JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

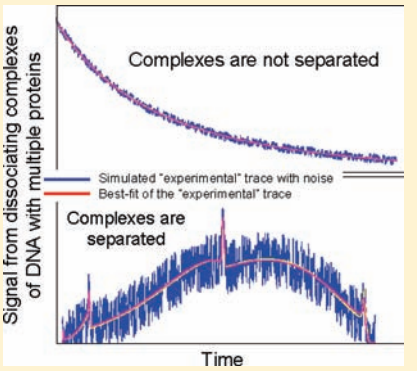
Separation-Based Approach to Study Dissociation Kinetics of Noncovalent DNA−Multiple Protein Complexes

Alexander P. Petrov, Leonid T. Cherney, Bryan Dodgson, Victor Okhonin, and Sergey N. Krylov*

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

Supporting Information

ABSTRACT: Noncovalent binding of DNA with multiple proteins is pivotal to many regulatory cellular processes. Due to the lack of experimental approaches, the kinetics of assembly and disassembly of DNA−multiple proteins complexes have never been studied. Here, we report on a first method capable of measuring disassembly kinetics of such complexes. The method is based on continuous spatial separation of different complexes. The kinetics of multiple complex dissociation processes are also spatially separated, which in turn facilitates finding their rate constants. Our separation-based approach was compared with a conventional no-separation approach by using computer simulation of dissociation kinetics. It proved to be much more accurate than the no-separation approach and to be a powerful tool for testing hypothetical mechanisms of the disassembly of DNA−multiple proteins complexes. An experimental implementation of the separation-based approach was finally demonstrated by using capillary electrophoresis as a separation method. The interaction between an 80 nucleotide long single-stranded DNA and single-stranded DNA binding protein was studied. DNA−protein complexes with one and two proteins were observed, and rate constants of their dissociation were determined. We foresee that a separation approach will be also developed to study the kinetics of the formation of DNA−multiple protein complexes.



INTRODUCTION

Noncovalent binding of a single DNA molecule with multiple proteins is common in biology and plays a pivotal role in regulation of gene expression, DNA replication, DNA integrity control, and virus replication.[1,2] In order to understand the dynamics of these fundamental biological processes, it is important to know kinetic parameters for all steps involved in the formation and dissociation of the relevant DNA−multiple protein complexes.[3−5] Proteins in these complexes can be bound to the DNA directly or indirectly through other proteins. Our knowledge of DNA−multiple protein complexes is typically limited to the identities of the DNA and proteins involved.[6] Complete kinetic analyses are rarely performed for DNA interaction with a single protein,[7] and to the best of our knowledge, kinetics of formation and/or dissociation of DNA−multiple protein complexes have never been measured. The lack of comprehensive kinetic studies is solely due to a lack of experimental approaches capable of distinguishing kinetics of the multiple interconnected processes involved in assembly/disassembly of DNA−multiple protein complexes. The present work was motivated by the insight that, in general, the kinetics of processes occurring during the formation and/or dissociation of DNA−multiple proteins complexes in vitro can be distinguished if different complexes move with different velocities or in other words are continuously spatially separated. Here, we present the first separation-based approach for studying kinetics of dissociation of DNA−multiple protein complexes.

In general, to study the dissociation kinetics of DNA−-multiple protein complexes, the following two-step operation should be performed. In the first step (equilibration step), multiple DNA−protein complexes are formed by incubating free DNA with proteins, ideally, long enough to approach equilibrium. In the second step (dissociation step), unbound proteins are continuously removed from the complexes so that the rates of complex formation become zero and the complexes are forced to dissociate. The difficulty of analyzing dissociation kinetics originates from multiple kinetic processes (i) occurring simultaneously and (ii) being indistinguishable from the detection standpoint. When dissociation is initiated by removing free proteins, all DNA−protein complexes start dissociating simultaneously, which results in multiple overlapping single-exponential kinetic curves of the same nature. The resulting signal is the sum of all individual kinetics as conceptually shown in Figure 1, left. Many processes can be described by similar modeles.[8,9] It is a well-known problem that the sum of single-exponential curves has a shape close to exponential, and in many cases, such a sum cannot be used to reliably determine powers of individual exponents comprising it. The inaccuracy of the no-separation approach is caused by the inherent instability of the inverse problem of finding powers of individual exponents from their sum. This instability is with respect to small perturbations of the initial exponential sum.[10,11] Such instability and the presence of

Received: July 29, 2010
Published: July 16, 2011

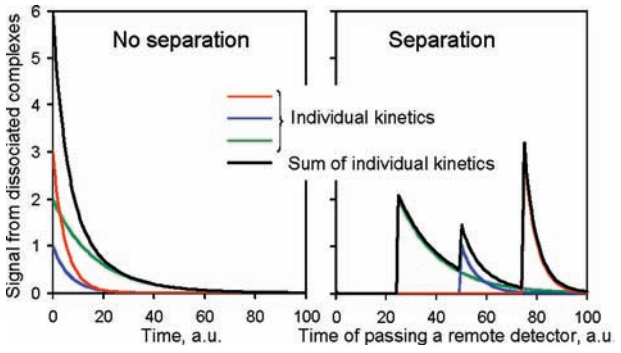


Figure 1. Schematic illustration of simultaneous dissociation kinetics of DNA−multiple protein complexes without separation (left) and with separation of the complexes. In each panel, three color traces show three individual single-exponential kinetics, and a black trace shows their sum.

experimental noise can lead to significant errors in the determination of the powers (rate constants). Errors can become unacceptably large if some components have close (or very different) powers and/or intensities and if the number of components is unknown (in the other words, the mechanism of reaction is unknown).

