


An T. H. Le
Svetlana M. Krylova
Sergey N. Krylov 

Department of Chemistry and
Centre for Research on
Biomolecular Interactions,
York University, Toronto,
Ontario, Canada

Received January 14, 2019
Revised April 16, 2019
Accepted May 2, 2019

Research Article

Ideal-filter capillary electrophoresis: A highly efficient partitioning method for selection of protein binders from oligonucleotide libraries

Selection of affinity ligands for protein targets from oligonucleotide libraries currently involves multiple rounds of alternating steps of partitioning of protein-bound oligonucleotides (binders) from protein-unbound oligonucleotides (nonbinders). We have recently introduced ideal-filter capillary electrophoresis (IFCE) for binder selection in a single step of partitioning. In IFCE, protein-binder complexes and nonbinders move inside the capillary in the opposite directions, and the efficiency of their partitioning reaches 10^9 , i.e., only one of a billion molecules of nonbinders leaks through IFCE while all binders pass through. The condition of IFCE can be satisfied when the magnitude of the mobility of EOF is smaller than that of the protein-binder complexes and larger than that of nonbinders. The efficiency of partitioning in IFCE is 10 million times higher than those of solid-phase-based methods of partitioning typically used in selection of affinity ligands for protein targets from oligonucleotide libraries. Here, we provide additional details on our justification for IFCE development. We elaborate on electrophoretic aspects of the method and define the theoretical range of EOF mobilities that support IFCE. Based on these theoretical results, we identify an experimental range of background electrolyte's ionic strength that supports IFCE. We also extend our interpretation of the results and discuss in-depth IFCE's prospective in practical applications and fundamental studies.

Keywords:

DNA aptamers / Ideal-filter capillary electrophoresis / Ionic strength of background electrolyte / Nonbinder background / Oligonucleotide libraries

DOI 10.1002/elps.201900028

1 Introduction

Screening molecular libraries for protein binders of non-protein nature is the mainstream approach for identifying drug leads [1–3]. The probability of finding a suitable protein binder increases with increasing structural diversity of the library. [4] Libraries-collections, in which every molecule is synthesized, stored, and tested for binding to the target protein separately, are limited to approximately a million different molecules. [5, 6] In contrast, libraries-mixtures synthesized via the split and mix approach can provide virtually unlimited diversity. However, *in vitro* selection of protein binders from such libraries requires physical partitioning of

protein-bound molecules (binders) from protein-unbound molecules (nonbinders) before their identification. [7, 8] Increasing library diversity reduces the number of copies of unique molecules in the library and makes binder identification more challenging [4].

Oligonucleotide libraries-mixtures, such as random-sequence RNA and DNA libraries as well as DNA-encoded libraries (DELs) [9–11], address the challenge of identification of a limited number of copies of the binders. Random-sequence libraries are used to select DNA and RNA aptamers, while DELs are used to select small molecules capable of binding target proteins. Oligonucleotide moieties of binders from such libraries can be amplified by the PCR and their sequences can be identified via DNA sequencing. These sequences represent binder “identities.” For DNA and RNA aptamers, knowing their sequences allows making their replicas via chemical synthesis of DNA and RNA. For binders selected from DELs, sequence information is used to deconvolute the structural information of the binders (e.g., structures of small-molecule binders); the binders are then synthesized without DNA tags for further development (e.g., as drug leads) [11].

Correspondence: Professor Sergey N. Krylov, Department of Chemistry, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3

E-mail: skrylov@yorku.ca

Abbreviations: DEL, DNA-encoded library; IFCE, ideal-filter capillary electrophoresis; NECEEM, nonequilibrium capillary electrophoresis of equilibrium mixture; qPCR, quantitative polymerase chain reaction; SELEX, systematic evolution of ligands by exponential enrichment

Color online: See the article online to view Figs. 1, 3–5, 7, and 8 in color.

The ability to identify a limited number of binders makes oligonucleotide libraries very attractive for *in vitro* selection of drug leads [11–13].

The abundance of binders in an oligonucleotide library is typically very low, e.g., one binder per million or per billion of library molecules. At such low abundancies, the binders cannot be partitioned from nonbinders in a single step of partitioning. The reason for this is the “leakage” of nonbinders through the partitioning process; this leakage creates a “non-binder background” [14–17]. As a result, multiple consecutive rounds of partitioning are typically used for *in vitro* selection of binders from oligonucleotide libraries [18]. The multi-round selection is resource-consuming and time-consuming. In multiround selection of oligonucleotide aptamers via the systematic evolution of ligands by exponential enrichment (SELEX) process (repetition of binding/partitioning and PCR amplification), the number of rounds is theoretically unlimited, but a large number of rounds can lead to selection failure due to sequence biases of polymerases [19–21]. In selection of binders from DELs, the number of rounds is limited to three or four due to (i) binder loss in partitioning and (ii) the inability to PCR-amplify the binders, which are not oligonucleotides in this case (e.g., small molecules) [11, 22–24].

An ultimate solution for problems originating from multiround selection would be a partition method that could enrich binders to the required level (e.g., 99% of binders in the binder-enriched library) in one step of partitioning. There have been several reports of one-step selection of oligonucleotide aptamers [25–31]; however, they have not been independently confirmed raising doubts in method practicality, transferability, and reliability. We think there are two major reasons for slow progress in creating a practical way of single-step selection of binders from oligonucleotide libraries. The first reason is technological: it is extremely difficult to achieve high efficiencies of partitioning, e.g., due to adsorption of nonbinders to surfaces in solid-phase selection methods [14–17], and non-uniform migration of the nonbinders in homogeneous CE-based methods [32]. The second reason is methodological: while high efficiencies of partitioning are the goal, the efficiency of partitioning was not used to guide developments or substantiate claims of one-step selection; moreover, it was rarely measured [25–30].

In our recent work, we addressed both the technological and methodological issues. First, we estimated theoretically the efficiency of partitioning required for one-step selection to be roughly 10^9 , which is at least a million times higher than efficiencies of typical solid-phase methods of partitioning [14–17, 33–35], and 10 000 times higher than the efficiency of classical CE-based selection methods [31]. Second, we developed a new partitioning method, ideal-filter capillary electrophoresis (IFCE), in which target-binder complexes and nonbinders move in the opposite directions inside the capillary. Third, the efficiency of partitioning was adjusted by changing the background electrolyte’s ionic strength (I_{BGE}) to reach the required value of 10^9 . Finally, we demonstrated

one-step selection of a high-affinity aptamer pool to a protein target (MutS protein) to confirm that the estimated efficiency of 10^9 required for one-step selection was correct. Our first report on IFCE has been recently published [36]. Here, we elaborate on the justification for IFCE development and important electrophoretic aspects of the method. We derive simple relations between the electrophoretic mobility of EOF (μ_{EOF}) and those of the nonbinders (μ_{N}) and target–binder complexes (μ_{TB}). Based on experimentally determined dependencies of μ_{EOF} , μ_{N} , and μ_{TB} on I_{BGE} , we determine the range of I_{BGE} , which corresponds to IFCE. We also provide further interpretation of the results and methods prospective in practical applications and fundamental studies.

2 Materials and methods

2.1 Chemicals and reagents

All chemicals were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. Fused-silica capillaries with inner and outer diameters of 75 and 360 μm , respectively, were purchased from Molex Polymicro (Phoenix, AZ, USA). Recombinant *Thermus aquaticus* MutS protein (MW \approx 90 kDa, pI 6.0) was expressed and purified as described previously [37]. All DNA molecules were custom synthesized by Integrated DNA Technologies (Coralville, IA, USA). Bodipy (4,4-difluoro-4-bora-3a,4a-diaza-s-indacene) was purchased from Life Technologies Inc. (Burlington, ON, Canada). The BGE was 50 mM Tris-HCl with NaCl ranging from 0 to 150 mM at pH 7.0. The sample buffer was always identical to BGE to prevent adverse effects of buffer mismatch. Accordingly, all dilutions of sample components used in CE experiments were done by adding the corresponding BGE.

2.2 DNA sequences

The DNA aptamer with affinity toward MutS protein was previously selected in our laboratory (clone 2–06) [38], and its fluorescein amidite (FAM)-labeled version was used here: 5'-FAM-CTT CTG CCC GCC TCC TTC CTG GTA AAG TCA TTA ATA GGT GTG GGG TGC CGG GCA TTT CGG AGA CGA GAT AGG CGG ACA CT-3'. For aptamer-selection study, a synthetic FAM-labeled DNA library (N40) with a 40-nt random region was used: 5'-FAM-AGC CTA ACG CAG AAC AAT GG-(N40)-CGA TGC CAG GTT AAA GCA CT-3'. The following primers were used for PCR amplification of the MutS aptamer: forward primer (MutS_uF), 5'-CTT CTG CCC GCC TCC TTC C-3'; reverse primer (MutS_uR), 5'-AGT GTC CGC CTA TCT CGT CTC C-3'. Two sets of primers were used to amplify binders selected from the naïve library. The first set of primers was an unlabeled forward primer (N40_uF), 5'-AGC CTA ACG CAG AAC AAT GG-3', and an unlabeled reverse primer (N40_uR), 5'-AGT GCT TTA ACC TGG CAT CG-3'. The second set contained a FAM-labeled forward primer (N40_famF), 5'-FAM-AGC CTA ACG CAG AAC AAT GG-3',

and a biotin-labeled reverse primer (N40-biotinR), 5'-Biotin-TEG-AGT GCT TTA ACC TGG CAT CG-3'.

