Title: Label-free selection of RNA aptamers for metabolic engineering

Chuhern Hwang and James M. Carothers*

Departments of Chemical Engineering and Bioengineering, Molecular Engineering and Sciences Institute, and Center for Synthetic Biology, University of Washington, Seattle WA 98195, United States

*Corresponding author: James M. Carothers, jcaroth@uw.edu

Abstract

RNA aptamers can be assembled into genetic regulatory devices that sense and respond to levels of specific cellular metabolites and thus serve an integral part of designing dynamic control into engineered metabolic pathways. Here, we describe a practical method for generating specific and high affinity aptamers to enable the wider use of *in vitro* selection and a broader application of aptamers for metabolic engineering. Conventional selection methods involving either radioactive labeling of RNA or the use of label-free methods such as SPR to track aptamer enrichment require resources that are not widely accessible to research groups. We present a label-free selection method that uses small volume spectrophotometers to track RNA enrichment paired with previously characterized affinity chromatography methods. Borrowing techniques used in solid phase peptide synthesis, we present an approach for immobilizing a wide range of metabolites to an amino PEGA matrix. As an illustration, we detail laboratory techniques employed to generate aptamers that bind *p*-aminophenylalanine, a metabolic precursor for bio-based production of plastics and the pristinamycin family of antibiotics. We focused on the development of methods for ligand immobilization, selection via affinity chromatography, and nucleic acid quantification that can be performed with common laboratory equipment.

Keywords

Label-free aptamer selection; ligand immobilization for *in vitro* selection

Highlights

- Small-volume spectrophotometers used for label-free *in vitro* RNA aptamer selection
- Solid phase peptide synthesis chemistry used for ligand immobilization
- Method is applied for selection of an RNA aptamer against *p*-aminophenylalanine

1. Introduction

Naturally occurring metabolic pathways have evolved mechanisms to sense and respond dynamically to changes in metabolite flux. Programming dynamic control by sensing and responding to cellular conditions may thus improve performance and titers of engineered metabolic pathways [1,2] . Sensitive and specific biosensors that can function in physiological conditions are necessary to implement dynamic control. Ligand-binding RNA aptamers are biosensors that can be incorporated into programmed metabolic pathways and can be linked to genetic control parts by selection or rational design [3,4] . Thus, generating specific and high affinity aptamers can be useful for evolving or engineering metabolic pathways to yield optimal titers [5] .

RNA aptamers can be identified from large nucleic acid pools of 10^{14} to 10^{16} sequences through iterative *in vitro* selection [6] . Since the development of *in vitro* selection, or the systematic evolution of ligands by exponential enrichment (SELEX), for the identification of aptamers, many variants on SELEX, such as capillary electrophoresis SELEX (CE SELEX) or surface plasmon resonance SELEX (SPR SELEX) have been developed [7–9] . However, performing SELEX using many of these variants is either insufficient for generating high affinity RNA aptamers for metabolic engineering or requires costly equipment or facilities that are out of reach for many research groups intending to select aptamers for their ligand of choice. We detail an approach for *in vitro* selection of high affinity RNA aptamers that utilizes well established affinity chromatography techniques, uses a ligand immobilization approach that can be easily applied to many different classes of small molecules, and is label-free while avoiding other costly methods for tracking enrichment for binders.

The progress and success of an aptamer selection is monitored by quantifying the relative amount of RNA that is eluted with the ligand after each round of selection. Many current approaches for nucleic acid quantification require radioactive labeling [6] ; label-free methods, such as monitoring the progress of a selection by surface plasmon resonance (SPR), often require expensive or specialized equipment [7] . Although these methods have high sensitivity for tracking enrichment for binders, they are impractical for many researchers due to cost or lack of access. While not as sensitive as radioactive labelling or SPR, quantifying nucleic acid using a small volume spectrophotometer such as a Thermo Scientific Nanodrop can still accurately track the progress of an RNA aptamer selection as long as sufficient quantities of RNA are transcribed and applied to the ligand-immobilized column.

