Selection of Smart Aptamers by Methods of Kinetic Capillary Electrophoresis

Andrei P. Drabovich, Maxim Berezovski, Victor Okhonin, and Sergey N. Krylov*

Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada

We coin the term "smart aptamers"−aptamers with predefined binding parameters (k_{on}, k_{off}, K_d) of aptamer− target interaction. Aptamers, in general, are oligonucle- otides, which are capable of binding target molecules with high affinity and selectivity. They are considered as potential therapeutic targets and also thought to rival antibodies in immunoassay-like analyses. Aptamers are selected from combinatorial libraries of oligonucleotides by affinity methods. Until now, technological limitations have precluded the development of smart aptamers. Here, we report on two kinetic capillary electrophoresis tech- niques applicable to the selection of smart aptamers. Equilibrium capillary electrophoresis of equilibrium mix- tures was used to develop aptamers with predefined equilibrium dissociation constants (K_d), while nonequi- librium capillary electrophoresis of equilibrium mixtures facilitated selection of aptamers with different dissociation rate constants (k_{off}). Selections were made for MutS protein, for which aptamers have never been previously developed. Both theoretical and practical aspects of smart aptamer development are presented, and the advantages of this new type of affinity probes are described.

Aptamers are single-stranded DNA or RNA oligonucleotides, which are selected from large combinatorial libraries based on their ability to bind target molecules. Aptamers can bind proteins, small-molecule compounds, and even living cells with high affinity and specificity.[1,2] They are thought to be very promising affinity ligands with a potential to revolutionize the market of affinity probes and replace classical antibodies as diagnostic, analytical, and therapeutic reagents.[3−5]

Aptamers were first described by Szostak and Gold's groups in 1990.[6,7] The conventional approach to select aptamers from combinatorial oligonucleotide libraries is named SELEX (system- atic evolution of ligands by exponential enrichment) and includes multiple rounds of alternating affinity separation and polymerase chain reaction (PCR) amplification to yield a pool of aptamers with high affinity (Figure 1). We have recently introduced a new approach for the selection of aptamers, which excludes amplifica-
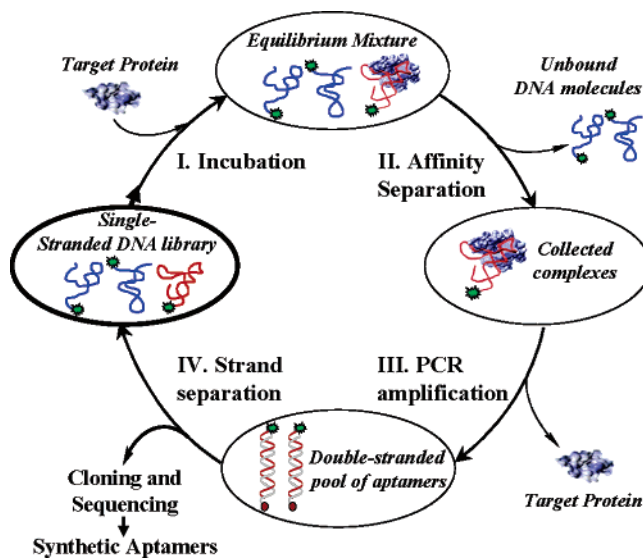

Figure 1. Schematic representation of aptamer selection procedure.

tion steps between rounds of affinity selection; the approach was named Non-SELEX.[8]

Primary attributes of aptamers, namely, outstanding affinity and binding specificity, allowed aptamers to find their niche in a number of applications including the following: development of bioanalytical assays,[9] inhibition of enzymes and receptors,[10,11] target validation,[12] drug screening,[13−15] imaging of cellular or- ganelles,[16] and development of biosensors.[17] Aptamers have indisputable advantages over antibodies in terms of simplicity and cost of production as well as robustness in applications. Now, aptamers are increasingly accepted as therapeutic reagents for the treatment of multiple pathologies.[18] Biomedical applications

* To whom correspondence should be addressed. E-mail: skrylov@yorku.ca.
(1) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999,** *68,* 611−647.
(2) Famulok, M.; Mayer, G.; Blind, M. *Acc. Chem. Res.* **2000,** *33,* 591−599.
(3) Rimmele, M. *Chembiochem.* **2003,** *4,* 963−971.
(4) Jayasena, S. D. *Clin. Chem.* **1999,** *45,* 1628−1650.
(5) Ellington, A. D.; Conrad, R. *Biotechnol. Annu. Rev.* **1995,** *1,* 185−214.
(6) Tuerk, C.; Gold, L. *Science* **1990,** *249,* 505−510.
(7) Ellington, A. D.; Szostak, J. W. *Nature* **1990,** *346,* 818−822.

(8) Berezovski, M.; Musheev, M.; Drabovich, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2006,** *128,* 1410−1411.
(9) Bock, C.; Coleman, M.; Collins, B.; Davis, J.; Foulds, G.; Gold, L.; Greef, C.; Heil, J.; Heilig, J. S.; Hicke, B.; Hurst, M. N.; Husar, G. M.; Miller, D.; Ostroff, R.; Petach, H.; Schneider, D.; Vant-Hull, B.; Waugh, S.; Weiss, A.; Wilcox, S. K.; Zichi, D. *Proteomics* **2004,** *4,* 609−618.
(10) Pestourie, C.; Tavitian, B.; Duconge, F. *Biochimie* **2005,** *87,* 921−930.
(11) Davidson, E. A.; Ellington, A. D. *Trends Biotechnol.* **2005,** *23,* 109−112.
(12) Blank, M.; Blind, M. *Curr. Opin. Chem. Biol.* **2005,** *9,* 336−342.
(13) Mayer, G.; Jenne, A. *BioDrug* **2004,** *18,* 351−359.
(14) Nutiu, R.; Yu, J. M.; Li, Y. *Chembiochem* **2004,** *5,* 1139−1144.
(15) Famulok, M.; Mayer, G. *Chembiochem* **2005,** *6,* 19−26.
(16) Ulrich, H.; Martins, A. H.; Pesquero, J. B. *Cytometry A* **2004,** *59,* 220−231.
(17) Tombelli, S.; Minunni, M.; Mascini, M. *Biosens. Bioelectron.* **2005,** *20,* 2424− 2434.
(18) Fine, S. L.; Martin, D. F.; Kirkpatrick, P. *Nat. Rev. Drug Discovery* **2005,** *4,* 187−188.

also include gene therapy and drug delivery to specific sites.[19−21] Thus, innovation in the development and applications of aptamers is essential for progress in analytical and biomedical sciences.

Here, we present the concept of smart aptamers, which are defined as aptamers with predefined parameters of interaction with a target molecule.