2.3 Default conditions for CE and fraction collection

All CE experiments were performed with a P/ACE MDQ apparatus (SCIEX, Concord, ON, Canada) equipped with a 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 [39]. Uncoated fused-silica capillaries, with a total length of 50 cm and a 10.2 cm distance from one of the ends to the detection zone were used. The two capillary ends were used as inlets interchangeably in experiments requiring different separation distances. The separation distance was defined as the distance in the capillary from the sample position at the time of electric field application to the detection point. Prior to every run, the capillary was rinsed successively with 0.1 M HCl, 0.1 M NaOH, deionized H₂O, and a run buffer for 3 min each. Mixtures of various concentration combinations were prepared, as described in the Results and Discussion section. Prior to the mixing, the DNA stock solution was incubated at 90°C for 2 min and gradually cooled down to 20°C at a rate of 0.5°C/s. For mobility study, the sample contained 100 nM MutS protein, 100 nM MutS aptamer and 150 nM Bodipy (EOF marker). The sample mixture was incubated for 30 min at a room temperature (22–24°C) and then injected with a pressure pulse of 0.5 psi × 10 s to yield a 10 mm long sample plug. The injected sample plug was propagated through the uncooled part of the capillary at the inlet by injecting a 5.7 cm long plug of the BGE with a pressure pulse of 0.3 psi × 90 s. CE was carried out at an electric field of 200 V/cm (10 kV over 50 cm). The duration of electrophoretic runs without fraction collection was 50 min. The duration of electrophoretic runs with fraction collection was 64 min. Collection vials contained 20 µL of the BGE each and were switched every 2 min; 32 fractions were collected for each run.

2.4 Quantitative PCR

DNA in the collected fractions was amplified and quantitated by quantitative polymerase chain reaction (qPCR) using a CFX Connect™ instrument from Bio-rad. q-PCR reagent mixture was prepared by combining IQ SYBR Green Supermix from Bio-Rad (Mississauga, ON, Canada) with unlabeled DNA primers at final concentrations of 1× SYBR Green Supermix, 100 nM MutS_{uF}, and 100 nM MutS_{uR}. qPCR reaction mixture was prepared by adding 18 µL of the qPCR reagent mixture to a 2 µL aliquot of each fraction immediately before thermocycling. The thermocycling protocol was 95°C for 3 min, 95°C for 10 s (denaturation), 56°C for 10 s (annealing), 72°C for 10 s (extension), followed by a plate read at 72°C and a return to the denaturation step (bypassing the 95°C × 3 min step) for a total of 43 cycles. All reactions were performed in duplicates.

2.5 Specifics of single-round aptamer selection

Fraction collection and qPCR detection were similar to the procedures described in the previous two sections with a few modifications specified below. The equilibrium mixture contained: 10 µM N40 library, 100 nM MutS protein, and 150 nM Bodipy. For qPCR, 1× SYBR Green Supermix and the unlabeled primers for N40 library (100 nM of each N40_{uF} and N40_{uR}) constituted the qPCR reagent mixture. The fraction that eluted at minute 29 and contained the highest amount of MutS-DNA complexes was subjected to preparative PCR. The procedure of preparative PCR involved two rounds of amplification. In the first round, the fraction was amplified by qPCR in quintuplicates as previously described. An S-shaped amplification curve was plotted, and the PCR product was removed two cycles into the exponential phase of the curve. After qPCR, 100 µL of the five combined PCR reactions was purified using MinElute PCR purification kit (Qiagen, Mississauga, ON, Canada) as per manufacturer's instructions. DNA was then eluted using 20 µL of 50 mM Tris-HCl, pH 7.0. Once product's purity was verified by native PAGE, it was subjected to asymmetric PCR. Five microliters of DNA was added to 45 µL of asymmetric PCR reagent mixture from New England Biolabs (MA, USA). Final concentrations of PCR reagents in the reaction mixture were as follows: 1× Standard *Taq* Reaction Buffer, 1 µM N40-famF, 50 nM N40-biotinR, 2.5 units/µL *Taq* DNA Polymerase, and 200 µM dNTPs mix. The reaction was performed in duplicates with the following temperature protocol: 94°C for 30 s (initial denaturation, performed once), 94°C for 10 s (denaturation), 56°C for 10 s (annealing), and 72°C for 10 s (extension). Seventeen cycles of asymmetric PCR were run. Ten microliters of MagnaBind streptavidin beads suspension (Thermo Scientific, IL, USA) was washed three times and resuspended in bead washing/binding buffer (10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA pH 8.0). Once amplified, the duplicate PCR reactions were combined and incubated with streptavidin magnetic beads for 30 min at room temperature. The beads were magnetized, discarded, and the PCR product was then purified using the MinElute PCR purification kit. The final product was eluted using 20 µL of 50 mM Tris-HCl, pH 7.0, and 2 µL of 1 M NaCl was added to bring NaCl concentration to 100 mM.

To determine DNA concentration in the enriched library pool, serial dilutions of N40-famF (2, 1 µM, 500, 250, 125, 62.5, and 31.25 nM) were prepared to build a standard curve by measuring fluorescence intensity at 519 nm with NanoDrop 3300 Fluorospectrometer (Thermo Scientific). DNA concentration in the enriched library pool was found to be 1.2 µM.

For a pressure-aided IFCE-based binding test of the enriched pool (used to determine $K_{d,app}$), a 47 nL plug of the equilibrium mixture containing 20 nM enriched library and 100 nM MutS in 50 mM Tris-HCl, 100 mM NaCl, pH 7.0 (default BGE for IFCE) was injected into 50 cm-long capillary by a 0.5 psi × 10 s pressure pulse. The sample mixture was propagated through the noncooled portion of the capillary by injecting a 5.7 cm-long plug of BGE with a pressure

pulse of $0.3 \text{ psi} \times 90 \text{ s}$. CE was carried out at an electric field of 200 V/cm with “+” at the inlet. In addition to applying voltage, a pressure of 0.20 psi was applied to the capillary inlet to supplement the electric field and ensure that the nonbinders reach the detector. The pressure-aided IFCE allowed detection of target-binder complexes and nonbinders, which is required for determination of $K_{d,\text{app}}$ [40].

3 Results and discussion

3.1 Efficiency of partitioning required for one-step selection of binders

If the library contains molecules capable of binding the target protein under the chosen selection conditions, then the equilibrium mixture of the library and the target contains target-bound library molecules (binders) and target-unbound library molecules (nonbinders). The purpose of partitioning in binder selection is to separate binders from nonbinders physically. Thus, the partitioning process is conceptually filtration that should let binders through and cut nonbinders. To describe partitioning quantitatively, we adopt a term of “transmittance,” which is used for optical filters. Transmittance of partitioning for binders, k_B , is the ratio between quantities of binders at the output, B_{out} , and input, B_{in} , of partitioning, respectively: $k_B = B_{\text{out}}/B_{\text{in}}$. Transmittance of partitioning for nonbinders, k_N , is the ratio between the quantities of nonbinders at the output, N_{out} , and input, N_{in} , of partitioning,

respectively: $k_N = N_{\text{out}}/N_{\text{in}}$. Note, that k_N is a fraction of nonbinders that penetrates through the step of partitioning; this fraction contaminates binders and creates nonbinder background in the selection process. Thus, k_N is a quantitative measure of nonbinder background.

One step of partitioning is typically not sufficient for reaching the desired binder-to-nonbinder ratio ($B_{\text{out}}/N_{\text{out}} = Q$ at the output. It is preferable to have $Q > 1$; $Q = 100$ corresponds to 99% purity of binders in the final post-selection binder-enriched library. There are two basic approaches for binder selection from oligonucleotide libraries: non-SELEX and SELEX (Fig. 1) [41]. Non-SELEX does not involve PCR amplification of binders while SELEX does.