The present work was motivated by the idea that the problem of resolving individual kinetics could be solved by simply making different complexes move with different velocities in the same direction and by placing a detector at a distance x from the site of initiation of movement. Single-exponential curves generated at the detection point by the dissociation of complexes moving with different velocities would be shifted with respect to each other as conceptually shown in Figure 1, right. It is clear that the sum of such shifted kinetics is more "informative" about its components than the sum of the nonshifted kinetics. It is not clear however whether or not this gain can be utilized for the determination of rate constants and for testing hypothetical mechanisms of dissociation. To answer this question, in this proof-of-principle work, we used extensive computer simulation to compare no-separation and separation-based approaches. We examined our method's ability to test hypothetical mechanisms of dissociation. We then demonstrated experimental use of our approach in the study of dissociation kinetics of complexes of DNA with multiple molecules of the single-stranded DNA-binding protein.[12] As a practical means of introducing differential mobilities of different DNA−protein complexes, we used capillary electrophoresis (CE).[13] CE simply provides an efficient way to accomplish the separation-based analysis of simultaneous dissociation processes of DNA−multiple proteins complexes. Our separation approach can potentially be used to study the disassembly kinetics of complex protein machines attached to DNA. We foresee that a separation approach will be also developed to study the kinetics of the formation of DNA−multiple protein complexes.

RESULTS AND DISCUSSION

Mathematical Model. Our goal is to compare the proposed separation-based approach with the no-separation approach in studying dissociation kinetics of DNA−multiple protein complexes. We first specify the difference between the no-separation and separation approaches. In the no-separation approach, all complexes are spatially colocated and dissociation is detected due to "mass loss" upon the dissociated components' leaving the point of localization. Sensor-based methods, such as surface plasmon resonance (SPR),[14] fall in the category of the no-separation approach. To study dissociation of DNA−multiple protein complexes with SPR, one would immobilize DNA on the surface and form different complexes on the surface by adding proteins in solution. Dissociation would be initiated and maintained by continuously removing free proteins from the solution and observing mass loss on the surface due to proteins leaving the surface. The resulting signal is a sum of nonshifted exponential curves similar to those depicted in Figure 1, left.

The separation approach is based on (i) continuous spatial separation of different DNA−multiple protein complexes in solution by introducing their differential mobility in one direction, (ii) dissociation of complexes during their separation, and (iii) following special distribution of DNA. While differential mobility can be caused by different means, electrophoresis is the most practical and best-developed way of mobility shift for DNA−protein complexes. The resulting signal will be conceptually similar to spatially shifted exponential curves depicted in Figure 1, right.

The major disassembly mechanism considered in this work was sequential dissociation from the state of equilibrium (we also considered branched dissociation from the state of equilibrium; mechanism 3 in the Supporting Information). In the first step, N DNA−protein complexes (DP_1, ..., $DP_1...P_N$) are formed by incubating free DNA (D) with N proteins (P_1, ...P_N), ideally, long enough to approach equilibrium:

$$D \overset{P_1, K_{d,1}}{\rightleftharpoons} DP_1 \overset{P_2, K_{d,2}}{\rightleftharpoons} DP_1P_2... \overset{P_N, K_{d,N}}{\rightleftharpoons} DP_1P_2...P_N \qquad (1)$$

where $K_{d,1}$, $K_{d,1}$, ..., $K_{d,N}$ are equilibrium dissociation constants of N sequential processes and the index also denotes the reverse order of dissociation (see eq 2 below). If all proteins are the same, indexes 1, 2, ..., N can be omitted as it is done in mechanisms 1 and 2 below. In the second step, unbound proteins are continuously removed from the complexes so that the rates of the forward processes in reaction 1 become zero and the complexes are forced to dissociate:

$$DP_1P_2...P_N \xrightarrow{k_{off,N}} DP_1P_2P_{N-1} \xrightarrow{k_{off,N-1}} ...DP_1 \xrightarrow{k_{off,1}} D \qquad (2)$$

where $k_{off,N}$, $k_{off,N-1}$, ..., $k_{off,1}$ are dissociation rate constants for the N DNA−protein complexes. The exact mechanism of complex assembly was not a subject of investigation in this work. Therefore, the sole purpose of equilibrium reaction 1 was to define the initial concentrations of complexes before dissociation. Reaction 2 was instead our major concern, and it was investigated using a system of differential equations. To write such a system we define the following terms for complexes:

$$C_0 = D, C_1 = DP_1, C_2 = DP_1P_2, ..., C_N = DP_1P_2...P_N \qquad (3)$$

Then, the no-separation approach is described by a system of ordinary differential equations:

$$\begin{aligned} d_tC_N &= -k_{off,N}C_N, \\ d_tC_{N-1} &= k_{off,N}C_N - k_{off,N-1}C_{N-1} \\ &... \\ d_tC_1 &= k_{off,2}C_2 - k_{off,1}C_1 \\ d_tC_0 &= k_{off,1}C_1 \end{aligned} \qquad (4)$$

The separation-based approach is described by a system of partial differential equations:

$$\begin{aligned}
&\partial_t C_N + \nu_N \partial_x C_N = -k_{\mathrm{off},N} C_N \\
&\partial_t C_{N-1} + \nu_{N-1} \partial_x C_{N-1} = k_{\mathrm{off},N} C_N - k_{\mathrm{off},N-1} C_{N-1} \\
&\ldots \\
&\partial_t C_1 + \nu_1 \partial_x C_1 = k_{\mathrm{off},2} C_2 - k_{\mathrm{off},1} C_1 \\
&\partial_t C_0 + \nu_0 \partial_x C_0 = k_{\mathrm{off},1} C_1
\end{aligned} \qquad (5)$$

In eqs 4 and 5, C_n and ν_n are a concentration and a migration velocity of a complex with n proteins ($0 \le n \le N$), respectively; d_t is ordinary derivation by time t; ∂_t and ∂_x are partial derivations by time t and spatial coordinate x, respectively. It is easy to see that eq 4 is a degenerate case of eq 5 when all velocities are equal to zero.

