

# Carnegie Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

## THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Doctor of Philosophy

**TITLE** Direct MicroRNA Detection Using Kilobase DNA Nanotags

In Rapid Gel-Free Electrophoresis

**PRESENTED BY** Johnathan Goldman

**ACCEPTED BY THE DEPARTMENT OF**

Chemical Engineering

James Schneider

5/4/15

ADVISOR

DATE

Lorenz Biegler

5/4/15

DEPARTMENT HEAD

DATE

**APPROVED BY THE COLLEGE COUNCIL**

Vijayakumar Bhagavatula

DEAN

DATE



**Direct MicroRNA Detection Using Kilobase DNA Nantoags In Rapid Gel-Free Electrophoresis**

Submitted in partial fulfillment of the requirements for  
the degree of  
Doctor of Philosophy  
in  
Chemical Engineering

Johnathan M. Goldman

B.S., Chemical Engineering, Northeastern University

Carnegie Mellon University  
Pittsburgh, PA

May, 2015



# Acknowledgements

I would like to begin by thanking my advisor, James Schneider, for helping me become the researcher I am today. As I was putting this thesis together, I went back through all of my work over the past 4 years. I could see Jim's influence improving my critical thinking, ability to successfully execute experiments, organization, and presentation/writing style. I have taken away so much from our time together and cannot thank you enough for molding me into the successful researcher I am today.

I would also like to thank all the past and present graduate students in the Schneider group; Stephen Istivan, Angela Jones, Max Farhenkopf, Daniel Bishop, and Randall Gamble. I gained not only insightful views on experiments and results from our conversations, but also a place I looked forward to going to work every day thanks to you. I also could not have completed more than half of the work in this thesis without the help from the extraordinarily talented undergraduate researchers who I worked with at CMU; Heather Dolan, Li Ang Zhang, Sarah Narburgh, Kristina Jang, Soyoung Kim, Palak Pujara, and Ariana Zito-Wolf. Thank you for teaching me how to become a successful manager and being researchers I could heavily rely on when the goings got tough.

I am also thankful for being part of The Center for Nucleic Acids Science and Technology (CNAST). CNAST not only funded my research through the David Scaife Family Charitable Foundation, but also granted me unique access to researchers from chemistry, biology, and physics all interested in nucleic acids.

I'd like to specifically thank Bruce Armitage from CNAST who acted as my unofficial co-advisor. Your enthusiasm and encouragement truly inspired me to work hard for the results presented in this thesis.

I'd like to thank the committee members, Robert Tilton, Kathryn Whitehead, Aditya Khair, and Bruce Armitage, chaired by my advisor James Schneider, for taking the time to read and listen to my dissertation.

One of the things I'm most thankful for is the new family I've gained from my time here. So many great people have entered my life and welcomed me into theirs over the years. Between Sunday night dinners at the ChemE house, to flag football, floor hockey and softball (IM and GSA champions, both by walk-offs) with the Sudsuckers, to the office comradery, these were the reasons I have truly cherished my time at CMU. Although there are many names I could list, I'd like to thank my fellow graduating classmates Stephen Spagnol and Patrick Boyer. We have been through it together every step of the way, starting with quals, to proposal, and now dissertations. Without having you guys in the office with me to work late nights or grab some drinks to celebrate each of our achievements, I would have had a much more difficult time here. Congratulations guys.

Finally, I'd like to thank Rebecca Ball for putting up with all of my stress while writing this thesis. Without your love and support, this could not have been done. I'm looking forward to being there for you when you begin writing your thesis and stressing out as much as I did.

# Abstract

The abnormal expression of microRNA (miRNA) within tissues has been linked to the onset of many diseases and cancers. The ability to employ miRNA as biomarkers for disease requires sensitive, high-throughput detection, ideally without expensive indirect enzymatic processing. We present our development of a direct, sensitive sandwich hybridization miRNA detection assay. We achieve hybridization of two probes to the short miRNA targets using high affinity PEG  $\gamma$ -carbon modified peptide nucleic acid amphiphiles ( $\gamma$ PNAA). The  $\gamma$ PNAA enables hybridization of a second, highly fluorescent DNA probe stained with intercalating dye, termed a nanotag, for sensitive detection of the low abundance miRNAs. Upon hybridization of both probes, an electrophoretic mobility shift is measured via interaction of the hydrocarbon modification of the  $\gamma$ PNAA with non-ionic surfactant micelles in a capillary electrophoresis running buffer, a technique known as micelle end-labeled free-solution electrophoresis (ELFSE). We demonstrate multiplexed detection of 6 *let-7* miRNAs in 4 minutes, excellent selectivity against G-T wobble base-pair single base mismatches between *let-7* miRNAs, and 100pM detection limits using our method.

In an effort to increase detection sensitivity closer to those required for trace miRNA concentrations (fM), we have investigated the use of isotachophoretic injections and longer nanotags. Although longer nanotags can accommodate a greater amount of fluorescent dye for enhanced signal, they are a challenge to separate by micelle ELFSE. The longer DNA lengths overcome the additional

friction of the transiently attached micelle and return to the inseparable free-solution limit. We find that *n*-alkyl polyoxyethylene ether surfactant ( $C_iE_j$ ) wormlike micelles yield high drag forces on the electrophoresing DNAs, appearing to be ideal for large nanotag separations. However, micelle-micelle interaction at moderate concentrations leads to a background sieving network that reduces ELFSE separation efficiency for long DNAs. We propose new constitutive models that accurately capture the behavior of long DNA lengths in these micelle networks. We demonstrate optimization of surfactant buffers to achieve separation of 1-10 kilobase DNA in 3 minutes and increase DNA sequencing length of reads by refining the degree of micelle-micelle interaction. Finally, we demonstrate micelle ELFSE separations on a microchip format for dramatically reduced separation times.

# Table of Contents

<b>Acknowledgements .....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>xiv</b>
<b>List of Figures.....</b>	<b>xvi</b>
<b>Chapter 1 – Motivation, Background, and Thesis Objectives .....</b>	<b>1</b>
1.1 – Motivation: MicroRNA Disease Biomarkers .....	1
1.1.1 – The Discovery of MicroRNA: From Trash to Treasure .....	1
1.1.2 – MicroRNA’s Role in Cancers and Diseases .....	2
1.1.3 – Challenges Facing miRNA Detection.....	3
1.1.4 – Gold Standards for miRNA Detection.....	5
1.1.5 – Alternative miRNA Detection Techniques.....	8
1.2 – Background.....	9
1.2.1 – Peptide Nucleic Acids (PNAs).....	9
1.2.2 – Ultrafluorescent, DNA Nanotags.....	11
1.2.3 – DNA Electrophoresis .....	14
1.2.4 – Capillary Gel Electrophoresis .....	15
1.2.5 – End-Labeled Free-Solution Electrophoresis .....	16
1.2.6 – ELFSE with Covalently Bound Drag-Tags .....	20
1.2.7 – ELFSE with Transient Micelle Interaction.....	22
1.3 – Thesis Objectives.....	26
1.4 – References.....	28
<b>Chapter 2 – High Affinity <math>\gamma</math>PNA Sandwich Hybridization Assay for Rapid MicroRNA Detection with Single Base Mismatch Discrimination.....</b>	<b>39</b>
2.1 – Introduction.....	39
2.2 – Materials and Methods.....	43
2.2.1 – Reagents .....	43

2.2.2 – $\gamma$ PNA Amphiphile ( $\gamma$ PNAA) Synthesis and Purification.....	43
2.2.3 – DNA Target Sandwich Detection with Cy5-labeled DNA Probe .....	44
2.2.4 – RNA Target Sandwich Detection with YOYO-1 DNA Probe .....	45
2.3 – Results and Discussion .....	47
2.3.1 – UV Melting Experiments on Full $\gamma$ PNAA-DNA Complements .....	47
2.3.2 – Design of Sandwich Complex .....	49
2.3.3 – Coaxial Stacking Stabilization.....	53
2.3.4 – Single Base Mismatch Discrimination.....	54
2.3.5 – Application Towards miRNA Detection .....	57
2.4 – Conclusion .....	62
2.5 – References.....	63
<b>Chapter 3 – Practical Improvements to MicroRNA <math>\gamma</math>PNA Sandwich Hybridization Assay: Multiplexing, PicoGreen Intercalating Dye, and Protein Contaminant Tolerance .....</b>	<b>67</b>
3.1 – Introduction.....	67
3.2 – Materials and Methods.....	70
3.2.1 – Reagents .....	70
3.2.2 – $\gamma$ PNA Amphiphile ( $\gamma$ PNAA) Synthesis and Purification.....	71
3.2.3 – Sandwich Hybridization Conditions .....	73
3.2.4 – Capillary Electrophoresis.....	73
3.2.5 – Goat Serum Spike-in Experiments.....	74
3.3 – Results and Discussion .....	74
3.3.1 – Multiplexed <i>let-7</i> miRNA Detection .....	74
3.3.2 – Theoretical Multiplexing Estimates.....	79
3.3.3 – Single Base Mismatch Discrimination.....	84

3.3.4 – Detection Limits.....	87
3.3.5 – Detection of miRNA in a Biological Matrix.....	88
3.4 – Conclusion .....	90
3.5 – References.....	90
<b>Chapter 4 – Coupled Transient Isotachopheresis Injections and Micelle ELFSE for Ultrasensitive DNA Detection .....</b>	<b>95</b>
4.1 – Introduction.....	95
4.2 – Materials and Methods.....	99
4.2.1 – Reagents .....	99
4.2.2 – DNA Alkylation Reaction.....	100
4.2.3 – PCR Amplification of End-Alkylated DNA.....	101
4.2.4 – Preparation of ITP Buffers.....	102
4.2.5 – Capillary Electrophoresis.....	102
4.2.6 – $\lambda$ DNA Dilutions .....	103
4.2.7 – tITP End-Alkylated DNA Injections and Separations .....	103
4.2.8 – tITP miRNA Sandwich Hybridization Detection .....	104
4.3 – ITP Theory.....	104
4.4 – Results and Discussion .....	106
4.4.1 – tITP $\lambda$ DNA Detection Limits.....	106
4.4.2 – tITP Injections with Micelle ELFSE.....	109
4.4.3 – tITP miRNA Sandwich Detection .....	114
4.5 – Conclusion .....	115
4.6 – References.....	116
<b>Chapter 5 – Unmodified DNA Separation in Entangled Wormlike Micelle Networks.....</b>	<b>119</b>
5.1 – Introduction.....	119
5.2 – Materials and Methods.....	121
5.2.1 – Reagents .....	121

5.2.2 – Lower Critical Solution Temperature (LCST) Determination .....	121
5.2.3 – Viscosity Measurements .....	122
5.2.4 – Capillary Electrophoresis .....	122
5.3 – Separation Theory .....	123
5.3.1 – Ogston Sieving .....	124
5.3.2 – Reptation .....	126
5.3.3 – Transient Entanglement Coupling / Post Collisions .....	130
5.4 – Results and Discussion .....	132
5.4.1 – Determination of $c^*$ .....	132
5.4.2 – Separation of DNA in C <sub>12</sub> E <sub>5</sub> Solutions .....	133
5.4.3 – Ogston Sieving Evaluation .....	134
5.4.4 – Reptation Evaluation .....	139
5.4.5 – TEC Comparison .....	140
5.5 – Conclusion .....	142
5.6 – References .....	143
<b>Chapter 6 - Free-solution Electrophoresis of End-alkylated DNA in Entangled Wormlike Micelle Networks.....</b>	<b>147</b>
6.1 – Introduction .....	147
6. 2 – Materials and Methods .....	150
6.2.1 – Reagents .....	150
6.2.2 – DNA Alkylation Reaction .....	151
6.2.3 – Lower Critical Solution Temperature (LCST) Determination .....	153
6.2.4 – PCR Amplification of End-alkylated DNA Ladder .....	153
6.2.5 – Capillary Electrophoresis .....	154
6.2.6 – Viscosity Measurements .....	155
6.3 – Results and Discussion .....	155

6.3.1 – Determination of $c^*$ .....	155
6.3.2 – Separation of End-alkylated DNA .....	156
6.3.3 – End-alkylated DNA ELFSE Parameters .....	157
6.3.4 – Resolution of End-alkylated DNA .....	161
6.3.5 – Modified ELFSE-Sieving Model .....	163
6.4 – Conclusion .....	171
6.5 – References .....	172
<b>Chapter 7 – Segregation Effects in Rapid Kilobase DNA Separations.....</b>	<b>175</b>
7.1 – Introduction .....	175
7.2 – Materials and Methods .....	177
7.2.1 – Reagents .....	177
7.2.2 – DNA Alkylation Reaction .....	178
7.2.3 – Lower Critical Solution Temperature (LCST) Determination .....	179
7.2.4 – PCR Amplification of C16B End-alkylated DNA Ladder .....	180
7.2.5 – Capillary Electrophoresis .....	180
7.2.6 – Viscosity Measurements .....	181
7.3 – Segregation Theory .....	181
7.3.1 – Segregated Electrophoretic Mobility .....	181
7.3.2 – Causes for Segregation .....	185
7.4 – Results and Discussion .....	188
7.4.1 – Determination of $c^*$ .....	188
7.4.2 – Long DNA Separations .....	189
7.4.3 – Effect of $E$ on Segregation .....	192
7.4.4 – Effect of $L$ on Segregation .....	193
7.4.5 – Rapid Long DNA Separations .....	195
7.5 – Conclusion .....	197