In non-SELEX, a single round of partitioning produces (Fig. 1A):

$$\frac{B_{\text{out}}}{N_{\text{out}}} = \frac{B_{\text{in}} k_B}{N_{\text{in}} k_N} \quad (1)$$

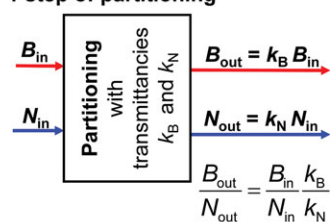
If one step of partitioning is not sufficient to reach the desired Q , then $m > 1$ consecutive steps of partitioning are employed (Fig. 1A) resulting in:

$$\frac{B_{\text{out}}}{N_{\text{out}}} = \frac{B_{\text{in}}}{N_{\text{in}}} \left(\frac{k_B}{k_N} \right)^m \quad (2)$$

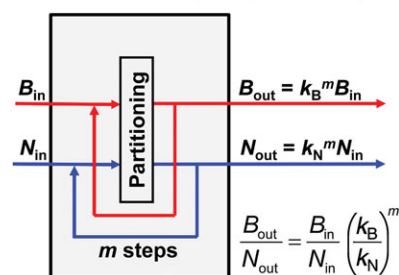
Non-SELEX is the only option for selection of binders from DELs as binders cannot be amplified by PCR. Non-SELEX can also be used as an option for selection of DNA and RNA aptamers. Non-SELEX appears to be robust, but it will typically fail when the number of required consecutive steps of partitioning exceeds $m = 3$. The failure is due to

A Non-SELEX: selection of binders without PCR amplification between steps of partitioning

1 step of partitioning

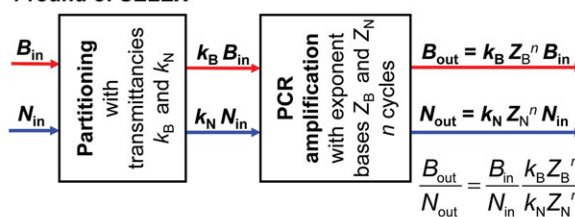


m consecutive steps of partitioning



B SELEX: selection of binders with n cycles of PCR amplification after every step of partitioning

1 round of SELEX



m rounds of SELEX

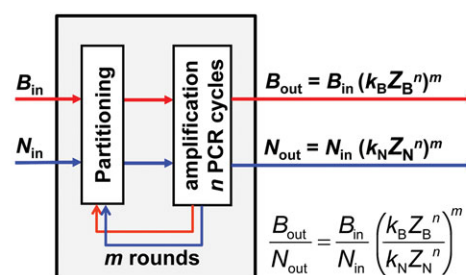


Figure 1. Schematic depiction of non-SELEX (A) and SELEX (B) processes of binder selection from oligonucleotide libraries. See text for details.

the trivial loss of binders at every step of partitioning. Fraction collection into a vial with the BGE leads to binder dilution and the following injection of only a small part of this diluted solution inevitably results to binder loss. We previously estimated that the loss of binders is approximately $100\times$ per round [41].

In SELEX, which can only be used for selection of binders from random-sequence oligonucleotide libraries but not for selection from DELs, binders (and leaking nonbinders) are amplified by PCR after every step of partitioning (Fig. 1B). This amplification maintains the binders in the selection process and appears like an ultimate solution for the binder loss in non-SELEX. However, as we will see below, the amplification of binders does not make the selection robust. One round of PCR results in:

$$\frac{B_{\text{out}}}{N_{\text{out}}} = \frac{B_{\text{in}}}{N_{\text{in}}} \frac{k_{\text{B}}}{k_{\text{N}}} \left(\frac{Z_{\text{B}}}{Z_{\text{N}}} \right)^n \quad (3)$$

where Z_{B} and Z_{N} are the exponents of the base of exponential function describing the amplification process and n is the number of cycles of PCR. In m consecutive rounds of SELEX the resulting binder-enriched library has:

$$\frac{B_{\text{out}}}{N_{\text{out}}} = \frac{B_{\text{in}}}{N_{\text{in}}} \left\{ \frac{k_{\text{B}}}{k_{\text{N}}} \left(\frac{Z_{\text{B}}}{Z_{\text{N}}} \right)^n \right\}^m \quad (4)$$

Then, according to the recently introduced Algebraic Model of binder selection, reaching $(B_{\text{out}}/N_{\text{out}}) \geq Q$, requires that the number of rounds, m , be:

$$m \geq \left\lceil \frac{\log \left\{ \frac{B_{\text{in}}/N_{\text{in}}}{Q} \right\}}{\log \left\{ \frac{k_{\text{B}}}{k_{\text{N}}} \left(\frac{Z_{\text{B}}}{Z_{\text{N}}} \right)^n \right\}} \right\rceil \quad (5)$$

where $\lceil x \rceil$ represents a mathematical function that rounds x up to the nearest integer. (Note that in case of non-SELEX this formula still stands by assuming $Z_{\text{N}} = Z_{\text{B}} = 1$ and, thus, replacing the multiplier $(Z_{\text{N}}/Z_{\text{B}})^n$ with unity.)

At one extreme, according to Eq. (5) SELEX fails unconditionally (m approaches $+\infty$ or becomes negative) when:

$$\frac{k_{\text{B}}}{k_{\text{N}}} \left(\frac{Z_{\text{B}}}{Z_{\text{N}}} \right)^n \leq 1 \quad (6)$$

Ideally, $k_{\text{B}} = 1$, $k_{\text{N}} = 0$, and $Z_{\text{B}} = Z_{\text{N}} = 2$; in reality, however, $k_{\text{B}} < 1$, $k_{\text{N}} > 0$, and $Z_{\text{B}} < Z_{\text{N}}$. As an example, for typical values of $k_{\text{B}}/k_{\text{N}} = 10$ and $n = 30$, SELEX fails when the amplification bias towards nonbinders is only about 7% ($Z_{\text{B}}/Z_{\text{N}} \approx 0.93$). Increasing n leads to SELEX failure at even a lesser amplification bias.

At the other extreme, according to Eq. (5), a single step of partitioning is sufficient for completing binder selection, i.e., reaching $(B_{\text{out}}/N_{\text{out}}) \geq Q$, when:

$$\frac{k_{\text{B}}}{k_{\text{N}}} \geq Q \frac{N_{\text{in}}}{B_{\text{in}}} \quad (7)$$

Assuming that $k_{\text{B}} \approx 1$ (which can be satisfied in principle), we can define a simple requirement for binder selection is a single step of partitioning:

$$k_{\text{N}} \leq \frac{B_{\text{in}}/N_{\text{in}}}{Q} \quad (8)$$

Values of $(B_{\text{in}}/N_{\text{in}})$ are hard to estimate for real random-sequence oligonucleotide libraries. Our rough estimate, done via binder selection from a DNA library in three consecutive steps of partitioning without PCR amplification between them, gave $(B_{\text{in}}/N_{\text{in}}) = \sim 10^{-7}$ [41]. According to Eq. (8), this estimate suggests $k_{\text{N}} \sim 10^{-9}$ as a rough value, which should be typically sufficient for binder selection in a single step of partitioning. Furthermore, we use this value as a guide in our development of a partitioning method expected to be suitable for one-step selection of binders.

Here we need to return to a more detailed analysis of available partitioning methods for k_{N} values that can be achieved in them. All practical partitioning methods are certainly characterized by $k_{\text{N}} \gg 10^{-9}$ [10, 14–17, 26, 33–35, 42–50]. Surface-based methods (e.g., partitioning on filters and magnetic beads) are most widely used in non-SELEX and SELEX, but typically have $k_{\text{N}} > 10^{-3}$ due to nonspecific adsorption of nonbinders onto the surface and contamination of binders. There have been successful attempts to reduce the nonbinder background in surface-based partitioning, e.g., via digestion of surface-bound nonbinders by a nuclease [25, 26], and by using microfluidics to precisely manipulate magnetic beads [30]. However, these methods were not adopted by other users and, thus, did not become practical. Also, the results were not substantiated by measurements of k_{N} to put them in quantitative context with other partitioning methods.

Homogeneous CE-based partitioning was introduced to overcome the limitations of surface-based methods in separation of binders from nonbinders [31, 38, 41, 50–53]. Theoretical estimates indicate that $k_{\text{N}} \sim 10^{-9}$ could be easily achieved in CE-based partitioning if the peaks had classical Gaussian shapes [54]. However, experiments revealed only $k_{\text{N}} \approx 10^{-5}$, which is two orders lower than the lowest k_{N} of practical solid-phase methods but four orders of magnitude higher than $k_{\text{N}} \sim 10^{-9}$ required for one-step selection [31]. The main reason for lower-than-expected k_{N} of CE-based partitioning lies in an unusual pattern of oligonucleotide migration in CE [32]. There is always a small part of oligonucleotides, which tails from the main part of oligonucleotides towards the protein-binder complexes. In the context of CE-based partitioning, there is a small part of nonbinders that tails towards binders and creates nonbinder background in the binder-collection time window. The previous study of this background suggested that it was a result of slow dissociation of DNA-counterion complexes. Increasing the concentration of NaCl in a solution of DNA led to increased background when such DNA was sampled for electrophoresis (Fig. 2). The background has a peak that migrates with the velocity similar to those of typical protein–DNA complexes. This peak can be even confused with that of the protein–DNA

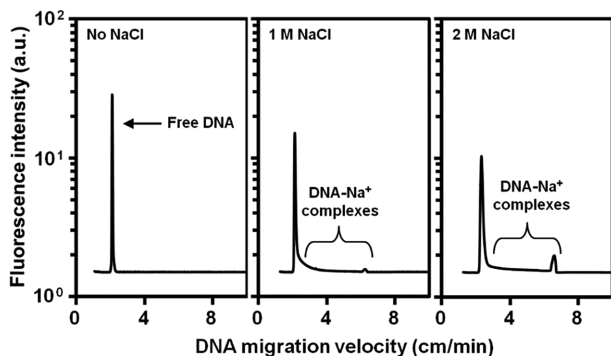


Figure 2. The influence of NaCl concentration in the sample of DNA on the DNA background in CE. No background is visible at no NaCl added to the ssDNA sample, but this absence of the background is only apparent as the background is present but is below the limit of fluorescence detection. The increase in concentration of NaCl added to the DNA sample leads to the increase of the background and even the appearance of the second peak. This graph was adopted from Ref. 32 with permission of the American Chemical Society.

complex if a proper negative control is not run. The DNA background is fundamental (instrument-independent) in its nature; thus, finding a way to decrease it proved to be highly challenging. Our unpublished attempts to simply increase the resolution between the peaks of binder-target complex and nonbinders proved to be inefficient in decreasing the background.

