Analytical Biochemistry 390 (2009) 203–205

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Notes & Tips

# Aptamer evolution for array-based diagnostics

Mark Platt [a,b,*,1], William Rowe [a,b,1], David C. Wedge [a,b], Douglas B. Kell [a,b], Joshua Knowles [a,c], Philip J.R. Day [a,d]

[a] Manchester Interdisciplinary Biocenter, University of Manchester, Manchester M1 7DN, UK
[b] School of Chemistry, University of Manchester, Manchester M13 9PL, UK
[c] School of Computer Science, University of Manchester, Manchester M13 9PL, UK
[d] School of Translational Medicine, University of Manchester, Manchester M13 9PT, UK

ARTICLE INFO

*Article history:*
Received 27 January 2009
Available online 15 April 2009

ABSTRACT

Closed loop aptameric directed evolution, (CLADE) is a technique enabling simultaneous discovery, evolution, and optimization of aptamers. It was previously demonstrated using a fluorescent protein, and here we extend its applicability with the generation of surface-bound aptamers for targets containing no natural fluorescence. Starting from a random population, in four generations CLADE produced a new aptamer to thrombin with high specificity and affinity. The best aptameric sequence was void of the set of four guanine repeats typifying thrombin aptamers and, thus, highlights the benefits of evolution performed in an environment closely mimicking the final diagnostic application.



Oligonucleotides known as aptamers are powerful reagents that bind target ligands with affinities comparable to antibodies. Conventionally, aptamers are generated through the process known as SELEX (systematic evolution of ligands by exponential enrichment)[2] [1,2], where strong binding sequences are evolved/enriched from extensive libraries. Since the development of SELEX, aptamers have seen applications ranging from biosensors to therapeutics [3]. However, transfer of this technology to planer surfaces for the purpose of multiplexed assay can be difficult, requiring immobilization and tethering strategies, often followed by "on-chip" optimization to experimental conditions [4].

Evolving aptamers in an environment where they are intended to be used would simplify the process. Surface-based evolutions of aptamers have been demonstrated using inherently fluorescent targets, namely, dye [5] and protein [6]. These targets, however, fail to represent the challenges of raising aptamers to proteins, or other more complex targets, because determining DNA binding events is simpler with fluorescent analytes. Producing aptamers to targets labeled with a fluorochrome (prior to or after hybridization) [7] is problematic because fluorescent dyes can interact strongly via intercalation and other π-stacking interactions [8]. Therefore, it is difficult to ascertain whether the aptamer is binding the analyte or the dye.

The power of SELEX lies within the size of the initial search library (up to $10^{16}$ sequences), vastly exceeding today's highest density arrays. Knight and coworkers produced aptamers on-chip with high affinity and specificity to the fluorescent protein allophycocyanin (APC) [6], with a relatively small initial library size (6000 sequences), using a genetic algorithm (GA) to optimize the binding strength. The ability to systematically generate aptamers through this method is likely to be dependent on the sequence binding profile of the target aptamer. Aptamers that are highly immutable may prove to be difficult to generate through iterative optimization; this is analogous to the "needle in a haystack" class of problems that have previously proven to be difficult for GAs [9].

The current study demonstrates the on-chip evolution for fluorescently tagged protein targets, highlighting aspects that are unique to surface-based processes. Thrombin has been widely studied and used in array experiments [7,10], and was chosen here as a target, due to the known reliance on binding to the motif: GG(A/T)TGGN$_{2-5}$GGT(A/T)GG. If viable aptamers are constrained to this motif, this situation should prove to be challenging to a GA. Using no prior information from existing aptamers, and with an initial library size of only $4.6 \times 10^4$ 30mers, the population was optimized within four evolutionary cycles (each taking 24 h), producing a population of aptamers with nanomolar (nM) affinity.

Bovine serum albumin (BSA), Tween 20, and phosphate-buffered saline (PBS: 0.137 M NaCl, 3 mM KCl, and 10 mM phosphate buffer, pH 7.4) were purchased from Sigma–Aldrich. Biotinylated thrombin was purchased from Novagen (cat. no. 69672-3), and streptavidin–Cy5 was purchased from Invitrogen (cat. no. 43-4316). Water was obtained from a Milli-Q purification system (18 MΩ, Millipore). Aptamers for surface plasmon resonance (SPR) analysis were synthesized with a 3′ biotin–triethylene glycol

* Corresponding author.
*E-mail address:* mark.platt@manchester.ac.uk (M. Platt).

[1] These authors contributed equally to this work.

[2] *Abbreviations used:* SELEX, systematic evolution of ligands by exponential enrichment; APC, allophycocyanin; GA, genetic algorithm; BSA, bovine serum albumin; PBS, phosphate-buffered saline; SPR, surface plasmon resonance; TEG, triethylene glycol.

