



Communication

Method for Rapid Analysis of Mutant RNA Polymerase Activity on Templates Containing Unnatural Nucleotides

Tatiana Egorova ¹, Ekaterina Shuvalova ¹, Sabina Mukba ¹, Alexey Shuvalov ¹, Peter Kolosov ^{2,*} 
and Elena Alkalaeva ^{1,*} 

¹ Engelhardt Institute of Molecular Biology, The Russian Academy of Sciences, 119991 Moscow, Russia; tatvladegorova@gmail.com (T.E.); hritova_katia@mail.ru (E.S.); sabina.mukba1996@gmail.com (S.M.); laursen1243@mail.ru (A.S.)

² Institute of Higher Nervous Activity and Neurophysiology, The Russian Academy of Sciences, 117485 Moscow, Russia

* Correspondence: kolosov@ihna.ru (P.K.); alkalaeva@eimb.ru (E.A.);
Tel.: +7-495-334-4113 (P.K.); +7-499-135-9977 (E.A.)

Abstract: Pairs of unnatural nucleotides are used to expand the genetic code and create artificial DNA or RNA templates. In general, an approach is used to engineer orthogonal systems capable of reading codons comprising artificial nucleotides; however, DNA and RNA polymerases capable of recognizing unnatural nucleotides are required for amplification and transcription of templates. Under favorable conditions, in the presence of modified nucleotide triphosphates, DNA polymerases are able to synthesize unnatural DNA with high efficiency; however, the currently available RNA polymerases reveal high specificity to the natural nucleotides and may not easily recognize the unnatural nucleotides. Due to the absence of simple and rapid methods for testing the activity of mutant RNA polymerases, the development of RNA polymerase recognizing unnatural nucleotides is limited. To fill this gap, we developed a method for rapid analysis of mutant RNA polymerase activity on templates containing unnatural nucleotides. Herein, we optimized a coupled cell-free translation system and tested the ability of three unnatural nucleotides to be transcribed by different T7 RNA polymerase mutants, by demonstrating high sensitivity and simplicity of the developed method. This approach can be applied to various unnatural nucleotides and can be simultaneously scaled up to determine the activity of numerous polymerases on different templates. Due to the simplicity and small amounts of material required, the developed cell-free system provides a highly scalable and versatile tool to study RNA polymerase activity.



Citation: Egorova, T.; Shuvalova, E.; Mukba, S.; Shuvalov, A.; Kolosov, P.; Alkalaeva, E. Method for Rapid Analysis of Mutant RNA Polymerase Activity on Templates Containing Unnatural Nucleotides. *Int. J. Mol. Sci.* **2021**, *22*, 5186. <https://doi.org/10.3390/ijms22105186>

Academic Editor: Kensaku Sakamoto

Received: 26 April 2021

Accepted: 10 May 2021

Published: 14 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: unnatural nucleotides; T7 RNAP; cell-free translation

1. Introduction

The genetic code can be modified at different levels, namely nucleotides, codons, and amino acids. At each of these levels, unnatural components can be used. Thus, unnatural amino acids can be encoded by unnatural nucleotides within the three-letter codons [1–5], four-letter codons [6], or by reassignment of sense or stop codons [7–9].

The development of unnatural base pairs (UBPs) is one of the most remarkable advances in synthetic biology. UBPs, which function as a third base pair in replication, transcription, and translation, expand the genetic alphabet and can be used in various applications such as site-specific fluorescence labeling, nucleic acid immobilization, and DNA and RNA analysis. Originally synthesized to expand the genetic alphabet, UBPs have become powerful tools for monitoring the structural changes in nucleic acids and their interactions with other molecules [10–22].

All UBPs obtained to date can be divided into two groups: (1) UBPs, which form hydrogen bonds between bases (hydrogen complementarity) and (2) UBPs with hydrophobic interactions within the pair (structural complementarity) [23–36]. The first synthesized

UBPs comprised hydrogen bonds. Their topology differed from that found in natural nucleotides, which provided specificity for enzymatic synthesis [23–31]. Later, hydrophobic UBPs were obtained [32–36]. This proved that hydrogen bonding is not the only mechanism capable of storing and transmitting genetic information. Presently, both types of UBPs are being actively developed and improved; however, for the amplification and transcription of templates containing unnatural nucleotides, the presence of DNA and RNA polymerases capable of their recognition is required. Under favorable conditions, in presence of modified nucleotide triphosphates, DNA polymerases synthesize unnatural DNA with high efficiency [17,37–41]; then with RNA polymerases, everything is not so good. At present, the available RNA polymerases for *in vitro* studies are specific to natural nucleotides and do not recognize unnatural nucleotides [31,42–45].

Bacteriophage T7 RNA polymerase (T7 RNAP) is one of the widely used polymerases for incorporating UBPs into RNA. It is a one subunit protein with 883 amino acid residues (98 kDa). T7 RNAP is capable of performing a full transcription cycle without involving any additional protein factors [46]. For a better understanding of the mechanisms underlying transcription, experiments on T7 RNA polymerase mutagenesis were carried out. As a result of this mutational analysis, it was found, for example, that the Y639F mutant has an amazing ability to use both rNTP and dNTP as substrates [47]. The additional mutation S641A markedly enhanced this property; therefore, the “double” mutant Y639F/S641A was able to synthesize long mixed polynucleotides, where one, two, or three types of rNTPs were replaced by the corresponding dNTPs [48,49]. Moreover, mutant RNA polymerase is used in the development of synthetic nucleotides. For example, some mutant variants of T7 RNAP, such as Y639F, Y639F/H784A, and G542V/H772R/H784S, were obtained to increase the efficiency of incorporating pyrimidine-2'-modified nucleoside triphosphates used for site-specific labeling of RNA [50].

