Jeffrey M. Savard
Shane T. Grosser
James W. Schneider

Department of Chemical Engineering,
Carnegie Mellon University,
Pittsburgh, PA, USA

Received August 1, 2007
Revised February 16, 2008
Accepted February 18, 2008

Research Article

Length-dependent DNA separations using multiple end-attached peptide nucleic acid amphiphiles in micellar electrokinetic chromatography

End-labeled free-solution electrophoresis (ELFSE) is an alternative approach to gel-based methods for size-based electrophoretic separation of DNA. In ELFSE, an electrically neutral "drag-tag" is appended to DNA to add significant hydrodynamic drag, thereby breaking its constant charge-to-friction ratio. Current drag-tag architecture relies on covalent attachment of polymers to each DNA molecule. We have recently proposed the use of micellar drag-tags in conjunction with sequence-specific hybridization of peptide nucleic acid amphiphiles (PNAAs). This work investigates the effect of multiple PNAA attachment on DNA resolution using MEKC. Simultaneous PNAA hybridization allows for the separation of long DNA targets, up to 1012 bases, using micellar drag-tags. Each PNAA handle independently interacts with the micellar phase, reducing the overall mobility of this complex relative to individual PNAA binding. The sequence- and size-based dependence of this separation technique is maintained with multiple PNAA binding over a range of DNA sizes. Results are accurately described by ELFSE theory, yielding $\alpha = 54$ for single-micelle tagging and $\alpha = 142$ for dual-micelle tagging. This method is the first example of a non-covalent drag-tag used to separate DNA of 1000 bases based on both size and sequence.

Keywords:
Capillary electrophoresis / DNA electrophoresis / DNA sequencing / End-labeled free-solution electrophoresis / Micellar electrokinetic chromatography
DOI 10.1002/elps.200700580

1 Introduction

End-labeled free-solution electrophoresis (ELFSE) is an electrophoretic separation method proposed as a faster alternative to capillary gel electrophoresis (CGE) for the length-based separation of DNA oligomers [1]. While such separations are commonplace in molecular biology labs, the main application envisioned for ELFSE is DNA sequencing. Here, oligomers up to 700 bases in length, and differing by only a single base must be accurately resolved. This technique relies on the addition of an uncharged group to DNA, commonly referred to as a "drag-tag", to break the constant charge-to-friction ratio of DNA and facilitate size-based separations in the absence of a sieving matrix [2]. The electrophoretic mobility (μ) of the resulting polyampholyte is given by the following expression:

$$\mu = \mu_0 \frac{L}{L + \alpha} \tag{1}$$

where L is the length of the DNA oligomer, μ_0 is its mobility in free solution, and α is a measure of the additional drag imparted on the DNA as a result of drag-tag addition, quantified in units of equivalent or "neutral" DNA bases. Since diffusion and other processes give rise to band broadening in free-solution electrophoresis, the resolution of populations of oligomers with nearly the same length decreases with the absolute length of those oligomers. This sets an upper bound on the length of read that can be analyzed in a DNA-sequencing context.

Correspondence: Professor James W. Schneider, Department of Chemical Engineering, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh PA 15213-3890, USA
E-mail: schneider@cmu.edu
Fax: +1-412-268-7139

Abbreviations: ELFSE, end-labeled free-solution electrophoresis; **PNAA**, peptide nucleic acid amphiphile; **TX-100**, Triton X-100 surfactant

To increase the length of read accessible by ELFSE, considerable effort has focused on the use of larger drag-tags to maximize the parameter α. Current ELFSE drag-tags designs consist of covalently attached, uncharged, water-soluble macromolecules including natural proteins, polymers, synthetic polypeptides and polypeptoids, and genetically engineered polypeptides [3–6]. To avoid loss of resolution between DNA oligomers of similar but unequal length, ELFSE drag-tags must be very monodisperse in size and not give rise to band broadening by adsorption to capillary walls.

Recently, we proposed the use of surfactant micelles as ELFSE drag-tags [7]. Micelles are water-soluble and form fairly monodisperse populations of structures with a tunable size and morphology. To encourage interaction with micelles, DNA oligomers are alkylated by hybridization of a nonpolar compound to a specific sequence on the oligomer of interest. The nonpolar tag is a peptide nucleic acid amphiphile (PNAA), an alkylated form of peptide nucleic acid (PNA) [8]. PNA are very specific, stable DNA binders [9] and are readily derivatized by peptide chemistry. These properties are retained when the PNA peptide is derivatized to form PNA amphiphiles [8, 10–12]. Importantly, the timescale for demicellization of Triton X-100 (TX-100) is in the order of milliseconds to seconds [13, 14], so PNAA-tagged DNA is expected to interact with many thousands of micelles during the run. Therefore, micelles are expected to impart an average drag on PNAA-tagged DNA that is highly uniform, even if the micelles are not perfectly monodisperse.

We have recently demonstrated the use of nonionic micelles of TX-100 as ELFSE drag-tags in CE [7]. To maintain the interaction of micelles with alkylated DNA, it is necessary to include micelles in the running buffer in a manner similar to MEKC methods [15, 16]. PNAA is hybridized to DNA in solution, and the mixture is injected into an uncoated CE capillary pre-filled with a buffer containing TX-100 micelles. The CE capillary is uncoated, and therefore a significant EOF opposes the electrophoresis of negatively charged objects in the capillary. On application of electric field, this combination of flows causes TX-100 micelles to perfuse through the sample plug containing PNAA-tagged DNA oligomers. Interaction of the micelles with the PNAA-tagged DNA brings about elution first of the short, tagged DNA followed by the longer, tagged DNA. Any untagged, free DNA has the highest electrophoretic mobility and therefore elutes last. The result is a length-based separation of DNA, with shortest DNA emerging first and longest DNA emerging last.

The goal of this work is to determine what effect end-attachment of two TX-100 micelles (Fig. 1) has on the PNAA-MEKC separation, and to determine the maximum length of tagged DNA that can be discriminated from untagged DNA. This would be particularly useful in the identification of specific nucleic-acid components from a biological sample, given both their length and the presence of a short binding sequence. Quantification of mRNA and microRNA are possible applications, and depending on the resolution provided, micellar drag-tags may be useful for DNA sequencing.

2 Materials and methods

2.1 Reagents

Streptavidin-agarose and 3,3′-diethylthiacarbocyanine iodide [$DiSC_2(5)$] were purchased from Aldrich (St. Louis, MO). TX-100 was purchased from Fluka (Milwaukee, WI). Bodipy 493/503 and DH5α *Escherichia coli* cells were purchased from Invitrogen (Carlsbad, CA). All PCR-related reagents were from Promega (Madison, WI). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA) and were of the highest grade available. PNAA were manually synthesized as described elsewhere [8, 10] and characterized on a PerSeptive Voyager STR MALDI-TOF mass spectrometer (Applied Biosystems, Foster City, CA). Observed and theoretical molecular weights were PNAA1 – 3510.3 (3509.5), PNAA2 – 3483.8 (3481.5). All synthetic DNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA), dissolved in either nuclease-free water or 50 mM Tris MES (pH 8.0), and used without further purification. Table 1 features all PNAA and DNA used in this study.

