Iraqi Journal of Medical and Health Sciences, ISSN: XXXX-XXXX, Vol. 1, No. 1, 2024, Pages 14-21
DOI: https://doi.org/XXXXXX

Iraqi Journal of Medical and Health Sciences
Journal Homepage
https://ijmhs.mtu.edu.iq/ijmhs/index.php/home

RESEARCH ARTICLE – GENERAL BIOCHEMISTRY, GENETIC AND MOLECULAR BIOLOGY

Efficient Correction of DNA Hetero–Duplexes Formed During SELEX Procedure

Adil Sabr Al-Ogaili[1,*], Billy M. Hargis[1,2] and Young Min Kwon[1,2]

[1] Department of Cell and Molecular Biology, University of Arkansas, Fayetteville, AR 72701, USA

[2] Center of Excellence for Poultry Science, University of Arkansas, Fayetteville, AR 72701, USA

[*] Corresponding author E-mail: adel.saber@mtu.edu.iq

Article Info.

Article history:

Received
1 Oct. 2024

Accepted
5 Nov. 2024

Publishing
10 Nov. 2024

Abstract

Background: Aptamers, oligos-based (ssDNA or RNA) and peptide-based, are now emerging as the first-choice biomolecules in various applications in medicine and biology. DNA aptamers selection process includes repeatable steps of DNA amplification and digestion. The amplification process of the enriched DNA aptamers includes using PCR.

Objective of study: In this article, we sought to describe an efficient and robust method that corrects or permanently removes the PCR by–products during the cycles of DNA aptamers selection.

Materials and Methods: Our method is simple, efficient, and requires no enzymatic treatment. In this method, we applied repetitive cycles of thermal treatment to the enriched aptamers final products during the enrichment cycles. The procedure starts with loading the PCR-induced DNA on 1% agarose gel to evaluate the presence of ladder-like DNA or smearing band-like DNA. Next, we cut and purified these by–products and applied thermal treatment. The thermal treatment was by heating the DNA to 65ºC for 10 minutes and then cool them down to 20ºC for 1 minute.

Results: PCR artifacts evaluation has been revealed two types of unwanted DNA sequences; the ladder-like and the smearing bands. The next–Generation illume sequencing specified the presence of long and short sequences. Upon the thermal treatment of the PCR products, the next–generation illume sequencing method results showed that the longer sequences and the shorter sequences of the DNA have been reduced considerably and effectively in the DNA pool over the cycles of the systemic evolution of ligands by exponential enrichment (SELEX) procedure. Furthermore, we could reduce the formation of the hetero–duplexes in the next cycles of (SELEX) procedure.

Conclusion: This method suggests practical and routine application to prevent or reduce DNA aptamer by–products during the selection process. This treatment could reduce the mis-shaped or irrelevant non-functional aptamers in SELEX procedure.

Publisher: Middle Technical University

Keywords: ssDNA Aptamer; PCR Artifacts; Hetero–Duplexes; PCR Artifact Correction; SELEX.

1. Introduction

DNA aptamers selection starts to emerge as a leading technique that is now replacing monoclonal antibodies in many biological disciplines including medicine, chemistry, and nanotechnology [1-3]. Aptamers (DNA, RNA, or peptide) are selected using a method that had been developed by two laboratories in 1990 independently and simultaneously [4, 5]. The technique of aptamers selection (single-stranded DNA or RNA) is carried out by a procedure called systemic evolution of ligands by exponential enrichment (SELEX). SELEX has been extensively utilized to develop ssDNA aptamers with high affinity and specificity to a vast array of targets including proteins, whole cells, ions, and molecules. In essence, SELEX procedure relies on designing an oligonucleotide library as the milestone of the target capturing. This library has a central randomized sequence that is flanked by two constant sequence regions at both 5-prime end and 3-prime end. The constant regions act as primer/primer-binding sites that are required for the subsequent PCR amplification. In the variable oligonucleotide sequence design, all the four bases (i.e., A, T, C, G) are incorporated equally at any position. Therefore, any predicted diversity of a variable library depends on the length of its random sequence. In general, the randomness of a library is 4^n where 4 represents the four nucleotides (C, T, G, A) and *n* represents the number of the oligos in the random region [6]. SELEX method also depends on using PCR to amplify the enriched aptamers exponentially after each cycle. However, the PCR produces double-stranded DNA (dsDNA) and this dsDNA needs to be converted into single ssDNA. This is carried out by the digestion of the dsDNA with an exonuclease enzyme (for example, lambda exonuclease) to resume the original ssDNA format, which in turn, starts a new cycle of aptamers enrichment [7].

