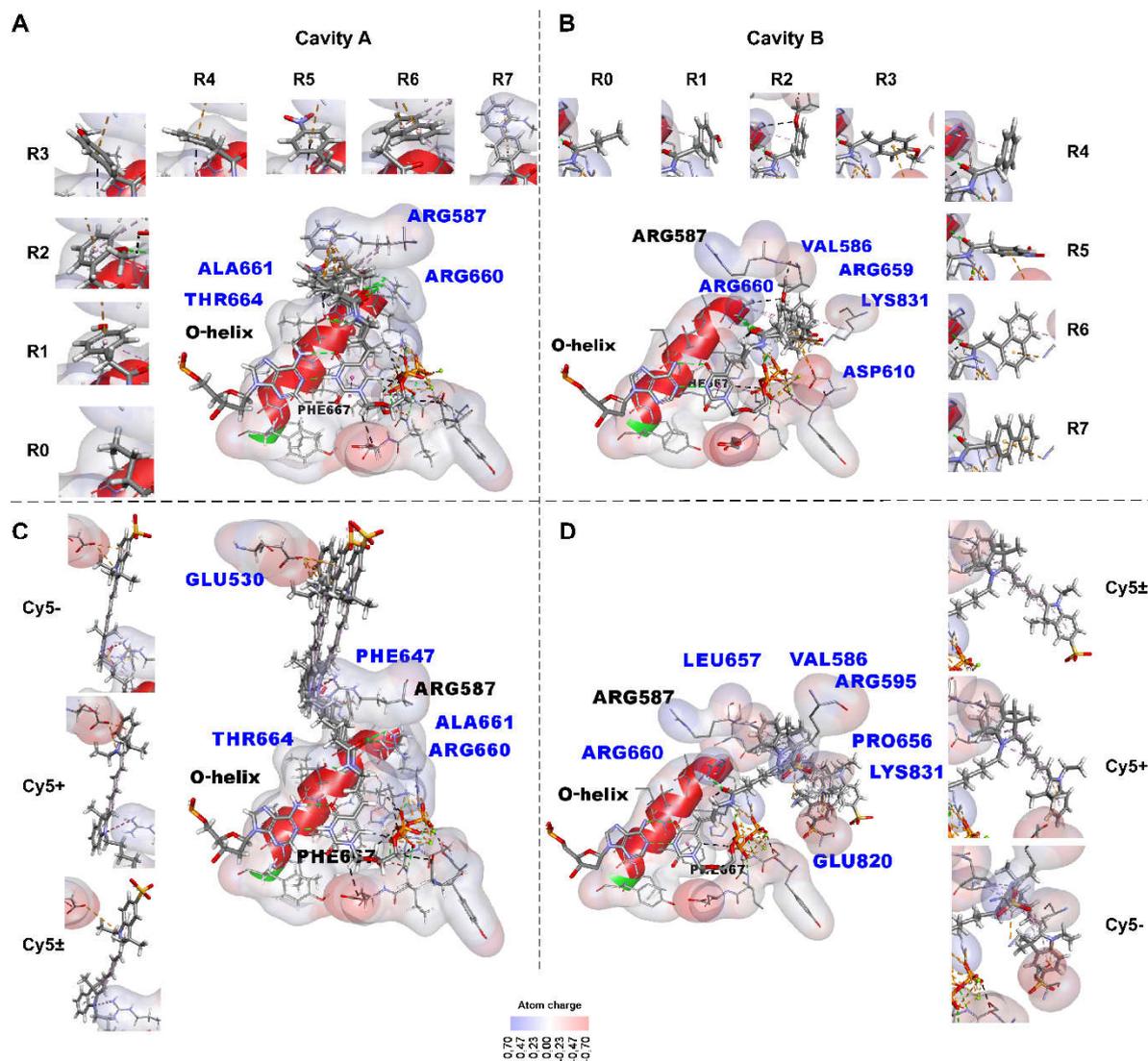
3D structures of the linker (-CH=CH-CH<sub>2</sub>-NHCO-CH<sub>2</sub>-) and the attached functional groups (R0 - R7) were optimized by minimizing the relative energies of their conformations and interactions with surrounding chemical groups of polymerase and DNA. Meanwhile the remaining parts of the KlenTaq polymerase–DNA–(modified dUTP) complex maintain rigid conformation.

The optimized conformations and the localization of small substituents of the modified dUTPs within the fragments of KlenTaq polymerase–DNA–(modified dUTP) 3D complexes are shown in Figure 4A,B. The substituents of the modified dUTPs can be oriented in two directions— via cavity A or cavity B, which are adjacent to the space of the active center of the KlenTaq polymerase [40]. We obtained that in the case of the 5E41 structure, the substituents extend through cavity A (Figure 4A), and in the case of the 1QTM structure, the substituents extend through cavity B (Figure 4B).

Figure 4A,B also show various types of non-covalent bonds, namely:

- hydrogen,
- carbon–hydrogen,
- $\pi$ -donor hydrogen,
- $\pi$ -cation,
- $\pi$ -anion, and

$\pi$ -alkyl bonds between the chemical groups of the substituents and the chemical groups of the amino acid residues of the KlenTaq polymerase obtained using the Discovery Studio program. These interactions are listed in Table S2. Table S2 also presents the mean number of non-covalent interactions between each substituent, and KlenTaq polymerase averaged over two types of localization (via cavity A and cavity B).



**Figure 4.** Fragments of three-dimensional structures of the KlenTaq polymerase–DNA–(modified dUTP or ddTTP) complexes (based on X-ray structures, PDB IDs 5E41 and 1QTM [33,40,50]). Computer-simulated conformations and localization of the low-molecular-weight substituents attached at the C5 position of the pyrimidine rings of the modified nucleotides incorporated in the active center of the enzyme are illustrated (.dsv files with these 3D structures are located in Supplementary data).

A and B—the small substituents (126–251 Da) containing the functional groups R0 – R7 attached using the  $-\text{CH}=\text{CH}-\text{CH}_2-\text{NHCO}-\text{CH}_2-$  linker.

C and D—the bulky substituents (536–694 Da) containing the analogs of Cy5 dye (zwitterionic neutrally charged  $\text{Cy}5^\pm$ , positively charged  $\text{Cy}5^+$ , or negatively charged  $\text{Cy}5^-$ ) attached using the  $-\text{CH}=\text{CH}-\text{CH}_2-\text{NHCO}-\text{(CH}_2)_5-$  linker.

A and C—computer-simulated superpositions of the substituents are extended through cavity A of the 5E41 structure [50] of the enzyme (crossing over the O-helix).

B and D—computer-simulated superpositions of the substituents are extended through cavity B of the 1QTM structure [50] of the enzyme (partially parallel to the O-helix).

The interactions between the functional groups of each substituent with the neighboring amino acid residues are also shown in eight (A, B) or three (C, D) separate images detailing the corresponding superimposed structures.

The amino acid residues of the KlenTaq polymerase forming non-covalent bonds with the substituents are indicated by blue characters. The charge distribution over the surfaces bounding the van der Waals radii of the

atoms is shown for amino acid residues of the active center of the KlenTaq polymerase using the standard scale of colors (plus—blue, minus—red).

The non-covalent bonds are shown using the Discovery Studio program [49] as follows:

- Hydrogen bonds by green dotted lines;
- Carbon and  $\pi$ -donor hydrogen bonds by black dotted lines;
- Electrostatic interactions, including  $\pi$ -cationic and  $\pi$ -anionic, by orange dotted lines;
- Hydrophobic  $\pi$ -alkyl interactions by purple dotted lines;
- Electrostatic attractive charge bonds by orange dotted lines;
- $\pi$ -lone pair bond by green dotted line;
- $\pi$ -sulfur bond by yellow dotted line;
- Hydrophobic  $\pi$ - $\pi$  stacking bond is denoted by a pink dotted line, alkyl bonds by purple dotted lines, and the  $\pi$ -sigma bond by a dark purple dotted line. The bonds can be seen more in the Discovery Studio program [49] using .dsv files with these 3D structures in Supplementary data. An unfavorable positive-positive interaction between the nitrogen atom close to the linker indolyl ring of the substituent and the nitrogen atom of Arg587 is marked by a red dotted line in C.

### 2.2.2. Localization of Bulky Aromatic Hydrocarbon Substituents (Containing Cy5 Dye Analogs) in KlenTaq Polymerase–DNA–(Modified dUTP) 3D Complexes

Figure 4C,D show sterically resolved and optimized conformations of the bulky substituents and their localization in KlenTaq polymerase–DNA–(modified dUTP) 3D complexes after energy minimization via covalent docking procedure. The orientation of the bulky aromatic substituents carrying Cy5 dye analogs is the same as for the small low-molecular-weight substituents shown in Figure 4A,B. Similar to the small aromatic substituents, the bulky aromatic substituents extend via cavity A for the 5E41 structure (Figure 4C) or via cavity B for the 1QTM structure (Figure 4D).

