

Bulk Affinity Assays in Aptamer Selection: Challenges, Theory, and Workflow

Eden Teclmichael, An T. H. Le, Svetlana M. Krylova, Tong Ye Wang, and Sergey N. Krylov*



Cite This: *Anal. Chem.* 2022, 94, 15183–15188



Read Online

ACCESS |



Metrics & More

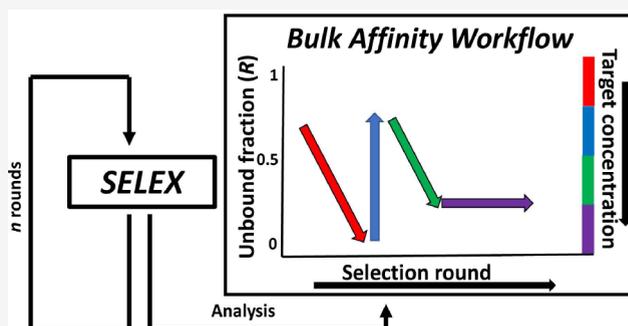


Article Recommendations



Supporting Information

ABSTRACT: Selection of oligonucleotide aptamers involves consecutive rounds of affinity isolation of target-binding oligonucleotides from a random-sequence oligonucleotide library. Every next round produces an aptamer-enriched library with progressively higher fitness for tight binding to the target. The progress of enrichment can only be accurately assessed with bulk affinity assays in which a library is mixed with the target and one of two quantitative parameters, the fraction of the unbound library (R) or the equilibrium dissociation constant (K_d), is determined. These quantitative parameters are used to help researchers make a key decision of either continuing or stopping the selection. Despite the importance of this decision, the suitability of R and K_d for bulk affinity assays has never been studied theoretically, and researchers rely on intuition when choosing between them. Different approaches used for bulk affinity assays expectedly hinder comparative analyses of selections. Our current work has two goals: to give bulk affinity assays a thorough theoretical consideration and to propose a scientifically justified and practical bulk-affinity-assay approach. We postulate a formal criterion of suitability: a quantitative parameter must satisfy the principle of superposition. R satisfies this principle, while K_d does not, suggesting R as a theoretically preferable parameter. Further, we propose a solution for two limitations of R : its dependence on target concentration and narrow dynamic range. Finally, we demonstrate the use of this algorithm in both computer-simulated and experimental aptamer selection. This study sets a cornerstone in the theory of bulk affinity assays, and it provides researchers with a scientifically sound and instructive approach for conducting bulk affinity assays.



Aptamers are single-strand oligonucleotides capable of tightly binding to targets for which they have been selected.^{1,2} Aptamers selected for protein targets can serve as affinity probes and therapeutic agents.^{3–10} Aptamers are selected from random-sequence oligonucleotide libraries using their ability to bind to the target as a driver of selection. A common aptamer-selection process involves repetitive rounds of four major steps. In step 1, a ssDNA library is reacted with the target to form target–DNA complexes (target–binder complexes).^{11,12} In step 2, the complexes are partitioned from the unbound oligonucleotides and collected. In step 3, the collected oligonucleotides are amplified and purified to obtain a large amount of the binder-enriched ssDNA library. Finally, in step 4, the progress of selection is assessed.^{13,14} While being a “service” step, the assessment of the selection progress is of critical importance as its results are used to decide on whether to proceed to the next round or to stop.

Methods used for assessment of the selection progress can be categorized into two types: nonaffinity assays and bulk affinity assays. Nonaffinity assays follow the change in melting temperature during the selection process to assess library diversity.^{15,16} These assays assume that the decreasing diversity unconditionally correlates with increasing fitness of the library

for binding to the target, which is not true. In contrast, bulk affinity assays assess library binding to the target and, thus, are a preferable analytical tool for assessing the progress of aptamer selection.¹⁷ Despite their importance, bulk affinity assays lack any theoretical foundation, and this limitation imposes a problem: theoretical and experimental uncertainties that can lead to misinterpretation of the selection results. This work focuses on bulk affinity assays in aptamer selection.

Bulk affinity assays are procedurally similar to the classic affinity assay used for the determination of the equilibrium dissociation constant K_d of a target–ligand complex (TL) formed in the binding reaction of a target (T) and a ligand (L):



Received: July 21, 2022

Accepted: October 17, 2022

Published: October 27, 2022



Yet, there is a fundamental difference between these two classes of assays. Classic affinity assays analyze a single complex characterized by a single K_d value, which is a thermodynamic constant independent of ligand and target concentrations. In contrast, bulk affinity assays are to assess the affinity of a highly heterogeneous pool of oligonucleotide ligands characterized by a wide scope of K_d values. A single parameter determined for a heterogeneous pool using the formal rules of K_d determination is not a thermodynamic constant; for example, it depends on the ligand and target concentrations. Therefore, we call it here an equilibrium pseudoconstant \bar{K}_d .