7.6 – References .....	198
<b>Chapter 8 – Improved DNA Sequencing by Reducing Micelle Entanglement.....</b>	<b>201</b>
8.1 – Introduction .....	201
8.2 – Materials and Methods.....	203
8.2.1 – Reagents .....	203
8.2.2 – DNA Alkylation Reaction.....	204
8.2.3 – Lower Critical Solution Temperature (LCST) Determination .....	205
8.2.4 – PCR Amplification of C16B End-alkylated DNA Ladder.....	206
8.2.5 – Sanger Sequencing C18 End-alkylated DNA Amplification .....	206
8.2.6 – Capillary Electrophoresis of C16B PCR Products.....	207
8.2.7 – Capillary Electrophoresis of C18 Sanger Sequencing Products.....	208
8.2.8 – Viscosity Measurements .....	209
8.3 – Results.....	209
8.3.1 – Determination of $c^*$ .....	209
8.3.2 – C16B End-alkylated DNA Separations .....	210
8.3.3 – Reduced Temperature DNA Sequencing.....	214
8.3.4 – C <sub>10</sub> E <sub>5</sub> End-Capping DNA Sequencing.....	216
8.4 – Conclusion .....	222
8.5 – References.....	223
<b>Chapter 9 – Rapid, High Resolution Microchip Micelle ELFSE DNA Separations .....</b>	<b>225</b>
9.1 – Introduction.....	225
9.2 – Materials and Methods.....	228
9.2.1 – Reagents .....	228
9.2.2 – DNA Alkylation Reaction.....	229

9.2.3 – PCR Amplification of End-alkylated DNA Ladder.....	231
9.2.4 – Microchip Electrophoresis Detection System.....	232
9.2.5 – Microchip Preparation .....	234
9.2.6 – Pull-back Voltage Imaging .....	235
9.2.7 – Microchip Electrophoretic Separation .....	235
9.3 – Results and Discussion .....	236
9.3.1 – Optimization of Pull Back Voltage.....	236
9.3.2 – Microchip Micelle ELFSE Separations .....	238
9.3.3 – Microchip Resolution Factors and Theory .....	242
9.3.4 – Theoretical Estimates for ME DNA Sequencing.....	246
9.4 – Conclusion .....	248
9.5 – References.....	249
<b>Chapter 10 – Conclusions, Future work, and Outlook.....</b>	<b>253</b>
10.1 – Conclusions.....	253
10.2 – Future Work.....	258
10.2.1 – Longer Nanotags for fM Detection Sensitivity.....	259
10.2.2 – Single Molecule Imaging of DNA in Entangled Micelle Networks.....	262
10.3 – Outlook .....	263
10.4 – References.....	264



## List of Tables

<b>Table 1.1</b> – Surfactant Exchange and Micelle Relaxation Time Scales.....	24
<b>Table 2.1</b> – $\gamma$ PNAA, DNA Probe, and Target Sequences and $T_m$ s.....	50
<b>Table 2.2</b> – Sequences of Single Base Mismatch Off-Targets.....	54
<b>Table 3.1</b> – List of miRNA Sequences, $\gamma$ PNAAs, and DNA Probes .....	72



# List of Figures

<b>Figure 1.1</b> – Illustration of miRNA gene regulation via mRNA recognition by RISC. Repression of LIN-14 protein synthesis by partial complementary binding of <i>lin-4</i> miRNA RISC to 7 sites within the 3' UTR of the <i>lin-14</i> gene. Adapted from [13] .....	3
<b>Figure 1.2</b> – Illustrations of the protocols for A) RT-PCR stem-loop primer miRNA amplification and detection by TaqMan probes adapted from [50], B) enzyme modification and fluorophore labeling of miRNA for microarray analysis, and C) a representative result from a scanned microarray where each spot corresponds to a different miRNA target, adapted from [68].....	7
<b>Figure 1.3</b> – Structure of the sugar-phosphate backbone of DNA/RNA, LNA, the uncharged peptide backbone of PNA, and the PEG $\gamma$ -modified <sup>MP</sup> $\gamma$ PNA. The letter B represents nucleic acid base placement on each backbone type. Adapted from [98].....	10
<b>Figure 1.4</b> – Chemical structure of bis-intercalating YOYO-1 and mono-intercalating PicoGreen.....	12
<b>Figure 1.5</b> – Modes of DNA electrophoresis. Each system represents the initial injection of a mixture of short (red) and long (green) DNA lengths. Upon imposing an electric field, we observe either no separation due to the length-invariant free-solution electrophoretic mobility in capillary zone electrophoresis (Top), elution of the short strands first by CGE (Middle), or elution of the long strands first by micelle ELFSE (Bottom), where aDNA refers to end-alkylated DNA.....	17
<b>Figure 1.6</b> – Illustration of blob theory for a DNA strand (red) end-attached to A) spherical drag-tag or B) polymeric drag-tag (black). The blob sizes for each segment are equivalent in hydrodynamic radii, as outlined by the dotted circles .....	18
<b>Figure 1.7</b> – Chemical structure of hydrophobic modifications to A) (PEG) $\gamma$ -carbon substituted peptide nucleic acid amphiphile ( $\gamma$ PNAA) and B) BODIPY C16 hydrophobically modified DNA that are used to encourage interaction of DNA with non-ionic surfactant micelles for ELFSE separations.....	23
<b>Figure 1.8</b> – Illustration of the fast monomer exchange $\tau_1$ , and the slower micelle formation and dissolution $\tau_2$ . Adapted from [162] .....	24

**Figure 1.9** – The entanglement of wormlike micelles as surfactant concentration increases from  $c < c^*$  to  $c > c^*$ , where  $c^*$  represents the overlap concentration of the surfactant system when micelles just begin to interact. Adapted from [169]..... 26

**Figure 2.1** – Structure of  $\gamma$ PNAA (C to N) where **B** = A, T, G, or C used in this study for sandwich hybridization of short nucleic acids and interaction with micelle drag tags .....42

**Figure 2.2** – UV melting profiles of  $\gamma$ PNAA bound to DNA. 5 $\mu$ M Strand concentrations in 1xTBE. Sequences:  $\gamma$ PNAA, C18 - MP<sub>3</sub> - Fluo – CCTACTACCTCA-Lys-NH<sub>2</sub>; DNA, 5' TGA GXT AGT AGG 3' **X** = A, C, G, T. Both heating and cooling profiles were performed and were identical, only heating curves are shown .....48

**Scheme 2.1** – Detection mechanism for reported sandwich hybridization assay .....49

**Figure 2.3** – Micelle CE detection for each component of the sandwich hybridization assay. Run conditions 20kV, 20/30cm capillary, 25°C, 24mM Triton X-100 running buffer in 1xTBE. A) 512nm fluorescence detection of the fluorescein-labeled  $\gamma$ PNAA. B) 675nm fluorescence detection of the Cy5-labeled DNA probe. Curves are labeled with hybridized components using abbreviations according to Table 2.1. Sandwich complex peaks are marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes. RFU = Relative Fluorescence Units .....51

**Figure 2.4** – A) 512nm fluorescence detection and B) 675nm fluorescence detection electropherograms of sandwich complexes under different binding configurations. Curves are labeled with hybridized components using abbreviations according to Table 2.1. Separation conditions the same as Figure 2.3. Sandwich complex peaks are marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes. ....52

**Figure 2.5** – Added stability from adjacent probe hybridization for sandwich complex formation due to coaxial stacking. Percentage of target bound in sandwich complex (light gray) or by only the  $\gamma$ PNAA (dark gray). Average of three measurements.....53

**Figure 2.6** – Percent of target bound by both probes at different capillary temperatures for SBM off-targets. Sequences of each target are seen in Table 2.2. A) 25°C, B) 30°C, C) 35°C, D) 40°C. Average of three measurements .....56

<b>Figure 2.7</b> – Detection of perfectly matched target from pool of single base mismatch off-targets. Run conditions 20kV, 40/50cm capillary, 50µm, 40°C, 24mM Triton X-100 running buffer in 1xTBE. Sandwich peak is marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.....	57
<b>Figure 2.8</b> – 512nm fluorescence detection of the 22nt RNA target (5' CGA UCA CAU GGC AUU GAG GUA G 3') by sandwich complex formation between the γPNAA and a 66 base-pair duplex DNA probe (sequences in experimental section). a) Detection of fluorescein-labeled γPNAA only, b) Detection of the fluorescein-labeled γPNAA and YOYO-1 stained DNA probe. Sandwich complex peaks are marked by an asterisk. Same run conditions as Figure 2.3. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes .....	59
<b>Figure 2.9</b> – Cartoon representation of γPNAA, 22nt RNA target, and 66 base-pair YOYO-1 stained DNA duplex sandwich complex .....	60
<b>Figure 2.10</b> – Observed fluorescent signal of sandwich complex at different target concentrations. Constant probe concentration of 1µM. Solid line represents linear fit with $R^2 = 0.999$ , averaged over three measurements .....	61
<b>Figure 3.1</b> – Cartoon representation of miRNA sandwich hybridization complex.....	70
<b>Figure 3.2</b> – Separation of 6 <i>let-7</i> miRNA sandwich complexes by CE with end-attached non-ionic surfactant micelles, γPNAA, and PicoGreen stained DNA nanotags .....	76
<b>Figure 3.3</b> – Integrated peak areas for the 6 <i>let-7</i> miRNAs complexes separated in Figure 3.2. DNA nanotag lengths of 20, 30, 40, 50, 60, and 70 bases corresponding to <i>let-7</i> miRNAs <i>e, f, c, d, a, and b</i> , respectively .....	77
<b>Figure 3.4</b> – Inverse sandwich complex length ( $1/L$ ) versus scaled elution time ( $t/t_0$ ) for peaks in Figure 3.2. According to Equation 3.1, the slope can be used to calculate $\alpha$ for the TX-100 ( $\alpha = 35$ ).....	78
<b>Figure 3.5</b> – Experimental baseline resolution factors from peaks in Figure 3.2 (black circles) vs. theoretical model Equation 3.8 (dotted line).....	82
<b>Figure 3.6</b> – Theoretical DNA nanotag lengths required to achieve baseline resolution between sequential miRNA complexes .....	83

**Figure 3.7** – Integrated peak areas for single base mismatch complexes for different run temperatures. A) *let-7c* or *let-7b* incubated with *let-7c* hybridization probes. B) *let-7f* or *let-7a* incubated with *let-7f* hybridization probes. 250nM miRNA in 2.5 $\mu$ M of respective probe. miRNA,  $\gamma$ PNAA, and nanotag sequences shown above each figure. Error bars represent standard deviation from  $n = 3$  measurements.....84

**Figure 3.8** – Representative electropherograms of competitive binding for single base mismatch target discrimination in solutions of both perfect match and mismatched hybridization probes. Separations performed at 25°C with 25nM of indicated miRNA and 250nM of *let-7a* and *let-7f* hybridization probes .....86

**Figure 3.9** – Integrated peak area for different concentrations of *let-7b* miRNA target in constant excess of sandwich hybridization probes (2.5 $\mu$ M probes). Solid line represents linear fit. Error bars represent  $n = 3$  measurements.....87

**Figure 3.10** – Detection of *let-7* miRNA complexes in the presence of protein contamination from dilutions of goat serum. Total protein concentration is indicated on each electropherogram trace. Electropherograms are arbitrarily shifted on the y-axis for display purposes.....88

**Figure 3.11** – Effect of total protein concentration in goat serum on the peak height of *let-7* miRNA sandwich complex peaks .....89

**Figure 4.1** – Illustration of tITP injection process for sensitive micelle ELFSE detection. The capillary is filled with LE buffer containing non-ionic surfactant micelles. A DNA sample prepared in TE buffer is then placed at the inlet of the capillary. A potential is applied to the system, driving electrophoretic motion of LE, TE, and DNA towards the detector. Since the micelles are uncharged, they remain stationary within the capillary. As the TE DNA injection enters the capillary, it becomes a micelle saturated zone. However, due to the ITP conditions within the capillary, the DNA will rapidly concentrate at the interface between the LE and TE buffers without separation. After a substantial injection of DNA in TE, the inlet buffer is replaced by LE buffer containing non-ionic surfactant micelles. A potential is applied and the components migrate towards the detector. The DNA will continue to concentrate at the LE-TE interface until the high mobility LE ions overtake the low mobility TE ions, disrupting the ITP interface. The capillary returns to a single component buffer system and DNA will now separate according to micelle ELFSE .....98

**Figure 4.2** – Elution peak height for titrations of  $\lambda$ DNA stained with YOYO-1 dye for single buffer system injections (black) or tITP injections (red). Error bars are standard deviation for  $n = 3$  .....107

**Figure 4.3** – Representative elution peak morphology for 1pM YOYO-1 stained  $\lambda$ DNA from A) single buffer system injections or B) tITP injections .....108

**Figure 4.4** – Representative electropherograms of C16B end-alkylated DNA stained with PicoGreen dye for either single buffer system injections or tITP injections. Injections were 1kV for 15sec. DNA lengths are indicated above each peak. Electropherograms are arbitrarily shifted on the y-axis for display purposes .....109

**Figure 4.5** – Effect of increasing tITP injection time on peak morphology of long C16B end-alkylated DNA stained with PicoGreen dye. Total injection time is represented as Volt-sec, indicated on each trace. DNA lengths are indicated above each peak. Electropherograms are arbitrarily shifted on the y-axis for display purposes .....110

**Figure 4.6** – Relative peak height for tITP injected DNA vs. single buffer injections .....111

**Figure 4.7** – Resolution factors, calculated using Equation 4.11, for long DNA lengths as a function of tITP injection time .....112

