Systematic Comparison of the Capability of Modified Nucleoside Triphosphates to Act as a Substrate for Thermostable DNA Polymerases during PCR

Yury Nikolaevich Lezhnin1,2, Anna Yurievna Krystenichenko1, Dmitry Sergeyevich Kravchenko1,2, Elena Ivanovna Frolova4,5, Julia Evgenyevna Kravchenko1,2, Stepan Petrovich Chumakov1,2

1 Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, 16/10, Miklukho-Maklaya St., GSP-7, Moscow 117997, Russia
2 Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 119991, Russia

Abstract

Aptamer-based molecules are becoming increasingly popular for both research applications and clinical therapy. However, advantages of this family, namely high affinity and specificity, are partially offset by a disadvantage: existence of “difficult” targets. For these targets, selection of a highly-specific aptamer is hindered by low diversity of existing functional groups – nucleotides, comprised in an aptamer. To overcome this drawback, aptamers could be selected from non-natural libraries that contain nucleotides with various functional groups attached to nitrogenous bases. This increases the chances of selecting good aptamer against targets that have high affinity towards these functional groups. The simplest way to connect new functional groups to nucleotides is to use click-chemistry or allylamine-NHS joining, which means that a minimal blocker, making it possible to carry out post-amplification modification of an aptamer library, is a nucleotide with an amine or alkyne linker attached to a nitrogenous base. The critical stage that determines whether the modified library is suitable for aptamer selection, is enzymatic amplification of sequences by a thermostable DNA polymerase. In this study we compared capabilities of the most widely spread types of thermostable DNA polymerases to incorporate modified nucleoside triphosphates into DNA strands, which can be used to introduce additional functional groups after amplification of the library.

Keywords

Modified aptamers; DNA polymerases; Non-natural nucleotides

Introduction

For a long time nucleic acids were regarded as a coding matrix containing information on the structure of protein molecules, which, in turn, carry out all the main functions of a cell. However, during the last 20 years of studies numerous examples demonstrated that molecules of nucleic acids can act directly in a numerous cellular processes. For example, in the course of the study of gene expression, the process of suppression of gene expression at a transcription stage by means of small RNAs was discovered. This process was named RNA-interference [1]. Later, many other examples emerged, showing that single-stranded DNAs and RNAs are capable to form complicated 3D structures, which allows them to specifically bind certain targets and even directly catalyze certain reactions. Thus, nucleic acids can also be used as specific ligands similar to antibodies. This kind of single-stranded RNA and DNA molecules capable to recognize target molecules were named aptamers [2]. Generally, single-stranded RNA and DNA molecules are used as aptamers; these molecules consist of the following three parts: two constant end sequences and a central variable part. The length of the variable part is 20-30 nucleotides, which comprises a huge combinatorial variety of possible variants of molecules capable to form diverse secondary and tertiary structures. Among that pool of molecules, there are almost always one or few that are capable to bind to a certain target. Numerous experiments with aptamer selection demonstrated that aptamers can be obtained for proteins, polysaccharides, viruses and a wide variety of other targets [3]. Due to their unique properties, aptamers can be used as effective substitutes for antibodies. By the majority of their characteristics, aptamers are on par or even better than monoclonal antibodies. Their advantages include higher stability, simplicity and speed of production, and smaller size of functional molecules. Small size of aptamers is the main advantage of that kind of molecules, as compared to monoclonal antibodies. The size of a monoclonal antibody is around 150 kDa; at the same time the size of aptamer molecules is around 3 kDa. The size of molecule is crucial for therapeutic applications, since it directly influences the bioavailability of the therapeutic agent, its capability to permeate into tissues and cells of an organism. Moreover, small size of aptamers makes them invisible for an immune system of an organism. Nowadays, both RNA and single-stranded DNA aptamers are used in various applications. The analysis of different selected aptamers shows that the main characteristics, that is, the specificity and affinity, are the same for both RNA and DNA aptamers. However, there are certain differences. An advantage of RNA-aptamers is that they can be expressed inside cells (same with miRNAs). This characteristic is important for experiments in vivo. At the same time, DNA aptamers are stable in a wide range of conditions, which is even more important in case of in vivo administration of active molecules. In that regard, aptamers on the basis of deoxyribonucleic acid (ssDNA aptamers) are becoming increasingly popular.