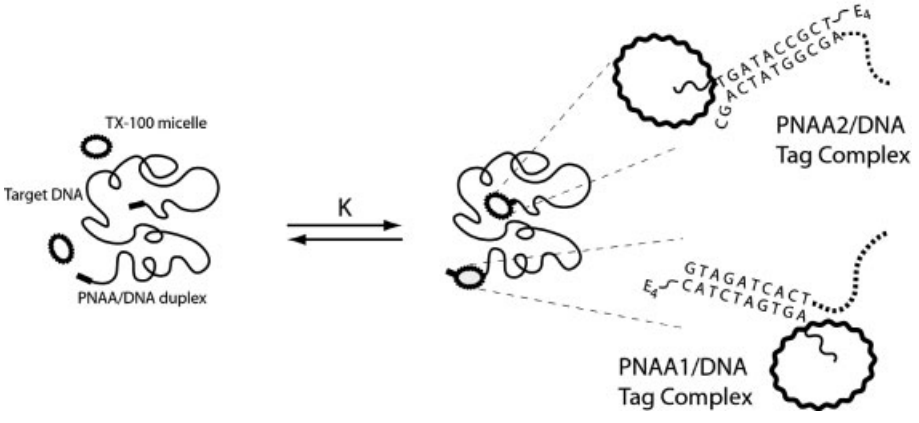

Figure 1. Schematic of TX-100 micelle interaction with a DNA target, mediated by PNAA/DNA duplex formation. Binding sites and sequences for PNAA1 and PNAA2 are featured. For analytical convenience, the TX-100 micelle and the PNAA/DNA duplex have been lumped together into a "tag complex". E = glutamic acid.

Table 1. PNAA and DNA used in this study

Type	Name	Sequence
PNAA	PNAA1	C_{18}-agtgatctac-E_4
	PNAA2	C_{18}-tgataccgct-E_4
Forward primer DNA	fp3-F	5′-F-GTAGATCACTCCGAATTCGTAATCATGTCATAGC
Reverse primer DNA	B-rp0	5′-B-TGATACCGCTGCGGATAACAATTTCACACAGG
	B-rp1	5′-B-GCTGATACCGCTAGCTCACTCATTAGGCACC
	B-rp2	5′-B-GCTGATACCGCTTCGTATGTTGTGTGGAATTGTG
	B-rp3	5′-B-GCTGATACCGCTTGGCCGATTCATTAATGCAGC
	B-rp4	5′-B-GCTGATACCGCTTGGATAACCGTATTACCGCCT
	B-rp5	5′-B-GCTGATACCGCTCGGAGCCTATGGAAAAACG
	B-rp6	5′-B-GCTGATACCGCTACTGAGATACCTACAGCGTGAGC
	B-rp7	5′-B-GCTGATACCGCTTGCCGCTAATCTGCTGCTTGCA

PNAA-binding sites are underlined. F = 6-carboxyfluorescein; B = Biotin.

2.2 PNAA/DNA hybridization

To ensure maximum hybridization, all PNAA-DNA mixtures were heated to *ca.* 85°C for 10 min in a dry bath incubator and allowed to cool to room temperature over the course of 75 min. Hybridization buffer was 50 mM Tris MES (pH 8.0) for CE experiments. These PNAA are incapable of binding to dsDNA due to their mixed base makeup and experimental pH utilized; therefore, all sequence-specific binding is attributed to PNAA and ssDNA.

2.3 Primer design

The forward primer contained a 6-carboxyfluorescein (FAM) appended to the 5'-end of the primer *via* a flexible hexa-ethyleneglycol spacer, which enabled DNA detection *via* fluorescence detection. Each reverse primer contained a biotin appended to the 5'-end. PNAA recognition sequences, denoted in Table 1, were appended to the 5'-end of all ssDNA products by inclusion in the forward and reverse primer design, rather than cloning it into the PCR template itself. This allowed for consistent PNAA-binding site location (relative to DNA strand end), regardless of overall target length.

2.4 PCR conditions

All PCR reactions were conducted in a SmartCycler (Cepheid, Sunnyvale, CA) at a total volume of 100 μL. Symmetric PCR reactions included the following reagents: 1000 nM forward and reverse primer, 0.2 mM dNTP mixture, 2.5 mM $MgCl_2$, 1 × PCR reaction buffer (50 mM KCl, 10 mM Tris-HCl pH 9, 0.1% TX-100), 0.04 ng/μL plasmid pSP64 template, and 0.04 U/μL *Taq* DNA polymerase. Plasmid pSP64 was produced in Subcloning Efficiency DH5α Competent *Escherichia coli* cells. An initial template denaturation step of 120 s at 95°C was conducted on each reaction. This was followed by 30 cycles of 95°C for 10 s (71, 96, 142, 224 bp) or 30 s (352 bp), 57°C for 15 s or 30 s, and 72°C for 15 s or 60 s. The purity of each PCR reaction was assessed with TAE-agarose gel electrophoresis. Full-length product bands were subsequently excised from the gel using a clean scalpel, then purified and resuspended in nuclease-free water using a Wizard SV-Gel PCR cleanup kit (PromegaI).

Asymmetric PCR used the following reagents: 1000 nM forward primer, 50 nM reverse primer, 0.2 mM dNTP mixture, 2.5 mM $MgCl_2$, 1 x GoTaq Flexi buffer, 0.04 ng/μL plasmid pSP64 template, and 0.04 U/μL GoTaq DNA Polymerase. An initial template denaturation step of 120 s at 95°C was conducted on each reaction. This was followed by 50 cycles of 95°C for 30 s, 57°C for 30 s, and 72°C for 60 s. DNA of lengths 71, 96, 142, 224, 352, 455, 660, and 1012 bases were used in this study. With these conditions, we were unable to generate a sufficient quantity of ssDNA longer than 1012 bases.

2.5 ssDNA generation

The ssDNA was generated and isolated using one of two methods, depending on target length. For shorter PCR products (71, 96, 142, 224, 352 bp) the 5'-biotin moiety, which was incorporated into the full-length PCR amplicon, was leveraged to immobilize full length dsDNA by introducing it to a packed column containing 400 μL streptavidin-agarose slurry pre-equilibrated in 10 mM Tris/1 mM EDTA/50 mM NaCl, pH 8.0. The mixture was allowed to bind for 20 min at room temperature, after which the column was washed four times with the aforementioned buffer to remove unbound primers and PCR reaction constituents. The ssDNA elution was accomplished by adding 600 μL of 0.2 M NaOH, and incubating the solution for 15 min at 37°C. Alkaline conditions have been shown to severely decrease dsDNA stability, inducing a helix-to-coil transition near room temperature for shorter strand lengths [17]. Samples then underwent three buffer exchanges with 50 mM Tris MES

(pH 8.0) using a 30 000 Da cutoff Microcon spin filter (Millipore, Milford, MA). Each sample underwent volume reduction to *ca.* 25 μL. Final nucleic acid concentration was between 10 and 500 nM (strand basis) for all samples, as judged by fluorescence measurements. To each DNA sample 1 μM C_{18}-agtgatctac-E_4 PNAA (PNAA1), 1 μM C_{18}-tgataccgct-E_4 PNAA (PNAA2), or both were added and the mixture was hybridized as previously described.