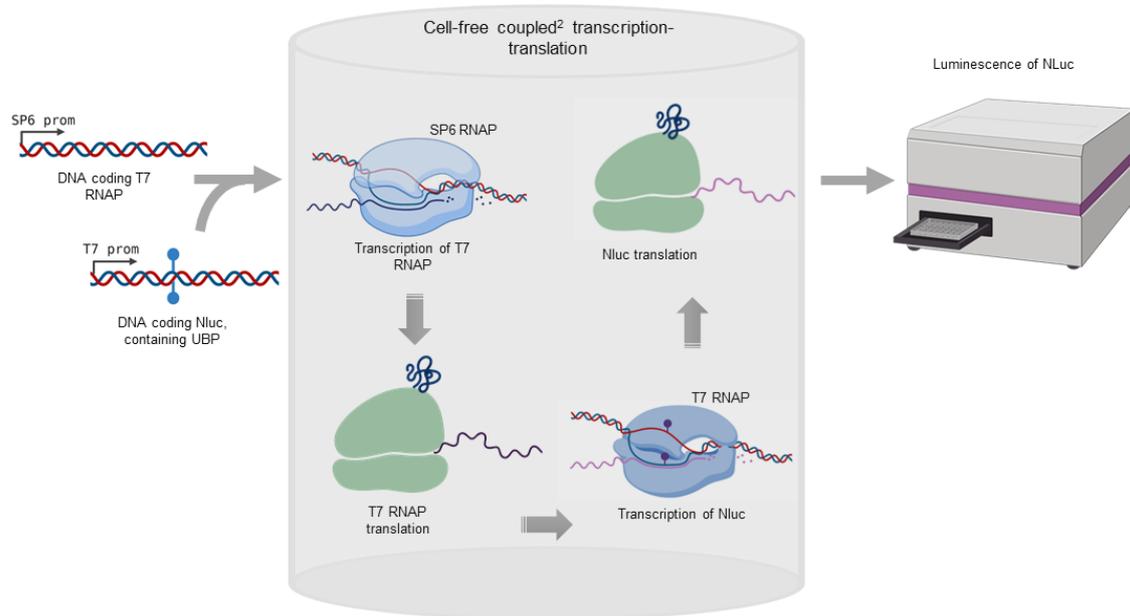
Nevertheless, presently, no simple methods are available for determining the activity of mutant forms of RNA polymerase, synthesizing RNA on a template containing UBPs. In the present study, we proposed a new method based on a double-coupled (coupled²) cell-free transcription–translation system, in which we determined the activity of mutant T7 RNA polymerases by the luminescence of synthesized nanoluciferase (Nluc). As a result, we evaluated the transcription efficiency of unnatural templates with different mutants of T7 RNA polymerases. In our system mutant T7 RNAPs used natural ribonucleotides to transcribe DNA templates containing UBPs. Furthermore, we revealed high sensitivity and simplicity of the developed method and demonstrated the ability of the studied mutants of T7 RNA polymerase to recognize UBPs. Additionally, we demonstrated that the approach we developed for determining the activity of RNA polymerases is applicable to a wide range of UBPs.

2. Results

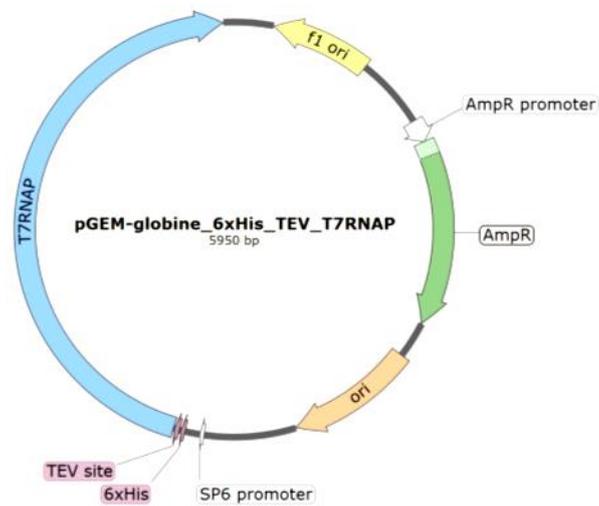
2.1. Overview of the Method

For simple and efficient testing of the functional activity of mutant T7 RNA polymerases, we developed a method of a coupled² transcription–translation system. The scheme of the method is illustrated in Figure 1. SP6 RNA polymerase together with a linear template for T7 RNA polymerase containing the SP6 promoter and linear template for Nluc containing the T7 promoter was added to a wheat germ extract (WGE) cell-free system for coupled transcription and translation. NanoGlo, a substrate for Nluc, was also added to the reaction. First, SP6 RNA polymerase from the SP6 promoter synthesized mRNA for translation of T7 RNA polymerase, thereby producing T7 RNAP. The newly formed T7 RNAP enzyme transcribed Nluc mRNA from the T7 promoter, and thereafter, Nluc was translated in the same lysate leading to the formation of a luminescent product. This allowed us to determine the activity of T7 RNA polymerase in one test tube, starting from the DNA encoding it, without using difficult methods to obtain and purify RNA and proteins. The addition of the Nluc DNA template containing UBPs allowed us to study

the ability of T7 RNAP to use an unnatural template to synthesize mRNA with natural ribonucleotide triphosphates.



(a)



(b)

Figure 1. Cell-free “coupled²” transcription–translation system. (a) Scheme of the cell-free “coupled²” transcription–translation system. (b) Map of the plasmid encoding the human beta-globin 5’UTR, the T7 RNAP coding sequence containing a 6xHis tag, and a subsequent TEV protease site at the N-terminus.

We selected Nluc as a reporter protein for determining the functional activity of mutant T7 RNA polymerases for the following reasons: (1) luciferases induce luminescence in the presence of substrate immediately after their translation because they do not need a maturation time and they leave the ribosome in a catalytically active form [51]; (2) among all known luciferases, we selected Nluc, since it not only has three times smaller size than the widely used firefly and Renilla luciferases, but also causes 100 times higher luminescence

levels than classical luciferases [52]. Thus, the RNA of this protein is translated faster than that of other luciferases, while the method for determining the activity of mutant T7 polymerases is more sensitive by two orders of magnitude, which is crucial for our task.

2.2. Testing of T7 RNAP Mutants on a Natural Template

To investigate the transcriptional efficiency of T7 RNAP, we constructed a plasmid encoding the human beta-globin 5'UTR, T7 RNAP coding sequence containing a 6xHis tag, and a subsequent TEV protease site at the N-terminus (Figure 1b). The globin leader provides an efficient translation of mRNA even in the absence of a 5'-cap. The 6xHis tag allows the isolation of T7 RNA polymerase from the lysate. The mRNA obtained from this construct was tested in various lysates and revealed high activity (data not shown).

Along with wild-type T7 RNA polymerase (T7wt RNAP), we obtained by site-directed mutagenesis and tested the functional activity of 10 previously published mutant T7 RNAPs with different specificities (Table 1). As a negative control in our experiments, we used the mutant T7 RNAP containing Leu229UAG replacement (Δ T7 RNAP). Such polypeptide can be synthesized in a cell-free system but cannot transcribe RNA.

Table 1. Mutants of T7 RNAP, tested in the study.

Mutant	Transcription Activity	References
P266L	Increased thermal stability and promoter clearance	[53]
N433Q	Increased thermal stability	[54,55]
G542V	Involved in interactions with the 2'-hydroxyl moiety of ribonucleotides	[50]
R627S	Increased thermal stability, interactions with phosphate groups of NTPs, incorporation of ddNTP	[54,55]
S633P	Increased thermal stability	[54–56]
H772A		[56]
F849I	Increased thermal stability	[54–56]
F849A	Increased thermal stability	[54–56]
F849Y	Increased thermal stability	[54–56]
FAL: P266L/Y639F/H784A	Increased thermal stability, can accept 2'-O-methyl triphosphates	[31,54,57]

For the initial testing of the approach, we first used a Nluc coding template containing only natural dNTPs. To determine the optimal conditions for the coupled² system, we added 1–10 ng/ μ L of linear DNA coding T7 RNAP to TNT WGE along with ribonucleotide triphosphates (NTPs), SP6 RNA polymerase, and 1 ng/ μ L DNA encoding Nluc. We observed the maximum luminescence signal using 10 ng/ μ L linear DNA coding T7 RNAP. The optimized concentrations of the coupled² system are summarized in Table 2.