Due to the presence of this region, there have always been a possibility to develop artifacts in the PCR products. Many sequences from the previous PCR might hybridize partially at the randomized region of the library to form PCR by–products or artifacts with variable lengths [8-10]. One of the well-explained issues during the amplification of random DNA libraries is that either or both primers completely or partially-hybridize to random regions other than the constant primer annealing sites. Here, the normal PCR products are usually mixed with multiple types of by–products such as smearing band and ladder–like bands. These PCR by–products might interfere negatively with the final products of the SELEX process, and could affect the diversity of the enriched aptamers. There are several studies that suggest practical methods to resolve this issue, or reduce its effect. These include PCR reconditioning [8-12]. In this article, we describe a practical procedure that corrects the shape and the size of the PCR artifacts that occur extensively during the process of DNA aptamer selection. We assumed that these by–products are forming due to the structure of the template. To reduce abnormal or undesirable hybridization, we applied thermal treatment that partially denatures these dsDNA at the site of hetero-duplexes, sustaining homo-duplexes formation. No enzymatic treatment is required therein. With this method, the fully hybridized homo-duplexes would not be affected due to their high stability (very low free energy change, ΔG) value [13, 14]. In other words, very high affinity between the primer-template couple. The thermal cycles we applied involved heating the PCR by–products up to 65°C for 10 minutes and cooling them down to 20°C for 1 minute. The procedure includes repeating this cycle for 6X. The visible outcome showed significant reduction in the ladder–like and smearing bands in PCR by–products. Illumina sequencing of the corrected PCR products revealed a marked reduction in the proportion of the shorter and longer sequences over the course of the treatment method in comparison to control, untreated PCR by–products.

2. Materials and Methods

2.1. Library and primers

A set of primers, forward and reverse, and a single-stranded DNA (ssDNA) library, were designed (https://eu.idtdna.com/pages/tools/oligoanalyzer) and ordered from Integrated DNA Technologies (IDT, Coralville, Iowa, USA). The ssDNA library was designed to have two parts; two constant regions (20 nucleotide sequences each) that are flanking a variable region in the middle of 40 nucleotides random sequences. The fixed regions of known sequences at upstream and downstream were to anneal the forward primer at the 5-prime (5') end and reverse primer binding site at the 3-prime (3') end. The overall library sequence design was 5'-CCGAATTCG AAGGACAAGAG $(N)_{40}$ TCTTTTATGCTACGTCCCGC-3', where the randomized region (N)40, is with randomly-incorporated nucleotide bases (i.e., A, C, G, T) at equal opportunity in each position. The sense and the antisense primers sequences were: 5'-CCGAATTCGAAG GAC AAGAG-3' and 5'-GCGGGACGTAGCATAAAAGA-3, respectively as shown in Table 1. The antisense primer was modulated by phosphorylation at the 5' end to sustain digestion with λ–exonuclease enzyme during the repeated steps of the enrichment SELEX cycles. The oligos (the ssDNA library and the primers) were diluted in 1 X tris–EDTA buffer (TE buffer, Thermo Fisher Scientific, USA) to have 100µM final concentration as recommended by the manufacturer. These diluted oligos were stored in -20°C.

2.2. Target protein

For our experiment, we used the extra–domain of the chicken CD40 receptor (CD_{40}EDR) as the target protein. The extra–domain motif of the CD_{40}EDR protein was ordered as expressed in the HEK293 Free–Style cells expression system in protein-free medium by Antagene (Santa Clara, CA, USA). The stock of the CD_{40}EDR protein was diluted with 1X sterile phosphate buffer saline (PBS 1 X, w/v) as recommended by the manufacturer, to final concentration of 15µg/L. Several aliquots in 1 ml sterile test tubes have been made and was stored at –80°C. aliquots in 1 ml sterile test tubes have been made and was stored at -80°C.

Table 1. The common by–product sequences in PCR after 10 rounds of aptamers selection procedure. **A)** The common oligos design with barcode for Illumina included. **B)** Reverse primer (RP) shows a diverse sequence. Many aptamers' sequences showed extra pieces carrying partial RP sequences. **C)** The extra piece of a long aptamers carrying the original sequence of the forward primer binding site (FPBS). **D)** The long extra piece showing overlapping in both FPBS and RP sequences. **E)** Normal aptamers length (40nt) showing partial RP sequence. All sequences are in 5'→3' direction

A				
Barcode	FPBSsequence	ssDNA Library	Aptamers	RPsequence
CTT CGG GTA TCC TCT T	TCCGAATTCGAAGGA CAAGAG	5'-CCGAATTCGAAG GACAAGAG$(N)_{40}$TCTTTT ATGCTACGTCCCGC-3'	CCACCGCTGTCCGTA AACACAGTTACTTAC GTCCACCTAG	TCTTTTATGCTA CGTCCCGC
B				
FPBS	Aptamers	RP*	Extra–piece with RP sequence	
TCCGAATTCGAAGGA CAAGAG	N40	TCTTTTTTTCTTCCTCCC CC	AGATCGGAAGCGCCCCTTCTGACCTCCAT TCCCATCTCTTTTTCCTTCTTCTT	
*RPmodified sequence				
ǂAptamers with extra piece that carrying partial RP sequence (bold font, highlighted gray)				
C				
FPBS	Aptamer	RP	Extra–piece with FB* sequence	
TCCGAATTCGAAGGA CAAGAG	N40	TCTTTTATGCTACGT CCCGC	TTCATCCGAATTCGAAGGACAAGAGGTTAT TTGAGCTCTATTCACGAGTTCACT	
*Normal aptamers with extra piece carrying RP sequence (bold font-highlighted gray)				
D				
FPBS	Aptamer	RP	Extra–piece with overlapped FP-RP* sequence	
TCCGAATTCGAAGGA CAAGAG	N40	TCTTTTATGCTACGT CCCGC	AGTCATGCATTCCGAATTCTTTTATGCTACGTC CCGC	

