analytical chemistry

ARTICLE

pubs.acs.org/ac

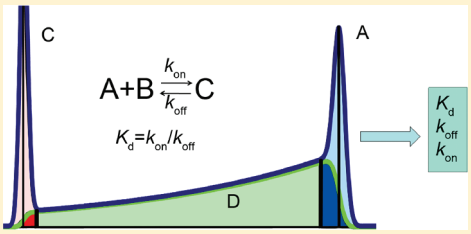
Method for Determination of Peak Areas in Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures

Leonid T. Cherney, Mirzo Kanoatov, and Sergey N. Krylov*

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

S Supporting Information

ABSTRACT: Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) facilitates determination of both the kinetic constants (k_{off}) and the equilibrium constants (K_d) of complex dissociation from a single experiment. A typical NECEEM electropherogram consists of two peaks and an "exponential bridge" between them, smoothly merging into the peaks. The values of k_{off} and K_d are usually calculated with simple algebraic formulas, by utilizing the areas of the peaks and the bridge. Accurate determination of the two constants requires accurate positioning of the two boundaries separating the bridge from the peaks. Here, we propose a more systematic method for the determination of boundaries between the peaks and the bridge. The method involves a simple geometrical analysis of a NECEEM electropherogram based on an assumption of symmetry in ordinary electrophoretic peaks. To test the method, we (i) constructed a series of computer-simulated NECEEM electropherograms, (ii) determined the two boundaries with our method, and (iii) calculated the values of k_{off} and K_d. We found that the deviation of the calculated values from those used to simulate the electropherograms did not exceed 15% for k_{off} and 25% for K_d, as long as the peaks and the bridge were visually identifiable. We finally applied the method to the determination of K_d and k_{off} values for the interaction between AlkB protein and its DNA aptamer. The developed method for rational boundary determination in NECEEM will facilitate accurate data analysis in a simple and efficient manner.



Many biological macromolecules assemble into noncovalent complexes for their proper function. Such complexes play an important role in key cellular processes including DNA replication, gene expression, signal transduction, transport across membranes, cell movement, enzyme maturation, etc.[1] Uncontrolled aggregation of misfolded proteins is known to result in pathological conditions (for example, in prion and amyloid diseases).[2] The kinetic and equilibrium stability of noncovalent complexes is characterized by the rate constant (k_{off}) and the equilibrium constant (K_d) of complex dissociation. Knowledge of both parameters is important for understanding the dynamics of the interaction and the role of noncovalent macromolecular complexes in cellular processes.

Several approaches have been developed to measure the rate constants of complex formation and dissociation.[3−10] Some of them, such as fluorescence correlation spectroscopy and nanopore amperometry, are microscopic in their nature and require observation of a statistically insignificant number of molecules.[4] Others, such as surface-immobilized binding sensors[5] and kinetic capillary electrophoresis (KCE) methods,[6−10] are macroscopic in their nature and require the measurement of changes in concentrations. Each of these approaches has different advantages, limitations, and major applications. KCE methods are particularly attractive when one of the reactants can be fluorescently labeled for detection without affecting the interaction. KCE methods can be used to measure a wide range of rate and equilibrium constants with only minute consumption of materials.

Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) is one of the KCE methods. It is used more often than others, because of the simplicity of both the experiment and the data processing.[6−8] Other KCE methods (such as SweepCE, plug−plug KCE, short SweepCEEM, and continuous NECEEM) require implementation of more-complex initial conditions or larger amounts of reactants than NECEEM does.[9,10] NECEEM was successfully applied to kinetic studies of protein−DNA,[11] protein−peptide,[12] and other macromolecular complexes.[13] It was also used for calibration-free quantitative analysis of proteins, using aptamers as affinity probes.[14−16] In addition, NECEEM has been proven to serve as a highly efficient partitioning approach in the selection of DNA aptamers.[14,16,17] Conceptually, NECEEM starts with mixing and equilibration of reactants A and B to allow the formation of their noncovalent complex, C:

$$\mathrm{A} + \mathrm{B} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{C} \qquad K_d = \frac{k_{off}}{k_{on}} = \frac{A_0 B_0}{C_0} \tag{1}$$

where A_0, B_0, and C_0 represent the equilibrium concentrations of A, B, and C, respectively. A short plug of the equilibrium mixture is then injected into a capillary and an electric field is applied to separate the three components via capillary electrophoresis. It is assumed that electrophoretic zones of A and B are separated fast so that the forward process in eq 1 is negligible, with respect to the reverse one. As a result, A and/or B are separated from C under nonequilibrium conditions and C continuously dissociates

Received: August 6, 2011
Accepted: October 13, 2011
Published: October 13, 2011

ACS Publications © 2011 American Chemical Society

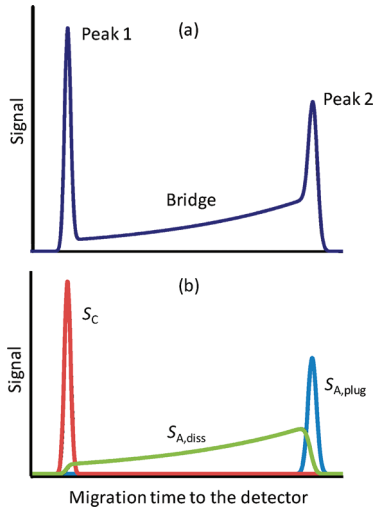


Figure 1. Graphical analysis of NECEEM data: (a) signal in a typical NECEEM electropherogram, and (b) signal deconvolution into its three components.

into A and B during the course of electrophoresis. Detection is typically arranged so that only one reactant (for example, A) and the complex C (C = AB) are visible (e.g., by fluorescently labeling A). In this case, a NECEEM electropherogram, which is, per se, signal versus migration time to a detector, contains three characteristic features: two peaks and a bridge between them that smoothly merges into the peaks (see Figure 1a). This cumulative signal (S) is a sum of three component signals: (i) signal generated by A that was unbound in the injected plug ($S_{A,plug}$), (ii) signal generated by A which was dissociated from C during electrophoresis ($S_{A,diss}$), and (iii) signal generated by C that reached the detector intact (S_C) (see Figure 1b).[8] $S_{A,plug}$ and S_C are major contributors to the two peaks, while $S_{A,diss}$ is the major contributor to the bridge in the electropherogram. Importantly, $S_{A,diss}$ overlaps at its flanks with $S_{A,plug}$ and S_C and thus merges the bridge flanks into the peaks. This overlap can create a challenge for accurately finding k_{off} and K_d from a NECEEM electropherogram and, therefore, is a reason for this study and its major subject.

The overlap does not impose a problem if a pattern-based approach is used to find k_{off} and K_d. In this approach, the constants are determined by fitting computer-simulated electropherograms into experimental ones.[8] While, arguably, being very accurate, the pattern-based approach is complicated and, thus, has not been utilized by practical users of NECEEM. Indeed, curve fitting is not a computatively transparent procedure that, generally speaking, depends on the initial curve used in fitting and may require significant computing time.[8]

In contrast, the overlap does create a challenge when a parameter-based approach is used to calculate k_{off} and K_d from a NECEEM electropherogram. The parameter-based approach utilizes the distinct features in a NECEEM electropherogram: areas under $S_{A,diss}$, $S_{A,plug}$, and S_C and the time position of maximum of S_C.[7,8] The areas and time should be determined from a NECEEM electropherogram, as shown in Figure 1a. Since the area under $S_{A,diss}$ partially overlaps with the areas under $S_{A,plug}$ and S_C (Figure 1b), determination of these three areas is not trivial. Generally, it requires deconvolution of the cumulative signal $S(t)$ into three components. Inaccurate deconvolution can result in inaccurate determination of k_{off} and K_d in the parameter-based approach. So far, the overlap of the areas in NECEEM electropherograms has been neglected, and the determination of the three areas has been done by visually placing vertical boundaries between the bridge and the peaks.[7] Subjective boundary positioning is influenced by individual judgment and experience of the analyst and leads to uncertainties in the determined k_{off} and K_d values. To exclude these uncertainties, a method is needed for rational and accurate determination of the overlapping areas. There are many studies on peak deconvolution in chromatography and electrophoresis.[18,19] Peak deconvolution has certain limitations and requires some preliminary knowledge about the individual peak shapes that are yet to be determined. For example, the application of integral transformations can be useful in peak deconvolution if individual signals are not considerably different in shape and are relatively narrower than the cumulative signal.[18] Evidently, this is not true for NECEEM signals (Figure 1). Deconvolution by curve fitting is based on known or hypothetical forms of the individual peaks.[19] In fact, the incorporation of such technique in NECEEM amounts to the use of the pattern-based approach. Indeed, the latter itself is based on fitting of the sum of three analytical solutions of the mass-transfer equations (see expressions (S5)−(S7) in the Supporting Information) into the cumulative signal. Therefore, a simpler and more straightforward method for the analysis of overlapping zones is needed for the parameter-based approach.