Affinity chromatography also holds advantages over newer developments such as CE SELEX or microfluidic SELEX (M SELEX). Large pool sizes of 10^{14} to 10^{16} are necessary for the selection of high affinity aptamers for metabolic engineering. The informational complexity of high affinity aptamers for small molecules makes the expected number of such aptamers extremely low even in pools of over 10^{14} sequences [10] . Volume limitations make M SELEX and CE SELEX impractical for selecting from pools of larger than about 10^{13} - 10^{14} sequences [8,9] ; affinity chromatography on columns does not face a similar limitation, making it more suitable for selection of high affinity aptamers. Affinity chromatography on a column also provides a straightforward method of generating specific aptamers through counterselection. High affinity aptamers are not necessarily highly specific aptamers, so counterselection using chemically similar compounds that occur physiologically can screen out non-specific binders [11] . Affinity chromatography also allows for off-rate selection in physiological buffer conditions.

The immobilization chemistry described here has been applied for immobilizing fluorophores such as tetramethylrodamine (TMR) and amino acid derivatives such as *p*-aminophenylalanine (pAF) and *p*-amino cinnamic acid (pACA) to an amino polyethylene glycol polyacrylamide copolymer (PEGA) resin for affinity chromatography. It could be further applied for immobilizing other amino acid derivatives, fluorophores, or aromatics, including compounds relevant to recent engineered metabolic pathways [12,13]. Immobilization of compounds lacking amines skips the requirement for fluorenylmethyloxycarbonyl (Fmoc) protection of amines and subsequent Fmoc deprotection, while the presence of multiple amines as in the case of pAF requires not just Fmoc but also tert-butyloxycarbonyl (Boc) protection. Fmoc and Boc protected amino acid derivatives are widely available from suppliers such as Bachem and Anaspec. The immobilization chemistry described derives from the use of Fmoc-protected amino acids for solid phase peptide synthesis [14,15] and can be implemented by most labs that conduct chemical or biological work. Here, we detail the immobilization of pAF, a precursor for bio-based production of plastics and the pristinamycin family of antibiotics [1,16]. The method detailed below successfully generated two high affinity aptamers for pAF and was also used to generate high affinity aptamers for TMR [17].

2. Approach

2.1. Ligand immobilization onto PEGA column.

Ligand immobilization of pAF was achieved through three steps: coupling of the ligand, acetylation of unreacted amines, and Fmoc/Boc deprotection. Initially, Fmoc and Boc protected pAF was coupled onto an amino PEGA resin (Figure 1A). The unreacted amines on the resin were then acetylated with acetic anhydride to prevent RNA from sticking to the unreacted amines (Figure 1B). After the unreacted amines were capped, the Fmoc and Boc groups protecting the amines on pAF were deprotected, leaving pAF immobilized on the amino PEGA resin with all other amines on the resin acetylated and capped (Figure 1C).

2.1.1. Coupling.

200 μ L selection columns were made by coupling Fmoc-*p*-aminophenylalanine-Boc (Fmoc-pAF-Boc) (Bachem B-1995) to an amino PEGA matrix. 0.11 g amino PEGA beads (Novabiochem 855015) were swelled in 40 column volumes (CVs) water in Bio-Rad Poly-Prep Chromatography Columns (Bio-Rad 7311550), where a column volume was defined as 200 μ L. Amino PEGA was then washed with 40 CVs dimethylformamide (DMF). 50 μ L of 12 mM Fmoc-pAF-Boc in DMF, 100 μ L of 12 mM diisopropylethylamine (DIPEA) in DMF, and 50 μ L of 12 mM PyBOP in DMF, were added to the 200 μ L of amino PEGA resin for coupling. The column was covered with foil to decrease exposure to light and was shaken on a small rotator at 30 rpm at room temperature overnight (12-18 hours). In this reaction, Fmoc-pAF-Boc couples with the amino PEGA resin by the forming a peptide bond.