* Normal aptamers with extra piece showing overlapped FPBS binding site (underlined) and RP sequence (bold font, highlighted gray)

E		
FPBS	Aptamers with partial RP sequence	RP
TCCGAATTCGAAGGACAAGAG	TCTGAATGGCGATCTTAATTGTCTTTCCTGATACCTCCCG	TCTTTTATGCTACGTCCCGC

2.3. SELEX procedure

Our SELEX procedure included ten enrichment cycles of developing specific ssDNA aptamers to the CD_{40}EDR as the target protein. As previously described [4, 5], the SELEX procedure was performed. Briefly, we started by denaturing the ssDNA by heating up to 95°C for 10 min, cooling down to 4°C for 5 min and incubating the oligos at room temperature (RT) for 10 min. Next, 35µL from the oligos library was mixed with a 115 µL binding buffer ([100 mM NaCl], [10 mM Tris], [5 mM KCl], [2 mM $MgCl_2$], pH=7.5). The mixture was filtered through a nitrocellulose filter membrane to exclude non-specific bounders. 50 µL of the target protein was incubated with ssDNA library–in–binding buffer mixture under rotation for 1hr at RT. A processes of filtration, elution was carried out to specify the bound aptamers. The resulting first round enriched aptamers were precipitated by ethyl alcohol precipitation procedure. To amplify these aptamers, PCR was used. To repeat this procedure, the PCR products were digested into ssDNA by λ–exonuclease (Cat no. M0262S, NEB, USA). The rest nine rounds of the procedure and the buffers used were as described elsewhere [15]. However, the ratio of ssDNA to protein was shifted by 75% in each round toward the ssDNA to increase the affinity and the specificity of the aptamers.

2.4. Amplification and digestion of the DNA

As mentioned above, the SELEX enrichment aptamers are ssDNA and these aptamers are amplified by PCR. Therefore, we applied 25 thermal cycles using aptamers as the template, sense and antisense primers, and GoTaq® master mix (G2 Hot–Start Green Master Mixes, Cat no. M7422, Promega) as the amplification and buffer solution. The reaction conditions of the PCR were as recommended by the manufacturer. The resulted dsDNA was monitored on 1% agarose gel. The dsDNA was subjected to digestion with exonuclease enzyme (λ–exonuclease, Cat no. M0262S, NEB, USA) to be converted into ssDNA. This step is critical to start the next cycle of the SELEX. The digestion procedures were conducted according to the manufacturer's manual. After each treatment, the DNA is precipitated and purified by ethanol precipitation protocol.

2.5. Correcting thermal cycles

PCR products were loaded on 1% agarose gel, and the undesirable DNA–in–gel that were showing different length and appearance were cut and purified using DNA extraction kit (QIAquick Gel Extraction kit, Cat no. 28704, Qiagen, USA). All the procedures were according to the manufacturer's manual. The resulting DNA was re–suspended in 1XTE buffer. This DNA was subjected to six thermal cycles consisting of heating up the DNA to 65°C for 10 minutes and then cooling down the suspension to 20°C for 1 minute.

2.6. *Next-generation illumina sequencing*

Illumina sequencing was performed to specify the PCR artifacts depending on the DNA sequence. The procedure as was previously described by Scoville and coworkers [2]. The output of the illumina sequencing was analyzed using bioinformatics analysis by a free online tool called "Galaxy Bioinformatics" at: https://usegalaxy.org/ [16]. Briefly, PCR products from SELEX 1^{st}, 6^{th}, 8^{th} and 10^{th} rounds were prepared for Illumina sequencing. This included barcode insertion into sequences Table 1A. The DNA from these rounds was used as a template for amplification using specific primers with (Illumina adapters). The DNA was diluted using nuclease-free deionized distilled H_2O (ddH_2O) to final concentration of 10 µM. The illumina adapters, which were tagged with unique barcode index sequences for each sample, were included in the amplification of the templates. Accordingly, 25 PCR cycles were applied to prepare the sample for Illumina using (Q5® hot–start, high–fidelity DNA polymerase and buffer, Cat no. M 0493S, NEB, USA). The reaction conditions were as recommended by the manufacturer. The resulting DNA products were combined at equal molar ratio and sent for High throughput Illumina sequencing center (University of Wisconsin Biotechnology Center, UWBC, WI, USA).

Results from the illumina sequencing were analyzed using Galaxy software (https://usegalaxy.org/), with multiple pipelines. Illumina sequencing has produced about 16 million reads. These reads were tracked and sorted according to their indexed barcodes. The quality of the DNA, their lengths and the conserved DNA sequences were evaluated as described previously [16, 17, 18].