Here, we propose a systematic method for the determination of the areas under $S_{A,plug}$, $S_{A,diss}$, and S_C through the rational positioning of boundaries to account for the overlap of $S_{A,diss}$ with $S_{A,plug}$ and S_C. This analysis uses a general assumption of axial symmetry of $S_{A,plug}$ and S_C. Such an assumption is typically valid if neither A nor C adsorbs to inner capillary walls, and if the sample and run buffers are identical. To determine the accuracy of our method, we constructed a series of over 30 computer-simulated NECEEM electropherograms with a wide range of k_{off} and K_d values, determined the boundary positions with our method, and calculated K_d and k_{off} using the algebraic formulas previously derived. We found that the constants determined in this way deviated from those used in construction of the simulated electropherograms by no more than 15% for k_{off} and 25% for K_d for 90% of the electropherograms. We expect that the proposed method will be very useful in fast data analysis of NECEEM experiments.

RESULTS AND DISCUSSION

Basic Equations for Parameter-Based Approach to k_{off} and K_d Determination in NECEEM. Since the injected plug contains the equilibrium mixture of A, B, and C, the following relations occur in the injected plug before separation begins:

$$a_0 b_0 = V K_d c_0, \quad a_0 + c_0 = a_{tot}, \quad b_0 + c_0 = b_{tot} \tag{2}$$

Here, a_0, b_0, and c_0 are total equilibrium amounts of A, B, and C, respectively, in the plug; V is the plug volume; and a_{tot} and b_{tot} are total amounts of A and B used to prepare the plug (without adding any amount of C). The first relation in Expression 2 represents the law of mass action. The other two relations in Expression 2 express the mass balance between A, B, and C in reaction 1. The total amount $c(t)$ of complex C decreases with time, according to the following law:[8]

$$c(t) = c_0 \exp(-k_{off} t) \tag{3}$$

Equation 3 results from general equations of mass transfer in the NECEEM approximation of negligible forward process (see the Supporting Information). Expressions 2 and 3 allow us to express K_d and k_{off} and in terms of quantities that either are known or can be easily measured:[8]

$$K_d = \frac{(b_{tot}/V)[1 + (a_0/c_0)] - (a_{tot}/V)}{1 + (c_0/a_0)}, \quad k_{off} = \left(\frac{1}{t_C}\right) \ln\left(\frac{c_0}{c(t_C)}\right) \tag{4}$$

where t_C is the migration time of the maximum concentration of intact C to the detector. Ratios a_{tot}/V and b_{tot}/V are usually known, since they are the initial concentrations of A and B immediately after mixing A and B and before the appreciable formation of C. These nonequilibrium concentrations should not be confused with the equilibrium ones (A_0 and B_0) that are unknown in the NECEEM setup. Expressions 4 allows one to find K_d and k_{off}, since the value of a_0/c_0, $c_0/c(t_C)$, and t_C can be determined using a NECEEM electropherogram. Provided that values of K_d and k_{off} are found, the rate constant k_{on} can be calculated directly, using the second relation in eq 1.

Expressions 4 contains total amounts (a_0, c_0, and $c(t_C)$) of components A and C, whereas the experimental data, such as electropherograms, operate with signals (optical, electrochemical, etc.). To express a_0, c_0, and $c(t_C)$ in terms of measurable areas of characteristic regions of the electropherograms, we assume that signals of visible components are proportional to their concentrations:

$$S_A = g_A A, \quad S_C = g_C C \left(\text{where } g_A = \frac{Q_A}{\chi_A}, g_C = \frac{Q_C}{\chi_C}\right) \tag{5}$$

In relations 5, S_A and S_C are the signals generated by A and C, respectively; A and C represent the local concentrations of A and C (defined as their amounts per unit length of the capillary), respectively. In the case of fluorescence detection, the coefficients g_A and g_C are given by the expressions in the brackets in relations 5, where Q_A and Q_C are the absolute quantum yields of A and C, and χ_A and χ_C are proportionality coefficients.[20,21] As previously noted, the cumulative signal S in a NECEEM electropherogram (see Figure 1a) can be presented as a sum of the three component signals (see Figure 1b):

$$S = S_{A,plug} + S_{A,diss} + S_C \tag{6}$$

Relations 5 and 6 allow the ratios a_0/c_0 and $c_0/(c(t_C))$ to be expressed in terms of measurable characteristic features of the cumulative signal S (see the Supporting Information):

$$\frac{a_0}{c_0} = \frac{\Omega(A) - \Omega(A^*)}{G(\Omega(C) - \Omega(C^*)) + \Omega(D) + \Omega(A^*) + \Omega(C^*)} \tag{7}$$

$$\frac{c_0}{c(t_C)} = 1 + \frac{\Omega(D) + \Omega(A^*) + \Omega(C^*)}{G(\Omega(C) - \Omega(C^*))} \tag{8}$$

Here, $\Omega(A)$ is the area of the A peak, and $\Omega(A^*)$ is the area of the overlapping zone A* between the A peak and the zone under $S_{A,diss}$ (see Figures 2a and 2b). Similarly, $\Omega(C)$ is the area of the C peak, and $\Omega(C^*)$ is the area of the overlapping zone C* between the C peak and the zone under $S_{A,diss}$. Finally, $\Omega(D)$ is the area of zone D under $S_{A,diss}$ without the overlapping zones (see Figures 2a and 2b). The A peak and zone A* start at the beginning of signal $S_{A,plug}$, whereas the C peak and zone C* end at the end of signal S_C.

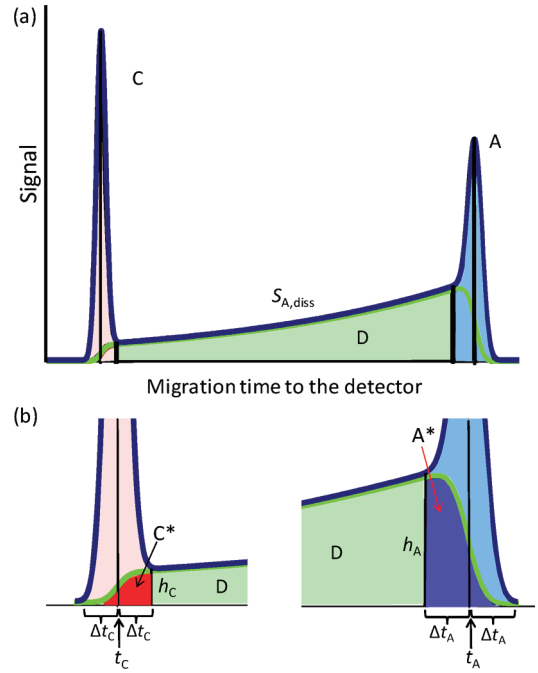


Figure 2. Identification of areas used for calculation of K_d and k_{off}: (a) typical NECEEM electropherogram from Figure 1a with areas $\Omega(A)$, $\Omega(C)$, and $\Omega(D)$ shown in blue, pink, and green colors, respectively (upper boundaries of the overlapping zones A* and C* are shown by the green color); and (b) details of the overlapping zones. Areas $\Omega(A^*)$ and $\Omega(C^*)$ are shown by dark blue and dark red colors, respectively. The time interval from the beginning of the C peak to its maximum (t_C) is denoted by Δt_C. Similarly, Δt_A is the time interval from the maximum of the A peak (t_A) to the its end. The bases of the overlapping zones C* and A* are equal to $2\Delta t_C$ and $2\Delta t_A$, respectively. The heights of the overlapping zones are denoted by h_C and h_A. They coincide with boundaries between zone D and the C and A peaks, respectively.

In eqs 7 and 8, parameter G is determined by the following relation (also see the Supporting Information):

$$G = \frac{g_A v_C}{g_C v_A} = \frac{\chi_C Q_A t_A}{\chi_A Q_C t_C} \tag{9}$$

where the migration times of the A and C peaks (t_A and t_C, respectively) can be readily determined from their positions in NECEEM electropherograms (see the Supporting Information).

Now we need to determine the boundaries of the A and C peaks and $\Omega(A^*)$ and $\Omega(C^*)$. To do this, we assume that peaks $S_{A,plug}$ and S_C are symmetrical (see Figure 1b) and that the time positions of their maxima approximately coincide with those of the maxima of the A and C peaks. In this case, the bases of the A peak and overlapping zone A* are equal to $2\Delta t_A$ (see Figure 2b). Here, Δt_A is a time interval between the end of the A peak and the time t_A of the maximum of A. Similarly, the bases of the C peak and overlapping zone C* are equal to $2\Delta t_C$ (see Figure 2b). Here, Δt_C is the time interval between the time t_C of the maximum of the C peak and its beginning. Obviously, the upper boundaries of the overlapping zones A* and C* coincide with the parts of the signal $S_{A,diss}$ located inside the A and C peaks, respectively (see Figure 2a). Polynomial approximation of these boundaries gives

the following expressions (see the Supporting Information):

$$\Omega(A^*) = \left(\frac{2}{3}\right) h_A \Delta t_A, \quad \Omega(C^*) = \left(\frac{2}{3}\right) h_C \Delta t_C \tag{10}$$

where h_A and h_C are the heights of zones A* and C*, respectively.