Table 2. Components of the coupled² system.

Components	Final Concentration
TNT [®] Wheat Germ Extract (Promega, Madison, WI, USA)	50%
TNT buffer (Promega, Madison, WI, USA) 25 \times	1 \times
Amino Acid mix (Promega, Madison, WI, USA) 1 mM	0.02 mM
RiboLock RNase Inhibitor (Thermo Fisher Scientific, Waltham, MA, USA) 40 U/ μ L	0.8 U/ μ L
Nluc PCR product 10 ng/ μ L	1 ng/ μ L
SP6 RNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) 200 U/ μ L	16 U/ μ L
T7 Polymerase PCR product	10 ng/ μ L
Nano-Glo [®] (Promega, Madison, WI, USA)	1%

After optimizing the experimental conditions, we characterized 10 mutant T7 RNAPs (Table 1) by their ability to transcribe Nluc RNA on a natural DNA template. Following our protocol (Table 2), we incubated linear DNA templates for mutant T7 polymerases in a coupled² system. The T7 RNAP translation was controlled by western blotting with anti-His-tag antibodies (Figure 2a). We have shown that in all samples where PCR products encoding T7wt RNAP and T7 RNAP mutants were added, 101.1 kDa proteins were detected, corresponding to the estimated size of T7 RNAP with tags. In the sample containing the DNA of Δ T7 RNAP, comprising the first 228 codons, the corresponding 28.1 kDa protein

product was detected. In the assay where mQ was added instead of T7 RNAP DNA, only proteins nonspecifically interacting with anti-His-tag antibodies were detected. Thus, we have proved that T7wt RNAP and its mutants are specifically synthesized in a cell-free system.

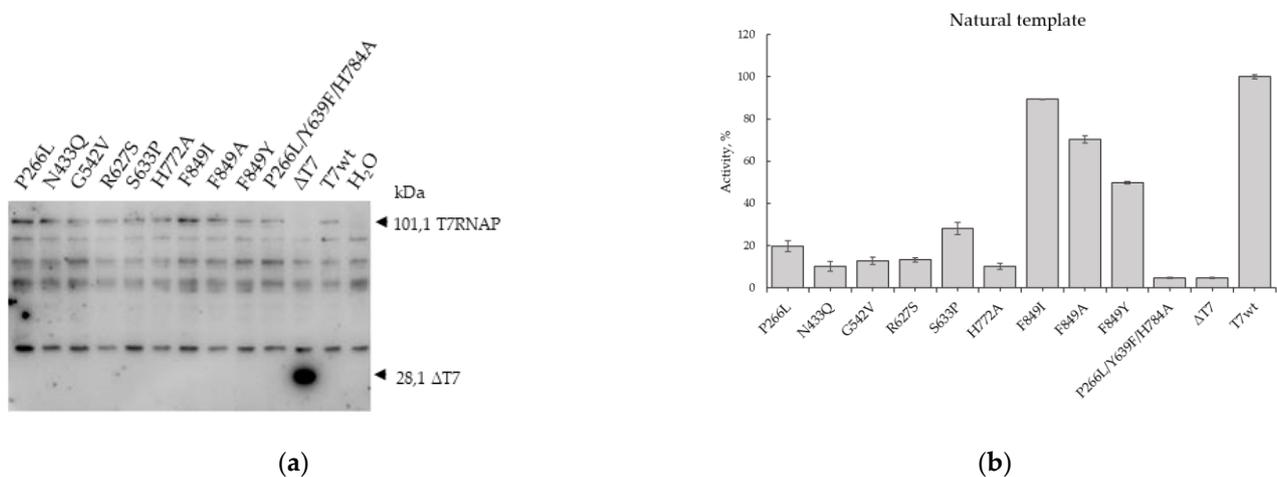


Figure 2. Activity of T7 RNAP mutants on the natural template. (a) Expression level of T7 RNAP mutants in the coupled² cell-free system. Western blot of the WGE lysate where T7 RNAP mutants were synthesized. (b) Activity of the mutants of T7 RNAP normalized on the amount of T7 RNAP synthesized in the coupled² cell-free system using natural template, $n = 3$, mean \pm SD.

The activities of the mutants during transcription of Nluc mRNA are illustrated in Figure 2b. The amount of synthesized Nluc reflects the amount of transcribed mRNA encoding Nluc, which in turn reflects the activity of the T7 RNAP mutant. All activities of T7 RNAP mutants were normalized to the amounts of the synthesized proteins in the coupled² system determined by western blotting of lysates. We revealed that in this system, one mutant form P266L/Y639F/H784A completely lost its activity on natural mRNA, and the luminescence level in the presence of this mutant was similar to the luminescence induced by the truncated form Δ T7 RNAP. Previously, it was reported that this mutant has increased thermal stability and can accept 2'-O-methyl triphosphates and dNTPs [54]. Moreover, this mutant of T7 RNAP was used to transcribe DNA containing eight nucleotide letters, four natural letters (dG, dC, dA, and dT), and four additional letters (dP, dZ, dS, and dB) [31]. Thus, we revealed that the increased specificity for dNTPs and unnatural nucleotides of this mutant markedly reduced its ability to transcribe natural templates. Presumably, Nluc mRNA synthesized by this mutant contains deoxynucleotides and cannot be utilized by the ribosome during translation in a cell-free system; hence, no luminescence was detected. Most of the other tested mutant forms (P266L, N433Q, G542V, R627S, S633P, H772A) decreased their activity in the coupled² system to 10%–30%. For all these mutants, increased thermal stability was demonstrated earlier [50,54–56]. Furthermore, for G542V the violation of interaction with 2'-OH of ribonucleotides was demonstrated [50] and for R627S, the violation of specificity expressed in the incorporation of ddNTP to RNA was shown [54,55]. Additionally, for the mutants S633P and H772A, relative activities in vitro by fluorescent beacon were determined, which were 56% and 70%, respectively [56]. In the coupled² system, we observed the lower activities of these mutants by 30% and 10%, respectively; however, mutants with replacements in position F849 retained more than half of the activity of the T7wt RNAP in the coupled² system—F849I, F849A, and F849Y. For F849I and F849Y mutants, the relative in vitro activities were also determined by fluorescent beacon as 88% and 37%, respectively [56], which is in accordance with our data of 90% for F849I and 50% for F849Y (Figure 2b). Thus, characterizing the activity of T7 RNAP mutants, we confirmed the efficiency of the developed method on the natural DNA

template apart from certain discrepancies with the published data, which are presumably bound to the greater proximity of the coupled² system to the natural conditions.

2.3. Obtaining DNA Templates Containing UBPs and Testing of T7 RNAP Mutants Activities

The linear DNA template for testing mutant forms of T7 RNA polymerase comprised a promoter for SP6 RNA polymerase, human beta-globin 5'UTR, a sequence encoding an Nluc protein of approximately 600 nucleotides in size, and a polyT tract for obtaining a polyA tail in mRNA. The total size of the PCR product was approximately 800 nucleotides. While choosing a codon for introducing UBP, we focused on the data for critical amino acids in Nluc [58]. Therefore, we selected the second position in the Val29 codon. Additionally, we chose the second position of the valine codon as it can be recognized by four different tRNAs with any nucleotide in the third position. So insertion of inappropriate nucleotide in the second position will change Val codon to another one (Ala, Asp, Glu, or Gly) what will lead to disruption of Nluc activity that can be detected as luminescence decrease. Two types of UBPs were tested in our study: the hydrophobic pair 5-Me-isodC–isodG and the hydrophilic pair dNaM–d5SICS.