N equations in eqs4 and 5 describe N independent pathways of sequential multistage dissociations of N complexes. For example, if DNA can bind at most three proteins, there are three complexes (plus unbound DNA) in the initial mixture produced by reaction 1. Hence, the three equations in either system of differential equations describe three independent pathways starting at $N = 3$, 2, and 1, respectively.

The solution for eq 5 can be found by an approach described in detail in the Supporting Information. This solution depends on the migration velocities of complexes, ν_n, that are present in eq 5. By assuming that all ν_n are equal to zero, we obtain a solution for eq 4 that describes the no-separation approach. Such a solution can be also found directly by solving ordinary differential equations in eq 4.

By using the found solutions of eqs 4 and 5, we simulated a sum of kinetic traces (also called simulated experimental traces) for the no-separation and separation approaches for two hypothetical mechanisms of disassembly. We also added noise to the simulated experimental traces to mimic real experimental data. We then fit simulated experimental traces with model traces by varying rate constants and amounts of initial complexes using the method of least-squares that minimizes the sum of squared differences between the model and simulated experimental traces.[15] The values of velocities in the model traces were determined independently (not from the fitting procedure) making sure the peaks in the simulated experimental traces and model traces match up. Nonlinear regression was utilized to obtain the best fit. The rate constants and amounts of initial complexes that led to the best fit were considered the sought ones. The comparison of fundamental properties of the no-separation and separation approaches requires that only fundamental processes (dissociation and migration) are considered. Hence, no other processes (e.g., diffusion, convection, interaction with reactor walls, etc.) were taken into account when building simulated experimental traces for either approach. Accordingly, no additional processes were taken into account when building the model traces used to fit the simulated experimental traces.

Comparison of No-Separation and Separation Approaches. Figure 2 shows one example of a comparison between the no-separation and separation approaches for a reaction mechanism in which one DNA molecule sequentially binds up to three molecules of the same protein:

$$\begin{aligned}
&\mathrm{D} \overset{\mathrm{P},K_{\mathrm{d},1}}{\rightleftharpoons} \mathrm{DP} \overset{\mathrm{P},K_{\mathrm{d},2}}{\rightleftharpoons} \mathrm{DPP} \overset{\mathrm{P},K_{\mathrm{d},3}}{\rightleftharpoons} \mathrm{DPPP} \\
&\mathrm{DPPP} \xrightarrow{k_{\mathrm{off},3}} \mathrm{DPP} \xrightarrow{k_{\mathrm{off},2}} \mathrm{DP} \xrightarrow{k_{\mathrm{off},1}} \mathrm{D}
\end{aligned} \qquad \text{(Mechanism 1)}$$

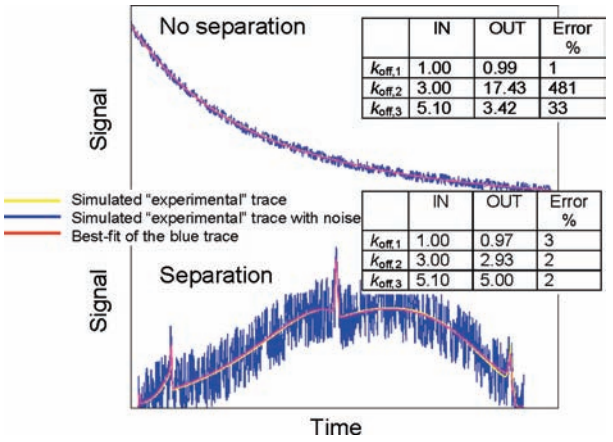


Figure 2. Numerical illustration of rate constant determination for simultaneous dissociation of three complexes of DNA with the same protein P by no-separation and separation-based approaches. DNA is assumed to be detected. Accordingly, in the no-separation approach, the signal is the cumulative concentration of intact DNA−protein complexes. Such a signal would be generated if the dissociation kinetics of DNA−multiple protein complexes was studied by sensor-based techniques, such as surface plasmon resonance (SPR), with DNA immobilized on the sensor. In the separation-based approach, the signal is the concentration of DNA free or within the complexes. The separation trace has a noise level of 30%. The no-separation trace has a noise level of 5%. The tables in the panels show the actual (IN) and determined (OUT) rate constants as well as the deviations of the determined from the actual ones.