In these calculations, we define time positions of the end of the A peak ($t_{A,end}$) and the start of the C peak ($t_{C,start}$) using the following conditions:

$$S - S_0 = 10^{-2} \max_A (S - S_0) \quad \text{at } t = t_{A,\,end}$$
$$= t_A + \Delta t_A \tag{11}$$

$$S - S_0 = 10^{-2} \max_C (S - S_0) \quad \text{at } t = t_{C,\,start}$$
$$= t_C - \Delta t_C \tag{12}$$

where S_0 is the level of the background in a NECEEM electropherogram for S. The determination of Δt_A, Δt_C, h_A, and h_C can be performed manually by measuring these values from graphically recorded (printed) NECEEM electropherograms. However, this determination can be made much more precisely using Microsoft Excel software to process the digitally recorded data (i.e., to find maxima of peaks, their starts and ends, etc.), as was done for all examples in this paper. Finally, $\Omega(A)$, $\Omega(C)$, and $\Omega(D)$ can be easily found from the total signal, since the boundaries between the A and C peaks and the zone D are determined by vertical lines (see Figure 2a). As long as all of the necessary areas can be found, relations 4 and eqs 7−9 allow one to determine K_d and k_{off} easily. A tutorial example of such calculations for a real experimental electropherogram of the AlkB-DNA complex is given in the Supporting Information.

Accuracy of the New Systematic Method for Boundary Positioning. It is instructive to estimate the accuracy of the developed systematic method for boundary positioning. The most effective way for this is to study relative errors in the calculations of a_0, c_0, and $c(t_C)$. Indeed, these component amounts are directly expressed through the characteristic areas (in signals) defined by boundary positioning, whereas values of K_d and k_{off} themselves are determined by a_0, c_0, and $c(t_C)$. To carry out such accuracy analysis, we must analyze NECEEM electropherograms for reaction 1 at various *known* values of K_d, k_{off}, a_0, and c_0. The best way to produce such electropherograms is to simulate them using the exact NECEEM solution for equations of mass transfer (see the Supporting Information).[8,9] This solution holds true as long as the forward reaction is negligible, with respect to the reverse reaction in eq 1. A computer program, utilizing the exact analytical NECEEM solution, allows one to calculate $S_{A,plug}$, $S_{A,diss}$, and S_C, as well as their sum S for varying K_d, k_{off}, a_0, and c_0.[8] Then, the simulated electropherogram (S versus time) can be used to back-calculate a_0, c_0, and also $c(t_C)$, using expressions (S9)−(S12) in the Supporting Information. Since the true values of a_0, c_0, and $c(t_C)$ (where $c(t_C) = \int C(t_C)\,dx$) are known in this case, the corresponding relative errors (an absolute value of the ratio between the deviation from the true value and the true value) are easily determined. After that, relative errors for K_d and k_{off} can be calculated, as shown in the Supporting Information. Alternatively, we can determine values of a_0, c_0, $c(t_C)$, K_d, and k_{off} and the corresponding relative errors, if the visual method for boundary positioning is employed. In this method, overlapping zones are neglected and the boundaries are assumed to be vertical lines positioned visually.

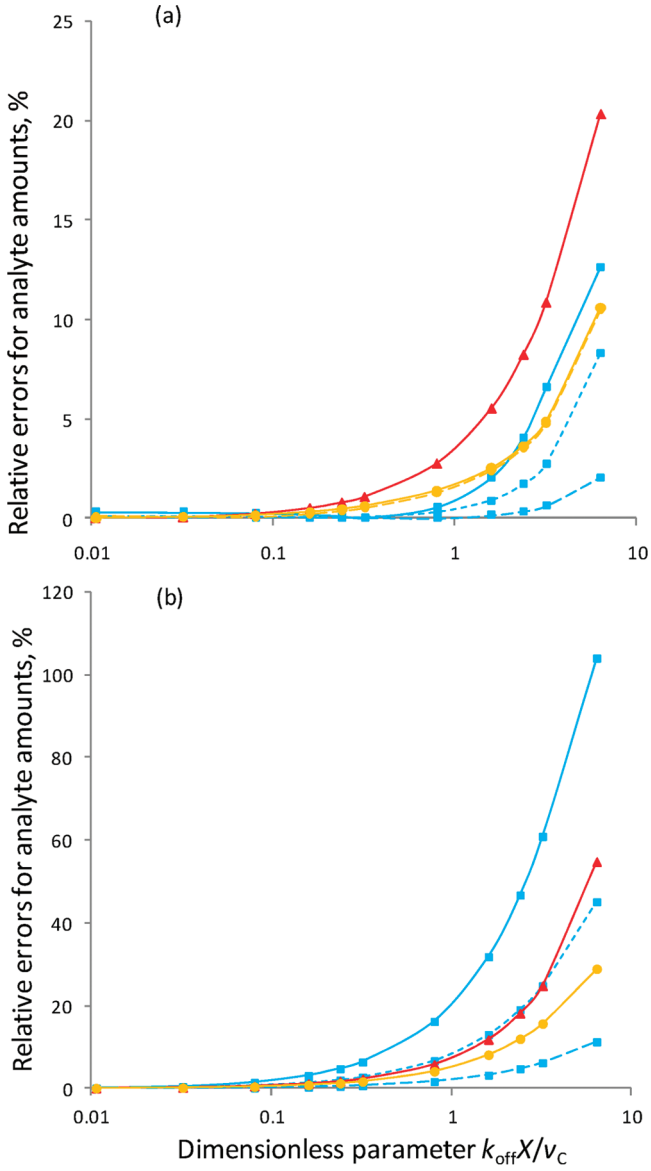


Figure 3. Relative errors in the determination of a_0, c_0, and $c(t_C)$ versus dimensionless parameter $k_{off}X/v_C$ for the (a) systematic and (b) visual methods of boundary positioning. Dependencies for a_0, c_0, and $c(t_C)$ are depicted by blue, orange, and red colors, respectively. Solid lines show errors for ratio $K_d/B_0 = 0.2$. Short and long dashed lines correspond to values $K_d/B_0 = 0.5$ and $K_d/B_0 = 2.0$, respectively. Dashed lines for c_0 and $c(t_C)$ are almost indistinguishable from the corresponding solid lines.

We simulated a total of 33 NECEEM electropherograms mimicking experimental ones. They corresponded to various values of K_d, k_{off}, a_0, and c_0. The values of the velocities, the initial plug width W, and the detector position X were given as

$$v_C = 0.12\text{cm/s}$$

$$v_A = 0.08\text{cm/s}$$

$$W = 0.36\text{cm}$$

$$X = 40\text{cm}$$

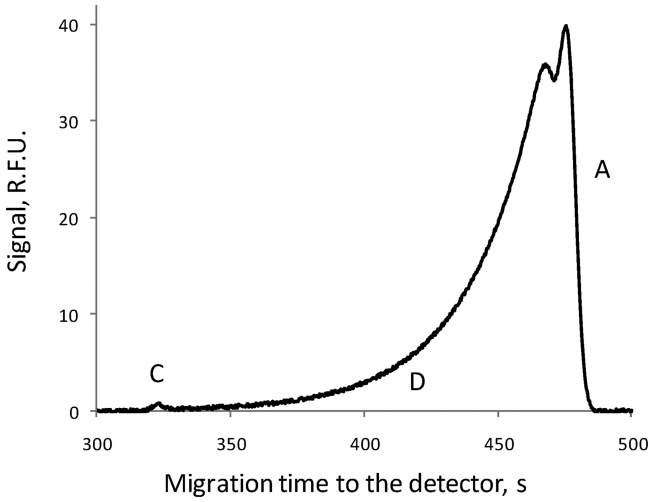


Figure 4. Signal simulated with the exact NECEEM solution at $k_{off}X/v_C$ = 5.75 and K_d/B_0 = 0.2.

which are typical for NECEEM experiments. For simplicity, we assumed that the coefficients in signal−concentration relations are the same: $g_A/g_C = 1$. Some representative simulated electropherograms are shown in the Supporting Information.

Figure 3a shows relative errors in determination of the equilibrium amounts a_0 and c_0 of components A and C in the initial plug if the systematic method for boundary positioning is applied (overlapping zones are taken into account). This figure also demonstrates the relative errors in the determination of the amount $c(t_C)$ of complex C that reaches the detector. The accuracy of k_{off} determination is better than 4% in 32 simulations, and it reaches 14% in only one simulation (see Figure S5 in the Supporting Information). The accuracy of K_d determination is better than 15% in 29 simulations, and it lies in the range of 25%−50% only in four simulations (see Figure S5 in the Supporting Information). In contrast, Figure 3b shows relative errors in the determination of a_0, c_0, and $c(t_C)$ if the visual method for boundary positioning is used (overlapping zones are neglected). In this case, the corresponding values of k_{off} and K_d are found with significantly higher errors that can reach 500% for K_d (see Figure S5 in the Supporting Information). Figure 3 and Figure S5 in the Supporting Information clearly demonstrate the superiority of the developed systematic method to the visual one. At small and moderate values of $k_{off}X/v_C$, the former leads to errors that are several times smaller than those produced by the latter. At higher values of $k_{off}X/v_C$, the systematic method can give up to 10-fold reduction in errors for K_d: from 100%−500% to 25%−50% (Figure S5 in the Supporting Information). Furthermore, the systematic method is not dependent on a visual estimation of boundary position and, therefore, is less prone to user bias.