Initially, we tested the activity of T7wt RNAP at each of the three obtained templates carrying unnatural nucleotides and compared it with the activity of T7wt RNAP at the natural template (Figure 3a). It is important to note that in the transcription reaction with T7 RNAP we used only natural ribonucleotide triphosphates. So RNAP recognized UBP in the DNA template and inserted into RNA natural ribonucleotide. Thus, RNA polymerase was mistaken in reading the unnatural nucleotide in DNA and used available natural ribonucleotides to build RNA. Therefore, the efficiency of transcription with unnatural templates was significantly lower than with natural ones. We observed that on the template containing isodG, the activity of T7wt RNAP was 30%, compared to that of the natural template. Moreover, T7wt RNAP is practically unable to transcribe the template containing 5-Me-isodC. The T7wt RNAP activity on this template was less than 5% compared to the natural one. T7 RNAP activity on the template containing dNaM was approximately 20% of its activity on the natural template.

Further, the activity of the selected mutant RNAPs (Table 1) was determined using unnatural templates. The activity of the mutants was normalized to the activity of T7wt RNAP on the same unnatural templates. On the template with isodG in the second position of the Val29 Nluc codon, the activities of most of the tested mutant RNAPs were lower than that of T7wt RNAP (Figure 3b); however, those of the mutants F849I and F849Y were comparable to the activity of T7wt RNAP on this template. A similar pattern was observed for the unnatural template comprising 5-Me-isodC in the same position. Most of the mutants demonstrated low activity, and only F849I and F849Y mutations had relatively high transcriptional activity (Figure 3c). On the template containing dNaM, the activity of mutants F849I, F849A, and F849Y was comparable to that of T7wt RNAP, and the mutants R627S, N433Q, S633P, G542V, and H772A were less active (Figure 3d). The activity of the mutant P266L/Y639F/H784A was similar to the level of the negative control on all templates; thus, this mutant was unable to transcribe functional mRNA. The high activities of the mutants F849I and F849Y compared to T7wt RNAP on all tested unnatural nucleotides indicated a lower selectivity of these polymerases, which can be used for further designing the mutant RNAPs for obtaining the protein utilizing unnatural nucleotides with high specificity.

Thus, we have demonstrated the efficiency of the developed method and compared the activities of the ten mutant T7 RNAPs on different unnatural templates.

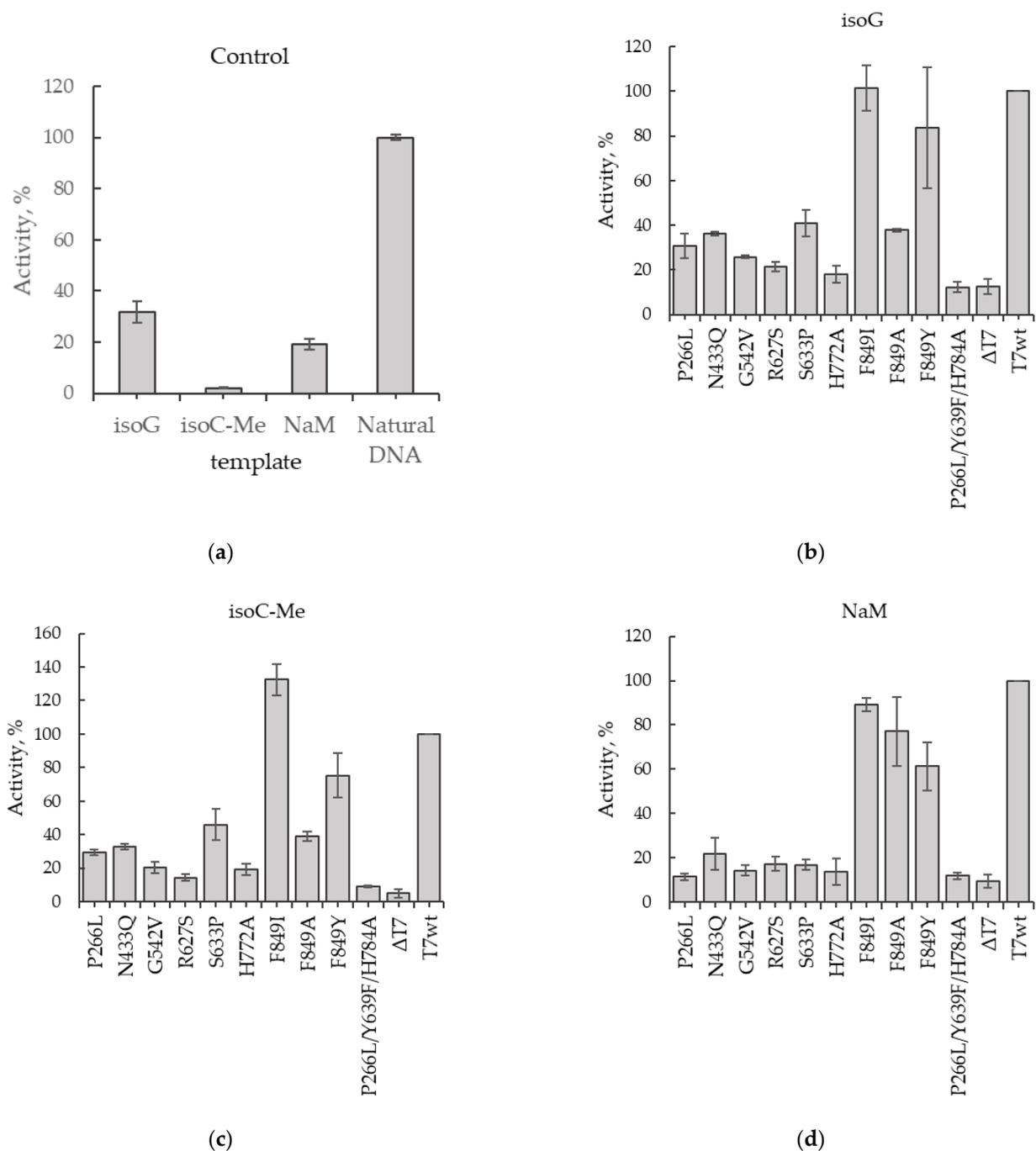


Figure 3. Activity of the T7 RNAP mutants on the templates containing unnatural nucleotides in the second position of V29 codon of Nluc normalized on the amount of T7 RNAP synthesized in the coupled² cell-free system, $n = 3$, mean \pm SD. (a) Activity of T7 RNAP wt on different templates. (b) Activity of T7 RNAP mutants on the template containing isodG. (c) Activity of T7 RNAP mutants on the template containing 5-Me-isodC. (d) Activity of T7 RNAP mutants on the template containing dNaM.

