





A system for multiplexed selection of aptamers with exquisite specificity without counterselection

Alex M. Yoshikawa^{a,1}, Leighton Wan^{b,1} , Liwei Zheng^c, Michael Eisenstein^{c,d}, and H. Tom Soh^{c,d,e,2} 

Edited by William DeGrado, University of California, San Francisco, CA; received November 1, 2021; accepted January 13, 2022

Aptamers have proven to be valuable tools for the detection of small molecules due to their remarkable ability to specifically discriminate between structurally similar molecules. Most aptamer selection efforts have relied on counterselection to eliminate aptamers that exhibit unwanted cross-reactivity to interferents or structurally similar relatives to the target of interest. However, because the affinity and specificity characteristics of an aptamer library are fundamentally unknowable a priori, it is not possible to determine the optimal counterselection parameters. As a result, counterselection experiments require trial-and-error approaches that are inherently inefficient and may not result in aptamers with the best combination of affinity and specificity. In this work, we describe a high-throughput screening process for generating high-specificity aptamers to multiple targets in parallel while also eliminating the need for counterselection. We employ a platform based on a modified benchtop sequencer to conduct a massively parallel aptamer screening process that enables the selection of highly specific aptamers against multiple structurally similar molecules in a single experiment, without any counterselection. As a demonstration, we have selected aptamers with high affinity and exquisite specificity for three structurally similar kynurenine metabolites that differ by a single hydroxyl group in a single selection experiment. This process can easily be adapted to other small-molecule analytes and should greatly accelerate the development of aptamer reagents that achieve exquisite specificity for their target analytes.

aptamer | affinity reagent | SELEX | counterselection

Aptamers are a class of oligonucleotide-based affinity reagents that have proven highly effective for the detection of small molecules in applications including molecular diagnostics (1, 2), live-cell imaging (3, 4), and real-time monitoring of drugs in live animals (5, 6). This is because the aptamer selection process can be tailored to achieve desired levels of specificity, and there are even examples of aptamers that can distinguish molecules that differ by only a single methyl group (7). To increase the likelihood of generating aptamers with such high specificity, researchers typically incorporate a process known as counterselection into their standard positive selection-based systematic evolution of ligands by exponential enrichment (SELEX) workflow. Here, the aptamer pool isolated after a given round of selection for binding to the target is incubated with one or more “counter targets” that represent common interferents or other structurally similar molecules, and aptamers that bind to the counter targets are removed from the pool.

While conceptually straightforward, the implementation of effective counterselection is extremely challenging because most of the key variables that govern the outcome of the selection cannot be readily measured (8–10). For example, researchers must determine the number of counterselection steps to perform and when to implement them within the overall selection process. Furthermore, the stringency of each individual counterselection step must be optimized in terms of incubation time and counter-target concentration. Since the initial affinity and range of specificities for an aptamer library are fundamentally unknowable a priori, optimal counterselection parameters cannot be identified in advance and can only be determined through time-consuming trial-and-error testing. The consequences of poor counterselection conditions can include selection of low-quality aptamers or the outright failure of the selection (8, 9). For example, excessive counterselection may remove desirable aptamers that offer an ideal balance of affinity and specificity but are present only at low copy numbers in the aptamer pool (11). On the other hand, inadequate counterselection may result in cross-reactive aptamers with poor specificity. Furthermore, the process of characterizing aptamers for specificity requires significant effort due to the well-known challenges of measuring binding interactions between small molecules and aptamers (12). As such, an alternative strategy to counterselection that can facilitate the efficient and reliable generation of high-specificity aptamers is urgently needed.

Significance

Aptamers have the capacity to discriminate between structurally similar small molecules. However, generating such highly specific aptamers has proven challenging using conventional processes based on counterselection against nontarget molecules. In this work, we describe a high-throughput screening platform that can characterize the specificity of millions of aptamers toward a group of structurally related molecules in a single experiment and generate exquisitely specific aptamers without any counterselection. As exemplars, we generated aptamers with high affinity and specificity toward three structurally related kynurenine metabolites using our platform. Our platform can be readily adapted to other small-molecule targets and should therefore accelerate the development of aptamer reagents with exquisite specificity.

Author contributions: A.M.Y., L.W., L.Z., and H.T.S. designed research; A.M.Y., and L.W. performed research; A.M.Y. and L.W. contributed new reagents/analytic tools; A.M.Y., L.W., L.Z., and H.T.S. analyzed data; and A.M.Y., L.W., M.E., and H.T.S. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹A.M.Y. and L.W. contributed equally to this work.

²To whom correspondence may be addressed. Email: tsoh@stanford.edu.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2119945119/-/DCSupplemental>.

Published March 15, 2022.

In this work, we describe a high-throughput screen that can characterize the specificity of millions of aptamers toward a group of structurally related molecules in a single experiment and generate exquisitely specific aptamers without any counterselection process. Our approach builds upon our previously published nonnatural aptamer array (N2A2) system, in which we modified an Illumina MiSeq sequencer such that vast numbers of aptamer clusters can be generated and characterized for binding on a sequencing flow cell (13). Our screening procedure consists of three steps: a multiplexed enrichment step, high-throughput sequencing, and a high-throughput specificity screen. First, a conventional capture-SELEX technique is used to enrich a DNA aptamer library toward a combined pool of structurally similar molecules. Next, the enriched library is sequenced on an Illumina MiSeq, generating millions of aptamer clusters on the surface of the flow cell. Finally, we characterize the binding of the aptamer clusters to each of the target molecules to identify highly specific aptamers. To demonstrate the utility of our platform, we characterized the specificity of an enriched aptamer library toward the tryptophan metabolite kynurenine and four structurally related kynurenine metabolites. Without any counterselection, we were able to identify high-specificity aptamers for three of the metabolites, including multiple aptamers that can differentiate molecules differing by only a single hydroxyl group. The entire selection, sequencing, and screening process is faster than a traditional selection campaign and typically requires ~ 2 wk from the beginning of the experiment to the isolation and characterization of multiple aptamers with high specificity for closely related molecules.