**Figure 4.8** – A) Peak spacing ( $dt/dL$ ) between DNA lengths as a function of tITP injection time. B) Peak broadness ( $FWHM$ ) of elution peaks as a function of tITP injection time .....113

**Figure 4.9** – tITP focusing of *let-7a* miRNA sandwich complex. 250nM *let-7a* miRNA incubated with 250nM of  $\gamma$ PNA and DNA nanotag in either LE (single buffer injection) or TE (tITP injection). miRNA sandwich peak marked with an asterisk .....114

**Figure 5.1** – Schematic representations of gel electrophoresis DNA separations. A) Ogston sieving, reptation, and biased reptation through network entanglements. B) TEC of DNA (green) colliding with an individual polymer (red), adapted from [18].....124

**Figure 5.2** – Specific viscosity of  $C_{12}E_5$  surfactant in 1xTBE buffer at 20°C. Dashed line represents linear fit to data below  $c^*$ . Solid line represents linear fit to data above  $c^*$ . Error bars represent standard deviation with  $n = 3$ . The intersection of these two linear fits marks the  $c^*$  .....132

**Figure 5.3** – Representative electropherograms showing the separation of the unmodified 2-Log DNA ladder at different  $C_{12}E_5$   $c/c^*$  ratios in 1xTBE. DNA lengths are indicated above peaks in kilobases.  $c/c^* =$  A) 0.67, B) 1.90, C) 3.33, D) 6.0 .....133

**Figure 5.4** – Ferguson plots for different lengths of DNA based on C<sub>12</sub>E<sub>5</sub> wt%. Lines represent linear regression fits to the data according to Equation 5.3 .....134

**Figure 5.5** – Plot of  $K_R^{1/2}$  vs.  $R_g$  of DNA for data taken at  $E = 33$  V/cm. Deviation from linear fit may indicate a transition from Ogston sieving to reptation .....136

**Figure 5.6** – Effective wormlike micelle network pore size as a function of C<sub>12</sub>E<sub>5</sub> wt% calculated via extrapolation of Ferguson plot data according to Equations 5.3. Red squares represent DNA  $R_g$  calculated by the WLC model Equation 5.16.....137

**Figure 5.7** – Calculated fractional volume available to electrophoresis vs. C<sub>12</sub>E<sub>5</sub> concentration .....138

**Figure 5.8** – Elution time for different DNA lengths and concentrations all at  $f = 0.75$  .....139

**Figure 5.9** – Reptation plots,  $\log \mu$  vs.  $\log (l/L)$ , at different concentration of C<sub>12</sub>E<sub>5</sub> for unmodified DNA. Straight line represent linear fits to lengths of DNA possibly undergoing reptation, where a slope of 1 would be predicted by  $\mu \propto l/L$ .....140

**Figure 5.10** – Predicted elution time by TEC theory (dotted line) and actual elution time (black circles) vs. DNA length for 2.5wt% C<sub>12</sub>E<sub>5</sub> separation buffer .....142

**Figure 6.1** – Effect of increasing micelle size ( $\alpha$ ) on elution time of miRNA complexes for different DNA nanotag lengths. Theoretical run conditions:  $l_d = 20$ cm,  $l_t = 30$ cm,  $V = 20$ kV,  $\mu_0 = 2.7 \times 10^{-4}$  cm<sup>2</sup> V<sup>-1</sup> s<sup>-1</sup>.....148

**Figure 6.2** – Chemical structure of fluorophore labeled, end-alkane modified DNA. BODIPY C16 fatty acid (C16B) is covalently linked to 5' C<sub>6</sub> amine modified DNA through an amide bond. The modified DNA is used as a primer for PCR amplification.....152

**Figure 6.3** – Specific viscosity of C<sub>12</sub>E<sub>5</sub> surfactant in 1xTBE buffer at 20°C. Dashed line represents linear fit to data below  $c^*$ . Solid line represents linear fit to data above  $c^*$ . Error bars represent standard deviation with  $n = 3$ . The intersection of these two linear fits marks the  $c^*$  .....156

**Figure 6.4** – Representative electropherograms showing the separation of C16B end-alkylated DNA ladders at different C<sub>12</sub>E<sub>5</sub>  $c/c^*$  ratios in 1xTBE.  $c/c^* =$  A) 1.0, B) 2.7, C) 4.4, D) 6.7. Numbers above peaks correspond DNA lengths (1) 2,840,

(2) 2,093, (3) 1165, (4) 986, (5) 909, (6) 842, (7) 808, (8) 787, (9) 701, (10) 647, (11) 603, (12) 551, (13) 488, (14) 421, (15) 375, (16) 265.....158

**Figure 6.5** –  $t/t_0$  vs  $l/L$  for the C16B DNA ladder separations at different C<sub>12</sub>E<sub>5</sub>  $c/c^*$  ratios. Linear fits demonstrate behavior according to ELFSE theory, Equation 6.1. Error bars represent standard deviation with  $n = 3$ . A) Full data set. B) Zoom in on long DNA length deviations .....159

**Figure 6.6** – ELFSE parameters from linear fits on Figure 6.5 according to Equation 6.2 at different C<sub>12</sub>E<sub>5</sub>  $c/c^*$  ratios. A) Effective micelle drag-tag size,  $\alpha$ . B) Measured  $\mu_0$ . Error bars represent standard deviation with  $n = 3$  .....160

**Figure 6.7** – A) Experimental (symbols) and theoretical (solid lines) resolution factors based on the FWHM measured for peaks in Figure 6.4. Theoretical resolution factors were found using  $dt/dL$  calculated using Equation 6.4 from the ELFSE parameters found in Figure 6.6. B) A comparison of actual  $dt/dL$  vs. the calculated  $dt/dL$  using Equation 6.3.....162

**Figure 6.8** –  $t/t_0$  scaled by fractional volume available to electrophoresis in modified ELFSE equation vs.  $l/L$ . Straight lines represent fits to Equation 6.9.167

**Figure 6.9** – C<sub>12</sub>E<sub>5</sub> micelle network pore size calculated by Equation 6.9 using a non-linear optimization algorithm to maximize the  $R^2$  value for the linear fits in Figure 6.8 .....168

**Figure 6.10** – Effective micelle drag term  $\alpha$  calculated using Equation 6.1, the ELFSE only model, or Equation 6.9, the modified ELFSE-sieving model .....169

**Figure 6.11** – A) Experimental (symbols) and theoretical (solid lines) resolution factors based on the FWHM measured for peaks in Figure 6.4. Theoretical resolution factors were found using  $dt/dL$  calculated using Equation 6.11 from the ELFSE parameters found in Figure 6.6. B) A comparison of actual  $dt/dL$  vs. the theoretical  $dt/dL$  using Equation 6.11 .....170

**Figure 7.1** – Schematic representation of DNA drag-tag conjugates under A) standard, unsegregated conformation, and B) segregated conformation. DNA represented as solid red line, drag-tag as black circle.....183

**Figure 7.2** – Relative mobility of unsegregated and segregated DNA-micelle complexes plotted against  $L/\alpha$  .....185

**Figure 7.3** – Critical segregation electric field strength ( $E_C$ ) calculated using Equation 7.12 for different  $\alpha$  values .....187

**Figure 7.4** – Specific viscosity of mixed C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> surfactant buffers in 1xTBE at 25°C. Solid line represents linear fit to data below  $c^*$ . Dotted line represents linear fit to data above  $c^*$ . Intersection of the fits indicates  $c^*$ . .....189

**Figure 7.5** – Representative electropherograms of 1-6kB C16B end-alkylated DNA separated in C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> mixed surfactant 1xTBE buffers at 25°C. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.  $V = 5\text{kV}$  .....190

**Figure 7.6** – A)  $t$  vs  $l/L$  for the C16B end-alkylated DNA lengths separated in Figure 7.5. Linear fits demonstrate behavior according to standard, unsegregated ELFSE theory, Equation 7.14. B)  $t$  vs  $L^{-0.5}$  for the C16B end-alkylated DNA lengths separated in Figure 7.5. Linear fits demonstrate behavior according to segregated ELFSE theory, Equation 7.15 .....191

**Figure 7.7** – Effective drag-tag parameters for unsegregated and segregated ELFSE models vs.  $c/c^*$  of ternary C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> surfactant .....192

**Figure 7.8** – Scaled elution time vs. A)  $l/L$  (unsegregated model) and B)  $1/\text{sqrt}(L)$  (segregated model) for  $c/c^* = 2.33$  surfactant buffer at different E .....193

**Figure 7.9** – Comparison of A) unsegregated and B) segregated ELFSE models for separation of 86 – 1165 C16B end-alkylated DNA in ternary C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> surfactant buffers at  $c/c^* = 2.33$ .....194

**Figure 7.10** – Rapid separation of kB C16B DNA lengths using mixed surfactant system. Run Conditions: 12mM C<sub>16</sub>E<sub>6</sub>, 8mM C<sub>12</sub>E<sub>5</sub>, 1mM C<sub>10</sub>E<sub>5</sub> ( $c/c^* = 2.33$ ) in 1xTBE, 25°C, 25kV,  $l_d = 20\text{cm}$ ,  $l_t = 30\text{cm}$ ,  $d = 50\mu\text{m}$ , 4kV 30sec injection from formamide .....195

**Figure 7.11** – Resolution factors for peaks in Figure 7.10, with predictions from unsegregated  $dt/dL$  in Equation 7.17 (red squares, red line) and segregated  $dt/dL$  in Equation 7.18 (blue diamonds, blue line) .....197

**Figure 8.1** – Specific viscosity of C<sub>12</sub>E<sub>5</sub> surfactant in 7M urea 1xTBE at 40°C. Solid line represents linear fit to data below  $c^*$ . Dotted line represents linear fit to data above  $c^*$ . Intersection of the fits indicates  $c^*$  .....210

**Figure 8.2** – Representative electropherograms of C16B end-alkylated DNA separated in C<sub>12</sub>E<sub>5</sub> 7M urea 1xTBE buffers at 40°C. Numbers above peaks correspond to lengths (1) 1165, (2) 986, (3) 909, (4) 842, (5) 808, (6) 787, (7) 701, (8) 647, (9) 603, (10) 551, (11) 488, (12) 421, (13) 375. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes .....211

**Figure 8.3** – A) Elution time ( $t$ ) vs  $l/L$  for the C16B end-alkylated DNA lengths separated in C<sub>12</sub>E<sub>5</sub> 7M urea buffers in Figure 8.2. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Effective micelle drag-tag size,  $\alpha$ , vs. C<sub>12</sub>E<sub>5</sub> surfactant concentration, calculated from the slope and intercept of Figure 8.3A, according to Equation 8.1 .....212

**Figure 8.4** – Resolution factors calculated for different C16B end-alkylated DNA lengths using Equation 8.2 from peaks in Figure 8.2 .....213

**Figure 8.5** – Relative viscosity of 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea (black circles) and 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> (red squares). A higher relative viscosity represents a greater extent of micelle interaction. Error bars represent standard deviation for  $n = 3$  .....214

**Figure 8.6** – A) Elution time ( $t$ ) vs  $l/L$  for T channel BigDye<sup>®</sup>C18 end-alkylated DNA sequencing products separated in 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea 1xTBE buffer on ABI310 at different temperatures. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Resolution factors calculated for or T channel BigDye<sup>®</sup>C18 end-alkylated DNA sequencing products using Equation 8.2. Dotted line represents resolution limit for DNA sequencing, corresponding to the  $LOR$ .....215

**Figure 8.7** – A) Elution time ( $t$ ) vs  $l/L$  for T channel BigDye<sup>®</sup>C18 end-alkylated DNA sequencing products separated in 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE buffers on ABI310 at different temperatures. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Resolution factors calculated for or T channel BigDye<sup>®</sup> C18 end-alkylated DNA sequencing products using Equation 8.2. Dotted line represents resolution limit for DNA sequencing, corresponding to the  $LOR$ .....217

**Figure 8.8** – Four channel DNA sequencing of C18 end-alkylated BigDye<sup>®</sup> sequencing products in 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE. A = green curve, C = blue curve, G = black curve, T = red curve. Numbers above peaks correspond to T channel DNA lengths. Elution times between different channels have been adjusted using Equation 8.4.....218

**Figure 8.9** – Elution time ( $t$ ) vs  $l/L$  for all four channels in Figure 8.8. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1.....220

**Figure 9.1** – Illustration of finish-line or snap-shot detection modes in microchip ELFSE. All DNA lengths (blue lines) are separated from one another by the time the longest DNA reaches the detector. The shorter lengths have barely migrated down the channel by the time the longest lengths reach the finish-line detector

(top cartoon). This creates a situation where most of the run time is spent waiting for the short DNA to pass the detector. In snap-shot mode (bottom cartoon), the entire channel is imaged once the longest DNA lengths are resolved. The separation of all DNA lengths is captured even though the short lengths are still far from the end of the channel.....227

**Figure 9.2** – Schematic of NS114 microfluidic chip from Perkin Elmer with dimensions of each channel .....231

**Figure 9.3** – Schematic of microchip electrophoresis laser-induced fluorescence detection system, described in Section 9.2.4 .....234

**Figure 9.4** – Effect of pull-back voltage applied to injection channels to prevent leakage of into separation channel. Images are taken at the channel intersection with the separation channel down the page, right after the separation voltage is applied. Pull-back voltages represent potential applied to channel 2 and 3 wells. Channel 4 was grounded ( $V = 0$ ) and channel 3 has a potential of 2,220V .....238