Retrospectively, the idea of exploiting a set of ligands with different affinities for the same target (protein) was initially raised in the proteomic research and development of protein miroarrays.[22] As concentrations of proteins in real samples vary significantly, affinity ligands need to be modified so that lower-affinity ligands are used for highly expressed proteins and higher-affinity ligands are used for proteins with low expression levels.[22] There are no easy ways to modify the affinity of antibodies. Besides, the use of polyclonal antibodies may fail due to their heterogeneity and wide range of physicochemical properties; the optimization of analytical conditions for all proteins assayed is a very challenging task. Smart aptamers with predefined binding parameters are a very attractive alternative to antibodies.[23] It was estimated that in vivo generation of antibodies develops the affinity of not lower than 10^{-10} M, while the rate constant of complex formation is limited to the order of 10^6 M^{-1} s^{-1}, and the first-order rate constant of complex dissociation is usually higher than 10^{-4} s^{-1}.[24,25] Aptamers can potentially go beyond these limits by at least 1 order of magnitude.[26]

A number of applications may require smart aptamers and smart ligands in general. An accurate quantitative analysis of a target in a range of concentrations from 1 pM to 1 mM requires a panel of probes with a similar range of affinities (K_d).[22] Diverse analytical and biomedical applications may also require ligands with different rate constants of complex dissociation (k_{off}). Therapeutic agents that act over different periods of time are one of the possible applications of such smart ligands. Ligands with fast k_{on} and fast k_{off} could be used as pharmaceuticals for acute disorders, while ligands with slow k_{on} and slow k_{off} are preferable for treating chronic diseases.

Smart aptamers promise to contribute significantly to the improvement of purification techniques. Affinity purification requires ligands that first bind a substance, but then release it easily under the elution conditions. In this case, aptamers with moderate affinities are preferable; otherwise tightly bound ligands cannot be eluted from the column under mild nondenaturing conditions.

Competitive aptamer-based drug screening assays rely on the displacement of a fluorescently labeled aptamer moiety by a small-molecule binder from the active site of a target. Such a displacement can be readily monitored and, thus, serves as an indirect indicator of the binding of a small molecule to the target.[27] To facilitate the displacement and reach the maximum signal, the affinity of an aptamer to the target should be similar to that of a small-molecule binder.

Until recently, aptamer selection relied mostly on heterogeneous separation techniques such as affinity chromatography and filter-binding assays. Heterogeneous separation methods suffer from a number of inherent drawbacks that limit their efficiency in aptamer selection. Capillary electrophoresis (CE), in contrast, is a homogeneous technique, which can facilitate efficient separation of chemicals and biopolymers in free solution.[28] In our recent works, we devoted a considerable effort to developing CE-based methods for selection of aptamers.

Capillary electrophoresis with laser-induced fluorescence is a powerful analytical technique with low mass and concentration detection limits. Homogeneous free-solution separation has also made CE the technique of choice for studies of biomolecular interactions.[29] The known examples include protein−DNA, protein−peptide, and protein−drug interactions.[30,31] Our laboratory has introduced a generic concept for studies of biomolecular interactions by CE; it is termed kinetic capillary electrophoresis (KCE).[32] Being a homogeneous kinetic approach, KCE allows accurate determination of equilibrium and kinetic parameters of biomolecular interactions.[32,33] The basic knowledge of kinetics sheds light on the processes that occur during separation. In a simple case, a target and a ligand interact with one-to-one stoichiometry. The interaction includes the bimolecular reaction of complex formation and a monomolecular reaction of complex dissociation. The assumption of simple stoichiometry allows us to accurately describe the laws of processes and unambiguously predict the migration of species with different binding parameters along the capillary.

The recent advances in library screening by KCE have permitted the rapid generation of fully functional pools of aptamers within a few days.[34,35] The new Non-SELEX selection of aptamers with KCE reduced this time to a few hours.[8] Uniquely, KCE facilitates selection of smart ligands−ligands with predefined binding parameters.[23,36] In a recently published work on nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM)-based selection of DNA aptamers, we hypothesized on selecting smart aptamers with predetermined k_{off} values.[34] We later reported, in a short communication format, on the selection of smart aptamers with predefined K_d values.[23]

Here, we present the application of two KCE methods, namely, equilibrium capillary electrophoresis of equilibrium mixtures (ECEEM) and NECEEM, to the selection of smart aptamers from DNA libraries. We also describe the theoretical considerations of smart aptamer selection with ECEEM and NECEEM. To experimentally prove the concept, we selected smart aptamers with

(19) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555−583.
(20) Patil, S. D.; Rhodes, D. G.; Burgess, D. J. *AAPS J.* **2005**, *7*, E61−77.
(21) Proske, D.; Blank, M.; Buhmann, R.; Resch, A. *Appl. Microbiol. Biotechnol.* **2005**, *69*, 367−374.
(22) Barry, R.; Soloviev, M. *Proteomics* **2004**, *4*, 3717−3726.
(23) Drabovich, A.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 11224−11225.
(24) Maynard, J.; Georgiou, G. *Annu. Rev. Biomed. Eng.* **2000**, *2*, 339−376.
(25) Foote, J.; Eisen, H. N. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1254−1256.
(26) Nimjee, S. M.; Rusconi, C. P.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555−583.
(27) Green, L. S.; Bell, C.; Janjic, N. *Biotechniques* **2001**, *30*, 1094−1100.
(28) Hu, S.; Dovichi, N. J. *Anal. Chem.* **2002**, *74*, 2833−2850.
(29) Heegaard, N. H. *Electrophoresis* **2003**, *24*, 3879−3891.
(30) Gayton-Ely, M.; Pappas, T. J.; Holland, L. A. *Anal. Bioanal. Chem.* **2005**, *382*, 570−580.
(31) Yang, P.; Whelan, R. J.; Jameson, E. E.; Kurzer, J. H.; Argetsinger, L. S.; Carter-Su, C.; Kabir, A.; Malik, A.; Kennedy, R. T. *Anal. Chem.* **2005**, *77*, 2482−2489.
(32) Petrov, A.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 17104−17110.
(33) Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124*, 13674−13675.
(34) Berezovski, M.; Drabovich, A.; Svetlana, M.; Krylova, M. M.; Okhonin, V.; Petrov, A.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 3165−3171.
(35) Mendonsa, S. D.; Bowser, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 20−21.
(36) Krylov, S. N. *J. Biomol. Screening* Published on Web 01/17.

predefined K_d and k_{off} for MutS protein. We believe that the presented KCE methods will greatly enhance and complement the traditional technologies available for the selection of aptamers and screening small-molecule libraries for drug candidates.