Results and Discussion

The High-Throughput Specificity-Screening Platform. Our workflow consists of three stages: multitarget enrichment, MiSeq sequencing, and a high-throughput specificity screen (Fig. 1 *A–C*). In the first step, the DNA aptamer library (*SI Appendix, Table S1*) is enriched for sequences that bind a group of small-molecule targets using a previously established capture-SELEX protocol (14, 15) (Fig. 1*A*; see *Methods* and *SI Appendix, Fig. S1* for more details). Briefly, a DNA library with a 30-nt variable region is hybridized to complementary 15-nt biotinylated capture strands that are coupled to a column of streptavidin-functionalized agarose resin. The column is then washed with buffer to remove unbound sequences, after which the pooled target solution is added to the column. We then collect the flow-through, which contains structure-switching aptamers that are released from the complementary capture strand upon binding to a small-molecule ligand. As proof of principle, we chose five molecules from the kynurenine pathway (KP)—kynurenine (Kyn), kynurenic acid (KA), 3-hydroxykynurenine (3HK), xanthurenic acid (XA), and 3-hydroxyanthranilic acid (3HA) (Fig. 1*D*)—with 100 μ M of each metabolite in the pooled target solution. The aptamers eluted from the column were then PCR amplified and converted to single-stranded DNA for use in a subsequent round of screening. We performed seven rounds of this enrichment process. During the PCR amplification step, we utilized a fluorescein-labeled forward primer, which enabled us to monitor the relative amount of the aptamer pool that survived each round of capture-SELEX (16).

Selections conducted against pools of structurally similar targets tend to favor the selection of cross-reactive aptamers, and it can be exceedingly difficult to identify rare aptamers that are

truly specific to individual targets. Our previously described N2A2 system gives us a platform to identify these “needle-in-a-haystack” target-specific aptamers, in which we can measure the affinity and specificity of millions of distinct aptamer sequences on the flow cell of a high-throughput sequencer in a single experiment (Fig. 1 *B* and *C*) (13). While the aptamer pool is being sequenced, this also results in aptamer clusters being generated on the MiSeq flow cell. Following the first read of sequencing, the sequence following the reverse primer is removed by exploiting an EcoRI cut site within the reverse-primer region in order to minimize target interactions with Illumina adapter sequences.

The final, specificity-screening stage of our assay employs a strand displacement–based readout mechanism. Every aptamer cluster on the flow cell is hybridized to a Cy3-labeled displacement strand, which uses the same complementary sequence as the capture strand from the selection step. As in the enrichment step, aptamers that bind to a given target molecule undergo a conformational structure switch that results in the displacement of the fluorescently labeled strand, resulting in a reduction in signal. We interrogate the specificity of each aptamer cluster by conducting a series of “buffer cycles” followed by “target cycles.” In the buffer cycles, aptamer clusters are annealed with the labeled displacement strand, washed with buffer, and then imaged on the flow cell. In the target cycles, aptamer clusters are incubated with one of the targets at a 100 μ M concentration for 15 min, washed with buffer, and imaged again. These cycles were repeated for each of the five KP metabolites, with at least one additional replicate cycle for each target. By comparing the relative reductions in fluorescence between the five different KP targets, we can identify highly specific aptamers that do not show appreciable cross-reactivity toward the other metabolites.

Generating a Map of Aptamer–Metabolite Specificity. We analyzed 2.8 million aptamer clusters during our high-throughput specificity screen of the five KP metabolites, 87% of which represented unique sequences. We processed the fluorescence intensity data to remove outliers, extremely low- or high-intensity clusters, and clusters that exhibited high variance between cycles. We measured the intensity data in terms of percent change from each buffer cycle to the subsequent target cycle, in which a higher percent change indicates a greater drop in fluorescence intensity. To normalize for cycle-to-cycle differences, the percent changes for each cluster in response to each target were converted to Z-scores. This was done by subtracting the mean percent change for all clusters in a cycle from each individual percent change measurement; this value was then divided by the SD of the percent change for all clusters in that cycle. Finally, we averaged these Z-scores across all replicates for each sequence–target combination. We set a cutoff Z-score of 2.576—corresponding to a 0.5% chance of false-positive binding—to identify target-binding sequences. These aptamers were further designated as target specific if their “specificity ratio”—that is, the ratio of their target Z-score to the largest off-target Z-score—was greater than 3.

We then created a “map” to compare the overall binding of the aptamer pool to each pair of targets (Fig. 2). We achieved the greatest success with 3HK, with the largest number of monospecific aptamers (902 candidates)—indeed, the majority of the 3HK aptamers were monospecific to 3HK (Fig. 2 *K–O*, gray diamonds). In contrast, we only identified six candidates that were monospecific for XA (Fig. 2 *U–Y*, orange squares) and one that was monospecific for KA (Fig. 2 *F–J*, red inverted