Note that the appearance of the exponential-like increase in errors in Figure 3 is caused by the logarithmic scale used for the horizontal axes. The actual dependences of the relative errors shown in this figure on $k_{off}X/v_C$ are practically linear. The higher end of the range of $k_{off}X/v_C$ in Figure 3 corresponds to the case of almost complete dissociation of the complex before it reaches the detector. Indeed, eq 3 yields a value of $c(t_C)/c_0 \approx 10^{-4}$ at $k_{off}X/v_C \approx 10$. As a result, the C peak in the simulated signal becomes many times smaller than the A peak that, in turn, acquires the second maximum (see Figure 4). Any further significant increase in $k_{off}X/v_C$ leads to the merger of the maxima of the A peak and to a fast decrease in the height of zone D at almost all distances from the A peak. As a result, the A peak tends to become almost symmetrical. Such behavior of the electropherogram leads to an increase in relative error in the determination of a_0, c_0, and $c(t_C)$ at the upper limit of the $k_{off}X/v_C$ range (recall Figure 3). In general, errors can result from (i) the approximate nature of modeling of the overlapping zone boundaries by the quadratic functions, (ii) the uncertainty in positioning the start and the end of the C and A peaks, respectively, and (iii) errors in the determination of peak maxima based on a single cumulative-signal electropherogram.

CONCLUDING REMARKS

In this work, we introduced a systematic method for finding boundary positions between three characteristic parts of nonequilibrium capillary electrophoresis of equilibrium mixture (NECEEM) electropherograms (i.e., between two peaks and the bridge between them). This method also takes into account the overlap of signals that are responsible for the presence of the peaks and bridge. The developed approach is based on (i) an assumption of the symmetry of peaks if they were not affected by the complex dissociation and (ii) an approximation of the upper boundaries of overlapping zones by quadratic functions. These assumptions allow one to determine all areas in NECEEM electropherograms that are required in a parameter-based approach to calculate the K_d and k_{off} values from the expressions given in relations 4. We tested the accuracy of the method of boundaries positioning in NECEEM by applying it to 33 label-propagation patterns simulated with the exact NECEEM solution. We found that the method's accuracy in the determination of k_{off} and K_d is better than 4% and 15%, respectively, except one simulation that gave a 14% error for k_{off} and four simulations that gave errors in the range of 25%−50% for K_d (from a total of 33 simulations). The developed method for rational boundary determination in NECEEM will facilitate accurate data analysis in this simple and efficient method.

ASSOCIATED CONTENT

Supporting Information. Mathematical derivations, results, and materials and methods. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 416-736-2100, ext 22345. Fax: 416-736-5936. E-mail: skrylov@yorku.ca.

REFERENCES

(1) (a) Bell, S.; Dutta, A. *Annu. Rev. Biochem.* **2002**, *71*, 333–374. (b) Kohn, K. *Mol. Biol. Cell* **1999**, *10*, 2703–2734. (c) Nooren, I.; Thornton, J. *EMBO J.* **2003**, *22*, 3486–3492. (d) Ullrich, A.; Schlessinger, J. *Cell* **1990**, *61*, 203–212. (e) Vonhippel, P.; Bear, D.; Morgan, W.; McSwiggen *J. Annu. Rev. Biochem.* **1984**, *53*, 389–446.

(2) (a) Hashimoto, M.; Rockenstein, E.; Crews, L.; Masliah, E. *Neuromolecular Med.* **2003**, *4*, 21–36. (b) Prusiner, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13363–13383.

(3) Wilson, W. D. *Science* **2002**, *295*, 2103–2105.

(4) (a) Hornblower, B.; Coombs, A.; Whitaker, R. D.; Kolomeisky, A.; Picone, S. J.; Meller, A.; Akeson, M. *Nat. Methods* **2007**, *4*, 315–317. (b) Al-Soufi, W.; Reija, B.; Novo, M.; Felekyan, S.; Kuhnemuth, R.;

Seidel, C. A. M. *J. Am. Chem. Soc.* **2005**, *127*, 8775–8784. (c) Li, Y.; Augustine, G. J.; Weninger, K. *Biophys. J.* **2007**, *93*, 2178–2187.

(5) (a) Abdiche, Y. N.; Malashock, D. S.; Pinkerton, A; Pons, J. *Anal. Biochem.* **2008**, *377*, 209–217. (b) Rich, R. L.; Myszka, D. G. *Anal. Biochem.* **2007**, *361*, 1–6.

(6) Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124*, 13674–13675.

(7) Berezovski, M.; Nutiu, R.; Li, Y.; Krylov, S. N. *Anal. Chem.* **2003**, *75*, 1382–1386.

(8) Okhonin, V.; Krylova, S. M.; Krylov, S. N. *Anal. Chem.* **2004**, *76*, 1507–1512.

(9) (a) Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2004**, *126*, 7166–7167. (b) Petrov, A.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 17104–17110. (c) Okhonin, V.; Petrov, A. P.; Berezovski, M.; Krylov, S. N. *Anal. Chem.* **2006**, *78*, 4803–4810.

(10) (a) Okhonin, V.; Berezovski, M. V.; Krylov, S. N. *J. Am. Chem. Soc.* **2010**, *132*, 7062–7068. (b) Cherney, L. T.; Krylov, S. N. *Anal. Chem.* **2011**, *83*, 1381–1387.

(11) (a) Wang, H. L.; Li, T. *Anal. Chem.* **2009**, *81*, 1988–1995. (b) Petrov, A. P.; Cherney, L. T.; Dodgson, B.; Okhonin, V.; Krylov, S. N. *J. Am. Chem. Soc.* **2011**, *133*, 12486−12492.

(12) Yang, P.; Mao, Y.; Lee, A. W.-M.; Kennedy, R. T. *Electrophoresis* **2009**, *30*, 457–464.

(13) (a) Sloat, A. L.; Roper, M. G.; Lin, X; Ferrance, J. P.; Landers, J. P.; Colyer, C. L. *Electrophoresis* **2008**, *29*, 3446–3455. (b) Carrasco-Correa, E. J.; Beneito-Cambra, M.; Herrero-Martinez, J. M.; Ramis-Ramos, G. *J. Chromatogr. A* **2011**, *1218*, 2334–2341. (c) Xu, Y. Z.; Feng, X. J.; Du, W.; Liu, X.; Luo, Q. M.; Liu, B. F. *Anal. Chem.* **2008**, *80*, 6935–6941.

(14) Krylov, S. N. *Electrophoresis* **2007**, *28*, 69–88.

(15) Drabovich, A. P.; Okhonin, V.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2007**, *129*, 7260–7261.

(16) 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.

(17) (a) Mendonsa, S. D.; Bowser, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 20–21. (b) Mendonsa, S. D.; Bowser, M. T. *J. Am. Chem. Soc.* **2005**, *127*, 9382–9383. (c) Mosing, R. K.; Mendonsa, S. D.; Bowser, M. T. *Anal. Chem.* **2005**, *77*, 6107–6112. (d) Li, S. F. Y.; Tok, J.; Lai, J.; Leung, T. *Electrophoresis* **2010**, *31*, 2055–2062. (e) Javaherian, S.; Musheev, M. U.; Kanoatov, M.; Berezovski, M. V.; Krylov, S. N. *Nucleic Acids Res.* **2009**, *37*, e62. (f) Kanoatov, M.; Javaherian, S.; Krylov, S. N. *Anal. Chim. Acta* **2010**, *681*, 92–97. (g) Tran, D. T.; Janssen, K. P. F; Pollet, J.; Lammertyn, E.; Anné, J.; Van Schepdael, A.; Lammertyn, J. *Molecules* **2010**, *15*, 1127–1140. (h) Turner, D. J.; Tuytten, R.; Janssen, K. P. F.; Lammertyn, J.; Wuyts, J.; Pollet, J.; Eyckerman, S.; Brown, C.; Kas, K. *Anal. Chem.* **2011**, *83*, 666–670.

(18) (a) Allegri, D.; Mori, G.; Seeber, R. *Analyst* **1996**, *121*, 1359–1365. (b) Shao, X.; Cai, W.; Sun, P.; Zhang, M.; Zhao, G. *Anal. Chem.* **1997**, *69*, 1722–1725. (c) Zhang, Y.; Mo, J.; Xie, T.; Cai, P.; Zou, X. *Anal. Chim. Acta* **2001**, *437*, 151–156.