EXPERIMENTAL SECTION

Chemicals and Materials. Thermostable DNA mismatch binding protein (MutS) from *Thermus aquaticus* was purchased from InterSciences (Markham, ON). Fused-silica capillaries were purchased from Polymicro (Phoenix, AZ). Buffer components were obtained from Sigma-Aldrich (Oakville, ON, Canada). All solutions were made using the Milli-Q-quality deionized water and filtered through a 0.22-μm filter (Millipore, Nepean, ON, Canada). Nonlabeled primers, a biotin-labeled primer, a 6-carboxyfluorescein-labeled primer, and a synthetic random DNA library were obtained from IDT (Coralville, IA). The polyacrylamide gel-purified DNA library contained a central randomized sequence of 39 nucleotides flanked by 19- and 22-nt primer hybridization sites (5′-CTTCTGCCCGCCTCCTTCC-(39N)-GGAGACGAGATAGGC GGACACT-3′). Recombinant *Taq* DNA polymerase and all other chemicals were from Sigma-Aldrich unless otherwise stated.

ECEEM Separation and Fraction Collection. All experiments were conducted with an MDQ-PACE instrument (Beckman-Coulter) with either LIF or diode-array detector modules. The inner diameter and length of the capillary were 75 μm and 80 cm, respectively. The bare fused-silica capillary was used. The electrophoresis and incubation buffer was 50 mM Tris-HCl at pH 8.2. All MutS−DNA equilibrium mixtures were prepared in the incubation buffer using the following two-step procedure. First, 2.5 μL of the solution of fluorescently labeled DNA in the incubation buffer was denatured by heating at 95 °C for 3 min with subsequent cooling to 20 °C at a rate of 7.5 °C/min. Second, 2.5 μL of MutS solution in the incubation buffer was mixed with the DNA sample and incubated at 20 °C for 30 min. The first round of selection was implemented with an unlabeled 50 μM DNA library (final concentration); however, the DNA library was PCR-labeled with a fluorescein-labeled primer in the subsequent rounds. To maintain equilibrium conditions in ECEEM, the electrophoresis run buffer was supplemented with 100 nM MutS protein, and the capillary was prefilled with the run buffer before injection of the equilibrium mixture. A plug of equilibrium mixture was injected into the capillary and subjected to electrophoretic separation at an electric field of 365 V/cm with a positive electrode being at the injection end of the capillary. Three fractions were collected in different time windows. Laser-induced fluorescence detection was used to record all the electropherograms in this work, except for Figure 2, where a diode-array detector was used to find the migration time of free MutS and nonlabeled DNA library.

NECEEM Selections. The only difference between NECEEM and ECEEM was the absence of MutS protein in the run buffer for NECEEM. Two fractions were collected in different time windows.

PCR Amplification. DNA in the collected fractions was PCR-amplified in a thermocycler (MasterCycler 5332, Eppendorf, Germany). In addition to the collected DNA ligands, the PCR mixtures contained 50 mM KCl, 10 mM Tris-HCl (pH 8.6), 2.5 mM MgCl$_2$, all four dNTPs at 200 mM each, primers (300 nM each), and 0.05 unit/μL *Taq* DNA polymerase. An HPLC-purified
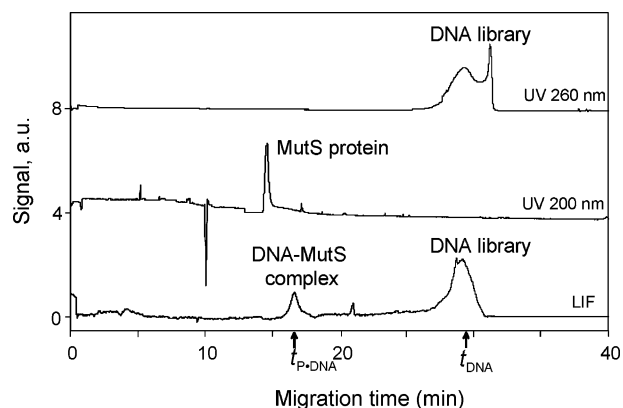

Figure 2. Determination of migration times of the naïve DNA library (top curve), MutS protein (middle curve), and nonspecific MutS-DNA complex (bottom curve).

forward primer (5′-6-carboxyfluorescein-CTTCTGCCC GCCTC-CTTCC-3′) and a reverse primer (5′-biotin-AGTGTCCG CCTA-TCTCGTCTCC-3′) were used. The total volume of the PCR reaction mixture was 50 μL. From 12 to 30 thermal cycles were conducted with every cycle consisting of melting at 94 °C for 10 s, annealing at 56 °C for 10 s, and extension at 72 °C for 10 s. PCR mixtures were analyzed by CE to find the optimal number of cycles and exclude the effect of overamplification of libraries.[37,38]

Strand Separation. The 6-carboxyfluorescein-labeled ssDNA was separated from the complementary biotinylated ssDNA strand on streptavidin-coated super paramagnetic iron oxide particles (Sigma-Aldrich) according to the supplier's instructions.

Cloning and Sequencing. Three pools of DNA ligands obtained from the third round of ECEEM selection were PCR-amplified using unlabeled primers and cloned into Nova Blue Singles Competent cells (*Escherichia coli*) using the pT7 Blue-3 Perfectly Blunt Cloning Kit (Novagen, Madison, WI). Colonies were grown on three separate agar plates overnight, and then 90 colonies (30 from each plate) were picked randomly and reseeded on fresh plates. All selected colonies were screened for the ligand sequence insert into the plasmid by real-time PCR, and 86 colonies were found to have such an insert. Then, 33 individual sequences were investigated by NECEEM for the affinity to MutS. Twenty clones were chosen for sequencing of DNA. A plasmid from each colony was prepared using a GenElute Plasmid Miniprep Kit and sequenced at the Core Molecular Biology Facility at York University.

Measurements of K_d and k_{off}. A MutS-DNA equilibrium mixture was prepared as described above. A plug of the equilibrium mixture was injected into the capillary and subjected to NECEEM at an electric field of at 500 V/cm. The equilibrium constant (K_d) and rate constant (k_{off}) of protein−DNA complex dissociation were found from a NECEEM electropherogram (see Supporting Information).

RESULTS AND DISCUSSION

Differences between NECEEM and ECEEM. First, we describe the difference between the ECEEM and NECEEM methods. Even though the methods are applicable to selection of

(37) Drabovich, A.; Krylov, S. N. *J. Chromatogr., A* **2004**, *1051*, 171−175.
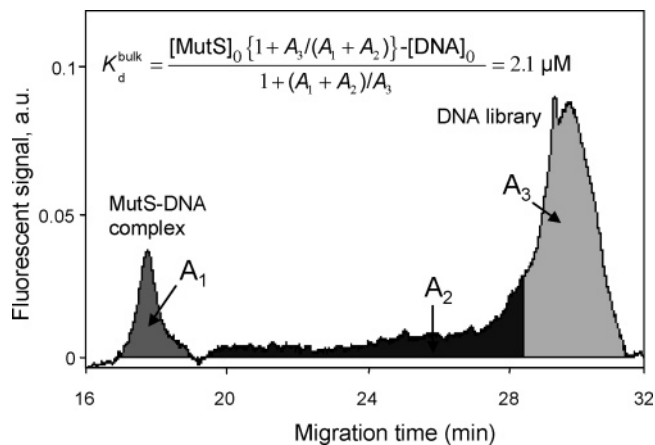(38) Musheev, M. U.; Krylov, S. N. *Anal. Chim. Acta*. In press.