The biotin-streptavidin scheme was ill-suited for purification of ssDNA from longer PCR amplicons (455, 660, 1012 bp). These products displayed poor column retention, and required higher temperatures for strand melting, which compromised the agarose matrix integrity. For these amplicons, reverse primers were still labeled with a 5'-biotin moiety for experimental consistency, but went unused in the ssDNA capture scheme. Instead, ssDNA was directly created using asymmetric PCR. Asymmetric PCR utilizes unequal primer concentrations to generate both ss- and dsDNA in a single reaction; once the limiting primer is fully incorporated into dsDNA products, these serve as templates for the linear amplification of one DNA strand. In a typical reaction, ssDNA will be at a five- to tenfold higher concentration compared to dsDNA. Asymmetric PCR samples then underwent buffer exchange, volume reduction, and PNAA hybridization as previously described. If desired, asymmetric PCR can also be used to generate the short single-stranded products created with the biotin-streptavidin scheme.

2.6 MEKC

MEKC was performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped with an LIF detector. The capillary used was a 50-μm id fused-silica capillary (Polymicro Technologies, Phoenix, AZ), 31 cm total length, 21 cm length to detector. A suite of TX-100 concentrations in Tris MES (pH 8.0) were used as the running buffer for all samples. TX-100 concentrations tested were 0, 1.2, 2.4, 3.6, 4.8, 6.4, 8, 12, 16, 24, 36, and 48 mM. Hydrodynamic injection (0.5 psi for 5 s) was used to introduce sample into the capillary. Electrophoretic separation was conducted under normal polarity (from anode to cathode) with an electric field strength of 700 V/cm. The capillary and sample chambers were maintained at 22 and 10°C, respectively. LIF detection was performed at 488/520 nm excitation/emission. Bodipy 493/503 was added to the sample to determine the electroosmotic velocity of the running buffer in all MEKC experiments. Data collection and analysis were performed using 32 Karat software (Beckman Coulter).

To account for EOF drift and systematic variations of EOF with buffer viscosity, we have included intensity *versus* electrophoretic mobility plots, in addition to traditional time-domain plots. Migration time, t, of each species was converted to apparent mobility, μ_{app}, and then normalized with respect to the apparent mobility of the neutral marker ("EOF mobility," μ_{EOF}) to calculate the effective mobility, μ_{eff}:

$$\mu_{eff} = \mu_{app} - \mu_{EOF},\ \mu_{app} = \frac{IL}{tV} \tag{2}$$

Here, l is the length to the detector, L is the total capillary length and V is the applied voltage.

The suitability of Bodipy 493/503 for EOF determination was confirmed by comparing the calculated effective mobilities of DNA oligomers of several different lengths using this dye with that of methanol and benzyl alcohol, two accepted CE neutral markers. Effective oligomer mobilities were identical, within error, for all three neutral markers. The mobility of the neutral marker ("EOF mobility," μ_{EOF}) varied from $4.5–5.0 \times 10^{-4}$ cm^2/Vs, depending on the micelle concentration used. A full tabulation of the EOF mobilities observed in this work is given in the Supporting Information.

2.7 MEKC model implementation

All MEKC datasets were fit according to the model proposed by Grosser *et al.* [7]. In this treatment, the effective mobility of the PNAA-tagged DNA in the presence of micelles (μ_{eff}), is simply a weighted average of the intrinsic mobility of the bound target, μ^0_{hDNA}, and mobility of a bound target-micelle complex, μ^0_{mDNA}. This average mobility is controlled by both the partition coefficient K and the micelle concentration [M]. The working form of the model equation is:

$$\mu_{eff} = \frac{\mu^0_{hDNA} + K[M]\mu^0_{mDNA}}{1 + (K + C_{visc})[M] + KC_{visc}[M]^2} \tag{3}$$

The superscript "0" indicates the mobility in micelle-free buffer. K is defined as:

$$K = \frac{[mDNA]}{[hDNA][M]} \tag{4}$$

Here, "mDNA" represents micellated DNA, to include the PNAA-tagged DNA and its attached TX-100 micelle. The "hDNA" is the DNA with only its PNAA tag hybridized to it. A true partitioning process would therefore have the concentration of mDNA increasing with [M]. The values of K and μ^0_{mDNA} are calculated by non-linear regression from plots of μ_{eff} *vs.* [M]. This treatment also includes a correction, C_{visc}, for the change in buffer viscosity imparted by the high micelle concentration used in the separations. The value of this constant was measured using capillary viscometry inside the P/ACE MDQ, and determined to be 0.53 ± 0.02 mM^{-1}. Micelle concentration [M] was calculated by:

$$[M] = \frac{[S] - CMC}{N} \tag{5}$$

The CMC and aggregation number N were determined using the solvochromatic fluorescent probe Nile Red as previously described [11]. The CMC and aggregation number were found to be 0.33 mM and 128, respectively.

3 Results

3.1 MEKC separation of long ssDNA targets

To test the effect of micelle attachment to both ends of an ssDNA strand, we used two different PNAA tags (PNAA1 and PNAA2), which are complementary to the forward and reverse PCR primers, respectively. As shown in Fig. 1, PNAA1 binds to a 10-base-binding site at the 5'-end of the forward primer, while PNAA2 binds to a 10-base-binding site two bases removed from the 3'-end of the reverse primer. The ssDNA strands of varying length were produced by symmetric PCR followed by streptavidin purification or asymmetric PCR (aPCR) method using the same forward primer (fp1) and a series of reverse primers (Brp0-7) located at various distances from the forward primer on a pSP64 plasmid template (Table 1). The forward primer is linked to a 6-carboxyfluorescein moiety for LIF detection; hence only products (both ssDNA and dsDNA) incorporating the forward primer will be observed in LIF. The reverse primers are linked to a biotin moiety for purification purposes.

Prior to MEKC, a PCR product was incubated with either the PNAA1, PNAA2 tag, or both, at 85°C for 10 min to ensure maximal hybridization of the tags. A typical MEKC electropherogram for separation of a tagged, 142-base DNA target in a running buffer containing 12 mM TX-100 is depicted in Fig. 2. The same dataset is plotted as a function of electrophoretic mobility in Supporting Fig. S1. In Fig. S1, the time axis is converted to the apparent electrophoretic mobility by subtracting the EOF. This negates run-to-run variation in EOF velocity, determined by adding a neutral marker (Bodipy 493/503). This particular product was generated using symmetric PCR followed by streptavidin purification of ssDNA oligomers, and therefore does not contain dsDNA. A population of the untagged 142-base target elutes at an effective mobility (μ_{eff}) of -2.9×10^{-4} cm^2/Vs (3.2 min), in good agreement with published values. [18] Upon incubation of the DNA oligomers with either PNAA1 or PNAA2, the magnitude of the free DNA peak is significantly reduced in favor of a second population, centered at -2.3×10^{-4} cm^2/Vs (2.1 min) and -2.2×10^{-4} cm^2/Vs (2.0 min), respectively. When both probes are added, a majority peak appears at -1.7×10^{-4} cm^2/Vs (1.8 min), shifted away from each remaining peak. These shifts result from the transient association of multiple TX-100 micelles with the alkane chain of the PNAA tags during the MEKC run.