Nowadays, methods of selecting modified aptamers are becoming increasingly popular. Adding various modified nucleotides allows to considerably increase the combinatorial variation of oligonucleotides and thus facilitate in obtaining the most efficient and specific molecule. Moreover, introducing additional functional groups increases the number of possible contacts between an aptamer molecule and its target, which results in increased affinity and specificity of an aptamer.
even for complicated non-protein compounds, such as glycated parts of receptors of cells [4,5]. On the other hand, modifications of nucleotides make aptamers resistant to nucleases, which is also significant for in vivo applications [6,7]. Also, in some cases fluorescent groups are attached to an aptamer in order to visualize its interaction with a target molecule [8,9]. There are two alternative approaches for creating modified aptamers. In the first approach, modified nucleotides are used for selecting aptamers in the process of library amplification. The main limitation of this method is incompatibility of a modified nucleotide with the DNA (RNA) polymerase, that is, inability of a nucleotide to act as a substrate for a ferment that catalyzes PCR. In the second approach, already selected aptamer molecules undergo functionalization. In this case the number of available modifications significantly increases. However, most of these functionalizations result in decreased affinity of aptamers towards their targets, which is even a bigger problem than incompatibility of a nucleotide and a polymerase at initial stage of selecting specific aptamers.

Libraries of single-stranded DNA molecules are produced by a standard method of chemical synthesis of oligodeoxynucleotides using a mix of all four nucleotides during the synthesis of a variable part. Then the library is incubated with the immobilized target, with further separation of unbound oligonucleotides from oligonucleotide-target complexes. Purified oligonucleotide-target complexes are then dissociated and bound DNA undergoes amplification for the next round of selection. Generally, in order to produce specific aptamers with high affinity to their targets 5-20 selection rounds are carried out. That method of selecting aptamers that are highly specific towards their targets, was named SELEX (systematic evolution of ligands by exponential enrichment). In each round of SELEX, selection of single-stranded DNA molecules is performed; therefore, an important aspect of amplifying DNA molecules linked with a target is obtaining exactly single-stranded, rather than double-stranded molecules, due the fact that only single-stranded DNA molecules are capable to form 3D spatial structures, which is necessary for binding with a target, whereas double-stranded DNA molecules form a canonical double spiral and are incapable to interact with a target. There are several methods of producing single-stranded DNA molecules: asymmetric PCR [10], biotin-streptavidin separation [11], using exonuclease of phage λ [12] and separating DNA strands during electrophoresis in polyacrylamide gel in denaturing conditions. However, due to its simplicity and high reproducibility, the method of asymmetric PCR has become the most widely used.

Considering all the aspects of producing and selecting aptamers specific to targets, a DNA polymerase can be regarded as the most important factor that determines the success of creation of a library of variable ssDNA aptamers. Therefore, the structure and properties of a DNA polymerase determine the possibility of incorporating the modified bases into the created strand of a single-stranded oligonucleotide or the processivity of a PCR reaction in general. Generally, the structures of various DNA polymerases are considerably conservative. Subunits forming them have minor difference between each other in various types of organisms. This organization is explained by the importance of that type of enzymes for cells. However, the study of various DNA polymerases demonstrated that there are certain differences in a structure of their catalytic centers. Thus, thermostable DNA polymerases belonging to various families can differ in their capability to incorporate modified nucleoside triphosphates in a created daughter DNA strand in the course of a polymerase reaction. In the presented study, we analyzed the efficiency of incorporating modified nucleoside triphosphates by various thermostable DNA polymerases: Taq, Terasus(Pfu), Deep Vent, Phusion.