(19) Goodman, K. J.; Brenna, J. T. *Anal. Chem.* **1994**, *66*, 1294–1301.

(20) Chartier, A.; Georges, J.; Mermet, J. *Chem. Phys. Let.* **1990**, *171*, 347–352.

(21) Parker, C.; Rees, W. *Analyst* **1960**, *85*, 587–600.

SUPPORTING INFORMATION

Method for Determination of Peak Areas in Non-Equilibrium Capillary Electrophoresis of Equilibrium Mixtures

Leonid T. Cherney, Mirzo Kanoatov, and Sergey N. Krylov*

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

Mass transfer equations and their solution

In the NECEEM setup, the mass transfer is determined by the following approximate equations:[1]

$$\frac{\partial A}{\partial t} + v_{\mathrm{A}} \frac{\partial A}{\partial x} - \mu_{\mathrm{A}} \frac{\partial^2 A}{\partial x^2} = k_{\mathrm{off}} C \tag{S1}$$

$$\frac{\partial B}{\partial t} + v_{\mathrm{B}} \frac{\partial B}{\partial x} - \mu_{\mathrm{B}} \frac{\partial^2 B}{\partial x^2} = k_{\mathrm{off}} C \tag{S2}$$

$$\frac{\partial C}{\partial t} + v_{\mathrm{C}} \frac{\partial C}{\partial x} - \mu_{\mathrm{C}} \frac{\partial^2 C}{\partial x^2} = -k_{\mathrm{off}} C \tag{S3}$$

where A, B, and C are concentrations of components A, B, and C, respectively; v_{A}, v_{B}, and v_{C} are velocities of A, B, and C; μ_{A}, μ_{B}, and μ_{C} are diffusion coefficients of A, B, and C. Terms corresponding to the forward reaction A + B → C with rate constant k_{on} are omitted in (S1) – (S3) since they are negligible in NECEEM.[1] Consequent integration of (S3) first over the whole spatial coordinate x and then over time t gives relation (3) from the main text.

Equations (S1) – (S3) have an analytical solution found in the work where the NECEEM theory was originally developed:[1]

$$A = A_{\mathrm{plug}} + A_{\mathrm{diss}} \tag{S4}$$

$$A_{\mathrm{plug}} = \frac{A_0}{2}\left(\operatorname{erf} \frac{W - x + v_{\mathrm{A}} t}{\sqrt{4\mu_{\mathrm{A}} t}} - \operatorname{erf} \frac{v_{\mathrm{A}} t - x}{\sqrt{4\mu_{\mathrm{A}} t}} \right) \tag{S5}$$

$$A_{\mathrm{diss}} = \frac{\varepsilon_{\mathrm{A}} k_{\mathrm{off}} C_0}{v_{\mathrm{C}} - v_{\mathrm{A}}} \exp \frac{k_{\mathrm{off}} (x - v_{\mathrm{A}} t)}{v_{\mathrm{C}} - v_{\mathrm{A}}} \int_0^{1/\varepsilon_{\mathrm{A}}} \left(\int_{(v_{\mathrm{A}} t - x)/\varepsilon_{\mathrm{A}}}^{(v_{\mathrm{C}} t - x)/\varepsilon_{\mathrm{A}}} \exp\left(-(\varphi + \psi)^2 \right) \frac{d\varphi}{\sqrt{\pi}} \right) d\psi \tag{S6}$$

$$C = \frac{C_0 \exp(-k_{\text{off}} t)}{2} \left(\operatorname{erf} \frac{W - x + v_{\text{C}} t}{\sqrt{4 \mu_{\text{C}} t}} - \operatorname{erf} \frac{v_{\text{C}} t - x}{\sqrt{4 \mu_{\text{C}} t}} \right) \tag{S7}$$

$$\varepsilon_{\text{A}} = 2 \sqrt{\frac{\mu_{\text{C}} (x - v_{\text{A}} t) - \mu_{\text{A}} (x - v_{\text{C}} t)}{v_{\text{C}} - v_{\text{A}}}} \tag{S8}$$

Here A_{plug} is the concentration of A which was unbound in the injected plug; A_{diss} is the concentration of A generated by dissociation of C during electrophoresis; A_0, B_0, and C_0 are concentrations of A, B, and C, respectively, in the initial plug; W is the length of the initial plug. Symbol erf in (S5) and (S7) denotes the Gauss error functions. The integral in (S6) can also be solved through such functions.[1] Solution of equation (S2) for B is similar to (S4) – (S6). We do not present it here since it is not used in calculations.

Derivation of expressions for a_0/c_0 and $c_0/c(t_{\text{C}})$

Obviously, the total amount of component A cannot be changed by its migration. As a result, we obtain the following expression for a_0:

$$a_0 = \int A_{\text{plug}} dx = v_{\text{A}} \int A_{\text{plug}} dt = \frac{v_{\text{A}}}{g_{\text{A}}} \int_{\mathbf{A}} S_{\text{A,plug}} dt = \frac{v_{\text{A}}}{g_{\text{A}}} \int_{\mathbf{A}} (S - S_{\text{A,diss}}) dt = \frac{v_{\text{A}}}{g_{\text{A}}} \left(\Omega(\text{A}) - \Omega(\text{A*}) \right) \tag{S9}$$

where A_{plug} and A_{diss} denote linear concentrations of A corresponding to unbound A in the injected plug and A formed from dissociation of complex C during electrophoresis, respectively; the third and the fourth integrals are within the time limits of peak A; Ω(A) is the area of peak A; and Ω(A*) is the area of the overlapping zone A* between the peak A and the zone under $S_{\text{A,diss}}$ (**Figs. 2a, b**). In relations (S9), the peak A and zone A* start at the beginning of the signal $S_{\text{A,plug}}$. Similarly, we can derive relations:

$$c(t_{\text{C}}) = \int C dx = v_{\text{C}} \int C dt = \frac{v_{\text{C}}}{g_{\text{C}}} \int_{\mathbf{C}} S_{\text{C}} dt = \frac{v_{\text{C}}}{g_{\text{C}}} \int_{\mathbf{C}} \left(S - S_{\text{A,diss}} \right) dt = \frac{v_{\text{C}}}{g_{\text{C}}} \left(\Omega(\text{C}) - \Omega(\text{C*}) \right) \tag{S10}$$

$$a_{\text{diss}}(t_{\text{C}}) = \int A_{\text{diss}} dx = v_{\text{A}} \int A_{\text{diss}} dt = \frac{v_{\text{A}}}{g_{\text{A}}} \int_{\text{A+C+D}} S_{\text{A,diss}} dt = \frac{v_{\text{A}}}{g_{\text{A}}} \left(\Omega(\text{D}) + \Omega(\text{A*}) + \Omega(\text{C*}) \right) \tag{S11}$$

Here $c(t_C)$ is the total amount of C still intact at time t_C; $a_{diss}(t_C)$ is the total amount of A produced by dissociation of C until time t_C; the third and the fourth integrals in (S10) are within the time limits of peak C; Ω(C) is the area of peak C, and Ω(C*) is the area of the overlapping zone C* between the peak C and the zone under $S_{A,diss}$; Ω(D) is the area of the zone D under $S_{A,diss}$ without the overlapping zones; and the third integral in (S11) is within the time limits of peaks A and C and zone D (**Figs. 2a, b**). In relations (S10) the peak C and zone C* end at the end of the signal S_C. Since the total amount of component C decreases with time due to dissociation of C, a simple expression for c_0 takes place:

$$c_0 = c(t_C) + a_{diss}(t_C) \tag{S12}$$

Relations (S9) – (S12) express a_0, $c(t_C)$, $a_{diss}(t_C)$, and c_0 in terms of areas Ω(A), Ω(C), Ω(A*), Ω(C*), and Ω(D) that characterize a typical NECEEM electropherogram.

Evidently, components A and C cannot directly contribute to the D region because the latter, by definition, does not include the overlapping zones A* and C* (see Fig. 2). However, the indirect contribution of A to the D region is taken into account through A_{diss} that results from the dissociation of C. Finally, there cannot be an indirect contribution of C to the D region since the component C is entirely located in zone C (see Figs 1 and 2). It should also be emphasized that though the D region does not include overlapping the zones A* and C*, these zones are accounted for during the derivation of expressions (S9), (S10), and (S12) for a_0, b_0, and c_0 and, therefore, in calculations of K_d and k_{off}.