0003-2697/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2009.04.013

(TEG) modification, thrombin-69-18(29) (5′-AGTCCGTGGTAGGGC AGGTTGGGGTGACT-3′), and ThEV.G4.04422 (5′-GGTTGGTTTATTTT TTACTAGTGGCCAGG-3′) (Eurofins MWG Operon).

DNA arrays were synthesized in-house using the CombiMatrix B3 synthesizer (http://www.combimatrix.com). Briefly, 90 K chips contain 94,928 electrode loci, 25 μm in diameter, with electrodes spaced 20 μm apart. DNA sequences were directly synthesized over the electrodes in a 3′-to-5′ direction. A total of 2150 spots were employed to fulfill quality control and intergeneration normalization functions. All sequences were synthesized in duplicate within each chip, duplicate chips were hybridized, and signal intensities from all four replicate spots were averaged to give the final sequence score. Detailed explanations of data extraction, concentration gradient removal, and chip synthesis can be found elsewhere [6,11] and in the Supplementary material.

Unless stated otherwise, all hybridizations were performed in PBS at 37 °C for 1 h. Prior to hybridization, all chips were incubated with a prehybridization solution (5% BSA and 0.5% Tween 20 in PBS) for 30 min at 37 °C. The prehybridization buffer was removed, and the chip was washed once with PBS followed by incubation in hybridization solution (1 μM thrombin). Two control hybridizations were performed. In the first experiment, the hybridization solution contained both thrombin (1 μM) and BSA (1 μM). In the second, thrombin was removed and the chip was incubated only in PBS. After hybridization chips were washed twice using PBS at room temperature, incubated with streptavidin–Cy5 ($2 \times 10^{-5}$ mg/ml) for 2 min at room temperature, and washed twice with PBS immediately before imaging.

Binding constants were measured using SPR (Biacore 3000). Two DNA sequences, the best sequence from generation 4 and a known thrombin binder for comparison [12], were immobilized onto a Biacore SA chip. Solutions of biotinylated aptamers (66 μM) were flowed over individual channels at a rate of 15 μl/min for 10 min. Thrombin present in PBS at various concentrations (1 μM–10 nM) was passed over the Biacore chip (25 μl $s^{-1}$ at 37 °C), and between samples the chip was reconstituted with glycine buffer (pH 1.5). A blank reference cell value was subtracted from all sensograms, and binding was calculated via the BIA-evaluation software (version 4.1) using the Langmuir model with drifting baseline (see Table S1 in Supplementary material).

Key to developing aptamers on-chip is the ability to differentiate actual protein binding events from spurious nonspecific interactions [13]. Previously, biotin labeling followed by a posthybridization stage with a streptavidin–fluorophore conjugate was shown to have lower nonspecific and background interactions [11].

The starting population, G1, contained 46,389 randomly generated 30mers. After hybridization to the target and extraction of the binding scores (based on fluorescence), the sequences were subject to selection and mutation using a GA. A mu, lambda (μ,λ) algorithm [14] was employed and selected the top 1% of the population (in terms of binding) as the "parents" for the subsequent generation. The point mutation rate was set to a 1:30 probability that each base was mutated (a rate typical to GAs) [15]. In addition, insertion/deletion (indel) mutations were employed at the same rate [6]. A total of 45,158 new sequences were generated alongside 1075 controls uniformly selected (by score) from the previous generation, facilitating intergeneration normalization. In addition, several known thrombin aptamers were synthesized on the chip. The process was repeated until the evolved sequence scores matched those of the known thrombin aptamers.

Fig. 1 shows the distributions of normalized scores from each generation and the sequence logo generated from the top 100 sequences within generation 4. Scoring scales are arbitrary values after analysis using R and the marray package. The biphasic scoring distribution reflects the existence of both specific and nonspecific binding. During evolution, the proportion of specific aptamers increases. However, some nonspecific aptamers continue to be produced as the result of deleterious mutations causing sequences to lose their specific binding. Evidently, a key 5′ feature had developed for the top scoring sequences (i.e., 5′-GGTTGGTTTT ... -3′). Two control experiments were performed on generation 4. The first experiment was to ascertain the specificity of the aptamers to thrombin. Both thrombin and BSA were added in equal concentrations to the hybridization solution, with only the thrombin target having a biotin tag. Analysis of the raw intensities showed that BSA did not significantly decrease binding across the chip, with both scores and rank correlations relating back to the thrombin hybridizations at 0.95 and 0.82, respectively. Raw intensities alongside images taken from the arrays for all control experiments are shown in Fig. S1 of the Supplementary material. To demonstrate that insignificant binding occurred between streptavidin–Cy5 and the aptamers, a second control hybridization was carried out. Thrombin was removed from the hybridization solution while both the imaging and wash procedures remained constant. As can be seen in Fig. S1c, no significant fluorescence was observed.

