analytical chemistry

Article

pubs.acs.org/ac

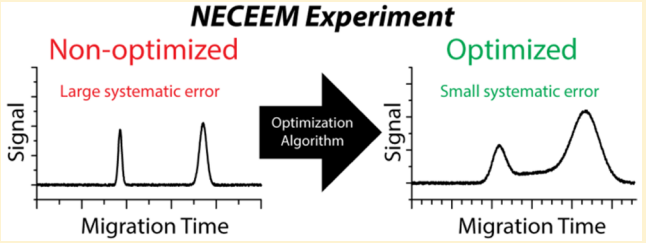
Systematic Approach to Optimization of Experimental Conditions in Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures

Mirzo Kanoatov, Sina Mehrabanfar, and Sergey N. Krylov*

Department of Chemistry and Centre for Research on Biomolecular Interactions, York University, Toronto, Ontario M3J 1P3, Canada

Supporting Information

ABSTRACT: Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) is an efficient method for studying intermolecular interactions. Optimization of NECEEM experiments is not a trivial task, due to the complex interrelation between numerous experimental parameters and their combined effects on the accuracy and precision of measurements. Here we present an "algorithmic" approach for NECEEM optimization, which eliminates all of the guesswork out of this process and allows researchers to approach it in a systematic manner. We have fully tested our approach using comprehensive in silico analysis and have showed its utility within a real experimental study. The new approach makes NECEEM more robust, resilient to errors, and easily approachable for researchers with varying experience in CE.



Studying equilibrium and kinetic properties of noncovalent biomolecular interactions is an imperative task in the fields of molecular biology, pharmacology, and medicine. Kinetic measurements are especially important,[1] as very few processes in living organisms occur at equilibrium (in fact, equilibrium is more characteristic of dead organisms). The list of widely applicable kinetic methods includes surface plasmon resonance (SPR),[2] biolayer interferometry (BLI),[3] stopped-flow approaches,[4] and kinetic capillary electrophoresis (KCE),[5] with each method offering its own set of advantages and inherent limitations. Regardless of the choice of an analytical technique, it is widely appreciated that proper design and optimization of kinetic experiments are essential for the accuracy of produced data.[6] Experimental conditions (e.g., analyte concentrations, time scale of measurement) must be chosen with comprehension and care in order to avoid introducing large systematic errors into the measurements. The need for optimization stems from the fact that molecular interactions underlying biological processes are highly diverse in their properties, with magnitudes of measured parameters varying over multiple orders: equilibrium dissociation constant (K_d) values of biomolecular interactions are known to range between 10^{-3} and 10^{-15} M,[7,8] rate constant of association (k_{on}) values between 10^3 and 10^9 M^{-1} s^{-1}, and rate constant of dissociation (k_{off}) values between 10^{-6} and 1 s^{-1}.[9] Experimental conditions must be selected such that the sensitive range of a given method matches the magnitudes of the measured parameters. Over the years, a variety of strategies and guidelines have been developed for design and optimization of SPR,[10] BLI,[11] and stopped-flow[12] experiments; however, prior to this work, no such strategies have been available for KCE methods.

In this work, we develop a systematic approach for optimization of experimental conditions of nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM), the most popular method within the "KCE toolbox". In NECEEM, the intermolecular complex between species of interest dissociates while being separated from its components by capillary electrophoresis (CE).[13] The resulting migration patterns can reveal K_d and k_{off}, allowing k_{on} to be calculated as $k_{on} = k_{off}/K_d$. When compared to other kinetic methods, NECEEM shows a remarkable balance between the accuracy of the produced results, sample consumption, and the required investment of time and effort. The accuracy of NECEEM stems from its homogeneous-phase nature: measurements are performed with analytes in free solution, without the use of sieving matrixes or surface immobilization.[14] This property gives NECEEM a major advantage over heterogeneous-phase kinetic methods like SPR and BLI, which depend on immobilization of one of the interacting components and are, thus, prone to bias due to disruption of molecular conformation, introduction of steric hindrance, and spatial exclusion.[15,16] Rapidity of NECEEM stems from the fact that a single experiment performed under suitable (optimized) conditions provides all of the required information to determine K_d and k_{off} with high accuracy.[17] While, in practice, several experiments are always performed for purposes of optimization and estimation of precision, NECEEM studies still involve a smaller number of experiments than studies with alternative methods, most of which necessitate creation of laborious isothermal binding curves. Furthermore, as a CE-based method, NECEEM facilitates exceptionally small sample

Received: July 27, 2016
Accepted: August 21, 2016
Published: August 21, 2016

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

consumption and is readily amenable for automation.[5] This combination of accuracy, processivity, and facility make NECEEM a powerful method for rapid evaluation of binder candidates for biomolecular targets, a crucial task in the drug lead screening stage of pharmacological research.[18]

Optimization of NECEEM experimental conditions is not a trivial task. The main challenge lies in the complex interrelation between experimental parameters that often produce counteracting effects. As an example, to perform NECEEM measurements in high-conductivity electrolytes (e.g., physiological buffers) the users are required to decrease the applied electric field strength to avoid excessive Joule heating; decreasing the electric field strength, in turn, results in longer analysis times; prolonged analysis may invalidate kinetic NECEEM measurements by rendering parts of electropherograms (e.g., the intermolecular complex peak) undetectable. Performing such adjustments properly often requires an effort of a well-trained specialist, with comprehensive understanding of the underlying physical and chemical phenomena. Without proper guidance, the optimization of NECEEM can become a prolonged task, significantly undermining the processivity of the method. What further aggravates the issue is that improperly optimized NECEEM experiments are often difficult to recognize, which degrades the resilience of the method against errors.

To alleviate this issue, a systematic approach for optimization of NECEEM experimental conditions is required. In this article, we describe such an approach, which takes the complex interrelation between NECEEM experimental conditions into account and deals with them in a stepwise manner. Our optimization "algorithm" allows the obtained results to be validated and provides objective recommendations on improving experimental conditions when the validation fails. We implement the developed algorithm in a form of a user-friendly software package which is able to automatically process experimental data. We demonstrate the practical utility of this "expert system" by applying it to a NECEEM experiment. Our developed optimization approach allows for the application of a wider range of experimental conditions in NECEEM, making this method more reliable, robust, and amenable for practical use by researchers with varying experience in CE.

MATERIALS AND METHODS

All chemicals and buffer components were purchased from Sigma-Aldrich (Oakville, ON, Canada) unless otherwise stated. Fused-silica capillaries were purchased from Molex (Phoenix, AZ). *Thermus aquaticus* MutS recombinant protein was expressed and purified as described previously.[19] The truncated version of the MutS aptamer (variant 5) was designed by Professor Philip E. Johnson (York University), based on a DNA aptamer previously selected (clone 2−06).[20] The aptamer variant was custom synthesized by Integrated DNA Technologies (Coralville, IA). The nucleotide sequence of the fluorescein-labeled, single-stranded DNA aptamer was 5′-/fluorescein/-GCC CGC CTC CTT CCT GGT AAA GTC ATT AAT AGG TGT GGG GTT TCG GAG ACG AGA TAG GCGG-3′.

NECEEM. All CE experiments were carried out using CESI-8000 instrument (Sciex, Concord, ON), equipped with a standard laser-induced fluorescence (LIF) detection system (488 nm excitation, 520 nm emission). Runs were performed in an uncoated fused-silica capillary, with inner radius of 10 μm and outer radius of 180 μm. The total length of the capillary was 30 cm, with the detection window placed 20 cm from the inlet.

Dilutions of all sample components were prepared with the electrophoresis run buffer: 50 mM Tris−HCl pH 7.4, 20 mM NaCl. Sample mixtures were incubated at room temperature for 20 min prior to analysis to achieve equilibration in the binding reaction. Prior to every run, the capillary was rinsed with the run buffer at 60.0 psi (414 kPa) for 3 min (to pump 10 capillary volumes). At the end of each run, the capillary was rinsed with a succession of 100 mM HCl, 100 mM NaOH, and deionized water, at the same pressure/time settings. The samples were injected into the capillary, prefilled with the run buffer, by a 1.0 psi (6.9 kPa) pressure for 16 s to yield a 5 mm long sample plug. Prior to applying voltage, the sample mixture was propagated 4 cm through the uncooled portion of the capillary by 2.0 psi (13.8 kPa) pressure for 1 min.

A total of four types of experiments were performed: three preliminary and one optimized. Each preliminary experiment was performed only once, while the optimized experiment was repeated 3 times. For all of the preliminary experiments, the sample mixture was injected from the inlet end of the capillary (20 cm distance to detector) and electrophoresis was carried out with the positive electrode at the inlet, at a constant voltage of 25.0 kV. The coolant temperature was set to 15 °C. The first preliminary experiment aimed at estimating the mobility (μ_L) and the response factor (γ_L) of the ligand and was performed without the addition of the target. Concentration of the aptamer used in this experiment was 10 nM. This experiment also allowed us to estimate the resistivity of the run buffer (ρ) and the average amplitude of the background noise (σ) in the obtained signal. The second preliminary experiment aimed at obtaining the mobility (μ_C) and the response factor (γ_C) of the complex and was performed at the highest possible concentration of the target to result in binding saturation. The concentration of the aptamer was 10 nM, and the concentration of the protein was 4.5 μM. The third preliminary experiment aimed at estimating the K_d and the k_{off} values and was performed according to the recommendations from section S9 in the Supporting Information. The concentration of the aptamer was 300 nM, and the concentration of the protein was 1.5 μM. The last of the preliminary experiments was subjected to validation and optimization using the further described algorithm, and the optimized set of experiments was performed using the algorithm-provided conditions. The sample mixture was injected from the outlet end of the capillary (10 cm distance to detector), and electrophoresis was carried out with the positive electrode at the outlet, at a constant voltage of 2.0 kV. The coolant temperature was set to 24 °C. In these experiments the concentration of the aptamer was 60 nM, while the concentration of the protein was 300 nM.