**Materials and Methods**

**Polymerase chain reaction (PCR)**

In order to verify the efficiency of incorporation of modified nucleotides in a DNA strand, we carried out a series of PCRs using various types of thermostable DNA polymerases. In the presented study, we used the following DNA polymerases: Taq (NEB, USA), Tersus (Evrogen, Russia), Deep Vent (NEB, USA), Phusion (NEB, USA). PCR was carried out in standard conditions for all polymerases. PCR included the following steps: preliminary denaturation of primers at 95°C for 4 min, denaturation of duplexes of daughter and matrix strand of DNA for each new cycle at 95°C for 30 s, hybridization of primers at 64°C for 40 s, elongation of daughter strands of DNA at 72°C for 8 min. Concentration of a matrix (primers) for one reaction was 40 pmol. Concentration of not modified nucleotides for one reaction was 250 µM, and concentration of modified nucleotides was 400 µM. Concentration of MgSO₄ and DMSO for each DNA polymerase was calculated separately according to the recommendation of the producer. The volume of the PCR mix for one polymerase was 30 µl.

Primers used in the study were supplied in lyophilized form. The synthesis of primers was carried out on a basis of Evrogen CJSC (Russia). Before using the primers were dissolved in mQ and heated at 95°C for 5 min.

**Polyacrylamide Gel Electrophoresis and the analysis of the results**

The specimens obtained as the result of PCR reaction were analyzed using electrophoresis separation of DNA molecules in a polyacrylamide gel. To ensure good quality of the visualization of electrophoresis results, SybrGold dye was added to specimens with ratio 1:10,000. Electrophoresis was carried out at a voltage of electric field of 80 V for 90 min at room temperature. For PAGE-electrophoresis, we used BioRad electrophoresis vertical system produced by BioRad Laboratories (USA). Electrophoresis separation of DNA molecules was carried out in 20% polyacrylamide gel using Tris-Borate buffer. Visualization of electrophoresis results was carried out by means of Safe Imager system produced by ThermoFisher Scientific Company (USA). To determine the molecular mass of DNA fragments, we used the marker of molecular length Generuler™ Ultra Low Range DNA Ladder produced by ThermoFisher Scientific Company (USA).

**Results and Discussion**

Various post-translation modifications have a significant influence on functions and activity of many proteins, including their capability to interact with antibodies and aptamers. However, many of proteins of a cell, including receptors of a cell, have significant modified domains. Those modifications comprise, first of all, glycosylation. Glycosylation is a complicated enzymatic process, directly reflecting the metabolic condition of a cell. For example, for normal and tumor cells the same receptor of a cell can have different variants of glycosylated part [13]. Thus, glycosylated domains are an excellent target for creating a wide range of therapeutic agents. Currently, both antibodies and methods of selecting aptamers based on oligonucleotides are used to develop agents that are capable to specifically bind to glycoproteins. Unfortunately, both methods are incapable to concentrate selection of an aptamer or antibody exactly to glycosylated part of a protein. Generally, that
problem is related to the limitations of combinatorial libraries or, in other words, to a small number of variants of selected molecules. However, a method of increasing the number of variants of selected molecules exists, and it allows to considerably increase the possibility of selecting a specific molecule, capable to recognize a glycosylated domain. This is the method of introducing chemical modifications in a structure of nucleotides, which are inserted in an aptamer. Generally, that method is used for creating aptamer libraries; chemical modifications of antibodies are scarcely used. Introducing chemical modifications in an aptamer can significantly increase variation of a selected library up to $10^4$ for aptamers with 30 random nucleotides. However, the main limitation of that approach is the impossibility of incorporating modified nucleotides by a certain DNA polymerase. In this study, we verified the possibility of incorporating a number of modified nucleotides by certain types of DNA polymerases. As modified nucleotides, we used nucleotides with uracil base linked by 5’ position with the following chemical radicals: aminoallyl radical $\text{C}_3\text{H}_6\text{N} (–\text{CH}−\text{CH}_2–\text{NH}_2)$, aminoallyl radical $\text{C}_9\text{H}_{17}\text{N}_2 (–\text{CH}−\text{CH}_2–\text{NH–CO}–(\text{CH}_2)_5–\text{NH}_2)$ and $\text{C}_8\text{H}_9$ alkyne radical $(-\text{C}≡\text{C}–(\text{CH}_2)_4–\text{C}≡\text{CH})$ (Figure 1).