3. **Results**

The results of the current study show that there are typical artifacts in PCR products when using ssDNA library as the template for the PCR procedure. We could specify two types of these artifacts according to the pattern that they take on 1%$ agarose gel; a ladder-like bands, and smearing band Fig.1.

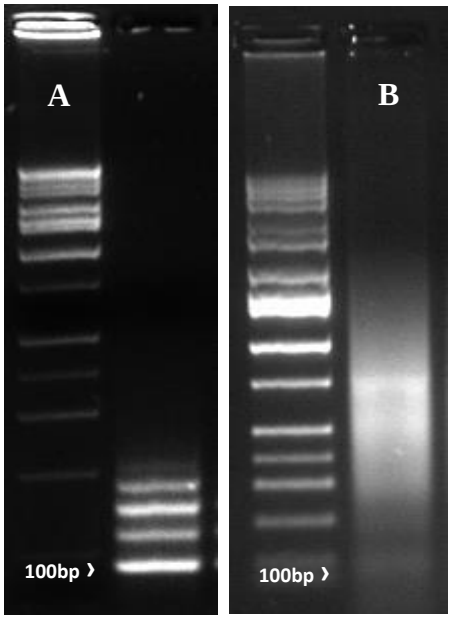

Fig. 1. Patterns of PCRartifacts during SELEX procedure. PCR products from SELEX rounds (8 and 10) on 1% agarose gel. Patterns of the PCR by–products take two different categories (A) ladder–like patterns. (B) Smearing–band pattern. According to our olig design, the correct size of the aptamers is 80 bp.

The bands in the ladder-like patterns are the doubling of the original size of the template (80bp according to our ssDNA library design). The smearing band could have the actual size of the aptamers included as a distinctive band. The ladder-like artifact pattern increases drastically with the increasing rounds of the SELEX procedure Fig. 2. In the later rounds of the SELEX, most of the PCR products are actually artifacts. Samples from rounds 1, 6, 8, and 10 showed that the proportion of the longer and the shorter reads are dominantly and linearly increasing with the rounds Table 2. To investigate the formation of these artifacts with the PCR, we carried out a n experiment using the outcome of the ladder-like and the smearing bands as templates with and without primers or with single primer for amplification. Surprisingly, we could amplify the ssDNA library with all of the possibilities Fig. 3.

Table 2. ssDNA aptamers normal, short and long sequences. The proportion of the short sequences (< 80 nt %) and long sequences (> 80 nt %) in comparison to normal length sequences reads (80 nt %). The total sequence shows that there is an increase in the percentage of the short and the long sequences drastically when the rounds proceed. In these data, no correction procedure was applied

Rounds	Reads (millions)	< 80 nt %	80 nt %	> 80 nt %
1st	4.441.562	0.657	95.678	3.659
6th	4.379.389	1	96.8	2.2
8th	4.087.274	4.331	90.008	5.661
10th	3.777.547	23.861	67.892	9.333

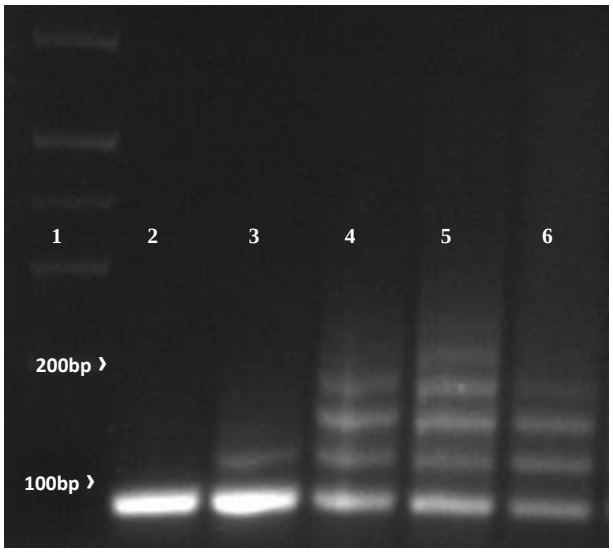

Fig. 2. Evolution of the PCRartifacts over the rounds of aptamers selection. Lane 1 represents Hi-Lo ladder. Lane 2 represents the ssDNA library, lanes 3, 4, 5 and 6 represent PCR products from SELEX rounds one, two, three, and four respectively. All the DNAs were loaded on 1% agarose gel.

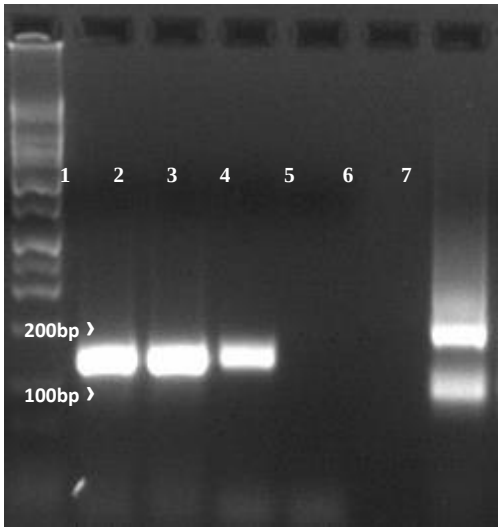

Fig. 3. Amplification of DNA library and aptamers pool with a single primer. Aptamers form round six was used as a template for PCR procedure. The process of amplification was carried out using only forward primer (lane 2) with 45 thermal cycles, reverse primer alone (lane 3), with 45 thermal cycles, forward primer (lane 4), 30 thermal cycles, or two primers, i.e., forward and reverse primers (lane 7), 45 thermal cycles. Lane 5 represents no template added and lane 6 represents no primer.

