Protein Labeling Enhances Aptamer Selection by Methods of Kinetic Capillary Electrophoresis

Stephanie de Jong and Sergey N. Krylov*

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

Supporting Information

ABSTRACT: Methods of kinetic capillary electrophoresis (KCE) facilitate highly efficient selection of DNA aptamers for protein targets. The inability to detect native proteins at low concentrations in capillary electrophoresis creates, however, a significant obstacle for many important protein targets. Here we suggest that protein labeling with new Chromeo dyes can help to overcome this obstacle. By labeling a number of proteins with Chromeo P503, we show that the labeling procedure enables accurate detection of proteins in CE without significantly affecting their electrophoretic mobility or their ability to bind DNA. Moreover, Chromeo P503 does not appear to label the amino-groups of buffer components to a significant extent, making the labeling procedure compatible with a large number of selection and run buffers. Fluorescent labeling of protein targets with Chromeo dyes empowers selection of aptamers by KCE methods and promises to increase the rate at which aptamers for new targets are being developed and introduced in various applications.

Aptamers are small ssDNA (RNA) oligonucleotides which can bind their targets with high selectivity and affinity.1,2 Their molecular recognition capacity has been shown to parallel that of antibodies,3 enabling the use of aptamers in several important applications as affinity probes, drug candidates, and drug delivery vehicles.4–6 Aptamers are selected from large libraries of random DNA sequences using the following iterative process. First, the library is combined with the target and left to incubate, allowing for suitable DNA sequences to bind. Then the target-bound DNA is isolated from the unbound DNA in a partitioning step. Because of limited separation efficiency, the partitioning step typically needs to be repeated multiple times to obtain aptamers with the desired binding parameters. Moreover, partitioning leads to unavoidable aptamer-loss due to collection inefficiencies associated with the separation technique. As a result, PCR amplification is typically needed after each partitioning step in order to compensate for the DNA loss. This general procedure which involves multiple rounds of alternating aptamer partitioning and PCR amplification is referred to as systematic evolutionary enrichment of ligands by exponential enrichment (SELEX).7 The development of highly efficient partitioning methods is critical, as this will increase the rate at which aptamers are selected for new targets.

Kinetic capillary electrophoresis (KCE) is a conceptual platform for homogeneous kinetic affinity methods which can, in particular, facilitate highly efficient partitioning of aptamers.8 One to four rounds of KCE-based partitioning is generally sufficient to complete the selection process,9 while other partitioning methods typically require more than 10 rounds. The high efficiency of KCE methods allows the intermediary PCR steps to be excluded during aptamer selection, by an approach entitled Non-SELEX.10 In addition, KCE methods uniquely facilitate selection of smart aptamers, aptamers with predefined binding parameters, which include the equilibrium and rate constants of interaction between the aptamer and target.11 KCE-based aptamer selection can be optimized in three distinct steps: (i) finding conditions under which the protein does not interact with the capillary inner wall, (ii) maximizing the separation between the target protein and the DNA library, and (iii) determining the aptamer-collection window (Figure 1).

To perform these three steps of optimization, both the DNA library and protein target need to be detected. DNA can be easily tagged with a fluorescent label while maintaining its native electrophoretic properties; therefore, both UV absorption and laser induced fluorescence can be used for library detection. As for the protein, early works on fluorescent labeling demonstrated peak broadening and a significant shift in CE migration times, making the protein labeling procedure unsuitable for optimizing KCE-based aptamer selection.12 Therefore, only a single means of protein detection, light absorption in UV, has been used to facilitate the selection process. This creates problems in all three steps of optimization. Light absorption has a poor detection limit in CE due to the short optical path-lengths of approximately 0.1 mm. Because of the multistep nature of protein purification, many proteins are unavailable at the detectable concentration of \(10^{-3}\) M or higher. In these cases, optimization steps 1 and 2 cannot be carried out using detection by UV absorbance.
The inability to detect the protein can also present a challenge in step 3, which typically relies solely on the fluorescently labeled DNA library. Under ideal circumstances, the labeled DNA is sufficient to detect both the complex and DNA, which will accordingly establish the aptamer collection window (see Figure 1C). However, if the protein has nuclease activity, the library may become partially degraded during the incubation period and these products of degradation can be confused with the complex (see the Supporting Information, Figures S1 and S2). It is also possible that the complex peak remains undetected during the bulk affinity assay even at the highest available target concentrations. In these situations, the mobility of the pure protein is used as a reference for boundary determination, which again relies on UV absorbance.