RESULTS AND DISCUSSION

NECEEM Concept. For a detailed conceptual description of the technique we refer to previous publications.[13,14] NECEEM is a method for studying noncovalent binding between a target (T) and a ligand (L), with the formation of an intermolecular complex (C), which is described by the following reaction equation:

$$\mathrm{T} + \mathrm{L} \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} \mathrm{C} \qquad (1)$$

In a NECEEM experiment, T and L are mixed together at initial concentrations $[T]_0$ and $[L]_0$ and incubated to achieve

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

equilibration of the binding reaction. The concentrations of components in the equilibrium mixture ($[T]_{eq}$, $[L]_{eq}$, and $[C]_{eq}$) are described by the following relationships:

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[T]_{eq}[L]_{eq}}{[C]_{eq}} \quad (2)$$

NECEEM analysis relies on CE to separate and quantitate the components of the equilibrium mixture. This is done by injecting a short zone of the mixture (with a width w) into a narrow bore capillary (with an inner radius r) prefilled with a background electrolyte (BGE) and subjecting the capillary to a uniform electric field (E). To avoid electrodispersive phenomena, the buffer used to prepare the sample mixture is matched with the BGE as closely as possible. The electric current (I) passing through the capillary causes generation of Joule heat, which often necessitates active cooling to maintain a desired in-capillary temperature. Most modern CE instruments are equipped with liquid or forced air cooling systems which are unable to cool the entire capillary efficiently: often stretches of the capillary at the inlet, outlet, and point of contact with the detector are not cooled due to engineering challenges.[21] Temperatures in the efficiently cooled (T_{ef}) and inefficiently cooled (T_{inef}) portions of the capillary can differ dramatically.[22] To avoid the detrimental effects of heating on the studied molecules and CE separation, the sample zone is pressure-propagated through the inefficiently cooled portion at the inlet prior to the application of the electric potential. This operation reduces the length of the capillary available for separation (l_{sep}), which is defined as

$$l_{sep} = l_{tot} - l_{out} - l_{prop} = l_{det} - l_{prop} \quad (3)$$

where l_{tot} is the total length of the capillary, l_{det} is the distance between capillary inlet and the detector, l_{out} is the distance between the detector and the capillary outlet (for instruments with on-column detection), and l_{prop} is the length of the initial pressure propagation zone (Figure 1A). When the electric field is established, molecules with different size-to-charge ratios migrate through the capillary with different constant velocities (v_T, v_L, and v_C) and separate from each other into distinct spatial zones. Mobilities of relevant analytes are defined as follows:

$$\mu_L = \frac{v_L}{E} \quad (4)$$

$$\mu_C = \frac{v_C}{E} \quad (5)$$

$$\Delta\mu = |\mu_L - \mu_C| \quad (6)$$

Analyte quantitation is achieved by any of the available detection methods, such as light absorption spectroscopy, LIF, electrochemical detection, or mass spectrometry. It is sufficient that only one of the components (L by convention) is detectable in both its free (unbound) and bound states. In this case, it is sufficient to achieve separation between the zones of L and C only. During CE separation, the equilibrium fractions of free T and L migrate as distinct zones, while C undergoes continuous dissociation as a result of disturbed equilibrium. Upon reaching the detector, L and C generate signal that is presented as a function of migration time to the detector, referred to as an electropherogram plot, and typically contains three distinct features: two peaks, which correspond to the zones of free L and intact C, and a smear-like region of C

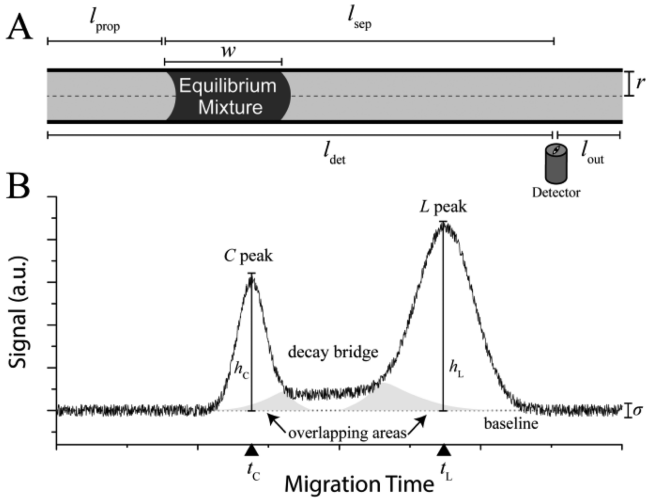


Figure 1. Schematic diagram of NECEEM initial conditions and resulting data. (A) Initial conditions in the capillary prior to application of the electric field. A plug of equilibrium mixture is injected into the capillary and propagated through the inefficiently cooled portion. (B) Example of a NECEEM electropherogram, which consists of three features: peaks of unbound (free) L and intact C, and the decay bridge. Overlapping areas between the peaks and the bridge necessitate the use of a deconvolution procedure for accurate measurement of feature areas.

dissociation products (decay bridge), which merges with both L and C peaks. The amplitude of the generated signal is defined by the local concentrations of the detected analytes and their response factors (γ_L and γ_C, respectively):

$$\gamma_L = \frac{h_L}{[L]_{eq}} \quad (7)$$

$$\gamma_C = \frac{h_C}{[C]_{intact}} \quad (8)$$

where h_L and h_C are the height of the free ligand and intact complex peaks, respectively, and $[C]_{intact}$ is the concentration of intact complex at the time of its detection, t_C, calculated as

$$[C]_{intact} = [C]_{eq} e^{-k_{off} t_C} \quad (9)$$

See section S2 in the Supporting Information on how to calculate the equilibrium concentrations of mixture components based on K_d. The areas under the three features of a NECEEM electropherogram (A_L, A_C, and A_D for peaks of L and C, and the decay bridge, respectively) are measured, with a deconvolution procedure applied to the overlapping areas between the features (Figure 1B),[23] and used to calculate the fraction of unbound ligand, R:

$$R = \frac{A_L}{A_L + (\gamma_L/\gamma_C)A_C + A_D} \quad (10)$$

K_d and k_{off} are then calculated as follows:

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

$$k_{off} = \frac{\ln\left(\frac{A_C + A_D}{A_C}\right)}{t_C} \quad (12)$$

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

The value of k_{on} is calculated using eq 2.

Parameters and Variables. NECEEM is a versatile tool that can be used for both analytical and preparative purposes. In analytical applications, NECEEM has been used for determination of unknown equilibrium and kinetic rate constants of intermolecular interactions,[24] as well as for affinity probe-based quantification of analytes.[25] In the preparative mode, NECEEM can be used to select affinity ligands from heterogeneous libraries (e.g., selection of aptamers).[26] To adapt NECEEM for the intended application (analytical or preparative), the users are free to decide on the values of four experimental parameters: (1) composition of the BGE, (2) intended temperature in the efficiently cooled portion of the capillary ($T_{ef,goal}$), (3) internal radius of the capillary (r), and (4) width of the sample injection zone (w). The choice of BGE composition, which includes the electrolyte's ionic strength, concentration of the buffering components, pH, presence of cofactors, and stabilizing agents, should imitate the environment in which the intermolecular interaction usually takes place. Similarly, the chosen in-capillary temperature should reflect either the natural conditions of the given interaction (e.g., core temperature of the host organism) or the conditions of intended use (e.g., room temperature for affinity probe-based detection kits). The choice of a capillary radius is dictated by the aims of the application, with smaller radii being preferential in analytical studies to reduce sample consumption and larger radii being preferential in preparative applications to increase processivity. Larger capillary radii may also be chosen to improve the system's limit of quantitation (LOQ), as wider cross sections enable higher analyte flux and, thus, yield a stronger signal. On the other hand, capillaries with smaller inner radii are better at dissipating heat and are, thus, more suitable for experiments that use high-conductivity run buffers. The width of the sample injection zone has an impact on the separation efficiency of CE, with narrower injection zones used when the resolution between the species is a priority. As the length of the injection zone defines the volume of the sample (along with the capillary radius), wider sample injection plugs are also used to improve method's processivity.

The user-selected values of the four experimental parameters, as well as the properties of the studied molecular interactions, will have an effect on the optimal values of three experimental variables: (1) initial analyte concentrations, $[L]_0$ and $[T]_0$, (2) length of capillary to detector, l_{det}, and (3) strength of the applied electric field, E. The purpose of the optimization procedure is to determine the combination of values of the three experimental variables which ensure accurate and bias-free NECEEM analysis.

The optimization is performed in a succession of refining rounds, where the apparent values of k_{off} and K_d (denoted by k_{off}* and K_d*, respectively) are measured and subjected to an experiment validation test. If the experiment validation fails, the apparent values are rejected, and the experiment is performed again, with the experimental conditions rationally adjusted to specifically address the shortcomings revealed by the validation test. Once the experiment validation passes, the apparent values are accepted as accurate. The optimization algorithm, thus, consists of two phases: validation of the experiment, followed by generation of recommendations for adjustment of conditions. A schematic overview of the algorithm is presented in Figure 2, while the detailed description is presented in the following sections.