In conclusion of this consideration, we would like to restate that the available methods of partitioning have high nonbinder background and can hardly facilitate reliable and practical one-step selection of binders from ultimately diverse oligonucleotide libraries. In our attempt to rationally develop a partitioning method with $k_N \approx 10^{-9}$, we decided to utilize CE as an instrumental platform since it provides a means of easily measuring the efficiency of partitioning and since it demonstrated the best efficiencies of partitioning on record [31].

3.2 Hypothesis for reducing nonbinder background in CE-based partitioning

Partitioning by CE is based on separation of target-binder complexes from nonbinders in an electric field. All oligonucleotides in the library have the same length and the same negative charge. Therefore, all nonbinders are characterized by a single value of electrophoretic mobility, μ_N , which is negative, i.e., vector μ_N has opposite direction to the vector of electric field E .

Typical targets for partition by CE are proteins capable of electrophoretic shift of oligonucleotides upon binding them. It is extremely unlikely that a protein has a positive charge of a magnitude greater than the negative charge of the oligonucleotide to make the complex positively charged. Accordingly, all target-binder complexes (for a single target) have identical

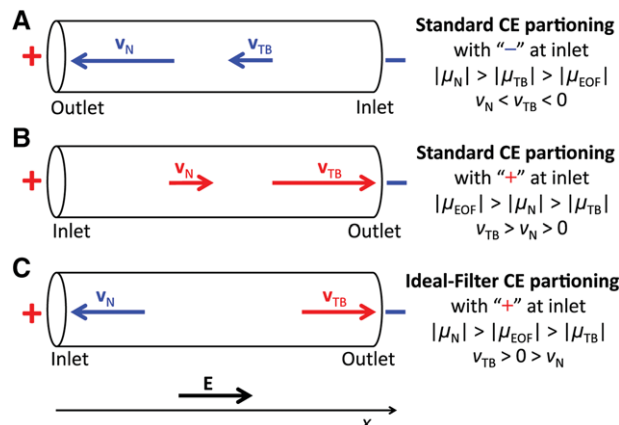


Figure 3. Schematics of CE-based partitioning of target-binder complexes (TB) from nonbinders (N). Panels A and B show standard CE-based partitioning with velocity vectors of complexes and nonbinders directed towards the capillary outlet biased at “+” and “-” respectively. Panel C shows ideal-filter capillary electrophoresis (IFCE)-based partitioning with the velocity vector of complexes directed towards the capillary outlet at “-” and with a counter-directed velocity vector of nonbinders. Adopted from Ref. 36 with permission of John Wiley and Sons.

negative charges and characterized by a single value of negative electrophoretic mobility μ_{TB} . The vector μ_{TB} is co-directed with μ_N and opposite in its direction to E . Furthermore, the hydrodynamic size of the protein–oligonucleotide complex is always greater than that of a lone oligonucleotide. Thus, target-binder complexes experience a greater drag force in electrophoresis than nonbinders. As a result, the magnitude of electrophoretic mobility of target-binder complexes is lower than that of nonbinders: $|\mu_{TB}| < |\mu_N|$.

Two known modes of CE-based partitioning differ by polarity, but in both of them nonbinders and target-binder complexes move in the same direction (Fig. 3A,B) [31, 51]. Partitioning of target-binder complexes from nonbinders is realized by setting up a correct cutoff time. This way of partitioning is in essence a nonideal filter. We hypothesize that k_N in CE-based partitioning could be drastically decreased if the target-binder complexes and nonbinders moved in the opposite directions (Fig. 3C). In other words, CE would function as a real filter giving it a name of IFCE. Hypothetically, in such a case, a small fraction of irregularly moving nonbinders that constitutes the nonbinder background in CE-based partitioning would be effectively prevented from exiting the capillary. Of course, it is still an open question how much (if any) decrease in the nonbinder background can be achieved with IFCE even if IFCE proves to be feasible (e.g., target-binder complexes and nonbinders move in the opposite directions).

We will consider a one-dimensional system of coordinates with axis x being co-directed with E . To develop IFCE, we chose CE with positive velocity (v_{TB} is co-directed with E and axis x) of the target-binder complexes, $v_{TB} > 0$, and negative velocity (v_N is opposite in direction to E and axis x) of nonbinders, $v_N < 0$ (Fig. 3C). These velocities depend on E as well as on μ_{TB} , μ_N , and the mobility vector of EOF, μ_{EOF} .

These dependences are scalar products of \mathbf{E} and sums of the corresponding mobility vectors:

$$\begin{aligned} v_{TB} &= \mathbf{E} \cdot (\boldsymbol{\mu}_{\text{EOF}} + \boldsymbol{\mu}_{\text{TB}}) \\ v_{\text{N}} &= \mathbf{E} \cdot (\boldsymbol{\mu}_{\text{EOF}} + \boldsymbol{\mu}_{\text{N}}) \end{aligned} \quad (9)$$

\mathbf{E} and $\boldsymbol{\mu}_{\text{EOF}}$ are co-directed except for very acidic pH values of BGE, which are not used in CE. For instance, typical conditions used for the partitioning mode shown in Fig. 3B are higher than physiological pH (approximately 7.4) and lower than physiological I_{BGE} (approximately 160 mM). These conditions result in $|\mu_{\text{EOF}}| > |\mu_{\text{N}}| > |\mu_{\text{TB}}|$ which, in turn, leads to undesirable $v_{\text{TB}} > v_{\text{N}} > 0$ (Fig. 3B) [31].

The required relation $v_{\text{TB}} > 0 > v_{\text{N}}$ can be achieved by decreasing $|\mu_{\text{EOF}}|$ to satisfy $|\mu_{\text{N}}| > |\mu_{\text{EOF}}| > |\mu_{\text{TB}}|$. Decreasing $|\mu_{\text{EOF}}|$ can, in turn, be made by decreasing the zeta potential of the negatively charged surface of the inner wall of a fused silica capillary via either coating the surface with a neutral layer or increasing I_{BGE} [55–57].

Coating the surface typically changes EOF abruptly from very strong to very weak resulting in $|\mu_{\text{N}}| > |\mu_{\text{TB}}| > |\mu_{\text{EOF}}|$ and, hence, $v_{\text{N}} < v_{\text{TB}} < 0$ (Fig. 3A) [55, 56]. In contrast, gradually increasing I_{BGE} can gradually change μ_{EOF} potentially reaching the required relation $|\mu_{\text{N}}| > |\mu_{\text{EOF}}| > |\mu_{\text{TB}}|$, but it also has a limitation. Increasing I_{BGE} leads to greater Joule heat generation and increased temperature inside the capillary, T_{cap} , that can lead to dissociation of target-binder complexes [58]. Timely, this limitation can be overcome with a recently introduced simplified universal method for predicting electrolyte temperatures (SUMET) which allows adjusting T_{cap} to a desirable value by rationally lowering the electric field strength, E [59]. Advantageously, increasing I_{BGE} makes BGE more physiological and can allow selection of binders intended for use in vivo (e.g., as detection probes, drugs, or drug-delivery vehicles). Therefore, we decided to explore varying I_{BGE} to test our main hypothesis that achieving $v_{\text{N}} < 0 < v_{\text{TB}}$ could reduce the nonbinder background in CE-based partitioning.

In order to solve the nonbinder background problem in the context of our hypothesis, we formulated five specific goals. Goal 1 was to find E that could guarantee a suitable

range of T_{cap} for I_{BGE} ranging from low values typically used in CE to high values approaching physiological. Goal 2 was to test whether a condition of IFCE ($v_{\text{TB}} > 0 > v_{\text{N}}$) could be achieved in CE by increasing I_{BGE} to near-physiological values while keeping elution time for the target-binder complexes under 1 h. Goal 3 was to test if satisfying this condition could drastically reduce the nonbinder background (reduce k_{N} to the desired value of $\sim 10^{-9}$). Goal 4 was to confirm that $k_{\text{B}} \approx 1$. Goal 5 was to examine if the obtained $k_{\text{B}}/k_{\text{N}} \approx 1/k_{\text{N}}$ was sufficient to facilitate selection of binders from a naïve oligonucleotide library in one step of IFCE-based partitioning. The following section describes how the five goals were achieved.

3.3 Rational development of IFCE-based partitioning with the required efficiency 10^9

3.3.1 Electric field to guarantee no overheating

The value of I_{BGE} was changed by introducing NaCl in concentration ranging from 25 to 150 mM to BGE. The value of I_{BGE} for 50 mM Tris-HCl is 46 mM and every additional 25 mM NaCl adds 25 mM to the I_{BGE} value. Accordingly, the values of I_{BGE} in our experiments ranged from 46 to 196 mM. To place these values in the context of biological relevance, note that I_{BGE} of phosphate buffered saline is 162 mM. A SCIEX CE instrument utilized in our study allowed capillary thermostabilization via its contact with a liquid coolant stabilized at T_{coolant} . The central longest part of the capillary was washed by the coolant and had temperature $T_{\text{cooled}} > T_{\text{coolant}}$; while the two short flanking parts of the capillary were not in contact with the coolant and their temperature $T_{\text{non-cooled}}$ was higher than T_{cooled} . The goal of this part of the study was to find E which guaranteed $T_{\text{cooled}} \leq 20^\circ\text{C}$ and $T_{\text{non-cooled}} \leq 42^\circ\text{C}$ at $T_{\text{coolant}} = 15^\circ\text{C}$ (20 and 42°C flank the range of vital temperatures of human body).