If not optimized, KCE-based aptamer selection has a much lower chance of success. The goal of this work was to find a means of fluorescently labeling the protein target so that it is compatible with KCE-based aptamer selection. Labeling reagents typically react with the positively charged ε-amine groups on the exposed lysine residues and alter the protein’s charge upon conjugation. If the proteins charge-to-size ratio is affected by the labeling procedure, its corresponding electrophoretic velocity will be shifted making the migration time unreliable. A new series of fluorescent dyes, termed Chromeo, have recently become commercially available. These dyes can facilitate fluorogenic labeling without changing the protein’s charge during the process. Dovichi and coauthors demonstrated that Chromeo dyes can be successfully used in capillary isoelectric focusing, as the protein’s pI is maintained.13,14 While not suggesting that protein mobility in CE would remain unaffected, this work had motivated us to explore this opportunity. Our results proved that Chromeo labeling of protein can facilitate the optimization of KCE-based aptamer selection for proteins.

### RESULTS AND DISCUSSION

Ideally, the labeling would have three properties. First, it cannot significantly shift the protein’s migration with respect to the time-window chosen for aptamer collection. Second, labeling should not prevent protein–DNA binding. Third, the labeling reaction should be compatible with commonly used amino-containing buffers such as Tris, glycine, etc. The goal of our work was to test these three properties of protein labeling using a dye from the Chromeo family. We chose Chromeo P503 to test this hypothesis given that its quantum yield increases by a factor of 50 when bound to the protein. This property effectively eliminates any background signal, allowing for sensitive detection. Chromeo P503 offers additional advantages over other Chromeo dyes, attributed to its unique spectral properties. The excitation wavelength at 503 nm is similar to both the FAM and Alexa488-labeled DNA libraries, while its emission is shifted to a...
much longer wavelength. This would allow both the protein and DNA to be excited simultaneously by the same laser source yet detected through different channels by using the appropriate filters.

We first confirmed that by fluorescently labeling proteins with Chromeo P503, detection in CE is significantly improved compared to light-absorbance of the native protein. Because of the significant variations that exist in the physical—chemical properties of the proteins, we decided to analyze a series of proteins as models; bovine serum albumin (BSA), human serum albumin (HSA), α-acid glycoprotein (AGP), myoglobin, and α-lactalbumin, which would allow us to make reliable conclusions with a reasonable power of prediction. Our results shown in Table 1 suggest that the lower limit of detection improves (for the five proteins studied) from an average value of 1620 to 35 nM when switching from UV light absorbance of unlabeled proteins to fluorescence detection of labeled proteins (see also the Supporting Information). This detection limit is suitable for KCE-based aptamer selection since proteins are typically available, and often used, at concentrations higher than 35 nM. Thus, for the majority of proteins, Chromeo labeling can facilitate the optimization of aptamer selection provided that the label does not influence the protein’s electrophoretic mobility to a great extent.

In order to quantify the mobility shift attributed to the Chromeo label, we measured the difference in the migration times of native and Chromeo-labeled proteins with respect to the separation window between the protein and DNA library. These experiments were performed using the five previously listed proteins, as they were available at concentrations high enough to produce a reliable signal when using UV absorbance for detection. The top left panel in Figure 2 shows the mobility of the library; the remaining panels demonstrate the CE migration of the unlabeled and Chromeo-labeled proteins, detected using UV absorbance and laser induced fluorescence (LIF), respectively. For each protein tested, only a marginal shift in electrophoretic mobility was observed following the labeling reaction. α-Lactalbumin expressed the greatest shift in mobility upon labeling, which was found to be approximately 10% of the protein-library separation window. Moreover, Chromeo P503 does not lead to labeling products with heterogeneous mobility since all labeled proteins migrate as individual peaks with widths similar to those of the native proteins, a result which could not be obtained by FQ-labeling. From these results, we conclude that Chromeo-labeled proteins can be used in the first two steps of optimization for KCE-based aptamer selection.