### 2.2.3. Non-Covalent Interactions between the Bulky Aromatic Substituents and the KlenTaq Polymerase Amino Acid Residues

Figures 4C,D and S1–S3 show various types of non-covalent bonds formed between the bulky aromatic substituents and the amino acid residues of the KlenTaq polymerase. These non-covalent bonds are

hydrogen,  
electrostatic (attractive charge bonds,  $\pi$ -anion bonds),  
 $\pi$ -lone pair bond,  
 $\pi$ -sulfur bond, and

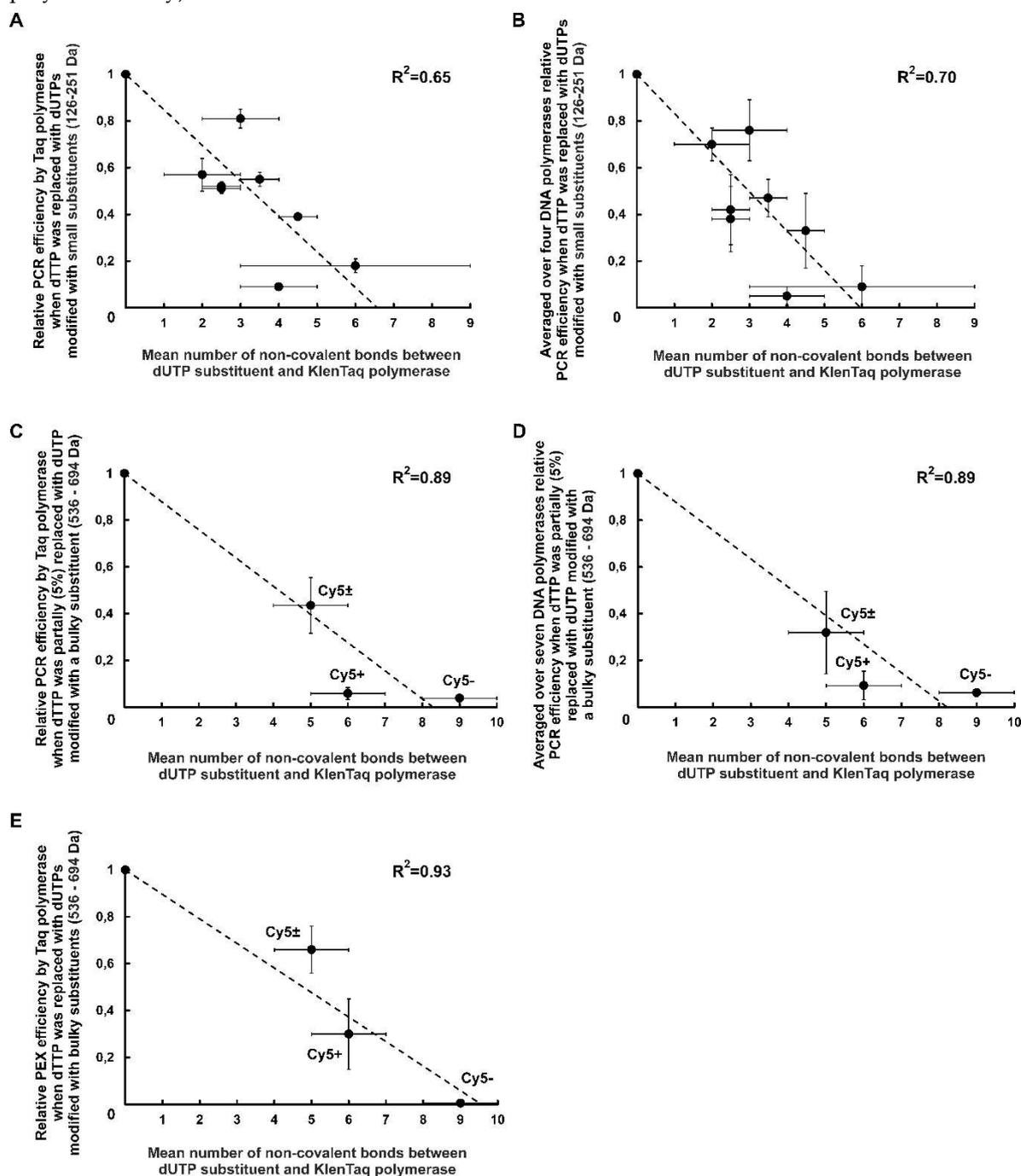
hydrophobic bonds ( $\pi$ - $\pi$ -stacked,  $\pi$ -sigma, alkyl, and  $\pi$ -alkyl bonds) which are also listed in Table S3. Table S3 also presents the mean numbers of non-covalent bonds between the chemical groups of the substituent, and the chemical groups of the KlenTaq polymerase amino acid residues averaged over two types of localization (via cavity A and cavity B) for each substituent.

### 2.2.4. Negative Correlation between PCR Efficiency in the Presence of Modified dUTPs and the Number of Non-Covalent Bonds between the dUTP C5-Substituents and DNA Polymerase

#### 2.2.4.1. PCR Compatibility of dUTPs C5-Modified with R0-R7 Functional Groups

Figure 5A shows the plot of experimentally estimated values of the relative PCR efficiency (for Taq polymerase in the presence of dUTPs C5-modified with R0 – R7 functional groups) depending on the mean number of non-covalent bonds between the substituents and the polymerase. These mean numbers were calculated for each substituent by averaging the numbers of substituent-polymerase bonds estimated by docking for two possible localizations of the substituent in the enzyme (through the A cavity or the B cavity). This plot demonstrates a negative correlation (characterized by a linear decrease of about 15% per one non-covalent bond,  $R^2=0.65$ ) between the relative efficiency of PCR amplification (in the presence of modified dUTP) and the mean number of non-covalent bonds (formed by the dUTP C5-substituent with the KlenTaq polymerase amino acid

residues). Notably, a similar negative correlation ( $R^2=0.70$ , see Figure 5B) takes place if the values of the relative PCR efficiency averaged over the polymerases studied in [28] (Taq, Tth, Vent (exo-), and Deep Vent (exo-)) are plotted on the y-axis (instead of the relative PCR efficiency values for the Taq polymerase only).



**Figure 5. Correlations between the relative PCR or PEX efficiency in the presence of C5-modified dUTPs and the number of non-covalent bonds formed by the dUTP substituents with the KlenTaq polymerase amino acid residues.** X-axis in all the plots shows the mean number of non-covalent bonds between dUTP substituents and KlenTaq polymerase. The numbers of non-covalent bonds were obtained by molecular modeling based on the known X-ray structures [33,40,50] and averaged on localizations within 5E41 and 1QTM structures (see also Tables S2 and S3 and Figures 4 and S1–S3). PCR efficiency is plotted as a percentage with that in the presence of only natural dNTPs. A. Y-axis shows the relative PCR efficiency of the Taq polymerase when all the dTTP molecules were replaced with dUTPs modified with small substituents (126-251 Da) carrying functional groups R0–R7. B. Y-axis shows the relative PCR efficiency averaged over four polymerases (Taq, Tth, Vent (exo-

), and Deep Vent (exo-)) when all the dTTP molecules were replaced with dUTPs C5-modified with small substituents ((126-251 Da) carrying functional groups R0-R7. C. Y-axis shows the relative PCR efficiency of the Taq polymerase when 5% dTTP molecules were replaced with the dUTPs modified with bulky aromatic substituents (536-694 Da) carrying Cy5 dye analogs of various total charges, i.e., Cy5±, Cy5+, or Cy5-. D. Y-axis shows the relative PCR efficiency averaged over seven polymerases (Taq, Tth, Pfu, Vent, Deep Vent, Vent (exo-), and Deep Vent (exo-)) when 5% dTTP molecules were replaced with dUTPs modified by bulky aromatic substituents. E. Y-axis shows the relative PEX efficiency using the Taq polymerase when all the dTTP molecules were replaced with the dUTPs modified by bulky aromatic substituents. The bars indicate mean absolute deviations. The dotted straight lines show the linear approximations for the corresponding values.

#### 2.2.4.2. PCR Compatibility of dUTPs C5-Modified with the Bulky Substituents

Figure 5C shows the experimentally obtained values of the relative PCR efficiency obtained for the Taq polymerase in the presence of dUTPs C5-modified by the bulky substituents (containing Cy5+, Cy5±, or Cy5-; see Figure 1 and Table S3) depending on the mean number of non-covalent bonds between the substituents and the polymerase. As for small low-molecular-weight aromatic substituents (see Figure 5A), the plot in Figure 5C demonstrates a linear decrease of the relative PCR efficiency for about 12% per one non-covalent bond,  $R^2=0.89$  for the case of bulky substituents.