Our SUMET method was utilized for determination of T_{cooled} and $T_{\text{non-cooled}}$ for different values of E . Since the highest

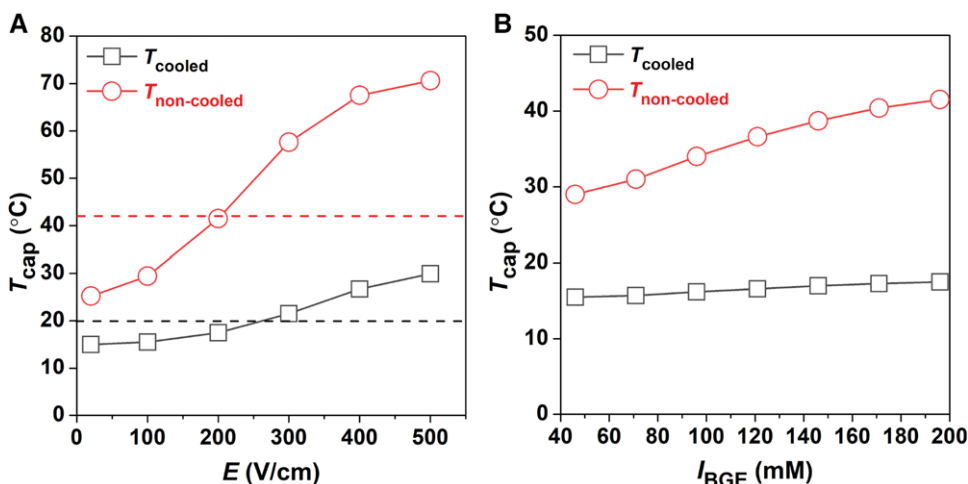


Figure 4. Determination of E that guarantees no overheating. Panel A shows the dependency of T_{cooled} and $T_{\text{non-cooled}}$ on E . The value of E satisfying both $T_{\text{cooled}} \leq 20^\circ\text{C}$ and $T_{\text{non-cooled}} \leq 42^\circ\text{C}$ was 200 V/cm. Panel B shows the values of T_{cooled} and $T_{\text{non-cooled}}$ for different I_{BGE} at $E = 200$ V/cm.

I_{BGE} corresponds to the greatest Joule heat, it was sufficient for us to determine E which satisfied these conditions for BGE containing 150 mM NaCl ($I_{\text{BGE}} = 196 \text{ mM}$); BGEs with lower I_{BGE} would then automatically satisfy the two conditions. Accordingly, the electric current was measured for BGE containing 150 mM NaCl during 1 min for each of six values of applied voltage and for six corresponding values of E : 20, 100, 200, 300, 400, and 500 V/cm. The collected current–voltage data and the SUMET program were used to determine the dependence of T_{cooled} and $T_{\text{non-cooled}}$ on E (Fig. 4A). The highest E satisfying both $T_{\text{cooled}} \leq 20^\circ\text{C}$ and $T_{\text{non-cooled}} \leq 42^\circ\text{C}$ was 200 V/cm; the sought temperatures were $T_{\text{cooled}} = 17.5^\circ\text{C}$ and $T_{\text{non-cooled}} = 41.5^\circ\text{C}$.

We then conducted current–voltage measurements at $E = 200 \text{ V/cm}$ for BGEs with lower ionic strengths and determined T_{cooled} and $T_{\text{non-cooled}}$ as functions of I_{BGE} (Fig. 4B). As expected all values of T_{cooled} were in a range of 15.0–17.5°C and all values of $T_{\text{non-cooled}}$ were higher than 17.5°C and lower than 41.5°C. Thus, an electric field of 200 V/cm guaranteed $T_{\text{cooled}} \leq 20^\circ\text{C}$ and $T_{\text{non-cooled}} \leq 42^\circ\text{C}$ and was used in the rest of this study.

To avoid sample exposure to higher temperatures in the inlet noncooled part of the capillary, the sample was propagated through this part by pressure at $E = 0$, and the voltage was applied only when the sample was 2 cm inside the cooled part of the capillary. The intact complexes were still exposed to the elevated $T_{\text{non-cooled}} \leq 41.5^\circ\text{C}$ while passing through the noncooled part at the capillary outlet.

3.3.2 Achieving $v_{\text{TB}} > 0 > v_{\text{N}}$

In a set of experiments testing whether $v_{\text{TB}} > 0 > v_{\text{N}}$ could be achieved in CE by increasing I_{BGE} , the sample was an equilibrium mixture of MutS protein with its fluorescently labeled DNA aptamer selected in a previous study [38]. MutS protein, which does not absorb on the capillary surface, is a good model protein for studying the mobility of protein–binder complexes. The equilibrium mixture contained the protein–aptamer (target–binder) complex and an unbound aptamer (nonbinder). Bodipy, a fluorescent dye with a zero net charge, was added to the equilibrium mixture to serve as a marker of EOF. Figure 5 shows the dependence of μ_{N} , μ_{TB} , and μ_{EOF} on $\sqrt{I_{\text{BGE}}}$. The choice of dimensions for the abscissa was dictated by a known near-linear dependence of μ_{EOF} on $\sqrt{I_{\text{BGE}}}$ [60]. The dependencies of μ_{N} , μ_{TB} , and μ_{EOF} are well approximated with a direct line in such coordinates; the extrapolations of this line for high I_{BGE} were used to define the right boundary of the I_{BGE} range of IFCE conditions. The values of μ_{EOF} , μ_{N} , and μ_{TB} were simply extrapolated to higher I_{BGE} ; this allowed us to avoid experimenting under conditions of extremely intense heat generation at high ionic strengths and a constant value of $E = 200 \text{ V/cm}$. High I_{BGE} values require significantly lower values of E to minimize Joule heating; the low values of E would result in impractically long migration times for N, TB, and EOF marker.

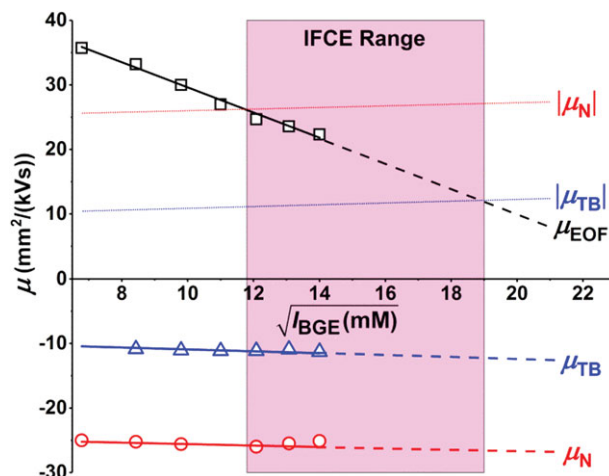


Figure 5. Dependence of μ_{N} , μ_{TB} , and μ_{EOF} on I_{BGE} . The values of I_{BGE} at which μ_{EOF} intersects $|\mu_{\text{N}}|$ and $|\mu_{\text{TB}}|$ defined the range of I_{BGE} in which ideal-filter capillary electrophoresis (IFCE) conditions are satisfied.

As expected, μ_{N} and μ_{TB} were both negative and changed slightly with changing I_{BGE} . In contrast, μ_{EOF} was positive and changed significantly with changing I_{BGE} . The relation $|\mu_{\text{N}}| > |\mu_{\text{EOF}}| > |\mu_{\text{TB}}|$, which is required for IFCE, was satisfied in the range of $140 \text{ mM} \leq I_{\text{BGE}} \leq 360 \text{ mM}$. Values of I_{BGE} greater than 140 mM and smaller than 360 mM resulted in $|\mu_{\text{EOF}}| > |\mu_{\text{N}}| > |\mu_{\text{TB}}|$ and $|\mu_{\text{N}}| > |\mu_{\text{TB}}| > |\mu_{\text{EOF}}|$, respectively, which correspond to CE-partitioning methods shown in Fig. 3A and B. The results of the described study led us to the conclusion that IFCE is feasible.