For step 3 of the optimization process (determining the aptamer selection window), the labeling process should not significantly perturb protein—DNA binding. Figure 3 illustrates how protein labeling influences aptamer binding by using two proteins for which aptamers were readily available: thrombin15 and AlkB.16 These binding studies were performed by non-equilibrium capillary electrophoresis of equilibrium mixtures (NECEEM) which can be used to determine equilibrium ($K_d$) and rate ($k_{off}$) constants of the protein—aptamer complex dissociation. With the use of a 520 nm filter, only the fluorescence emitted by the DNA library was detected and shown in each trace. We can easily see that the NECEEM electropherograms are similar for both the Chromeo-labeled and unlabeled proteins in both aptamer/protein models. This result was unexpected but is not counterintuitive. Indeed, while labeling can affect biomolecular interactions, the wide use of fluorescently labeled antibodies suggests that when it comes to binding other molecules, proteins have a relatively high tolerance toward the labeling of a limited number of residues. Our observation certainly does not prove that Chromeo labeling will be this inert to all protein—aptamer pairs. However, during step
The bottom trace is a second control which illustrates the peak obtained from a 10 mM Tris buffer. The middle trace is a control illustrating how this Tris buffer reacts with the Chromeo dye. The bottom trace is a second control which illustrates the peak obtained from the Chromeo dye alone.

Figure 5. Negligible effect of amino-containing buffer components on fluorescence signal from Chromeo-labeled Exo1 protein. The top trace shows the electropherogram obtained from Chromeo-labeled Exo1, which was stored in a 10 mM TrisHCl buffer. The middle trace is a control illustrating how this Tris buffer reacts with the Chromeo dye. The bottom trace is a second control which illustrates the peak obtained from the Chromeo dye alone.

3 of optimization, which relies on the interaction of the protein with the entire library (so-called bulk affinity assays), the influence of Chromeo labeling is expected to be negligible for the majority of proteins. The library provides an incredibly large number of DNA structures and even if labeling prevents the interaction of some structures, it may equally increase the protein’s affinity to others. Since the bulk affinity assay is done prior to the collection and amplifications steps, the labeled protein can be used initially for optimization and predictive purposes while the unlabeled protein is subsequently used for aptamer selection. This approach would enhance the selection process and ensure that the aptamers will bind tightly to the native protein. Therefore, it is possible to use Chromeo-labeled proteins in step 3 of the optimization process.

We then showed that Chromeo labeling can help to correctly identify the protein–DNA complex from other unknown peaks such as library degradation products. This distinction can be made by simultaneously detecting the protein and DNA at two different wavelengths. Both the Chromeo-labeled protein and Alexa-labeled DNA can be excited using the same 488 nm laser line; however, their emission wavelengths are unique, 610 nm for the protein and 520 nm for the DNA. Thus, with the use of a two channel detection system, the fluorescence emitted from both the DNA and protein can be separated into two distinct electropherograms. The principle of complex distinction is based on the fact that the protein–DNA complex will fluoresce at both wavelengths, whereas any peaks attributed to DNA degradation or impurities would fluoresce only at a single wavelength (unless bound to the protein). To illustrate this approach, we conducted an experiment in which we used either Chromeo-labeled or unlabeled thrombin and its Alexa-labeled aptamer, using fluorescence detection at both 520 and 610 nm. NECEEM experiments were conducted in which we expect to observe two main peaks: unbound aptamer and protein-bound aptamer.

When the unlabeled protein is used in the equilibrium mixture (upper panel in Figure 4), only DNA fluoresces and ideally we expect to see no signal in the 610 nm channel. However, the channel picks up the tail of Alexa fluorescence from the DNA and records a signal similar to that in the 520 nm channel but at a much lower intensity. The similarity can be confirmed by normalizing the signals from the two channels; the normalized signals are identical. The right-hand side peak is tentatively assigned to the protein–DNA complex but it may potentially be a product of DNA degradation that migrates differently from intact DNA.

To test whether or not the assignment is right, we use the Chromeo-labeled thrombin (lower panels in Figure 4). In this case, the signals at the two wavelengths are not similar; the right-hand side peak is higher relative to the left-hand side peak in the 610 nm channel. This becomes very obvious when the signals are normalized. Both the Chromeo label of the protein and the Alexa label of the DNA contribute to this peak while the left peak is solely the spectral tail of DNA fluorescence. This allows us to accurately conclude that the right-hand side peak is the complex and not any other product such as aptamer degradation. This simple model experiment proved that Chromeo labeling of the protein can facilitate the simple identification of the protein–aptamer complex from other possible peaks. This distinction is crucial during the determination of the aptamer-selection window, step 3 in the optimization process, and throughout the course of selection.