Figure 3. Measurement of "bulk" affinity for nonspecific interactions of the naïve DNA library with MutS protein.

both DNA and RNA aptamers, here we consider only DNA aptamers. While different conceptually, ECEEM and NECEEM have very similar experimental designs. In both methods, a DNA library is mixed with a target protein in a selection buffer and incubated to form the equilibrium mixture. Then, a plug of the equilibrium mixture is injected into the capillary, a high voltage is applied, and the separation of species in the equilibrium mixture starts. The difference between the methods is in the composition of the CE run buffer. In ECEEM, the target protein is present in the run buffer at the concentration equal to that in the equilibrium mixture, while in NECEEM, the run buffer is devoid of the target protein. As a result, the dynamic quasi-equilibrium is maintained in ECEEM, but disturbed during separation in NECEEM, which leads to conceptual differences between distributions of species along the capillary and observed concentration profiles. In both ECEEM and NECEEM, different DNA sequences in the library have similar electrophoretic mobilities and migrate as a single electrophoretic zone. Protein−aptamer complexes, however, have mobilities different from free DNA; they typically move faster. Eluted fractions are collected in specific time windows to facilitate selection of aptamers with desirable values of K_d and k_{off}. In subsequent steps, selected fractions of DNA are amplified in PCR, strands are separated to yield single-stranded DNA, and the enriched library is subjected to the next selection round of partitioning and amplification. Preliminary steps in the selections by ECEEM and NECEEM include the determination of initial selection parameters, which are the same in both methods.

Measuring Initial Selection Parameters. In the case of DNA aptamer selection, initial parameters imply separation parameters (migration times of the DNA library, t_{DNA}, a target protein, t_P, a protein−DNA complex, $t_{P \cdot DNA}$) and "bulk affinity" of the protein to the naïve DNA library. Migration times of species allow the calculation of favorable selection windows (see eqs 2 and 3 below). The bulk affinity, in general, can be described as the average affinity of the target to the entire library. It shows the affinity before the selection and serves as a reference for assessing the progress of selection in the consecutive rounds. All initial selection parameters can be found with NECEEM (Figures 2 and 3).[33] To implement this, a fluorescently labeled random DNA library is mixed with a target protein at a concentration $10^2−10^3$ times higher than the one used for the selection and incubated to form the equilibrium mixture. The equilibrium mixture is then sub-
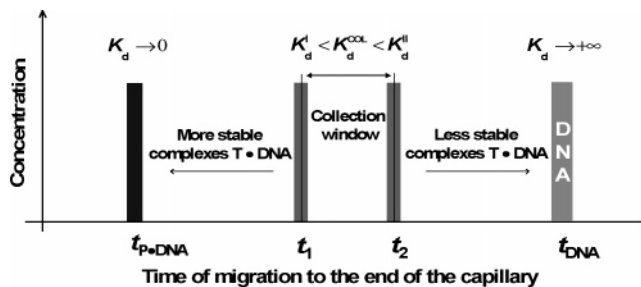
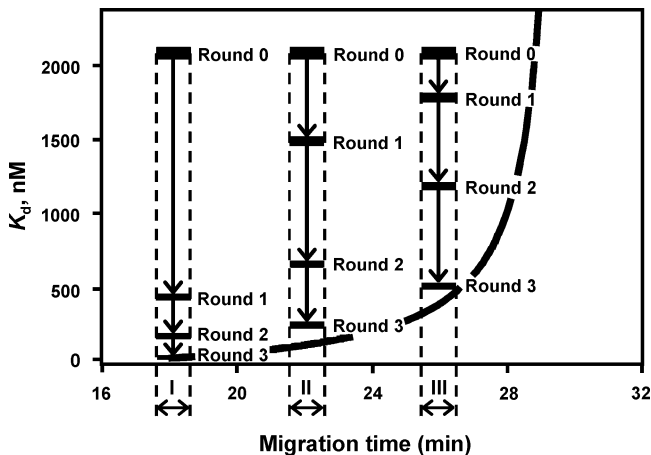**Figure 4.** Concept of ECEEM-based selection of smart aptamers with predefined K_d.

Figure 5. Theoretical distribution of K_d in the capillary (solid curve) and experimental K_d (horizontal bars) obtained in three rounds of ECEEM-based selection.

jected to NECEEM. At the end of NECEEM separation, three areas are present in the electropherogram. Area A_1 corresponds to the intact protein−DNA complexes. Area A_2 corresponds to free DNA, which is formed from the dissociation of the complex during the separation. Area A_3 corresponds to the fraction of free DNA in the initial equilibrium mixture. Ratios between these areas allow us to determine K_d and k_{off} constants (see Supporting Information). Accurate measurement of the bulk affinity by NECEEM is limited to the certain range of K_d values, which is determined by the LOD (the lowest K_d) and the solubility of the target protein (the highest K_d).

Selection of Smart Aptamers with Predefined K_d. The equilibrium dissociation constant K_d is the most important and universal parameter that describes the interaction and determines the advantages of a ligand in the analysis. Here, we describe in detail the selection of K_d-predetermined smart aptamers by ECEEM.

Conceptually, the dynamic quasi-equilibrium in this method means that aptamers spend some time within the complex and some time as free molecules. The effective migration time t of the ligand in ECEEM depends on K_d and the concentration of free protein $[P]_{free}$:

$$\frac{1}{t} = \frac{1}{t_{DNA}} \frac{K_d}{[P]_{free} + K_d} + \frac{1}{t_{P \cdot DNA}} \frac{[P]_{free}}{[P]_{free} + K_d} \quad (1)$$

where $t_{P \cdot DNA}$ and t_{DNA} are migration times of the target−DNA complex and free DNA, respectively. The effective migration time