Incorporating additional radicals in a structure of nitrogen base allows to expand the spectrum of existing nucleotides, which increases the variation of aptamer libraries and its combinatorial capabilities.

Incorporation of modified nucleotides was carried out in the process of PCR. As an enzyme for PCR, we used the following types of DNA-dependent DNA polymerases: Taq DNA polymerase, Tersus DNA polymerase, Deep Vent DNA polymerase and Phusion DNA polymerase. As a matrix, we used two overlapping primers (Figure 2).

The theoretical length of PCR-product after successful completion of DNA strand is 40 pairs of bases.

The efficiency of incorporation of nucleotides was analyzed using electrophoresis in 20% polyacrylamide gel. It is worth mentioning that in the beginning of PCR the first nucleotides, incorporated in a created strand of DNA, are deoxyuridine nucleotides with 5’ radical groups. Thus, if tested DNA polymerase is incapable to use some of modified nucleotides as a substrate, the formation of full 40 nucleotide duplex does not occur, which significantly facilitates the visualization of the results of PCR. The results of the comparison of full 40 nucleotide duplex and 27 nucleotide primers are presented in Figure 3.

PCR for all types of DNA polymerases was carried out in the same conditions. In order to achieve full saturation of reaction mix by 40 nucleotide duplex DNA molecules, five cycles of PCR were carried out. The results of PCR with various types of DNA polymerases and modified nucleotides are presented in Figure 4.

As it can be seen from the results of electrophoresis, Taq polymerase is capable to effectively incorporate in the sequence of a strand only
intermediate variant, when part of a DNA strand was completed and
Deoxyuridine-5'-Triphosphate allows to see differences in using polymerases from various
accuracy and processivity of PCR of DNA strand. Using Tersus
the family B of thermostable DNA polymerases, which also contains
as Tfi, HotTub and Tth. DeepVent and Phusion polymerases belong to
polymerases belong to various families. Taq polymerase belongs to the
One of the main reasons is that the studied thermostable DNA
modified nucleoside triphosphates can be due to a number of reasons.
DNA polymerases. These differences in the efficiency of incorporating
nucleotide cannot be quantitatively incorporated by any of the verified
nucleotide is effectively used only by DeepVent and Phusion DNA
polymerase, as well as Taq polymerase, is capable to effectively
incorporate modified 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate. M – marker of molecular length
GeneRuler Ultra Low Range DNA Ladder (ThermoFisher Scientific, USA).
Control – unamplified primer mixture

Figure 4: PAGE electrophoresis analysis of the efficiency of incorporating modified nucleotides with various types of DNA polymerases: Taq DNA polymerase, Tersus DNA polymerase, DeepVent DNA polymerase, Phusion-polymerase: (a) positive control with natural deoxythymidine triphosphates. (b) 5-(C3-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate, (c) 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate and (d) 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate. M – marker of molecular length