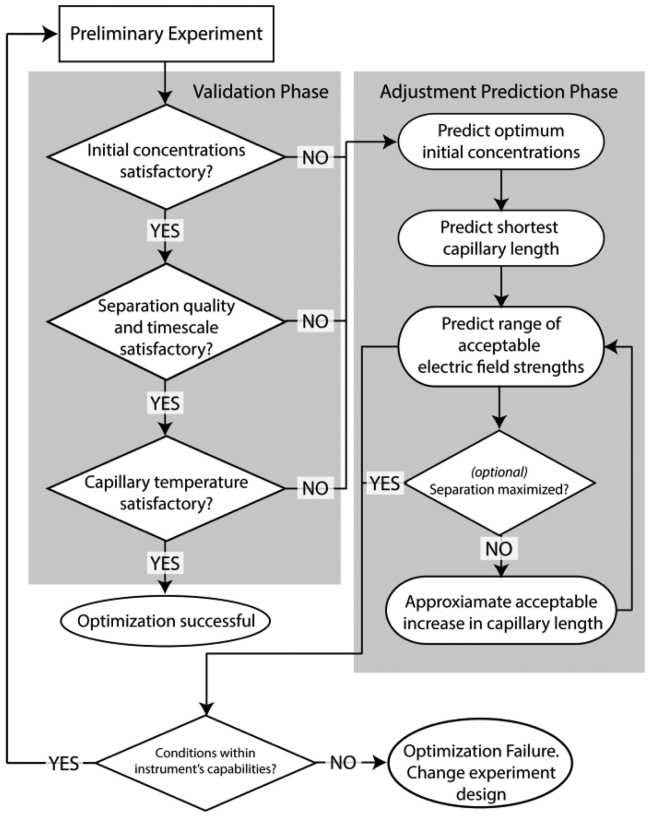


Figure 2. Schematic diagram of the optimization algorithm. Information from a preliminary experiment is first subjected to a validation test. If the validation fails, necessary adjustments to experimental variables are predicted based on values from the preliminary experiment. If the predicted values are within capabilities of the CE instrument, another experiment is performed using the newly defined conditions. This refinement of preliminary data continues until experiment validation succeeds.

Phase 1: Experiment Validation. Upon deciding on the values of the experimental parameters, and performing a preliminary experiment at a nonoptimized set of conditions (see section S9 in the Supporting Information for the best way to choose initial conditions), the users must validate the experiment based on the following criteria: (i) suitability of $[L]_0$ and $[T]_0$, (ii) suitability of separation quality and time scale, and (iii) suitability of in-capillary temperatures.

A small experimental error in measurement of K_d can become greatly amplified if $[T]_0$ and $[L]_0$ are improperly chosen. We have previously derived a relationship between variability in experimental measurements of R and errors in the calculated final K_d (see section S3 in the Supporting Information).[19] Analysis of this relationship shows that the most accurate determination of K_d requires the measurements to be performed at $[L]_0$ that is smaller than or equal to $2K_d$ (Figure S3A), the point at which concentrations of equilibrium components are equal (i.e., $[L]_{eq} = [T]_{eq} = [C]_{eq} = K_d$), and that $[T]_0$ is chosen so that $R = 0.5$ (Figure S3B), a concentration that can be calculated by rearranging eq 11. Thus, values of $[L]_0$ and $[T]_0$ can be considered as suitable if they satisfy the following conditions:

$$[L]_0 \leq 2K_d^* \tag{13}$$

$$(K_d^* + 0.5[L]_0) \leq [T]_0 \leq 2K_d^* \tag{14}$$

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

The extent and the rate of separation between L and C play an important role in our ability to accurately extract information from NECEEM electropherograms. A useful measure of separation efficiency in CE is the characteristic time of separation, t_{sep}, which is the time required for the initially superimposed zones of L and C to completely separate from each other:

$$t_{sep} = \frac{w}{\Delta\mu E} \tag{15}$$

By dividing the total analysis time, t_{run} (the time required for the slowest detectable component to reach the detector), by t_{sep} we can calculate the number of separation events, S, which took place during a CE experiment:

$$t_{run} = \frac{l_{sep}}{\min(\mu_L, \mu_C)E} \tag{16}$$

$$S = \frac{t_{run}}{t_{sep}} = \frac{\Delta\mu l_{sep}}{\min(\mu_L, \mu_C)w} \tag{17}$$

S is a parameter similar to chromatographic resolution, with the exception that it neglects zone broadening. Since zone broadening is difficult to predict in a generic way, S is a more amenable measure for a predictive optimization algorithm. For S to be an accurate predictor of separation quality, however, zone broadening effects must be minimized by matching the sample and run buffers (to avoid electrodispersive phenomena) and by avoiding excessively long experiments, in which the effects of longitudinal diffusion become non-negligible (usually longer than 5 h for macromolecules). Insufficient separation between L and C peaks will obscure the decay bridge; thus, a required minimum number of separation events must occur during an experiment to facilitate accurate extraction of A_L, A_C, and A_D. This minimum is defined by the capabilities of the area deconvolution procedure used for analysis of NECEEM electropherograms.[23] The best way to study the limits of the area deconvolution procedure is to apply it to a set of computer-simulated NECEEM electropherograms, extract the K_d and k_{off}, and compare them to the original values predefined in the simulations. We have previously developed an in silico NECEEM modeling tool in COMSOL Multiphysics environment, which can generate realistic and accurate simulated electropherograms.[19] To find the minimum S required for accurate area analysis, we have simulated a set of NECEEM experiments with a range of S values (by varying l_{sep}), but with otherwise constant and optimized conditions, and applied the area deconvolution approach to them. Figure S4 shows that accurate determination of both K_d and k_{off} requires a minimum value of S = 5. On the other hand, unnecessarily long separation, characterized by a large t_{run}, may lead to a scenario where the complex peak becomes too small for accurate quantitation, i.e., when its height becomes smaller than the LOQ of the instrument, commonly defined as 10 times the amplitude of the background noise, σ.[27,28] From eq 9, the time that it takes for the C peak to dissociate to this level, t_{dis}, can be calculated:

$$t_{dis} = \frac{\ln\left(\frac{[C]_{eq}\gamma_C}{10\sigma}\right)}{k_{off}} \tag{18}$$

Thus, to ensure proper quantitation of intact C, t_{run} may not exceed t_{dis}.

Lastly, the prominence of the decay bridge relative to the other features of an electropherogram influences the accuracy with which the overlapping areas can be deconvoluted. The relative size of the decay bridge is defined by τ, the ratio of two characteristic times: t_{sep}, and the characteristic time of equilibration, t_{eq}:

$$t_{eq} = \frac{1}{(k_{on}[T] + k_{off})} \tag{19}$$

$$\tau = \frac{t_{sep}}{t_{eq}} = \frac{w(k_{on}[T]_{eq} + k_{off})}{E\Delta\mu} \tag{20}$$

where $[T]$ is the local concentration of T at a given point in the capillary, which is usually substituted by its maximum value of $[T]_{eq}$. Small values of τ will lead to an undetectable decay bridge, which will prevent accurate determination of k_{off}, while large values will cause the decay bridge to obscure the peaks, which prevents accurate determination of K_d. Thus, a valid experiment must be characterized by an intermediate range of τ values, the boundaries of which can be determined by an in silico study. We have simulated a set of NECEEM electropherograms with a range of τ values (by varying k_{off} and k_{on}), but with otherwise optimized conditions, and used them to test the accuracy of both k_{off} and K_d measurements (Figure S5). By defining a relative systematic error threshold at 20%, we can establish the optimum range of τ values to lie between 0.15 and 0.6. It should be noted that all of the error thresholds described in this work may be redefined to satisfy specific requirements of a given study. Our recommended error thresholds only serve the purpose of making the algorithm applicable to a general case. For NECEEM separation conditions to be considered as suitable, they must satisfy the following conditions:

$$S \geq 5 \tag{21}$$

$$t_{run} \leq t_{dis} \tag{22}$$

$$0.15 \leq \tau \leq 0.6 \tag{23}$$

It should be noted that, for studies which aim at measuring only the value of K_d, and not k_{off}, a smaller τ is preferable so that the overlaps between the decay bridge and the peaks are minimal. For such studies, condition 23 may be substituted by the following:

$$\tau \leq 0.6 \tag{24}$$

Temperatures in both efficiently and inefficiently cooled portions of the capillary can be measured using the simplified universal method for determining electrolyte temperatures (SUMET).[29]

$$T_{ef} = T_{cool} + \frac{c(EI_{av})^{n+1}}{g + (EI_{av})^n} \tag{25}$$

$$T_{inef} = T_{amb} + ka(EI_{av})^2 \tag{26}$$

where T_{cool} is the temperature of the coolant, T_{amb} is the ambient temperature surrounding the inefficiently cooled portions of the capillary, I_{av} is the average electric current during electrophoresis, and a, c, g, k, and n are constants defined by r (see section S7 in the Supporting Information).

For the T_{ef} to be considered as valid it must be equal to the chosen value of $T_{ef,goal}$. As long as the sample mixture is propagated through the inefficiently cooled portion of the capillary, the value of T_{inef} does not have a significant effect on

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

accuracy of the experiment and is only limited by the boiling temperature of the electrolyte, T_{boil}. However, depending on the specifics of the study, the users may define a more conservative limit of T_{inef} based on some other considerations (e.g., maximum allowable change in BGE viscosity). Thus, in-capillary temperatures can be considered as suitable if they satisfy the following conditions:

$$T_{ef} = T_{ef.goal} \tag{27}$$

$$T_{inef} < T_{boil} \tag{28}$$

If an experiment satisfies conditions described in formulas 13, 14, 21−23, 27, and 28, it can be considered as valid and the values of K_d* and k_{off}* as accurate; otherwise, more suitable experimental conditions should be found, using the steps described below.

Phase 2: Recommendations for Experimental Condition Adjustment. To generate practicable recommendations for adjusting experimental conditions, the limitations of the employed CE instrument must be taken into account. Table 1 lists the information regarding the experimental system, which needs to be ascertained by the user.

Table 1. List of Information about the Limitations of the CE Instrument Required for Generation of Recommendations

parameter	description
$V_{instr.min}$	minimum electrophoresis voltage supplied by the instrument
$V_{instr.max}$	maximum electrophoresis voltage supplied by the instrument
$I_{instr.max}$	maximum electrophoresis current supplied by the instrument
$l_{det.instr.min}$	minimum capillary length to detector
$l_{det.instr.max}$	maximum capillary length to detector
$T_{cool.min}$	minimum coolant temperature
$t_{run.max}$	maximum analysis time[a,b]
σ	average amplitude of noise at baseline

[a]Usually limited by data-acquisition time or buffer depletion time. [b]The minimum time of analysis is subsumed by S and, thus, does not need to be considered separately.