We believe that the free DNA population results from the stripping of some PNAA tags from their DNA targets during the initial phases of the run, where the tagged DNA may traverse through sharp electric field gradients. The free DNA population persisted even with the addition of excess PNA tags; furthermore, colorimetric assays performed in our lab indicated a 1:1 binding stoichiometry for the PNAA/DNA pairs studied here. There is also the small, varied presence of minor peaks throughout the electropherogram (second curve, μ_{eff} of -1.6×10^{-4} cm^2/Vs is 1.7 min, for example). They

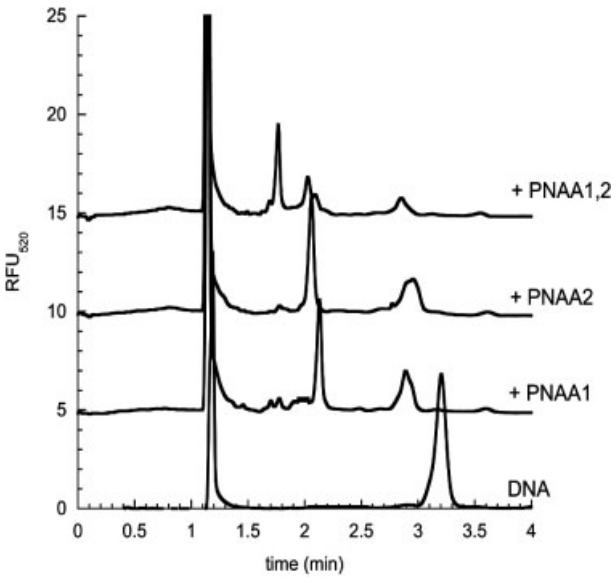

Figure 2. Electropherogram for elution of a 142-base DNA target and the same target bound with PNAA1, PNAA2, or both. The large peak between 1 and 1.5 min represents the elution of a neutral EOF marker (Bodipy 493/503). [TX-100] = 12 mM (corresponding to a [micelle] of approximately 0.094 mM). [DNA] = 10–100 nM (strand basis), [PNAA] = 1 μM. The top three electropherograms have been arbitrarily shifted in the positive *y* direction for display purposes.

presumably represent non-specific products generated during PCR and are minimized by raising the annealing temperature, albeit at the expense of product generation. The elution times for the putative "free DNA" varied between 2.9 and 3.3 min for the runs in Fig. 2. This is due entirely to EOF drift, and each gives the same electrophoretic mobility when EOF is accounted for (Supporting Fig. S1). The effects of EOF drift are largest for components with the longest elution times, hence the free DNA populations show the largest variability in elution time.

To properly gauge the drag imparted by the attachment of one or two micelles to tagged DNA, we must determine the mobility of the tagged DNA when attached to a micelle ("micellated DNA", μ^0_{mDNA}). This quantity is not the observed electrophoretic mobility, but rather a limiting case achieved as the fraction of time the DNA spends attached to a micelle approaches unity. That fraction (*fmic*) increases with the concentration of micelles in the running buffer ([M]) [7]:

$$f_{Mic} = \frac{K[M]}{1 + K[M]} \quad (6)$$

where K is the partition coefficient of the tagged DNA with the TX-100 micelles in the running buffer: Fig. 3 is a series of seven electropherograms with the TX-100 concentration in the running buffer increasing from 0 to 24 mM. As expected, the mobility of the tagged DNA shifts dramatically with [M]. There is also a small positive shift in the mobility of the

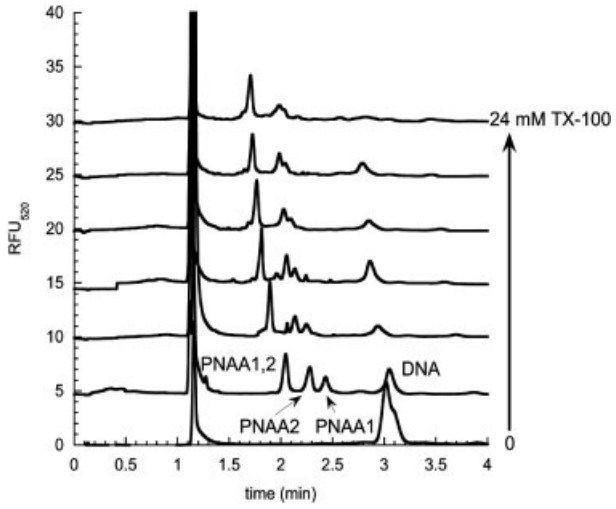

Figure 3. The effect of [TX-100] on the migration of unbound and PNAA-bound 142-base target DNA in MEKC. The large peak between 1 and 1.5 min represents the elution of a neutral EOF marker (Bodipy 493/503). TX-100 concentrations used were 0, 1.2, 2.4, 4.8, 8, 12, and 24 mM. Increasing TX-100 concentrations are displayed in ascending order, bottom to top. Electropherograms have been arbitrarily shifted in the positive-y direction for display purposes.

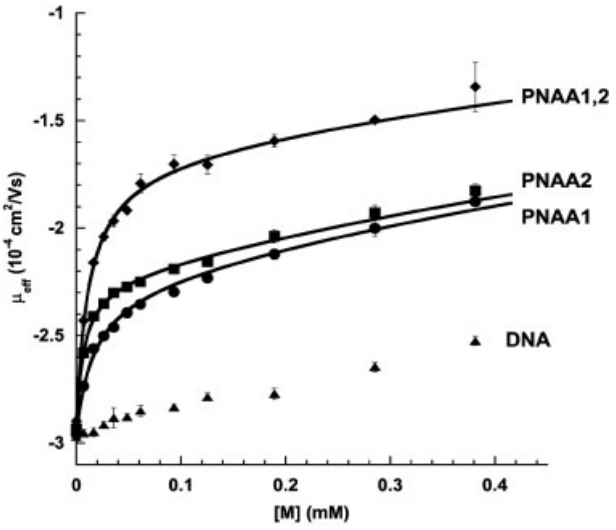

Figure 4. Measured effective mobility (μ_{eff}) comparison of free and PNAA-tagged DNA 142 bases in length as a function of micelle concentration [M]. Each data point represents an average of a minimum of three independent experiments. Datasets were fit using the MEKC model equation (Eq. 3), as previously described. Fit parameters μ^0_{mDNA} and *K* are listed in Table 2.

untagged DNA due to increasing running buffer viscosity. The quantities *K* and μ^0_{mDNA} are obtained by a two-parameter fit of mobility *vs.* [M] data to Eq. (3) (Fig. 4). The fit parameters are listed in Table 2. There is a large disparity in partition coefficient *K* between the PNAA1 and PNAA2 tag, which can be traced to the orientation of the PNAA alkane. As shown in Fig. 1, the tail of PNAA2 points away from the DNA interior, 2 bases from the chain end, while PNAA1 faces the DNA interior, 10 bases from chain end. Presumably, the tail of PNAA2 is more accessible for interaction with the micellar phase. Values for the mobility of micellated DNA (μ^0_{mDNA}) are comparable for both PNAA1 and PNAA2 tags. Using these K and μ^0_{mDNA} values, the unique mobility trends of PNAA1 and PNAA2-tagged ssDNA can readily be explained. At low to moderate TX-100 concentrations, a higher *K* value for PNAA1 relative to PNAA2 will result in a less negative mobility (faster elution time), as this species experiences a heightened micellar phase interaction. At higher TX-100 concentrations, differences in partitioning are mitigated as samples migrate at or near their μ^0_{mDNA} values, which are equivalent for PNAA1 and PNAA2-tagged ssDNA. As expected, the doubly tagged DNA has a mobility much closer to zero.