**Figure 9.5** – Representative electropherograms for microchip micelle ELFSE separations of C16B end-alkylated DNA ladders. A) 24mM TX-100 in 1xTBE. B) 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M urea 1xTBE. Run conditions as described in Section 9.2.7.....239

**Figure 9.6** – Elution time ( $t$ ) vs.  $1/L$  from Figure 9.5. C16B end-alkylated DNA ladders separated by ME in 24mM TX-100 (black circles) or 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M Urea (red squares). Solid lines are linear fits according to Equation 9.1.....241

**Figure 9.7** – Resolution factors for C16B end-alkylated DNA ladders separated by ME in 24mM TX-100 (black circles) or 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M urea (red squares). Solid lines represent fits to theoretical resolution model for ELFSE in Equation 9.7 .....245

**Figure 9.8** – Theoretical estimates of microchip channel length and run-time to detect 450 bases with single base resolution and an  $\alpha = 500$  micelle buffer according to Equation 9.8. Diffusion only model (black line) and diffusion + wall adsorption with  $W = 18\mu\text{s}$  (red line) .....248

**Figure 10.1** – The CuAAC click chemistry reaction between an alkyne modified PCR amplicon and azide modified miRNA specific oligomer. Under simple reaction conditions, the long DNA probes with miRNA specific cohesive ends can be produced.....261

# Chapter 1 – Motivation, Background, and Thesis Objectives

## 1.1 – Motivation: MicroRNA Disease Biomarkers

### 1.1.1 – The Discovery of MicroRNA: From Trash to Treasure

In the early 1980s, the *lin-4* gene in *C. elegans* was found to control the timing of larval development.<sup>[1]</sup> It wasn't until 10 years later that a surprising discovery was made; the *lin-4* gene did not encode for a protein, but rather a pair of small RNAs.<sup>[2], [3]</sup> These approximately 22 nucleotide (nt) single stranded RNAs were found to have partial antisense complementarity to the 3' untranslated region (UTR) of the *lin-14* gene in *C. elegans*.<sup>[2]</sup> The 22nt RNAs down regulated the expression of the LIN-14 protein, acting as post-transcriptional gene regulators.

For decades, many researchers discarded the short, non-protein coding RNAs thinking they were simply superfluous transcripts. Soon after the work of Lee, Feinbaum and Ambros was published on *lin-4* in 1993, the short RNAs were more closely studied.<sup>[2]</sup> It wasn't until the early 2000s that a second, non-protein coding *let-7* gene in *C.elegans* was discovered that solidified the evidence for short RNA gene regulation.<sup>[4], [5]</sup> Furthermore, the *let-7* 21nt RNA was found to be conserved in sequence and function across a range of animal species, such as mollusks, annelids, arthropods, and vertebrates, including humans.<sup>[6]</sup> Soon, an entirely new

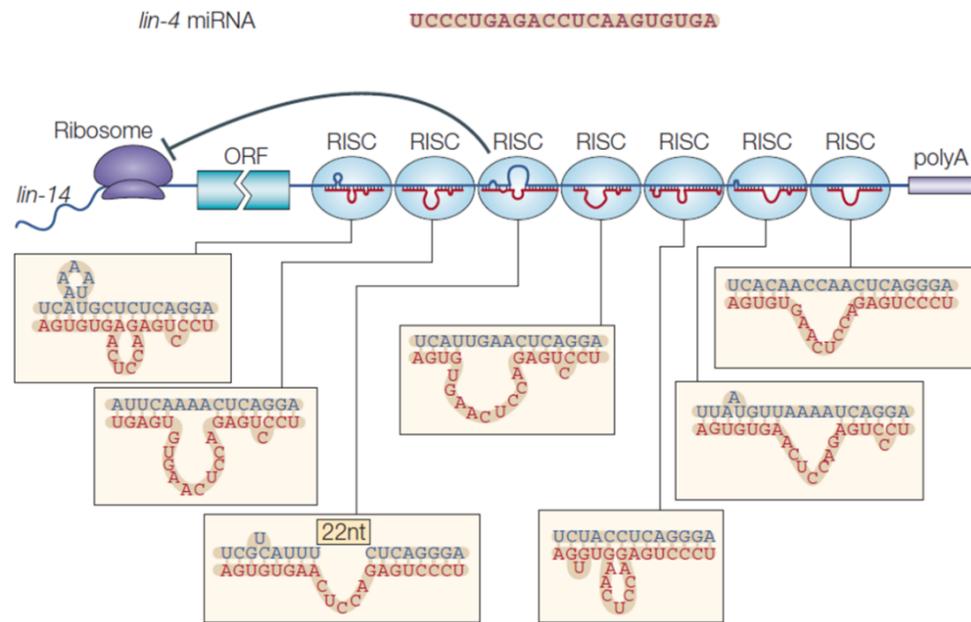
class of small gene regulating RNAs was established, now known as microRNAs (miRNAs).<sup>[7]-[9]</sup>

### **1.1.2 – MicroRNA’s Role in Cancers and Diseases**

MiRNA play critical roles in regulation of cellular processes, such as cell proliferation, differentiation, and apoptosis.<sup>[10]-[13]</sup> They bind to the 3’UTR of messenger RNA (mRNA) to induce mRNA cleavage or repression of ribosome protein translation. MiRNA biogenesis is characterized by sequential cleavages of the large miRNA gene transcripts, and nuclear transport to the cytoplasm to form an 18-25nt double stranded RNA.<sup>[13], [14]</sup> These short RNA duplexes are recognized by many different proteins to form a ribonucleoprotein RNA induced silencing complex (RISC). One strand of the RNA duplex is degraded, allowing the miRNA RISC to recognize mRNA binding sites with partial complementarity for gene regulation. Figure 1.1 shows the 7 different recognition sites of *lin-4* miRNA RISC on the 3’UTR of the *lin-14* mRNA. Bioinformatic approaches have predicted that 30% of our genes are regulated by miRNA through this partial complementary.<sup>[15]-[17]</sup>

With such a large role in gene regulation, it was soon discovered that deviations from normal miRNA expression altered normal cell growth and development, leading to a variety of cancers and diseases.<sup>[18]-[24]</sup> For example, the deregulation of miRNA-125b, 145, 21, and 155 have been directly linked to breast cancer tumor formation.<sup>[23], [25]-[27]</sup> The overexpression of miRNA-10b in non-metastatic breast tumors has also been associated with increased cell migration and

metastasis of cancers cells.<sup>[28]</sup> Over 50% of miRNA genes are located in cancer-associated genomic regions or fragile sites, indicating they may have a crucial role in cancer progression.<sup>[29]</sup> Furthermore, the existence of extracellular miRNAs in body fluids such as saliva, blood serum/plasma, and urine has increased the screening capabilities for miRNA expression level analysis.<sup>[21], [30], [31]</sup> Thus, the differential miRNA expression between normal and diseased tissues offers tremendous potential as an early stage cancer and disease diagnostic.



**Figure 1.1** – Illustration of miRNA gene regulation via mRNA recognition by RISC. Repression of LIN-14 protein synthesis by partial complementary binding of *lin-4* miRNA RISC to 7 sites within the 3' UTR of the *lin-14* gene. Adapted from [13]

### 1.1.3 – Challenges Facing miRNA Detection

MiRNA detection is rather challenging because of their short length, high sequence homology between miRNA family members, low natural abundance, and high number of different miRNAs that require detection. Their short length makes detection by conventional nucleic acid hybridization methods difficult, as

hybridization probes are typically the same length as miRNA. Many detection assays employ a sandwich hybridization approach, where target sequences are bound by two separate probes.<sup>[32]–[41]</sup> Normally one probe provides target segregation from the bulk solution (the capture probe), while the second probe imparts a measurable signal to the hybridization event (the reporter probe). In the case of miRNA, binding of probes to either half of the target is too weak. As such, miRNA hybridization detection is normally carried out using enzymatic methods to effectively lengthen the miRNA sequence for stable binding.<sup>[42]–[48]</sup> These additional methods are laborious, time-consuming, and expensive compared with simpler direct, one-step miRNA detection.

MiRNA sequences are also closely related, differing by a single nucleotide in many cases. Many assays have poor discrimination of these single base mismatches, leading to false positive results when non-specific hybridization probes are used. The stability of miRNA also varies dramatically due to the large differences in GC content for different miRNAs. The melting transition temperature ( $T_m$ ), defined as the temperature required for a duplex to exist as half single stranded and half double stranded, may range from 45°C to 74°C for miRNA hybridization probes.<sup>[49]</sup> The increased temperatures required to achieve selectivity against single base mismatches for one set of miRNA may lead to reduced hybridization stability for another due to these large stability differences.

MiRNA copy numbers are relatively low, as low as tens of copies per cell.<sup>[50]</sup> Many detection methods rely on reverse transcriptase polymerase chain reaction

(RT-PCR) to amplify the miRNA to detectable concentrations of complementary DNA.<sup>[50]–[52]</sup> Although RT-PCR has high sensitivity, the short length of miRNAs requires the use of complicated and expensive stem-loop primers to achieve PCR amplification.<sup>[53]</sup>

Finally, there are 2,588 mature human miRNAs listed in the miRNA database miRBase 21 (<http://www.mirbase.org/>, July 2014) and thousands of reported miRNA-disease associations.<sup>[53]–[56]</sup> Additionally, identification of tissue and condition specific diseases is typically associated with a unique miRNA signature composed of 2-20 miRNAs.<sup>[22], [57]–[59]</sup> Therefore, the ability to detect multiple miRNAs in a single, high-throughput assay is crucial towards implementing miRNA as potent biomarkers.

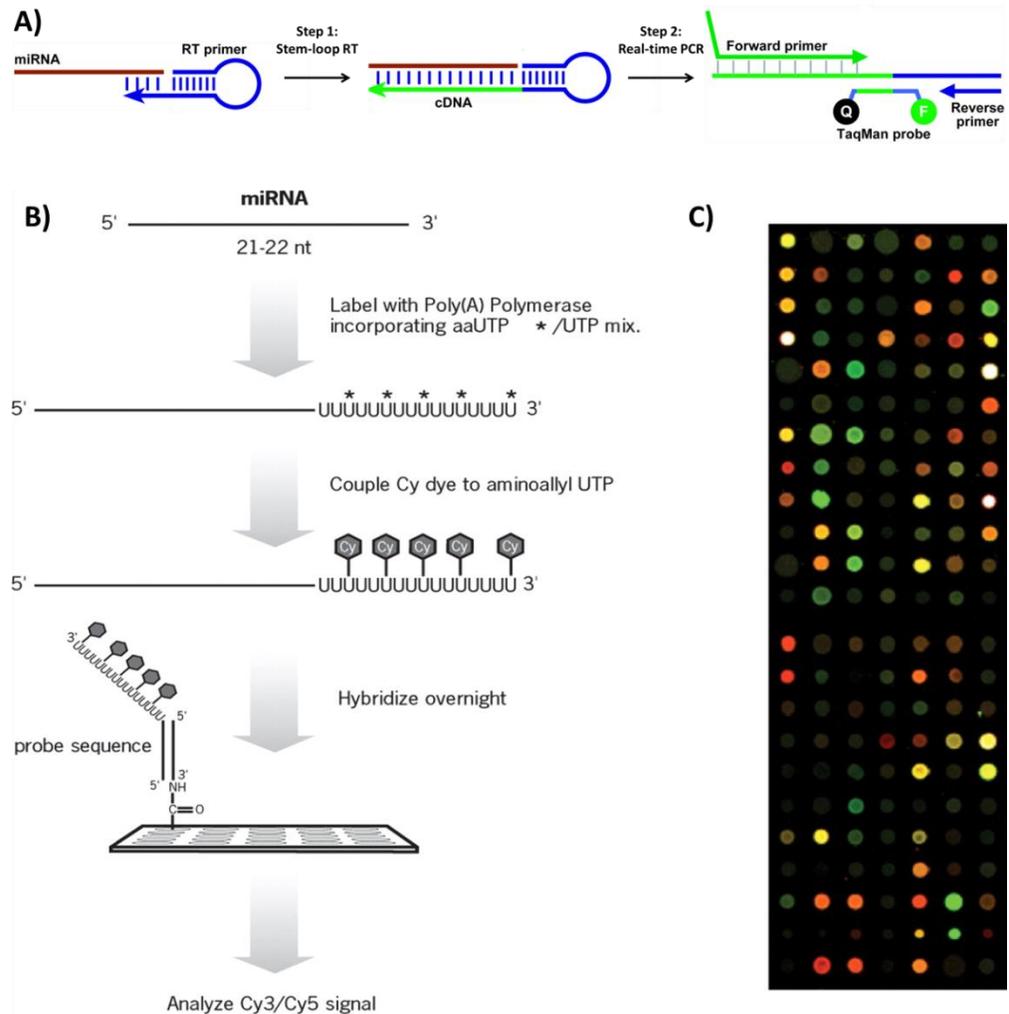
### **1.1.4 – Gold Standards for miRNA Detection**

The three main techniques used to detect and quantify miRNA expression levels are northern blotting<sup>[60], [61]</sup>, RT-PCR<sup>[50], [62], [63]</sup>, and oligonucleotide microarrays<sup>[64], [65]</sup>. The hybridization of radioactive probes to miRNA in northern blotting enables sensitive detection limits. Single miRNA expression levels can be compared between wild-type and mutant cell lines, as well as across different tissues. However, northern blotting yields poor quantification of miRNA concentrations and is quite time-consuming and laborious. As the number of known miRNAs has expanded, northern blotting became unrealistic for large scale detection of hundreds of miRNA.