Thus, dUTPs C5-modified with bulky substituents can be characterized as follows:

For dU(Cy5±)TP:

relative PCR efficiency –  $0.44\pm 0.12$ ; mean number of non-covalent bonds –  $(5\pm 1)$ ;

For dU(Cy5+)TP:

relative PCR efficiency –  $0.06\pm 0.03$ ; mean number of non-covalent bonds –  $(6\pm 1)$ ;

For dU(Cy5-)TP:

relative PCR efficiency –  $0.04\pm 0.01$ ; mean number of non-covalent bonds –  $(9\pm 1)$ .

The plots in Figure 5A,C demonstrate that PCR amplification efficiency by Taq polymerase in the presence of modified dUTPs is decreased with increasing the number of non-covalent interactions formed between the dUTP C5-substituent and the polymerase.

Figure 5D shows PCR efficiency values averaged over all the polymerases used in our work (Taq, Tth, Pfu, Vent, Deep Vent, Vent (exo-), and Deep Vent (exo-)) and plotted on the y-axis. Similar for small substituents (Figure 5B), the negative correlation ( $R^2=0.89$ ) was found for the PCR efficiency values averaged over all the tested polymerases for bulky substituents (Figure 5D).

Summing the obtained results one can conclude that the number of non-covalent bonds between the dUTP C5-substituent and the amino acid residues of polymerase characterizes the ability of polymerase to incorporate modified nucleotide. The established pattern can be used to design and synthesize dNTP derivatives characterized by new polymerase-specific properties.

#### 2.2.5. Negative Correlation between the Number of Non-Covalent Bonds Formed by dUTP C5-Substituents with the Taq Polymerase and PEX Efficiency in the Presence of dUTPs Modified with Bulky Chromophore-Containing Substituents

Figure 5E shows the relative efficiency of PEX by Taq polymerase in the presence of dUTPs C5-modified with bulky aromatic substituents on the number of non-covalent bonds between the substituents and the amino acid residues of the KlenTaq polymerase. The relative efficiencies of PEX were taken from our recent paper [47]. In that paper, PEX experiments were carried out with dU(Cy5+)TP, dU(Cy5±)TP, and dU(Cy5-)TP and a similar DNA template.

The dependences of the relative PCR (Figure 5C) and PEX (Figure 5E) efficiencies on the docking-obtained quantity of non-covalent interactions between the dUTP C5-substituent and Taq polymerase looks quite similar. Thus, both PCR and PEX approaches, combined with molecular modeling, are applicable for analyzing non-covalent interactions between the dNTP substituent and the polymerase affecting the latter's functioning.

### 2.2.6. The Similarity of the Local Environments of the dUTP C5-Substituents Localized in the Active Center of Different Polymerases

Based on the similarity of the plots shown in Figure 5A,B and the plots shown in Figure 5C,D, one can propose that the chemical groups of amino acids facing the spaces of active centers of DNA polymerases, as well as those facing the adjacent regions (possibly forming cavities similar to the A and B KlenTaq polymerase cavities), form similar local environments. These local environments are similar not only for dUTP (that is well known) but also for its substituents attached at the C5 position of the pyrimidine ring.

It can be assumed that modified nucleotide under incorporation by different polymerases is localized in a similar environment of chemical groups despite the varieties in the amino acid chains forming the active centers of different polymerases. For details, see the corresponding paragraph "Part E: The similarity of local environments in the active centers of various polymerases to the dUTP substituents attached at the C5 position of the pyrimidine ring" in Supplementary data.

### 2.2.7. Analysis of the Linker Parts of the dUTP-Attached Substituents in Non-Covalent Interactions with Taq Polymerase Amino Acid Residues

Analysis of the  $-\text{CH}=\text{CH}-\text{CH}_2-\text{NHCO}-\text{CH}_2-$  and  $-\text{CH}=\text{CH}-\text{CH}_2-\text{NHCO}-(\text{CH}_2)_5-$  linker parts of the modeled substituents shows that mainly the NH group and oxygen atom of the "inverted" peptide group, NHCO, of both linkers are involved in the formation of non-covalent (hydrogen) bonds with the atoms of the polymerase amino acid residues (Figure 4). NH group or oxygen atom or both of them are involved in interactions. Meanwhile, for the substituent-carrying R6 functional group, a single carbon-hydrogen bond is formed between the carbon atom (next to the nitrogen atom of the NHCO group) of the linker and the oxygen atom of the Arg660 CONH group.

### 2.2.8. Analysis of the Functional Groups R0-R7 of the dUTP-Attached Substituents in Non-Covalent Interactions with Taq Polymerase Amino Acid Residues

Structural models presented in Figure 4A,B show that the aromatic rings of R1-R7 functional groups form non-covalent ( $\pi$ -donor hydrogen, electrostatic and/or hydrophobic) interactions with the amino acid residues of the KlenTaq polymerase. Meanwhile, the chemical groups attached to the aromatic rings (see functional groups R1-R7 in Figure 1) do not participate in the interactions with the polymerase chemical groups.

The exception is the 2-OMe group in R2. The oxygen atom of the 2-OMe group in R2 is involved in the hydrogen bond with the nitrogen atom of Arg660 in the case of the 5E41 structure (Figure 4A) and the carbon-hydrogen bond with the carbon atom of Arg660 in the case of the 1QTM structure (Figure 4B). The carbon atom of the 2-OMe group in R2 is also involved in the carbon-hydrogen bond with the oxygen atom of Val586 in the case of the 1QTM structure.

The formation of the non-covalent bonds between 2-OMe group of R2 and the polymerase can be the reason of 1.5–2 times lower PCR efficiency of Taq, Vent (exo-), and Deep Vent (exo-) polymerases in the presence of R2-modified dUTPs as compared with dUTPs conjugated with a 4-methoxyphenyl residue (R3) [28] whose 4-OMe group does not form non-covalent interactions.

### 2.2.9. Analysis of Bulky dUTP-Attached Substituents in Non-Covalent Interactions with Taq Polymerase

Figure S3 shows that the negatively charged bulky substituent of dU(Cy5-)TP forms more non-covalent bonds with the polymerase, ( $9\pm 1$  bonds), than neutral ( $5\pm 1$  bonds) or positively charged ( $6\pm 1$  bonds) bulky substituents (Cy5 dye analogs). The negatively charged substituent differs from neutral or positively charged ones by the additional  $\text{SO}_3$  group attached to the chromophore indolyl ring closest to the linker. This  $\text{SO}_3$  group forms two extra non-covalent bonds with the amino acid residues of the Taq polymerase in both structures: 5E41 and 1QTM. In the case of the 5E41 structure, the aromatic ring of Phe647 forms a  $\pi$ -sulfur bond with the sulfur atom of the  $\text{SO}_3$  group and a  $\pi$ -lone pair bond with the oxygen atom of the  $\text{SO}_3$  group (see Figure S3A). In the case of the 1QTM structure,

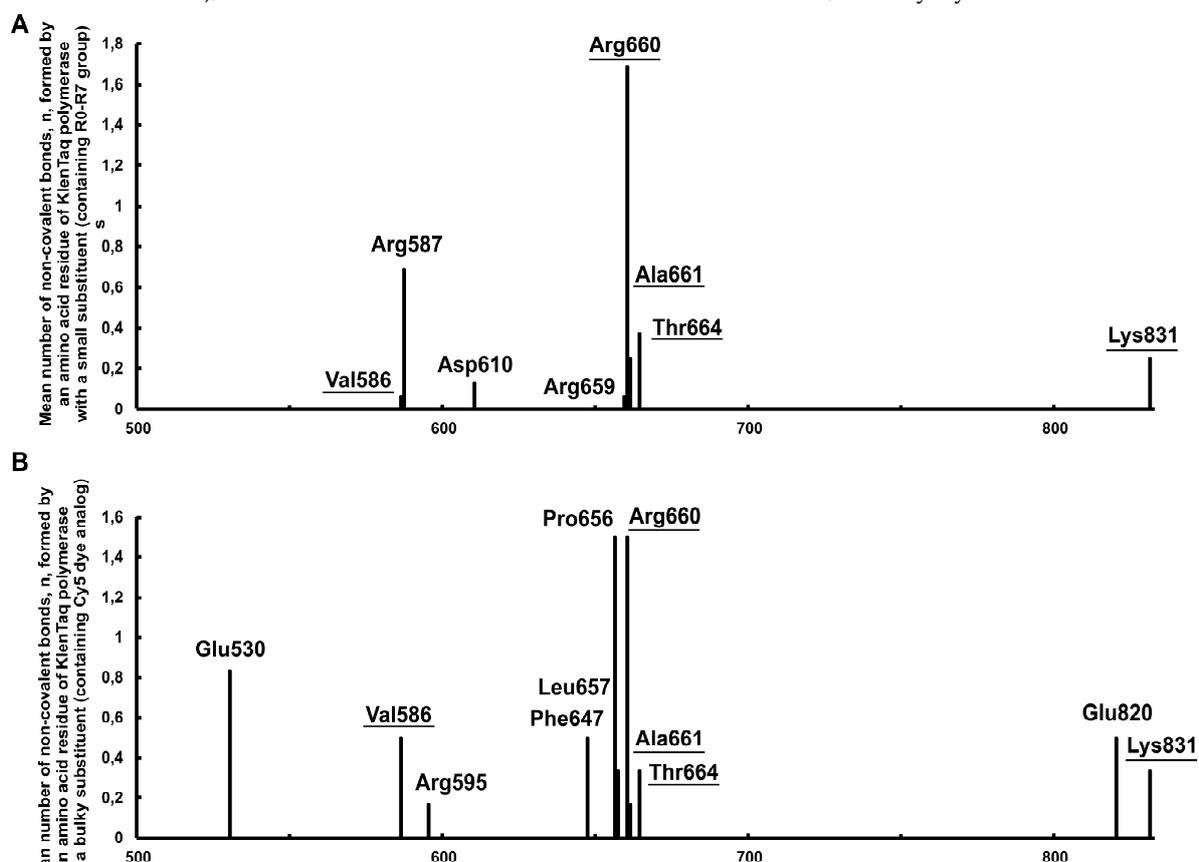
the nitrogen atoms of Arg595 and Lys831 form two electrostatic attractive charge bonds with the sulfur atom of the SO<sub>3</sub> group (see Figure S3B). Due to these extra bonds and the other non-covalent bonds (6 in the case of 5E41 structure and 8 in the case of 1QTM structure), dU(Cy5-)MP probably has lower efficiency of inclusion during PCR or PEX.