Aptamer selection by KCE methods is often conducted in amino-group containing buffers, such as Tris and glycine. Since Chromeo dyes react with primary and terminal amino groups present on the proteins, they can potentially react with amino groups present in buffer components. The labeled buffer can, in turn, interfere with protein detection. To test this potential problem, we performed a labeling reaction using Exo1, a protein that was stored in a Tris-containing buffer. With comparison of the electropherogram of the labeled protein to that of the labeled storage buffer containing the Tris-component, the migration of the pure protein can be inferred. It is evident from Figure 5 that the Tris buffer, which is present at a much higher concentration than the Exo1, produced a much smaller peak relative to the protein. This suggests that Chromeo-labeling is compatible with amino-containing buffers.

## CONCLUDING REMARKS

To conclude, we proved that labeling proteins with Chromeo can facilitate sensitive detection of proteins in CE and does not significantly affect protein mobility or DNA-binding properties. Along with low reactivity toward amino-groups present in buffer solutions, these findings make Chromeo labeling an important addition to the tool-box of KCE-based methods for aptamer selection. This new opportunity promises to significantly increase the rate of success in KCE-based aptamer selection and thus contribute to further development of aptamer-based diagnostic and therapeutic tools.

## MATERIALS AND METHODS

**Materials.** Uncoated fused-silica capillaries with 75 μm inner diameter (375 μm outer diameter) were purchased from Polymicro (Phoenix, AZ). Chromeo P503 pyrylium dye was purchased from Active Motif (Burlington, ON, Canada). AlkB and Exo1, both native to *E. coli* were provided by Oxford University and Oxford Nanopores, respectively. Human-α-thrombin was...
μsolutions were prepared in deionized water and filtered through a m filter (Millipore, Nepean, ON, Canada). The DNA was controlled at 15 °C for 1 min followed by a controlled decrease in temperature at a rate of 0.5 °C per second until the sample reached 25 °C. This temperature treatment, or “annealing”, is necessary as it promotes the proper folding of the ssDNA aptamers. A mixture containing 0.5 μM of labeled/unlabeled thrombin and 0.1 μM thrombin aptamer was prepared in a 50 mM Tris-acetate (pH 8.2) buffer supplemented with 5 mM KCl and 1 mM MgCl2. The mixture was left to incubate for 15 min at room temperature prior to separation in a 50 cm PVA coated capillary, which was used reduce protein adsorption to the inner capillary surface. A voltage of 10 kV was used for the first 5 min of separation in order to avoid sample overheating in the uncooled region of the capillary. The voltage was then increased to 20 kV for the remainder of the run. A mixture containing 80 nM labeled/unlabeled AlkB and 0.3 nM AlkB aptamer was prepared in a 50 mM HEPES (pH 7.5) buffer containing 50 mM NaCl. Following a 15 min incubation, the equilibrium mixture was injected into a 50 cm uncooled capillary. A voltage of 10 kV was applied for the first 5 min of the run, which was subsequently increased to 20 kV.

**ASSOCIATED CONTENT**

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

**AUTHOR INFORMATION**

Corresponding Author

E-mail: skrylov@yorku.ca.

**ACKNOWLEDGMENT**

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

**REFERENCES**


Supporting Information

Protein Labeling Enhances Aptamer Selection by Methods of Kinetic Capillary Electrophoresis

Stephanie de Jong and Sergey N. Krylov

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

Exonuclease I (Exo1) degradation of the ssDNA library during bulk affinity assay

An equilibrium mixture containing 3 μM Exo1 and 100 nM of an annealed DNA library were prepared in a 50 mM Tris-acetate (pH 8.2) incubation buffer containing 100 mM NaCl, 1 mM EDTA, and 0.5 mM MgCl2. The mixture was left at room temperature for 15 minutes to ensure the protein and DNA had reached equilibrium. 50 nL of the mixture was injected into a 50 cm uncoated capillary and separated by applying 20 kV. Figure S1 depicts multiple peaks. Although these peaks may be mistakenly identified as Exo1-DNA complex, they are products of DNA degradation.