3. Discussion

We describe the “coupled²” cell-free transcription–translation system, a simple in vitro method to determine RNAP activity on templates containing unnatural nucleotides using a luminescence microplate reader. This method has a major advantage over other approaches, as it does not require expression and purification of proteins to test their activity during transcription. Moreover, in this system, all manipulations are performed in one tube/well for only 1 h. This method allows rapid measurement of activities of different

ferments simultaneously in one 364 well microplate. Importantly, the “coupled²” system can provide comparable results with other methods such as classical *in vitro* transcription, thereby suggesting that it is a robust and convenient framework for developing more complex methods.

“Coupled²” system has certain advantages over the classical approach, in which a purified recombinant polymerase is used for *in vitro* transcription. Indeed, classical approaches might not be convenient due to the additional difficult and time-consuming steps of bacterial expression and purification of RNA polymerases. This issue was resolved by using a coupled² cell-free transcription–translation system. As a result, the first round of transcription and translation RNA polymerases can immediately transcribe mRNA in the same tube. A key analytical advantage of this method is that it enables in determining the relative activity of RNAPs based on the luminescence induced by synthesized Nluc, which represents the amount of active RNAP in solution. Thus, this approach allowed easy and reproducible quantitative evaluation of RNAP activities. Furthermore, the use of a microplate reader remarkably simplified the method, which in turn allowed massive parallel analysis of dozens of samples in one day.

Another significant advantage of this method is the possibility to use templates containing different unnatural nucleotides or the use of different transcription activators/suppressors in parallel experiments. This approach might be useful in comparing the RNAP preferences of different unnatural nucleotides and the effects of other transcription regulators. The high sensitivity of the method allows comparison of a high range of RNAP activities, which is particularly important in experiments with components that significantly reduce the transcription activity.

Nevertheless, the coupled² system has certain limitations. Firstly, this is an indirect method; the determination of the activity of RNA polymerases occurs through the translation of the Nluc. This means that in order to obtain unambiguous results, it is necessary to control three parameters: translation of RNA polymerase, Nluc mRNA synthesis, and Nluc mRNA quality. Translation of RNA polymerase could be controlled by the Western blot hybridization, as we did in our work (Figure 2a). The synthesis and the amount of Nluc mRNA can be controlled using either Northern blot hybridization, or by an affinity chromatography purification of newly synthesized RNA polymerase and following *in vitro* transcription of mRNA. This approach also can be used to control the quality of transcribed mRNA. The quality of mRNA can be analyzed using cDNA sequencing. Therefore, described limitations are not critical and after the appropriate solution of them, they will provide additional information about the activity of the RNAP mutants.

In the future, this approach can be used to quickly determine the activity of RNA polymerases on new substrates and templates, which is presently lacking the development of useful unnatural base pairs within the synthetic approach.

4. Materials and Methods

4.1. Obtaining Constructs Encoding Mutant Forms of T7 RNAP

T7 RNAP cds was amplified by PCR with Q5 polymerase (NEB, Ipswich, MA, USA) from the plasmid pAR1219 using primers T7_HindIII_F and T7_XhoI_R (Table 3). The resulting PCR product of T7 RNAP and globine 5'UTR were inserted into the pGEM vector (Promega, Madison, WI, USA) at the HindIII and XhoI restriction sites. So we got the basic construction pGEM-globine_6xHis_TEV_T7RNAP. Based on this construct, mutants with single amino acid substitutions and a mutant with a premature stop codon were obtained using the QuikChange Site-Directed Mutagenesis Kit (Agilent, Santa Clara, CA, USA), as well as the P266L_Y639F_H784A mutant using the QuikChange Multi Site-Directed Mutagenesis Kit (Agilent, Santa Clara, CA, USA).

the amplification of the template, d5SICSTP was added to the mixture of nucleotide triphosphates to a final concentration of 0.4 mM. To adjust the concentration of $MgCl_2$, it was increased in the reaction mixture from 2 mM to 2.4 mM. Amplification was performed in a 50 μ L mixture using Q5 DNA polymerase (NEB, Ipswich, MA, USA) and 25 pmol of each primer (glob-Nluc forward primer and Nam-Seq1 reverse primer). The amplification program consisted of 1 cycle of initial denaturation at 98 °C for 30 s of the plasmid with the Nluc gene; 25 amplification cycles—primer annealing at 60 °C for 10 s, elongation at 72 °C for 80 s and denaturation at 98 °C for 10 s. At the final stage, one cycle of DNA elongation was performed at 72 °C for 5 min. Thereafter, the 200 bp PCR product was purified in 1% agarose gel using a Qiagen kit (Qiagen, Hilden, Germany). This PCR product was added to the second PCR to produce a full-length PCR product with a reverse primer Nluc-A50 complementary to the 3' end of the noncoding region comprising the T50 sequence. The volume of the mixture was 50 μ L and it contained 25 μ L ($\frac{1}{2}$ V) of previously purified PCR product. Amplification program: 1 cycle—98 °C, 30 sec, initial denaturation; 28 cycles—98 °C, denaturation for 10 sec. Annealing of the primer 69 °C, 10 s. Elongation at 72 °C for 160 s. Final elongation—5 min at 72 °C. Since the NaM-5SICS pair is specific [39,40], it was not eliminated during the amplification process due to DNA polymerase errors.

Pair of unnatural nucleotides 5-Me-isodC–isodG does not reveal enough specificity which leads to the accumulation of errors with every round of amplification of approximately 2% [26,29,59]. This indicates that after each round, the unnatural pair will be partially eliminated, and after 20 cycles, a significant loss of UBP will occur, which will hinder unambiguous interpretation of the results. To overcome this limitation, we synthesized pairs of primers complementary to each other, so that the forward primer carried 5-Me-isodC in the second position of the V29 codon, and the reverse primer carried isodG, and vice versa (Figure 4). In the first round of PCR, two fragments were amplified using primers flanking both ends of the template and mutant primers carrying unnatural nucleotides (Figure 4). PCR was performed in the presence of complementary nucleotide triphosphates—in the reaction with oligonucleotides containing 5-Me-isodC, isodGTP (Jena Bioscience, Jena, Germany) was added, and vice versa, 5-Me-isodCTP (MyChem, San Diego, CA, USA) was added with a primer containing isodG. As the amplified PCR product contained a modified primer, the total loss of the template with unnatural bases was $\leq 2\%$, that is, 98% of the templates carry an unnatural nucleotide.

Amplification of fragments was performed in 50 μ L of the mixture using Q5 DNA polymerase (NEB, Ipswich, MA, USA) and 25 pmol of each primer. First pair: glob-Nluc forward primer and isoC reverse primer; second pair: isoG forward primer, Nluc-A50 reverse primer. In such a case, the working chain for the T7 RNAP contains isoC. For a template with isoG decoding, the same steps were carried out, but the first pair contained the isoG reverse primer, and the second pair contained the isoC forward primer. The amplification program consisted of 1 cycle of initial denaturation of the plasmid with the Nluc gene at 98 °C for 30 s; 25 amplification cycles—primer annealing at 60 °C for 10 s, elongation at 72 °C for 80 sec and denaturation at 98 °C for 10 s. At the final stage, 1 cycle of DNA elongation was performed at 72 °C for 5 min. Then, PCR products of 200 bp and 600 bp were purified in 1% agarose gel using a Qiagen kit (Qiagen, Hilden, Germany).