Figure 2 shows simulated experimental and model traces for this mechanism of dissociation. Good-quality fitting could be obtained for both no-separation and separation approaches. However, the no-separation approach resulted in good fitting when the errors in rate constant determination were very large (upper table in Figure 2). Indeed, only the rate constant of the slowest process was found accurately. The other two rate constants were determined with poor accuracy, which is a result of the above-mentioned instability of the ill-posed inverse problem. Even in the presence of a relatively small noise, this instability leads to large errors in the determination of rate constants if they are close to each other. The no-separation approach also led to large errors in initial concentrations of the complexes (see Table S1 in the Supporting Information). In contrast, the separation approach determined all three rate constants with high accuracy (lower table in Figure 2) even though the level of computer-generated random noise was setup at a much higher level. The separation approach also accurately determined the initial concentrations of the complexes (see Table S1 in the Supporting Information). It is instructive to compare the imaginary separation-based trace in Figure 1 with the simulated separation-based trace in Figure 2. Due to the interplay between the three processes in dissociation Mechanism 1, the simulated trace in Figure 2 does not show explicit exponential regions that one could intuitively expect. The three kinetic processes in Mechanism 1 sum up into a trace of more complex shape. Numerical approaches are required in general for the extraction of kinetic information from such complex kinetic shapes.

To confirm that the above example is not unique, we studied a simpler mechanism with two complexes and two rate constants (see Mechanism 2 in the Supporting Information). Even for this

simple mechanism, the no-separation approach led to large error in one rate constant and two initial concentrations despite good quality of fitting. The separation approach correctly determined all unknown parameters.

To further test the separation approach, we challenged it with a very complex branched mechanism with 16 unknown parameters and 30% noise (see Mechanism 3 in the Supporting Information). The separation approach allowed us to accurately determine eight rate constants out of nine and five initial concentrations out of seven. As any practical approach, the separation-based one has its limitations. Accurate determinations of all parameters may be impossible with a single experiment; a number of experiments with different concentrations of proteins may be required for complex mechanisms. Consistent results from the multiple experiments would also ensure that the hypothetical mechanism is correct.

The determination of rate constants by any method is very difficult even for known mechanisms. It would be very useful if the same method could also test a hypothetical mechanism for correctness. Therefore, as our final simulation experiment, we compared the no-separation and separation approaches in their ability to distinguish between correct and incorrect mechanisms. For this purpose, simulated experimental traces obtained with Mechanism 1 were used. As an incorrect mechanism, we chose the worst-case-scenario mechanism with the same number of complexes and the same migration velocities as in Mechanism 1. In this case, the incorrect mechanism cannot be distinguished from the correct one based on a pattern of peaks in the model trace. The incorrect mechanism was the following:

$$\mathrm{D} \overset{\mathrm{P},K_{\mathrm{d},1}}{\rightleftharpoons} \mathrm{DP}, \quad DP^{*}\mathrm{P}^{*} \overset{\mathrm{P},K_{\mathrm{d},3}}{\rightleftharpoons} DP^{*}\mathrm{P}^{*}\mathrm{P} \qquad \text{(Mechanism 1}')$$
$$DP^{*}\mathrm{P}^{*}\mathrm{P} \xrightarrow{k_{\mathrm{off},3}} DP^{*}\mathrm{P}^{*}, \quad \mathrm{DP} \xrightarrow{k_{\mathrm{off},1}} \mathrm{D}$$

Here, protein molecules P* are assumed to form a very stable complex ($K_d \rightarrow 0$) with DNA; as a result, the P* molecules cannot dissociate from DNA. Even though an incorrect model does not have to be biologically relevant for our test, Mechanism 1′ may describe a real mechanism in cells when two protein molecules become cross-linked due to a free radical reaction.

The result of using the incorrect Mechanism 1′ for fitting the simulated experimental trace constructed with Mechanism 1 is shown in Figure 3. Though the peaks in the model trace perfectly match those in the simulated experimental one, the best fit is strikingly poor and results in a 10-fold increase in the sum-of-least-squares (compare with fitting the same trace with the correct model in Figure 2). We also applied the same incorrect mechanism to simulated experimental traces built with Mechanism 2 in the Supporting Information. The quality of fitting was also evidently unacceptable revealing that the tested mechanism was not feasible. The results allow us to conclude that, in general, the separation approach allows both testing hypothetical mechanisms of dissociation for their feasibility and finding rate constants of dissociation. Thus, the separation approach can be used to create a practical method for studying dissociation kinetics of DNA−multiple protein complexes. In addition, our numerical experiments also confirmed that the mathematical approach used for rate constant determination was also correct and could be reliably applied to rate constant determination from the experimental data even if such data have high level of noise.

Test of Separation-Based Method in Experiment. To create a practical separation-based method for studying dissociation kinetics of DNA−multiple protein complexes, three technical issues should be addressed. First, different DNA−protein complexes should be separated. Second, free proteins should be continuously removed from the vicinity of the complexes. Third, the concentration of the complexes and free DNA should be measured at a detection point distant from the point of initiation of dissociation. The three conditions can be satisfied by capillary electrophoresis (CE).[16] Due to the high negative charge of DNA, its velocity in gel-free electrophoresis changes significantly upon binding to a protein.[17] Binding to additional proteins should introduce additional changes in velocity. When the equilibrium mixture of DNA and proteins is subjected to an electric field, the DNA−protein complexes are separated from each other and from free proteins. Finally, if DNA is labeled fluorescently, the concentrations of all complexes and free DNA can be measured. We used a commercial CE instrument with fluorescence detection.