![Figure S1](image_url)

Figure S1. Equilibrium Mixture containing 3 μM Exo1 and 100 nM DNA library in a 50 mM Tris-acetate incubation buffer containing 100 mM NaCl, 1 mM EDTA, and 0.5 mM MgCl2. The mixture was separated in a 50 mM Tris-acetate (pH 8.2) run buffer with an applied voltage of 20 kV. The DNA library is shown as the rightmost peak while the products of degradation are indicated with arrows.

The identity of these peaks was confirmed by performing capillary gel electrophoresis (CGE) using the same incubation buffer. CGE can accurately discriminate between DNA that differs in length by only one nucleotide and, therefore, confirm if any DNA degradation has occurred. A DNA ladder containing 20 nucleotides (nt), 37 nt, 56 nt and 80 nt DNA markers was used to assess the length of the DNA present in the sample. By performing a co-injection of an equilibrium mixture and the DNA ladder, degradation can be accurately assessed. Figure S2 illustrates the DNA digestion following a 15 min incubation of 100 nM Exo1 and 5μM of the DNA library that was diluted in the same 50 mM Tris-acetate (pH 8.2), 100
mM NaCl, 1 mM EDTA, 0.5 mM MgCl₂ buffer. It is clear that degradation products are present, as ssDNA containing less than 20 nt are observed in the electropherogram.

**Figure S2.** CGE results illustrating Exo1 activity in a 50 mM Tris-acetate (pH 8.2) buffer supplemented with 100 mM NaCl, 1 mM EDTA, and 0.5 mM MgCl₂. The DNA ladder and an equilibrium mixture containing 100 nM Exo1 and 5 μM of the DNA library were co-injected into the capillary. Products of DNA degradation having less than 20 nt in length are indicated with arrows.

This observation proves that DNA degradation occurs under these conditions, and the multiple peaks seen in Figure S1 are likely degraded ssDNA.

**Establishing the Limit of Detection (LOD)**

Solutions of labeled and unlabeled BSA, HSA, AGP, myoglobin and α-lactalbumin were prepared as serial dilutions using 100 mM NaHCO₃ (pH 8.3) dilution buffer. 50 nL of the sample was introduced into a 50 cm uncoated capillary (75 μm i.d.) and protein migration was induced by applying 20 kV. A PDA detector was used to detect the protein at 214 nm, while fluorescence emission at 610 nm was used to detect the labeled protein. The standard deviation of the background signal height (within the 1 to 2.5 minutes time frame) was determined for each sample. The standard deviation was then multiplied by 5 to determine the LOD in relative fluorescence units. This value was converted into units of concentration by normalizing it to the protein peak intensity at a known concentration. For each protein, detection limits obtained with UV absorbance and fluorescence were averaged and tabulated in Table 1 of the main text. LOD experiments are shown in figures S3, S4, S5, S6, and S7 with fluorescence detection shown in the top panel and UV detection shown in the lower panel.
Figure S3. Electropherograms illustrating the improvement in CE detection of BSA by labeling protein with Chromeo P503. The top panel shows the data obtained using fluorescence detection while the lower panel illustrates the data obtained using UV absorbance detection. The respective concentrations are indicated in the right corners.

Figure S4. Electropherograms illustrating the improvement in CE detection of HSA by labeling protein with Chromeo P503. The top panel shows the data obtained using fluorescence detection while the lower panel illustrates the data obtained using UV absorbance detection. The respective concentrations are indicated in the right corners.
Figure S5. Electropherograms illustrating the improvement in CE detection of AGP by labeling protein with Chromeo P503. The top panel shows the data obtained using fluorescence detection while the lower panel illustrates the data obtained using UV absorbance detection. The respective concentrations are indicated in the right corners.

Figure S6. Electropherograms illustrating the improvement in CE detection of myoglobin by labeling protein with Chromeo P503. The top panel shows the data obtained using fluorescence detection while the lower panel illustrates the data obtained using UV absorbance detection. The respective concentrations are indicated in the right corners.
Figure S7. Electropherograms illustrating the improvement in CE detection of α-lactalbumin by labeling protein with Chromeo P503. The top panel shows the data obtained using fluorescence detection while the lower panel illustrates the data obtained using UV absorbance detection. The respective concentrations are indicated in the right corners.