It is worth noting, that expressions (4) for K_d and k_{off} depend only on ratios a_0/c_0 and $c_0/c(t_C)$, respectively, rather than on total amounts a_0, c_0, and $c(t_C)$. Using relations (S9), (S11), and (S12) we readily obtain:

$$\frac{a_0}{c_0} = \frac{\Omega(\mathrm{A}) - \Omega(\mathrm{A}^*)}{G\left(\Omega(\mathrm{C}) - \Omega(\mathrm{C}^*)\right) + \Omega(\mathrm{D}) + \Omega(\mathrm{A}^*) + \Omega(\mathrm{C}^*)} \tag{S13}$$

$$\frac{c_0}{c(t_C)} = 1 + \frac{\Omega(\mathrm{D}) + \Omega(\mathrm{A}^*) + \Omega(\mathrm{C}^*)}{G\left(\Omega(\mathrm{C}) - \Omega(\mathrm{C}^*)\right)} \tag{S14}$$

where the parameter G is determined by relations:

$$G = \frac{g_A v_C}{g_C v_A} = \frac{\chi_C Q_A t_A}{\chi_A Q_C t_C} \tag{S15}$$

The first equality in (S15) is just a definition of G, whereas the second equality follows from expressions (5) for g_A and g_C and expressions for the migrations times t_A and t_C of peaks A and C:

$$t_A = \frac{X}{v_A}, \quad t_C = \frac{X}{v_C} \tag{S16}$$

Here X is the distance to the detector from the position of the plug of equilibrium mixture at time zero. The values of t_A and t_C can be determined from position of peak A and C in NECEEM electropherograms. According to expression (S15), G depends on the relative quantum yield Q_A/Q_C rather than on absolute quantum yields of A and C. Furthermore, coefficients χ_A and χ_C cancel each other out if the same detector is used in measurements of signals from A and C (in this case $\chi_A = \chi_C$).

Derivation of expressions for Ω(A*) and Ω(C*)

Given that zones A* and C* do not have bell-like Gaussian shape, we can approximate the upper boundaries of A* and C* in Fig. 2 by polynomials. Our calculations show that the use of polynomials of the third and higher orders does not result in significant improvement of accuracy in the determination of overlapping zones, though it does make the procedure more complicated. Thus, we can approximate the upper boundaries of A* and C* by quadratic functions. According to Fig. 2, such functions should reach their minima at the end of peak A and at the beginning of peak C, and can be written as in equations (S17).

$$S_A^*(t) = \frac{h_A}{4\Delta t_A^2}(t - t_A - \Delta t_A)^2, \quad S_C^*(t) = \frac{h_C}{4\Delta t_C^2}(t - t_C + \Delta t_C)^2 \tag{S17}$$

Here h_A and h_C are the heights of zone A* and C*, respectively. The value of h_A is equal to the total signal S at $t = t_A - \Delta t_A$ since $S_{A,plug} = 0$ at the start of A*. Similarly, h_C is equal to the total signal S at $t = t_C + \Delta t_C$ since $S_C = 0$ at the end of C*. The function $S_A^*(t)$ is the simplest polynomial reaching a value of h_A at the beginning of peak A and smoothly vanishing at its end. Similarly, $S_C^*(t)$ is the simplest polynomial reaching a value of h_C at the end of peak C and smoothly vanishing at its start. As a

result, we finally obtain simple expressions for Ω(A*) and Ω(C*) in terms of easily measurable quantities Δt_A, h_A and Δt_C, h_C:

$$\Omega(\mathrm{A}^*) = \int_{t_A-\Delta t_A}^{t_A+\Delta t_A} S_A^*(t) = \frac{2}{3} h_A \Delta t_A, \quad \Omega(\mathrm{C}^*) = \int_{t_C-\Delta t_C}^{t_C+\Delta t_C} S_C^*(t) = \frac{2}{3} h_C \Delta t_C \tag{S18}$$

Example of boundary positioning and determination of characteristic areas and constants K_d and k_{off}.

Below, we illustrate the application of boundaries determination for an experimental electropherogram shown in **Fig. S1**.

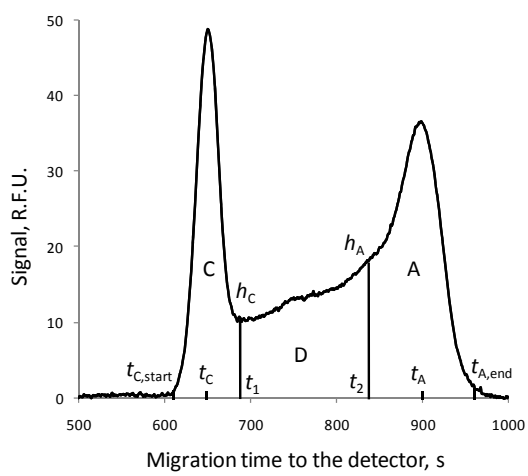


Figure S1. Electropherogram used in an example of boundary determination and area calculations (S19) – (S25). Complex C is formed by DNA (component A) and AlkB protein (component B).

The electropherograms was obtained for the interaction between AlkB protein and its fluorescently-labeled aptamer, developed elsewhere.[16] For peaks A and C in this electropherogram the procedure described in **Fig. 2** gives at $S_0 = 0.6$ (hereafter we omit signal units, R.F.U.):

$$\max_{\mathrm{A}}(S - S_0) = 36.07 \quad \text{at} \quad t = t_A = 899 \text{ s} \tag{S19}$$

$$10^{-2} \max_{\mathrm{A}}(S - S_0) = 0.36 \quad \text{at} \quad t = t_{\mathrm{A,end}} = 967 \text{ s} \tag{S20}$$

$$\max_{\mathrm{C}}(S - S_0) = 48.28 \quad \text{at} \quad t = t_{\mathrm{C}} = 651 \text{ s} \tag{S21}$$

$$10^{-2} \max_{\mathrm{C}}(S - S_0) = 0.48 \quad \text{at} \quad t = t_{\mathrm{C,start}} = 612 \text{ s} \tag{S22}$$

Therefore, we have:

$$\Delta t_{\mathrm{A}} = t_{\mathrm{A,end}} - t_{\mathrm{A}} = 68 \text{ s}, \quad \Delta t_{\mathrm{C}} = t_{\mathrm{C}} - t_{\mathrm{C,start}} = 39 \text{ s} \tag{S23}$$

The left and right boundaries, t_1 and t_2, of the zone D (i.e. boundaries between the zone D and the peaks C and A, respectively) are positioned at (**Fig. S1**):

$$t_1 = t_{\mathrm{C}} + \Delta t_{\mathrm{C}} = 690 \text{ s}, \quad t_2 = t_{\mathrm{A}} - \Delta t_{\mathrm{A}} = 831 \text{ s} \tag{S24}$$

The height h_{A} and h_{C} are determined as signals at these times:

$$h_{\mathrm{A}} = S(t_2) = 16.84, \quad h_{\mathrm{C}} = S(t_1) = 9.63 \tag{S25}$$

For the left and right boundaries of zone D positioned in **Fig. S1** at times determined in (S24), calculations of areas in this figure give $\Omega(\mathrm{A}) = 2688.09$, $\Omega(\mathrm{C}) = 1721.46$, and $\Omega(\mathrm{D}) = 1786.77$. Substituting values (S23) and (S25) into expressions (S18) we find areas of the zones A* and C*: $\Omega(\mathrm{A}^*) = 763.53$ and $\Omega(\mathrm{C}^*) = 250.28$. After $\Omega(\mathrm{A})$, $\Omega(\mathrm{C})$, $\Omega(\mathrm{A}^*)$, $\Omega(\mathrm{C}^*)$, and $\Omega(\mathrm{D})$ are found, expressions (S13)–(S15) allow us to determine ratios $a_0/c_0 = 0.40$ and $c_0/c(t_{\mathrm{C}}) = 2.39$. In these

calculations, the values of t_A and t_C were defined by (S19) and (S21), respectively. We used a single detector and, therefore, $\chi_A = \chi_C$. A value of $Q_A/Q_C = 0.99$ was measured in an independent experiment (see the last section). As a result, (S15) leads to $G = 1.37$. Finally, substitutions of obtained ratios a_0/c_0, and $c_0/c(t_{C,max})$ into expressions (4) from the main text give $K_d = 39$ nM and $k_{off} = 1.3\times10^{-3}$ s^{-1} for the following values of two other parameters present in (4): $a_{tot}/V = 30$ nM, $b_{tot}/V = 120$ nM. The latter values are defined by the equilibrium mixture preparation and injection stages of NECEEM.

Simulation of NECEEM electropherograms

Explicit expressions (S5) - (S7) for concentrations A_{plug}, A_{diss}, and C were employed to simulate signals

$$S_{A,plug} = g_A A_{plug}, \quad S_{A,diss} = g_A A_{diss}, \quad S_C = g_C C \tag{S26}$$

Calculations were performed in Excel software. We simulated a total of 33 NECEEM electropherograms mimicking experimental ones. They were used in studies of the accuracy of finding a_0, c_0, $c(t_C)$, K_d, and k_{off} in a NECEEM parameter-based approach that employs the developed method for boundary positioning. Below we show a few typical simulated electropherograms in the presence of noise (2 % of the signal maximum) that was added to the calculated cumulative signal.