2.1.2. Acetylation.

At the end of overnight coupling, the column was washed with 40 CVs of DMF by filling the column with 8 mL DMF, capping it, agitating the column, and then draining the DMF after the beads settled. The column was then acetylated with a mix of 16.8 μ L acetic anhydride, 31.2 μ L DIPEA, and 152 μ L DMF. The column was covered with foil and the acetylation was shaken at room temperature for 2 hours. Acetylation caps the amine ends of the amino PEGA resin and prevents aptamers that can interact with the amine end of pAF from interacting with the amines on the PEGA resin. At the end of acetylation, the column was washed with 40 CVs of DMF.

2.1.3. Deprotection.

Boc groups were deprotected from the coupled Fmoc-pAF-Boc using a solution of 4.75 ml trifluoroacetic acid (TFA), 125 μ L water, and 125 μ L triisopropylsilane (TIPS) at room temperature, covered, for 180 minutes. After Boc deprotection, the column was washed with 40 CVs of DMF. Fmoc deprotection was performed by shaking the column with 1 mL 20% piperidine in DMF for 30 minutes. The absorbance of the deprotection flow through at 301 nm was measured to determine the amount of Fmoc released from the column and thus the extent of coupling; most columns exhibited a final concentration of 0.5 mM pAF immobilized to the amino PEGA resin. Fmoc and Boc deprotection removed the protecting groups from the amines on pAF, leaving just pAF conjugated to the amino PEGA resin by a peptide bond.

2.1.4. Column preparation for affinity chromatography.

The column was washed again with 40 CVs of DMF, then soaked with 40 CVs of water and 20 CVs of selection buffer (SB) for half an hour each. The water and selection buffer washes were important due to the different swelling properties of the PEGA resin in DMF and water. Selection buffer was made with 130 mM K glutamate (pH 7.5), 15 mM NaCl, 10 mM dithiothreitol (DTT), and varying levels of MgCl_2 from 1 to 10 mM. After incubation with SB, the column was ready for affinity chromatography.

2.2. Preparation of DNA pool.

A library of 10^{14} DNA oligonucleotide sequences was generated using a previously described pool design strategy at the Keck Facility [17]. Similar randomized DNA pools can be generated by Integrated DNA Technologies or the vendor of your choice. The pool was resuspended in water and size purified using an 8% PAGE gel and a Whatman Elutrap (GE Healthcare). Primer extension was used to convert the oligonucleotides into dsDNA.

2.3. *In vitro* transcription and preparation of RNA.

In each round, candidate RNA sequences were produced from the DNA pool in 0.5 μ M *in vitro* transcriptions. Transcriptions were performed in 25 mM MgCl_2 , 10 mM DTT, 2.5 mM spermidine, 40 mM Tris (pH 7.8), and 0.01% Triton X-100. 9 mM GTP and 5 mM each of the other rNTPs were added along with 1 U/mL Thermostable Inorganic Pyrophosphatase (NEB M0296), 2.5 U/ μ L T7 RNA Polymerase (NEB M0251), and 1 U/ μ L RiboLock RNase inhibitor (ThermoFisher EO0381). Transcription of the DNA pool in the first round was determined by the size of the DNA pool; a nanomole pool (10^{14} sequences) would require a 2 mL transcription volume and transcribing a larger pool requires scaling accordingly. Subsequent transcriptions were performed at 200 μ L volumes to produce enough RNA for label-free detection. Transcription buffer should be made fresh to ensure high transcription yields; all buffer components should be stored at -20 °C prior to use.