Table 1. Unique Clones Obtained with ECEEM Selection

	Clone ID	K_d, nM	k_{off}, 10^{-3} s^{-1}	k_{on}, $M^{-1} s^{-1}$*	Sequence (5'→3')
Region I Theoretical K_d = 0 - 11 nM Experimental bulk K_d = 33 nM	2-05	29	2.5	9×10^4	F-CTGGAGTGACGGGCTTATTTGAATCGGGCGAACTGGACT-R
	2-06	15	0.4	3×10^4	F-TGGTAAAGTCATTAATAGGTGTGGGGTGCCGGGCATTTC-R
	2-07	15	0.7	6×10^4	F-GACTGTTGGTTCTTGGGAGGAGTGCGGACTTTGCCACGC-R
	2-11	87	1.0	1×10^4	F-GACGGTGATTGTTATCTCAGAGCCAGTTTGAACAACCCG-R
	2-15	30	0.8	3×10^4	F-GTGGGGATGCCATTGGTCCACAGGGTTATGGGAGCATCA-R
	2-16	29	0.8	3×10^4	N/a
	2-18	170	8.2	5×10^4	N/a
	2-20	26	1.0	4×10^4	N/a
	2-26	98	4.6	5×10^4	F-GTAAGTTATTTTCCTCGCGTTGTATGATGGAATTGACTC-R
	2-27	27	2.7	1×10^5	F-GGAATGATGCGTAAAGGACCGAGGGTGCTTGCCATAATC-R
Region II Theoretical K_d = 73 - 105 nM Experimental bulk K_d = 340 nM	4-05	3.6	0.5	9×10^4	F-GTCTTATGTCGTTAGTCGCAGGGTGATGAGTGAGGCAAG-R
	4-08	780	3.3	4×10^3	F-CTAGTGCAGGGGTTCACTCAGGACCTTACGAGCTTTTTC-R
	4-13	17	1.5	9×10^4	F-CGGGTCATATGGGAACGCTACAGTGTTTCACGCTGAGTT-R
	4-14	160	1.1	7×10^3	N/a
	4-18	97	3.1	3×10^4	N/a
	4-26	67	3.0	4×10^4	F-AGCCGTCACAGGTCGCTCATTCATTACGGTCTGTTTGCG-R F-AACGGGTACTGCCTATCAGACATGGAAACGGTGCCGTTG-R
	4-28	22	1.0	5×10^4	N/a
	4-33	550	5.8	1×10^4	F-TCGGGGTTAGAACGTCGTGTAGGACTCCTATCGGTTTAT-R
Region III Theoretical K_d = 284 - 410 nM Experimental bulk K_d = 550 nM	10-04	520	3.0	6×10^3	F-AGAGTAGCCAAGCGGAGAGTATTTTTGACTGCTTCCCCC-R
	10-11	190	3.8	2×10^4	F-GTTTTCCTGCGGCAGCGTTTTGTTGCTTTCGTTCCTCCG-R
	10-12	820	3.0	4×10^3	F-CGCACTTGCAAACCTTCATGGTCTCGTGCTTGAGTGGTT-R
	10-17	41	5.8	1×10^5	F-CTAGTCACTATTGTCTGGCGTGGGTAGGTGCCAATGGTA-R
	10-28	860	5.7	7×10^3	N/a
	10-31	160	2.7	2×10^4	F-TTCTTGAATTTTGAGATGAGGTCACTAGGTCCTTCGCGG-R
	10-33	240	6.7	3×10^4	F-GAGCATTCGTCAGATAGGTGGCAAACAGTTATGGGAACG-R

* calculated as k_{off}/K_d

Constant primer sequences:
F = CTTCTGCCCGCCTCCTTCC
R = GGAGACGAGATAGGCGGACACT

of the ligand can change between $t_{P \cdot DNA}$ and t_{DNA} depending on K_d and $[P]_{free}$. $[P]_{free}$ is assumed to be equal to the overall concentration of the protein $[P]$ because of the constant supply of the protein from the run buffer. As a result, the interaction of the library with a constant flow of the protein distributes DNA molecules along the capillary according to their K_d values:

$$K_d(t) = [P] \frac{t_{DNA}}{t_{P \cdot DNA}} \frac{t - t_{P \cdot DNA}}{t_{DNA} - t} \qquad (2)$$

Ideally, aptamers with the same K_d values should migrate as a single peak with the infinitely small width (Figure 4). Thus, eq 2 allows us to calculate a theoretical range of K_d values in a fraction collected from time t_1 to t_2. Experimentally, peak broadening due to diffusion, slow equilibrium, or nonspecific interactions with other components in the system may introduce a bias of nonbinding DNA into the collected fraction. That is why the selection procedure may require a few rounds of selection to approach the theoretically predicted K_d values.

The unique feature of ECEEM for the selection of ligands arises from its simple mathematical description. The range of the affinity distribution along the capillary is determined by $[P]$, $t_{P \cdot DNA}$, and t_{DNA}. By changing these parameters, it is feasible to screen the DNA library for the ligands with affinities ranging from picomolar to millimolar values of K_d, which covers 9 orders of magnitude.

In our experiment, we chose three collection windows shown in Figure 5 (compare with Figure 3) and completed three rounds of selection. Each round consisted of ECEEM separation, fraction collection, PCR amplification, separation of strands, and measuring the bulk affinity of the enriched library with NECEEM. After the third round of selection, three pools of aptamers were cloned into bacteria. Selected bacterial clones were screened for the aptamer insert into the plasmids, and 33 individual aptamers were obtained. As the next step, K_d and k_{off} constants of individual sequences were measured with NECEEM, and representative clones were sequenced (Table 1). The precisions (RSD) of K_d and k_{off} measurements with NECEEM were 20% and 5%, respectively (see Supporting Information). The LOD in NECEEM with instrumentation used here was 5×10^7 molecules or 500 pM. Theoretically, the range of K_d values, which can be measured with NECEEM, spans from 1 pM to 1 mM and depends on the concentration LOD from the lower side and on the concentration of protein available from the upper side.[36] The rate constant of complex dissociation, k_{off}, was proven to be reliably measurable in NECEEM if it was within the range of $10^{-4}-1$ s^{-1}.[31,36] ECEEM can also be used to calculate K_d of aptamers to MutS (eq 1). However, it requires a set of experiments to be performed and cannot determine k_{off}. It also requires knowledge of the migration time $t_{P \cdot DNA}$, which can only be directly found with NECEEM. This shows the importance of both methods for successful selection.
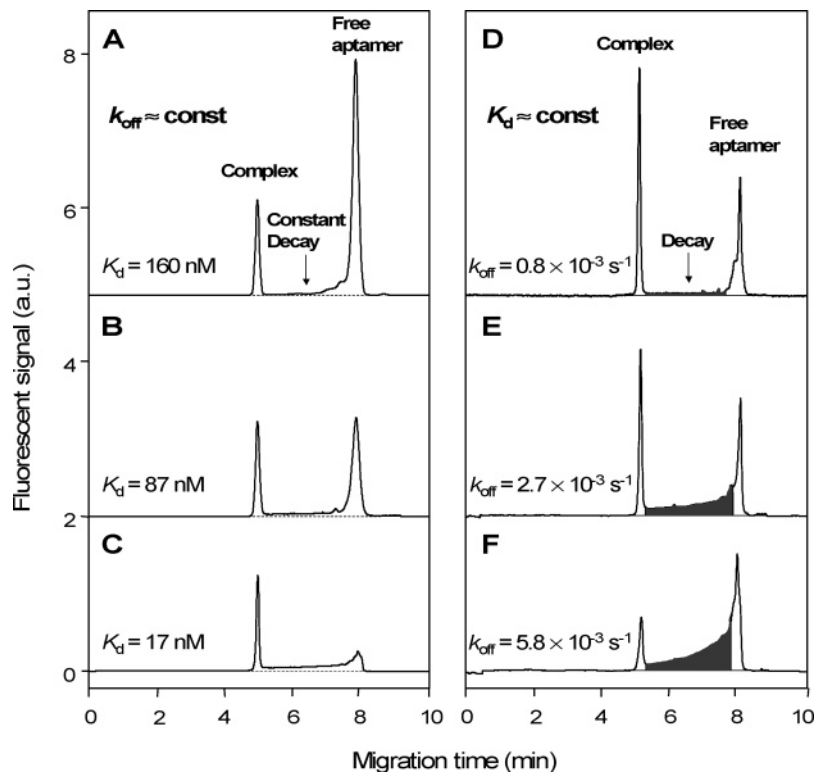
Figure 6. Binding of six different aptamers to MutS protein. Panels A–C correspond to aptamers with similar k_{off} constants, but varying K_d. Clones used are 4-14, 2-11, and 4-13 (Table 1), respectively. Panels D–F correspond to aptamers with similar K_d constants, but different k_{off}. Clones used are 2-16, 2-27, and 10-17, respectively. Binding conditions are kept identical. Concentration of MutS and aptamers were 100 and 10 nM, respectively.