In some cases, it may be impossible to generate recommendations which fall within the instrument's limitations. In these cases, the optimization algorithm aborts with a "failure" output and provides a suggestion message on how the overall design of the experiment must be modified to facilitate the measurement. Recommendations for adjusting experimental variables are generated in four steps: (i) prediction of optimum $[L]_0$ and $[T]_0$, (ii) prediction of minimum l_{det}, (iii) prediction of optimum range of values of E, and (iv) iterative maximization of S.

Step 1. The choice of $[L]_0$ and $[T]_0$ should aim to satisfy conditions 13 and 14. The recommended value of $[L]_0$ must be smaller than or equal to $2K_d$*, but this difference should not exceed 10-fold, as this would lead to an unnecessarily small $[C]_{eq}$. Often, the choice of $[L]_0$ is limited by the LOQ of the instrument. As the signal for L has to be split between three electropherogram features, $[L]_0$ must be larger than the concentration of free L that results in a peak with a height 100 times larger than the amplitude of background noise, σ (Figure S6).

$$[L]_0 = \begin{cases} 0.2K_d^* & \text{if } \dfrac{100\sigma}{\gamma_L} \le 0.2K_d^* \\ \dfrac{100\sigma}{\gamma_L} & \text{if } 0.2K_d^* \le \dfrac{100\sigma}{\gamma_L} \le 2K_d^* \end{cases} \tag{29}$$

Optimization of $[L]_0$ fails if $(100\sigma/\gamma_L) > 2K_d$*, providing a suggestion to improve the LOQ of the system (e.g., by increasing r). Once the value of $[L]_0$ is determined, the optimum value of $[T]_0$ can be calculated:

$$[T]_0 = K_d^* + 0.5[L]_0 \tag{30}$$

Step 2. The choice of l_{det} should aim to satisfy condition 21. The shortest l_{sep} that satisfies $S \ge 5$, $l_{sep.min}$, can be calculated by rearranging formula 17, which can then be substituted into formula 3 to calculate the corresponding distance to detector, $l_{det.min}$:

$$l_{sep.min} = \frac{5\min(\mu_L, \mu_C)w}{\Delta\mu} \tag{31}$$

$$l_{det.min} = l_{sep.min} + l_{prop} \tag{32}$$

Between $l_{det.min}$ and $l_{det.instr.min}$, the larger value is used as the recommendation:

$$l_{det} = \max(l_{det.min}, l_{det.instr.min}) \tag{33}$$

Optimization of l_{det} fails if $l_{det.min} > l_{det.instr.max}$, with a suggestion to improve $\Delta\mu$ by modifying separation conditions (e.g., by choosing a BGE which results in a slower electroosmotic flow).

Step 3. Once l_{det} is defined, the range of optimum values of E can be determined. The choice of E should satisfy all of the remaining conditions, presented in formulas 22, 23, 25, and 26. The maximum acceptable E is limited by four parameters: $V_{instr.max}$, $I_{instr.max}$, $T_{cool.min}$, and τ:

$$E_{max.V} = \frac{V_{instr.max}}{l_{det} + l_{out}} \tag{34}$$

$$E_{max.I} = \frac{\rho I_{instr.max}}{\pi r^2} \tag{35}$$

$$E_{max.Tef} = \frac{-m_1}{2m_2} + \frac{\sqrt{m_1^2 - 4m_2(m_0 - T_{ef.goal} + T_{cool.min})}}{2m_2} \tag{36}$$

$$E_{max.\tau} = \frac{w(k_{on}[T]_{eq} + k_{off})}{0.15\Delta\mu} \tag{37}$$

where m_2, m_1, m_0 are coefficients determined by fitting the relationship between E and T_{ef} with a second-order polynomial (see section S7 in the Supporting Information) and ρ is the experimentally measured resistivity of the BGE, calculated as

$$\rho = \frac{E\pi r^2}{I_{av}} \tag{38}$$

The minimum acceptable E is limited by four parameters: $V_{instr.min}$, $t_{run.max}$, t_{dis}, and τ:

$$E_{min.V} = \frac{V_{instr.min}}{l_{det} + l_{out}} \tag{39}$$

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

$$E_{\text{min.trun}} = \frac{l_{\text{sep}}}{\min(\mu_{\text{C}}, \mu_{\text{L}})t_{\text{run.max}}} \quad (40)$$

$$E_{\text{min.dis}} = \frac{l_{\text{sep}}}{\mu_{\text{C}} t_{\text{dis}}} \quad (41)$$

$$E_{\text{min}.\tau} = \frac{w(k_{\text{on}}[\text{T}]_{\text{eq}} + k_{\text{off}})}{0.6\Delta\mu} \quad (42)$$

The most conservative range of acceptable values of E is defined by the smallest of the maximum values and the largest of the minimum values.

$$E_{\text{max}} = \min(E_{\text{max.V}}, E_{\text{max.I}}, E_{\text{max.Tef}}, E_{\text{max}.\tau}) \quad (43)$$

$$E_{\text{min}} = \max(E_{\text{min.V}}, E_{\text{min.trun}}, E_{\text{min.dis}}, E_{\text{min}.\tau}) \quad (44)$$

$$E_{\text{range}} = E_{\text{max}} - E_{\text{min}} \quad (45)$$

The interval between E_{min} and E_{max} represents the set of all acceptable values of E. Since choosing a value from this interval smaller than E_{max} will lead to an unnecessarily prolonged experiment, E_{max} is used as the recommendation. Optimization of E fails if E_{range} is smaller than zero, followed by a suggestion message that is defined by considering the combination of variables which define E_{max} and E_{min} (see section S8 in the Supporting Information). There is one special case: when E_{min} is defined by $E_{\text{min.dis}}$ (i.e., if C becomes undetectable before it can reach the detector even at the highest E), it may be possible to find an acceptable set of conditions by shortening l_{sep}. This can be achieved by increasing l_{prop} beyond the value required to propagate the sample through the inefficiently cooled portion of the capillary. The longest l_{sep} which allows for the complex to be detected, $l_{\text{sep.dis}}$, as well as the adjusted l_{prop} can be calculated as follows:

$$l_{\text{sep.dis}} = t_{\text{dis}} \min(\mu_{\text{L}}, \mu_{\text{C}}) E_{\text{max}} \quad (46)$$

$$l_{\text{prop}} = l_{\text{det.instr.min}} - l_{\text{sep.dis}} \quad (47)$$

The resulting separation quality must still satisfy condition 21; thus, the algorithm fails if the calculated $l_{\text{sep.dis}} < l_{\text{sep.min}}$. Optimization of E also fails if the determined E_{max} leads to $T_{\text{inef}} \geq T_{\text{boil}}$, with the suggestion to decrease the generation of Joule heating (e.g., by increasing the resistivity of the BGE) or improve heat dissipation (e.g., by decreasing r).

Step 4. If the user wishes to perform analysis in the shortest possible time (a common goal in analytical applications), the algorithm completes at the end of step 3. However, if the user wishes to achieve the best possible separation, i.e., to maximize S (a common goal in preparative applications), then the algorithm proceeds with iterative maximization of S, by increasing l_{sep}. Maximum l_{sep} is limited by the smaller of the values between $l_{\text{det.instr.max}}$ and $l_{\text{sep.dis}}$.

$$l_{\text{sep.max}} = \min(l_{\text{det.instr.max}}, l_{\text{sep.dis}}) \quad (48)$$

Due to a circular dependence between $l_{\text{sep.dis}}$, E_{max}, and l_{det}, the value of $l_{\text{sep.max}}$ cannot be calculated directly, but must be approached iteratively with "half-step" increments. Every iteration begins by estimating an updated value of l_{det} ($l_{\text{det.new}}$) which will yield increased separation, while being limited by $l_{\text{sep.dis}}$. An estimate value of $l_{\text{sep.dis}}$ is calculated by formula 46, using the current value of E_{max}, but to avoid overshooting the goal, $l_{\text{det.new}}$ is increased by half of the difference between l_{det} value from the previous iteration ($l_{\text{det.previous}}$) and the latest estimate of $l_{\text{sep.dis}}$.

$$l_{\text{det.new}} = l_{\text{det.previous}} + 0.5(l_{\text{det.previous}} - (l_{\text{sep.dis}} + l_{\text{prop}})) \quad (49)$$

Once $l_{\text{det.new}}$ is defined, the algorithm proceeds to step 3, where a new set of values of E_{max} and E_{min} are calculated. The closer the estimated $l_{\text{sep.max}}$ is to its true value, the smaller is the interval between E_{max} and E_{min}. The iterations are repeated until E_{range} collapses to a value of 0, or until $l_{\text{det.new}} = l_{\text{det.instr.max}}$. At the end of the iterative procedure, the value of l_{det} will yield the best quality of separation (maximum S), within the limits of condition 22.

Modifications for Preparative Applications. For preparative applications (such as affinity purification or aptamer selection) complex detectability may not serve as a useful optimization criterion. Instead, it may be more advantageous to define the smallest allowable fraction of intact complex, f, and apply it to calculate t_{dis} using eq 9:

$$t_{\text{dis}} = \frac{-\ln(f)}{k_{\text{off}}} \quad (50)$$

For example, to improve the yield of affinity purification the users might decide that no more than 20% of the initial C may dissociate during an experiment, thus defining $f = 0.8$. For aptamer selection, on the other hand, it may be useful to define f as a small fraction, and substitute the value of k_{off} in formula 50 with its maximum desired value, thus defining the stringency of the selection process. It may also be advantageous to modify conditions 13 and 14 to increase the yield of preparative applications by maximizing complex formation: $[\text{L}]_0$ may be set at the highest available concentration, while $[\text{T}]_0$ may be calculated so that $R > 0.9$ using eq 11.