Transcriptions were performed overnight (12-18 hours) at 37 °C. Transcriptions were terminated by adding DNaseI (NEB M0303), 0.5 U/ μ g DNA, for 30 minutes at 37 °C. At the end of DNaseI incubation, 25 mM EDTA was added and DNaseI was heat inactivated at 75 °C for 10 minutes. 200 μ L transcriptions were then desalted using NAP-5 Sephadex columns (GE Healthcare 17085302) and RNA was eluted in water. RNA in water was ethanol precipitated by combining up to 400 μ L RNA with 1 μ L GlycoBlue (ThermoFisher AM9516) and enough KCl for a combined concentration of 0.3 M. 3 volumes of 100% ethanol were then added and the mixture was set at -80 °C for 30 minutes, then spun down at 14000 g in a microcentrifuge at 4 °C for 10 minutes. The supernatant was discarded and the RNA pellet was washed with 70% ethanol and then centrifuged at 14000 g for 3 minutes. Wash supernatant was discarded and RNA pellets were left to dry in a chemical fume hood for 20 minutes before the RNA was resuspended in water. Resuspended transcription RNA totaled between 120-200 μ g of RNA.

The resuspended RNA was performed on a PAGE gel at 20 W for 1 hour; the correct bands were identified using UV shadowing by comparison with xylene cyanol and bromophenol blue dye bands and then cut from the gel. Typically, a 100-base RNA migrates slightly slower than the xylene cyanol band on a 8% PAGE gel; RNA size can also be determined by comparison with a RNA ladder. RNA was recovered from the cut gel slices with 0.5X TBE using a Whatman Elutrap; for recovery, the Elutrap was operated at 200 volts for 2 hours. Eluted RNA was extracted from the Elutrap, ethanol precipitated with KCl, and then resuspended in water. Resuspended electroeluted RNA totaled between 60-120 µg of RNA. The total mass of RNA must not be lower than this range to ensure reliable quantification of eluted RNA from the column after affinity chromatography.

2.4. *In vitro* selection.

Ethanol precipitated RNA was resuspended in 200 µL water and then incubated at 37 °C for 5 minutes; 4 µL 200 µM yeast tRNA, 4 µL 1 mg/mL BSA, and 200 µL of 2X SB were then added. The RNA mixture was then applied to the ligand-immobilized column and incubated for 5-30 minutes. Columns were washed with 20 CVs SB, then eluted with 2 CVs 1 mM pAF in SB three times. Another 2 CVs 1 mM pAF in SB were applied to the column and collected 1 hour later. All selection steps were performed at room temperature unless otherwise noted. Off-rate selection can be performed by varying the amount of time the ligand in SB is incubated on the column.

2.5. Reverse transcription and pool amplification.

Elution fractions were pooled and desalted in NAP-25 Sephadex columns (GE Healthcare 17085201) and eluted in water. Desalted RNA was volume reduced by butanol extraction and then cleaned up again by ethanol precipitation with KCl. RNA was resuspended in 50 µL water and the nucleic acid concentration was measured by Nanodrop. RNA was reverse transcribed using illustra Ready-to-Go RT-PCR beads (GE Healthcare 27925901) at 42 °C for 1 hour. RNA was mixed with RT primer and heated to 75 °C for 5 min, then chilled on ice prior to being added to the RT-PCR reaction. Although Ready-to-Go RT-PCR beads were advertised to reverse transcribe 1 µg mRNA/bead, we found that the beads tended to be most efficient when loaded with less than 250 ng/bead, which required us to use 16 RT-PCR beads to RT the RNA collected from the first round of the selection; in later rounds of the selection 2 RT-PCR beads were used per round.

RT reactions were then pooled and cleaned up using the Qiagen PCR Purification Kit (Qiagen 28104). Purified cDNA was then PCR amplified using DreamTaq DNA Polymerase (ThermoFisher K1081) to a total of 10 µg DNA. The Qiagen PCR Purification Kit was used again to clean up the PCR-amplified DNA, which was then used for transcription of the next round of the selection. It was essential to monitor the progress of PCR reactions when amplifying the pool: overamplification of the pool allows for mispriming and the amplification of incorrect products. We monitored amplification by performing Taq PCR incrementally and using agarose gels to qualitatively assess PCR progress, although monitoring using qPCR should be effective as well. After each round, half of the amplified DNA was used for transcription in the next round while half was stored at -20 °C for later sequencing and as a backup in case of unexpected failure in a subsequent round.