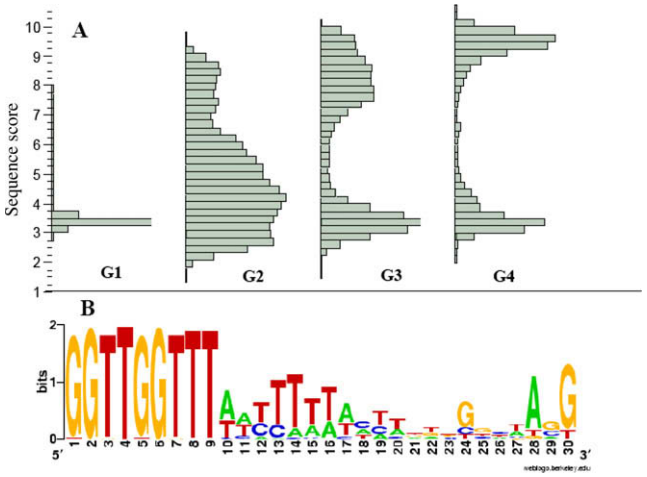


**Fig. 1.** (A) Distribution of scores from each generation. G1 has a population size of 46,655, and G2 to G4 contain 45,158 sequences. Scores are arbitrary numbers from all replicates based on raw intensity values having undergone post-image analysis. (B) Sequence motif taken from the top 100 sequences in generation 4 (http://weblogo.berkeley.edu).

The strength of the binding was ascertained using SPR analysis from the top scoring aptamer in generation 4 (Fig. 2), and the $K_D$ was measured at 24 nM. The evolved sequences appeared to be surprising at first glance given that the typical quadruplex motif (see

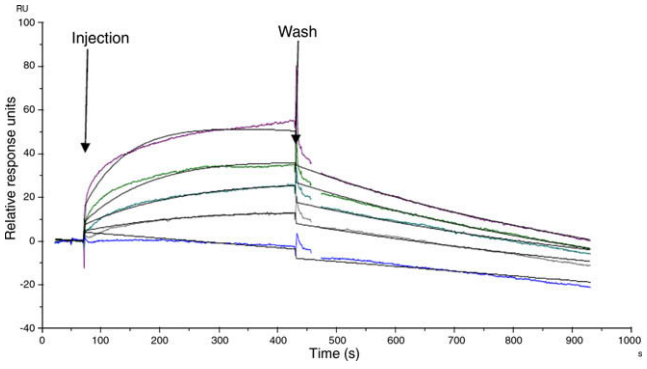


**Fig. 2.** Sensograms from the best aptamer in generation 4 (G4.04422) as determined by SPR. Thrombin was injected onto the aptamer-immobilized SA chip at a range of concentrations between 1 μM and 16 nM (130, 65, 32, 16, and 0 nM shown here) under a buffer flow rate of 25 μl/min (spikes in data from injection times have been removed for data fitting). The on-rate was measured as $8.7 \times 10^4$ and the off-rate as $2.04 \times 10^{-3}$ binding with a $\chi^2$ value of 0.557. The black line represents the best fit.

Fig. S3a in Supplementary material) was not generated [16]. The 5′ end 6mer GGTTGG, however, has previously been shown in solution to bind and inhibit thrombin activity [10]. Given the unique nature of surface-based evolution, one aspect that cannot be discounted is the potential of multiple sequences forming an active aptamer [17,18]. It is possible that two sequences combine to form a quadruplex structure capable of binding to thrombin (Fig. S3b). No mechanism of binding can be given here. However, it is clearly discernible from the SPR data that the new aptamer has a strong affinity for thrombin that rivals existing sequences (Fig. S2). The polyT chain appearing directly before this 6mer may have been generated simply as a linker chain. PolyT chains are widely used as spacer sequences in array experiments, known for their low nonspecific binding and acting to reduce the interaction with the chip's surface.

Microarrays are increasingly being used to study sequence binding relationships. A recent example is that of immunoglobulin E, which highlighted the immutability of the core structure [19]. From this observation, the authors proposed that the sequence space resembled a rugged landscape with sharp peaks resulting from highly restricted base compositions. Similarly, the thrombin aptamer has been shown to possess a constrained sequence motif owing to its complex three-dimensional structure. However, even in this instance the probability of randomly generating the GG(A/T)TGGN$_{2-5}$GGT(A/T)GG consensus sequence at least once within a population of 40,000 30mers is 58% (see Supplementary material). If any flexibility exists around this motif, the probability of generating a sequence that can be optimized via a GA is greatly increased, not discounting multiple aptamers being developed to a single target [19]. No conclusion can be drawn regarding the capability of the GA and its ability to identify a typical core quadruplex motif. Given the relatively small starting population, this study does demonstrate the adaptability of aptamers and how experimental conditions can affect the optimal sequence.