To reduce these effects, we apply our protocol of thermal treatment of the PCR artifacts. The thermal treatment could reduce the presence, the pattern and the concentration of the artifacts drastically Fig. 4. It found that 6X cycles is the least–required number for best results.

To follow these by–products after the application of our thermal protocol, Next–Generation Illumina sequencing method has been applied. Accordingly, the correction was significant in later rounds (round 10) and obvious by reduction of the number and percentage of the shorter and the longer reads Table 3.

Table 3. Effect of thermal treatment on the PCR by–products. The proportion of short sequences (< 80 nt %) and long sequences (> 80 nt %) in comparison to normal length sequences (80 nt %) with and without the course of the thermal correction. Samples from round ten were used to evaluate the method. Samples form round one was used as the reference round for all lengths of sequences.

Rounds	Reads (millions)	< 80 nt %	80 nt %	> 80 nt %
1st	4,441,562	0.657	95.681	3.659
10th /not–corrected	3,777,547	23.861	67.892	8.253
10th /corrected	4,549,398	0.875	95.564	3.561

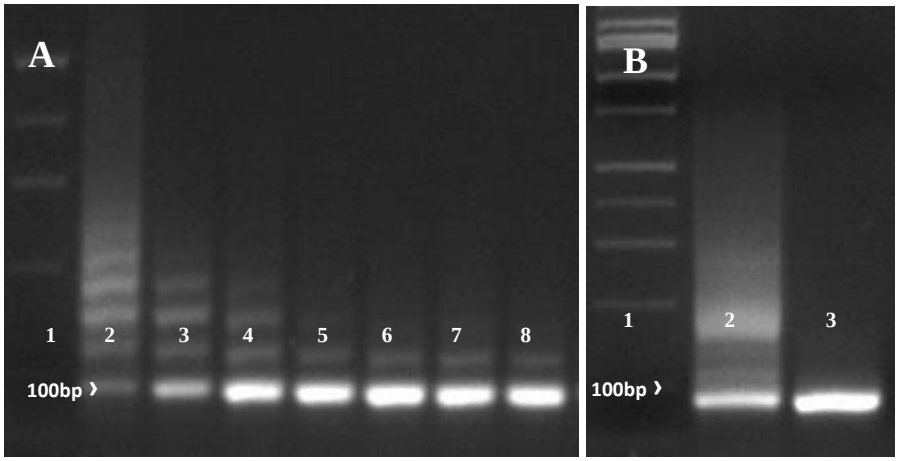

Figure 4. Application of thermal treatment for PCR by–products correction. (A): PCR products from the 10th enrichment cycle of SELEX procedure. Lane 2, the PCR by–products appear as a ladder–like pattern. Lanes 3–8, represent the correction protocol which includes application of 65ºC for 10 minutes, followed by 20ºC for 1 minute for one, two, three, four, five and six cycles, respectively. (B): correction of the smearing–band of PCR products by applying the same 6 cycles of the thermal treatment.

4. Discussion

There is a selection pressure that is usually applied on dsDNA formation due to the combination of randomized nucleotide library and SELEX procedure. These two elements are the ideal elements required for mutation in resulting aptamers typically. Being a genetic material, DNA aptamers show a tendency for mutation to have stable structure thermodynamically. For this reason, the SELEX procedure could favor the amplification of certain sequences over others due to thermos-stability and shape–to–function features [19, 20]. In addition, amplification of the DNA with PCR is a key player that shows a significant bias issue when the template DNA molecule is a random library. The main resulting