After a sustained increase in the percentage of RNA recovered in the pAF elutions was observed, PCR products were cloned using TA cloning and sequenced; alternatively, the PCR amplified DNA can be assessed with the sequencing technology of your choice.

2.6. RNA quantification and label-free tracking.

During the selection, we quantified nucleic acid amount at four points in each round:

1. DNA pool quantification prior to *in vitro* transcription
2. RNA pool quantification following the ethanol precipitation after transcription
3. RNA pool quantification following the ethanol precipitation after electroelution
4. Quantification of pooled ligand-eluted RNA from affinity chromatography column following ethanol precipitation

Quantification of the DNA pool was performed at the beginning of each round to determine the volume of DNA to load in 0.5 μ M *in vitro* transcription reactions. Typically, 100 μ moles of DNA were loaded into each 200 μ L transcription. Post-transcription RNA pool quantification allowed us to monitor transcription yield. Post-electroelution RNA pool quantification was necessary to determine the amount of correct-size transcription product and thus the amount of RNA being applied to the column. After Round 1, most 200 μ L transcription yielded between 120-200 μ g RNA, of which 60-120 μ g would be correct length product.

Monitoring ligand-eluted RNA allowed us to track the progress of the selection. In most early rounds, less than 1% of the RNA applied to the column would be eluted with the ligand. GlycoBlue co-precipitant used in ethanol precipitation contributes to nucleic acid absorbance measurements on a small volume spectrophotometer and can cause overestimation of RNA amount, especially when the total eluted RNA amount is less than 500 ng. While this problem can be circumvented by ensuring that sufficient RNA (greater than 60 μ g) is applied to the column each round, the lower sensitivity of tracking using spectrophotometry is not problematic for tracking the selection as a whole: once enrichment was measurable and the binders form a large portion of the pool, RNA recovered from the column reached higher than 28%. Even if only 60 μ g of RNA were applied to the column, 28% recovery would correspond to 16.9 μ g RNA, which is easily detectable on a Nanodrop (Figure 2). It is impractical to perform the selection with lower amounts of RNA, as using less RNA may cause the concentration of resuspended ligand-eluted RNA to fall below the limit of detection for the Nanodrop, which is between 2-3 ng/ μ L of nucleic acid for most Nanodrop models.

2.7. Safety.

The immobilization procedure, selection steps, PAGE gel casting, and butanol extraction should be performed in a chemical fume hood with standard personal protective equipment and face and eye protection. Hazardous waste from organic solvent washes and waste from butanol extraction should be collected separately and disposed of in accordance with institutional regulations.

3. Conclusions

Since high affinity aptamers for metabolites are rare in sequence space due to their high informational complexity, affinity chromatography is still the appropriate approach for generating aptamers for metabolic engineering. Furthermore, the Fmoc immobilization chemistry borrowed from solid phase peptide synthesis can be used for functionalizing many metabolites. By transcribing a large mass of RNA of at least 60 μ g during each round of the selection, we were able to accurately quantify both RNA applied to the affinity chromatography column and RNA eluted by the ligand from the column using a small volume spectrophotometer. This approach removes the need for either radioactive labeling or label-free tracking methods such as SPR to monitor the enrichment of binders in our pool. Our modifications on *in vitro* selection methods eliminate barriers associated with conventional methods that make aptamer selections impractical for many research groups. This label-free approach will make *in vitro* selections more widely accessible for research groups looking to generate aptamers for metabolic engineering.