To obtain reliable values for the ELFSE drag parameter α (Eq. 1), we repeated the above analysis for DNA target lengths of 71, 96, 142, 224, 352, 455, 660, and 1012 bases. Plots of mobility *vs.* [M] for the doubly tagged DNA, along with fits to Eq. (3), are given in Fig. 5. Corresponding values for μ^0_{mDNA} and *K* are given in Table 3 and an electropherogram analogous to Fig. 2 for 455-base DNA targets is given in Fig. S3 (Supporting Information). Importantly, we find that targets up to 1012 bases in length have a mobility that is significantly shifted from that of untagged DNA. In principle, this makes possible the separation of mRNA con-

Table 2. Calculated mobility of micellated DNA (μ^0_{mDNA}) and partition coefficient (*K*) for PNAA-bound 142-base target, as computed by fitting Eq. (3) to the data of Fig. 4

Tag	μ^0_{mDNA} (10^{-4} cm^2/Vs)	*K* (mM^{-1})
PNAA1	-2.28 ± 0.05	62 ± 3
PNAA2	-2.24 ± 0.02	144 ± 8
Both	-1.69 ± 0.03	84 ± 4

Table 3. Calculated mobility μ^0_{mDNA} and partition coefficient *K* for micellated DNA interacting with doubly tagged target DNA (both PNAA1 and PNAA2 attached), as computed by fitting Eq. (3) to the data of Fig. 5

DNA length (bases)	μ^0_{mDNA} (10^{-4} cm^2/Vs)	*K* (mM^{-1})
71	-1.35 ± 0.10	145 ± 58
96	-1.51 ± 0.07	59 ± 5
142	-1.69 ± 0.03	83 ± 3
224	-1.94 ± 0.04	111 ± 3
352	-2.10 ± 0.02	94 ± 9
455	-2.24 ± 0.02	155 ± 11
660	-2.35 ± 0.02	408 ± 58
1012	-2.65 ± 0.02	240 ± 105

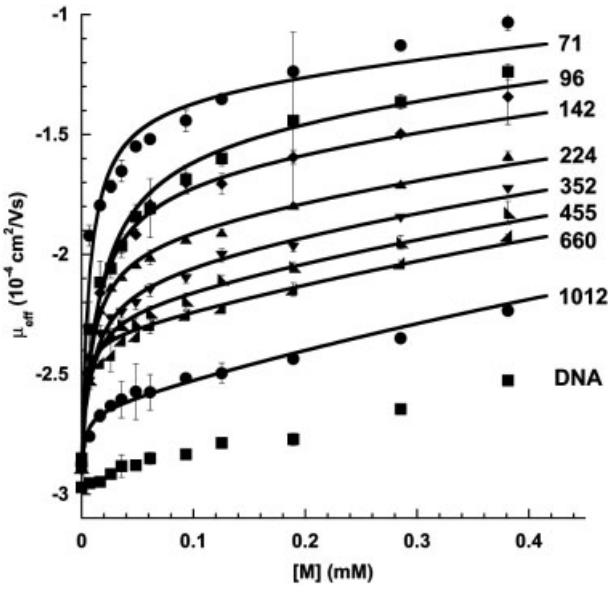

Figure 5. Measured effective mobility (μ_{eff}) comparison of doubly tagged (PNAA1 + PNAA2) target DNA as a function of micelle concentration. DNA lengths tested were 71, 96, 142, 224, 352, 455, 660, and 1012 bases. Each data point represents an average of a minimum of three independent experiments. Datasets were fit using the MEKC model equation (Eq. 3), as previously described. Fitted μ^0_{mDNA} values are given in Table 3.

taining a specific sequence from non-targets by a dual tagging/MEKC approach. Variations in K among doubly tagged DNA strands suggest changes in the DNA coil conformation that promote or hinder tail-micelle interactions, depending on tail orientation (Fig. 1).

3.2 ELFSE implementation

Nearly all the published ELFSE methods utilize some uncharged, covalently attached drag-tag [3–6, 19]. We need to account for two physical effects in order to fairly compare our values of α to those using covalently attached drag-tags that are uncharged (or nearly so). First, we must use the fitted value of μ^0_{mDNA} to describe the electrophoretic mobility of the micellated DNA. This quantity represents the mobility of the tagged DNA in the micelle-containing running buffer as f_{mic} approaches unity. While this value can never reach unity, values above 0.9 are easily obtained in practice. For example, the data collected for 142-base oligomers tagged with PNAA1 in 24 mM TX-100 has K = 62.3 mM^{-1} and therefore f_{mic} = 0.92. Under these conditions, the tagged DNA, interacting with myriad TX-100 micelles during the elution in MEKC, has a mobility similar to that expected if the micelle were permanently attached.

Another consideration is that the PNAA/DNA duplex formed at the PNA-binding site (Fig. 1) has a markedly different mobility than the longer stretch of unbound, ssDNA. As such, Eq. (1) will not give a good fit to the mobility *vs.* length data, particularly at short lengths where the lengths of unbound, ssDNA and duplexed DNA are comparable. Good fits are obtained by defining a drag-tag complex that includes the TX-100 micelle as well as the PNAA/DNA duplex (including any overhanging bases proximal to the micelle) as shown in Fig. 6. Using this definition, the remainder of the polyampholyte contains only ssDNA, and the α values calculated can be assigned an equivalent hydrodynamic radius (R_h) using established models. We then recast Eq. (1) to account for a drag-tag with a nonzero mobility (μ^0_{tag}):

$$\mu^0_{mDNA} = \mu_0 \frac{L}{L+\alpha} + \mu^0_{tag} \frac{\alpha}{L+\alpha} \quad (7)$$

Within this new definition, L is the length of ssDNA extending from the PNAA/DNA duplex on the side distal to the micelle. In the case of PNAA1, this will decrease L by 10 bases, and by 12 bases for PNAA2 (Fig. 1). The mobility of this newly defined drag-tag is readily obtained by the above MEKC analysis, using short DNA oligomers that do not extend beyond the PNAA-binding site (L = 0). In the case of PNAA1, the oligomer is 10-bases long, and for PNAA2 it is 12-bases long. As shown in Table 4, the values for both are similar, and significantly negative (-0.82 and -0.74×10^{-4} cm^2/Vs, respectively). This is due to the charge on the duplexed DNA and the glutamic acid residues (Fig. 1) on the PNAA, now considered part of the drag-tag.