issues are incorrect or unwanted primer–template, or template–to–template hybridization events. Therefore, when the template is a randomized ssDNA library with certain length, certain DNA by–products is considerable, and when the length of the randomized sequence is increased, the artifacts could increase significantly. Moreover, the artifacts increase linearly with the number of the PCR cycles [21, 22, 23, 24, 25]. Evidently, there is an evolution of selection pressure, which is in favor of highly-fitted amplifying DNA over the highly affiant aptamers in later cycles of the SELEX. Many researchers are required to deal with the formation of the PCR undesirable by–products in order to shift this balance toward the specificity of the aptamers and to reduce the development of non-specific aptamers [22, 26, 27]. In the process of SELEX procedure, there are two noticeable artifacts in the PCR products; these are when the DNA takes ladder–shaped in multiple sizes and smearing band Fig. 1. These artifacts may appear as early as in the 2nd round of selection and may continue to increase linearly over the whole procedure Fig. 2. In explanation, these by–products are formed because of a series of undesirable events during the amplification of the aptamers. According to the randomness of the variable region of the library, there are multiple potential sites where primers can fit according to Watson–Crick base pairing. This sustains the annealing of the primer(s) other than the constant regions on the library [10, 23]. According to the Illumina sequencing results, there is a partial annealing of two DNA strands other than primers. This could form hetero–duplexes. Each strand could work as a primer for its cognate DNA. These two DNA strands act synergistically with the primer(s) to produce the amplicon [22, 28, 29, 30]. To test these hypotheses, we could amplify the ssDNA library with either or both primers Fig. 3. With one of the primers only, we amplified the library with no development of ladder–like DNA or smear band. However, the length of the PCR products of this experiment was not identical with the typical library length. This finding could be attributed to the fact that there are specific sequences that have a high tendency to suit each other to form hetero–duplexes more than the others. In the presence of a single primer, the complementary sequence resulting from the previous round of the PCR acts as the second primer and *vice versa*. Therein, the size of the PCR by–products depend on the site of where the two templates hybridize more than the actual site where the primer–template hybridizes. We concluded that this by–product could be <2X the library size at least over the first round of the PCR cycle. When the cycles continue, the by–products size increases linearly due to the formation of more casts of the DNA [28].

According to our hypothesis, we sought to apply thermal treatment to decrease the potential of hetero–duplexes formation in PCR products. The highest temperature we applied was meant to be lower than the normal denaturing temperature of the fully-hybridized templates and above the annealing temperature of the primer-template. In this direction, we tested several different denaturing temperatures (data not shown) and we spotted 65ºC as the ideal temperature for partial denaturation of the DNA. The lower limit, i.e., 20ºC, over the thermal course, is to sustain the homo–duplexes to re-anneal Fig. 4. Therefore, there is no full and complete denaturation for the homo–duplexes due to their very low ΔG value. Instead, the hetero–duplexes are applied under the pressure to find their cognate complementary sequences due to the relatively high ΔG value. The number of thermal cycles required were evaluated and we found that six cycles of the thermal treatment is the least significant procedure that gives the best results. Both the ladder–like shape and smearing band apparent are sensitive to the correction with this method and only traces of either by–products could be detected in the aftermath Fig. 4 [13, 14]. Illumina sequencing results revealed the presence of several artifacts over the PCR products. There are stretches of DNA over the aptamers sequences where they were partially matching the reverse primer sequence (Table 1D). This agrees with our explanation for the origin of the PCR by–products and agrees with what concluded by [10]. In addition, there were extra-piece of oligo sequences in the downstream of the aptamers where it, in most cases, expresses partial or full forward primer sequence, partial or full reverse primer sequence or both. Also, it appears as an overlapping DNA sequence over the aptamers Table 1. The format of illumine sequencing was analyzed with bioinformatics software, and the typical illumine sequence results are as listed in Table 1. After ten rounds of ssDNA aptamers selection, the common PCR artifacts and by–products are listed in Table 1. In these results, some aptamers showed incorrect sequence of the reverse primer (RP) included. This could be attributed to the polymerase enzyme proofreading capacity or presence of highly consensus sequences [29, 30].

Some aptamers sequences appear longer than designed one, and the extra piece is actually carrying partial or complete RP sequence or forward primer (FP) binding site sequence. This result matches with the conclusion of previous publication by [10]. In some instances, the long extra piece shows overlapped FP binding site and overlapped RP sequences Table 1.

There were short stretches of DNA sequences that did not carry any matching primer or primer binding site sequences. Both short and long sequences were increasing linearly over the SELEX rounds. However, PCR by–products that have been treated with the thermal correction protocol showed significant tendency to repair these artifacts. In such PCR products, the short reads proportion was with steep increase in comparison to the total reads. This could be explained as the short reads are the results from the amplification of homo-duplexes and hetero-duplexes, whereas the longer reads are the results of the hetero-duplexes only Table 2. Our finding suggests that the application of the correction thermal cycle for the PCR artifacts in the 10th round was successful in shifting the proportion of short and long reads toward the normal sequence size Table 3. Our observation was that the smearing bands of the PCR by–products were more sensitive to the thermal correction protocol than the ladder–like products Fig. 4. This could be explained that the smearing bands are due to the formation of ssDNA more often and these oligos tends to form three–dimensional structure hairpin, therefore, they are more sensitive to thermal treatment. The ladder–like by–products, on the other hand, are formed as hetero-duplexes, and these structures are more refractory to elevated temperature. Unfortunately, we were facing the ladder–like by–products in our experiment mostly [31].