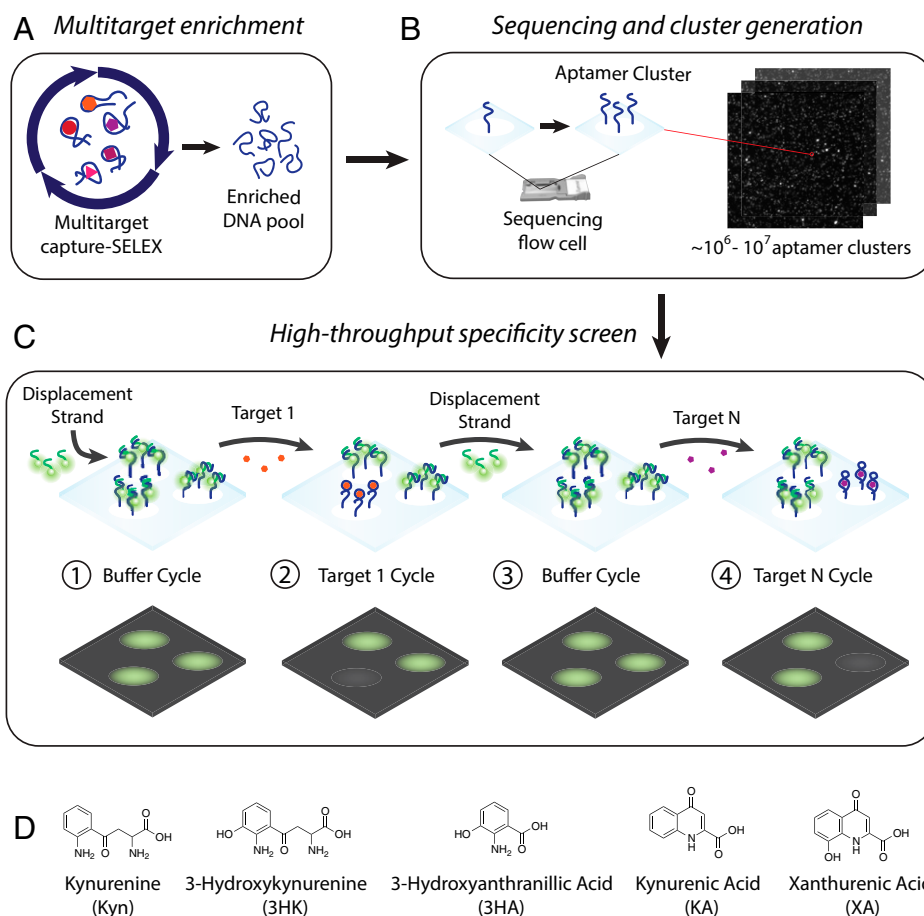


Fig. 1. Overview of the high-throughput specificity-screening platform. (A) Multitarget aptamer enrichment. In the first step of the aptamer-selection pipeline, Capture-SELEX is performed against five pooled kynurenine metabolites (Kyn, 3HK, 3HA, KA, and XA) to enrich the aptamer library for binders prior to high-throughput screening. (B) Sequencing and cluster generation. The enriched aptamer pool is sequenced on a modified Illumina MiSeq sequencer, during which $\sim 10^6$ to 10^7 aptamer clusters are generated on the sequencing flow cell. (C) High-throughput specificity screen. After sequencing is complete, a “buffer cycle” is conducted in which a Cy3-labeled displacement strand is annealed onto the aptamer clusters in buffer and imaged. Next, a “target cycle” is conducted, in which the flow cell is incubated with a single metabolite target for 15 min and then imaged. This process of buffer cycles and target cycles is repeated multiple times for each target. The flow cell is imaged at each step, capturing the fluorescence of each cluster. (D) Structures of the five KP metabolites used in this study.

triangle) and were unable to identify any monospecific aptamers for either 3HA or Kyn. In order to rule out the possibility that the large number of monospecific 3HK sequences were primarily mutants of a small number of aptamers that were present in the initial naïve library by chance, we grouped these sequences into families using a Levenshtein edit distance of ≤ 5 . We identified 784 families among the 902 candidate clusters, indicating that the majority of the 3HK-specific aptamers were not closely related and thus arose independently. These results clearly indicate that 3HK is more amenable to aptamer recognition than the other KP metabolites.

In contrast, only a small proportion of the XA aptamer candidates (6/77) and KA aptamer candidates (1/90) were classified as target specific. Most of the cross-reactive KA and XA aptamers bound both molecules with similar affinity, which was not surprising, since the two metabolites only differ by a single hydroxyl group. We did note that additional monospecific aptamers for these two targets could be identified by using a less stringent specificity ratio cutoff. For example, with a specificity ratio cutoff of 2, we identified 47 more 3HK candidates and two more KA candidates (*SI Appendix*, Fig. S3).

We were not able to identify any monospecific Kyn and 3HA aptamer candidates using our stringent screening thresholds (Kyn, Fig. 2 A–E; 3HA, Fig. 2 P–T), although there were some cross-reactive aptamer candidates that displayed at least some degree of

binding to each target (six sequences for Kyn, 13 sequences for 3HA). This result was unexpected—particularly for Kyn due to its structural similarity to 3HK, for which selection was highly successful. Although we expected many aptamers to be cross-reactive to both targets, as with XA and KA, the majority of 3HK binders did not show any cross-reactivity with Kyn (Fig. 2K). We hypothesized that there might be extremely rare Kyn- or 3HA-binding aptamers present in the multitarget enrichment pool that were not expressed as clusters on the flow cell during sequencing. To address this possibility, we conducted two additional rounds of selection with that pool using only 3HA and Kyn. When we repeated the specificity screen on this pool, we noted that 3HK-, XA-, and KA-specific aptamers were far less abundant, as expected. Even though some highly cross-reactive 3HA and Kyn aptamers were enriched this time around, we were still unable to identify any monospecific aptamers for these targets (*SI Appendix*, Fig. S4). Subtle chemical features may account for the variable success rates of these screens. For example, although 3HA and 3HK share a similar aromatic moiety in their neutral forms, 3HA almost solely exists in its carboxylate form at pH 7.5 (17).

Identification and Characterization of Monospecific KP Metabolite Aptamers. We next identified the most promising monospecific aptamer candidates for 3HK, XA, and KA from the specificity screen, selecting the highest copy-number