In the second round of amplification, we mixed these two PCR products without adding flanking primers and performed 10 amplification cycles without using unnatural nucleoside triphosphates (Figure 4). The volume of the mixture was 100 μ L. Amplification program: 1 cycle—98 °C, 30 s, initial denaturation; then 10 cycles—98 °C, 10 sec—denaturation. Annealing of the primer—69 °C, 10 s. Elongation at 72 °C for 90 s. The final extension of the DNA ends was carried out for 5 min at 72 °C. As a result, we obtained a full-length PCR product encoding Nluc containing a 5-Me-isodC–isodG pair. Since we did not use primers flanking ends of the amplified sequence, the percentage of unnatural nucleotides in the template did not change and remained at 98%. Single-stranded DNA molecules and unreacted fragments remaining in the reaction were removed by agarose gel electrophoresis (Figure 4). The resulting templates with substitutions for dNaM, 5-Me-

isodC, and isodG in the lower chain were used for functional analysis of mutant T7 RNAPs in the coupled² cell-free system.

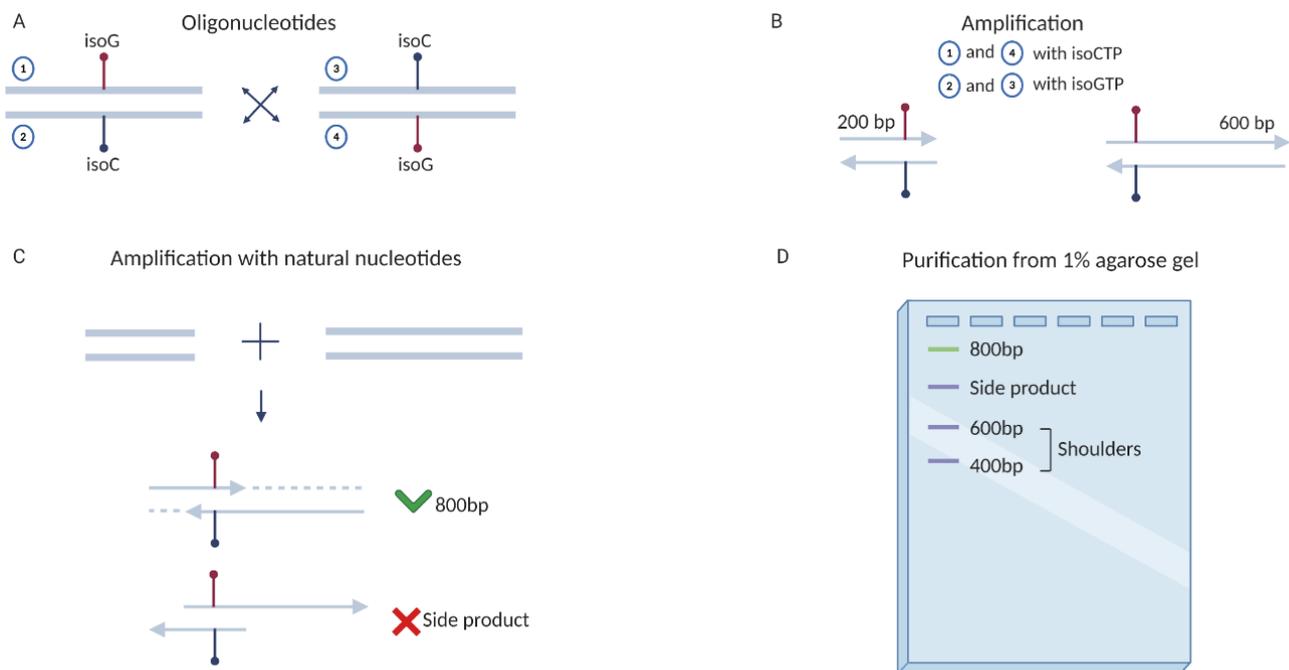


Figure 4. Scheme of the approach used to produce templates containing unnatural nucleotides isodG and 5-Me-isodC. (A) Pairs of primers complementary to each other, so that the forward primer carried 5-Me-isodC in the second position of the V29 codon, and the reverse primer carried isodG, and vice versa. (B) In the first round of PCR, two fragments were amplified using primers flanking both ends of the template and mutant primers carrying unnatural nucleotides. (C) In the second round of amplification, these PCR products were mixed without adding primers flanking ends of the template and performed 10 amplification cycles without using unnatural nucleoside triphosphates. (D) Single-stranded DNA molecules and unreacted fragments remaining in the reaction were removed by agarose gel electrophoresis.

Author Contributions: Conceptualization, E.A. and P.K.; methodology, P.K.; investigation, T.E., E.S., S.M., A.S.; writing—original draft preparation, T.E., E.S., S.M., A.S.; writing—review and editing, P.K. and E.A.; supervision, E.A.; funding acquisition, E.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RFBR, grant number 18-29-08044.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We are grateful to Ilya Terenin for providing plasmid encoding Nluc with globin leader sequence. Sequencing of plasmids, coding mutant proteins were performed by the center of the collective use “Genome” of EIMB RAS. Figures 1a and 4 were created with [BioRender.com](https://www.biorender.com) (accessed on 26 April 2021).

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

Abbreviations

CDS	CoDing Sequence
dNTPs	deoxyribonucleotide triphosphate
isoC	isocytosine (2-amino-4-ketopyrimidine)
isoG	isoguanine (6-amino-2-ketopurine)
Nluc	nanoluciferase
NTPs	ribonucleotide triphosphate
PCR	Polymerase Chain Reaction
RNAP	RNA polymerase
UBP	unnatural base pair
5'UTR	5'-untranslated regions
WGE	wheat germ extract
Wt	wild type