RT-PCR based techniques offer the advantages of high sensitivity, low sample requirement, sequence specificity, and a dynamic detection range over 7 logs, all which enables extremely reliable quantification of miRNA expression levels.<sup>[50], [66]</sup> RT-PCR has even been used to quantify miRNA expression from a single cell (220 different miRNAs).<sup>[67]</sup> However, each miRNA requires an individual set of custom stem-loop primers and TaqMan probes. These primers and probes are extremely expensive, reducing their applicability towards full miRNA profiling. An example of the RT-PCR process is seen in Figure 1.2A.

Solid-phase microarrays have become the gold standard for multiplexed miRNA detection.<sup>[64], [65], [68]</sup> These arrays are silica substrates patterned with small spots of surface-bound miRNA complementary probes. Each spot is  $\sim 100\mu\text{m}^2$ , enabling the high-throughput quantification of thousands of miRNA on a single  $\text{cm}^2$  chip. However, the kinetics of miRNA target hybridization to a surface-bound probe is dramatically slower, and more complex, than hybridization in solution.<sup>[69], [70]</sup> A high probe density is required to achieve sensitive detection limits on the chip. Too high of a coverage reduces miRNA surface hybridization, caused by the increased electrostatic repulsion between the negatively charged strands.<sup>[71]</sup> Additionally, the  $T_m$  of surface based hybridization is a function of probe density, leading to reduced selectivity on the surface.<sup>[72], [73]</sup> The reduced kinetics of surface based hybridization leads to 18 hour target incubations that make analysis times extremely long. MiRNAs also require enzymatic modification to enable hybridization to surface-based capture probes, as well as labeling with fluorescent reporter probes. Additionally, high chip costs ( $\sim \$250/\text{array}$ ) and highly specific

microarray detection capital equipment costs (~\$150,000) reduce widespread clinical adaptation of microarrays. A typical miRNA microarray procedure is seen in Figures 1.2B and C.



**Figure 1.2** – Illustrations of the protocols for A) RT-PCR stem-loop primer miRNA amplification and detection by TaqMan probes adapted from [50], B) enzyme modification and fluorophore labeling of miRNA for microarray analysis, and C) a representative result from a scanned microarray where each spot corresponds to a different miRNA target, adapted from [68].

### 1.1.5 – Alternative miRNA Detection Techniques

Northern blotting, RT-PCR, and microarrays have dramatically increased our understanding of miRNA-disease relations over the past decade. However, their application towards miRNA diagnostics has been limited due to their high costs, extensive sample preparation (miRNA isolation and enzymatic amplification/labeling), and the trade-off between high sensitivity (RT-PCR) and multiplexed detection (microarray).

To fill the niche for viable miRNA diagnostics, a new class of direct (enzyme-free) miRNA detection assays has emerged. These direct techniques cut down on assay time and cost by removing any complex enzymatic steps for detection, while also eliminating potential enzyme sequence-specific biases that reduce reliability.<sup>[74]</sup> These techniques include bioluminescence<sup>[75]</sup>, atomic force microscopy (AFM)<sup>[76]</sup>, surface plasmon resonance<sup>[77]</sup>, electro-chemical biosensors<sup>[78], [79]</sup>, electro-mechanical devices<sup>[80]</sup>, molecular beacons<sup>[81], [82]</sup>, isotachopheresis<sup>[83], [84]</sup>, two-color fluorescence coincident detection<sup>[85]</sup>, and capillary electrophoresis (CE)<sup>[86]–[90]</sup>. Many of these methods demonstrate exceptional detection sensitivity in the aM-fM range, but are only capable of detecting one miRNA at a time. Only the CE methods have demonstrated promise for multiplexed miRNA detection, albeit with less than desirable nM-pM sensitivity. Additionally, many clinical labs already possess CE instruments, which should make adaptation of new miRNA detection methods by CE much more amenable for new users. New CE methods that are capable of fM detection

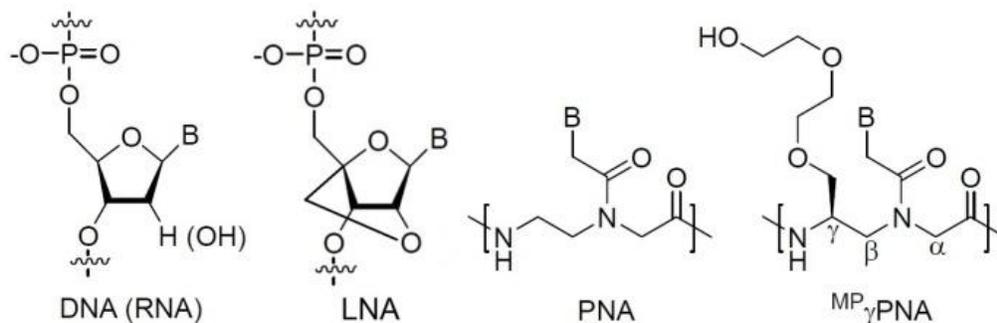
limits coupled with high-throughput, multiplexed detection stand to make a dramatic impact on the application of miRNA as disease biomarkers.

## 1.2 – Background

### 1.2.1 – Peptide Nucleic Acids (PNAs)

In order to improve miRNA hybridization, nucleic acid analogs and mimics are commonly used that exhibit increased binding stability. The two most common synthetics are locked nucleic acids (LNAs) and peptide nucleic acids (PNAs). Backbone structures for DNA/RNA, LNA, PNA, and the modified PNA used in this work are seen in Figure 1.3.

LNAs have nucleic acid backbones whose ribose ring is locked by a methylene bridge across the 2'-O and the 4'-C. This modification forces the single stranded LNA oligomers to take on an A-form like conformation of RNA duplexes, increasing LNA-RNA duplex stability by reducing the entropic costs of hybridization.<sup>[91], [92]</sup> LNA modified oligomers have been used to improve miRNA detection efficiency of northern blotting, solid-phase microarrays, *in situ* hybridization, and single molecule fluorescence coincidence detection.<sup>[61], [85], [93]–</sup><sup>[95]</sup> However, design of LNA probes is complicated by a propensity for LNA oligomers to form hairpins and LNA-LNA duplexes, as well as expensive LNA oligomer synthesis.<sup>[91], [96], [97]</sup>



**Figure 1.3** – Structure of the sugar-phosphate backbone of DNA/RNA, LNA, the uncharged peptide backbone of PNA, and the PEG  $\gamma$ -modified  $MP_{\gamma}$ PNA. The letter B represents nucleic acid base placement on each backbone type. Adapted from [98]

PNAs are synthetic DNAs where the negatively charged sugar-phosphate backbone has been replaced by an uncharged, peptide-like *N*-(2-aminoethyl) glycine backbone.<sup>[99]</sup> PNAs bind DNA and RNA with remarkable affinity and selectivity due to the reduced electrostatic repulsion between the two strands.<sup>[100]</sup> <sup>[101]</sup> They were originally recognized for their excellent strand-invasion capabilities, forming triplex DNA-PNA-DNA structures.<sup>[99], [102]–[104]</sup> More recently, their high binding affinity has found application to direct DNA/RNA hybridization assays.<sup>[78], [79], [105]–[108]</sup>

One major issue for PNAs is that since they are uncharged, they have poor water solubility and a tendency to aggregate.<sup>[109]–[111]</sup> Charged residues are typically incorporated on the PNA terminus to increase water solubility.<sup>[111]–[115]</sup> However, these charged residues may reduce the PNA's high affinity or selectivity by introducing small amounts of charge to the system. Our group has demonstrated that model predictions of  $T_m$  for PNAs modified with negatively charged glutamate groups are higher than experimentally observed (NH<sub>2</sub>-AGTGATCTAC-(Glu)<sub>4</sub>,  $T_{m,obs} = 38.9^{\circ}\text{C}$ ,  $T_{m,theo} = 53.7^{\circ}\text{C}$ ).<sup>[111]</sup> Conversely,  $T_m$

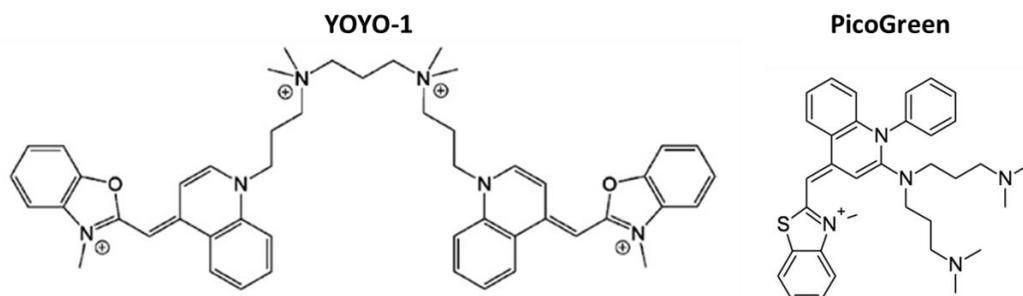
predictions were lower than experimentally observed for PNAs modified with positive lysine groups (NH<sub>2</sub>-TTTCCG-(Lys)<sub>2</sub>,  $T_{m,obs} = 27.4^{\circ}\text{C}$ ,  $T_{m,theo} = 24.3^{\circ}\text{C}$ ).<sup>[111]</sup> The non-specific electrostatic interaction of the lysine groups with the negatively charged DNA/RNA backbone may lead to reduced discrimination of single base mismatches.

Sahu *et al.* have recently reported a polyethylene glycol (PEG)  $\gamma$ -carbon substituted peptide nucleic acid ( $\gamma$ PNA) that exhibits excellent water solubility.<sup>[116]</sup> Not only do the PEG modifications increase water solubility, but the  $\gamma$ -carbon modification preorganizes the  $\gamma$ PNA oligomer into a B-form like conformation of a DNA duplex.<sup>[98], [116], [117]</sup> This organization leads to similarly reduced entropic costs for hybridization like the LNA probes. By combining the benefits of an uncharged oligomer with a preorganized conformation, the  $\gamma$ PNA offers tremendous improvements in affinity compared to unmodified PNA and LNA probes. The  $\gamma$ PNA should enable stable hybridization with as little as 8 bases of a miRNA target, allowing stable hybridization of a second probe for sandwich hybridization detection.

### **1.2.2 – Ultrafluorescent, DNA Nanotags**

The low expression level of miRNA inhibits the use of conventional fluorescence detection. The most straightforward way to increase fluorescent signal is to increase the number of fluorophores available for detection. Many methods increase detection sensitivity by utilizing DNA probes covalently labeled with multiple fluorophore systems, typically streptavidin phycoerythrin.<sup>[68], [118], [119]</sup> These assemblies utilize excitation light more efficiently than single

fluorophores, as the extinction coefficient of the assembly is assumed to be the sum of the individual molecules ( $\epsilon_{\text{phycoerythrin}} = \sim 10^6 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>[120]</sup> However, at high degrees of fluorophore labeling, fluorescence intensity can decrease due to self-quenching. Thus, in order to preserve fluorescence performance, the multifluorophore assembly must maintain spatial segregation of individual fluorophores.



**Figure 1.4** – Chemical structure of bis-intercalating YOYO-1 and mono- intercalating PicoGreen.

The double-helix of a DNA duplex creates an ideal scaffold for fluorophore segregation. Individual molecules can be segregated between base pairs. Cyanine intercalating dyes are widely used for their ability to form fluorescent complexes with nucleic acids.<sup>[121], [122]</sup> The structure of the commonly implemented YOYO-1 (a  $+4$  charge state bis-unsymmetrical cyanine dye) and PicoGreen (a  $+1$  charge state mono-unsymmetrical cyanine dye) are shown in Figure 1.4. Although it appears that the dimethylamino groups on the PicoGreen dye may become protonated, leading to a  $+3$  charge state, the salt dependence interaction between PicoGreen and DNA estimates only a single electrostatic constant.<sup>[123]</sup> Thus, PicoGreen is thought to maintain a  $+1$  charge state in solution.

These dyes insert themselves non-covalently between adjacent base pairs of DNA through multiple mechanisms.<sup>[121], [123]–[125]</sup> The aromatic rings of the dye participate in  $\pi$ -stacking with the nucleotide base pairs. The cationic charges interact with the negatively charged phosphates on the DNA backbone. In the case of PicoGreen, the dimethylamino branches act as DNA minor groove binders, greatly increasing the affinity of the mono-intercalator with the DNA backbone.<sup>[123]</sup> Upon intercalation, both dyes undergo >1,000 fold increases in fluorescent intensity.<sup>[124], [126]</sup> The dyes will self-quench in solution, but  $\pi$ -stacking with DNA base pairs results in fluorescent conformations of the intercalating dyes. They also have extremely high binding constants, on the order of  $10^{-10}$  M in gels and  $10^{-8}$  M in solution, enabling DNA staining in a few minutes with minimal dissociation.<sup>[123], [127], [128]</sup>

Due to steric constraints upon intercalation, a single base pair cannot accommodate more than a single dye molecule. From this nearest neighbor exclusion, YOYO-1 and PicoGreen dyes associate roughly 1 dye molecule per 4 base pairs.<sup>[124], [126]</sup> Since the dimethylamino arms on PicoGreen interact with the minor groove of the DNA duplex, the mono-intercalating PicoGreen dye occupies a similar number of base pairs as the bis-intercalating YOYO-1 dye.<sup>[123]</sup> We term these intercalating dye stained DNA strands “DNA nanotags” and view their high fluorescence intensity as ultra-bright probes for direct miRNA detection.