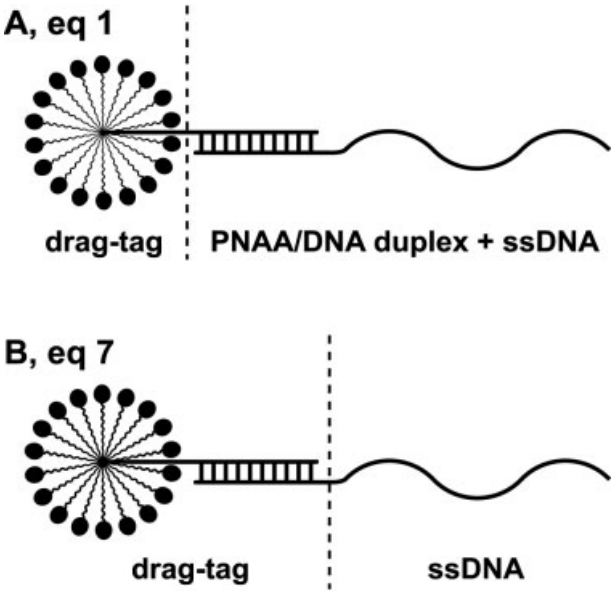

Figure 6. Assignment of drag-tag mobility in the micelle-PNAA-DNA polyampholyte complex. In (A), we define the drag-tag to be the TX-100 micelle alone, which we presume to be uncharged, and Eq. (1) can be used to describe the polyampholyte mobility. In (B), we define the drag-tag to be the micelle along with the PNAA tag and the 10–12-base section of DNA bound to the PNAA (including any overhanging bases proximal to the micelle). In this case, the more general form of the ELFSE equation (Eq. 7) more accurately describes the polyampholyte mobility, as the drag-tag mobility is nonzero.

Table 4. Effective friction coefficient (α) measured for PNAA1, PNAA2, and simultaneous addition

Drag-tag	μ^0_{tag} 10^{-4} cm²/Vs	Effective friction coefficient α Eq. (1)	Effective friction coefficient α Eq. (7)
PNAA1	−0.82 ± 0.05	54 ± 9	75 ± 13
PNAA2	−0.74 ± 0.06	58 ± 9	77 ± 12
PNAA1 + PNAA2	−0.78 ± 0.08	142 ± 9	193 ± 20

Equation (1) considers the tag to be uncharged, and consists only of the Triton X-100 micelle (Fig. 6A). Equation (7) considers the tag to be charged, and consist of the micelle along with the bound portion of the DNA (Fig. 6B). The μ^0_{tag} value for simultaneous addition case was estimated by averaging the individual PNAA1, PNAA2 tag values. Error was estimated by standard error propagation techniques.

For fitting purposes, Eq. (7) can be linearized as follows:

$$\frac{\mu_0 - \mu^0_{mDNA}}{\mu^0_{mDNA} - \mu^0_{tag}} = \alpha\left(\frac{1}{L}\right) \quad (8)$$

Plots of the left hand side of Eq. (8) *versus* $1/L$ should therefore be linear with slope equal to the ELFSE drag parameter α. The results of fitting the MEKC experimental data to Eq. (8) are displayed in Fig. 7 and appear to be linear except for the shortest oligomers with PNAA1 and PNAA2 both bound. The data for L = 1012-oligomers did not fall on the linear fit, and this may be due to the formation of secondary structure for these longer populations. We therefore did not include the L = 1012 data in further analysis.

The fitted values of α are listed in Table 4, and compare very favorably with those in the literature. To our knowledge, the largest value of α that has been reported is α = 70, observed using an end-attached, genetically engineered protein polymer [5]. Our MEKC results using either the PNAA1 tag or the PNAA2 tag are in the range of α = 54–58. The dual-tagging result, α = 142, is greater than the sum of the singly tagged results as has been reported for other systems [6].

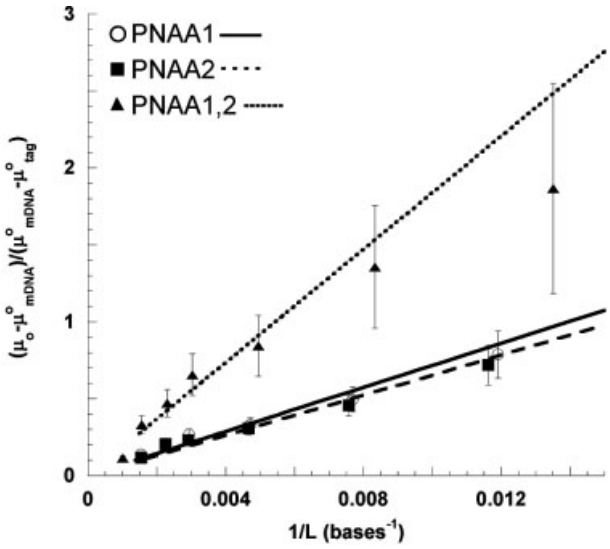

Figure 7. Plot of $(\mu_o - \mu^0_{mDNA}/\mu^0_{mDNA} - \mu^0_{tag})$ *versus* $1/L$, measured in DNA bases. μ^0_{mDNA} was estimated by fitting each dataset with Eq. (3). Representative fits are featured in Figs. 4 and 5. μ^0_{tag} was estimated by fitting as described in the text. The slope and intercept of each line were used to determine the effective friction coefficient α for each system, according to Eq. (8). Values are listed in Table 4.

4 Discussion

4.1 Hydrodynamic drag imparted by transiently attached micelles

The lifetime of a TX-100 micelle, or slow dissociation mode, has been measured to be between thousandths of a second to 3.5 s, with TX-100 monomer exchange, or fast dissociation mode, occurring at even faster rates [13, 14, 20]. As such, we expect that PNAA-tagged DNA interacts with an individual TX-100 micelle for only a short time, approximately milliseconds to seconds. During that time, the TX-100 micelle will be exchanging its monomers with those in solution. This dynamic association allows the partitioning to proceed under quasi-equilibrium conditions. It is possible, however, to assign an average, equivalent hydrodynamic drag to transiently attached micelles by converting the α values to their hydrodynamic radius. Following the procedure of Desruisseaux *et al.* [21] we first calculate the hydrodynamic radius (R_h) of α ssDNA bases under the solution conditions:

$$R_h^2 \cong \left(\frac{2}{3}\right)^2 \frac{b\alpha p}{3}\left[1 - 3\left(\frac{p}{b\alpha}\right) + 6\left(\frac{p}{b\alpha}\right)^2 - 6\left(\frac{p}{b\alpha}\right)^3\left(1 - e\frac{-b\alpha}{p}\right)\right] \quad (9)$$

where p is the persistence length of the DNA in 50 mM Tris HCl (3.3 nm [22]), and b is the contour length of ssDNA (0.43 nm [22]). Equation (9) gives a value of R_h = 3.4 nm for α = 75 (PNAA1) and R_h = 3.5 nm for α =77 (PNAA2). Dynamic light scattering experiments performed in our lab using the same buffer conditions yielded R_h = 3.87 ± 0.05 nm, in reasonable agreement with the value predicted by Eq. (9). We conclude that the transiently attached micelles yield about the same hydrodynamic drag as an isolated TX-100 micelle in the same buffer.