Notably, indolyl rings of the substituents which are closest to the linker in both dU(Cy5+)TP and dU(Cy5±)TP do not contain SO<sub>3</sub> groups in the contrary to dU(Cy5-)TP. Respectively, neutral and positively charged substituents form a lower number of non-covalent bonds, (6±1 bonds) and (5±1 bonds) correspondingly (see Figures S1 and S2 and Table S3).

The considered numbers of non-covalent bonds are the extra ones compared to their absence in dUTP. These extra bonds and the corresponding energies of interactions should be overcome to release the polymerase molecule for the movement along the DNA primer/template just after the formation of the phosphodiester bond.

#### 2.2.10. Analysis of Amino Acid Residues Capable of Forming Non-Covalent Bonds with the dUTP-Attached Substituents

Figure 6A shows the result obtained for small low-molecular-weight aromatic substituents (126-251 Da). Eight amino acid residues are detected with the help of molecular modeling (Val586, Arg587, Asp610, Arg659, Arg660, Ala661, Thr664, and Lys831) as capable of forming non-covalent bonds with the dUTP-attached substituents. Five of them (Arg587, Arg659, Arg660, Ala661, and Thr664) are included in the active center of the enzyme, while other three (Val586, Asp610, and Lys831) face cavity B adjacent to the active center of the KlenTaq polymerase. The amino acid residues Arg587, Arg660, Ala661, and Thr664 found by this modeling were first observed by Marx A. et al. ([36–41]; see Introduction), who found as well as one more amino acid residue, namely Lys663.



**Figure 6.** The mean number of non-covalent bonds, *n*, between the KlenTaq polymerase amino acid residues and the substituents (attached at the C5 position of the pyrimidine base of the dUTPs) containing functional groups R0-R7 (A) or Cy5 dye analogs (B). X-axis shows the number of the amino acid residue in the Taq polymerase chain. Y-axis shows the mean number of non-covalent bonds per one structural solution. The data were obtained using molecular modeling and covalent

docking for the substituents of the modified dUTPs localized in the structures 1QTM and 5E41. The amino acid residues forming non-covalent interactions with both small and bulky substituents are underlined and correspondingly shown in the histograms A and B. The n values were calculated as described in Materials and Methods (paragraph 2.6.) and detailed in "Part F: Table S4" in Supplementary data.

Similarly, Figure 6B shows that the amino acid residues participating in the interactions with bulky aromatic hydrocarbon substituents are Glu530, Val586, Arg595, Phe647, Pro656, Leu657, Arg660, Ala661, Thr664, Glu820, and Lys831. Among these residues, Glu530, Val586, Arg595, Phe647, Pro656, Leu657, Glu820, and Lys831 extend beyond the active center of the polymerase. The amino acid residues Glu530 and Phe647 face cavity A, while the amino acid residues Val586, Arg595, Pro656, Leu657, Glu820, and Lys831 face cavity B.

Totally 14 KlenTaq polymerase amino acid residues (Glu530, Val586, Arg587, Arg595, Asp610, Phe647, Pro656, Leu657, Arg659, Arg660, Ala661, Thr664, Glu820, and Lys831) are detected with the help of molecular modeling and docking and identified as capable of forming non-covalent bonds with the chemical groups of both light (126-251 Da) or bulky (536 - 694 Da) low-molecular-weight aromatic substituents. Among these amino acid residues five, Val586, Arg660, Ala661, Thr664, and Lys831 (underlined in Figure 6A, B), can form non-covalent interactions with both small and bulky substituents (see also Materials and Methods, paragraph 4.6. and "Part F: Table S4" in Supplementary data).

### 3. Discussion

Analysis of the structures of DNA polymerase–DNA–(modified dNTP) 3D complexes aims to understand the mechanisms of enzymatic incorporation of modified nucleotides into the DNA chain. This is essential to find optimal structures of functional dNTP substituents necessary to meet many biotechnological and medical challenges [52]. Usually, this analysis is time consuming as it is rather empirical and based on direct experimental verification. The correlation between the experimental data and the molecular modeling results obtained in the present study might be applied to facilitate the search for specific substituent structures. The described structural–functional relationships would probably allow one to analyze the properties of modified dNTPs using preliminary computational methods. This approach aims to either increase the efficiency of polymerases or use modified dNTPs as their selective inhibitors. The regularities identified in this study are considered and discussed below.

#### *3.1. Negative Correlation between the PCR Amplification Efficiency in the Presence of the Modified Dutps and the Number of Non-Covalent Bonds between the dUTP-Attached Substituents and DNA Polymerase Amino Acid Residues*

A negative correlation between the number of non-covalent bonds formed by dUTP-attached substituents with the KlenTaq polymerase amino acid residues and the PCR amplification efficiency in the presence of the modified dUTPs was found by comparing the experimental data with the results of molecular modeling. Roughly, the PCR efficiency decrease is about 15%, with an increase by one non-covalent bond between the substituent and the polymerase. This correlation turns out to be true for all the studied polymerases.

Molecular modeling has revealed various types of non-covalent bonds responsible for the specificity of dNTP C5-substituent-polymerase interactions: hydrogen, carbon–hydrogen,  $\pi$ -donor hydrogen, electrostatic (attractive charge bonds,  $\pi$ -cation,  $\pi$ -anion bonds),  $\pi$ -lone pair bond,  $\pi$ -sulfur bond, and hydrophobic bonds ( $\pi$ - $\pi$ -stacked,  $\pi$ -sigma, alkyl, and  $\pi$ -alkyl bonds). This set of interactions obtained in molecular modeling can be helpful in the choice of structural solutions for new polymerase-specific dNTP derivatives in the future.

Thus the number of non-covalent interactions between the substituents and the polymerase amino acid residues can be a potentially variable parameter for regulating the enzyme activity.

### 3.2. Commonality of Seven Polymerases from the Families A and B to Use the dUTPs Modified with Small or Bulky Aromatic Substituents in PCR

It was found that various polymerases of the A and B families (Taq, Tth, Pfu, Vent (exo-), Deep Vent (exo-), Vent, and Deep Vent) possess a reduced ability (but not too different in values) to incorporate the dUTPs modified with bulky aromatic substituent into the growing DNA strand as compared with natural dTMP (see Figures 2C and 3, and Relationship (3)). This result demonstrates the similarity of all the tested DNA polymerases of the A and B families in the ability to use a substrate that mimics dTTP, the dUTP modified with a certain bulky aromatic substituent.

Our data suggest the similarity of the local environments of dNTP substituents in the active centers of various polymerases. For example, covalent docking used in the present study (see Table S2) show the formation of  $\pi$ -cation bonds between phenyl rings of functional groups R1-R7 of dU\*TP substituents and positively charged amino acid residues Arg587 (in the case of the 5E41 structure) or Lys831 and Arg659 (in the case of the 1QTM structure) of KlenTaq polymerase. Recently, Hocek M. et al. [42], also published the formation of  $\pi$ -cation interaction between the phenyl ring of dG\*TP substituent and Arg629 of another enzyme – Bst polymerase, using docking. Thus, molecular modeling shows that the positively charged polymerase amino acid residues of at least two DNA polymerases form the  $\pi$ -cation bonds with phenyl rings of dU\*TP, or dG\*TP aromatic substituents. The existence of  $\pi$ -cation interactions between the phenyl ring of the dT\*TP or dC\*TP substituent and Arg587 and Lys663 of the KlenTaq polymerase was previously shown by Marx A. et al. in X-ray studies [38].