The dependence of \bar{K}_d on concentrations questions both the validity and the utility of \bar{K}_d as a measure of bulk affinity. This dependence makes less obvious the advantage of \bar{K}_d over the other parameter, which is intrinsically concentration dependent: fraction R of the unbound (or bound) library. Both \bar{K}_d and R are used as affinity measures in the bulk affinity assays without scientific justification. Accordingly, the goal of this work was to consider a theoretical foundation for the bulk affinity assays. Specifically, we intended to compare theoretical grounds for using R and \bar{K}_d .

Our theoretical analysis suggests R as a preferable measure of bulk affinity. Because R is concentration dependent, we worked out an algorithm for choosing a suitable constant library concentration and for adjusting the target concentration with the progress of selection. The application of this bulk-affinity-assay algorithm was demonstrated experimentally in a de novo selection of DNA aptamers for MutS protein using capillary electrophoresis for partitioning of the target-bound oligonucleotides from free oligonucleotides.

MATERIALS AND METHODS

Chemicals and Materials. All chemicals and buffer components were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. All solutions were prepared in deionized water filtered through a 0.22- μm Milipore filter membrane (Nepean, ON, Canada). Fused-silica capillaries with inner and outer diameters of 75 and 360 μm , respectively, were purchased from Molex Polymicro (Phoenix, AZ). Recombinant His-tagged MutS protein (MW \approx 93 kDa, pI 5.67) was purchased from Prospec Protein Specialist (Ness-Ziona, Israel). All DNA molecules were custom synthesized by Integrated DNA Technologies (Coralville, IA). Polymerase chain reaction (PCR) reagents Q5 High-Fidelity 2 \times Master Mix, Q5 Reaction Buffer, and deoxynucleotide solution mix were purchased from New England BioLabs (Whitby, ON, Canada), and SYBR Green was purchased from Thermo Fisher Scientific (Rockford, IL). The MinElute PCR purification kit was purchased from Qiagen (Toronto, ON, Canada). The CE running buffer was 50 mM Tris-HCl pH 8.0. The sample buffer was always identical to the running buffer to avoid the adverse effects of buffer mismatch. Accordingly, all dilutions of sample components in CE experiments were done by adding the same running buffer.

DNA Sequences. All DNA stock solutions were subjected to annealing by incubation at 90 $^{\circ}\text{C}$ for 2 min before being cooled to 20 $^{\circ}\text{C}$ at a rate of 0.5 $^{\circ}\text{C}/\text{s}$, prior to the dilution and preparation of the equilibrium mixtures. We used a synthetic FAM-labeled DNA library (N40) with a 40-nt random region: 5'-FAM-CTC CTC TGA CTG TAA CCA CG-N40-GCA TAG GTA GTC CAG AAG CC-3'. For quantitative PCR (qPCR), the sequence of forward and reverse primers was as follows: 5'-CTC CTC TGA CTG TAA CCA CG-3' and 5'-

GGC TTC TGG ACT ACC TAG GC-3', respectively. For asymmetric PCR (aPCR), the fluorescently labeled version of the forward primer was used instead, 5'-Alexa Fluor488-CTC CTC TGA CTG TAA CCA CG-3'.

Capillary Electrophoresis (CE) Instrumentation. All CE experiments were performed with a P/ACE MDQ apparatus from SCIEX (Concord, ON, Canada) equipped with a laser-induced fluorescence (LIF) detection system. Fluorescence was excited with a blue line (488 nm) of a solid-state laser and detected at 520 nm using a spectrally optimized emission filter system.¹⁸ The poly(vinyl alcohol) (PVA)-coated capillaries were prepared as described elsewhere.¹⁹ The total lengths of the capillaries were 50 cm for bulk affinity assays and 80 cm for fraction collections; the distances to the detection window were 40 and 70 cm, respectively. Prior to every CE run, the PVA-coated capillaries were rinsed with the running buffer at 20 psi (138 kPa) for 3 min. The coolant temperature was set at 15 $^{\circ}\text{C}$.