Heller *et al.* [4] reported an α value of 25–30 for the protein streptavidin and Won *et al.* [5] reported α = 70 for a genetically engineered, protein polymer. Our values of 75–77 are certainly competitive, but because our drag-tag carries some negative charge it would not be expected to shift the mobility of its attached DNA as much as an uncharged equivalent used in previous studies. We can approximate a value of α for a hypothetical uncharged drag-tag with the same mobility shift by equating Eqs. (1) and (7) in the limit of long L (valid for L >100, comparison plotted in Fig. 8):

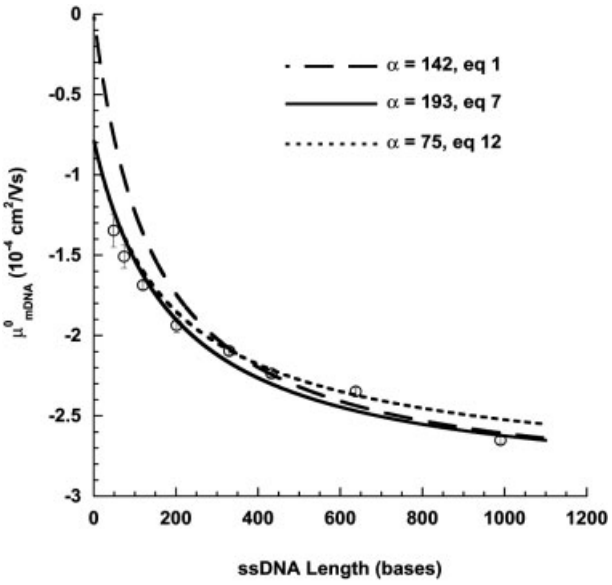

Figure 8. Mobilities of micellated DNA (μ^0_{mDNA}) tagged with both PNAA1 and PNAA2. Fits are to Eq. (1) (drag-tag definition A, Fig. 6), Eq. (7) (drag-tag definition B, Fig. 6), and Eq. (12).

$$\alpha_1 \cong \alpha_2\left(1-\frac{\mu^0_{tag}}{\mu_0}\right) \tag{10}$$

where α_1 is the approximate ELFSE parameter calculated by Eq. (1) (uncharged drag tag), and α_2 is the exact ELFSE parameter for the negatively charged drag-tag complex. Using Eq. (1), we find that this micelle-tagging system provides about the same mobility shift as an uncharged, covalently attached drag-tag having α = 54-58, which is still a large value. It could be increased further by using larger, nonionic micelles as drag-tags, or by using an uncharged linker for the attachment of ssDNA to the micelle.

4.2 Non-additive effect of dual-micelle tagging

It has been demonstrated that covalent attachment of drag-tags to either end of DNA oligomers provides a greater drag than twice the drag imparted by each tag individually [6]. This "end effect" arises from difference in hydrodynamic interactions among monomers in the chain, such that monomers near the chain ends are subject to higher hydrodynamic friction than those in the chain interior [23]. The net result is that the mobility of the polyampholyte is a weighted average of the mobilities of each monomer unit, with monomer units near chain ends receiving the highest weight. Tagging on both ends not only adds a second tag, it displaces charged DNA monomers to the interior of the coil, reducing their weighting factor. A recent study [6], using a wide range of drag-tags covalently attached to ssDNA, concluded a 6–9% enhancement of the drag when tagging both ends, compared to twice the value observed when tagging one end.

We observe a similar effect in our system. Dual tagging yielded α = 193, and the sum of the α from PNAA1 tagging alone (α = 75) and that from PNAA2 tagging alone (α = 77) is 152. This represents a 21% enhancement, which is somewhat larger than results reported for the covalent tagging approach. However, it does appear to be consistent with the end-effect theory. Following the approach of Meagher *et al.* [6], we approximate the weighting function Ψ as follows:

$$\Psi\left(\frac{n}{N}\right) = -0.65 + 0.62\left(\frac{n}{N}\right)^{-1/4} + 0.62\left(1-\frac{n}{N}\right)^{-1/4} \tag{11}$$

where n is the monomer in question and N is the total number of monomers in the polyampholyte. Next, we regard the drag-tag to be a Gaussian chain with α monomers, each giving the same drag contribution as an ssDNA monomer. For accuracy, we choose to define the drag-tag to include the PNAA tag plus its hybridized DNA (Scheme B, Fig. 6). As such, we need to include the mobility of the negatively charged drag tag ($\mu^0_{tag} = -0.78\times10^{-4}$ cm^2/Vs). The mobility of ssDNA with micelles attached to either end is given by:

$$\mu^0_{mDNA} = \frac{1}{L+2\alpha}\left[\int_0^{\alpha} 2\mu^0_{tag}\Psi\left(\frac{n}{L+2\alpha}\right)dn + \int_{\alpha}^{\alpha+L}\mu_0\Psi\left(\frac{n}{L+2\alpha}\right)dn\right] \tag{12}$$

Here, L represents the length of ssDNA between the drag-tags defined in Fig. 6B, 22 bases shorter than pre-hybridized length. Substituting the value α = 76 (obtained from ssDNA tagged only with the average of PNAA1 and PNAA2, Table 4), we obtain the solid line plotted in Fig. 8, which agrees very well with the MEKC-obtained data. The result of Eq. (12) actually appears to fit data for the longer oligomers better than Eq. (7), which does not account for the end effect. For comparison purposes, the end effect was not considered in the determination of α, as is the common practice. However, the average value obtained for individual PNAA labeling (α = 76) provides a good fit of the dual-labeled data when used with Eq. (12), which does consider the end effect.

There appears to be a systematic deviation from the theoretical curves for the shortest tagged oligomers. This may be statistically insignificant as the error bars nearly overlap with the curves. It may also be due to an increased likelihood that a short, dual-tagged oligomer may bind a single TX-100 micelle, forming a looped structure with a more negative mobility than expected, skewing the average mobility in the negative direction. Dual-micelle tagging of synthetic oligomers 20, 32, and 46 bases in length do not result in a second shifted population (data not shown).

4.3 Applications of the method

The ssDNA, rather than RNA, was used in this study in order to directly compare results with prior ELFSE reports using covalent drag-tag modification. However, we believe this system may be directly applied to RNA detection, quantification,

and purification from biological samples. The use of CE for RNA analysis as an alternative to reverse-transcriptase PCR and microarray hybridization has been recently verified [24–27]. Micellar drag-tag separation of RNA offers a number of advantages, including extreme sequence specificity via PNAA hybridization, size-based purification, short analysis time, a sieving matrix-free separation environment, and direct quantification of RNA *via* probe fluorescence. Moreover, with the RNA size known *a priori*, this method will be capable of simultaneous RNA identification and quantification. Detection of biologically derived RNA will be the subject of a future study.

The method may also be applicable to DNA sequencing, generally considered the most promising application of ELFSE methods. In this application, Sanger extension products differing by only a single base in length must be discriminated. Certainly, one can envision hybridization of the PNAA tag to the forward primer sites on the extension products to give rise to interaction with micelles in MEKC. The electropherograms presented here do not have a sufficient resolution for separation of Sanger products, but this is most likely due to our use of hydrodynamic injection, rather than electrokinetic injection. The latter, which yields extremely sharp sample bands by a sample stacking method, cannot be implemented in bench-top CE instrumentation without EOF suppression. Other instrumentation, including microfluidic platforms, could be designed to address this shortcoming. Of course, to be useful for ELFSE, these drag-tags must behave as nearly monodisperse entities. Populations of TX-100 micelles are fairly monodisperse; dynamic light scattering experiments performed in our lab using the same buffer conditions yielded $R_h = 3.87 \pm 0.05$ nm for TX-100. Furthermore, as mentioned earlier, there is expected to be an enhanced monodispersity brought about by sampling many micelles during the elution in MEKC.