3.3.3 Lowering nonbinder background via increasing I_{BGE}

After confirming that IFCE was feasible at higher I_{BGE} values, we studied how I_{BGE} affected the nonbinder background which was quantitatively expressed by the value of k_{N} . The sample was the DNA aptamer of MutS without any target; the aptamer, thus, played a role of nonbinder. Two-minute fractions were collected and DNA amounts in them were determined *via* qPCR. Such experiments were carried out with different I_{BGE} , and the results were first presented as electropherograms with a decimal logarithm of number of DNA molecules in a corresponding fraction on the y axis. Then, the integral under each curve was calculated within a time window that corresponded to elution of the MutS–aptamer (target–binder) complex. The values of k_{N} were calculated as ratios between these integrals and the total quantity of the DNA aptamer sampled in CE ($47 \text{ nL} \times 10 \mu\text{M} \times 6.0 \times 10^{23} = 2.8 \times 10^{11}$). The dependence of $\log(k_{\text{N}})$ on I_{BGE} is shown in Fig. 6. The value of k_{N} reached the limit of detection of qPCR at $I_{\text{BGE}} = 146 \text{ mM}$; therefore, k_{N} values for $I_{\text{BGE}} > 146 \text{ mM}$ could not be determined accurately without increasing the quantity of sampled DNA. We did not want to sample a sample of DNA with higher than $10 \mu\text{M}$

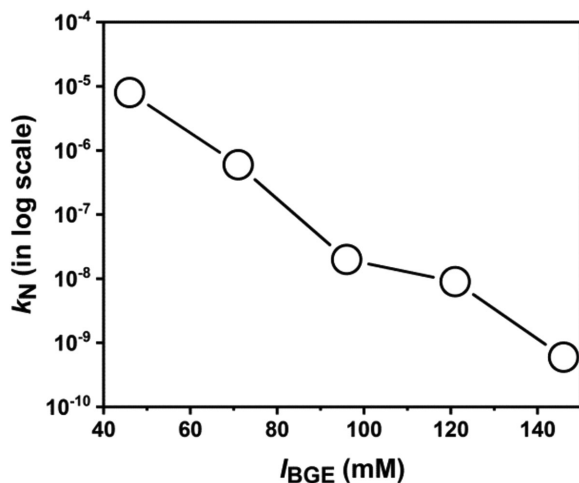


Figure 6. The dependence of $\log(k_N)$ on I_{BGE} .

concentration to avoid contamination of the CE instrument with highly concentrated DNA. The obtained results, however, were perfectly conclusive. The nonbinder background decreased 10 000 times when I_{BGE} grew from 46 to 146 nM. It likely decreased further with increasing I_{BGE} above 146 mM, but the value of $k_N \approx 10^{-9}$ achieved at $I_{BGE} = 146$ mM was sufficient for one-step selection according to our estimate described in Section 3.1.

3.3.4 Efficiency of binder collection

In our theoretical consideration above, we assumed that $k_B \approx 1$, and, accordingly, the value of k_N was anticipated to predominantly define k_B/k_N . It should be noted, however, that k_B can be much lower than unity due to binder loss in partitioning. For instance, in solid-phase partitioning, the strongest binders can be lost due to the inability to soft-dissociate them from the surface-immobilized target. In CE-based partitioning, the binders can be lost due to an incorrectly determined binder-collection time window. In partitioning by IFCE, k_B must be equal to unity for a target that does not adsorb onto the inner capillary wall unless the binder-collection time window is chosen incorrectly.

In this work, we validated all our procedures quantitatively and found k_B experimentally as $k_B = B_{out}/B_{in}$, where B_{in} is the number of target-binder complexes sampled and B_{out} is the number of binders collected in the binder-collection time window corresponding to the elution time window of the target-binder complexes. We first used NaCl-free BGE ($I_{BGE} = 46$ mM) in which the target-binder complexes and nonbinders migrate in the same direction (Fig. 3B) allowing for very accurate determination of B_{in} by using fluorescence detection (Fig. 7A) [31, 41]. A known volume of the equilibrium mixture containing MutS and its DNA aptamer was sampled for a CE run, in which both fluorescence and qPCR detections were used leading to two electropherograms for each CE run (Fig. 7A,B).

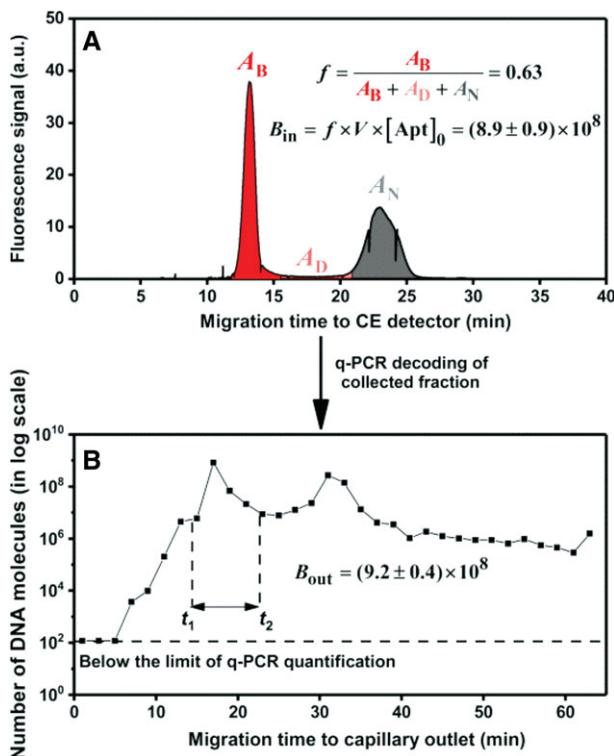


Figure 7. Determination of B_{in} (A) and B_{out} (B) for ideal-filter capillary electrophoresis (IFCE)-based partitioning of MutS-binder complexes and nonbinders. A known volume, $V = 47$ nL, of the equilibrium mixture containing 200 nM MutS and 50 nM fluorescently labeled aptamer ($[Apt]_0 = 50$ nM) was injected and CE was carried out with an electric field of 200 V/cm with a separation distance of 34 cm. Fluorescence detection was used to obtain an ordinary electropherogram shown in Panel A. Two-min fractions were collected and the number of aptamer copies was determined in each fraction by qPCR to produce an electropherogram in Panel B. A relative amount of MutS-bound aptamer (f) was found from the shaded areas in Panel A, and B_{in} was determined with formulas shown within the panel. B_{out} was calculated as an integral under the curve in Panel B within the binder-collection time window $t_1 - t_2$. Adopted from Ref. 36 with permission of John Wiley and Sons.

The value of $B_{in} = (8.9 \pm 0.9) \times 10^8$ was found from an electropherogram with fluorescence detection shown in Fig. 7A as the multiplication product of (i) the total concentration of aptamer in the equilibrium mixture, (ii) the volume of the sampled equilibrium mixture, and (iii) a relative amount of aptamer that was bound to MutS in the equilibrium mixture (relative to the total sampled amount of aptamer). We made no correction for complex dissociation during CE as less than 1% of the MutS-aptamer complex dissociated during complex migration to the capillary exit.

The value of $B_{out} = (9.2 \pm 0.4) \times 10^8$ was then determined from the electropherogram with qPCR detection shown in Fig. 7B by calculating an integral under the DNA curve within the binder-collection time window of 15–23 min. Knowing B_{out} and B_{in} allowed us to calculate $k_B = B_{out}/B_{in} = 1.0 \pm 0.1$.

We then conducted an IFCE experiment (with $I_{\text{BGE}} = 146 \text{ mM}$) and found $B_{\text{out}} = (7.3 \pm 0.5) \times 10^8$, which is slightly lower than the value of $(9.2 \pm 0.4) \times 10^8$ obtained for NaCl-free BGE. The difference was presumably due to additional ions' in IFCE affecting the strength of ionic bonds in protein–DNA complexes. These experiments demonstrated that by choosing a proper binder-collection time window we are able to satisfy our assumption of $k_{\text{B}} \approx 1$.

It should be noted that if IFCE can facilitate selection of binders in a single step of partitioning, the task of determining the binder-collection time window will be greatly simplified. Consecutive fractions eluting from IFCE can be collected and analyzed by qPCR for the presence of oligonucleotides. If the amount of oligonucleotides in a collected fraction exceeds the background measured in a control experiment, the fraction contains binders. Multiple fractions containing binders can be combined for further preparative PCR or another type of analysis.

3.3.5 One-step aptamer selection

To test if the achieved value of $k_{\text{B}}/k_{\text{N}} > 1.7 \times 10^9$ in partitioning by IFCE could facilitate one-step selection of aptamers we conducted partitioning of MutS binders from a naïve DNA library (the library that was not biased towards binding any specific target). Forty-seven nanoliters of the equilibrium mixture of the library ($B_{\text{in}} + N_{\text{in}} = 2.8 \times 10^{11}$) and 100 nM MutS was subjected to IFCE (with $I_{\text{BGE}} = 146 \text{ mM}$). Two-minute fractions were collected and analyzed by qPCR to build an electropherogram; the control experiment was similar, but the Target in the equilibrium mixture was replaced by BGE (Fig. 8). The collected fractions were analyzed within the wide binder-collection time window of 13–41 min. The numbers of binders and nonbinders at the output of selection were calculated as integrals under the corresponding curves within the binder-collection time window in IFCE-based partitioning and the control experiment, respectively: $B_{\text{out}} \approx 1.0 \times 10^6$ and $N_{\text{out}} \approx 4 \times 10^2$.