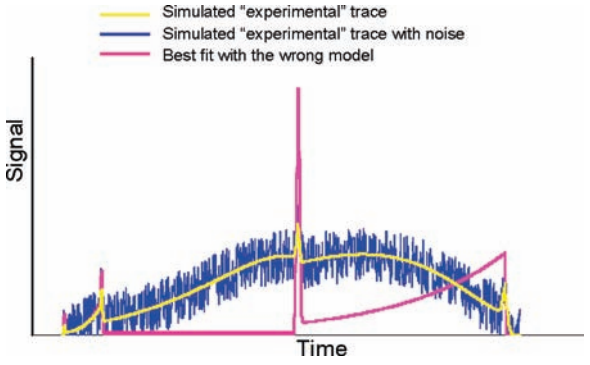


Figure 3. Numerical illustration of the best fit with an incorrect model. Simulated signal corresponds to Mechanism 1 with a noise level of 10% (it is similar to the signal at the bottom of Figure 2) whereas the model corresponds to Mechanism 1′.

As a test experimental model, we used the interaction between an 80 nucleotide long single-stranded DNA and a single-strand DNA binding (SSB) protein from *Escherichia coli*. SSB proteins bind to single-stranded DNA with high affinity and are important in DNA replication, recombination, and repair.[18−20] SSB from *E. coli* is composed of four identical subunits and is implicated in DNA metabolism, and it was shown to stimulate DNA polymerase activity by interacting with it.[18,21] To facilitate the equilibration step, DNA was mixed with the protein and incubated for 10 min (longer incubation did not influence the results). A short plug of the equilibrated mixture was then injected into the capillary by pressure. To facilitate the dissociation step, a high voltage was then applied to continuously remove free protein from the vicinity of the DNA−protein complexes and to move the complexes with different velocities. DNA was labeled with fluorescein at the 5′ end to allow for detection. A single-point detector was used to record an electropherogram with the sum of dissociation kinetics of all complexes. Sometimes the quantum yield of fluorescently labeled DNA changes upon binding with proteins. This fact can be taken into account by multiplying a modeled concentration of each complex by its corresponding quantum yield. Then, such modeled signals from all complexes are combined to produce a modeled electropherogram that is used in fitting the experimental electropherogram. In the present study, we calculated the quantum yields for DPP and DP based on the change in total fluorescence, and they were found to be 0.3 and 0.6, respectively. Nonlinear regression was used to find the best fit of the experimental electropherogram

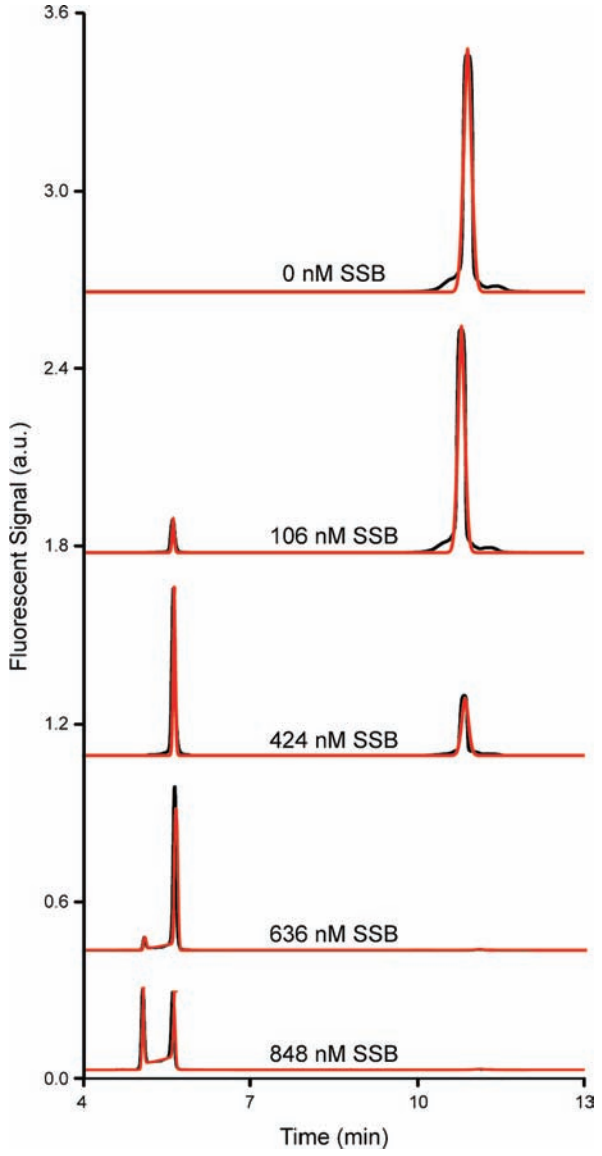


Figure 4. Experimental (black lines) and best-fit model (red lines) electropherograms. Experimental electropherograms were obtained for the interaction of fluorescently labeled dsDNA (200 nM) with SSB at varying concentrations. Experimental values of the total protein concentration (including bound and free protein) are shown. The calculated values of the free protein are 0, 79, 289, 449, and 455 nM, respectively.

with a modeled electropherogram (similar to how it was done in the numerical experiments described above). As the protein concentration mixed with DNA was known, the calculations revealed both k_{off} and K_d values.