It is desirable to maintain EOF in order to reduce run times for DNA sequencing. As noted by McCormick and Slater [28], in ELFSE methods the longest DNA (*i.e.* the most difficult to resolve), has the highest electrophoretic mobility. As such, once those peaks are resolved and detected, the remaining peaks are also resolved but cannot be detected until they reach the end of the capillary. EOF flows can be used to convect these "over-resolved" peaks past the detector more quickly, resulting in shorter run times. An important advantage of our system is that EOF is not disrupted by the presence of TX-100 micelles. Triton X-100 does adsorb to silica as hemi-micelles, but with no effect on silica's surface potential [29] or on the magnitude of EOF. It may also have a beneficial passivating effect that protects the capillary from nonspecific adsorption and concomitant EOF drift and/or peak broadening. We have been able to obtain reliable results from the same, bare capillary after thousands of consecutive runs. Still, to be competitive with current sequencing methods, it will be necessary to transition to larger micelles to realize substantially greater α values. Moderate gains in TX-100 micelle size can be realized by increasing operating temperature [30]. We are currently in the process of identifying microemulsions [31, 32] and liposomes [33] that are uncharged and compatible with the PNAA-MEKC method.

4.4 Conclusions

We have presented a new implementation of ELFSE that hybridizes PNAA to specific sequences on target DNA to encourage their binding to nonionic micelles of Triton X-100. The micelles bind transiently to the PNAA tag, providing a drag on the DNA oligomer roughly equivalent to the hydrodynamic drag of an isolated micelle. The tags are slightly negatively charged and give an effective drag coefficient of $\alpha = 75$–77 for single tagging and $\alpha = 193$ for double tagging. These transiently bound drag-tags would give the same resolving power as an uncharged, covalently attached drag-tag with $\alpha = 54$–58 ($\alpha = 142$ for double tagging). The latter represents that highest value of α reported for ELFSE methods. The substantial friction imparted by micelle tagging makes discrimination of oligomers up to 1012 bases in length possible. The PNAA-MEKC method emerges as a promising method for DNA sequencing and quantification of RNA.

The authors acknowledge the National Science Foundation (BES-0093538), the Arnold and Mabel Beckman Foundation, the Air Force Office of Scientific Research, and the DARPA SIMBIOSYS program for financial support of this work. J. M. S. was supported in part by an NIH training grant (5 T32 GM065100-03).

The authors have declared no conflict of interest.

5 References

[1] Mayer, P., Slater, G. W., Drouin, G., *Anal. Chem.* 1994, *66*, 1777–1780.

[2] Meagher, R. J., Won, J.-I., McCormick, L. C., Nedelcu, S. *et al.*, *Electrophoresis* 2005, *26*, 331–350.

[3] Haynes, R. D., Meagher, R. J., Won, J.-I., Bogdan, F. M., Barron, A. E., *Bioconjug. Chem.* 2005, *16*, 929–938.

[4] Heller, C., Slater, G. W., Mayer, P., Dovichi, N. *et al.*, *J. Chromatogr. A* 1998, *806*, 113–121.

[5] Won, J.-I., Meagher, R. J., Barron, A. E., *Electrophoresis* 2005, *26*, 2138–2148.

[6] Meagher, R. J., McCormick, L. C., Haynes, R. D., Won, J.-I. *et al.*, *Electrophoresis* 2006, *27*, 1702–1712.

[7] Grosser, S. T., Savard, J. M., Schneider, J. W., *Anal. Chem.* 2007, *79*, 9513–9519.

[8] Vernille, J. P., Kovell, L. C., Schneider, J. W., *Bioconj. Chem.* 2004, *15*, 1314–1321.

[9] Ratilainen, T., Holmen, A., Tuite, E., Nielsen, P. E., Norden, B., *Biochemistry* 2000, *39*, 7781–7791.

[10] Vernille, J. P., Schneider, J. W., *Biotechnol. Prog.* 2004, *20*, 1776–1782.

[11] Lau, C., Bitton, R., Bianco-Peled, H., Schultz, D. G. *et al.*, *J. Phys. Chem. B* 2006, *110*, 9027–9033.

[12] Marques, B. F., Schneider, J. W., *Langmuir* 2005, *21*, 2488–2494.

[13] Rillaerts, E., Joos, P., *J. Phys. Chem.* 1982, *86*, 3471–3478.

[14] Patist, A., Kanicky, J. R., Shukla, P. K., Shah, D. O., *J. Coll. Int. Sci.* 2002, *245*, 1–15.

[15] Terabe, S., Otsuka, K., Ichikawa, K., Tsuchiya, A., Ando, T., *Anal. Chem.* 1984, *56*, 111–113.

[16] Terabe, S., *Trends Anal. Chem.* 1989, *8*, 129–134.

[17] Lando, D. Y., Haroutiunian, S. G., Kul'ba, A. M., Dalian, E. B. *et al.*, *J. Biomol. Struct. Dyn.* 1994, *12*, 355–366.

[18] Stellwagen, N. C., Gelfi, C., Righetti, P. G., *Biopolymers* 1997, *42*, 687–703.

[19] Nedelcu, S., Slater, G. W., *Electrophoresis* 2005, *26*, 4003–4015.

[20] Fainerman, V. B., Makievski, A. V., *Colloids Surf.* 1993, *69*, 249–263.

[21] Desruisseaux, C., Long, D., Drouin, G., Slater, G. W., *Macromolecules* 2001, *34*, 44–52.

[22] Tinland, B., Pluen, A., Sturm, J., Weill, G., *Macromolecules* 1997, *30*, 5763–5765.

[23] Long, D., Dobrynin, A. V., Rubinstein, M., Ajdari, A., *J. Chem. Phys.* 1998, *108*, 1234–1244.

[24] Al-Mahrouki, A. A., Krylov, S. N., *Anal. Chem.* 2005, *77*, 8027–8030.

[25] Zhong, W., Yeung, E. S., *Anal. Chem.* 2003, *75*, 4415–4422.

[26] Goldsmith, J. G., Ntuen, E. C., Goldsmith, E. C., *Anal. Biochem.* 2007, *360*, 23–29.

[27] Kolesar, J. M., Allen, P. G., Doran, C. M., *J. Chromatogr. B* 1997, *697*, 189–194.

[28] McCormick, L. C., Slater, G. W., *Electrophoresis* 2006, *27*, 1693–1701.

[29] Gu, T., Zhu, B., *Colloids Surf.* 1990, *44*, 81–87.

[30] Streletzky, K., Phillies, G. D. J., *Langmuir* 1995, *11*, 42–47.

[31] Watarai, H., *Chem. Lett.* 1991, 391–394.

[32] Huie, C. W., *Electrophoresis* 2006, *27*, 60–75.

[33] Nakamura, H., Sugiyama, I., Sano, A., *Anal. Sci.* 1996, *12*, 973–976.