Software implementation. All parts of the described algorithm, including the iterative step, can be easily implemented in any of the available spreadsheet software, such as Microsoft Excel, or OpenOffice Calc. Our Excel implementation may be downloaded at http://www.yorku.ca/skrylov/resources.html. Besides the described algorithm, our application also includes an automated electropherogram processing and deconvolution subroutine, described previously,[19] which further reduces user involvement and makes the method more convenient and less prone to bias.

To test the performance of our software package, we have generated a number of NECEEM electropherograms with nonoptimized conditions, subjected them to automatic optimization, and then regenerated a new set of electropherograms following the recommendations from the software. Electropherograms before and after optimization are presented in Figure 3, along with the measure of accuracy in k_{off} and K_{d} extraction. The optimization essentially "standardizes" the shape of the electropherogram, ensuring that useful information can be extracted with the smallest possible error. This is confirmed by the fact that the optimized set of electropherograms consistently yields a more accurate combination of values than the nonoptimized ones.

Practical Application. To provide a demonstrative example of how the developed optimization algorithm can be used within a real experimental study, we used a model system of MutS protein (target) and a truncated version of its DNA aptamer (ligand). The goal of our analytical study was to measure both K_{d} and k_{off} accurately. First, we defined the study

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

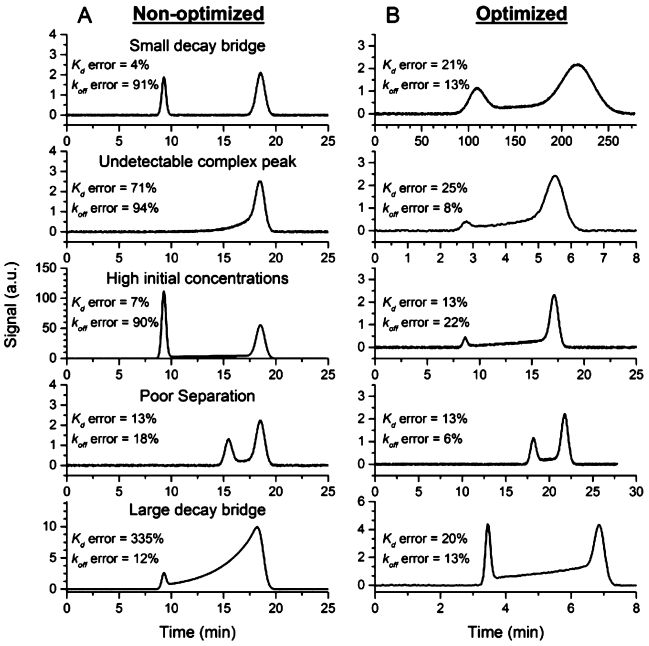


Figure 3. Application of the algorithm to computer-simulated experiments. Electropherograms were simulated at various nonoptimized conditions, representing some of the more commonly occurring problematic cases, and were subjected to optimization by the developed algorithm. Panel A shows electropherograms before the optimization, while panel B shows the results after algorithm-guided adjustment of experimental variables.

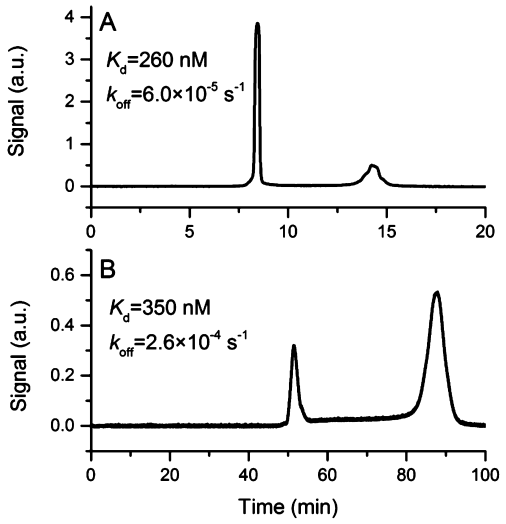


Figure 4. Application of the algorithm within an experimental study of interaction between MutS protein and a truncated aptamer. (A) Electropherogram resulting from a nonoptimized experiment. This electropherogram was subjected to the software implementation of the algorithm and failed the validation phase. (B) Recommendations from the software were applied to a new experiment, which resulted in the presented electropherogram. The electropherogram in panel B passed the validation test.

parameters: (i) we decided to use the protein storage buffer, 50 mM Tris−HCl pH 7.5 with added 20 mM NaCl, as the BGE for our measurements in order to avoid any buffer mismatch phenomena; (ii) we were interested in studying the interaction at 25 °C; (iii) as minimization of sample consumption is often desirable in analytical studies, we chose to perform the measurement in a capillary with an internal radius of 10 μm; (iv) as processivity was not a priority, we used the shortest reproducible width of injection plug for our instrument, which is 5 mm. Prior to analysis of the equilibrium mixture, we have performed a CE run with only the aptamer at 10 nM, which allowed us to measure $\sigma = 4.6 \times 10^{-9}$ arbitrary units (a.u.), $\gamma_L = 1.1 \times 10^7$ a.u. M^{-1}, and $\rho = 4.1$ Ωm. We then performed another preliminary experiment, where we analyzed an equilibrium mixture prepared with $[L]_0 \approx 100\sigma/\gamma_L$ (10 nM) and the highest possible $[T]_0$ (4.5 μM), which saturated the binding between T and L and allowed us to measure $\gamma_C = 0.9\gamma_L$. The initial analytical experiments were performed by following the recommendations described in section S9 in the Supporting Information: the first experiment was performed at $[L]_0 \approx 100\sigma/\gamma_L$, but no complex peak was detected; thus, $[L]_0$ was increased 5−10-fold until a prominent complex peak was observed. A resulting electropherogram, presented in Figure 4A, contained both peaks, but did not contain a discernible decay bridge. The electropherogram was subjected to validation and optimization using the algorithm and was rejected based on the unsatisfactory combination of initial concentrations and a small value for τ. The software provided us with the following recommendation: decrease initial concentrations to $[L]_0 = 60$ nM and $[T]_0 = 300$ nM, decrease l_{det} to 10 cm, and decrease E to 66 V cm^{-1}. These recommendations were applied to a second analytical experiment, along with the desired values for the experimental parameters. The resulting electropherogram (Figure 4B) contained all of the expected NECEEM features at correct proportions, and when subjected to the algorithm, passed the validation test. As a result, the obtained K_d = 350 nM ± 14% and $k_{off} = 2.6 \times 10^{-4}$ s^{-1} ± 8% were accepted as accurate (the presented percentage relative error is the variability in experimental measurements based on three parallel trials). This demonstrates that the developed algorithm allows an experiment to be optimized and the results validated within a few experimental runs.

Concluding Remarks. The use of KCE methods is still predominantly limited to academic settings, where the opportunities of an educational environment allow the operators to become highly skilled CE specialists. To promote the use of KCE in industrial and clinical settings, it must be made more accessible to less experienced users, such that investments in operator training can be converted into productive work within the shortest possible time. We believe that the presented optimization algorithm can be an efficient tool in the hands of novice NECEEM experimenters, which will allow them to produce reliable data while their experience with the methodology grows. We also believe that this optimization algorithm may be an important step toward greater automation of NECEEM analysis for applications in the industry. Lastly, the presented algorithm may serve as a useful tool during the establishment of the analytical design space,[30] as part of implementation of the "quality by design" principles in pharmaceutical assay development. Other KCE methods stand to benefit from development of similar optimization approaches. It should be noted that the provided optimization approach does not eliminate some sources of error inherent to KCE methods. As such, the users are encouraged to keep in mind that the NECEEM-studied molecular interactions can be influenced by the strength of the applied electric field,[31] by chemical modification during fluorescent labeling, and by the interaction of the sample with the walls of the bare-silica capillaries.

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.6b02882.

List of symbols, additional calculations and recommended values for variables, use of SUMET for optimization of in-capillary temperatures, suggestion messages for change in experimental design during E_{range} failure, and recommended conditions for initial experiment (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: skrylov@yorku.ca.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by Natural Sciences and Engineering Research Council of Canada (Discovery Grant 238990). M.K. was supported by Alexander Graham Bell Canada Graduate Scholarship.