Figure 8 clearly shows a peak of target-binder complexes in IFCE which is absent in the control “no-target” experiment. We calculated that the enriched library had binder abundance of $B_{\text{out}}/N_{\text{out}} = (2.9 \times 10^6)/(1.1 \times 10^3) \approx 2.6 \times 10^3$. Thus, $B_{\text{out}}/N_{\text{out}}$ is much greater than $Q = 100$, which confirms that the selection was completed according to the criterion of selection completion defined above: $B_{\text{out}}/N_{\text{out}} > Q$. The knowledge of the amount of the sampled library $B_{\text{in}} + N_{\text{in}} = 2.8 \times 10^{11}$ (which can be assumed to be equal to N_{in}) and the amount of binders at the output of partitioning $B_{\text{out}} \approx 2.9 \times 10^6$ (which can be used to define B_{in}) uniquely allows estimation of the initial binder abundance, $B_{\text{in}}/N_{\text{in}}$ to be approximately $(2.9 \times 10^6)/(2.8 \times 10^{11}) \approx 1.0 \times 10^{-5}$. In other words, approximately 1 of every hundred thousand molecules of the naïve library were bound to MutS in the equilibrium mixture containing 100 nM MutS and stayed bound for the duration of the IFCE partitioning run. It is worthwhile noting that these results prove that IFCE can uniquely facilitate studies

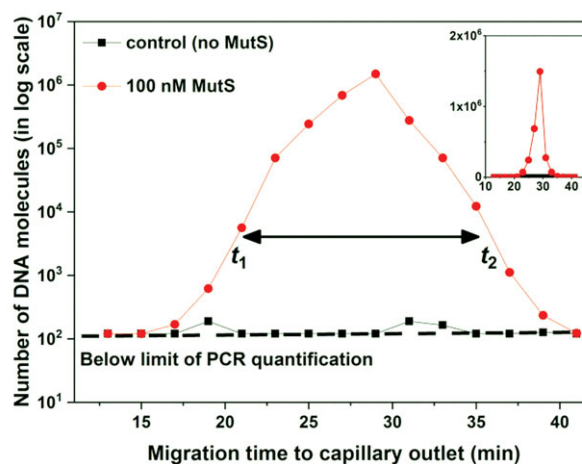


Figure 8. Determination of $B_{\text{out}}/N_{\text{out}}$ for ideal-filter capillary electrophoresis (IFCE)-based partitioning of MutS binder and the unbound library. A known volume, $V = 47 \text{ nL}$, of the equilibrium mixture containing $10 \mu\text{M}$ random-sequence DNA library and 100 nM MutS protein was injected and IFCE was carried out with an electric field of 200 V/cm with a separation distance of 34 cm. The control experiment was similar but without MutS. B_{out} and N_{out} were calculated as integrals under the curve within the target-binder complex collection time window, $t_1 - t_2$, in IFCE and control experiment, respectively. The inset shows the same data but with a linear ordinate. Adopted from Ref. 36 with permission of John Wiley and Sons.

which are needed to understand how $B_{\text{in}}/N_{\text{in}}$ changes with changing natures of the library and the target as well as their concentrations, incubation times, etc.

Finally, we amplified a fraction containing the highest amount of complexes (a fraction that eluted at 29 min) by PCR using a fluorescently labeled primer. Alternatively, multiple fractions within a 15–40 min time window could be combined and amplified to increase the amount of DNA material, but we decided to amplify a single fraction for simplicity. After amplifying DNA in this fraction, we performed a classical CE-based binding test with fluorescence detection [40]. This test revealed an apparent equilibrium dissociation constant of the enriched library of $K_{\text{d,app}} \approx 40 \text{ nM}$ and confirmed successful selection of a high-affinity aptamer pool in a single step of IFCE partitioning. For comparison, selecting a pool with similar $K_{\text{d,app}}$ by classical CE-based partitioning required three rounds of SELEX [38]. This successful selection, in turn, confirmed that our estimate of the required value for k_{N} of 10^{-9} for one-step selection was correct. Cumulatively, this study confirmed that our approach of rationally designing a partitioning method for one-step selection of binders from oligonucleotide libraries does work.

4 Concluding remarks

In this study, we proved that the condition of IFCE, i.e., the movement of target-binder complexes and nonbinders in the opposite directions inside the capillary, could be achieved

by simply raising I_{BGE} to approximately 150 mM at pH 7.0. We demonstrated that IFCE allows reaching uniquely low nonbinder background of $k_N < 6 \times 10^{-10}$. This value is 10^4 times lower than the lowest previously reported value of k_N . Furthermore, it is 10^7 times lower than the best values achievable with practical solid-phase partitioning methods used in binder selection from oligonucleotide libraries. Importantly, such extremely low k_N can be reached without sacrificing the efficiency of collection of binders: k_B can be as high as unity. We demonstrated that the resulting k_B/k_N of approximately 10^9 was sufficient for selection of a potent aptamer pool for MutS protein in a single step of partitioning. For comparison, the original selection of aptamers for this target required three rounds of CE-based partitioning/PCR amplification [38]. Finally, we showed how IFCE-based one-step partitioning can be used for estimating the abundance of binders in the naïve library. Such information is virtually impossible to obtain in a multiround selection procedure, such as SELEX, due to exponentially accumulating errors.

We expect that partitioning by IFCE will change the field of binder selection from oligonucleotide libraries. The simple procedure with collection of multiple fractions and analyzing the electropherogram described in Section 3.3.5 promises to become a basis for robust one-step selection of oligonucleotide aptamers by IFCE-based partitioning. Importantly, this procedure does not require a priori knowledge of the migration time of the target-binder complexes. An additional advantage is that the very high efficiency of partitioning can be achieved by simply increasing the ionic strength of BGE to the physiological value of 146 mM, which makes IFCE applicable to selection of binders intended for use in vivo. A side-bonus of the increased ionic strength is the anticipated decrease in nonspecific binding of the protein–target complex onto the capillary wall, which was a serious limitation of CE-based partitioning at low I_{BGE} . If sufficient pressure is applied to the capillary inlet in IFCE, then nonbinders also reach the capillary outlet and can be quantitated. Quantitatively detecting binders and nonbinders in a single run of such pressure-assisted IFCE facilitates measurements of K_d and k_{off} of the target-binder complex [40]. Thus, IFCE constitutes a “Swiss Army Knife” for both highly efficient selection of binders and determining their binding parameters. It should be noted that the accuracy of such pressure-assisted measurements are likely affected by the long pressure driven propagation which leads to significant peak dispersion. Accurately finding K_d and k_{off} in IFCE would require the development of an analytical scheme different from simple pressure-assisted measurements used in this work.

While the proof of IFCE was done here with a random-sequence DNA library, we foresee that IFCE will be directly applicable to selection of binders from other anionic libraries with a uniform charge. For example, IFCE should greatly benefit selection of binders from DNA-encoded libraries to which SELEX is not applicable due to inability to PCR-amplify such libraries [11, 22–24, 61]. Finally, one-step selection of binders effectively excludes quantitative bias of PCR and represents a quantitative tool to study binding properties of

various libraries to various targets. We expect that IFCE will allow “apple-to-apple” comparison of binder selection from different types of libraries, e.g., aptamers from libraries of modified DNA and RNA [13, 62]. To conclude, we foresee multiple applications of IFCE and expect that the use of this method by others will bring applications that we cannot anticipate at this time. We look forward to seeing the adaptation and use of IFCE by a diverse “in vitro selection” research community.

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

The authors have declared no conflict of interest.

5 References

- [1] Brenner, S., Lerner, R. A., *Proc. Natl. Acad. Sci. U. S. A.* 1992, *89*, 5381–5383.
- [2] Melkko, S., Scheuermann, J., Dumelin, C. E., Neri, D., *Nat. Biotechnol.* 2004, *22*, 568–574.
- [3] Bleicher, K. H., Bohm, H. J., Muller, K., Alanine, A. I., *Nat. Rev. Drug Discovery* 2003, *2*, 369–378.
- [4] Li, J., Murray, C. W., Waszkowycz, B., Young, S. C., *Drug Discovery Today* 1998, *3*, 105–112.
- [5] Plunkett, M. J., Ellman, J. A., *Sci. Am.* 1997, *276*, 68–73.
- [6] Hughes, J. P., Rees, S., Kalindjian, S. B., Philpott, K. L., *Br. J. Pharmacol.* 2011, *162*, 1239–1249.
- [7] Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K., Liu, D. R., *J. Am. Chem. Soc.* 2010, *132*, 11779–11791.
- [8] Vant-Hull, B., Gold, L., Zichi, D. A., *Curr. Protoc. Nucleic Acid Chem.* 2000, *00*, 9.1.1–9.1.16.
- [9] Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., Toole, J. J., *Nature* 1992, *355*, 564–566.
- [10] Tuerk, C., Gold, L., *Science* 1990, *249*, 505–510.
- [11] Clark, M. A., Acharya, R. A., Arico-Muendel, C. C., Belyanskaya, S. L., Benjamin, D. R., Carlson, N. R., Centrella, P. A., Chiu, C. H., Creaser, S. P., Cuzzo, J. W., Davie, C. P., Ding, Y., Franklin, G. J., Franzen, K. D., Gefter, M. L., Hale, S. P., Hansen, N. J. V., Israel, D. I., Jiang, J. W., Kavarana, M. J., Kelley, M. S., Kollmann, C. S., Li, F., Lind, K., Mataruse, S., Medeiros, P. F., Messer, J. A., Myers, P., O’Keefe, H., Oliff, M. C., Rise, C. E., Satz, A. L., Skinner, S. R., Svendsen, J. L., Tang, L. J., van Vloten, K., Wagner, R. W., Yao, G., Zhao, B. G., Morgan, B. A., *Nat. Chem. Biol.* 2009, *5*, 772–772.
- [12] Ecker, D. J., Vickers, T. A., Hanecak, R., Driver, V., Anderson, K., *Nucleic. Acids. Res.* 1993, *21*, 1853–1856.
- [13] Kraemer, S., Vaught, J. D., Bock, C., Gold, L., Katilius, E., Keeney, T. R., Kim, N., Saccomano, N. A., Wilcox, S. K., Zichi, D., Sanders, G. M., *PLoS One* 2011, *6*, e26332–e26332.
- [14] Irvine, D., Tuerk, C., Gold, L., *J. Mol. Biol.* 1991, *222*, 739–761.
- [15] Wang, J., Rudzinski, J. F., Gong, Q., Soh, H. T., Atzberger, P. J., *PLoS One* 2012, *7*, e43940.
- [16] Papoulas, O., *Curr. Protoc. Mol. Biol.* 1996, *36*, 12.18.11–12.18.19.