Fraction Collection. In the first round of selection, the equilibrium mixture contains 10 μM annealed DNA library and 100 nM His-tagged MutS. For later rounds, the equilibrium mixtures contain 330 nM binder-enriched library and 100 nM His-tagged MutS protein. The mixtures were all incubated for at least 30 min to reach equilibrium. The equilibrium mixture was injected into the capillary by a pressure pulse of 1 psi (6.9 kPa) \times 28 s. The sample plug was propagated by a pressure pulse of 0.9 psi (6.2 kPa) \times 45 s (to yield a 5.4 cm-long buffer plug) to avoid the uncooled region of the capillary to the cooled region. Partitioning was carried out using reversed polarity (anode at the outlet) at 25 kV for 28 min followed by pressure propagation of buffer at 5 psi (34.5 kPa) for 1 min to elute the His-tagged MutS–DNA complex into a collection vial containing 20 μL of the running buffer. A total of five rounds of selection were conducted.

PCR Procedures and Generation of Binder-Enriched DNA Library. The collected binder-enriched library was amplified and quantitated by two rounds of qPCR using CFX Connect instrument (Bio-Rad, ON, Canada). The qPCR reagent mixture was prepared to obtain final concentrations of 1 \times Q5 High-Fidelity 2 \times Master Mix, 1 \times SYBR Green, unlabeled 500 nM forward primer, and unlabeled 500 nM reverse primer. Before thermocycling, the qPCR reaction mixture was prepared by adding a 2 μL aliquot of the collected fraction to 18 μL of the qPCR reagent mixture. The PCR thermocycling protocol is as follows: 98 $^{\circ}\text{C}$ for 30 s (initialization, performed once), 98 $^{\circ}\text{C}$ for 10 s (denaturation), 65 $^{\circ}\text{C}$ for 20 s (annealing), and 72 $^{\circ}\text{C}$ for 20 s (extension), followed by a plate read at 72 $^{\circ}\text{C}$ and a return to the denaturation step for a total of 40 cycles. All qPCR reactions were performed in duplicate. In the first round of qPCR, the collected fraction was quantitated using an eight-point standard curve. An S-shaped amplification curve was plotted, and in the second round of the qPCR, the qPCR product was removed two cycles into the exponential phase of the amplification curve. After qPCR, 100 μL of the qPCR product was later purified using the MinElute PCR purification kit (see note S1).

The purified double-stranded DNA product was then subjected to aPCR strand separation using the following procedure. Initially, the aPCR reagent mixture was prepared to obtain the final concentrations of 1 \times Q5 Reaction Buffer, 1 μM Alexa Fluor488 labeled forward primer, 50 nM label-free reverse primer, and 200 μM deoxynucleotide solution mix.

Before thermocycling, the aPCR reaction mixture was prepared by adding a 5 μL aliquot of the qPCR product to 45 μL of the aPCR reagent mixture. The aPCR thermocycling protocol is as follows: a single step of initiation at 98 $^{\circ}\text{C}$ for 30 s followed by 18 cycles of denaturation at 98 $^{\circ}\text{C}$ for 10 s, annealing at 65 $^{\circ}\text{C}$ for 20 s, extension at 72 $^{\circ}\text{C}$ for 20 s, and fluorescence plate reading at 72 $^{\circ}\text{C}$. The fluorescently labeled single-stranded DNA product of aPCR was purified using the MinElute PCR purification kit (see note S1).

The concentration of the purified DNA product was determined by measuring its fluorescence intensity with the NanoDrop 3300 fluorospectrometer (Thermo Fisher Scientific, Wilmington, DE) at 519 nm and converting the fluorescence intensity into DNA concentration using a standard curve built using serial dilutions of fluorescently labeled forward primer (2000, 1000, 500, 250, 125, 62.5, and 31.25 nM). The purified aPCR product was then used for the next round of partitioning.

Bulk Affinity Assays. Equilibrium mixtures of 1 nM of DNA library and varying target concentrations were prepared and incubated at room temperature for a minimum of 30 min prior to injection. The sample was injected into the capillary inlet by a pressure pulse of 0.5 psi (3.4 kPa) \times 20 s. The sample plug was propagated by a pressure pulse of 0.9 psi (6.2 kPa) \times 45 s (to yield a 5.4 cm-long buffer plug) to avoid the uncooled region of the capillary.²⁰ Separation was carried out at 25 kV with reversed polarity (anode at the outlet) for 15 min.

RESULTS AND DISCUSSION

Definitions of R and \bar{K}_d . Fraction R of the unbound library is defined as

$$R = [L]/[L]_0 \quad (2)$$

where $[L]$ is the concentration of the unbound library at equilibrium and $[L]_0$ is the total concentration of the library.