Moreover, as was published by Marx A. et al. [53], the same substituent of modified dATP (7-deaza-modified adenosine triphosphate) forms three similar non-covalent bonds for both the KOD DNA polymerase and KlenTaq DNA polymerase. In the case of the KOD polymerase, the substituent forms a hydrogen bond and a carbon-hydrogen bond with Lys487 and an alkyl bond with Ile488 (6Q4T [50,53]). In contrast, in the case of the KlenTaq polymerase, the same substituent forms two hydrogen bonds with Thr664 and an alkyl bond with Lys663 (6Q4U [50,53]).

Thus, based on the abovementioned results, as well as on the similarity in the secondary structures of DNA polymerases [34,35], one can assume that, despite the varieties in the amino acid chains forming the active centers of different polymerases, the chemical groups of the amino acid residues facing the spaces of the active centers, as well as those facing the adjacent regions (possibly forming cavities similar to cavities A and B of the KlenTaq polymerase), form similar local environments not only for dNTPs (that is well known), but also for their substituents.

### 3.3. The Role of Non-Covalent Interactions between Low-Molecular-Weight Substituents Attached to the C5 Position of the Pyrimidine ring of dUTPs and the Amino Acid Residues of Taq DNA Polymerase in PCR Efficiency

Marx A. et al. [36,38–40] performed X-ray crystallography for many structures. In the paper [38], he wrote that concerning the modified dNTPs, "hydrogen-bonding capability might improve their substrate properties" and "the introduction of an aromatic ring enables new interactions as cation- $\pi$  interaction to positively charged amino acid side chains like arginine or lysine. This fact might explain the efficient processing of dT\*TP and dC\*TP".

Meanwhile, molecular modeling and comparing it with experimental data lead us to the opposite conclusion: the relative efficiency of PCR or PEX in the presence of modified dNTPs decreases with an increase in the number of non-covalent bonds between the substituent and polymerase. These non-covalent bonds may be hydrogen, carbon-hydrogen,  $\pi$ -donor hydrogen, electrostatic (attractive charge bonds,  $\pi$ -cation,  $\pi$ -anion bonds),  $\pi$ -lone pair bond,  $\pi$ -sulfur bonds, and hydrophobic bonds ( $\pi$ - $\pi$ -stacked,  $\pi$ -sigma, alkyl, and  $\pi$ -alkyl bonds). The data presented in Figure 5A–E for the modified dUTPs may be an example.

### 3.3.1. Incorporation Efficiency of dT<sup>spin</sup>MP and dT<sup>dend</sup>MP and X ray Structural Data Obtained by Marx A. et al. Analyzed in the Discovery Studio Program

Our conclusion concerning the decrease of incorporation efficiency due to the number of non-covalent bonds between the dUTP substituent and polymerase is in agreement with the previous X-ray data and the results of competitive incorporation of modified dNTPs in the presence of their natural counterparts [36–40]. Kinetic experiments published by Marx A. et al. [36] showed that dT<sup>spin</sup>MP (where the small "spin" substituent had a mass of 163 Da) was incorporated by the KlenTaq polymerase with surprisingly less efficiency than dT<sup>dend</sup>MP (where the bulkier "dend" substituent had a mass of 577 Da). It would be interesting to understand why a substituent with a significantly larger mass ("dend") inhibits the inclusion of the corresponding nucleotide to a lesser extent than a substituent with a smaller mass ("spin").

In attempting to explain this apparent discrepancy, we analyzed X-ray structures obtained by Marx A. et al. [36], namely KlenTaq polymerase–DNA–(dT<sup>spin</sup>TP or dT<sup>dend</sup>TP) complexes using the Discovery Studio program [49]. We downloaded the corresponding structure files 3OJU.PDB and 3OJS.PDB from the Protein Data Bank [50] and compared the number of non-covalent bonds formed by "spin" and "dend" substituents with the amino acid residues of the KlenTaq polymerase. It was obtained that two non-covalent carbon–hydrogen bonds were formed by the small "spin" substituent (between the oxygen atom of the "spin" substituent and the carbon atoms of Arg587 and Arg660, respectively). In contrast, only one non-covalent hydrogen bond was formed by the bulky "dend" substituent (between the oxygen atom of the "dend" substituent and the nitrogen atom of Arg587). This fact can be considered in favor of the proposal that a lower number of dNTP substituent–polymerase non-covalent interactions facilitates the enzymatic incorporation of such modified dNMP.

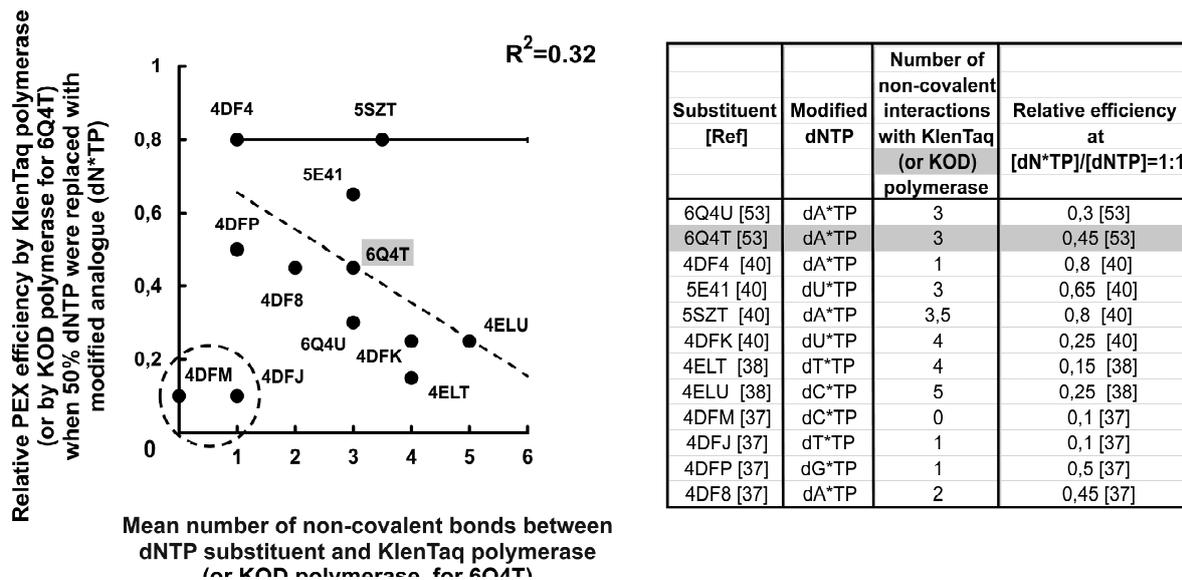
Meanwhile, using the Discovery Studio program, we could ascertain that the "dend" substituent additionally forms five non-covalent bonds with DNA (two hydrogen bonds with a previous nucleotide, dCMP; one  $\pi$ –lone pair bond; and two carbon–hydrogen bonds with the complimentary dAMP). We suppose that these substituent–DNA bonds improve incorporation efficiency at the first step of modified dNTP binding to the enzyme's active center. But at the same time these DNA–substituent interactions do not decrease incorporation efficiency since they do not need to be broken for the subsequent one-step translocation of the polymerase for attaching the next nucleotide.

### 3.3.2. Competitive Incorporation of the Modified dNTPs and the X-ray Structural Data for KlenTaq Polymerase–DNA–(Modified dNTP) Complexes

To more widely test the usefulness of the developed approach and explain the results recently obtained in different works, we used the data from X-ray structural studies published in recent papers for six substituents in the corresponding twelve X-ray structures, 4DFM, 4DMJ, 4DFP, 4DF8 [37], 4ELT, 4ELU [38], 4DFK, 5SZT, 5E41, 4DF4 [40], 6Q4U, and 6Q4T [53], obtained by Marx A. et al. The structure files were downloaded from the Protein Data Bank [50].

Figure 7 shows the relative PEX efficiency values estimated by Marx A. et al. [37,38,40,53] at dN\*TP/dNTP=1:1 (y-axis) plotted against the structure-determined numbers of non-covalent substituent–polymerase bonds (x-axis). Ten out of twelve points on the plot form a negative correlation, while two points corresponding to the structures 4DFM and 4DFJ (circled) are outside. The resulting correlation is characterized by a linear decrease of about 10% per one non-covalent bond, roughly similar to the case of a decrease in the relative efficiency of PCR (15% per one non-covalent bond), as demonstrated in Figure 5A. The corresponding values of the parameters are included in the table near the plot in Figure 7.

Notably, each point of the plot was obtained in independent experiments using different nucleotides modified with different substituents. Therefore, one can conclude that non-covalent interactions between dNTP substituents and polymerases really decrease enzymatic incorporation efficiency.