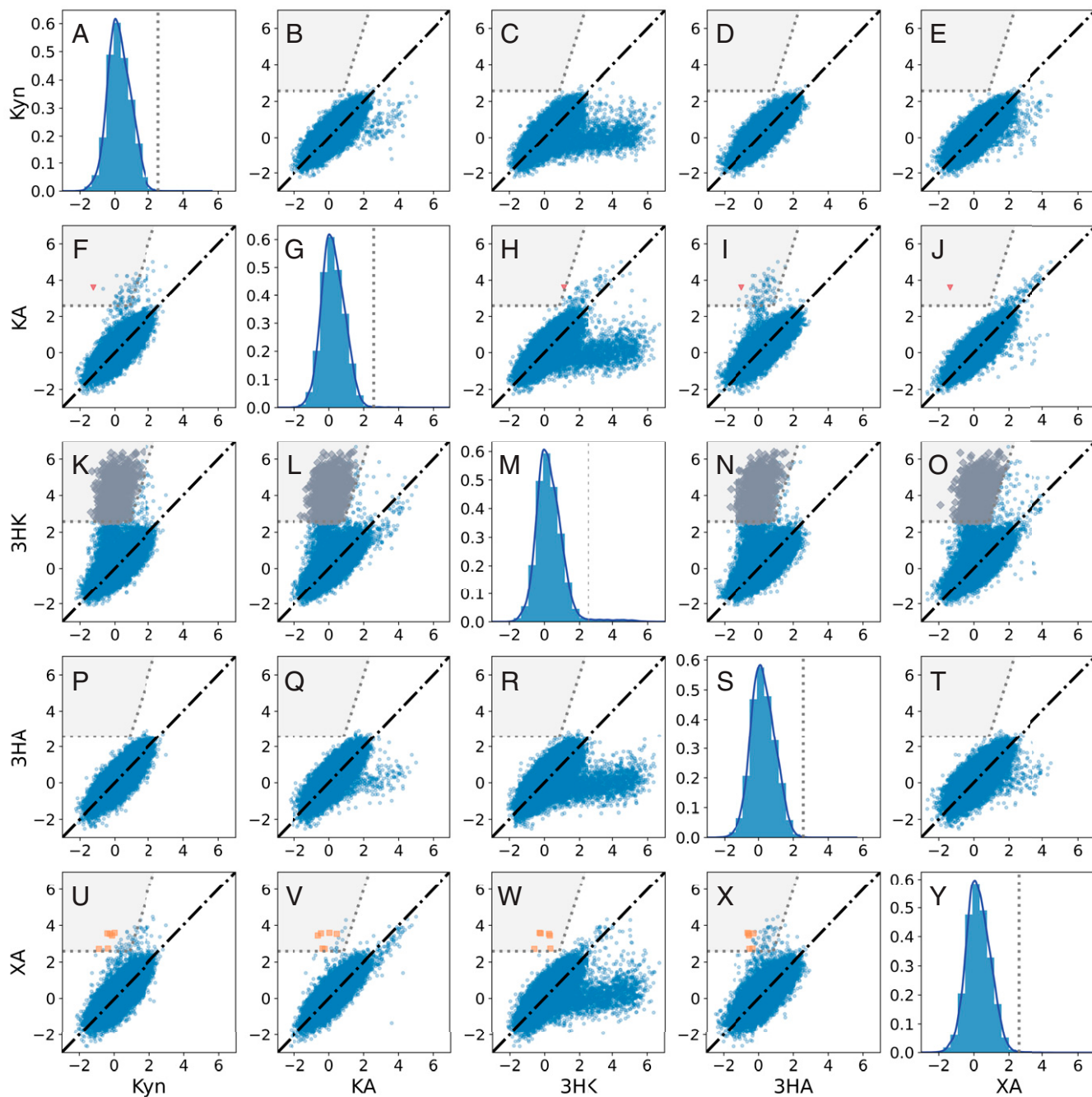


Fig. 2. Specificity map of the aptamer pool toward the five KP metabolites. The metabolites in order from top row to bottom row are Kyn (A–E), KA (F–J), 3HK (K–O), 3HA (P–T), and XA (U–Y). A, G, M, S, and Y show probability density functions and histograms of the Z-scores for Kyn, KA, 3HK, 3HA, and XA, respectively. The y-axis describes the probability density, indicating the likelihood that any given sequence would yield a particular Z-score. B–F, H–L, N–R, and T–X show comparisons of Z-scores between pairs of targets for all sequences with two or more replicates. Each point represents the average Z-score for a single sequence across multiple replicates. Sequences with equivalent scores for both targets fall on the black diagonal line in each plot. Sequences that are likely to bind to the target shown on the y-axis (Z-score ≥ 2.576) and are specific for that same target (specificity ratio ≥ 3) fall into the gray regions on the plots. Putative monospecific sequences were highlighted as red inverted triangles for KA, gray diamonds for 3HK, and orange squares for XA. No monospecific sequences were found for Kyn or 3HA.

aptamer candidate from each monospecific pool. We selected sequences 3HK-1 (Z-score = 4.90 for 3HK versus a maximum of 0.29 for other targets; Fig. 3A), KA-1 (Z-score = 3.60 for KA versus a maximum of 1.11 for other targets; Fig. 3B), and XA-1 (Z-score = 2.72 for XA versus a maximum of 0.77 for other targets; Fig. 3C). Inspection of the flow-cell images for these aptamer clusters during the buffer and target cycles provided further evidence of their specificity and agreed with the extracted intensity values (SI Appendix, Fig. S5). Analysis of the secondary structures of each aptamer revealed a single loop

stabilized by the prestructured eight- to nine-nucleotide hairpin (SI Appendix, Fig. S6). Interestingly, we generally observed that the highest copy-number sequence from the enrichment stage did not exhibit significant binding to any of the five targets (maximum average Z-score = -0.51) (SI Appendix, Fig. S7A). Furthermore, a Pearson correlation coefficient of 0.037 between copy number and Z-score of monospecific 3HK candidates suggests that copy number is not correlated with affinity. These results demonstrate that our screen could overcome the effects of PCR biases, contamination, or other experimental artifacts

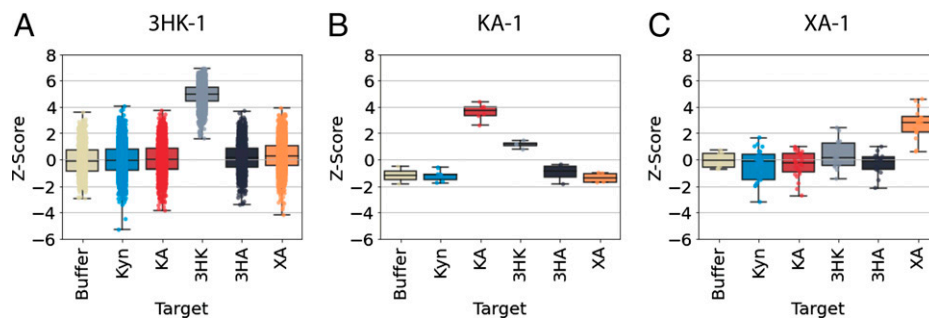


Fig. 3. Results of the high-throughput specificity screen. Z-scores for cluster intensities for buffer and the five KP metabolites for aptamers (A) 3HK-1, (B) KA-1, and (C) 3XA-1. The discrete Z-score values are overlaid for all measurements on a box plot. The middle of the box plot represents the median value, and the top and bottom of the box are the upper and lower quartiles. The whiskers are the minimum and maximum values. 3HK-1, KA-1, and 3XA-1 were respectively represented by 847, 2, and 8 aptamer clusters.

that might lead to the unwanted enrichment of low-quality sequences in other SELEX-based selection strategies.