Equilibrium pseudoconstant \bar{K}_d is a parameter determined using the formal rules utilized for the determination of a true K_d value. Finding K_d and, thus, \bar{K}_d requires a set of R values obtained for single $[L]_0$ but different $[T]_0$ (total concentration of the target). The unknown \bar{K}_d is found by varying it while fitting a theoretical expression for R into an experimental dependence of R on $[T]_0$ (typically termed a binding isotherm):

$$R = -\frac{\bar{K}_d + [T]_0 - [L]_0}{2[L]_0} + \sqrt{\left(\frac{\bar{K}_d + [T]_0 - [L]_0}{2[L]_0}\right)^2 + \frac{\bar{K}_d}{[L]_0}} \quad (3)$$

It is important to re-emphasize that \bar{K}_d is calculated on the basis of the values of R . Hence, R is a primary parameter determined experimentally, while \bar{K}_d is a secondary parameter.

Requirement for Complete Separation of L from TL in Bulk Affinity Assays. The general procedure of a bulk affinity assay starts with preparing the equilibrium mixture of the library and the target. Two signals then are measured from this mixture: one is a cumulative signal from all unbound ligands (S_L), and the other one is a cumulative signal from all of the bound ligands (S_{TL}). This measurement is done via physical or spectral separation of L from TL. The separation is complete if the peaks or spectra do not overlap, and the

separation is incomplete if they do overlap. The choice of a signal processing approach depends on whether or not L and TL are completely separated.

Complete separation of L from TL allows one to express R for given total concentrations of the target, $[T]_0$, and the library, $[L]_0$, through the two signals:²¹

$$R = \frac{[L]}{[L]_0} = \frac{S_L}{S_L + S_{TL}/\varphi} \quad (4)$$

where φ is a coefficient of signal changes when L binds to T, for example, the quantum yield of TL relative to that of L.

If separation of L from TL is incomplete, then only a cumulative signal S from them can be measured. If the signals from L and TL do not interfere (which is true in most detection modes), then the cumulative signal follows the principle of superposition:^{22,23}

$$S = S_L \frac{[L]}{[L]_0} + S_{TL} \frac{[TL]}{[L]_0} \quad (5)$$

In this case, the fraction of unbound ligand can still be determined, but with a formula that includes three signals:

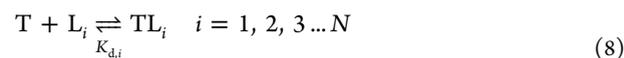
$$R = \frac{S - S_{TL}}{S_L - S_{TL}} \quad (6)$$

Using this formula requires measuring signals from pure L (S_L) and pure TL (S_{TL}) along with the signal from their mixture (S). Measuring S_L is trivial; it is the signal from the ligand in the absence of the target. Measuring S_{TL} requires that $[TL] \gg [L]$, which is achieved (when a single ligand is studied instead of a heterogeneous library) via using a saturating total concentration of the target:

$$[T]_0 > K_d > [L]_0 \quad (7)$$

where K_d characterizes this ligand. The problem is that in a bulk affinity assay the saturating concentration is theoretically unachievable as the library may contain individual ligands with $K_d \rightarrow \infty$ (nonbinders). The above consideration leads to an important practical conclusion: bulk affinity assays require that L and TL be completely separated from each other either spectrally or physically.

Suitability Criterion for Quantitative Measures of Bulk Affinity. It is useful to define a criterion that a quantitative measure of bulk affinity should satisfy to adequately characterize the affinity of the library to the target. We assume that the library is composed of N unique oligonucleotides, and that every oligonucleotide in the library is a ligand capable of binding the target and forming a complex. K_d values of such complexes theoretically range from 0 to ∞ to cover the entire library. The library is thus composed of N unique ligands. In a bulk affinity assay, the library is mixed and incubated with the target to reach equilibrium in a complete set of N binding reactions for N unique ligands:



We postulate that, to be suitable for a bulk affinity assay, a quantitative measure of affinity X must satisfy the principle of superposition:

$$X_{\text{sup}} = X_{L_1} \frac{[L_1]_0}{[L]_0} + X_{L_2} \frac{[L_2]_0}{[L]_0} + \dots + X_{L_N} \frac{[L_N]_0}{[L]_0} \quad (9)$$

In other words, X for the library must be equal to a weighted sum of X values for each individual ligand.