**Figure 7.** Correlation between the relative PEX efficiency and the number of non-covalent bonds that each of the dNTP substituents forms with amino acid residues of the KlenTaq polymerase [37,38,40,50,53] or KOD polymerase (for 6Q4T [50,53]). PEX efficiency was measured when 50% dNTP was replaced with dNTP modified with low-molecular-weight substituent (82-577 Da) (dN\*TP) [37,38,40,53]. The ordinate axis shows the relative PEX efficiency when 50% dNTP was replaced with dN\*TP using the KlenTaq polymerase [37,38,40,53] or KOD polymerase [53]. The abscissa axis shows the number of non-covalent bonds between dNTP substituents and the KlenTaq or KOD polymerase obtained from X-ray studies [37,38,40,50,53]. For the structure 5SZT, the number of non-covalent bonds is averaged over extensions through cavities A and B, and the bars indicate the mean absolute deviations. The dotted straight line shows linear approximations for the corresponding values. The table shows the plot data.

Thus, it can be concluded that non-covalent interactions between low-molecular-weight substituents of the modified dNTPs and the DNA polymerase stabilize the localization of the modified dNTPs in the active center of the enzyme and increase the efficiency of the phosphodiester bond formation. However, the energy of non-covalent bonds should be overcome by their breaking for the subsequent one-step translocation of the polymerase to attach the next nucleotide.

To complete the entire cycle of enzymatic nucleotide incorporation, the energy loss due to the breaking of non-covalent bonds to continue the polymerase movement along the DNA template is more important than the energy gain of localization of the modified dNTP in the enzyme's active center.

### 3.4. Analysis of Amino Acid Residues of the KlenTaq Polymerase Capable of Forming Non-Covalent Bonds with the dUTP C5-Attached Substituents

The molecular modeling using the Discovery Studio program [49] and X-ray structures 1QTM and 5E41 [33,40,50] revealed 14 KlenTaq polymerase amino acid residues (Glu530, Val586, Arg587, Arg595, Asp610, Phe647, Pro656, Leu657, Arg659, Arg660, Ala661, Thr664, Glu820, and Lys831) that can potentially form non-covalent bonds with the chemical groups of comparatively light (126-251 Da) or rather bulky (536 - 694 Da) aromatic substituents attached to the C5 position of the pyrimidine ring of the dUTPs localized in the active center of the enzyme.

Amino acid residues Arg587, Arg659, Arg660, Ala661, and Thr664 belong to the active center of the enzyme [32,36-41] while Glu530 and Phe647 extend beyond the active center facing cavity A while Val586, Arg595, Asp610, Pro656, Leu657, Glu820, and Lys831 extend beyond the active center facing cavity B. The non-covalent interactions with 4 of the 14 amino acid residues (Arg587, Arg660, Ala661, and Thr664) were detected experimentally and published previously [36-41]. The ten amino

acid residues (Glu530, Val586, Arg595, Asp610, Phe647, Pro656, Leu657, Arg659, Glu820, and Lys831) are first predicted in this paper as capable of forming non-covalent bonds with dNTP substituents. The abovementioned amino acid residues can be used as a target in directed mutagenesis to create new polymerases with desirable specificity to the modified dNTP substituents.

#### 4. Materials and Methods

##### 4.1. dUTPs Modified at C5-Position with Bulky or Small Aromatic Substituents

The chemical structures of dUTPs C5-modified with bulky aromatic substituents (536 - 694 Da, including linker -CH=CH-CH<sub>2</sub>-NHCO-(CH<sub>2</sub>)<sub>5</sub>- and bulky aromatic groups which are Cy5<sup>+</sup>, Cy5<sup>±</sup>, or Cy5<sup>-</sup>) or with lighter hydrocarbon substituents R0-R7 (126-251 Da, including linker -CH=CH-CH<sub>2</sub>-NHCO-CH<sub>2</sub>- and -CH<sub>2</sub>-CH<sub>3</sub> or small aromatic groups) are shown in Figure 1. The synthesis and spectroscopic characteristics of the corresponding modified dUTPs (dU(Cy5<sup>±</sup>)TP, dU(Cy5<sup>+</sup>)TP, and dU(Cy5<sup>-</sup>)TP), as well as PCR efficiency of dUTPs modified with lighter low-molecular-weight substituents R0-R7, were described earlier [28,47].

##### 4.2. The DNA Template and Primers.

The 68-nucleotide DNA template and 18- and 17-nucleotide-long primers labeled with Cy3 dye were the same as those used in [28]:

Template:

5' - TCTCTTGCCCTTTCGTCTCTAAATTGTCTTAATCTCTTCTATCCTTCTCTCACCACCTT  
ACATCCGC - 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. Adenines subject to complementary binding with modified deoxyuridines of the enzymatically synthesized DNA strand are underlined by wavy lines. Thymines that can be replaced with modified deoxyuridines during PCR are shown italic. The embedding of modified nucleotides complementary to the template is significantly difficult in sites containing clusters of complementary bases. Therefore, the sequence of the DNA template specifically included both the isolated A and isolated AA and AAA repeats to study the ability of the polymerase to use modified dUTPs complementary to these A-repeats. In addition, the template sequence was chosen to avoid the formation of self-complementary structures as much as possible.

##### 4.3. PCR and Electrophoresis

PCR was performed following a protocol similar to the one described previously [28]. DNA polymerases of different organisms were used, namely

*Thermus aquaticus* (Taq),

*Thermus thermophilus* (Tth),

*Pyrococcus furiosus* (Pfu) DNA polymerases (Sileks, Badenweiler, Germany), as well as

*Thermococcus litoralis* (Vent),

*Pyrococcus* GB-D strain (Deep Vent) DNA polymerases, and forms of the native DNA polymerases from

*Thermococcus litoralis* (Vent (exo-)) and

*Pyrococcus* GB-D strain (Deep Vent (exo-)) that had been genetically engineered to eliminate the 3'→5' proofreading exonuclease activity (New England Biolabs, Ipswich, UK).

Each polymerase's PCR conditions were selected for efficient amplification with natural dNTPs. Herein, the amount of PCR products amplified with the polymerases was the same for all the considered polymerases within an accuracy of 25%, which allowed for facile observation, registration, and comparison of the electrophoregrams.

The PCR conditions for each of the polymerases and the detailed protocols for PCR amplification, gel electrophoresis, and obtaining gel images are presented in "Part G: PCR and electrophoresis" in Supplementary data.

#### 4.4. Quantitative Analysis of Electrophoretic Bands Containing PCR-Amplified Full-Length DNA Fragments

The relative quantities of the PCR-amplified full-length DNA fragments containing modified dUMPs were estimated by the relative fluorescence intensity of the Cy3 dye conjugated to oligonucleotide primers ( $\lambda^{\text{abs}}_{\text{max}}=551$  nm,  $\lambda^{\text{em}}_{\text{max}}=567$  nm [47]). These quantities were calculated separately for each of the electrophoregram lanes. For this purpose, the summed Cy3 fluorescence intensity of the bands containing full-length DNA fragments with incorporated modified dUMPs was quantified using a virtual rectangular frame that surrounded the bands. The fluorescence intensities of all the pixels surrounded by the frame were summed to obtain the total fluorescence intensity within the frame. The total fluorescence intensity of the blank gel region within the same frame of the same size was then subtracted from the obtained value.

To obtain the relative values of PCR efficiency, the fluorescence intensity of the DNA products containing incorporated modified dUMPs was normalized to the product obtained in the presence of only natural dNTPs.

The relative fluorescence intensity of the Cy5 labeled dUMPs incorporated into DNA strands ( $\lambda^{\text{abs}}_{\text{max}}=648$  nm,  $\lambda^{\text{em}}_{\text{max}}=670$  nm [48]) also characterized the relative quantity of the PCR-amplified full-length DNA fragments. Cy5 fluorescence intensity of the bands containing full-length DNA product in each lane was summarized with the same approach, which was used to calculate PCR efficiency in the Cy3 fluorescence range. The obtained quantities were normalized to the maximum value.

#### 4.5. Computer Modeling of 3D Structures of Triple Complexes Consisting of KlenTaq Polymerase–DNA–(Modified dUTP)

3D molecular modeling was carried out using the Small Molecules and Receptor–Ligand Interactions modes found in the Discovery Studio program [49] and based on known X-ray structures [33,40,50] to localize the substituents at the C5 position of the pyrimidine ring of dUTPs in the active center of the KlenTaq polymerase.

3D structures of the complexes (1QTM and 5E41) determined by X-ray crystallography [33,40,50] were rigidly fixed. The structures of the corresponding deoxy pyrimidine triphosphates localized in the active center were also rigidly fixed in accordance with the covalent docking procedure. Water molecules were deleted. Structural solutions were obtained as a result of optimizing the energy of the conformation of the substituents (linker–functional groups, see Figure 1) and the energy of its interactions with polymerase amino acid residues and DNA.