To date, only one miRNA detection mechanism has utilized intercalating dyes for high sensitivity detection.<sup>[129]</sup> In this method, single miRNA molecules were detected by total internal reflection microscopy after hybridization to

complementary DNA probes and YOYO-1 staining. Since YOYO-1 has a much higher propensity for staining the duplex over single stranded oligomers and RNAs, the detection of miRNA was possible in a total RNA system.<sup>[129]</sup> However, multiplexing detection is difficult as different miRNA-DNA duplexes would both yield the same fluorescent signal. Other methods have not implemented intercalating dyes for miRNA detection due to the non-specific labeling of all double stranded nucleic acids. Separation of bound miRNAs from one another, and from unbound probes, is challenging. Our proposed electrophoretic miRNA detection mechanism yields such a separation.

### 1.2.3 – DNA Electrophoresis

DNA is a highly-charged polyelectrolyte that will migrate upon placement in an electric field. It will move with a velocity ( $v$ ) set by the electric field strength ( $E$ ) and its electrophoretic mobility ( $\mu$ ), generally defined as the ratio of its total charge ( $Q$ ) to friction factor ( $f$ ).

$$v = \mu E = \frac{QE}{f} \quad (1.1)$$

However, DNA longer than its persistence length (~17 bases for single and ~120 for double stranded DNA) has a length invariant free-solution electrophoretic mobility ( $\mu_0$ ), making electrophoretic separation difficult.<sup>[130], [131]</sup> This scenario is caused by counterions present in the running buffer freely penetrating the hydrodynamic sphere of the DNA coil. The counterions hydrodynamically interact with DNA on the scale of a Debye length ( $\kappa^{-1}$ ),

screening hydrodynamic interactions between DNA segments. In the ionic strength buffers used for electrophoresis,  $\kappa^{-1}$  is roughly the length of a DNA monomer. Thus, the DNA behaves like a free-draining polymer where electrical and frictional forces are balanced on the monomer level. DNA cannot be separated by length in free-solution electrophoresis due to this length independent  $\mu_0$ .

### 1.2.4 – Capillary Gel Electrophoresis

Length dependent electrophoretic separation of DNA is typically carried out using polymer gels, such as linear polyacrylamide<sup>[132]</sup>, PEO<sup>[133], [134]</sup>, pDMA<sup>[135]</sup>, and cellulose based polymers<sup>[136], [137]</sup>. In capillary gel electrophoresis (CGE), the polymer fibers create an entangled network that the DNA must find its way through. The network adds frictional forces to the migrating DNA in a length dependent manner, where shorter DNAs can move through the network more easily than long ones. Small DNAs migrate through the pore spaces of the network unperturbed, following an Ogston sieving mechanism.<sup>[138], [139]</sup> Longer fragments have a more difficult time moving through the network, as they are too large to fit through the pores unperturbed. They will reptate through the network head first like a snake, termed biased reptation.<sup>[140], [141]</sup> As the DNA length or electric field strength increases, the DNA will orient itself in the direction of the electric field and separation is no longer possible. This is known as the biased reptation with orientation limit of CGE. It reduces the applicability of CGE to DNA lengths < 20kB, as well as low  $E$ , which greatly extends the time required to complete the separation.

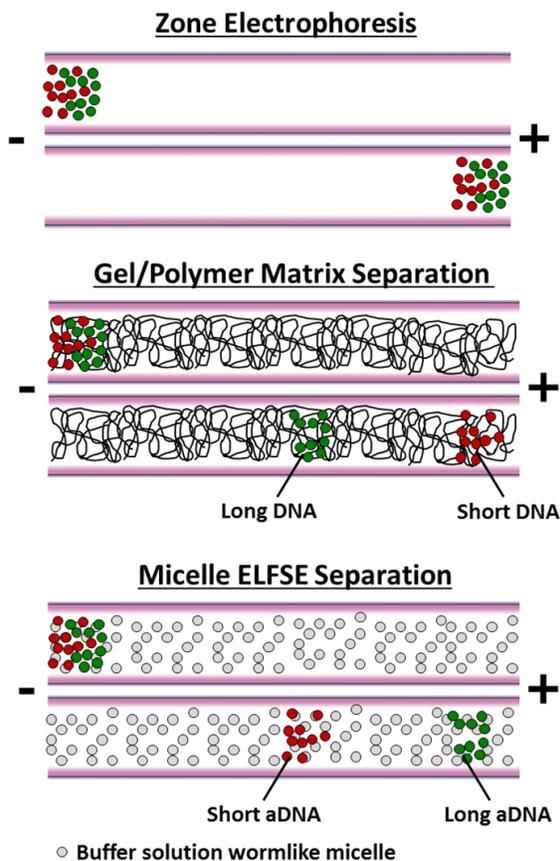
### 1.2.5 – End-Labeled Free-Solution Electrophoresis

End-labeled free-solution electrophoresis (ELFSE) was proposed in 1994 by Mayer *et al.* as a more efficient DNA separation method than gel electrophoresis. <sup>[142]</sup> In ELFSE, DNA strands are end-labeled with an uncharged, monodisperse molecule to generate additional hydrodynamic drag on the electrophoresing DNA. These molecules act as hydrodynamic parachutes, termed “drag-tags”, that return a length dependence to DNA’s  $\mu$ . Longer fragments are less retarded by the additional drag-tag friction and migrate faster than shorter fragments.

In order to predict the velocity of end-labeled DNA strands, the DNA-drag-tag conjugate is treated as a composite polyelectrolyte composed of charged and uncharged segments. The composite molecule’s electrophoretic mobility ( $\mu$ ) is an average of the mobilities of charged ( $\mu_C$ ) and uncharged ( $\mu_U$ ) segments, weighted by their frictional coefficients to the total friction of the composite molecule. <sup>[143]</sup>

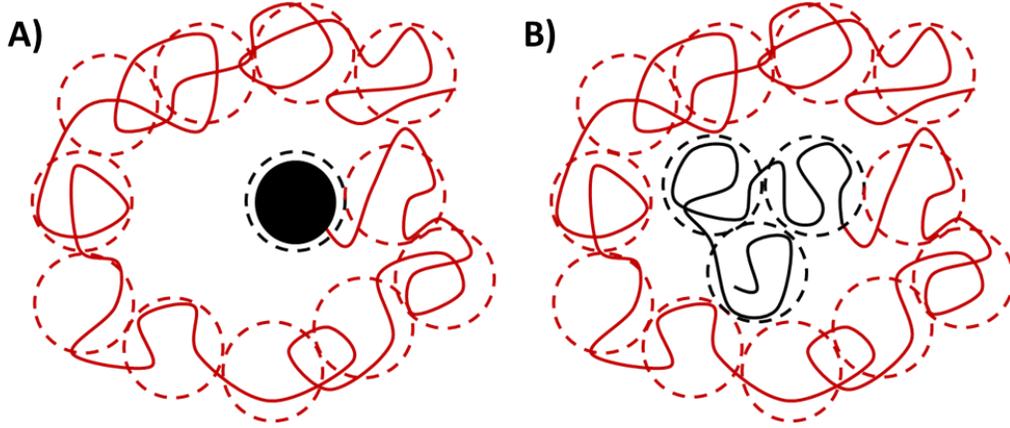
$$\mu = \mu_C \frac{\xi_C}{\xi_C + \xi_U} + \mu_U \frac{\xi_U}{\xi_C + \xi_U} \quad (1.2)$$

Where  $\xi_C$  and  $\xi_U$  are the friction coefficients for the charged and uncharged segments, respectively. Since the drag-tag is uncharged,  $\mu_U$  is negligible and we are left with just the weighted charged mobility segment in Equation 1.2.



**Figure 1.5** – Modes of DNA electrophoresis. Each system represents the initial injection of a mixture of short (red) and long (green) DNA lengths. Upon imposing an electric field, we observe either no separation due to the length-invariant free-solution electrophoretic mobility in capillary zone electrophoresis (Top), elution of the short strands first by CGE (Middle), or elution of the long strands first by micelle ELFSE (Bottom), where aDNA refers to end-alkylated DNA.

Equation 1.2 assumes that the composite polyelectrolyte has a homogeneous backbone, such that the uncharged and charged segments have the same hydrodynamic properties. In most cases, the DNA and drag-tag are not hydrodynamically equivalent such that a weighted average of friction coefficients is invalid.



**Figure 1.6** – Illustration of blob theory for a DNA strand (red) end-attached to A) spherical drag-tag or B) polymeric drag-tag (black). The blob sizes for each segment are equivalent in hydrodynamic radii, as outlined by the dotted circles.

In order for Equation 1.2 to be utilized, the segments are broken down into “blobs” of hydrodynamically equivalent monomers, see Figure 1.6.<sup>[144], [145]</sup> Regardless of size or shape of the drag-tag and DNA, each blob has the same hydrodynamic radius. Since charged and uncharged blobs now have the same friction coefficients, Equation 1.2 can be expressed as the number of charged ( $M_{B,C}$ ) and uncharged ( $M_{B,U}$ ) blobs.

$$\mu = \mu_C \frac{M_{B,C}}{M_{B,C} + M_{B,U}} \quad (1.3)$$

The number of blobs for charged and uncharged ( $i = C, U$ ) segments can be expressed as the ratio of the total number of Kuhn lengths per segment ( $M_{K,i}$ ) divided by the number of Kuhn lengths per blob ( $m_{K,i}$ )

$$M_{B,i} = \frac{M_{K,i}}{m_{K,i}} \quad (1.4)$$

The number of Kuhn segments can be expressed as

$$M_{K,i} = \frac{M_i b_i}{b_{k,i}} \quad (1.5)$$

Where  $M_i$  is the number of  $i$  component monomers,  $b_i$  is the  $i$  component monomer length, and  $b_{k,i}$  is the  $i$  component Kuhn length. Substituting Equations 1.4 and 1.5 into 1.3, the electrophoretic mobility of the composite molecule becomes

$$\mu = \mu_c \frac{M_c}{M_c + \frac{m_{K,C} b_U b_{K,C}}{m_{K,U} b_c b_{K,U}} M_U} = \mu_c \frac{M_c}{M_c + \alpha_1 M_U} \quad (1.6)$$

The term  $\alpha_1$  represents a dimensionless lumped parameter that relates the relative friction per uncharged monomer to charged monomers. Since the blobs have equivalent hydrodynamic radii ( $R_C = R_U = R_B$ ),  $R_B$  can be expressed as

$$R_{B,i} = \frac{2}{3} \sqrt{\frac{b_{K,i}^2 m_{K,i}}{6}} \quad (1.7)$$

We do not have *a priori* knowledge of  $m_{K,i}$ . Equation 1.7 allows us to represent the ratio of  $m_{K,C}$  to  $m_{K,U}$  as known Kuhn lengths of each segment.

$$\frac{m_{K,C}}{m_{K,U}} = \frac{b_{K,U}^2}{b_{K,C}^2} \quad (1.8)$$

Substituting 1.8 into 1.6 yields a new expression for  $\alpha_1$  in terms of the monomer length and Kuhn length of each polymer segment

$$\alpha_1 = \frac{b_U b_{K,U}}{b_c b_{K,C}} \quad (1.9)$$

Finally, the governing theory for DNA migrating under ELFSE is given by Equation 1.10, where we have substituted  $\mu_C$  for the length invariant free-solution electrophoretic mobility of DNA ( $\mu_0$ ),  $M_C$  as the DNA length  $L$ , and the expression  $\alpha_1 M_U$  as  $\alpha$  to account for the additional friction of the drag-tag in terms of hydrodynamically equivalent DNA bases.

$$\mu = \mu_0 \frac{L}{L + \alpha} \quad (1.10)$$

### 1.2.6 – ELFSE with Covalently Bound Drag-Tags

ELFSE was first experimentally realized using covalently attached globular streptavidin drag-tags.<sup>[146]</sup> Although the separation followed ELFSE theory, the resolution was quite poor. The polydispersity of the globular streptavidin drag-tags added peak broadening effects to the DNA during the separation. Attachment of drag-tags with differing sizes is one of the greatest hurdles when implementing ELFSE. In fact, a PEG polymer with a polydispersity index of 1.01 yields a broad smear of peaks for a single length of DNA.<sup>[147]</sup> Through significant size fractionation of streptavidin to produce a monodisperse drag-tag, DNA sequencing up to 110 bases in 18 min was achieved.<sup>[144]</sup>

The drag-tag size for streptavidin is quite small,  $\alpha = 24$  for the 52.8kDa protein.<sup>[144]</sup> As the DNA length increases, the additional friction from the drag-tag becomes negligible and the DNA returns to the inseparable free-solution mobility limit. A few groups have expanded the molecular conformations and chemistries of covalently bound drag-tags to increase  $\alpha$ . A general trend has seen the

transition from globular streptavidin to monodisperse, genetically engineered linear polypeptides.<sup>[86], [148], [149]</sup> The linear polymer conformation sweeps a larger volume than the spherical labels, creating a greater amount of hydrodynamic drag.<sup>[150]</sup> The drag-tag  $\alpha$  value was increased to 56 for a 267 amino acid (aa) protein polymer, capable of sequencing 265 bases in 30 min.<sup>[144], [151]</sup> Increasing the aa-protein polymer to 512 residues more than doubled  $\alpha$  to 152, however attachment to PCR primers inhibited the sequencing reaction, making it difficult to use as a drag-tag.<sup>[152]</sup>

Branched polypeptides were synthesized to test the effect of side chains on the drag-tags separation efficiency.<sup>[153]</sup> Due to limits of solid phase synthesis and monodispersity requirements, short 30mer backbone peptides with five 4mer or 8mer sidechains were produced. The drag-tag was only  $\alpha \sim 17$  due to the low molecular weight of the branched polypeptide. In fact, theoretical analysis revealed that a linear polymer of the same molecular weight would yield a higher  $\alpha$  than the branched polymer.<sup>[150], [154]</sup> This interesting result is due to the branched polymer having a more compact structure than the linear conformation, generating a smaller degree of drag.