R satisfies eq 9 by its very nature of being a fraction as illustrated in eq 2; in contrast, \bar{K}_d does not satisfy this criterion. Theoretical validation of the superposition principle for R and \bar{K}_d is demonstrated in note S2. Thus, R is a theoretically sound measure of bulk affinity, while \bar{K}_d is not. The theoretical preference of R over \bar{K}_d does not mean that \bar{K}_d is invalidated as an acceptable parameter for bulk affinity assays, but it suggests strongly that R should be chosen over \bar{K}_d unless R has limitations weighing more than the limitations of \bar{K}_d . Thus, while deciding on the practical preference, the advantages and limitations of R and \bar{K}_d in bulk affinity assays should be considered.

Advantages and Limitations of R and \bar{K}_d in Bulk Assays. R is a primary parameter while \bar{K}_d is a secondary parameter depending not only on R but also on $[L]_0$ and $[T]_0$. Hence, R can be determined, in general, more accurately than \bar{K}_d . On the downside, R has two basic limitations as a measure of affinity: it intrinsically depends on both $[L]_0$ and $[T]_0$, and it has a limited dynamic range (from 0 and 1). R can be made independent of $[L]_0$ by choosing an excess of the target, $[T]_0 \gg [L]_0$, but the dependence of R on $[T]_0$ remains. Thus, using R as a measure of bulk affinity for monitoring round-to-round library enrichment only makes sense if there is a sound algorithm for choosing a suitable $[T]_0$. Because it is impossible to choose a priori $[T]_0$, which satisfies bulk affinity assays for all progressively enriched libraries, varying $[T]_0$ will be required when using R as a measure of bulk affinity.

When eq 3 is used to determine a true thermodynamic constant, K_d , the resulting K_d convolutes values of R measured for different $[T]_0$ values, which makes K_d a constant theoretically independent of $[L]_0$ or $[T]_0$. This independence is a great advantage of K_d over R . However, \bar{K}_d is not a true thermodynamic constant and, thus, should depend on $[T]_0$, but, advantageously, this dependence must be much weaker than that of R . Another advantage of \bar{K}_d over R is its wide dynamic range (from 0 to $+\infty$). However, errors of true K_d can be very large when $[L]_0/K_d > 1$.²⁴ This disadvantage must translate into a similar disadvantage for \bar{K}_d .

Thus, using R and \bar{K}_d as measures of affinity in bulk assays have their advantages and limitations. A priori, we can summarize them as follows. Using \bar{K}_d is inferior from the theoretical rigorousness and accuracy standpoints, but provides a formalized way of convoluting data for R measured at different $[T]_0$ values. In contrast, R is a theoretically sound bulk affinity parameter that can be measured accurately, especially around $R = 0.5$. Using R potentially may require fewer experiments as no binding isotherms are needed (unlike calculation of \bar{K}_d with eq 3). Using R , thus, should be preferred over \bar{K}_d , provided that a suitable algorithm of choosing $[T]_0$ is found. Our next goal was, thus, to propose such an algorithm.

Algorithm of Choosing $[T]_0$ in the R -Based Bulk Affinity Assay. We propose the following criteria while designing the algorithm. The accuracy for R measurements is the highest near the point $R = 0.5$; therefore, it is ideal that $[T]_0$ is adjusted to keep R values close to $R = 0.5$. However, a reasonably large range of acceptable R values is required to keep the number of $[T]_0$ -adjustment experiments to the bare minimum.

The first bulk affinity assay is the one with the starting library (before enrichment). Typically, the bulk affinity of such libraries is low, and, therefore, it makes sense to start with the

highest attainable value of target concentration denoted as $[T]_{0,1}$ (1 stands for the first value of $[T]_0$). It is expected that $R > 0.5$ will be typically obtained for the starting library. Progressing enrichment will gradually lower the value of R measured at $[T]_{0,1}$ to $R < 0.2$, where R measurements become unacceptably inaccurate. R values tend to be inaccurate near both extremes: 0 and 1. When the R value is near 1, the fraction of bound library becomes too small to be measured accurately. Similarly, when R is near 0, the fraction of unbound library becomes too small to be measured accurately. At this stage, the target concentration should be decreased to $[T]_{0,2}$ (e.g., $[T]_{0,2} = 0.1[T]_{0,1}$) to increase R to the optimum range of R values designated by us as $0.3 < R < 0.7$. Bulk affinity assays are carried out with $[T]_{0,2}$ for the following rounds of selection until R reaches the level of $R < 0.2$ again, when the target concentration should be further decreased to $[T]_{0,3}$ (e.g., $[T]_{0,3} = 0.1[T]_{0,2}$). This process of a gradual decrease of $[T]_0$ should proceed with progress of selection until no change in R is detected in consecutive rounds of selection with the final R being in a range of $0.3 < R < 0.7$.