While our N2A2 assay can assess each aptamer's relative affinity for each target during the specificity screen, it cannot directly measure aptamer equilibrium dissociation constant (K_D). We therefore utilized a previously established plate-reader assay (15), in which 3HK-1, KA-1, and XA-1 were chemically synthesized and labeled at the 5' end with Cy3 and then combined with displacement strands tagged with a DABCYL quencher group at the 3' end. These strands were of various lengths for each aptamer (12 to 14 nucleotides), in which the length was selected to ensure that its K_D for the aptamer was ~ 100 to 500 nM, as previously recommended (see Methods for details) (15). After hybridizing the two strands, each aptamer–displacement strand complex was titrated with varying concentrations of the five different KP metabolites. Target binding results in ejection of the displacement strand, and the fluorescence intensity increases in the absence of the quencher group. The K_D can subsequently be derived from a quantitative equation developed for this competition assay (18). First, we characterized the binding interaction between the Cy3-labeled aptamer and DABCYL-labeled displacement strand (SI Appendix, Fig. S8). Next, we measured the interaction between the aptamer–displacement strand complex and the target of interest. Finally, these measurements were used to compute the K_D of the aptamer for each target.

The plate-reader assay confirmed that the selected aptamers exhibited strong affinity for 3HK (3HK-1 $K_D = 388.4$ nM), KA (KA-1 $K_D = 3.7$ μ M), and XA (XA-1 $K_D = 56.5$ μ M) (Fig. 4). As expected, these aptamers were also remarkably specific, displaying essentially no cross-reactivity to any of the other KP metabolites at the concentrations assayed. Nonspecific

quenching of Cy3 was observed at high concentrations of 3HK, leading to relative fluorescence unit (RFU) signals below baseline (SI Appendix, Fig. S9). We further confirmed aptamer affinity and specificity in a control experiment using scrambled versions of the sequences for 3HK-1, KA-1, and XA-1, in which we observed no meaningful target binding (SI Appendix, Fig. S10). There are no previously published aptamers for these molecules, but by way of comparison, the affinity of 3HK-1 for its target is nearly an order of magnitude higher than that of a previously published tryptophan aptamer [$K_D \sim 2$ μ M (19)], which is relevant given that Kyn is a metabolite of tryptophan. Furthermore, 3HK-1's target affinity is more than two orders of magnitude greater than the majority of previously reported DNA aptamers for amino-acid targets, which typically exhibit K_D s in the mid-micromolar to low-millimolar range (20–22).

Conclusion

In this work, we describe an aptamer generation pipeline that enables the efficient discovery of highly specific aptamers for multiple structurally similar molecules in a single experiment without counterselection. We were able to obtain specific aptamers for three of the five kynurenine metabolites included in the selection, in which the selected aptamers exhibited essentially no off-target binding and can differentiate between molecules that differ by only a single hydroxyl group. The elimination of a counterselection procedure greatly simplifies the selection workflow, bypassing the time-consuming trial-and-error process of optimization that is typically required. As demonstrated by the poor target-binding properties we observed for the most abundant sequence from our enrichment-stage pool, the use of the N2A2 platform for screening can overcome inherent biases associated with multiround SELEX in order to identify the top-performing aptamers in a vast

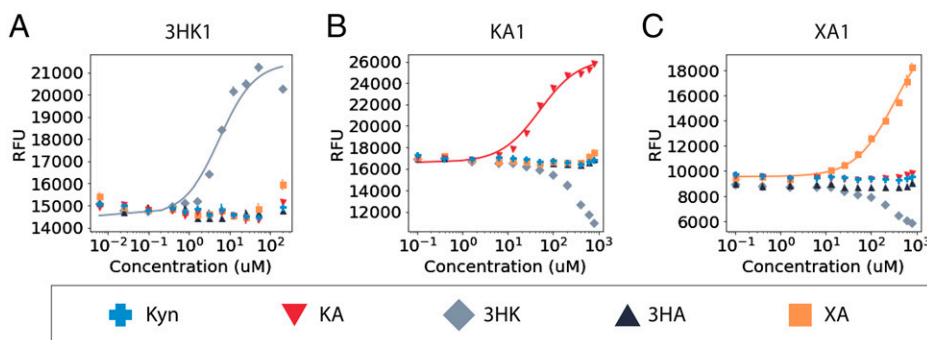


Fig. 4. Measuring the target affinity and specificity of the 3HK-1, KA-1, and XA-1 aptamers via plate-reader assay. Binding assays for aptamers (A) 3HK-1, (B) KA-1, and (C) XA-1 for the five KP metabolites. The points represent the mean of three independent experiments, and the error bars represent the SD.

sequence pool (*SI Appendix, Fig. S7*). Finally, our platform greatly increases the throughput of selection such that we can identify high-specificity candidate aptamers for multiple targets in a single experiment, with the entire process of selection and characterization of individual aptamers requiring just 2 wk from start to finish.