Figure 4 shows representative experimental electropherograms (black traces) for varying concentrations of the protein. The number and combination of peaks in the electropherograms change with changing protein concentration. The right-most peak corresponds to unbound DNA that migrates slower than the complexes. The two peaks to the left of it correspond to different DNA−protein complexes. The peak with the shortest migration time is that of the complex with the greatest number of proteins molecules per DNA, which is confirmed by both faster migration time and increase in peak prominence with increasing concentration of the protein. There is pronounced exponential tailing between the left and middle peaks, which suggests that there is significant dissociation of the complex corresponding to left peak. The exponential tailing in Figure 4 is directed from the smaller complex to the larger complex it is formed from. This can be briefly explained in the following way. The highest rate of complex dissociation is in the very beginning, when the concentration of the complex is the highest. A smaller complex (or free DNA), which is generated as a result of the dissociation of a larger complex, will be produced in the largest amount in the very beginning and will migrate close to the peak of the smaller complex (or free DNA). The amount of the smaller complex produced decreases with time while these smaller "portions" will migrate closer to the peak of larger complex. This behavior defines the direction of the tail from the smaller complex to the larger one. There is no detectable tailing between the middle and right peaks even when the dissociation time is increased 10 times by decreasing the electric field strength. This fact indicates that the complex corresponding to middle peak does not dissociate significantly in the time scale of our experiment.

SSB is known to exist in solution predominantly as a tetramer.[18] Depending on conditions (buffer composition, pH, temperature), SSB can bind from 35 to 60 nucleotides per tetramer.[12,22,23] The DNA used in the experiments was 80 nucleotides long, so at least two SSB tetramers should theoretically be able to bind to it. Based on this information and on the observation of two peaks corresponding to DNA−protein complexes in Figure 4, we consider the following mechanism for the determination of rate constants:

$$\mathrm{D} \underset{}{\overset{\mathrm{P},K_{d,1}}{\rightleftharpoons}} \mathrm{DP} \overset{\mathrm{P},K_{d,2}}{\rightleftharpoons} \mathrm{DPP}$$
$$\mathrm{DPP} \xrightarrow{k_{off,2}} \mathrm{DP} \xrightarrow{k_{off,1}} \mathrm{D} \qquad \text{(Mechanism 2)}$$

where P represents a tetrameric protein.

Mechanism 2 does not include peak broadening, as in our particular case, the effect of various mechanisms of peak broadening was negligible (see the Materials and Methods Section), and therefore, the model with no peak broadening was used in fitting.

The procedure of building simulated traces was similar to that described above for Mechanism 1. The best fits of the experimental data by this model are shown in Figure 4 by red traces. The quality of the fits is high, which suggests that Mechanism 2 satisfactorily describes the experiment. Our calculations returned k_{off} and K_d values shown for convenience in the reaction mechanisms below:

$$\mathrm{D} \underset{}{\overset{K_d = 102 \pm 28\ \mathrm{nM}}{\rightleftharpoons}} \mathrm{DP} \overset{K_d = 90 \pm 57\ \mathrm{nM}}{\rightleftharpoons} \mathrm{DPP}$$
$$\mathrm{DPP} \xrightarrow{k_{off,2} = (1.45 \pm 0.13)\times 10^{-3}\ \mathrm{s}^{-1}} \mathrm{DP} \xrightarrow{k_{off,1} < 10^{-4}\ \mathrm{s}^{-1}} \mathrm{D} \qquad \text{(Mechanism 2a)}$$

The determined K_d values are in the range of K_d values typically observed for this protein.[24] The detection limit of the instrument and limited time of the experiment did not allow us to accurately measure the k_{off} value for the dissociation of DP; we could only estimate its upper limit. The entire experiment was repeated, and the new calculations returned identical values of the constants.

For these experiments to be successful, the effect of the reverse possesses of DNA−protein association on concentrations of complexes should be small. This effect will be small if each

complex is separated from the protein with a relative velocity $\Delta\nu = |\nu_c - \nu_p|$ that satisfies the $t_{sep} \ll t_{ass}$ condition. Here, ν_c and ν_p are velocities of the complex and protein, $t_{sep} = W/\Delta\nu$ and $t_{ass} = K_d/(k_{off}P_{free})$ are the characteristic times of separation and association; W is the width of the initial plug of complexes; P_{free} is the concentration of free protein. One can readily verify that the condition $t_{sep} \ll t_{ass}$ is valid for all found values of K_d and k_{off} (shown in Mechanism 2) and for corresponding concentrations P_{free} (mentioned in Figure 4) if $\Delta\nu \sim 0.1$ cm/s and $W < 0.5$ cm. Such values of $\Delta\nu$ and W are typical for CE experiments and lead to $t_{sep} < 5$ s. For example, we have $t_{ass} \sim 200$ s $\gg 5$ s in the case of dissociation of the DPP complex ($K_d \approx 100$ nM, $k_{off} \approx 1.5 \times 10^{-3}$ 1/s, $P_{free} \approx 450$ nM).

It should be emphasized that we do not claim that mechanism (eq 5) of the experimental example is adequately detailed. For instance, DNA can bind a single protein octamer instead of two protein tetramers.[12] Such DNA−protein complexes are identical in their size and charge and cannot be easily distinguished with electrophoresis-based experiments. Thus, the detailed study of complex mechanisms may require a combination of the method suggested here with other techniques.

To conclude, in this work, we proved in principle that separation-based kinetic methods can facilitate the study of dissociation kinetics of DNA−multiple protein complexes. Separation solves the virtually impossible problem of "extracting" individual single-exponential curves from their sum. Separation can potentially facilitate studying the assembly kinetics of DNA−multiple protein complexes. Further, our approach can be extended to studies of protein−protein interactions if a generic means of separation of multiple protein complexes is found.