5. Conclusion

The use of PCR is the most reliable method used for exponential amplification of the enriched ssDNA aptamers in the SELEX procedure. PCR, however, could diminish the efficiency and the potency of the SELEX method itself due to the presence of several artifacts over the enriched aptamers. In fact, the presence of these artifacts or by–products in the aptamers may truly interfere with optimum ligand-target interaction. This would result in low quality aptamers or failure in aptamers selection. Several labs have described methods to decrease these artifacts in PCR products with various outcomes. PCR reconditioning is a leading procedure in reducing these artifacts. In the current paper, we describe a simple yet, efficient protocol to reduce the most common PCR by–products confronting the developing functional aptamers. Our method treats the two more common artifacts, i.e., the ladder–like and the smearing band PCR by–products at the same efficiency. Visually, the agarose gel images showed that our protocol was capable of effective correction of these by–products. Similarly, the illumine sequencing results showed that there was marked reduction in proportion of the short and long sequences after the use of this protocol. The formation of these short reads

or long reads PCR by–products might be due to the presence of matched sequences within the random library sequence with that of FPBS or RP sequences. Additionally, the formation of hetero–duplexes between two random sequences within the library is common, and could act as a sensing primer to form multiple-length PCR products. Finally, the ladder–like by–products were more refractory to thermal correction than the smearing bands counterpart by–products.

Acknowledgement

First author was sponsored by the University of Arkansas, Fayetteville, AR 72701, USA. Department of Cell and Molecular Biology.

Nomenclature & Symbols			
ssDNA	Single-stranded DNA	SELEX	systemic evolution of ligands by exponential enrichment
dsDNA	Double-stranded DNA	ΔG	Free energy change
PCR	Polymerase Chain Reaction	FPBS	Forward primer-binding site
RP	Reverse primer	$CD_{40}EDR$	CD40 extra-domain receptor
FP	Forward primer	RT	Room temperature
EDTA	Ethylene-diamine-tetra acetic acid	TE	Tris-EDTA
ddH_2O	deionized distilled H_2O	Oligos	Oligonucleotides
nt	Nucleotides	NGS	Next–Generation Sequencing

References

[1] Dollins, C. M., Nair, S., & Sullenger, B. A. (2008). Aptamers in immunotherapy. Human gene therapy, 19(5), 443-450. https://doi:10.1089/hum.2008.045.

[2] Scoville, D. J., Uhm, T. K. B., Shallcross, J. A., & Whelan, R. J. (2017). Selection of DNA aptamers for ovarian cancer biomarker CA125 using one-Pot SELEX and high-throughput sequencing. Journal of nucleic acids, 2017(1), 9879135. http://dx.doi.org/10.1155/2017/9879135.

[3] Liu, M., Wang, L., Lo, Y., Shiu, S. C. C., Kinghorn, A. B., & Tanner, J. A. (2022). Aptamer-enabled nanomaterials for therapeutics, drug targeting and imaging. Cells, 11(1), 159. doi: 10.3390/cells11010159.

[4] Ellington, A. D., & Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. nature, 346(6287), 818-822. https://doi.org/10.1038/346818a0.

[5] Tuerk, C., & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. science, 249(4968), 505-510. DOI: 10.1126/science.2200121.

[6] Vorobyeva, M. A., Davydova, A. S., Vorobjev, P. E., Pyshnyi, D. V., & Venyaminova, A. G. (2018). Key aspects of nucleic acid library design for in vitro selection. International journal of molecular sciences, 19(2), 470. https://doi.org/10.3390/ijms19020470.

[7] Avci-Adali, M., Paul, A., Wilhelm, N., Ziemer, G., & Wendel, H. P. (2009). Upgrading SELEX technology by using lambda exonuclease digestion for single-stranded DNA generation. Molecules, 15(1), 1-11. https://doi.org/10.3390/molecules15010001.

[8] Thompson, J. R., Marcelino, L. A., & Polz, M. F. (2002). Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR'. Nucleic Acids Research, 30(9), 2083-2088. https://doi.org/10.1093/nar/30.9.2083.

[9] Smyth, R., Schlub, T., Grimm, A., Venturi, V., Chopra, A., Mallal, S., ... & Mak, J. (2010). Reducing chimera formation during PCR amplification. https://doi.org/10.1016/j.gene.2010.08.009.

[10] Tolle, F., Wilke, J., Wengel, J., & Mayer, G. (2014). By-product formation in repetitive PCR amplification of DNA libraries during SELEX. PloS one, 9(12), e114693. DOI:10.1371/journal.pone.0114693.

[11] Qiu, X., Wu, L., Huang, H., McDonel, P. E., Palumbo, A. V., Tiedje, J. M., & Zhou, J. (2001). Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. Applied and environmental microbiology, 67(2), 880-887. DOI: https://doi.org/10.1128/AEM.67.2.880-887.2001.

[12] Michu, E., Mráčková, M., Vyskot, B., & Žlůvová, J. (2010). Reduction of heteroduplex formation in PCR amplification. Biologia plantarum, 54, 173-176. https://doi.org/10.1007/s10535-010-0029-8.

[13] Golyshev, V. M., Pyshnyi, D. V., & Lomzov, A. A. (2021). Calculation of energy for RNA/RNA and DNA/RNA duplex formation by molecular dynamics simulation. Molecular Biology, 55(6), 927-940. https://doi.org/10.1134/S002689332105006X.

[14] Nüesch, M. F., Pietrek, L., Holmstrom, E. D., Nettels, D., von Roten, V., Kronenberg-Tenga, R., ... & Schuler, B. (2024). Nanosecond chain dynamics of single-stranded nucleic acids. Nature Communications, 15(1), 6010. https://doi.org/10.1038/s41467-024-50092-8.