Although this study focused on monospecific aptamers, which are the most desirable for many molecular detection applications, we would like to note that this same approach can also be used to identify cross-reactive aptamers with defined specificities (*SI Appendix, Fig. S11*). Technologies such as “Toggle-SELEX” have been previously utilized to purposely select such cross-reactive aptamers that can bind proteins expressed by different two different species of animal (23). Such species cross-reactivity enables the preclinical evaluation of potentially therapeutic aptamers in animal models. Furthermore, cross-reactive aptamers that can bind promiscuously to entire families of small molecules with a shared chemical scaffold could be useful in many contexts, such as detecting illicit drugs and their metabolites, for which detecting a family of small molecules is more efficient than the use of multiple highly specific assays (24). In our prior work, we have highlighted the use of N2A2 to map sequence determinants that influence target affinity (13), and we envision that data derived from this screening process could likewise uncover sequence and structural elements that inform target specificity. Given the central importance of target specificity in determining the practical utility of an affinity reagent for clinical or research applications, we believe our platform will deliver immediate value as a means for generating superior aptamers in a far more efficient manner than was possible before.

Materials

The DNA library, primers, capture strand, displacement strands, and adaptor sequencing strands were all chemically synthesized by Integrated DNA Technologies (IDT) (*SI Appendix, Table S1*). Fluorophore- or biotin-tagged strands were high-performance liquid chromatography (HPLC) purified, and all other sequences besides the library were purified via polyacrylamide gel electrophoresis. GoTaq DNA polymerase was purchased from Promega (No. M3005). The five kynurenine metabolites (3-hydroxyanthranilic acid, L-kynurenine, XA, KA, and 3-hydroxyl-DL-kynurenine) were all ordered from Sigma-Aldrich (Nos. 148776, K8625, D120804, 67667, and H1771, respectively). Pierce streptavidin agarose was ordered from Thermo Fisher Scientific (No. 20349), and the Micro Bio-Spin chromatography columns used during the selection were ordered from Bio-Rad (No. 7327204). Dynabeads MyOne Streptavidin C1 magnetic beads used for the single-strand generation of the amplified aptamer pool were purchased from Thermo Fisher Scientific (No. 65002). Cy3-labeled aptamer candidates were synthesized by the Stanford Protein and Nucleic Acid (PAN) Facility and cartridge purified (*SI Appendix, Table S1*). The DABCYL quencher-tagged displacement strands were ordered from IDT HPLC purified (*SI Appendix, Table S1*). The plate-reader assays were measured in Corning 96-well half-area black flat-bottom polystyrene microplates (Thermo Fisher Scientific, No. 07-201-205).

Methods

Multitarget Selection. The multitarget-selection protocol was adapted from previously published protocols (15). In each round, we included five times more biotinylated capture strand than aptamer DNA in a final volume of 250 μ L selection buffer (20 mM Tris-HCl, 120 mM NaCl, 5 mM KCl, 1 mM MgCl₂, 1 mM CaCl₂, and 0.01% Tween-20 in nuclease-free water). The first round used 1 nmol DNA library, and subsequent rounds used 100 to 500 picomoles of single-stranded DNA. At the start of each round, the aptamer pool was annealed to the biotin-capture strand and folded by heating to 95 $^{\circ}$ C for 5 min and then incubating at room temperature for at least 30 min. Two separate first-round batches were performed using 1 nmol DNA library each and combined as an input for the second round.

A total of 250 μ L streptavidin resin was added to a column and washed five times with selection buffer, discarding the flow-throughs. This and all other washes were performed with one column equivalent (250 μ L) of solution unless otherwise stated. The annealed library was then added to the column, and the library flow-through was collected and added back to the column three additional times to maximize library capture. The column with captured library was then washed with selection buffer to remove noncaptured strands. Over the seven rounds of selection, the number of washes was increased for stringency from 10 washes in rounds 1 through 4 to 12 washes in rounds 5 and 6 and, finally, 15 washes in round 7. Fifteen washes were used for rounds 8 and 9 with 3HA and Kyn. The five KP metabolites were pooled together at individual concentrations of 100 μ M in selection buffer (final volume 750 μ L) and then added to the column in three 250- μ L aliquots. Target concentrations were consistent across all rounds of selection. Each column flow-through was collected separately and measured for fluorescence on the Qubit Fluorometer (Thermo Fisher Scientific) from round 2 onward. After measurement, the flow-throughs from each round were pooled and concentrated in 3-kDa columns (spin at 14,000 \times g for 15 min) to a final volume of \sim 75 μ L.

The pool was then amplified for subsequent rounds. PCR reagents were added to the concentrated flow-through (20 μ L 100 μ M fluorescein isothiocyanate-conjugated forward primer (FITC FP), 20 μ L 100 μ M Biotin reverse primer (RP), 1 mL GoTaq Master Mix, and water to a final volume of 2 mL). This mix was divided over 20 tubes and subjected to 95 $^{\circ}$ C for 2 min followed by cycles of 95 $^{\circ}$ C for 15 s, 54 $^{\circ}$ C for 15 s, and 72 $^{\circ}$ C for 30 s, followed by 72 $^{\circ}$ C for 1 min and then holding at 4 $^{\circ}$ C. We performed 10 cycles for rounds 1 through 3, 12 for round 4, 10 for round 5, 8 for round 6, 7 for round 7, and 8 for rounds 8 and 9. The amplified material was then collected and cleaned using a MinElute PCR Purification Kit (Qiagen) and quantified on a NanoDrop 2000 (Thermo Fisher Scientific) spectrophotometer and Qubit Fluorometer.

The amplicons were then converted to single-stranded DNA on beads. A total of 300 μ L streptavidin beads were washed three times with 400 μ L 20 mM NaOH (with 15-min incubations for each wash) and then three times with 800 μ L selection buffer. The double-stranded DNA library was then added, with additional selection buffer to a total volume of 500 μ L. The DNA and beads were incubated for 1 h at room temperature and then washed twice with 800 μ L selection buffer and once with 800 μ L water. The supernatant was removed, and then 200 μ L 20 mM NaOH was added and incubated for 8 min at room temperature. A total of 35 μ L of 1 M Tris-HCl and 500 μ L selection buffer were then added to the supernatant, after which the solution was concentrated using 10-kDa size-exclusion columns from Amicon (No. UFC501096), with two buffer exchanges in water (400 μ L). Single-stranded DNA was measured using the NanoDrop and Qubit Fluorometer. This process was repeated for all rounds. The use of a fluorescein-labeled forward primer for PCR enabled us to monitor the convergence of the aptamer pool.