The types and quantities of the detected non-covalent bonds between the substituent and polymerase were compared with the experimental data regarding the efficiency of PCR amplification in the presence of the corresponding modified dUTPs in the PCR mixture to find possible correlations.

#### 4.6. KlenTaq Polymerase Amino Acid Residues' Ability to Form Non-Covalent Interactions with the dUTP Substituents Attached at the C5 Position of the Pyrimidine Ring Localized in the Active Center of the Enzyme

The capability of KlenTaq polymerase amino acid residues to non-covalently interact with the substituents attached at the C5 position of the pyrimidine ring of dUTP localized in the active center of the enzyme was estimated quantitatively. The mean number of non-covalent bonds formed by a particular amino acid residue per one structural solution,  $n$ , was calculated as

$$n = N / S$$

where N is the number of non-covalent bonds formed between a particular amino acid residue and the substituents by summing over all the structural solutions. S is the total number of structural solutions obtained in modeling.

## 5. Conclusions

The PCR efficiency in the presence of dUTPs modified with low-molecular-weight substituents decreases with an increase in the number of non-covalent bonds between the substituents and the DNA polymerase (about 15% decrease per one extra non-covalent bond). The number of non-covalent bonds between the substituents and polymerase amino acid residues can be used as a variable parameter for regulating enzyme activity. Combined with molecular modeling using the highly efficient Discovery Studio software, this approach can be used to design and synthesize dNTP derivatives characterized by new polymerase-specific properties.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

**Author Contributions:** Conceptualization, A.S.Z.; Methodology, O.A.Z., S.A.S., V.E.K. and V.E.S.; Investigation, O.A.Z. and S.A.S.; Data Curation, O.A.Z.; Writing – Original Draft Preparation, O.A.Z.; Writing—review and editing, O.A.Z., A.V.C. and V.E.B.; Supervision, A.V.C.; Funding Acquisition, A.V.C.. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Russian Science Foundation, grant number 20-14-00287.

**Acknowledgments:** We thank Dr. Proshkin S.A. (EIMB RAS) for a helpful discussion and valuable comments. The authors are grateful to BIOVIA and Dassault Systèmes for providing Discovery Studio software for free.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Walsh, J.M.; Beuning, P.J. Synthetic Nucleotides as Probes of DNA Polymerase Specificity. *J. Nucleic Acids.* **2012**, *2012*, 530963.
2. Hollenstein, M. Nucleoside triphosphates—building blocks for the modification of nucleic acids. *Molecules.* **2012**, *17*, 13569-13591.
3. Wu, W.-J.; Yang, W.; Tsai, M.-D. How DNA polymerases catalyse replication and repair with contrasting fidelity. *Nat. Rev. Chem.* **2017**, *1*, 0068.
4. Perlíková, P.; Hocek, M. Pyrrolo[2,3-d]pyrimidine (7-deazapurine) as a privileged scaffold in design of antitumor and antiviral nucleosides. *Med Res Rev.* **2017**, *37*, 1429-1460.
5. Tokarenko, A.; Lišková, B.; Smoleň, S.; Tábořská, N.; Tichý, M.; Gurská, S.; Perlíková, P.; Frydrych, I.; Tloušť'ová, E.; Znojek, P.; Mertlíková-Kaiserová, H.; Poštová Slavětínská, L.; Pohl, R.; Klepetářová, B.; Khalid, N.U.; Wenren, Y.; Laposa, R.R.; Džubák, P.; Hajdúch, M.; Hocek, M. Synthesis and Cytotoxic and Antiviral Profiling of Pyrrolo- and Furo-Fused 7-Deazapurine Ribonucleosides. *J Med Chem.* **2018**, *61*, 9347-9359.
6. Tichý, M.; Pohl, R.; Xu, H.Y.; Chen, Y.L.; Yokokawa, F.; Shi, P.Y.; Hocek, M. Synthesis and antiviral activity of 4,6-disubstituted pyrimido[4,5-b]indole ribonucleosides. *Bioorg Med Chem.* **2012**, *20*, 6123-6133.
7. Fleuti, M.; Bártová, K.; Slavětínská, L.P.; Tloušť'ová, E.; Tichý, M.; Gurská, S.; Pavliš, P.; Džubák, P.; Hajdúch, M.; Hocek, M. Synthesis and Biological Profiling of Pyrazolo-Fused 7-Deazapurine Nucleosides. *J Org Chem.* **2020**, *85*, 10539-10551.
8. Gold, L.; Ayers, D.; Bertino, J.; Bock, C.; Bock, A.; Brody, E.N.; Carter, J.; Dalby, A.B.; Eaton, B.E.; Fitzwater, T.; Flather, D.; Forbes, A.; Foreman, T.; Fowler, C.; Gawande, B.; Goss, M.; Gunn, M.; Gupta, S.; Halladay, D.; Heil, J.; Heilig, J.; Hicke, B.; Husar, G.; Janjic, N.; Jarvis, T.; Jennings, S.; Katilius, E.; Keeney, T.R.; Kim, N.; Koch, T.H.; Kraemer, S.; Kroiss, L.; Le, N.; Levine, D.; Lindsey, W.; Lollo, B.; Mayfield, W.; Mehan, M.; Mehler, R.; Nelson, S.K.; Nelson, M.; Nieuwlandt, D.; Nikrad, M.; Ochsner, U.; Ostroff, R.M.; Otis, M.; Parker, T.; Pietrasiewicz, S.; Resnicow, D.I.; Rohloff, J.; Sanders, G.; Sattin, S.; Schneide, D.; Singer, B.; Stanton, M.; Sterkel, A.; Stewart, A.; Stratford, S.; Vaught, J.D.; Vrkljan, M.; Walker, J.J.; Watrobka, M.; Waugh, S.; Weiss, A.; Wilcox, S.K.; Wolfson, A.; Wolk, S.K.; Zhang, C.; Zichi, D. Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One.* **2010**, *5*, e15004.
9. Kuwahara, M.; Sugimoto, N. Molecular evolution of functional nucleic acids with chemical modifications. *Molecules.* **2010**, *15*, 5423-5444.