Analysis of the distribution of K_d and k_{off} in each pool showed the anticipated trends. First, the closer the window was to the DNA library, the bigger the number of poor binders and also nonbinders ($K_d > 10 \mu$M) was found in that pool after cloning and measuring binding to MutS. Second, the average of K_d values of individual clones in each family was close to bulk affinity of the final pool and increased from pool I to pool III. Third, each family contained at least one clone with a theoretically predefined K_d value, even though only a few clones were screened for binding. Besides, each individual clone sequenced had a unique sequence. This may prove the initial assumption about the availability of oligonucleotides with different binding parameters in the naïve library. Interestingly, k_{off} values of the clones were distributed more or less randomly, proving that K_d, but not k_{off}, was a selection determinant in ECEEM selection. It is reasonable to assume that aptamers with different k_{off} and k_{on} constants, but with equal k_{off}/k_{on} ratio (which is equal to K_d), will appear the same fraction of collected DNA. However, exceptions also occur; for example, the aptamer sequence with the least K_d was selected in the second region. This clone (4-05) had an affinity of 3.6 ± 0.5 nM, which is 6×10^2 times lower than the bulk affinity of the naïve library.

It is noteworthy to emphasize that ECEEM-based selection does not yield 100% of sequences with desired values of K_d, but rather significantly increases the probability to discover such sequences. Indeed, even though we screened few individual clones, we obtained DNA sequences with K_d values predetermined at the beginning of selection. This unambiguously proves the power of ECEEM for the selection of smart ligands with predefined K_d values.

We also present the NECEEM-based analysis of individual smart aptamers selected by ECEEM. Two sets of aptamers with unique sequences were used for this purpose. One set of aptamers had similar k_{off} constants, but varying values of K_d, while the second set had similar K_d, but varying k_{off} (Figure 6). Changing K_d affected the ratios between free DNA and the complex (A–C), while changing k_{off} influenced the ratio between the intact and dissociated complex (D–F).

Selection of Smart Aptamers with Predefined k_{off}. In the next approach, a combinatorial DNA library is equilibrated with a protein target and the components of the equilibrium mixture are separated under nonequilibrium conditions, by NECEEM. The nonequilibrium conditions promote the dissociation of the complex during separation. Fractions collected in a time window preceding the DNA library yield pools of DNA sequences capable of binding the target and dissociating from the complex with specific rates (Figure 7). Being a homogeneous method with comprehensive kinetic features, NECEEM provides a means for selection of DNA aptamers with predefined ranges of all binding parameters of complex formation (K_d, k_{off}, k_{on}). First, the selection can be with respect to K_d values by varying the concentration of the protein target [P] in the equilibrium mixture. The ratio between protein-bound and unbound ligands changes according to classical equilibrium: the ligands where $K_d <$ [P] are preferentially bound to the protein and selected, while the ligands where $K_d >$ [P] are preferentially unbound and not selected. Second, the selection can be implemented with respect to k_{off} values by varying time windows in which fractions are collected. Finally, selection with respect to k_{on} values can be carried out by varying the time of
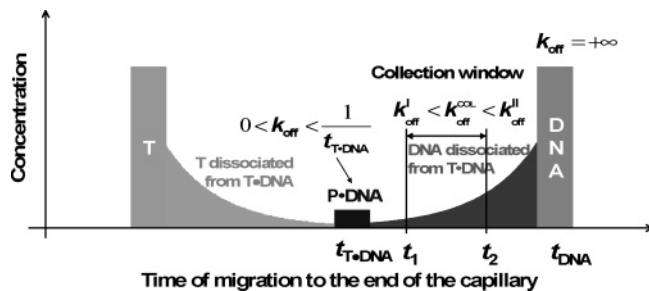
Figure 7. NECEEM concept for the selection of smart aptamers with predefined k_{off} constants.

Figure 8. NECEEM-based selection of two pools of smart aptamers with different and predetermined k_{off} values.

incubation of the library with the target. To select for a single binding constant, the parameters that control the other two binding constants should be kept unchanged. To select for k_{off}, we fix the protein concentration and the incubation time. In such a case, the time window in which a fraction is collected defines the range of k_{off} values in the following way:

$$k_{off}(t) = \frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}} \frac{1}{t_{DNA} - t} \quad (3)$$

Selection of DNA dissociated from the complex in the time window between t_1 and t_2 leads to DNA ligands with roughly the following k_{off} range:

$$\frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}} \frac{1}{t_{DNA} - t_1} < k_{off} < \frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}} \frac{1}{t_{DNA} - t_2} \quad (4)$$

Continuous selection within the $t_1 - t_2$ time window (Figure 7) leads to DNA molecules with k_{off} values being in the narrowing range around the following value:

$$k_{off}^{\infty} = \frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}(t_2 - t_1)} \ln\left(\frac{t_{DNA} - t_1}{t_{DNA} - t_2}\right) \quad (5)$$

where $t_{P \cdot DNA}$ and t_{DNA} are migration times in NECEEM of the target−DNA complex and free DNA, respectively (see Supporting Information for derivation).