To visualize this algorithm, we simulated progressive library enrichment by constructing a virtual starting library of 24 oligonucleotides with the semilog distribution of their K_d values.²⁵ The aptamer-enriched libraries would be progressively shifted to the left and become narrower while keeping the same semilog nature of the K_d distribution (Figure 1).

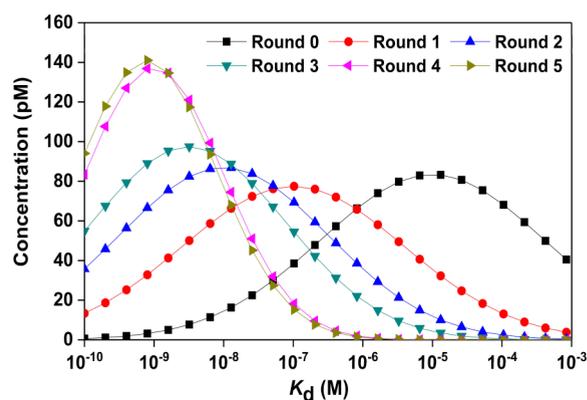


Figure 1. Modeled distributions of K_d values for progressive rounds of aptamer selection. The shape of all distributions was assumed to be Gaussian in semilog coordinates.²⁵

R values for every ligand in the library were calculated with the following equation:

$$R = \frac{K_d}{[T]_0 + K_d} \quad (10)$$

which is obtained from a basic equation for K_d :²⁶

$$K_d = \frac{[T]_0 - [L]_0(1 - R)}{(1/R - 1)} \quad (11)$$

under an assumption that $[L]_0 \ll [T]_0$. A bulk R value for the library was calculated using the superposition principle (eq 9). The details are shown in note S2. We then used our algorithm of choosing/changing $[T]_0$. The details are shown in note S3.

Using the theoretical model, we have simulated the dependency of R on multiple rounds of selection as shown in note S2. A wide range of target concentrations (0, 0.1, 1, 10, 100, 1000, and 10 000 nM) was considered to evaluate the

optimum target concentration needed for bulk affinity assays. Usually, the bulk affinity of the starting library tends to be in the micromolar range; hence, a target concentration of $[T]_{0,1} = 1000 \text{ nM}$ was chosen as a starting concentration with $R > 0.5$ as illustrated in Figure 2. After the second round, the R value

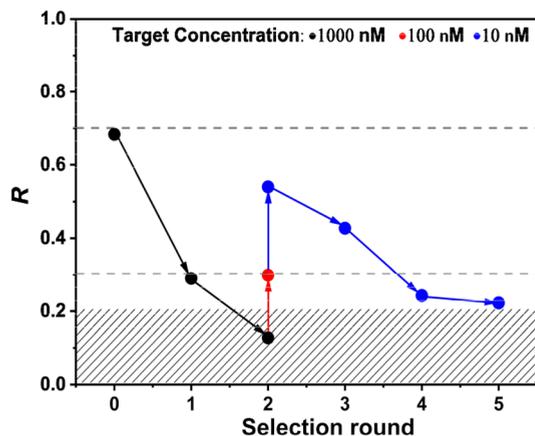


Figure 2. Proposed bulk affinity workflow. The R value in the shaded region indicates the point where the target concentration needs to be decreased to the point where the R value falls between the dashed lines. From left to right, the points represent 1 nM DNA with (i) 1000 nM protein in rounds 0, 1, and 2 (●), (ii) 100 nM protein in round 2 (red ●), and (iii) 10 nM protein in rounds 2, 3, 4, and 5 (blue ●).

for 1000 nM falls below 0.2 and is deemed to be no longer accurate. Hence, the target concentration was reduced gradually by 10-fold to $[T]_{0,2} = 100 \text{ nM}$ and later to $[T]_{0,3} = 10 \text{ nM}$ until the R value fell into the range of $0.3 < R < 0.7$. The enrichment of the libraries for the rest of the selection rounds after the second round was evaluated using $[T]_{0,3} = 10 \text{ nM}$ until saturation was reached.