10. Gold, L.; Walker, J.J.; Wilcox, S.K.; Williams, S. Advances in human proteomics at high scale with the SOMAscan proteomics platform. *N Biotechnol.* **2012**, *29*, 543-549.
11. Famulok, M.; Mayer, G. Aptamers and SELEX in Chemistry & Biology. *Chem Biol.* **2014**, *21*, 1055-1058.
12. Drolet, D.W.; Green, L.S.; Gold, L.; Janjic, N. Fit for the Eye: Aptamers in Ocular Disorders. *Nucleic Acid Ther.* **2016**, *26*, 127-146.
13. Gawande, B. N.; Rohloff, J. C.; Carter, J. D.; von Carlowitz, I.; Zhang, C.; Schneider, D. J.; Janjic, N. Selection of DNA aptamers with two modified bases. *Proc. Natl. Acad. Sci.* **2017**, *114*, 2898-2903.
14. Ladju, R.B.; Pascut, D.; Massi, M.N.; Tiribelli, C.; Sukowati, C.H. Aptamer: A potential oligonucleotide nanomedicine in the diagnosis and treatment of hepatocellular carcinoma. *Oncotarget.* **2017**, *9*, 2951-2961.
15. Zhu, G.; Chen, X. Aptamer-based targeted therapy. *Adv. Drug. Deliv. Rev.* **2018**, *134*, 65-78.
16. Rose, K.M.; Alves Ferreira-Bravo, I.; Li, M.; Craigie, R.; Ditzler, M.A.; Holliger, P.; DeStefano, J.J. Selection of 2'-Deoxy-2'-Fluoroarabino Nucleic Acid (FANA) Aptamers That Bind HIV-1 Integrase with Picomolar Affinity. *ACS Chem Biol.* **2019**, *14*, 2166-2175.
17. Minagawa, H.; Kataoka, Y.; Kuwahara, M.; Horii, K.; Shiratori, I.; Waga, I. A high affinity modified DNA aptamer containing base-appended bases for human  $\beta$ -defensin. *Anal Biochem.* **2020**, *594*, 113627.
18. Minagawa, H.; Kataoka, Y.; Fujita, H.; Kuwahara, M.; Horii, K.; Shiratori, I.; Waga, I. Modified DNA Aptamers for C-Reactive Protein and Lactate Dehydrogenase-5 with Sub-Nanomolar Affinities. *Int J Mol Sci.* **2020**, *21*, 2683.
19. Schmitz, A.; Weber, A.; Bayin, M.; Breuers, S.; Fieberg, V.; Famulok, M.; Mayer, G. A SARS-CoV-2 Spike Binding DNA Aptamer that Inhibits Pseudovirus Infection by an RBD-Independent Mechanism. *Angew Chem Weinheim Bergstr Ger.* **2021**, *133*, 10367-10373.
20. McKenzie, L.K.; El-Khoury, R.; Thorpe, J.D.; Damha, M.J.; Hollenstein M. Recent progress in non-native nucleic acid modifications. *Chem Soc Rev.* **2021**, *50*, 5126-5164.
21. Jager, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. A versatile toolbox for variable DNA functionalization at high density. *J. Am. Chem. Soc.* **2005**, *127*, 15071-15082.
22. Kuwahara, M.; Nagashima, J.; Hasegawa, M.; Tamura, T.; Kitagata, R.; Hanawa, K.; Hososhima, S.; Kasamatsu, T.; Ozaki, H.; Sawai, H. 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.* **2006**, *34*, 5383-5394.
23. Vaught, J.D.; Bock, C.; Carter, J.; Fitzwater, T.; Otis, M.; Schneider, D.; Rolando, J.; Waugh, S.; Wilcox, S.K.; Eaton, B.E. Expanding the chemistry of DNA for in vitro selection. *J. Am. Chem. Soc.* **2010**, *132*, 4141-4151.
24. Hocek M. 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.* **2014**, *79*, 9914-9921.
25. Hollenstein M. Generation of long, fully modified, and serum-resistant oligonucleotides by rolling circle amplification. *Org Biomol Chem.* **2015**, *13*, 9820-9824.
26. Ménéová, P.; Cahová, H.; Vrábek, M.; Hocek, M. Synthesis of Base-Modified dNTPs Through Cross-Coupling Reactions and Their Polymerase Incorporation to DNA. *Methods Mol Biol.* **2019**, *1973*, 39-57.
27. Röthlisberger, P.; Levi-Acobas, F.; Leumann, C.J.; Hollenstein, M. Enzymatic synthesis of biphenyl-DNA oligonucleotides. *Bioorg Med Chem.* **2020**, *28*, 115487.
28. Zasedateleva, O.A.; Surzhikov, S.A.; Shershov, V.E.; Miftakhov, R.A.; Yurasov, D.A.; Kuznetsova, V.E.; Chudinov, A.V. PCR incorporation of dUMP 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. *Bioorg Chem.* **2020**, *99*, 103829.
29. Ondruš, M.; Sýkorová, V.; Bednárová, L.; Pohl, R.; Hocek, M. Enzymatic synthesis of hypermodified DNA polymers for sequence-specific display of four different hydrophobic groups. *Nucleic Acids Res.* **2020**, *48*, 11982-11993.
30. Ollis, D.L.; Brick, P.; Hamlin, R.; Xuong, N.G.; Steitz T.A. Structure of the large fragment of Escherichia coli DNA polymerase I complexed with TMP. *Nature.* **1985**, *313*, 762-766.
31. Kim, Y.; Eom, S.H.; Wang, J.; Lee, D.S.; Suh, S.W.; Steitz, T.A. Crystal structure of Thermus aquaticus DNA polymerase. *Nature.* **1995**, *376*, 612-616.
32. Li, Y.; Korolev, S.; Waksman, G. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J.* **1998**, *17*, 7514-7525.
33. Li, Y.; Mitaxov, V.; Waksman G. Structure-based design of Taq DNA polymerases with improved properties of dideoxynucleotide incorporation. *Proc. Natl. Acad. Sci. USA.* **1999**, *96*, 9491-9496.
34. Steitz, T.A. DNA Polymerases: Structural Diversity and Common Mechanisms. *J Biol Chem.* **1999**, *274*, 17395-17398.

35. Rothwell, P.J.; Waksman, G. Structure and mechanism of DNA polymerases, in *Advances in protein chemistry*; Squire, J.M.; Parry, D.A.D. (Eds.); Elsevier, Amsterdam, Netherlands, 2005, Volume 71, pp. 401-440.
36. Obeid, S.; Baccaro, A.; Welteb, W.; Diederichs, K.; Marx, A. Structural basis for the synthesis of nucleobase modified DNA by *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. USA*. **2010**, *107*, 21327–21331.
37. Bergen, K.; Steck, A.-L.; Strütt, S.; Baccaro, A.; Welte, W.; Diederichs, K.; Marx, A. Structures of KlenTaq DNA polymerase caught while incorporating C5-modified pyrimidine and C7-modified 7-deazapurine nucleoside triphosphates. *J Am Chem Soc*. **2012**, *134*, 11840-11843.
38. Obeid, S.; Buskamp, H.; Welte, W.; Diederichs, K.; Marx, A. Interactions of non-polar and "Click-able" nucleotides in the confines of a DNA polymerase active site. *Chem Commun (Camb)*. **2012**, *48*, 8320-8322.
39. Hottin, A.; Marx, A. Structural Insights into the processing of nucleobase-modified nucleotides by DNA polymerases. *Acc Chem Res*. **2016**, *49*, 418-427.
40. Hottin, A.; Betz, K.; Diederichs, K.; Marx, A. Structural Basis for the KlenTaq DNA Polymerase Catalysed Incorporation of Alkene- versus Alkyne-Modified Nucleotides. *Chemistry*. **2017**, *23*, 2109-2118.
41. Kropp, H.M.; Durr, S.L.; Peter, C.; Diederichs, K.; Marx, A., Snapshots of a modified nucleotide moving through the confines of a DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 9992–9997.
42. Cahová, H.; Panattoni, A.; Kielkowski, P.; Jindřich Fanfrlík, J.; Hocek, M. 5-Substituted Pyrimidine and 7-Substituted 7-Deazapurine dNTPs as Substrates for DNA Polymerases in Competitive Primer Extension in the Presence of Natural dNTPs. *ACS Chem. Biol.* **2016**, *11*, 3165–3171.
43. Ramsay, N.; Jemth, A.S.; Brown, A.; Crampton, N.; Dear, P.; Holliger, P. CyDNA: synthesis and replication of highly Cy-dye substituted DNA by an evolved polymerase. *J. Am. Chem. Soc.* **2010**, *132*, 5096–5104.
44. Wynne, S.A.; Pinheiro, V.B.; Holliger, P.; Leslie, A.G. Structures of an apo and a binary complex of an evolved archeal B family DNA polymerase capable of synthesizing highly Cy-dye labelled DNA. *PLoS One*. **2013**, *8*, e70892.
45. Ivancová, I.; Leone, D.L.; Hocek, M. Reactive modifications of DNA nucleobases for labelling, bioconjugations, and cross-linking. *Curr Opin Chem Biol*. **2019**, *52*, 136-144.
46. Kuba, M.; Kraus, T.; Pohl, R.; Hocek, M. Nucleotide-Bearing Benzylidene-Tetrahydroxanthylum Near-IR Fluorophore for Sensing DNA Replication, Secondary Structures and Interactions. *Chemistry*. **2020**, *26*, 11950-11954.
47. Zasedateleva, O.A.; Vasiliskov, V.A.; Surzhikov, S.A.; Kuznetsova, V.E.; Shershov, V.E.; Guseinov, T.O.; Smirnov, I.P.; Yurasov, R.A.; Spitsyn, M.A.; Chudinov, A.V. dUTPs conjugated with zwitterionic Cy3 or Cy5 fluorophore analogues are effective substrates for DNA amplification and labelling by Taq polymerase. *Nucleic Acids Res*. **2018**, *46*, e73.
48. Mader, O.; Reiner, K.; Egelhaaf, H.J.; Fischer, R.; Brock, R. Structure property analysis of pentamethine indocyanine dyes: identification of a new dye for life science applications. *Bioconjug. Chem.* **2004**, *15*, 70–78.
49. BIOVIA, Dassault Systèmes, Discovery Studio, 2020, San Diego: Dassault Systèmes, 2020. Available online: <https://discover.3ds.com/discovery-studio-visualizer-download> (accessed on 31 July 2023)
50. <https://www.rcsb.org> Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N. and Bourne P.E. The Protein Data Bank. *Nucleic Acids Res*. **2000**, *28*, 235-242.
51. Shershov, V.E.; Kuznetsova, V.E.; Lysov, Y.P.; Guseinov, T.O.; Barsky, V.E.; Spitsyn, M.A.; Zasedateleva, O.A.; Vasiliskov, V.A.; Surzhikov, S.A.; Zasedatelev, A.S.; Chudinov, A.V. The effect of chromophore charge on the incorporation efficiency of fluorescence-labeled nucleotides catalyzed by Taq DNA polymerase in template synthesis. *Biophysika*, **2015**, *60*, 1013–1015.
52. Aschenbrenner, J.; Marx, A. DNA polymerases and biotechnological applications. *Curr Opin Biotechnol*. **2017**, *48*, 187-195.
53. Kropp, H.M.; Diederichs, K.; Marx, A. The Structure of an Archaeal B-Family DNA Polymerase in Complex with a Chemically Modified Nucleotide. *Angew Chem Int Ed Engl*. **2019**, *58*, 5457-5461.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.