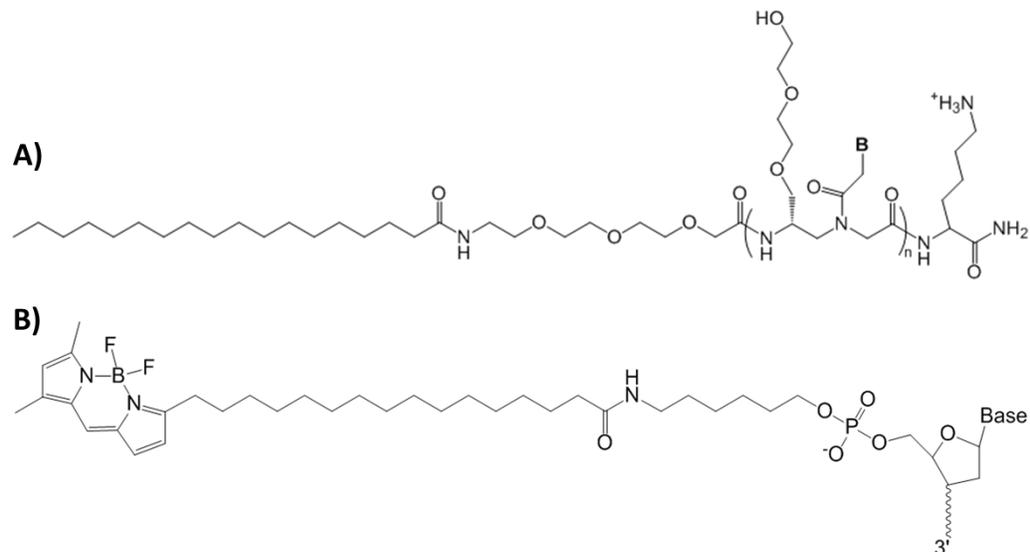
A final attempt to increase drag was made by adding positive residues to the aa-protein polymer.<sup>[155]</sup> The positive charges create an electrokinetic force opposing that of DNA, which is perceived as a higher drag force. A 116 residue aa-protein polymer with 6 positive charges achieved an  $\alpha$  of  $\sim 31$ , compared to  $\alpha = 35$  for a neutral 204 aa-protein tag. Increasing the number of positive charges led to poor

results from the electrostatic interaction of the positive drag-tag with the negative DNA strands and capillary walls.

Since these groups could not increase drag-tag size without encountering detrimental polydispersity limits, current applications for covalently attached drag-tags are limited to DNA < 200 bases.<sup>[86], [87], [156], [157]</sup>

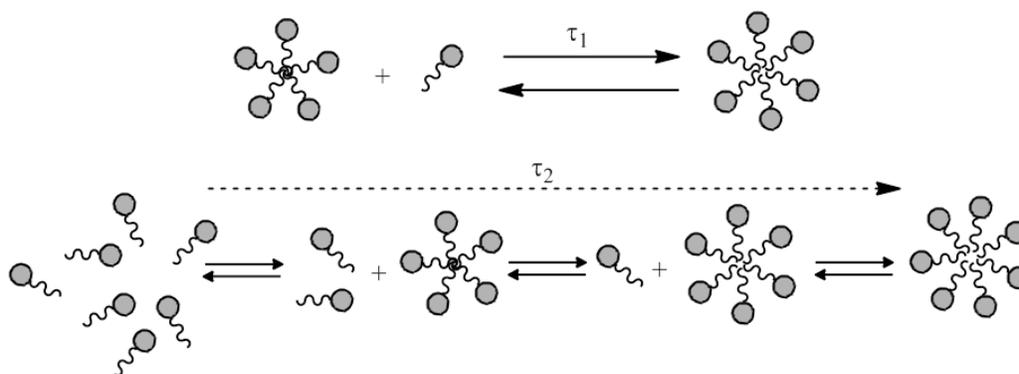
### **1.2.7 – ELFSE with Transient Micelle Interaction**

Above the critical micelle concentration (CMC), surfactant monomers will self-assemble into large micelle aggregates. We have previously demonstrated the use of these large, non-ionic surfactant micelles as successful ELFSE drag-tags.<sup>[113], [114], [158]–[160]</sup> DNA does not readily interact with non-ionic micelles in solution. Our group has utilized PNA amphiphiles (PNAAs) modified with hydrocarbon tails to recognize sequence specific DNAs, as well as end-alkane modified DNA PCR primers to achieve interaction with the non-ionic micelles in solution, see Figure 1.7. The hydrophobic modification partitions into the core of surfactant micelles present in the capillary electrophoresis (CE) running buffer. It is only through modification with the end-alkane that micelles can be implemented as ELFSE drag-tags.



**Figure 1.7** – Chemical structure of hydrophobic modifications to A) (PEG)  $\gamma$ -carbon substituted peptide nucleic acid amphiphile ( $\gamma$ PNAA) and B) BODIPY C16 hydrophobically modified DNA that are used to encourage interaction of DNA with non-ionic surfactant micelles for ELFSE separations.

The size and shape of these micelle aggregates depends on monomer geometry.<sup>[161]</sup> Surfactants with large hydrophilic head groups tend to form spherical micelles. Surfactants with large hydrophobic tail groups tend to form cylindrical micelles, or even bilayer vesicle structures. Micelles are also dynamic structures constantly undergoing exchanges with individual surfactant monomers between the bulk and micelle, as well as micelle formation and dissolution. This process is characterized by two time scales; a fast monomer exchange rate in and out of the micelles ( $\tau_1$ ), and a slower micelle lifetime related to micelle stability ( $\tau_2$ ), see Figure 1.8.



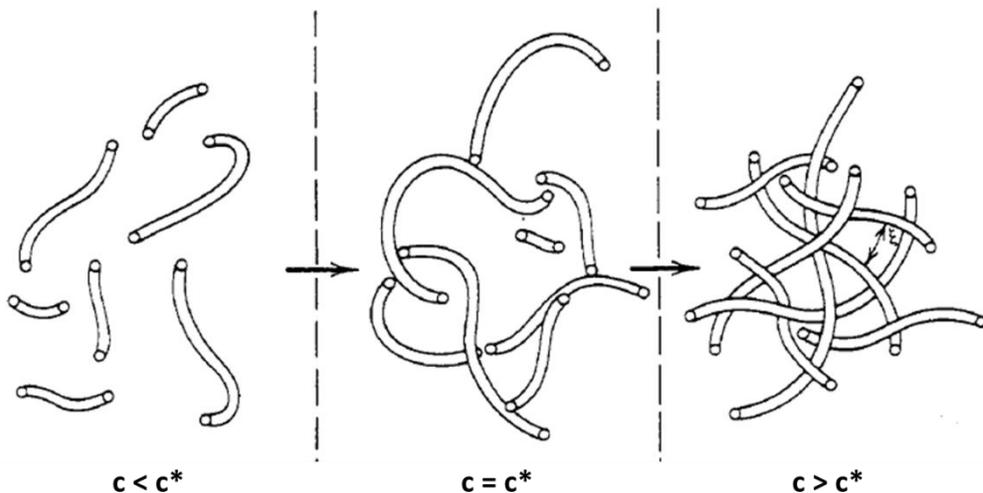
**Figure 1.8** – Illustration of the fast monomer exchange  $\tau_1$ , and the slower micelle formation and dissolution  $\tau_2$ . Adapted from [162]

These time scales are considered as step-wise addition reactions, limited by the mass transfer between individual monomers and intermediate micelle aggregates.<sup>[163]</sup> Table 1.1 lists  $\tau_1$  and  $\tau_2$  for the anionic surfactant sodium dodecyl sulfate (SDS), non-ionic Triton X-100, and *n*-alkyl polyoxyethylene ether wormlike micelle surfactants ( $C_iE_jS$ ). We can see that all three surfactant monomer systems share very similar  $\tau_1$  values, on the order of  $\mu\text{s}$ . This is due to their similar surfactant molecular weights, generating similar mass transfer limitations. The ionic SDS has a much faster  $\tau_2$  than the non-ionic surfactants. The electrostatic repulsion between SDS monomers in the micelle reduces its lifetime compared to non-ionic surfactants. Regardless, the slowest micelle time scale  $\tau_2$  is on the order of ms, much faster than any time scale present for DNA separations.

<b>Table 1.1</b> – Surfactant Exchange and Micelle Relaxation Time Scales		
<b>Surfactant</b>	<b>Monomer Exchange, <math>\tau_1</math> (<math>\mu\text{s}</math>)</b>	<b>Micelle Formation/Dissolution, <math>\tau_2</math> (ms)</b>
SDS <sup>[164]</sup>	29	0.0018
Triton X-100 <sup>[165]</sup>	50	3.5
$C_iE_jS$ <sup>[162], [166]</sup>	1 – 35	0.5 – 10

Although micelles are inherently polydisperse in size, the rapid  $\tau_1$  and  $\tau_2$  micelle dynamics leads to fluctuations in micelle drag-tag size. Over the course of the separation, the DNA is expected to interact with thousands of different micelle sizes. Thus, the effective drag-tag size of the micelle is an average between the populations of micelles sampled by the DNA. This sampling effect eliminates the requirement for perfectly monodisperse drag-tags, generating extremely efficient ELFSE separations.<sup>[159], [162], [167], [168]</sup>

Using  $\sim 2.5\text{nm}$  globular Triton X-100 micelles, we have achieved an  $\alpha = 58$  and DNA sequencing of 110 bases in 4 minutes.<sup>[113], [114], [158], [159]</sup> These results clearly demonstrate the effectiveness of micelle sampling for micelle drag-tags in ELFSE separations. Since the limitations of polydispersity are overcome by micelle sampling in place of covalently attached drag-tags, alternative surfactant systems were investigated to dramatically increase  $\alpha$ .<sup>[160]</sup> New surfactant formulations included liposomes formed by polar lipids, nanoemulsions, and  $C_iE_j$  wormlike micelle forming surfactants.<sup>[160]</sup> Of these systems, only the  $C_iE_j$  surfactant systems yielded efficient ELFSE separations, believed to be due to increased rates of micellization kinetics compared to liposomes and nanoemulsions. Using a binary  $C_iE_j$  with a small amount of urea, we achieved an  $\alpha = 502$  and DNA sequencing of 420 bases in 45 minutes.<sup>[159]</sup>



**Figure 1.9** – The entanglement of wormlike micelles as surfactant concentration increases from  $c < c^*$  to  $c > c^*$ , where  $c^*$  represents the overlap concentration of the surfactant system when micelles just begin to interact. Adapted from [169]

Efforts to increase  $\alpha$  and DNA sequencing length of reads were hindered by  $C_iE_j$  wormlike micelle entanglement. As the surfactant concentrations increases, the entangled network creates a system that mimics polymer gels, see Figure 1.9. These networks separate DNA by conventional gel electrophoresis mechanisms, counter to the elution order of ELFSE and resulting in poor separation efficiency.<sup>[170]–[172]</sup> Thus, network formation has limited effective micelle ELFSE separations to  $\alpha \sim 500$ .

### 1.3 – Thesis Objectives

The work presented in this thesis aims to develop a rapid CE based miRNA detection assay capable of direct, high-throughput detection with fM sensitivity using high affinity  $\gamma$ PNA probes, ultrafluorescent DNA nanotags, and rapid micelle ELFSE separations. We will address the following aspects of this goal:

- 1) We will confirm direct, stable hybridization of two probes to miRNA targets in a sandwich hybridization format using high affinity  $\gamma$ PNAAs and DNA nanotags. We will demonstrate the effective detection of bound miRNA from unbound probes via rapid micelle ELFSE separation. Using this technique, we determine run conditions to achieve single base mismatch discrimination between closely related miRNA family members.
- 2) We will extend this method towards high-throughput, multiplexed detection of a *let-7* miRNA family. We will present theoretical predictions on the multiplexing capabilities of this method by simply increasing the length of DNA nanotag hybridized to the target. We will present sensitive pM detection limits using short, oligomeric DNA nanotags stained with PicoGreen dye. In order to increase sensitivity, we explore optimized CE injections that can dramatically concentrate DNA nanotags to yield fM detection limits.
- 3) We find that longer DNA nanotags are necessary in order to achieve fM detection limits in our system. We explore the use of large, wormlike micelle running buffers for separation of these long DNA probes. We find that the presence of an entangled wormlike micelle network hinders the efficiency of our separation mechanism. We characterize the effects of micelle entanglement on the electrophoretic separation of both native DNA and end-alkylated DNA. We find that the entangled network leads to non-ELFSE migration of the end-alkylated DNAs. We offer a modified ELFSE-sieving model to better predict separation performance of end-alkylated DNA in

entangled micelle systems. We also investigate how temperature and small, micelle end-capping surfactants can reduce entanglement to recover efficient separations. We present optimized separation buffers for long DNA separations and improved DNA sequencing length of reads through these methods of entanglement reduction.