MATERIALS AND METHODS

Chemicals, Solutions, and Materials. All buffer components were from Sigma-Aldrich (Oakville, ON). All aqueous solutions were made with deionized water and filtered through a 0.22 μm filter (Millipore, Nepean, ON). SSB protein was from Interscience (Markham, ON, Canada). A fluorescently labeled 80-mer oligonucleotide (5-FAM-CTCCTCTGACTGTAACCACGTGCC TAGCGTTTCATTGTCC-CTTCTTATTAGGTGATAATAGCATAGGTAGTCCAGAAGCC-3) was custom synthesized by IDT Technologies Inc. (Coralwille, IA, USA). The protein and DNA stock solutions as well as equilibrium mixtures were prepared in the incubation buffer (25 mM Borax, pH 10). Fused silica capillaries were purchased from Polymicro (Phoenix, AZ).

Capillary Electrophoresis. CE experiments were performed with a CE instrument (P/ACE MDQ, Beckman-Coulter, USA) with thermostabilization of the capillary (the outer walls of the capillary were washed with a liquid heat exchanger maintained at 20 °C) and sample vials. The CE method used was Non-Equilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM).[13] The instrument employed laser-induced fluorescence detection with a 488 nm line of an argon-ion laser for fluorescence excitation. An uncoated fused silica capillary was used with the following dimensions: 50 cm total length/20 μm inner diameter/350 μm outer diameter. The length L from the injection end to the detection window was 40 cm. Electrophoresis was run with a positive electrode at the injection end and an electric field E of 600 V/cm. The run buffer for all NECEEM experiments was the same as the incubation buffer: 25 mM Borax, pH 10. The samples were injected into the capillary by a pressure pulse of 15 s × 2 psi; the length W of corresponding sample plug was ~0.6 cm. Prior to each run, the capillary was rinsed with deionized water for 2 min, 100 mM HCl for 2 min, 100 mM NaOH, and then, the run buffer solution for 2 min. All NECEEM experiments were performed in two repeats.

Equilibrium Mixtures. Equilibrium mixtures of the samples were prepared by mixing the protein and DNA in the incubation buffer and incubating at room temperature for 10 min. The fact that the equilibrium was reached was confirmed by not seeing changes in experimental results for incubation times longer that 10 min. The CE analysis was started immediately after that.

Peak Broadening Effects. Peak broadening δW caused by molecular diffusion can be estimated as $\delta W \sim \sqrt{(DL/\nu)}$, where $\nu \sim 0.1$ cm/s is the migration velocity and D is the molecular diffusion coefficient. The latter is of the order of $(2-10) \times 10^{-6}$ cm^2/s for DNA and $(3-10) \times 10^{-6}$ cm^2/s for proteins.[25,26] Hence, $\delta W/W < 10^{-1}$ (at $W \sim 0.6$ cm), and peak broadening due to pure molecular diffusion is negligible. Another source of a peak broadening could be Joule heating.[27] The amount of heat Q can be estimated as $Q = \sigma E^2 \sim 10^2$ J/(cm^3 s) where $\sigma \sim 10^{-3}$ A/(V cm) stands for buffer conductivity. Capillary cooling with a sufficient rate of heat transfer was provided to maintain the temperature at 15 ± 2 °C and to exclude potential peak broadening due to heating.[28,29] Peak broadening due to interactions of charged DNA−protein complexes with the wall was also negligible. Indeed, such interactions (adhesion to the walls) would lead to the peak tailing.[30−32] However, Figure 4 clearly shows that the peak tailing is much smaller than the peak fronting in experimental electropherograms. The pattern of observed peak fronting is consistent with broadening due to dissociation of DNA−protein complexes. Finally, pure DNA was ran as a control (Figure 4) to evaluate peak broadening that could be caused by stacking and antistacking. No such a phenomenon was observed (the peak of DNA was perfectly symmetrical), and thus, stacking and antistacking were not included in the model. The absence of stacking/antistaking is consistent with the fact that the sample buffer was not different from the run buffer.

Determination of Experimental Rate and Equilibrium Constants. To determine the rate and equilibrium constants of complex dissociation, experimental electropherograms were fitted with simulated electropherograms obtained using the mathematical model. The model took into account chemical equilibrium at the pre-electrophoresis stage and complex dissociation at the stage of electrophoresis. Minimum mean-square deviation between the experimental and simulated electropherograms was used as a criterion of acceptance of rate constants and equilibrium constant. A computer program, which built simulated electropherograms for varying rate constants and calculated mean-square deviation of the simulated and experimental electropherograms, was written in Excel using Visual Basic and Excel Solver. This program was not optimized for high-speed routine use but was sufficiently productive for the proof-of-principle work.

Assessment of Errors and Reproducibility for the Separation-Based Method. It is worth noting that systematic errors in determination of rate constants using the separation-based approach are relatively small in simulated experiments. In the presence of 30% noise, these errors are ~3% for Mechanism 1 as shown in the insert in Figure 2. They are ~10% for Mechanism 2 with two complexes and two rate constants (see Table S2 in the Supporting Information). Errors become smaller (2% for 10% noise and 1% for 5% noise) if the noise level decreases. Even in the case of very complicated (branched) Mechanism 3 with nine stages and a noise level of 30%, systematic errors remained less than 40% for seven rate constants (see Table S3 in the Supporting Information). These results were produced in single simulations where the starting values of the rate constants were determined by graphical guess-and-check. The guess-and-check procedure was performed by varying the rate constant and concentration parameters of the simulated electropherogram by orders of magnitude until the characteristic peaks and decays of the experimental data were reproduced as closely as possible. The starting values of the parameters need only be "good guesses", so the relative heights of the peaks and decays did not have to match the data exactly. To ensure that our results were reproducible and

did not depend on the choice of starting values, two "worst guess" simulations were run to fit Mechanism 1 to the simulated data from the separation-based trace in Figure 2. The first simulation underestimated all rate constants to be zero and the second overestimated them to be $>1 \times 10^2\ s^{-1}$. Convergence to the correct values was always observed, and it was found that underestimating the rate constants gave final errors <10% while overestimating them gave errors <7%. All simulated tests of the separation-based approach suggest that, in real experiments, errors in determination of rate constants should mainly depend on errors in experimental data used in the fitting procedure. To find such errors for electropherograms shown in Figure 4, these experiments were repeated 4 times with errors of ~27%, 63%, and 9% for $K_{d,1}$, $K_{d,2}$, and $k_{off,2}$, respectively.