Acknowledgments

We thank D. Sparkman-Yager and C.R. Burke for helpful discussions and comments. C.H. and J.M.C. were supported in part by funds from the University of Washington and an NSF Award MCB 1517052. J.M.C. was a fellow of the Alfred P. Sloan Foundation.

Figures:

Figure 1. Simplified reaction scheme for the immobilization of pAF onto amino PEGA resin using Fmoc/Boc. **A.** In the coupling reaction, DIPEA and PyBOP are used to form a peptide bond between Fmoc-pAF-Boc and the amino PEGA resin. **B.** Following coupling, some unreacted amine groups are left on the PEGA resin, which must be capped to prevent RNA from aggregating on the positively charged amines. Acetic anhydride and DIPEA are used to acetylate unreacted amines. **C.** All unreacted amines on PEGA resin are acetylated, while amines on pAF are protected by Fmoc and Boc. Boc and Fmoc are deprotected using TFA and piperidine, respectively. After deprotection, pAF is immobilized on amino PEGA resin and unreacted amines are capped and the ligand is prepared for affinity chromatography. All steps were performed at room temperature.

Figure 2. Ligand-eluted RNA yields over the course of the pAF selection normalized for applying 60 μg of RNA to the column per round. Each trace represents a separate selection for a pAF aptamer selection. Traces differ by concentration of Mg^{2+} in selection buffer; the effects of Mg^{2+} were discussed in [17]. Although small volume spectrophotometers do not provide much resolution when comparing RNA amounts in early rounds, enrichment for binders in the pool will be easily detectable once binders constitute a substantial portion of the pool as occurs between rounds 9 and 11 for these four parallel selections.

References:

- [1] J.T. Stevens, J.M. Carothers, *ACS Synth. Biol.* 4 (2014) 107–115.
- [2] F. Zhang, J.M. Carothers, J.D. Keasling, *Nat. Biotechnol.* 30 (2012) 354–9.
- [3] J.M. Carothers, J. a Goler, D. Juminaga, J.D. Keasling, *Science* 334 (2011) 1716–9.
- [4] M.N. Win, C.D. Smolke, *Science* 322 (2008) 456–60.
- [5] J.K. Michener, C.D. Smolke, *Metab. Eng.* 14 (2012) 306–316.
- [6] A.D. Ellington, J.W. Szostak, *Nature* 346 (1990) 818–22.
- [7] T.S. Misono, P.K.R. Kumar, *Anal. Biochem.* 342 (2005) 312–317.
- [8] R.K. Mosing, S.D. Mendonsa, M.T. Bowser, *Anal. Chem.* 77 (2005) 6107–6112.
- [9] X. Lou, J. Qian, Y. Xiao, L. Viel, A.E. Gerdon, E.T. Lagally, P. Atzberger, T.M. Tarasow, A.J. Heeger, H.T. Soh, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009).
- [10] J.M. Carothers, S.C. Oestreich, J.H. Davis, J.W. Szostak, *J. Am. Chem. Soc.* 63 (2004) 57–94.
- [11] J.M. Carothers, J.H. Davis, J.J. Chou, J.W. Szostak, (2006) 567–579.
- [12] J.S. Paige, K.Y. Wu, S.R. Jaffrey, *Science* (80-.). 333 (2011).
- [13] W.C. DeLoache, Z.N. Russ, L. Narcross, A.M. Gonzales, V.J.J. Martin, J.E. Dueber, *Nat. Chem. Biol.* 11 (2015) 465–471.
- [14] R.B. Merrifield, *J. Am. Chem. Soc.* 85 (1963) 2149.
- [15] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* 33 (2006).
- [16] J.M. Carothers, *Syst. Synth. Biol.* 7 (2013) 79–86.
- [17] J.M. Carothers, J.A. Goler, Y. Kapoor, L. Lara, J.D. Keasling, *Nucl. Acids Res.* 38 (2010) 2736–2747.