■ REFERENCES

(1) Pardue, H. L. *Anal. Chim. Acta* **1989**, *216*, 69−107.
(2) Homola, J. *Chem. Rev.* **2008**, *108*, 462−493.
(3) Auer, S.; Koho, T.; Uusi-Kerttula, H.; Vesikari, T.; Blazevic, V.; Hytönen, V. P. *Sens. Actuators, B* **2015**, *221*, 507−514.
(4) Wishnia, A.; Boussert, A.; Graffe, M.; Dessen, P.; Grunbergmanago, M. *J. Mol. Biol.* **1975**, *93*, 499−515.
(5) Galievsky, V. A.; Stasheuski, A. S.; Krylov, S. N. *Anal. Chem.* **2015**, *87*, 157−171.
(6) Crouch, S. R.; Scheeline, A.; Kirkor, E. S. *Anal. Chem.* **2000**, *72*, 53−70.
(7) Ohlson, S. *Drug Discovery Today* **2008**, *13*, 433−439.
(8) Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5076−5080.
(9) Nunez, S.; Venhorst, J.; Kruse, C. G. *Drug Discovery Today* **2012**, *17*, 10−22.
(10) Schasfoort, R. B.; Tudos, A. J. *Handbook of Surface Plasmon Resonance*; Royal Society of Chemistry: Cambridge, U.K., 2008; pp 275−310.
(11) Shah, N. B.; Duncan, T. M. *J. Visualized Exp.* **2014**, e51383.
(12) Ogilvie, I. R. G.; Sieben, V. J.; Mowlem, M. C.; Morgan, H. *Anal. Chem.* **2011**, *83*, 4814−4821.
(13) Petrov, A.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 17104−17110.
(14) Okhonin, V.; Krylova, S. M.; Krylov, S. N. *Anal. Chem.* **2004**, *76*, 1507−1512.
(15) Osmond, R. I. W.; Kett, W. C.; Skett, S. E.; Coombe, D. R. *Anal. Biochem.* **2002**, *310*, 199−207.
(16) Secundo, F. *Chem. Soc. Rev.* **2013**, *42*, 6250−6261.
(17) Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124*, 13674−13675.
(18) Krylov, S. N. *J. Biomol. Screening* **2006**, *11*, 115−122.
(19) Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. *Anal. Chem.* **2015**, *87*, 3099−3106.
(20) Drabovich, A. P.; Berezovski, M.; Okhonin, V.; Krylov, S. N. *Anal. Chem.* **2006**, *78*, 3171−3178.
(21) Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. *Anal. Chem.* **2010**, *82*, 8637−8641.
(22) Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. *Anal. Chem.* **2010**, *82*, 8692−8695.
(23) Cherney, L. T.; Kanoatov, M.; Krylov, S. N. *Anal. Chem.* **2011**, *83*, 8617−8622.
(24) Yang, P.; Mao, Y.; Lee, A. W. M.; Kennedy, R. T. *Electrophoresis* **2009**, *30*, 457−464.
(25) Pavski, V.; Le, X. C. *Anal. Chem.* **2001**, *73*, 6070−6076.
(26) Mosing, R. K.; Mendonsa, S. D.; Bowser, M. T. *Anal. Chem.* **2005**, *77*, 6107−6112.
(27) Galievsky, V. A.; Stasheuski, A. S.; Krylov, S. N. *Anal. Chim. Acta* **2016**, *935*, 58−81.
(28) Long, G. L.; Winefordner, J. D. *Anal. Chem.* **1983**, *55*, 712A−724A.
(29) Patel, K. H.; Evenhuis, C. J.; Cherney, L. T.; Krylov, S. N. *Electrophoresis* **2012**, *33*, 1079−1085.
(30) Rozet, E.; Lebrun, P.; Debrus, B.; Boulanger, B.; Hubert, P. *TrAC, Trends Anal. Chem.* **2013**, *42*, 157−167.
(31) Musheev, M. U.; Filiptsev, Y.; Okhonin, V.; Krylov, S. N. *J. Am. Chem. Soc.* **2010**, *132*, 13639−13641.

DOI: 10.1021/acs.analchem.6b02882
Anal. Chem. 2016, 88, 9300−9308

SUPPORTING INFORMATION

Systematic Approach to Optimization of Experimental Conditions in Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures

Mirzo Kanoatov, Sina Mehrabanfar and Sergey N. Krylov*

Department of Chemistry and Centre for Research on Biomolecular Interactions and
York University, Toronto, Ontario M3J 1P3, Canada

Table of contents:

S1. List of Symbols

a, c, g, k, n	empirical constants for determination of in-capillary temperature
A_C	area of complex peak on an electropherogram
A_D	area of the decay bridge on an electropherogram
A_L	area of free-ligand peak on an electropherogram
C	intermolecular complex between ligand and target
$[C]_{eq}$	equilibrium concentration of complex
E	strength of applied electric field
E_{max}	largest acceptable magnitude of applied electric field strength
$E_{max.I}$	largest magnitude of applied electric field strength, limited by $I_{instr.max}$
$E_{max.Tef}$	largest magnitude of applied electric field strength, limited by T_{ef}
$E_{max.V}$	largest magnitude of applied electric field strength, limited by $V_{instr.max}$
$E_{max.\tau}$	largest magnitude of applied electric field strength, limited by τ
E_{min}	smallest acceptable magnitude of applied electric field strength
$E_{min.dis}$	smallest magnitude of applied electric field strength, limited by t_{dis}
$E_{min.trun}$	smallest magnitude of applied electric field strength, limited by $t_{run.max}$
$E_{min.\tau}$	smallest magnitude of applied electric field strength, limited by τ
$E_{min.V}$	smallest magnitude of applied electric field strength, limited by $V_{instr.min}$
E_{range}	range of the interval between E_{max} and E_{min}
f	minimum desired fraction of complex to reach the detector intact
γ_C	signal response factor for complex
γ_L	signal response factor for ligand
h_C	height of complex peak
h_L	height of ligand peak
I	electric current
I_{av}	average electric current during electrophoresis
$I_{instr.max}$	maximum current supplied by the instrument
K_d	equilibrium dissociation constant
K_d^*	apparent (non-validated) value of K_d
k_{off}	rate constant of complex dissociation
k_{off}^*	apparent (non-validated) value of k_{off}
k_{on}	rate constant of complex formation
L	ligand molecule
$[L]_0$	initial (pre-equilibrium) concentration of ligand
$[L]_{eq}$	equilibrium concentration of ligand
l_{det}	length between capillary inlet and detector
$l_{det.instr.max}$	maximum capillary length to detector
$l_{det.instr.min}$	minimum capillary length to detector
$l_{det.min}$	shortest acceptable length between capillary inlet and detector
$l_{det.new}$	value of l_{det} used in the current iteration of step 4
$l_{det.previos}$	value of l_{det} used in the previous iteration of step 4
l_{out}	length between the detector and capillary outlet
l_{prop}	length of initial pressure propagation zone
l_{sep}	capillary length available for separation
$l_{sep.dis}$	capillary length available for separation which results in t_{run} equal to t_{dis}
$l_{sep.max}$	capillary length available for separation which results in highest acceptable value of S
$l_{sep.min}$	shortest acceptable capillary length available for separation
l_{tot}	total capillary length
m_2, m_1, m_0	coefficients determined by fitting the relationship between E and T_{ef} with a second order polynomial
μ_C	mobility of complex
μ_L	mobility of ligand
$\Delta\mu$	absolute difference in mobilities of ligand and complex
R	fraction of non-bound ligand
r	inner radius of capillary
ρ	capillary resistivity
S	number of zone separations events that occur during an experiment
σ	average amplitude of noise at baseline
T	target molecule
$[T]_0$	initial (pre-equilibrium) concentration of target
$[T]_{eq}$	equilibrium concentration of target
T_{amb}	ambient room temperature

T_{boil}	boiling temperature of the BGE
T_{cool}	coolant temperature used during the experiment
$T_{cool.min}$	minimum coolant temperature
T_{ef}	temperature in cooled portion
$T_{ef.goal}$	desired temperature in cooled portion of the capillary
T_{inef}	temperature in uncooled portion
t_C	time of migration of the complex peak to the detector
t_{dis}	time that it takes for the intact complex peak height to reduce to the value of the limit of quantitation
t_{eq}	characteristic time of equilibration
t_L	elution time of ligand
t_{run}	total analysis time, time it takes for the slowest detectable component to reach the detector
$t_{run.max}$	longest possible analysis time
t_{sep}	characteristic time of separation
τ	ratio of characteristic times t_{sep} and t_{eq}
$V_{instr.max}$	maximum electrophoresis voltage supplied by the instrument
$V_{instr.min}$	minimum electrophoresis voltage supplied by the instrument
v_C	velocity of complex
v_L	velocity of ligand
v_T	velocity of target
w	width of sample injection zone

S2. Calculating concentrations of equilibrium mixture components

Equilibrium concentrations of components can be related to their initial (pre-equilibrium) concentrations by the following equations:

$$[L]_{eq} = [L]_0 - [C]_{eq} \quad \text{(S1)}$$

$$[T]_{eq} = [T]_0 - [C]_{eq} \quad \text{(S2)}$$

$$[C]_{eq} = [T]_0 - [T]_{eq} \quad \text{(S3)}$$

$[C]_{eq}$ and $[L]_{eq}$ can be expressed through $[T]_{eq}$ and substituted into Equation 2 in the main text. The physically-meaningful positive root solution of the resulting quadratic equation can be used to calculate $[T]_{eq}$.

$$[T]_{eq} = \frac{([T]_0 - [L]_0 - K_d^*)}{2} + \frac{\sqrt{([L]_0 - [T]_0 + K_d^*)^2 + 4(K_d^*[T]_0)}}{2} \quad \text{(S4)}$$

Once $[T]_{eq}$ is calculated, $[C]_{eq}$ and $[L]_{eq}$ can be determined using Equations (S1) and (S3).

S3. Relationship between initial concentrations and error in measuring of K_d

We have previously derived a relationship between variability in experimental measurements (ΔR) and errors in the calculated final result (ΔK_d):[1]

$$\Delta K_d \approx \left(-[L]_0 + \frac{[T]_0}{(1-R)^2} \Delta R \right) \quad \text{(S5)}$$

Using Eq. (S5), dependence of relative error, $\Delta K_d/K_d$, on dimensionless ligand concentration $[L]_0/K_d$ at constant R and ΔR, can be calculated and plotted (**Fig. S3A**). Similarly, dependence of $\Delta K_d/K_d$ on value of R can be calculated and plotted at constant $[T]_0$ and ΔR (**Fig. S3B**).