modified nucleotide 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate. Also, Taq polymerase is capable to incorporate 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate; however, the efficiency of incorporation of that kind of modified nucleotide is significantly decreased. Tersus polymerase, as well as Taq polymerase, is capable to effectively incorporate modified 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate. At that, there are certain differences from Taq polymerase. The results of electrophoresis show that Tersus polymerase incorporates modified 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate, and the efficiency of incorporating such nucleotide is higher than the efficiency demonstrated by Taq polymerase. DeepVent and Phusion DNA polymerases use 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate and 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate nucleotides with similar efficiency; just as Taq and Tersus polymerases they are capable to use modified 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate nucleotide. Thus, on the basis of the results of electrophoresis, it can be concluded that nucleotide 5-(C8 Alkyne)-2'-Deoxyuridine-5'-Triphosphate can be effectively incorporated by all types of DNA polymerases; 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate nucleotide is effectively used only by DeepVent and Phusion DNA polymerases, and 5-(C3-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate nucleotide cannot be quantitatively incorporated by any of the verified DNA polymerases. These differences in the efficiency of incorporating modified nucleoside triphosphates can be due to a number of reasons. One of the main reasons is that the studied thermostable DNA polymerases belong to various families. Taq polymerase belongs to the family A of thermostable DNA polymerases, along with such polymerases as Tfi, HotTub and Tth. DeepVent and Phusion polymerases belong to the family B of thermostable DNA polymerases, which also contains Pfu and Pwo polymerases. On the other hand, Tersus polymerase is a mixture of several thermostable DNA polymerases, which provide accuracy and processivity of PCR of DNA strand. Using Tersus polymerase allows to see differences in using polymerases from various families. Efficiency of incorporating modified 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate nucleotide by Tersus polymerase is an intermediate variant, when part of a DNA strand was completed and a part is in incomplete 27 nucleotide form. It can be presumed that completed variants of DNA were formed due to the operation of Pfu polymerase, while Taq polymerase could not use modified nucleotide, which led to the incompletion of a part of sequences. The efficiency of incorporating modified nucleotides also directly depends on a structure of the nucleotide. In that case, nucleotides based on deoxyuridine modified in C5 position were used. Modifications consisted in linking additional radical groups in the following form: C3-Aminoallyl, C11-Aminoallyl and C8-Alkyne. Thus, during the initial analysis of the structure of radicals it is possible to identify a certain relationship between the efficiency of incorporating nucleotides with modified bases and characteristics of the modifications, such as the length of a strand of a radical and position of triple bond. In that case, the position of triple bonds could play a key role in the process of recognition of nucleotides by DNA polymerases. According to the results of PCR-analysis, it can be presumed that the presence of the triple bond between C7 and C8 allows to stabilize conformational capabilities of a radical, which, theoretically, can positively influence the capability of a nucleotide to correctly position itself in a catalytic center, and, thus, use a DNA polymerase as a substrate. A double bond is not enough for that kind of stabilization, therefore, a radical has a significantly larger number of the degrees of freedom, and it cannot be used as a substrate. On the other hand, a longer strand of a radical, such as in 5-(C11-Aminoallyl)-2'-Deoxyuridine-5'-Triphosphate, also increases the possibility of sterical complications, which can be clearly seen in the electrophrogram, where the nucleotide with that modification was virtually not recognized by the Taq-polymerase (the polymerase of A family). However, it was recognized and used by the polymerases DeepVent and Phusion (the polymerases of B family), which, in turn, is explained by the differences in the structure of catalytic centers of polymerases of families A and B.

Conclusion

It can be concluded that polymerases belonging to different classes are capable, with various efficiency, to incorporate modified nucleotides during PCR; at the same time, polymerases of B family show much better efficiency in incorporation of modified nucleotides. Moreover, the efficiency of incorporating modified bases during PCR is significantly influenced by the structure of the modifications: bases, which contain alkyne linker, are recognized as a substrate by a larger number of DNA polymerases, while bases with aminolinker are more difficult for incorporation. Surprisingly, the length of a linker does not have a negative influence on the efficiency of incorporating nucleotide. It should be noted that this type of modifications can be regarded as a basis for further incorporation of more complex chemical groups. Incorporation of additional chemical groups with certain properties allows to considerably expand variation of a library, which increases the chances to select a specific aptamer even for complex non-protein compounds, such as glycans.

Acknowledgment

The work was funded by MESR contract No. 14.604.21.0110 (RFMEFI60414X0110).

References