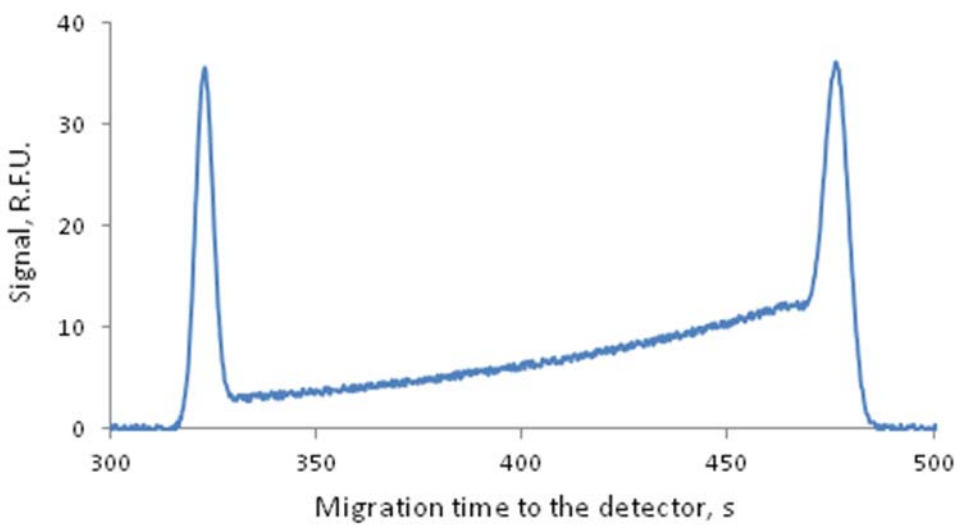


Figure S2. Simulated electropherogram at $k_{off} = 5.0 \times 10^{-3}$ s and $K_d/B_0 = a_0/c_0 = 0.2$. Values of $a_{tot}/V = 360$ nM and $b_{tot}/V = 360$ nM were used in calculations of K_d.

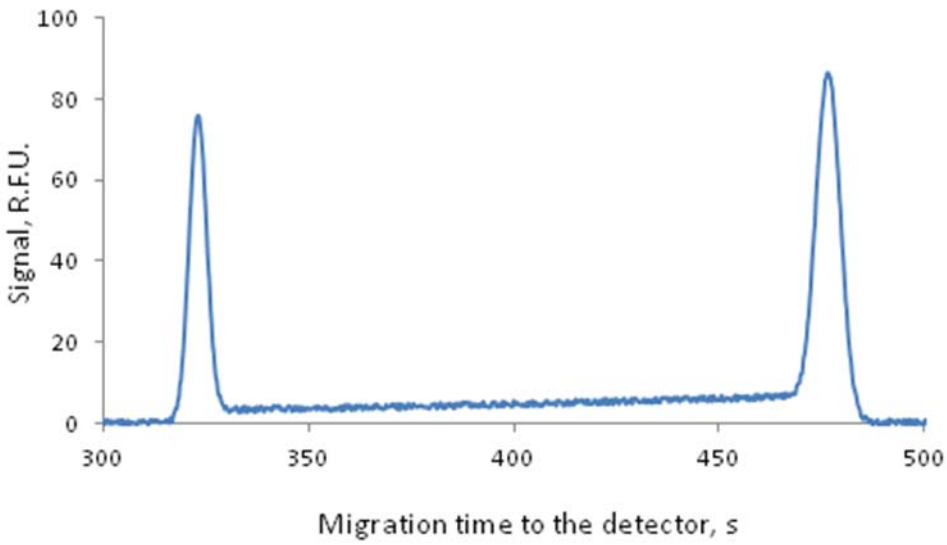


Figure S3. Simulated electropherogram at $k_{off} = 2.5 \times 10^{-3}$ s and $K_d/B_0 = a_0/c_0 = 0.5$. Values of $a_{tot}/V = 450$ nM and $b_{tot}/V = 450$ nM were used in calculations of K_d.

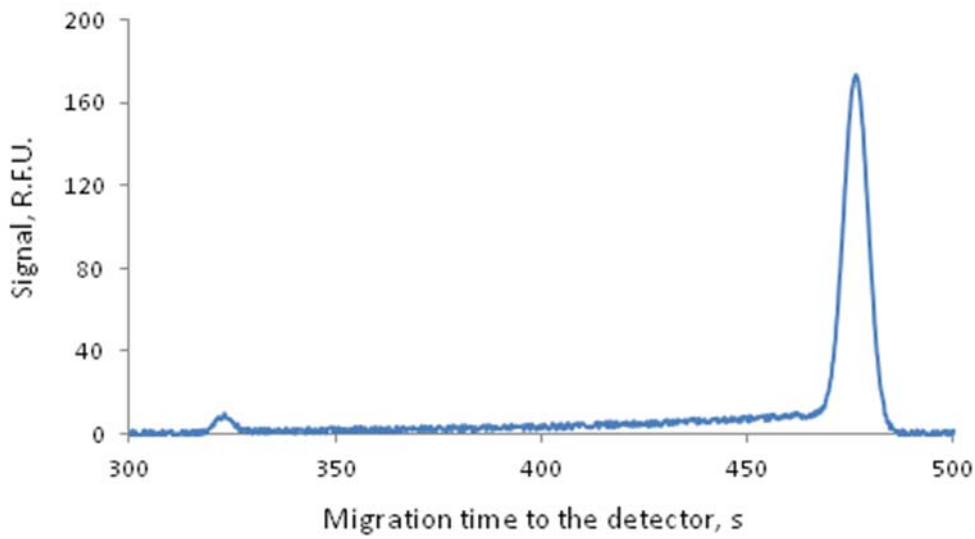


Figure S4. Simulated electropherogram at $k_{off} = 7.5 \times 10^{-3}$ s and $K_d/B_0 = a_0/c_0 = 2.0$. Values of $a_{tot}/V = 450$ nM and $b_{tot}/V = 250$ nM were used in calculations of K_d.

Accuracy of K_d and k_{off} determined with the new method for boundary positioning

Relative errors in determination of K_d and k_{off} based on (4) are shown in **Figure S5**. Calculations of relative errors were performed in a wide range of the dimensionless parameter $k_{off}X/v_C$ and at three different values of ratio K_d/B_0. The following values of concentrations a_{tot}/V and b_{tot}/V were used in

simulations and in expression (4) for K_d: $a_{tot}/V = b_{tot}/V = 360$ nM for $K_d/B_0 = 0.2$, $a_{tot}/V = b_{tot}/V = 450$ nM for $K_d/B_0 = 0.5$, and $a_{tot}/V = 450$ nM, $b_{tot}/V = 250$ nM for $K_d/B_0 = 2.0$.

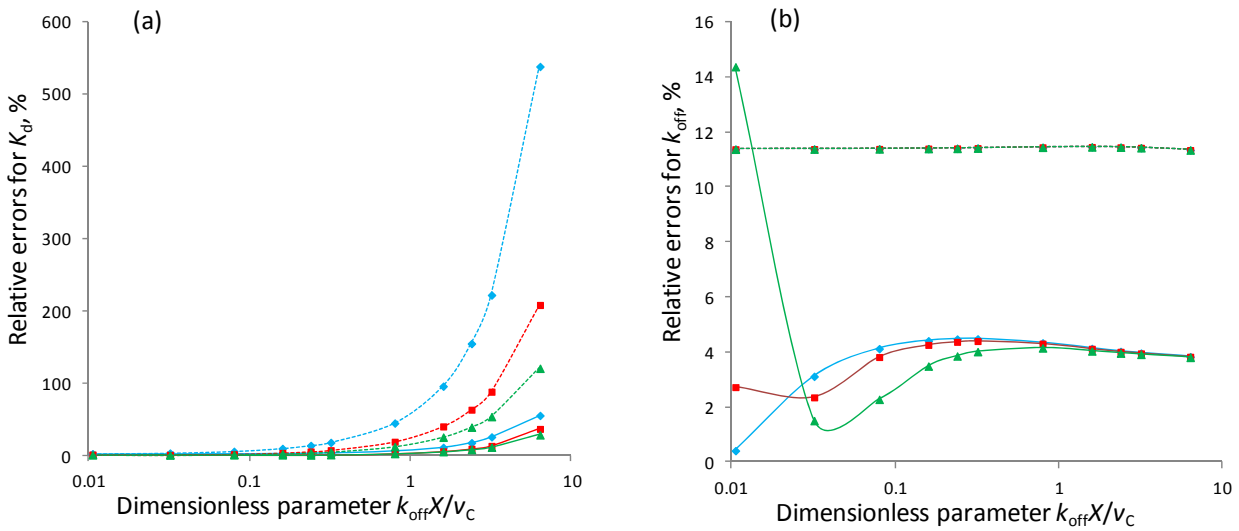


Figure S5. Relative errors in determination of K_d (**a**) and k_{off} (**b**) as functions of dimensionless parameter $k_{off}X/v_C$ at three different values of K_d/B_0 for the systematic (solid lines) and visual (dot lines) methods of boundary positioning. Blue, red, and green lines correspond to K_d/B_0 = 0.2, 0.5, and 2.0, respectively. Colored dot lines are practically undistinguishable in **Fig. S5b**. The appearance of the exponential-like increase in errors for K_d is caused by the logarithmic scale used for the horizontal axes. The actual dependences of the relative errors shown in **Fig.S5a** on $k_{off}X/v_C$ are practically linear. The higher end of the range of $k_{off}X/v_C$ corresponds to the case of almost complete dissociation of the complex before it reaches the detector ($c(t_C)/ c_0 \sim 10^{-4}$ at $k_{off}X/v_C \sim 10$).