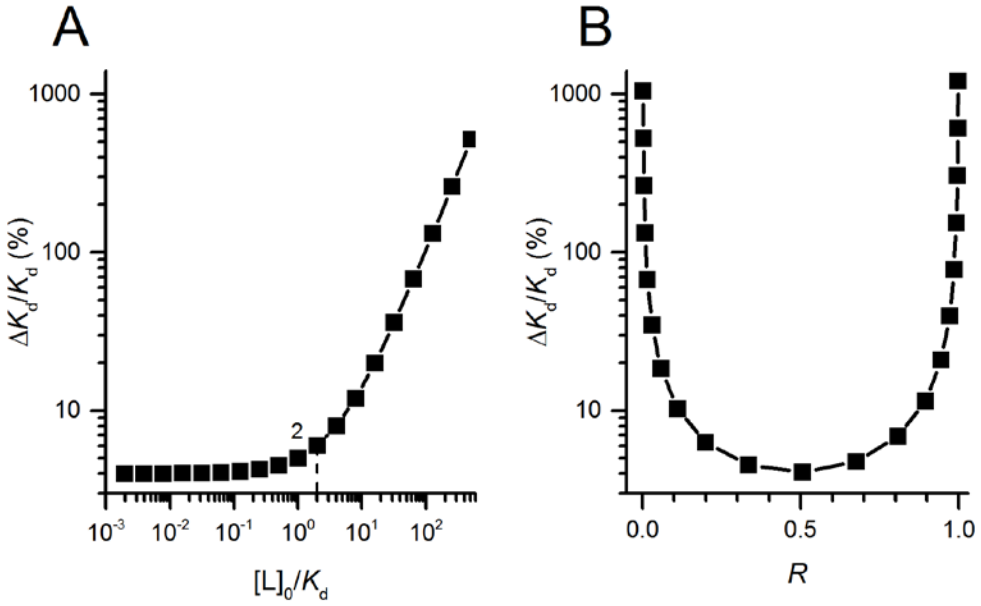


Figure S3: Relative error in determination of K_d, depending on the ratio of $[L]_0$ to K_d (panel A) and the value of R (panel B). **Panel A:** ΔK_d was calculated using Eq. (S5) by varying $[L]_0$ and $[T]_0$ to yield

R = 0.5. K_d was set to 1, and ΔR equaled 1% of R. **Panel B:** ΔK_d was calculated using Eq. (S5) by varying $[L]_0$ to yield R values in the range of 0.01 and 0.99. K_d was set to 1, $[T]_0 = K_d$, and ΔR equaled 1% of R. Note that both y-axes are presented in logarithmic scale.

Since ΔR may differ widely between instruments and experimental systems, it is difficult to define an error threshold which will be universally applicable. Thus, optimization criteria should be based on general properties of curves in **Figure S3**, such as their inflection points and minima. As a result, the most accurate determination of K_d requires the measurements to be performed at $[L]_0 \leq 2K_d$, and R=0.5.

S4. Minimum value of *S* for accurate deconvolution of peak areas

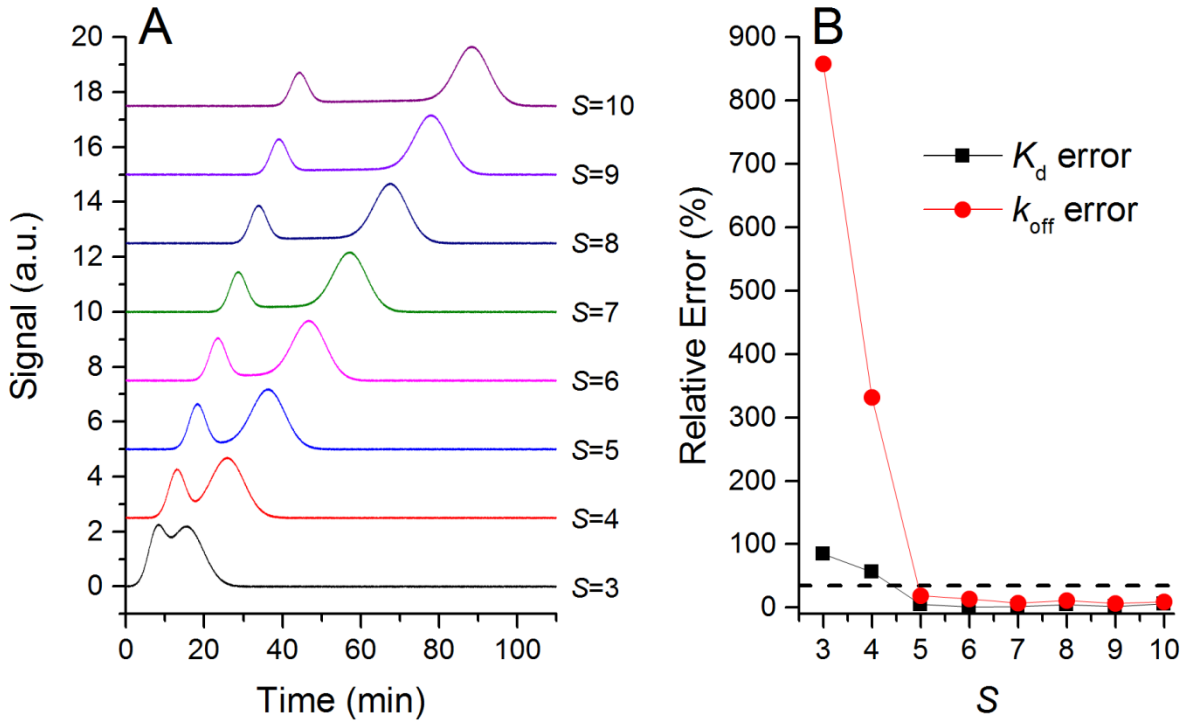


Figure S4: Panel A: Set of simulated NECEEM electropherograms at increasing values of S. For the simulations, all of the conditions were fixed, except for an increasing value of l_{det}. **Panel B:** Relative error in extracting K_d and k_{off} related to the value of S. By setting a maximum error threshold at 20%, we can establish the minimum value S = 5.

S5. Range of values of τ for accurate deconvolution of peak areas

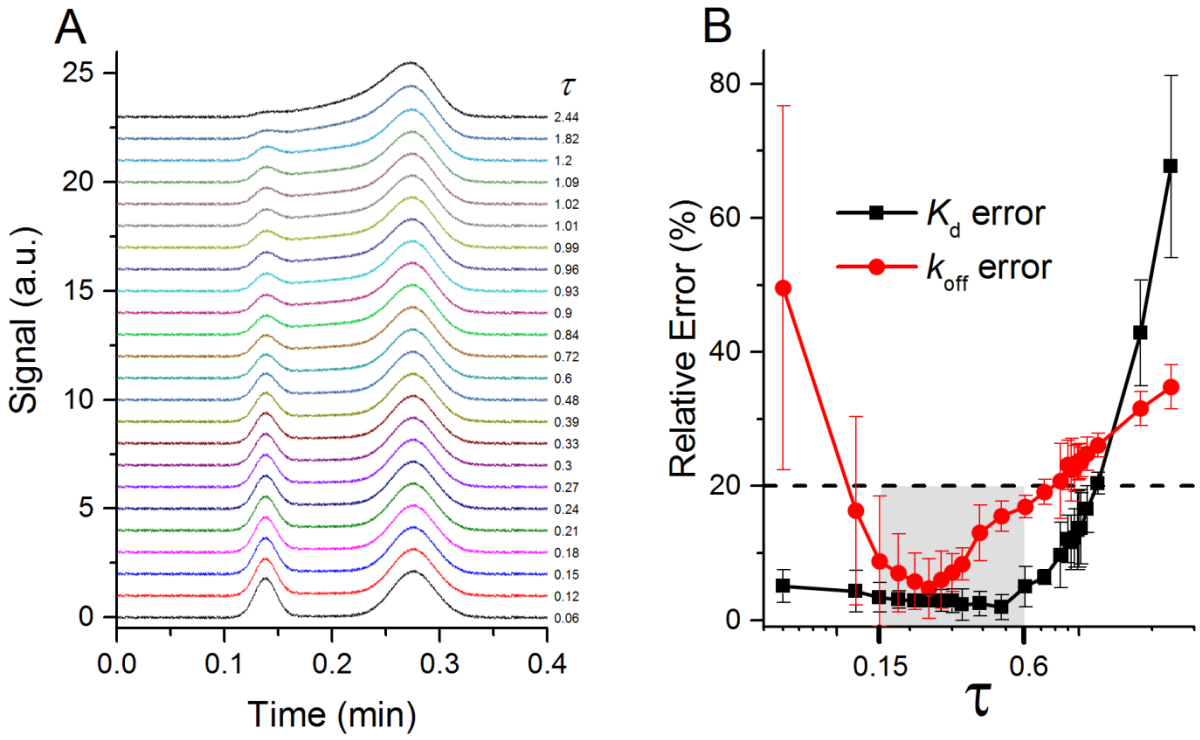


Figure S5: Panel A: Set of simulated NECEEM electropherograms at increasing values of τ. For the simulations, all of the conditions were fixed, except the values of k_{off} and k_{on}. **Panel B:** Relative error in extracting K_d and k_{off} related to the value of τ used to generate the electropherogram. As the random noise played a significant role in the magnitude of the error, each electropherogram was regenerated 5 times, with new values of the noise. The markers on the graph show the average value of the error at a given τ value, while the error bars show one standard deviation of the error based on 5 measurements. By setting a maximum error threshold at 20%, we can establish the optimum range of τ values to lie between 0.15 and 0.6.