Preparation for High-Throughput Sequencing and Screening. Approximately 125 ng of single-stranded DNA was added to a 300 μ L PCR volume with overhang adaptor RP and FP and amplified using the same protocol as above for the round DNA. The adaptor product was cleaned using an Axygen AxyPrep Mag PCR Clean-up Kit (Thermo Fisher Scientific). The pool then was indexed using the Nextera XT DNA Library Preparation Kit, cleaned with the AxyPrep kit, after which \sim 135 ng DNA was run on a 10% Tris-borate-EDTA (TBE) gel. The desired DNA bands were cut out and sheared, mixed with 400 μ L of Tris-EDTA (TE), and incubated overnight at room temperature. The supernatant was then separated from the gel using 0.2- μ m VWR filters (spun at 14,000 \times g for 3 min). The product was then concentrated using 10-kDa size-exclusion columns, with buffer exchange in TE.

High-Throughput Sequencing and Specificity Screen. The modified Illumina MiSeq sequencer previously developed by our laboratory for the N2A2 process was used to screen \sim 10⁷ aptamer clusters (13). The MiSeq first uses the aptamer pool and bridge amplification steps to generate monoclonal DNA clusters containing \sim 1,000 strands per cluster. In the first read, the MiSeq determines the sequence and location of every cluster on the flow cell. Instead of performing the second read, however, these cycles are used to introduce custom reagents including the complementary strands and fluorescently labeled target molecules. Before measuring cluster response to targets, the DNA strand beyond the reverse-primer sequence is removed using a built-in EcoRI cut site.

The high-throughput screening process measures cluster binding to each individual target over multiple cycles. Alternating buffer and target cycles are used to monitor displacement of the labeled strand upon aptamer-target binding. In buffer cycles, residual bound strands and targets are removed with 750 μL 0.05 M NaOH plus 0.25% sodium dodecyl sulfate (SDS). Next, the flow cell is washed with 500 μL selection buffer before adding 0.2 μM displacement strand in 515 μL selection buffer. The flow cell is then heated to 80 $^{\circ}\text{C}$ and slowly ramped down to 22 $^{\circ}\text{C}$ over ~ 30 min. Finally, the flow cell is washed with 6 mL selection buffer and then imaged. The target cycles begin with a 1.25-mL buffer wash before adding the target of interest. Over a period of 42 min, 605 μL 100 μM target in selection buffer is added, with 500 μL initially followed by seven periodic additions of 15 μL , with 5-min pauses in between. The flow cell is then washed with 6 mL selection buffer and imaged. We performed triplicate buffer/target cycle measurements for Kyn, KA, and 3HK and duplicate buffer/target cycle measurements for 3HA and XA. Fluorescence measurements from each cycle were automatically quantified using the MiSeq software, and the sequence-intensity data were linked using our previously published custom Python code (13) (<https://github.com/sohlab/non-natural-aptamer-array>).

Identification of Aptamer Candidates. Screening data were first cleaned and filtered to find sequences with consistent binding behavior. Clusters with < 80 RFU were removed to account for background signal, while clusters with $> 3,000$ RFU were removed to account for quantification errors or imaging artifacts. We used a constant cutoff based on the coefficient of variance ($c_v = \sigma/\mu$) to remove clusters that did not have consistent RFU across cycles with the same target. The filtered data were then converted to percent changes and Z-scores to improve consistency across cycles. The percent change from buffer cycles ($\%change = (I_{buffer} - I_{target})/I_{buffer}$) was first calculated, and these were shown to be normally distributed (SI Appendix, Fig. S2). These percent changes were normalized to Z-scores ($Z = (\%change_{cluster} - \mu_{cycle})/\sigma_{cycle}$), where μ_{cycle} is the mean percent change for the cycle and σ_{cycle} is the SD of the percent change for the cycle. Normalization using Z-scores was used to compensate for the gradual decrease in percent change with each subsequent cycle due to cluster degradation and other factors. Finally, the Z-scores for each target were used to calculate the average Z-score for each sequence. These were then averaged to derive a Z-score that reflects every replicate for a given sequence-target combination. This average Z-score served as a proxy for aptamer binding to the target, with higher Z-score indicating a higher likelihood of binding the target. To ensure a greater number of measurements for each sequence, only sequences present at two or more clusters were considered.

To identify sequences with relatively high binding for a target, we chose a critical value of $Z = 2.576$. For specificity, we assessed relative binding in terms of the ratio between the Z-score for a given target relative to the maximum Z-score for any of the off-target KP molecules: $Ratio_{z_{aptamer}} = \frac{Z_{target}}{Z_{off-target}}$. A Z-score ratio cutoff

of 3 was chosen to designate target-specific aptamers. Thus, for a target t , an aptamer was considered monospecific if both $z_t > 2.576$ and $\frac{z_t}{z_x} > 3$, where z_x reflects the maximum Z-score for all other targets and buffer. Monospecific aptamer candidates with the highest copy number and Z-score were ordered from the Stanford PAN facility, so that their K_D could be measured.