NECEEM appears to be a more complicated method than ECEEM, because each collection window may contain ligands with totally different k_{off} values; however, ligands with k_{off} defined by eq 5 are the most abundant and will be predominant after multiple rounds of selection. The concept is different in ECEEM, in which every window theoretically contains only ligands with calculated K_d values, and deviations occur only because of nonspecific interactions during separation, nonzero width of the peaks, and disturbed equilibrium. Thus, NECEEM is a more "evolutionary" method of iterative selection, which requires more than one round to select oligonucleotides with a narrow range of k_{off} constants.

Experimentally, aptamers for MutS protein were selected in two different windows as shown in Figure 8. In region I, the theoretically predicted range of k_{off} was $0-1.0 \times 10^{-3}$ s^{-1}, and in region II, it was $1.7 \times 10^{-3} - 2.5 \times 10^{-3}$ s^{-1}. After two rounds of selection in regions I and II, the pools of DNA had the experimental bulk k_{off} values of 0.4×10^{-3} and 1.7×10^{-3} s^{-1},
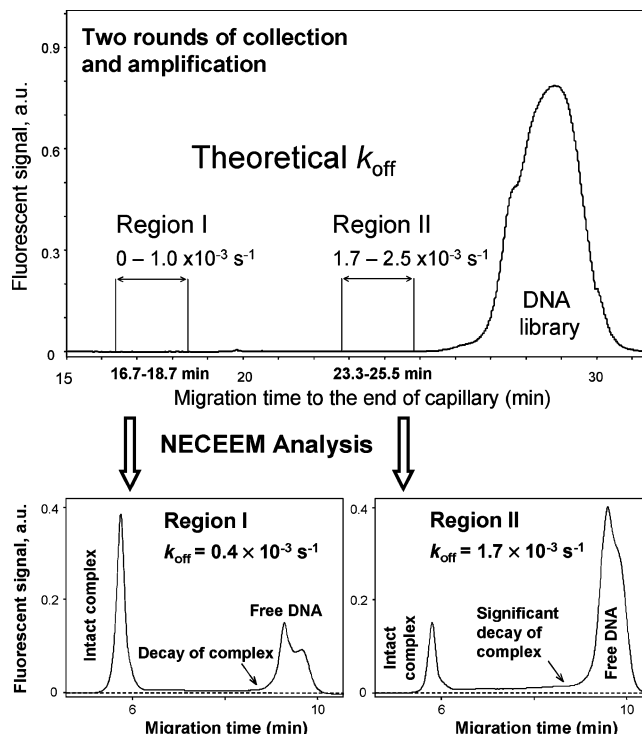
respectively. Thus, our theoretical consideration was proven experimentally.

CONCLUDING REMARKS

For the first time, we have selected smart aptamers, which are the ligands with predefined binding parameters. Besides, we selected them for MutS protein, for which aptamers have never been previously selected. The best aptamer selected had the affinity of 3.6 ± 0.5 nM. Our experimental results are in fine agreement with ECEEM and NECEEM theories for the selection of smart aptamers.

Certain comments should be made about the concept of smart aptamer selection with ECEEM and NECEEM. First, we assume that aptamers with different constants are present in the naïve library in considerable quantities. If aptamers with the desired constant are not present or scarce it may be impossible to select such ligands. Second, kinetics of interaction may be more complex if the stoichiometry of the protein−DNA complex is different from 1:1, requiring a more complex modeling.

Efficiency of aptamer selection determines the number of rounds required to generate the exceeding amounts of DNA binders over nonbinders. Homogeneous electrophoretic separation in free solution excludes the majority of nonspecific interactions, which usually occur on the solid−liquid interfaces. While the efficiency of partitioning in the common heterogeneous methods lies in the range of $10-100$, KCE partitioning achieves the values of 10^4-10^{5}.[34] The high efficiency reduces by far the number of selection rounds, so that KCE-based selection typically takes one to four rounds. Highly efficient selection of aptamers by KCE requires well-optimized PCR to gain full advantage of the high efficiency.[37,38]

Current KCE methodology provides a robust tool for generation of ligands that exhibit desired binding parameters. The next

challenge is to demonstrate the power of ECEEM and NECEEM for the selection of smart ligands from combinatorial libraries of antibodies, small molecules, peptides, and peptide—oligonucleotide conjugates.[39] Besides, KCE methods still retain their power as outstanding analytical techniques. For example, ECEEM was recently proposed as a method for the analysis of single-nucleotide polymorphisms with MutS protein as a mediator.[40] The quasi-equilibrium, which is the main feature of ECEEM, made it possible to unambiguously identify mismatched base pairs in double-stranded DNA as a function of their different K_d constants, which resulted in different mobilities in ECEEM. Although nucleic acid aptamers can be used as excellent ligands for diagnostic applications, their therapeutic potential is limited due to the instability in body fluids and lack of permeability through the cell mem-

branes. That is why peptide aptamers and small-molecule drugs with their structural diversity and enhanced physicochemical properties are the next step in the development of smart ligands. We believe that further development of KCE methods will provide a variety of methodological schemes for the screening of combinatorial libraries.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in test. This material is available free of charge via the Internet at http://pubs.acs.org.

(39) Takahashi, T. T.; Austin, R. J.; Roberts, R. W. *Trends Biochem. Sci.* **2003**, *28*, 159−165.

(40) Drabovich, A. P.; Krylov, S. N. *Anal. Chem.* **2006**, *78*, 2035−2038.

SUPPORTING INFORMATION

Selection of Smart Aptamers by Methods of Kinetic Capillary Electrophoresis

Andrei P. Drabovich, Maxim Berezovski, Victor Okhonin, and Sergey N. Krylov

Department of Chemistry, York University, Toronto, Ontario M3J 1P3, Canada

1. Determination of constants with NECEEM

The equilibrium constant (K_d) and rate constant (k_{off}) of protein-DNA complex dissociation were found from a NECEEM electropherogram using the areas of peaks of free DNA (A_3), DNA dissociated from the complex during electrophoretic separation (A_2), and DNA in the protein-DNA complex (A_1). To obtain correct values of A_3 and A_2, the apparent areas in the NECEEM electropherograms were divided by the migration time of free DNA, t_{DNA}. To obtain the correct value of A_1, its apparent area in the NECEEM electropherogram was divided by the migration time of the protein-DNA complex, $t_{P \cdot DNA}$. The value of K_d was calculated using the following formula:

$$K_d = \frac{[P]_0 \{1 + A_3/(A_1 + A_2)\} - [D]_0}{1 + (A_1 + A_2)/A_3} \tag{1}$$

where $[P]_0$ and $[D]_0$ are initial concentrations of protein and DNA, respectively. Rate constant k_{off} value was calculated using the following formula:

$$k_{off} = \frac{1}{t_{P \cdot DNA}} \ln \frac{A_1 + A_2}{A_1} \tag{2}$$

The precisions (RSD) of K_d and k_{off} measurements with NECEEM were 20% and 5%, respectively. The lower RSD value for k_{off} is explained by the independence of k_{off} on $[P]_0$ and $[D]_0$ (see equations), which excludes pipetting errors.