S6. Minimum signal-to-noise ratio for accurate measurement of peak areas

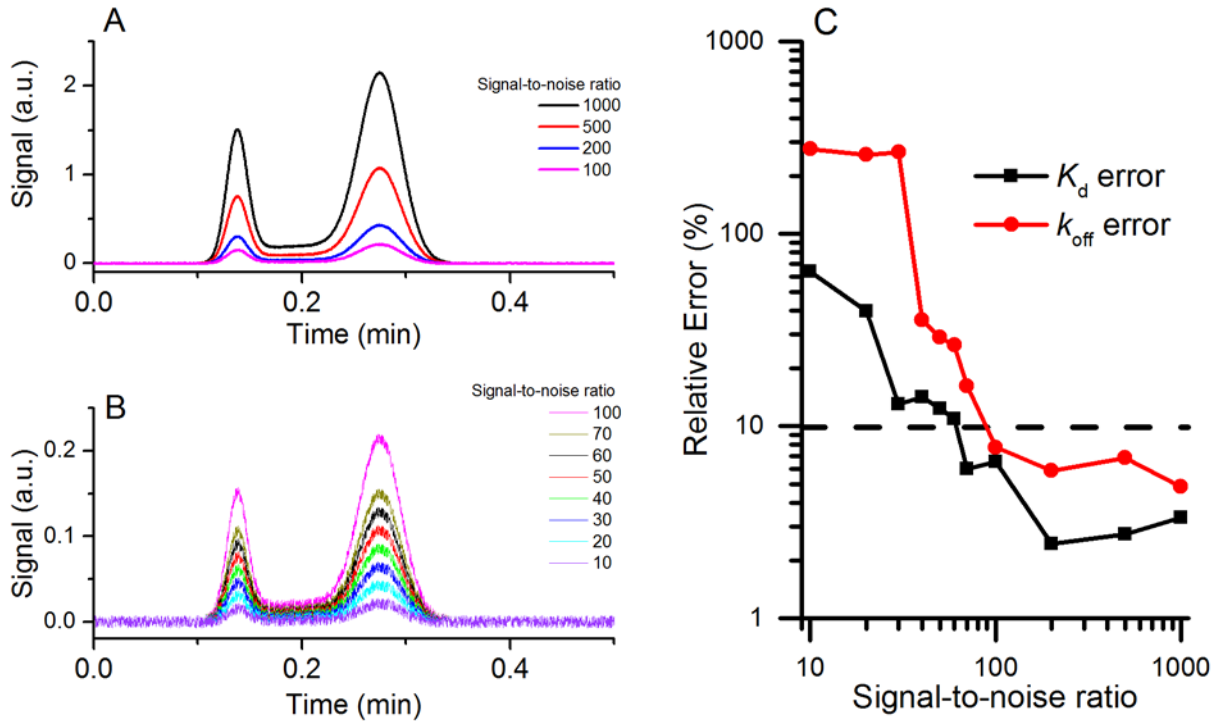


Figure S6. Effects of signal-to-noise ratio in ligand detection on accuracy of NECEEM data extraction. Signal-to-noise ratios are based on the height of the peak produced by a given concentration of ligand in absence of the target. **Panel A:** Set of simulated NECEEM electropherograms at decreasing signal-to-noise ratios ranging between 1000 to 100. **Panel B:** Set of simulated NECEEM electropherograms at decreasing signal-to-noise ratios ranging between 100 to 10. For the simulations, all of the conditions were fixed, except the values of γ_C and γ_L (which were always equal to each other). **Panel C:** Relative error in extracting K_d and k_{off} related to the signal-to-noise ratio used to generate an electropherogram. By setting a maximum error threshold at 10%, we can establish the minimum value of signal-to-noise ratio at 100.

S7. Using SUMET for optimization of in-capillary temperatures

The developed algorithm employs simplified universal method for determining electrolyte temperatures (SUMET) to estimate the temperatures in the efficiently-cooled and inefficiently-cooled portions of the capillary. To perform necessary calculations, the following coefficients are required:[2, 3]

Table S1. Values of coefficients *a*, *c*, *g*, *k* and *n*, required in calculations of T_{ef} and T_{inef}, depending on the choice of capillary inner-radius.

Capillary radius, *r* (μm)	***a***	***c***	***g***	***k***	***n***
10	1.452	0.558	0.851	0.958	7.39
25	1.412	0.454	0.382	0.949	7.54
37.5	1.437	0.881	0.250	0.950	6.56
50	7.091	8.873	0.186	0.959	6.54
75	1.339	1.584	0.494	0.968	5.07
100	10.15	12.33	0.157	0.951	7.37

These coefficients are applicable to liquid-cooled systems. For air cooled systems, please see the above provided reference.

To generate recommendations for optimization of *E*, it is required that we are able to predict values of *E* that correspond to a given T_{ef}. Equation 25 in the main text is difficult to solve for *E*, however, the relationship between the two parameters can be accurately approximated by fitting it with a second-order polynomial. This can be easily achieved by any of the available linear regression tools, for example, by using the LINEST function in Microsoft Excel. The best way to approach this problem is to generate a range of points for *E*, spanning values defined by $V_{instr.min}$ and $V_{instr.max}$ and the given l_{total}, and then using Equation 25 in the main text to calculate the corresponding T_{ef}. A linear regression tool of choice can be used to fit the calculated data points to an equation of a form:

$$m_2 E^2 + m_1 E + m_0 = T_{ef} \qquad \text{(S6)}$$

The physically meaningful positive root solution of Eq. (S6) can then be used to approximate *E* with a given T_{ef}. **Figure S7** depicts the typical accuracy of fitting of the relationship between *E* and T_{ef} by a second-order polynomial.

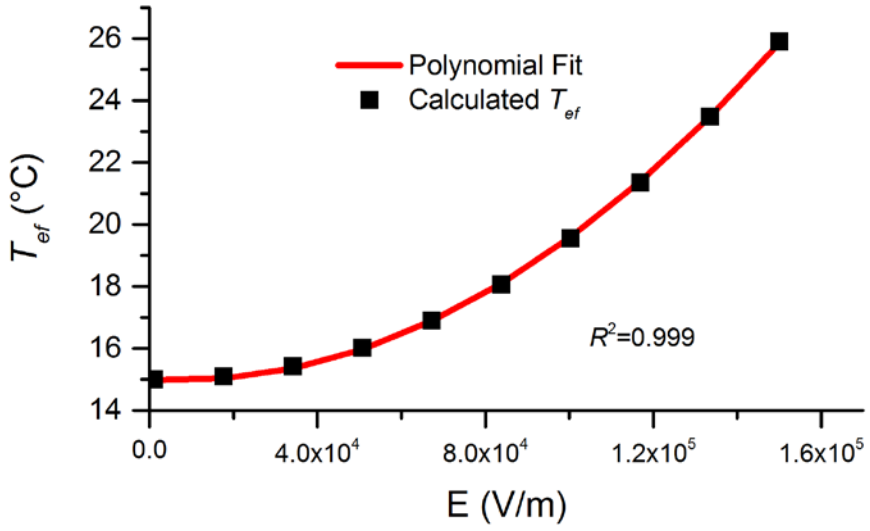


Figure S7: Example of fitting of the relationship between *E* and T_{ef} using a second-order polynomial.

S8. Suggestion messages for change in experimental design during E_{range} failure

Optimization of E in step 3 of phase 2 of the algorithm fails as a result of a negative value of E_{range}. The combinations of variables which define E_{max} and E_{min} during this failure determine which suggestion for modifying the experimental design is output by the algorithm. **Table S2** lists 16 possible combinations of the variables, and the message that accompanies the failure.

Table S2. All possible combinations of variables defining E_{min} and E_{max}, and the corresponding failure output message

Variable for E_{min}	Variable for E_{max}	Message
$E_{min.V}$	$E_{max.V}$	Information on instrument limitations is entered incorrectly
$E_{min.V}$	$E_{max.I}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.V}$	$E_{max.Tef}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.V}$	$E_{max.\tau}$	k_{off} is too slow for measurement, increase w
$E_{min.dis}$	$E_{max.V}$	t_{dis} is too small, increase $[T]_0$ to decrease R
$E_{min.dis}$	$E_{max.I}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.dis}$	$E_{max.Tef}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.dis}$	$E_{max.\tau}$	Unlikely case, check to see if all of the information is correct
$E_{min.trun}$	$E_{max.V}$	Migration is very slow, increase buffer concentration to improve buffer depletion time
$E_{min.trun}$	$E_{max.I}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.trun}$	$E_{max.Tef}$	BGE conductivity is too high, reduce r or increase ρ
$E_{min.trun}$	$E_{max.\tau}$	Migration is very slow, increase w
$E_{min.\tau}$	$E_{max.V}$	Separation is too slow, decrease w
$E_{min.\tau}$	$E_{max.I}$	Separation is too slow, decrease w
$E_{min.\tau}$	$E_{max.Tef}$	Separation is too slow, decrease w
$E_{min.\tau}$	$E_{max.\tau}$	Information on instrument limitations is entered incorrectly

S9. Recommended conditions for initial experiment

Novel studies of biomolecular interactions begin without any *a priori* knowledge of their properties. The conditions for a study-initiating experiment cannot be tailored for accurate determination of K_d and k_{off}, but instead should be aimed at minimizing the odds of producing uninformative electropherograms with indistinguishable or missing features. To facilitate subsequent optimization, the initial experiment should at the least yield estimates of analyte mobilities (μ_L and μ_C) and response factors (γ_L and γ_C), thus, it must produce an electropherogram with discernible peaks of intact C and free L. The best way to achieve this is to perform a short experiment (small t_{run}), to prevent complete dissociation of C, at the fastest-possible rate of separation (small t_{sep}), to prevent excessively large decay bridge or insufficient peak separation. Thus, the initial experiment should be performed using the conditions outlined in **Table S3**.

Table S3. Recommendations for choosing experimental conditions for a study-initiating experiments

Parameter or variable	Value
r	smallest available
$[L]_0$	$100\sigma/\gamma_L$
$[T]_0$	$5[L]_0$
l_{det}	shortest possible (usually 20-30 cm)
l_{prop}	to the edge of the cooled portion
w	shortest possible
E	maximum (voltage- or current-limited)
T_{cool}	smallest possible

If the preliminary experiment does not yield a detectable C peak and a decay bridge, then the $[L]_0$ and $[T]_0$ should be increased 10-fold to promote complexation. If the preliminary experiment does not yield a detectable C peak, but features a prominent decay bridge, then l_{det} should be shortened to decrease the degree of complex dissociation.

S10. Supporting References

(1) Kanoatov, M.; Galievsky, V. A.; Krylova, S. M.; Cherney, L. T.; Jankowski, H. K.; Krylov, S. N. *Anal. Chem.* **2015**, *87*, 3099-3106.

(2) Patel, K. H.; Evenhuis, C. J.; Cherney, L. T.; Krylov, S. N. *Electrophoresis* **2012**, *33*, 1079-1085.

(3) Evenhuis, C. J.; Musheev, M. U.; Krylov, S. N. *Anal. Chem.* **2011**, *83*, 1808-1814.