Characterization of Aptamers via Plate Reader. We used a previously established assay for measuring the affinity of small molecule-binding structure-switching aptamers to characterize each aptamer's K_D for each target (15). Overall, aptamer K_D is determined as the ratio of two K_D values ($K_{D,1}/K_{D,2}$) for the displacement strand ($K_{D,1}$) and for binding of target to the annealed aptamer and displacement strand ($K_{D,2}$). We first tested DABCYL-tagged displacement strands of various lengths to find a length that achieves $\sim 90\%$ quenching of the Cy3-labeled aptamer (50 nM) with K_D in the range ~ 100 to 500 nM. The 50-nM aptamer was annealed with the chosen displacement strand and then incubated with each target over a range of concentrations at a final volume of 100 μL selection buffer at room temperature. Fluorescence spectra for all samples were measured in an opaque black half-well plate at 25 $^{\circ}\text{C}$ on a Synergy H1 microplate reader (BioTek), with filter cube (emission: 590/35, excitation: 538/63, gain: 55). Fluorescence data were fitted using the Hill equation with $n = 1$: $RFU = B_{max}X^n / (K_D^n + X^n) + c$. Curve fitting was performed in Python using the "curve_fit" function from the "scipy" library. The final concentrations we chose based on titrations with each displacement strand (SI Appendix, Fig. S8) were 800 nM 13-mer displacement strand for 3HK-1, 400 nM 13-mer displacement strand for KA-1, and 200 nM 14-mer displacement strand v2 for XA-1. For the control experiments with scrambled aptamer sequences, the final concentrations were 200 nM 14-mer displacement strand v2 for 3HK-1 scramble, 400 nM 14-mer displacement strand v2 for KA-1 scramble, and 200 nM 14-mer displacement strand v2 for XA-1 scramble.

Data Availability. All study data are included in the article and/or supporting information. Data underlying the manuscript figures are provided in Dataset S1.

ACKNOWLEDGMENTS. This work was supported by the Chan-Zuckerberg Biohub, W. L. Gore & Associates, the Helmsley Trust, the Biomedical Advanced Research and Development Agency (grant 75A50119C00051), and the NIH (grants OT2OD025342 and R01GM129314-01). A.M.Y. was supported by the Stanford Bio-X Graduate Fellowship.

Author affiliations: ^aDepartment of Chemical Engineering, Stanford University, Stanford, CA 94305; ^bDepartment of Bioengineering, Stanford University, Stanford, CA 94305; ^cDepartment of Radiology, Stanford University, Stanford, CA 94305; ^dDepartment of Electrical Engineering, Stanford University, Stanford, CA 94305; and ^eChan Zuckerberg Biohub, San Francisco, CA 94158

- D. Ruiz Ciancio *et al.*, Aptamers as diagnostic tools in cancer. *Pharmaceuticals (Basel)* **11**, 1–23 (2018).
- X. Zou, J. Wu, J. Gu, L. Shen, L. Mao, Application of aptamers in virus detection and antiviral therapy. *Front. Microbiol.* **10**, 1462 (2019).
- M. You, J. L. Litke, S. R. Jaffrey, Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E2756–E2765 (2015).
- J. S. Paige, T. Nguyen-Duc, W. Song, S. R. Jaffrey, Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).
- N. Arroyo-Currás *et al.*, Real-time measurement of small molecules directly in awake, ambulatory animals. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 645–650 (2017).
- P. L. Mage *et al.*, Closed-loop control of circulating drug levels in live animals. *Nat. Biomed. Eng.* **1**, 0070 (2017).
- R. D. Jenison, S. C. Gill, A. Pardi, B. Polisky, High-resolution molecular discrimination by RNA. *Science* **263**, 1425–1429 (1994).
- M. McKeague *et al.*, Analysis of in vitro aptamer selection parameters. *J. Mol. Evol.* **81**, 150–161 (2015).
- F. Spill *et al.*, Controlling uncertainty in aptamer selection. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12076–12081 (2016).
- N. Komarova, A. Kuznetsov, Inside the black box: What makes Selex better? *Molecules* **24**, 1–23 (2019).
- Y. J. Seo, M. Nilsen-Hamilton, H. A. Levine, A computational study of alternate SELEX. *Bull. Math. Biol.* **76**, 1455–1521 (2014).
- A. Ruscito, M. C. DeRosa, Small-molecule binding aptamers: Selection strategies, characterization, and applications. *Front. Chem.* **4**, 14 (2016).
- D. Wu *et al.*, Automated platform for high-throughput screening of base-modified aptamers for affinity and specificity. *bioRxiv* [Preprint] (2020). <https://www.biorxiv.org/content/10.1101/2020.04.25.060004v2> (Accessed 20 February 2022).
- R. Stoltenburg, N. Nikolaus, B. Strehlitz, Capture-SELEX: Selection of DNA aptamers for aminoglycoside antibiotics. *J. Anal. Methods Chem.* **2012**, 415697 (2012).
- K. A. Yang, R. Pei, M. N. Stojanovic, In vitro selection and amplification protocols for isolation of aptameric sensors for small molecules. *Methods* **106**, 58–65 (2016).
- R. Stoltenburg, C. Reinemann, B. Strehlitz, FluMag-SELEX as an advantageous method for DNA aptamer selection. *Anal. Bioanal. Chem.* **383**, 83–91 (2005).
- A. Pérez-González, J. R. Alvarez-Idaboy, A. Galano, Dual antioxidant/pro-oxidant behavior of the tryptophan metabolite 3-hydroxyanthranilic acid: A theoretical investigation of reaction mechanisms and kinetics. *New J. Chem.* **41**, 3829–3845 (2017).
- N. Nakatsuka *et al.*, Aptamer-field-effect transistors overcome Debye length limitations for small-molecule sensing. *Science* **362**, 319–324 (2018).
- X. Yang *et al.*, Characterization and application of a DNA aptamer binding to L-tryptophan. *Analyst (Lond.)* **136**, 577–585 (2011).
- E. Vianini, M. Palumbo, B. Gatto, In vitro selection of DNA aptamers that bind L-tyrosinamide. *Bioorg. Med. Chem.* **9**, 2543–2548 (2001).
- K. Harada, A. D. Frankel, Identification of two novel arginine binding DNAs. *EMBO J.* **14**, 5798–5811 (1995).
- K. Ohsawa *et al.*, Arginine-modified DNA aptamers that show enantioselective recognition of the dicarboxylic acid moiety of glutamic acid. *Anal. Sci.* **24**, 167–172 (2008).
- R. White *et al.*, Generation of species cross-reactive aptamers using "toggle" SELEX. *Mol. Ther.* **4**, 567–573 (2001).
- W. Yang *et al.*, In vitro isolation of class-specific oligonucleotide-based small-molecule receptors. *Nucleic Acids Res.* **47**, e71 (2019).