[15] Wang, R., & Li, Y. (2013). Hydrogel based QCM aptasensor for detection of avian influenzavirus. Biosensors and Bioelectronics, 42, 148-155. https://doi.org/10.1016/j.bios.2012.10.038.

[16] Thiel, W. H., & Giangrande, P. H. (2016). Analyzing HT-SELEX data with the Galaxy Project tools–A web based bioinformatics platform for biomedical research. Methods, 97, 3-10. doi:10.1016/j.ymeth.2015.10.008.

[17] Thiel, W. H. (2016). Galaxy workflows for web-based bioinformatics analysis of aptamer high-throughput sequencing data. Molecular Therapy-Nucleic Acids, 5. http://doi:10.1038/mtna.2016.54.

[18] The Galaxy Community.(2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Research. 52(W1): W83–W94. https://doi.org/10.1093/nar/gkae410.

[19] Meyers, L. A., Lee, J. F., Cowperthwaite, M., & Ellington, A. D. (2004). The robustness of naturally and artificially selected nucleic acid secondary structures. Journal of Molecular Evolution, 58, 681-691. https://doi.org/10.1007/s00239-004-2590-2.

[20] Thiel, W. H., Bair, T., Wyatt Thiel, K., Dassie, J. P., Rockey, W. M., Howell, C. A., ... & Giangrande, P. H. (2011). Nucleotide bias observed with a short SELEX RNA aptamer library. Nucleic acid therapeutics, 21(4), 253-263. https://doi.org/10.1089/nat.2011.0288.

[21] Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., & Polz, M. F. (2005). PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Applied and environmental microbiology, 71(12), 8966-8969. DOI: https://doi.org/10.1128/AEM.71.12.8966-8969.2005.

[22] Shao, K., Ding, W., Wang, F., Li, H., Ma, D., & Wang, H. (2011). Emulsion PCR: a high efficient way of PCR amplification of random DNA libraries in aptamer selection. PloS one, 6(9), e24910. doi:10.1371/journal.pone.0024910.

[23] Yufa, R., Krylova, S. M., Bruce, C., Bagg, E. A., Schofield, C. J., & Krylov, S. N. (2015). Emulsion PCR significantly improves nonequilibrium capillary electrophoresis of equilibrium mixtures-based aptamer selection: allowing for efficient and rapid selection of aptamer to unmodified ABH2 protein. Analytical chemistry, 87(2), 1411-1419. https://pubmed.ncbi.nlm.nih.gov/25495441/.

[24] Takahashi, M., Wu, X., Ho, M., Chomchan, P., Rossi, J. J., Burnett, J. C., & Zhou, J. (2016). High throughput sequencing analysis of RNA libraries reveals the influences of initial library and PCR methods on SELEX efficiency. Scientific reports, 6(1), 33697. DOI: 10.1038/srep33697.

[25] Witt, M., Phung, N. L., Stalke, A., Walter, J. G., Stahl, F., von Neuhoff, N., & Scheper, T. (2017). Comparing two conventional methods of emulsion PCR and optimizing of Tegosoft-based emulsion PCR. Engineering in life sciences, 17(8), 953-958. https://doi.org/10.1002/elsc.201700047.

[26] Eaton, R. M., Shallcross, J. A., Mael, L. E., Mears, K. S., Minkoff, L., Scoville, D. J., & Whelan, R. J. (2015). Selection of DNA aptamers for ovarian cancer biomarker HE4 using CE-SELEX and high-throughput sequencing. Analytical and bioanalytical chemistry, 407, 6965-6973. https://doi.org/10.1007/s00216-015-8665-7.

[27] Blind, M., & Blank, M. (2015). Aptamer selection technology and recent advances. Molecular Therapy-Nucleic Acids, 4. doi:10.1038/mtna.2014.74.

[28] Cheng, L. Y., Dai, P., Wu, L. R., Patel, A. A., & Zhang, D. Y. (2022). Direct capture and sequencing reveal ultra-short single-stranded DNA in biofluids. Iscience, 25(10). https://www.cell.com/iscience/fulltext/S2589-0042(22)01318-9.

[29] McInerney, P., Adams, P., & Hadi, M. Z. (2014). Error rate comparison during polymerase chain reaction by DNA polymerase. Molecular biology international, 2014(1), 287430. http://dx.doi.org/10.1155/2014/287430.

[30] Musheev, M. U., & Krylov, S. N. (2006). Selection of aptamers by systematic evolution of ligands by exponential enrichment: addressing the polymerase chain reaction issue. Analytica chimica acta, 564(1), 91-96. https://doi.org/10.1016/j.aca.2005.09.069.

[31] Krause, N. M., Bains, J. K., Blechar, J., Richter, C., Bessi, I., Grote, P., ... & Schwalbe, H. (2024). Biophysical Investigation of RNA· DNA: DNA Triple Helix and RNA: DNA Heteroduplex Formation by the lncRNAs MEG3 and Fendrr. ChemBioChem, 25(10), e202400049. doi.org/10.1002/cbic.202400049.