Experimental Demonstration of the Proposed Bulk-Affinity-Assay Algorithm. We used the proposed algorithm of the bulk affinity assay to guide the selection of aptamers for MutS protein from a random-sequence DNA library by CE-based partitioning. The target was His-tagged MutS for which aptamers have not been previously selected. This His-tagged MutS was found to excessively adsorb to the fused-silica inner capillary wall (unlike the tagless MutS, which is no longer commercially available). We found that using PVA-coated capillaries can largely reduce the protein adsorption to capillary walls.²⁷ Because PVA coating suppresses the electroosmotic flow, we applied the “complex-last” NECEEM mode for the aptamer selection.²⁸

It is beneficial for affinity assays to use the lowest library concentration at which the signal-to-noise ratio (S/N) is still sufficiently high to ensure accurate R determination. To satisfy this condition, we chose $[L]_0 = 1 \text{ nM}$ for all of our bulk affinity assays. Five rounds (rounds 1–5) of aptamer selection were performed, and the random-sequence oligonucleotide library was considered as the product from round 0. For the library obtained from each round of selection, bulk affinity assays were conducted in accordance with the theoretical proposed workflow. Specifically, for rounds 0, 1, and 2, the binding experiments were performed using 1000 nM protein. In round 2, the R value fell below 0.3; hence, the protein concentration was decreased in a stepwise fashion to 100 and 10 nM subsequently to reach the desired range of R values ($0.3 < R < 0.7$). All experiments were performed in triplicate. The results

of the R values are summarized in Figure 3 (see note S4 for the detailed data analysis procedure). According to Figure 3, the

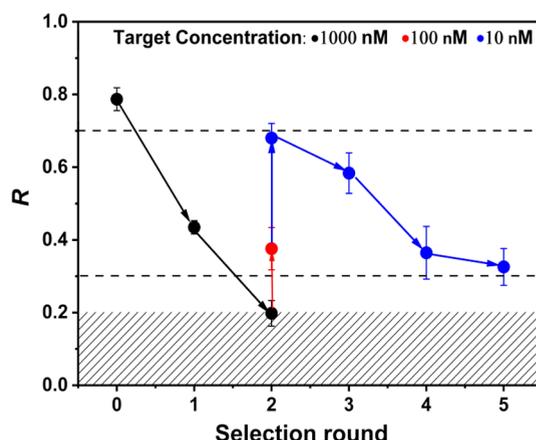


Figure 3. Experimental His-tagged MutS protein bulk affinity assay based on the proposed bulk affinity assay workflow. The R value in the shaded region indicates the point where the target concentration needs to be decreased to the point where the R value falls between the dashed lines. From left to right, the points represent 1 nM DNA with (i) 1000 nM protein in rounds 0, 1, and 2 (●), (ii) 100 nM protein in round 2 (red ●), and (iii) 10 nM protein in rounds 2, 3, 4, and 5 (blue ●).

saturation of selection was reached at rounds 4 and 5, because the R values obtained in these two rounds were greater than 0.2 and consistent within the uncertainties.

CONCLUDING REMARKS

Monitoring the progress of library enrichment is key to effective aptamer selection. Bulk affinity assays are the only analytical tool that can provide direct information about the fitness of the library for binding the target. Therefore, bulk affinity assays are preferred over nonaffinity assays. We demonstrated that a fraction R of the unbound (bound) library is theoretically preferred over an equilibrium pseudoconstant \bar{K}_d as a measure of bulk affinity. Yet, R has three limitations: it depends on the target concentration, its dynamic range is narrow (0–1), and its accuracy is poor when it is close to 0 or 1. To compensate these limitations, we propose an algorithm of target concentration change that keeps R within a range of 0.3–0.7. We demonstrated the use of this algorithm in a simulated aptamer selection as well as in experimental aptamer selection. Our approach allows one to avoid screening a wide range of target concentrations for every round and to avoid very large errors associated with using a single target concentration for bulk affinity assays in all rounds of selection. We suggest this approach as conventional. Having a single approach used by different aptamer-selection teams would allow comparative analysis of selection progress.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.2c03173>.

PCR purification procedure (note S1); theoretical validation of the superposition principle for R and \bar{K}_d (note S2); theoretical bulk affinity model (note S3); and data analysis for the bulk affinity assays (note S4) (PDF)

AUTHOR INFORMATION

Corresponding Author

Sergey N. Krylov – Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada; orcid.org/0000-0003-3270-2130; Email: skrylov@yorku.ca

Authors

Eden Teclemichael – Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada; orcid.org/0000-0001-7779-2884

An T. H. Le – Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada; orcid.org/0000-0002-3659-9938

Svetlana M. Krylova – Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada; orcid.org/0000-0002-3291-6721

Tong Ye Wang – Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada; orcid.org/0000-0001-9462-7194

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.2c03173>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the NSERC Canada (grant STPG-P 521331-2018).