The increase in relative errors in determination of a_0, c_0, and $c(t_C)$ at higher values of $k_{off}X/v_C$ (**Fig. 3**) leads to a corresponding increase in determination of K_d. Nevertheless, throughout the entire range of $k_{off}X/v_C$ the systematic method of boundary positioning gives significantly smaller errors for K_d than does the visual method (**Fig. S5a**). The relative errors for k_{off} remain fairly small at all values of $k_{off}X/v_C$ (**Fig. S5b**). This can be explained by a partial compensation of errors for c_0 and $c(t_C)$ in calculations of k_{off} based on relation (4) and also by the presence of the logarithmic function in expression (4) for k_{off}. Interestingly, errors for k_{off} do not practically depend on K_d/B_0 if the visual method is used for boundary positioning and these errors start to depend on K_d/B_0 only at small values of k_{off} if the systematic method is applied. Again, in all simulations except for one, the systematic method results in approximately three times smaller errors for k_{off} than the visual one does (**Fig. S5b**).

Materials and methods

All capillary electrophoresis (CE) procedures were performed with a P/ACE MDQ apparatus (Beckman Coulter, Mississauga, ON) equipped with a fluorescence detector and a 488-nm Solid-State

laser (JDSU, Santa Rosa, CA). An uncoated fused-silica capillary with an inner diameter of 75 µm and outer diameter of 360 µm was used (Polymicro, Phoenix, AZ). Experiments were performed in a 50-cm-long (40 cm to the detection window) capillary. Both the inlet and the outlet reservoirs contained the electrophoresis run buffer (25 mM Borax at pH 9.2). Prior to every run, the capillary was rinsed with the run buffer solution at 20 psi pressure for 1 min. At the end of each run, the capillary was rinsed with 100 mM HCl, 100 mM NaOH, and deionized water, with 20 psi for 1 minute each. The samples were injected into the capillary by a pressure pulse of 6.5 s at 0.5 psi. The length of the sample plug was calculated to be 6 mm. Before carrying out electrophoresis, a plug of running buffer was injected into the capillary for 80 s at 0.3 psi. This was done to move the equilibrium mixture plug by 5 cm along the capillary, to avoid the effects of the inefficiently cooled region of the capillary.[2] This step reduced the travel length to the detector from 40 cm to 35 cm. The electric field of 500 V/cm was then applied across the capillary. This is the point that corresponds to 0 time on the electropherogram in **Fig. S1**. Electrophoresis was carried out with a positive electrode at the injection end of the capillary; the direction of the electroosmotic flow was from the inlet to the outlet reservoir.

To obtain the NECEEM experimental data, the interaction between AlkB protein from *E. coli* and a DNA aptamer were studied. Details on the protein and aptamer selection process are described elsewhere.[3] The aptamer sequence was the following: 5'-/FAM/-CTCCTCTGACTGTAACCACGTGCCTAGCGTTTCATTGTCCCTTCTTATTAGGTGATAATAGC ATAGGTAGTCCAGAAGCC-3'. Equilibrium mixtures were prepared in 25 mM Borax at pH 9.2, and contained 120 nM AlkB protein and 30 nM DNA aptamer. The mixture was incubated at 15 °C for 5 minutes before being injected into the capillary, to allow for equilibration of the components. The DNA aptamer was fluorescently labelled for detection. A representative result of one of such experiments is shown in **Fig. S1**.

To determine the Q_A/Q_C ratio, Equilibrium Capillary Electrophoresis of Equilibrium Mixtures (ECEEM) experiments were performed.[4] In practice, ECEEM and NECEEM differ only in the fact that in ECEEM the unlabelled component (AlkB protein in this case) is present in the run buffer, throughout the capillary. This way, the labeled component (aptamer), is always in a state of pseudo-equilibrium. If the unlabelled component is taken in sufficient excess, all of the labeled component will be in the bound state, at all times, and will arrive to the detector as a single peak. Two types of sample mixtures were analyzed and compared. In the first set of experiments, mixtures containing 1 µM AlkB and 30 nM aptamer were analyzed. For these experiments, the CE run buffer also contained 1 µM AlkB. The peak area produced by these experiments allowed the determination of Q_C. The second set of sample mixtures only contained 30 nM of the aptamer. In these cases, neither the sample mixture, nor the running buffer contained any AlkB protein. All of the labeled component was thus unbound at all times and arrived at the detector as a single peak. The signal area produced by these experiments allowed the determination of Q_A. Due to the fact that only a single peak reached the detector in both cases, ECEEM allowed us to determine Q_A/Q_C ratio without assigning any inter-peak boundaries. All mixtures also contained 30 nM rhodamine 110, to act as a signal normalization standard between different electropherograms. Other than the described modifications, all of the other conditions were identical to ones used for NECEEM experiments. **Figure S6**. depicts a representative set of results. The blue trace shows the signal produced

by free aptamer, and the red trace shows the signal produced by aptamer fully bound to an excess of the AlkB protein. As revealed by the red trace, the aptamer sample contains a fluorescent impurity that does not interact with the protein, but contributes to the total fluorescence signal associated with the free-aptamer. The contaminant nature of this peak was confirmed by titration experiments. To find the Q_A/Q_C ratio, areas of fluorescent signals produced by free and bound aptamer were compared. The signal area contributed by the impurity was subtracted from total area of the free aptamer peak before obtaining the ratio. Q_A/Q_C for this molecular pair was found to be 0.99 ± 0.02.

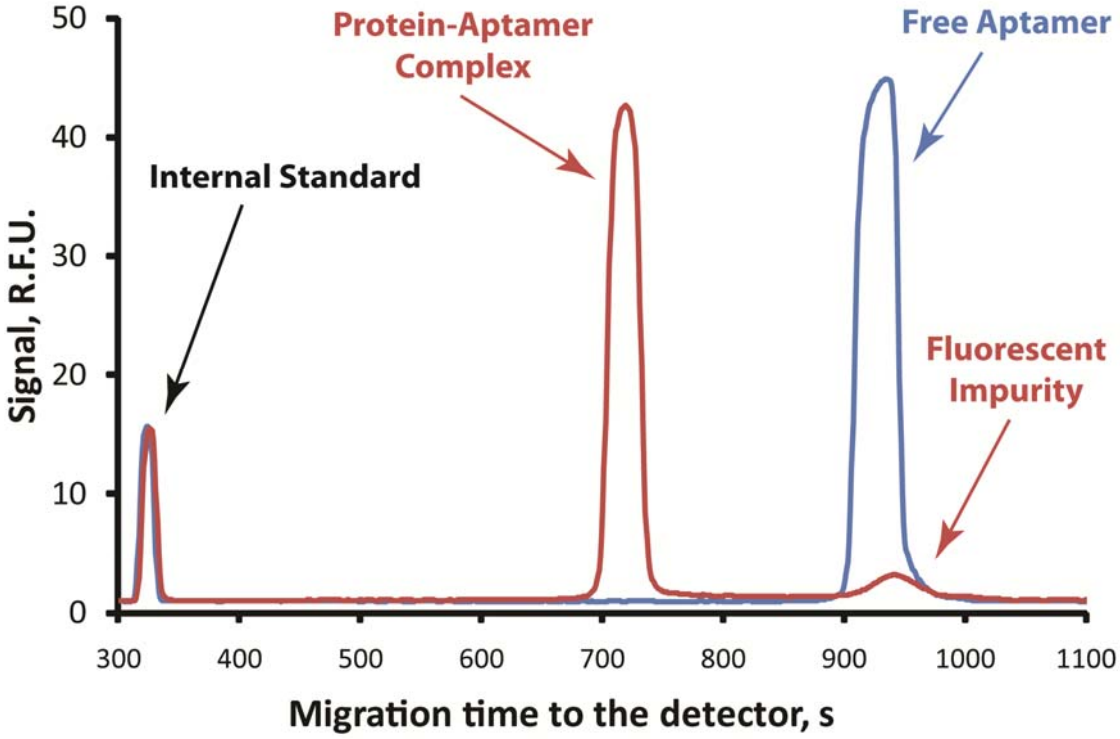


Figure S6. ECEEM measurements of samples containing 30 nM aptamer in the presence (red trace) and the absence (blue trace) of 1 μM AlkB protein in the run buffer. 30 nM of rhodamine 110 was used as an internal standard.

References

1. Okhonin, V.; Krylova, S.M., Krylov, S.N. *Anal. Chem.* **2004**, *76*, 1507-1512.
2. Krylov, S. N.; Musheev, M. U.; Filiptsev, Y. *Anal. Chem.* **2010**, *82,* 8692-8695.
3. 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.
4. Drabovich, A.; Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2005**, *127*, 11224-11225.