References

- Bain, J.D.; Switzer, C.; Chamberlin, A.R.; Benner, S.A. Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature* **1992**, *356*, 537–539. [[CrossRef](#)]
- Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; et al. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **2002**, *20*, 177–182. [[CrossRef](#)]
- Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S. A two-unnatural-base-pair system toward the expansion of the genetic code. *J. Am. Chem. Soc.* **2004**, *126*, 13298–13305. [[CrossRef](#)] [[PubMed](#)]
- Zhang, Y.; Ptacin, J.L.; Fischer, E.C.; Aerni, H.R.; Caffaro, C.E.; San Jose, K.; Feldman, A.W.; Turner, C.R.; Romesberg, F.E. A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **2017**, *551*, 644–647. [[CrossRef](#)]
- Fischer, E.C.; Hashimoto, K.; Zhang, Y.; Feldman, A.W.; Dien, V.T.; Karadeema, R.J.; Adhikary, R.; Ledbetter, M.P.; Krishnamurthy, R.; Romesberg, F.E. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat. Chem. Biol.* **2020**, *16*, 570–576. [[CrossRef](#)]
- Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J.W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464*, 441–444. [[CrossRef](#)]
- Berry, M.J.; Howard, M.T. Reprogramming the Ribosome for Selenoprotein Expression: RNA Elements and Protein Factors. In *Recoding: Expansion of Decoding Rules Enriches Gene Expression*; Atkins, J.F., Gesteland, R.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; p. 31.
- Hao, B.; Gong, W.; Ferguson, T.K.; James, C.M.; Krzycki, J.A.; Chan, M.K. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **2002**, *296*, 1462–1466. [[CrossRef](#)] [[PubMed](#)]
- Srinivasan, G.; James, C.M.; Krzycki, J.A. Pyrrolysine encoded by UAG in Archaea: Charging of a UAG-decoding specialized tRNA. *Science* **2002**, *296*, 1459–1462. [[CrossRef](#)] [[PubMed](#)]
- Tyagi, S.; Kramer, F.R. Molecular beacons: Probes that fluoresce upon hybridization. *Nat. Biotechnol.* **1996**, *14*, 303–308. [[CrossRef](#)]
- Collins, M.L.; Irvine, B.; Tyner, D.; Fine, E.; Zayati, C.; Chang, C.A.; Horn, T.; Ahle, D.; Detmer, J.; Shen, L.P.; et al. A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/mL. *Nucleic Acids Res.* **1997**, *25*, 2979–2984. [[CrossRef](#)]
- Sherrill, C.B.; Marshall, D.J.; Moser, M.J.; Larsen, C.A.; Daudé-Snow, L.; Prudent, J.R. Nucleic acid analysis using an expanded genetic alphabet to quench fluorescence. *J. Am. Chem. Soc.* **2004**, *126*, 4550–4556. [[CrossRef](#)]
- Johnson, S.C.; Marshall, D.J.; Harms, G.; Miller, C.M.; Sherrill, C.B.; Beaty, E.L.; Lederer, S.A.; Roesch, E.B.; Madsen, G.; Hoffman, G.L.; et al. Multiplexed genetic analysis using an expanded genetic alphabet. *Clin. Chem.* **2004**, *50*, 2019–2027. [[CrossRef](#)]
- Moser, M.J.; Christensen, D.R.; Norwood, D.; Prudent, J.R. Multiplexed detection of anthrax-related toxin genes. *J. Mol. Diagn.* **2006**, *8*, 89–96. [[CrossRef](#)] [[PubMed](#)]
- Kimoto, M.; Mitsui, T.; Yokoyama, S.; Hirao, I. A unique fluorescent base analogue for the expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2010**, *132*, 4988–4989. [[CrossRef](#)]
- Kimoto, M.; Mitsui, T.; Yamashige, R.; Sato, A.; Yokoyama, S.; Hirao, I. A new unnatural base pair system between fluorophore and quencher base analogues for nucleic acid-based imaging technology. *J. Am. Chem. Soc.* **2010**, *132*, 15418–15426. [[CrossRef](#)]
- Yamashige, R.; Kimoto, M.; Takezawa, Y.; Sato, A.; Mitsui, T.; Yokoyama, S.; Hirao, I. Highly specific unnatural base pair systems as a third base pair for PCR amplification. *Nucleic Acids Res.* **2012**, *40*, 2793–2806. [[CrossRef](#)]
- Someya, T.; Ando, A.; Kimoto, M.; Hirao, I. Site-specific labeling of RNA by combining genetic alphabet expansion transcription and copper-free click chemistry. *Nucleic Acids Res.* **2015**, *43*, 6665–6676. [[CrossRef](#)] [[PubMed](#)]
- Riedl, J.; Ding, Y.; Fleming, A.M.; Burrows, C.J. Identification of DNA lesions using a third base pair for amplification and nanopore sequencing. *Nat. Commun.* **2015**, *6*, 1–11. [[CrossRef](#)]
- Malyshev, D.A.; Romesberg, F.E. The expanded genetic alphabet. *Angew. Chemie Int. Ed.* **2015**, *54*, 11930–11944. [[CrossRef](#)] [[PubMed](#)]