■ ASSOCIATED CONTENT

Supporting Information. Solution of eq 5 and comparison of k_{off} determination by no-separation and separation-based approaches. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
skrylov@yorku.ca

■ ACKNOWLEDGMENT

The work was funded by the Natural Sciences and Engineering Research Council of Canada.

■ REFERENCES

(1) Schrofelbauer, B.; Hakata, Y.; Landau, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4130.
(2) Saiz, L.; Vilar, J. M. G. *IET Syst. Biol.* **2008**, *2*, 247.
(3) Verma, M.; Rawool, S.; Bhat, P. J.; Venkatesh, K. V. *Biosystems* **2006**, *84*, 39.
(4) Margeat, E.; Bourdoncle, A.; Margueron, R.; Poujol, N.; Cavailles, V.; Royer, C. *J. Mol. Biol.* **2003**, *326*, 77.
(5) Pavski, V.; Le, X. C. *Curr. Opin. Biotechnol.* **2003**, *14*, 65.
(6) Lin, J. J.-C.; Grosskurth, S. E.; Harlan, S. M.; Gustafson-Wagner, E. A.; Wang, Q. *Methods Mol. Biol.* **2007**, *366*, 183.
(7) Pollet, J.; Delport, F.; Janssen, K. P. F.; Jans, K.; Maes, G.; Pfeiffer, H.; Wevers, M.; Lammertyn, J. *Biosens. Bioelectron.* **2009**, *25*, 864.
(8) Wiscombe, W. J.; Evans, J. W. *J. Comput. Phys.* **1977**, *24*, 416.
(9) De Groen, P.; De Moor, B. *J. Comput. Appl. Math.* **1987**, *20*, 175.
(10) Zaikin, P. N.; Khoroshilova, E. V. *Comput. Math. Model* **1995**, *6*, 96.
(11) Lupu, M.; Todor, D. *Chemom. Intell. Lab. Syst.* **1995**, *29*, 11.
(12) Lohman, T. M.; Ferrari, M. E. *Annu. Rev. Biochem.* **1994**, *63*, 527.
(13) Berezovski, M.; Krylov, S. N. *J. Am. Chem. Soc.* **2002**, *124*, 13674.
(14) Schasfoort, R. B. M.; Tudos, A. J. *Handbook of Surface Plasmon Resonance*; RSC Publishing: Cambridge, U.K., 2008.
(15) Wolberg, J. R. *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*; Springer: Berlin, Germany, 2006.
(16) Berezovski, M.; Krylov, S. N. *Anal. Chem.* **2005**, *77*, 1526.
(17) Berezovski, M. V.; Musheev, M. U.; Drabovich, A. P.; Jitkova, J. V.; Krylov, S. N. *Nat. Protocols* **2006**, *1*, 1359.
(18) Molineux, I. J.; Friedman, S.; Gefter, M. L. *J. Biol. Chem.* **1974**, *249*, 6090.
(19) Cox, M. M.; Lehman, I. R. *J. Biol. Chem.* **1982**, *257*, 8523.
(20) Ryzhikov, M.; Koroleva, O.; Postnov, D.; Tran, A.; Korolev, S. *Nucleic Acids Res.* **2011**in press.
(21) Molineux, I. J.; Gefter, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3858.
(22) Bujalowski, W.; Lohman, T. M. *Biochemistry* **1986**, *25*, 7799.
(23) Bujalowski, W.; Overman, L. B.; Lohman, T. M. *J. Biol. Chem.* **1988**, *263*, 4629.
(24) Ferrari, M. E.; Lohman, T. M. *Biochemistry* **1994**, *33*, 12896.
(25) Nkodo, A. E.; Garnier, J. M.; Tinland, B.; Ren, H.; Desruisseaux, C.; McCormick, L. C.; Drouin, G.; Slater, G. W. *Electrophoresis* **2001**, *22*, 2424.
(26) Young, M. E.; Carroad, P. A.; Bell, R. L. *Biotechnol. Bioeng.* **1980**, *22*, 947.
(27) Xuan, X. C.; Li, D. Q. *J. Micromech. Microeng.* **2004**, *14*, 1171.
(28) Evenhuis, C. J.; Musheev, M. U.; Krylov, S. N. *Anal. Chem.* **2010**, *82*, 8398.
(29) Musheev, M. U.; Filiptsev, Y.; Krylov, S. N. *Anal. Chem.* **2010**, *82*, 8692.
(30) Ghosal, S. *J. Fluid Mech.* **2003**, *491*, 285.
(31) Shariff, K.; Ghosal, S. *Anal. Chim. Acta* **2004**, *507*, 87.
(32) Sun, Y.; Kwok, Y.; Nguyen, N.-T. *Microfluid. Nanofluid.* **2007**, *3*, 323.