We have demonstrated that aptamers can be developed on-chip using finite population sizes and employing fluorescently labeled targets. The aptamers produced are specific to surface-based analysis and, consequently, may hold many advantages for applied diagnostics. This technique is not limited to the ab initio development of sequences but rather would also allow the optimization of known aptamers to surfaces. Aptamers are developed using a wide range of experimental conditions; by their further evolution on-chip, they can be optimized to a single set of binding conditions, making them suitable for highly multiplexed assays.

## Acknowledgments

The authors thank Sanjay Nilapwar, Christopher Knight, and Farid Khan for their assistance. This work was supported by the Manchester Center for Integrative Systems Biology and was sponsored by the Biotechnology and Biological Sciences Research Council (BBSRC, PBB/D000203/1).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ab.2009.04.013.

## References

[1] A.D. Ellington, J.W. Szostak, In vitro selection of RNA molecules that bind specific ligands, Nature 346 (1990) 818–822.
[2] C. Tuerk, L. Gold, Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase, Science 249 (1990) 505–510.
[3] D.H.J. Bunka, P.G. Stockley, Aptamers come of age—at last, Nat. Rev. Microbiol. 4 (2006) 588–596.
[4] S. Balamurugan, A. Obubuafo, S. Soper, D. Spivak, Surface immobilization methods for aptamer diagnostic applications, Anal. Bioanal. Chem. 390 (2008) 1009–1021.
[5] R. Asai, S.I. Nishimura, T. Aita, K. Takahashi, In vitro selection of DNA aptamers on chips using a method for generating point mutations, Anal. Lett. 37 (2005) 645–656.
[6] C.G. Knight, M. Platt, W. Rowe, D.C. Wedge, F. Khan, P.J. Day, A. McShea, J. Knowles, D.B. Kell, Array-based evolution of DNA aptamers allows modelling of an explicit sequence–fitness landscape, Nucleic Acids Res. 37 (2009) e6.
[7] E.J. Cho, J.R. Collett, A.E. Szafranska, A.D. Ellington, Optimization of aptamer microarray technology for multiple protein targets, Anal. Chim. Acta 564 (2006) 82–90.
[8] B.A. Armitage, Cyanine dye–DNA interactions: intercalation, groove binding, and aggregation, Top. Curr. Chem. 253 (2005) 55–76.
[9] J. Perkel, SNP genotyping: six technologies that keyed a revolution, Nat. Methods 5 (2008) 447–453.
[10] L.C. Bock, L.C. Griffin, J.A. Latham, E.H. Vermaas, J.J. Toole, Selection of single-stranded DNA molecules that bind and inhibit human thrombin, Nature 355 (1992) 564–566.
[11] M. Platt, W. Rowe, J. Knowles, P.J. Day, D.B. Kell, Analysis of aptamer sequence activity relationships, Integr. Biol. 1 (2009) 116–122.
[12] D.M. Tasset, M.F. Kubik, W. Steiner, Oligonucleotide inhibitors of human thrombin that bind distinct epitopes, J. Mol. Biol. 272 (1997) 688–698.
[13] C.L. Warren, N.C.S. Kratochvil, K.E. Hauschild, S. Foister, M.L. Brezinski, P.B. Dervan, G.N. Phillips, A.Z. Ansari, Defining the sequence–recognition profile of DNA-binding molecules, Proc. Natl. Acad. Sci. USA 103 (2006) 867–872.
[14] B.L. Miller, D.E. Goldberg, Genetic algorithms, selection schemes, and the varying effects of noise, Evol. Comput. 4 (1996) 113–131.
[15] H. Muhlenbein, How genetic algorithms really work: mutation and hill-climbing, in: R. Manner, B. Manderick (Eds.), Parallel Problem Solving in Nature, Elsevier, Amsterdam, 1992.
[16] R.F. Macaya, P. Schultze, F.W. Smith, J.A. Roe, J. Feigon, Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution, Proc. Natl. Acad. Sci. USA 90 (1993) 3745–3749.
[17] S.C.B. Gopinath, T.S. Misono, P.K.R. Kumar, Prospects of ligand-induced aptamers, Crit. Rev. Anal. Chem. 38 (2008) 34–47.
[18] T.D. Jeffery, G-quartets 40 years later: from 5-GMP to molecular biology and supramolecular chemistry, Angew. Chem. Intl. Ed. 43 (2004) 668–698.
[19] E. Katilius, C. Flores, N.W. Woodbury, Exploring the sequence space of a DNA aptamer using microarrays, Nucleic Acids Res. 35 (2007) 7626–7635.