2. ECEEM selection of smart aptamers

The apparent mobility of DNA in ECEEM is an extensive function, which includes two boundary mobilities: $\mu_{DNA \atop P=0}$, which is the mobility of free DNA in the absence of protein, and $\mu_{P \cdot DNA \atop P \to +\infty}$, which is the mobility of protein-DNA complex, when protein is present in the run buffer at the infinite concentration:

$$\mu_{DNA}^{App} = \mu_{DNA \atop P=0} \times \frac{K_d}{[P]_{free} + K_d} + \mu_{P \cdot DNA \atop P \to +\infty} \times \frac{[P]_{free}}{[P]_{free} + K_d} \tag{3}$$

In case of DNA library we have about 10^{12} unique molecules in the injected plug; however, mobilities $\mu_{DNA \atop P=0}$ and $\mu_{P \cdot DNA \atop P \to +\infty}$ are the same for all molecules because of the same charge to size ratios for free DNA and protein-DNA complexes, respectively. As the mobility is inversely proportional to migration time of species through the expression:

$$\mu = \frac{\text{Capillary Length}}{\text{Time} \times \text{Electric Field Strength}} = \frac{L}{t \times E} \tag{4}$$

we can rewrite expression (3) using migration times:

$$\frac{1}{t} = \frac{1}{t_{DNA}} \times \frac{K_d}{[P]_{free} + K_d} + \frac{1}{t_{P \cdot DNA}} \times \frac{[P]_{free}}{[P]_{free} + K_d} \tag{5}$$

$[P]_{free}$ is assumed to be equal to the overall concentration of the protein [P] because of the constant supply of the protein from the run buffer. By rearranging expression (5) we can get the pre-defined affinity K_d as a function of migration time t in the capillary:

$$K_d(t) = [P] \times \frac{t_{DNA}}{t_{P \cdot DNA}} \times \frac{t - t_{P \cdot DNA}}{t_{DNA} - t} \tag{6}$$

3. NECEEM selection of smart aptamers with pre-defined k_{off} constants

To select aptamers with $k_{off}^{min} < k_{off} < k_{off}^{max}$, fractions are collected within the time window corresponding to the dissociation of protein-DNA complexes (see **Figure 7** in the article). It is essential to exclude from the fractioning the intact peak of protein-DNA complex, since it contains aptamers with $0 < k_{off} < 1/t_{P \cdot DNA}$. The rate of generation of free DNA during the dissociation of the complex is described by the following equation:

$$\frac{d[DNA]}{dt'} = \frac{d[P]}{dt'} = -\frac{d[P \cdot DNA]}{dt'} = [P \cdot DNA]_{eq} \exp(-k_{off}t') \tag{7}$$

where t' is time passed from the beginning of separation.

Our goal is to find a value of k_{off} for the ligands, which are preferably collected at time t from the beginning of separation in the infinite number of selection rounds. To accomplish this task, we have to figure out first for which k_{off} values the rate of decay is maximal at time t. The following equation has to be solved:

$$\frac{d}{dk_{off}}\frac{d[\text{DNA}]}{dt'} = 0$$

$$\frac{d}{dk_{off}}\frac{d[\text{DNA}]}{dt'} = \frac{d}{dk_{off}}\left\{[\text{P}\bullet\text{DNA}]_{eq}\,k_{off}\exp(-k_{off}t')\right\} =$$

$$[\text{P}\bullet\text{DNA}]_{eq}\exp(-k_{off}t') - k_{off}t'[\text{P}\bullet\text{DNA}]_{eq}\exp(-k_{off}t') =$$

$$[\text{P}\bullet\text{DNA}]_{eq}\exp(-k_{off}t')\left(1 - k_{off}t'\right) = 0$$

(8)

Equation (8) has the only one solution:

$$k_{off} = 1/t'$$

(9)

If the fraction is collected within time window between t_1 and t_2 (see **Figure 6** in the article), then instead of (7) we should use the difference:

$$[\text{DNA}]_{t_2} - [\text{DNA}]_{t_1} = \exp(-k_{off}t_2') - \exp(-k_{off}t_1')$$

(10)

Thus, k_{off} value is determined by the equation:

$$\frac{d}{dk_{off}}\left\{\exp(-k_{off}t_2') - \exp(-k_{off}t_1')\right\} = 0$$

(11)

Finally, instead of (9) we have:

$$k_{off} = \frac{1}{t_2' - t_1'}\ln(t_2'/t_1')$$

(12)

Now, we have to relate time t' in (9), which is time passed from the beginning of separation to complex dissociation, to the migration time, t of the dissociated DNA. The migration time, t, of DNA which dissociated from the complex after t' is:

$$t = t' + \frac{L - t'v_{\text{P}\bullet\text{DNA}}}{v_{\text{DNA}}} = t' + \frac{L - t'L/t_{\text{P}\bullet\text{DNA}}}{L/t_{\text{DNA}}} = t' + \frac{1 - t'/t_{\text{P}\bullet\text{DNA}}}{1/t_{\text{DNA}}} =$$

$$t' + t_{\text{DNA}}(1 - t'/t_{\text{P}\bullet\text{DNA}}) = t' + t_{\text{DNA}} - t't_{\text{DNA}}/t_{\text{P}\bullet\text{DNA}} = t'(1 - t_{\text{DNA}}/t_{\text{P}\bullet\text{DNA}}) + t_{\text{DNA}}$$

where L is a length of the capillary and v is a velocity of species in the capillary.
Solving this equation for t' we get:

$$t' = \frac{t - t_{\text{DNA}}}{1 - t_{\text{DNA}}/t_{\text{P}\bullet\text{DNA}}} = t_{\text{P}\bullet\text{DNA}}\frac{t - t_{\text{DNA}}}{t_{\text{P}\bullet\text{DNA}} - t_{\text{DNA}}} = t_{\text{P}\bullet\text{DNA}}\frac{t_{\text{DNA}} - t}{t_{\text{DNA}} - t_{\text{P}\bullet\text{DNA}}}$$

(13)

where t_{DNA} and $t_{P \cdot DNA}$ are migration times of free DNA and the protein-DNA complex, respectively. Expression (12) can now be rewritten as:

$$k_{off} = \frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}(t_1 - t_2)} \ln\left(\frac{t_{DNA} - t_2}{t_{DNA} - t_1}\right)$$ (14)

where t_1 and t_2 are the boundary times of the aptamer collection window.

If the finite number of selection rounds is accomplished k_{off} value roughly gets into the following region:

$$\frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}} \frac{1}{t_{DNA} - t_1} < k_{off} < \frac{t_{DNA} - t_{P \cdot DNA}}{t_{P \cdot DNA}} \frac{1}{t_{DNA} - t_2}$$ (15)