21. Yamashige, R.; Kimoto, M.; Okumura, R.; Hirao, I. Visual detection of amplified DNA by polymerase chain reaction using a genetic alphabet expansion system. *J. Am. Chem. Soc.* **2018**, *140*, 14038–14041. [[CrossRef](#)]
22. Feldman, A.W.; Dien, V.T.; Karadeema, R.J.; Fischer, E.C.; You, Y.; Anderson, B.A.; Krishnamurthy, R.; Chen, J.S.; Li, L.; Romesberg, F.E. Optimization of replication, transcription, and translation in a semi-synthetic organism. *J. Am. Chem. Soc.* **2019**. [[CrossRef](#)]
23. Switzer, C.; Moroney, S.E.; Benner, S.A. Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323. [[CrossRef](#)]
24. Piccirilli, J.A.; Krauch, T.; Moroney, S.E.; Benner, S.A. Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **1990**, *343*, 33–37. [[CrossRef](#)] [[PubMed](#)]
25. Sismour, A.M.; Benner, S.A. The use of thymidine analogs to improve the replication of an extra DNA base pair: A synthetic biological system. *Nucleic Acids Res.* **2005**, *33*, 5640–5646. [[CrossRef](#)] [[PubMed](#)]
26. Switzer, C.Y.; Moroney, S.E.; Benner, S.A. Enzymatic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry* **1993**, *32*, 10489–10496. [[CrossRef](#)]
27. Horlacher, J.; Hottiger, M.; Podust, V.N.; Hubscher, U.; Benner, S.A. Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns. *Proc. Natl. Acad. Sci. USA* **2006**, *92*, 6329–6333. [[CrossRef](#)]
28. Benner, S.A. Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.* **2004**, *37*, 784–797. [[CrossRef](#)]
29. Martinot, T.A.; Benner, S.A. Artificial genetic systems: Exploiting the ‘aromaticity’ formalism to improve the tautomeric ratio for isoguanosine derivatives. *J. Org. Chem.* **2004**, *69*, 3972–3975. [[CrossRef](#)]
30. Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A.M.; Benner, S.A. Artificially expanded genetic information system: A new base pair with an alternative hydrogen bonding pattern. *Nucleic Acids Res.* **2006**, *34*, 6095–6101. [[CrossRef](#)]
31. Hoshika, S.; Leal, N.A.; Kim, M.J.; Kim, M.S.; Karalkar, N.B.; Kim, H.J.; Bates, A.M.; Watkins, N.E., Jr.; SantaLucia, H.A.; Meyer, A.J.; et al. Hachimoji DNA and RNA: System with eight building blocks. *Science* **2019**, *363*, 884–887. [[CrossRef](#)]
32. Johar, Z.; Zahn, A.; Leumann, C.J.; Jaun, B. Solution structure of a DNA duplex containing a biphenyl pair. *Chemistry* **2008**, *14*, 1080–1086. [[CrossRef](#)]
33. Leconte, A.M.; Hwang, G.T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F.E. Discovery, characterization, and optimization of an unnatural base pair for expansion of the genetic alphabet. *J. Am. Chem. Soc.* **2008**, *130*, 2336–2343. [[CrossRef](#)]
34. Seo, Y.J.; Hwang, G.T.; Ordoukhanian, P.; Romesberg, F.E. Optimization of an unnatural base pair toward natural-like replication. *J. Am. Chem. Soc.* **2009**, *131*, 3246–3252. [[CrossRef](#)]
35. Matsuda, S.; Fillo, J.D.; Henry, A.A.; Rai, P.; Wilkens, S.J.; Dwyer, T.J.; Geierstanger, B.H.; Wemmer, D.E.; Schultz, P.G.; Spraggon, G.; et al. Efforts toward expansion of the genetic alphabet: Structure and replication of unnatural base pairs. *J. Am. Chem. Soc.* **2007**, *129*, 10466–10473. [[CrossRef](#)]
36. Seo, Y.J.; Romesberg, F.E. Major groove derivatization of an unnatural base pair. *Chembiochem* **2009**, *10*, 2394–2400. [[CrossRef](#)]
37. Yang, Z.; Sismour, A.M.; Sheng, P.; Puskar, N.L.; Benner, S.A. Enzymatic incorporation of a third nucleobase pair. *Nucleic Acids Res.* **2007**, *35*, 4238–4249. [[CrossRef](#)] [[PubMed](#)]
38. Yang, Z.; Chen, F.; Alvarado, J.B.; Benner, S.A. Amplification, mutation, and sequencing of a six-letter synthetic genetic system. *J. Am. Chem. Soc.* **2011**, *133*, 15105–15112. [[CrossRef](#)]
39. Malyshev, D.A.; Seo, Y.J.; Ordoukhanian, P.; Romesberg, F.E. PCR with an expanded genetic alphabet. *J. Am. Chem. Soc.* **2009**, *131*, 14620–14621. [[CrossRef](#)]
40. Malyshev, D.A.; Pfaff, D.A.; Ippoliti, S.I.; Hwang, G.T.; Dwyer, T.J.; Romesberg, F.E. Solution structure, mechanism of replication, and optimization of an unnatural base pair. *Chem. Eur. J.* **2010**, *16*, 12650–12659. [[CrossRef](#)] [[PubMed](#)]
41. Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. An efficient unnatural base pair for PCR amplification. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555. [[CrossRef](#)] [[PubMed](#)]
42. Morohashi, N.; Kimoto, M.; Sato, A.; Kawai, R.; Hirao, I. Site-specific incorporation of functional components into RNA by an unnatural base pair transcription system. *Molecules* **2012**, *17*, 2855–2876. [[CrossRef](#)] [[PubMed](#)]
43. Kimoto, M.; Hirao, I. Site-specific incorporation of extra components into RNA by transcription using unnatural base pair systems. *Methods Mol. Biol.* **2010**, *634*, 355–369. [[PubMed](#)]
44. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. An unnatural hydrophobic base pair system: Site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat. Methods.* **2006**, *3*, 729–735. [[CrossRef](#)] [[PubMed](#)]
45. Kimoto, M.; Meyer, A.J.; Hirao, I.; Ellington, A.D. Genetic alphabet expansion transcription generating functional RNA molecules containing a five-letter alphabet including modified unnatural and natural base nucleotides by thermostable T7 RNA polymerase variants. *Chem. Commun.* **2017**, *53*, 12309–12312. [[CrossRef](#)]
46. Cheetham, G.M.; Steitz, T.A. Insights into transcription: Structure and function of single-subunit DNA-dependent RNA polymerases. *Curr. Opin. Struct. Biol.* **2000**, *10*, 117–123. [[CrossRef](#)]
47. Sousa, R. Structural and mechanistic relationships between nucleic acid polymerases. *Trends Biochem. Sci.* **1996**, *21*, 186–190. [[CrossRef](#)]
48. Gudima, S. Deoxyribonucleotide-containing RNAs: A novel class of templates for HIV-1 reverse transcriptase. *Nucleic Acids Res.* **1997**, *25*, 4614–4618. [[CrossRef](#)]

49. Rusakova, E.E.; Tunitskaya, V.L.; Memelova, L.V.; Kochetkova, S.V.; Kostyuk, D.A.; Kochetkov, S. N. Mutant T7 RNA polymerase is capable of catalyzing DNA primer extension reaction. *FEBS Lett.* **1998**, *423*, 189–192. [[CrossRef](#)]
50. Chelliserrykattil, J.; Ellington, A.D. Evolution of a T7 RNA polymerase variant that transcribes 2'-O-methyl RNA. *Nat. Biotechnol.* **2004**, *22*, 1155–1160. [[CrossRef](#)]
51. Frydman, J.; Erdjument-Bromage, H.; Tempst, P.; Hartl, F.U. Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat. Struct. Biol.* **1999**, *6*, 697–705. [[CrossRef](#)]
52. England, C.G.; Ehlerding, E.B.; Cai, W. NanoLuc: A small luciferase is brightening up the field of bioluminescence. *Bioconjugate Chem.* **2016**, *27*, 1175–1187. [[CrossRef](#)]
53. Guillerez, J.; Lopez, P.; Proux, F.; Launay, H.; Dreyfus, M. A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5958–5963. [[CrossRef](#)] [[PubMed](#)]
54. Meyer, A.J.; Garry, D.J.; Hall, B.; Byrom, M.M.; McDonald, H.G.; Yang, X.; Yin, Y.W.; Ellington, A.D. Transcription yield of fully 2'-modified RNA can be increased by the addition of thermostabilizing mutations to T7 RNA polymerase mutants. *Nucleic Acids Res.* **2015**, *43*, 7480–7488. [[CrossRef](#)]
55. Rusakova, E.E.; Sugiyama, A.; Nishiya, Y.; Kawakami, B. RNA Polymerase Mutants with Increased Thermostability. U.S. Patent 7,507,567B2, 24 March 2009.
56. Boulain, J.C.; Dassa, J.; Mesta, L.; Savatier, A.; Costa, N.; Muller, B.; L'hostis, G.; Stura, E.; Troesch, A.; Ducancel, F. Mutants with higher stability and specific activity from a single thermostable variant of T7 RNA polymerase. *Protein Eng. Des. Sel.* **2013**, *26*, 725–734. [[CrossRef](#)] [[PubMed](#)]
57. Temiakov, D.; Patlan, V.; Anikin, M.; McAllister, W.T.; Yokoyama, S.; Vassylyev, D.G. Structural basis for substrate selection by T7 RNA Polymerase. *Cell* **2004**, *116*, 381–391. [[CrossRef](#)]
58. Hall, M.P.; Unch, J.; Binkowski, B.F.; Valley, M.P.; Butler, B.L.; Wood, M.G.; Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; et al. Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol.* **2012**, *7*, 1848–1857. [[CrossRef](#)]
59. Maciejewska, A.M.; Lichota, K.D.; Kusmierk, J.T. Neighbouring bases in template influence base-pairing of isoguanine. *Biochem. J.* **2003**, *369*, 611–618. [[CrossRef](#)]