- [17] Ciesiolka, J., Illangasekare, M., Majerfeld, I., Nickles, T., Welch, M., Yarus, M., Zinnen, S., *Methods Enzymol.* 1996, *267*, 315–335.
- [18] Darmostuk, M., Rimpelova, S., Gbelcova, H., Ruml, T., *Biotechnol. Adv.* 2015, *33*, 1141–1161.
- [19] Mutter, G. L., Boynton, K. A., *Nucleic. Acids. Res.* 1995, *23*, 1411–1418.
- [20] Warnecke, P. M., Stirzaker, C., Melki, J. R., Millar, D. S., Paul, C. L., Clark, S. J., *Nucleic. Acids. Res.* 1997, *25*, 4422–4426.
- [21] Takahashi, M., Wu, X. W., Ho, M., Chomchan, P., Rossi, J. J., Burnett, J. C., Zhou, J. H., *Sci. Rep.* 2016, *6*, 33697.
- [22] Goodnow, R. A., Dumelin, C. E., Keefe, A. D., *Nat. Rev. Drug Discovery* 2017, *16*, 131–147.
- [23] Chan, A. I., McGregor, L. M., Liu, D. R., *Curr. Opin. Chem. Biol.* 2015, *26*, 55–61.
- [24] Buller, F., Steiner, M., Frey, K., Mircsof, D., Scheuermann, J., Kalisch, M., Buhlmann, P., Supuran, C. T., Neri, D., *ACS Chem. Biol.* 2011, *6*, 336–344.
- [25] Liu, Y. M., Wang, C., Li, F., Shen, S. W., Tyrrell, D. L. J., Le, X. C., Li, X. F., *Anal. Chem.* 2012, *84*, 7603–7606.
- [26] Imashimizu, M., Takahashi, M., Amano, R., Nakamura, Y., *Biol. Methods Protoc.* 2018, *3*, bpy004.
- [27] Lauridsen, L. H., Shamaileh, H. A., Edwards, S. L., Taran, E., Veedu, R. N., *PLoS One* 2012, *7*, e41702.
- [28] Peng, L., Stephens, B. J., Bonin, K., Cubicciotti, R., Guthold, M., *Microsc. Res. Tech.* 2007, *70*, 372–381.
- [29] Nitsche, A., Kurth, A., Dunkhorst, A., Panke, O., Sielaff, H., Junge, W., Muth, D., Scheller, F., Stocklein, W., Dahmen, C., Pauli, G., Kage, A., *BMC Biotechnol.* 2007, *7*, 48.
- [30] Lou, X., Qian, J., Xiao, Y., Viel, L., Gerdon, A. E., Lagally, E. T., Atzberger, P., Tarasow, T. M., Heeger, A. J., Soh, H. T., *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 2989–2994.
- [31] Berezovski, M., Drabovich, A., Krylova, S. M., Musheev, M., Okhonin, V., Petrov, A., Krylov, S. N., *J. Am. Chem. Soc.* 2005, *127*, 3165–3171.
- [32] Musheev, M. U., Kanoatov, M., Krylov, S. N., *J. Am. Chem. Soc.* 2013, *135*, 8041–8046.
- [33] Chen, H., Gold, L., *Biochemistry* 1994, *33*, 8746–8756.
- [34] Nieuwlandt, D., Wecker, M., Gold, L., *Biochemistry* 1995, *34*, 5651–5659.
- [35] Tian, Y., Adya, N., Wagner, S., Giam, C. Z., Green, M. R., Ellington, A. D., *RNA* 1995, *1*, 317–326.
- [36] Le, A. T. H., Krylova, S. M., Kanoatov, M., Desai, S., Krylov, S. N., *Angew. Chem. Int. Ed.* 2019, *58*, 2739–2743.
- [37] Beloborodov, S. S., Bao, J., Krylova, S. M., Shala-Lawrence, A., Johnson, P. E., Krylov, S. N., *J. Chromatogr. B* 2018, *1073*, 201–206.
- [38] Drabovich, A. P., Berezovski, M., Okhonin, V., Krylov, S. N., *Anal. Chem.* 2006, *78*, 3171–3178.
- [39] Galievsky, V. A., Stasheuski, A. S., Krylov, S. N., *Anal. Chem.* 2017, *89*, 11122–11128.
- [40] Kanoatov, M., Krylov, S. N., *Anal. Chem.* 2016, *88*, 7421–7428.
- [41] Berezovski, M., Musheev, M., Drabovich, A., Krylov, S. N., *J. Am. Chem. Soc.* 2006, *128*, 1410–1411.
- [42] Ellington, A. D., Szostak, J. W., *Nature* 1990, *346*, 818–822.
- [43] Quang, N. N., Miodek, A., Cibiel, A., Duconge, F., *Methods Mol. Biol.* 2017, *1575*, 253–272.
- [44] Navani, N. K., Mok, W. K., Yingfu, L., *Methods Mol. Biol.* 2009, *504*, 399–415.
- [45] Stanlis, K. K. H., McIntosh, J. R., *J. Histochem. Cytochem.* 2003, *51*, 797–808.
- [46] Yang, K.-A., Pei, R., Stojanovic, M. N., *Methods* 2016, *106*, 58–65.
- [47] Oh, S. S., Qian, J., Lou, X., Zhang, Y., Xiao, Y., Soh, H. T., *Anal. Chem.* 2009, *81*, 2832–2839.
- [48] Platt, M., Rowe, W., Wedge, D. C., Kell, D. B., Knowles, J., Day, P. J., *Anal. Biochem.* 2009, *390*, 203–205.
- [49] Morris, K. N., Jensen, K. B., Julin, C. M., Weil, M., Gold, L., *Proc. Natl. Acad. Sci. U. S. A.* 1998, *95*, 2902–2907.
- [50] Mendonsa, S. D., Bowser, M. T., *J. Am. Chem. Soc.* 2004, *126*, 20–21.
- [51] Mendonsa, S. D., Bowser, M. T., *Anal. Chem.* 2004, *76*, 5387–5392.
- [52] Mosing, R. K., Mendonsa, S. D., Bowser, M. T., *Anal. Chem.* 2005, *77*, 6107–6112.
- [53] Krylova, S. M., Karkhanina, A. A., Musheev, M. U., Bagg, E. A. L., Schofield, C. J., Krylov, S. N., *Anal. Biochem.* 2011, *414*, 261–265.
- [54] Kochmann, S., Le, A. T. H., Hili, R., Krylov, S. N., *Electrophoresis* 2018, *39*, 2991–2996.
- [55] Lipponen, K., Tahka, S., Kostianen, M., Riekkola, M. L., *Electrophoresis* 2014, *35*, 1106–1113.
- [56] Melanson, J. E., Baryla, N. E., Lucy, C. A., *TrAC Trends Anal. Chem.* 2001, *20*, 365–374.
- [57] Ding, W. L., Thornton, M. J., Fritz, J. S., *Electrophoresis* 1998, *19*, 2133–2139.
- [58] Evenhuis, C. J., Musheev, M. U., Krylov, S. N., *Anal. Chem.* 2011, *83*, 1808–1814.
- [59] Patel, K. H., Evenhuis, C. J., Cherney, L. T., Krylov, S. N., *Electrophoresis* 2012, *33*, 1079–1085.
- [60] Geiser, L., Mirgaldi, M., Veuthey, J. L., *J. Chromatogr. A* 2005, *1068*, 75–81.
- [61] Bao, J. Y., Krylova, S. M., Cherney, L. T., Hale, R. L., Belyanskaya, S. L., Chiu, C. H., Shaginian, A., Arico-Muendel, C. C., Krylov, S. N., *Anal. Chem.* 2016, *88*, 5498–5506.
- [62] Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., Carter, J., Dalby, A. B., Eaton, B. E., Fitzwater, T., Flather, D., Forbes, A., Foreman, T., Fowler, C., Gawande, B., Goss, M., Gunn, M., Gupta, S., Halladay, D., Heil, J., Heilig, J., Hicke, B., Husar, G., Janjic, J., Jarvis, T., Jennings, S., Katilius, E., Keeney, T. R., Kim, N., Koch, T. H., Kraemer, S., Kroiss, L., Le, N., Levine, D., Lindsey, W., Lollo, B., Mayfield, W., Mehan, M., Mehler, R., Nelson, S. K., Nelson, M., Nieuwlandt, D., Nikrad, M., Ochsner, U., Ostroff, R. M., Otis, M., Parker, T., Pietrasiewicz, S., Resnicow, D. I., Rohloff, J., Sanders, G., Sattin, S., Schneider, D., Singer, B., Stanton, M., Sterkel, A., Stewart, A., Stratford, S., Vaught, J. D., Vrkljan, M., Walker, J. J., Watrobka, M., Waugh, S., Weiss, A., Wilcox, S. K., Wolfson, A., Wolk, S. K., Zhang, C., Zichi, D., *PLoS One* 2010, *5*, e15004.