REFERENCES

- (1) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
- (2) Tuerk, C.; Gold, L. *Science* **1990**, *249*, 505–510.
- (3) Quang, N. N.; Miodek, A.; Cibieli, A.; Duconge, F. *Methods Mol. Biol.* **2017**, *1575*, 253–272.
- (4) Navani, N. K.; Mok, W. K.; Yingfu, L. *Methods Mol. Biol.* **2009**, *504*, 399–415.
- (5) Keefe, A. D.; Pai, S.; Ellington, A. *Nat. Rev. Drug Discovery* **2010**, *9*, 537–550.
- (6) German, I.; Buchanan, D. D.; Kennedy, R. T. *Anal. Chem.* **1998**, *70*, 4540–4545.
- (7) Zhang, H. Q.; Wang, Z. W.; Li, X. F.; Le, X. C. *Angew. Chem., Int. Ed.* **2006**, *45*, 1576–1580.
- (8) Yi, M.; Yang, S.; Peng, Z.; Liu, C.; Li, J.; Zhong, W.; Yang, R.; Tan, W. *Anal. Chem.* **2014**, *86*, 3548–3554.
- (9) Zhu, C.; Li, L.; Wang, Z.; Irfan, M.; Qu, F. *Biosens. Bioelectron.* **2020**, *160*, 112213.
- (10) Gerasimova, Y. V.; Nedorezova, D. D.; Kolpashchikov, D. M. *Methods* **2022**, *197*, 82–88.
- (11) Berezovski, M.; Musheev, M.; Drabovich, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2006**, *128*, 1410–1411.
- (12) Berezovski, M. V.; Musheev, M. U.; Drabovich, A. P.; Jitkova, J. V.; Krylov, S. N. *Nat. Protoc.* **2006**, *1*, 1359–1369.
- (13) McKeague, M.; Bradley, C. R.; Girolamo, A. D.; Visconti, A.; Miller, J. D.; DeRosa, M. C. *Int. J. Mol. Sci.* **2010**, *11*, 4864–4881.
- (14) Valenzano, S.; De Girolamo, A.; DeRosa, M. C.; McKeague, M.; Schena, R.; Catucci, L.; Pascale, M. *ACS Comb. Sci.* **2016**, *18*, 302–313.
- (15) Mohammadinezhad, J. S. A. H.; Farahmand, H. *Anal. Meth.* **2020**, *12*, 3823–3835.
- (16) Luo, Z.; He, L.; Wang, J.; Fang, X.; Zhang, L. *Analyst* **2017**, *142*, 3136–3139.
- (17) Mencin, N.; Šmuc, T.; Vraničar, M.; Mavri, J.; Hren, M.; Galeša, K.; Krkoč, P.; Ulrich, H.; Šolar, B. *J. Pharm. Biomed.* **2014**, *91*, 151–159.
- (18) Galievsky, V. A.; Stasheuski, A. S.; Krylov, S. N. *Anal. Chem.* **2017**, *89*, 11122–11128.
- (19) De Jong, S.; Krylov, S. N. *Anal. Chem.* **2012**, *84*, 453–458.
- (20) Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. *Anal. Chem.* **2010**, *82*, 8692–8695.
- (21) Sisavath, N.; Rukundo, J.; Le Blanc, J. C. Y.; Galievsky, V. A.; Bao, J.; Kochmann, S.; Stasheuski, A. S.; Krylov, S. N. *Angew. Chem., Int. Ed.* **2019**, *58*, 6635–6639.
- (22) Rukundo, J.; Le Blanc, J. C. Y.; Kochmann, S.; Krylov, S. N. *Anal. Chem.* **2020**, *92*, 11973–11980.
- (23) Rukundo, J.; Kochmann, S.; Wang, T. Y.; Ivanov, N. A.; Le Blanc, J. C. Y.; Gorin, B. I.; Krylov, S. N. *Anal. Chem.* **2021**, *93*, 11654–11659.
- (24) Kanoatov, M.; Mehrabanfar, S.; Krylov, S. N. *Anal. Chem.* **2016**, *88*, 9300–9308.
- (25) Spill, F.; Weinstein, Z. B.; Shemirani, A. I.; Ho, N.; Desai, D.; Zaman, M. H. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 12076–12081.
- (26) Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. *Anal. Chem.* **2015**, *87*, 3099–3106.
- (27) Liyanage, R.; Krylova, S. M.; Krylov, S. N. *J. Chromatogr. A* **2013**, *1322*, 90–96.
- (28) Le, A. T. H.; Wang, T. Y.; Krylova, S. M.; Beloborodov, S. S.; Krylov, S. N. *Anal. Chem.* **2022**, *94*, 2578–2588.