- 4) Finally, a microfluidic based system for DNA electrophoresis is presented. We demonstrate micelle ELFSE microchip separations that perform very similarly to CE systems. We find that the reduced initial injection width creates a system that is almost completely limited by diffusional peak broadening. Using theory based on peak broadening and full-field chip imaging detection, we predict that microchip micelle ELFSE may grant DNA sequencing of 450 bases in 2 minutes.

## 1.4 – References

- [1] M. Chalfie, H. R. Horvitz, and J. E. Sulston, *Cell*, 1981, 24, 59–69.
- [2] R. C. Lee, R. L. Feinbaum, and V. Ambros, *Cell*, 1993, 75, 843–854.
- [3] B. Wightman, I. Ha, and G. Ruvkun, *Cell*, 1993, 75, 855–862.
- [4] F. J. Slack, M. Basson, Z. Liu, V. Ambros, H. R. Horvitz, and G. Ruvkun, *Mol. Cell*, 2000, 5, 659–669.
- [5] B. J. Reinhart, F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz, and G. Ruvkun, *Nature*, 2000, 403, 901–906.
- [6] A. E. Pasquinelli, B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, E. E. Ball, B. Degnan, P. Muller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson, and G. Ruvkun, *Nature*, 2000, 408, 86–89.
- [7] R. C. Lee and V. Ambros, *Science*, 2001, 294, 862–864.
- [8] A. E. Pasquinelli and G. Ruvkun, *Annu. Rev. Cell Dev. Biol.*, 2002, 18, 495–513.

- [9] J. Couzin, *Science*, 2002, 298, 2296–2297.
- [10] V. Ambros, *Nature*, 2004, 431, 350–355.
- [11] D. P. Bartel, *Cell*, 2004, 116, 281–297.
- [12] E. C. Lai, *Curr. Biol.*, 2003, 13, 925–936.
- [13] L. He and G. J. Hannon, *Nat. Rev. Genet.*, 2004, 5, 522–531.
- [14] V. N. Kim, *Nat. Rev. Mol. Cell Biol.*, 2005, 6, 376–385.
- [15] B. P. Lewis, C. B. Burge, and D. P. Bartel, *Cell*, 2005, 120, 15–20.
- [16] R. C. Friedman, K. K. H. Farh, C. B. Burge, and D. P. Bartel, *Genome Res.*, 2009, 19, 92–105.
- [17] A. Krek Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., Rajewsky, N., *Nat. Genet.*, 2005, 37, 495–500.
- [18] A. Esquela-Kerscher and F. J. Slack, *Nat. Rev. Cancer*, 2006, 6, 259–269.
- [19] J. P. Cogswell, J. Ward, I. A Taylor, M. Waters, Y. Shi, B. Cannon, K. Kelnar, J. Kemppainen, D. Brown, C. Chen, R. K. Prinjha, J. C. Richardson, A. M. Saunders, A. D. Roses, and C. A Richards, *J. Alzheimers. Dis.*, 2008, 14, 27–41.
- [20] E. Van Rooij, W. S. Marshall, and E. N. Olson, *Circ. Res.*, 2008, 103, 919–928.
- [21] E. E. Creemers, A. J. Tijssen, and Y. M. Pinto, *Circ. Res.*, 2012, 110, 483–495.
- [22] S. Volinia, G. A. Calin, C. G. Liu, S. Ambs, A. Cimmino, F. Petrocca, R. Visone, M. Iorio, C. Roldo, M. Ferracin, R. L. Prueitt, N. Yanaihara, G. Lanza, A. Scarpa, A. Vecchione, M. Negrini, C. C. Harris, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 2257–2261.
- [23] M. V Iorio, M. Ferracin, C. G. Liu, A. Veronese, R. Spizzo, S. Sabbioni, E. Magri, M. Pedriali, M. Fabbri, M. Campiglio, S. Menard, J. P. Palazzo, A. Rosenberg, P. Musiani, S. Volinia, I. Nenci, G. A. Calin, P. Querzoli, M. Negrini, and C. M. Croce, *Cancer Res.*, 2005, 65, 7065–7070.
- [24] G. A. Calin and C. M. Croce, *Nat. Rev. Cancer*, 2006, 6, 857–866.
- [25] S. Wang, C. Bian, Z. Yang, J. Li, L. Zeng, H. Zhou, and R. C. Zhao, *Int. J. Oncol.*, 2009, 34, 1461–1466.
- [26] X. Guo, Y. Wu, and R. Hartley, *RNA Biol.*, 2009, 6, 575–583.

- [27] R. Spizzo, M. S. Nicoloso, L. Lupini, Y. Lu, J. Fogarty, S. Rossi, B. Zagatti, M. Fabbri, A. Veronese, X. Liu, R. Davulutri, C. M. Croce, G. Mills, M. Negrini, and G. A. Calin, *Cell Death Differ.*, 2010, 17, 246–254.
- [28] L. Ma, J. Teruya-Feldstein, and R. A. Weinberg, *Nature*, 2007, 449, 682–689.
- [29] G. A. Calin, C. Sevignani, C. D. Dumitru, T. Hyslop, E. Noch, S. Yendamuri, M. Shimizu, S. Rattan, F. Bullrich, M. Negrini, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 2999–3004.
- [30] N. Kosaka, H. Iguchi, and T. Ochiya, *Cancer Sci.*, 2010, 101, 2087–2092.
- [31] W. Zhu, W. Qin, U. Atasoy, and E. Sauter, *BMC Res. Notes*, 2009, 2, 89.
- [32] X. Zhao, R. Tapeç-Dytioco, and W. Tan, *J. Am. Chem. Soc.*, 2003, 125, 11474–11475.
- [33] L. Chen, S. Lee, M. Lee, C. Lim, J. Choo, J. Y. Park, S. Lee, S. W. Joo, K. H. Lee, and Y.-W. Choi, *Biosens. Bioelectron.*, 2008, 23, 1878–1882.
- [34] S. Li, H. Liu, L. Liu, L. Tian, and N. He, *Anal. Biochem.*, 2010, 405, 141–143.
- [35] C. Y. Zhang and J. Hu, *Anal. Chem.*, 2010, 82, 1921–1927.
- [36] M. Dequaire and A. Heller, *Anal. Chem.*, 2002, 74, 4370–4377.
- [37] E. E. Ferapontova, M. N. Hansen, A. M. Saunders, S. Shipovskov, D. S. Sutherland, and K. V. Gothelf, *Chem. Commun. (Cambridge, United Kingdom)*, 2010, 46, 1836–1838.
- [38] J. Wang, G. Liu, and A. Merkoçi, *J. Am. Chem. Soc.*, 2003, 125, 3214–3215.
- [39] R. M. Franzini and E. T. Kool, *Chem. Eur. J.*, 2011, 17, 2168–2175.
- [40] D. J. Kleinbaum and E. T. Kool, *Chem. Commun. (Cambridge, United Kingdom)*, 2010, 46, 8154–8156.
- [41] Y. C. Cao, R. Jin, and C. Mirkin, *Science*, 2002, 297, 1536–1540.
- [42] J. Chen, J. Lozach, E. W. Garcia, B. Barnes, S. Luo, I. Mikoulitch, L. Zhou, G. Schroth, and J. B. Fan, *Nucleic Acids Res.*, 2008, 36, e87.
- [43] R. Shi and V. L. Chiang, *Biotechniques*, 2005, 39, 519–525.
- [44] J. Shingara, K. Keiger, J. Shelton, W. Laosinchai-Wolf, P. Powers, R. Conrad, D. Brown, and E. Labourier, *RNA*, 2005, 11, 1461–1470.
- [45] E. M. Harcourt and E. T. Kool, *Nucleic Acids Res.*, 2012, 40, e65.

- [46] Y. Cheng, X. Zhang, Z. Li, X. Jiao, Y. Wang, and Y. Zhang, *Angew. Chemie - Int. Ed.*, 2009, 48, 3268–3272.
- [47] S. C. Chapin and P. S. Doyle, *Anal. Chem.*, 2011, 83, 7179–7185.
- [48] J. Li, S. Schachermer, Y. Wang, Y. Yin, and W. Zhong, *Anal. Chem.*, 2009, 81, 9723–9729.
- [49] M. Castoldi, S. Schmidt, V. Benes, M. Noerholm, A. E. Kulozik, M. W. Hentze, and M. U. Muckenthaler, *RNA*, 2006, 12, 913–920.
- [50] C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak, and K. J. Guegler, *Nucleic Acids Res.*, 2005, 33, e179.
- [51] Y. Zhang and C. Zhang, *Anal. Chem.*, 2012, 84, 224–231.
- [52] T. D. Schmittgen, E. J. Lee, J. Jiang, A. Sarkar, L. Yang, T. S. Elton, and C. Chen, *Methods*, 2008, 44, 31–38.
- [53] Y. Li, C. Qiu, J. Tu, B. Geng, J. Yang, T. Jiang, and Q. Cui, *Nucleic Acids Res.*, 2014, 42, 1–5.
- [54] A. Kozomara and S. Griffiths-Jones, *Nucleic Acids Res.*, 2011, 39, D152–D157.
- [55] S. Griffiths-Jones, R. J. Grocock, S. van Dongen, A. Bateman, and A. J. Enright, *Nucleic Acids Res.*, 2006, 34, D140–D144.
- [56] S. Griffiths-Jones, H. K. Saini, S. van Dongen, and A. J. Enright, *Nucleic Acids Res.*, 2008, 36, D154–D158.
- [57] M. V. Iorio, M. Ferracin, C. G. Liu, A. Veronese, R. Spizzo, S. Sabbioni, E. Magri, M. Pedriali, M. Fabbri, M. Campiglio, S. Ménard, J. P. Palazzo, A. Rosenberg, P. Musiani, S. Volinia, I. Nenci, G. a. Calin, P. Querzoli, M. Negrini, and C. M. Croce, *Cancer Res.*, 2005, 65, 7065–7070.
- [58] R. Garzon, F. Pichiorri, T. Palumbo, R. Iuliano, A. Cimmino, R. Aqeilan, S. Volinia, D. Bhatt, H. Alder, G. Marcucci, G. A Calin, C. G. Liu, C. D. Bloomfield, M. Andreeff, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 5078–5083.
- [59] R. Liu, C. Zhang, Z. Hu, G. Li, C. Wang, C. Yang, D. Huang, X. Chen, H. Zhang, R. Zhuang, T. Deng, H. Liu, J. Yin, S. Wang, K. Zen, Y. Ba, and C. Y. Zhang, *Eur. J. Cancer*, 2011, 47, 784–791.
- [60] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, and T. Tuschl, *Science*, 2001, 294, 853–858.

- [61] A. Válczi, C. Hornyik, N. Varga, J. Burgyán, S. Kauppinen, and Z. Havelda, *Nucleic Acids Res.*, 2004, 32, e175.
- [62] Y. Chen, J. A. L. Gelfond, L. M. McManus, and P. K. Shireman, *BMC Genomics*, 2009, 10, 407.
- [63] N. Rosenfeld, R. Aharonov, E. Meiri, S. Rosenwald, Y. Spector, M. Zepeniuk, H. Benjamin, N. Shabes, S. Tabak, A. Levy, D. Lebanony, Y. Goren, E. Silberschein, N. Targan, A. Ben-Ari, S. Gilad, N. Sion-Vardy, A. Tobar, M. Feinmesser, O. Kharenko, O. Nativ, D. Nass, M. Perelman, A. Yosepovich, B. Shalmon, S. Polak-Charcon, E. Fridman, A. Avniel, I. Bentwich, Z. Bentwich, D. Cohen, A. Chajut, and I. Barshack, *Nat. Biotechnol.*, 2008, 26, 462–469.
- [64] W. Li and K. Ruan, *Anal. Bioanal. Chem.*, 2009, 394, 1117–1124.
- [65] C. G. Liu, G. A. Calin, B. Meeloo, N. Gamliel, C. Sevignani, M. Ferracin, C. D. Dumitru, M. Shimizu, S. Zupo, M. Dono, H. Alder, F. Bullrich, M. Negrini, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 9740–9744.
- [66] P. Mestdagh, T. Feys, N. Bernard, S. Guenther, C. Chen, F. Speleman, and J. Vandesompele, *Nucleic Acids Res.*, 2008, 36, e143.
- [67] F. Tang, P. Hajkova, S. C. Barton, K. Lao, and M. A. Surani, *Nucleic Acids Res.*, 2006, 34, e9.
- [68] J. Shingara, K. Keiger, J. Shelton, W. Laosinchai-Wolf, P. Powers, R. Conrad, D. Brown, and E. Labourier, *RNA*, 2005, 11, 1461–1470.
- [69] R. Levicky and A. Horgan, *Trends Biotechnol.*, 2005, 23, 143–149.
- [70] M. F. Hagan and A. K. Chakraborty, *J. Chem. Phys.*, 2004, 120, 4958–4968.
- [71] A. W. Peterson, R. J. Heaton, and R. M. Georgiadis, *Nucleic Acids Res.*, 2001, 29, 5163–5168.
- [72] J. H. Watterson, P. A. E. Piuanno, C. C. Wust, and U. J. Krull, *Langmuir*, 2000, 16, 4984–4992.
- [73] L. A. Goff, M. Yang, J. Bowers, R. C. Getts, R. W. Padgett, and R. P. Hart, *RNA Biol.*, 2005, 2, 93–100.
- [74] F. Sato, S. Tsuchiya, K. Terasawa, and G. Tsujimoto, *PLoS One*, 2009, 4.
- [75] K. A. Cissell, Y. Rahimi, S. Shrestha, E. A. Hunt, and S. K. Deo, *Anal. Chem.*, 2008, 80, 2319–2325.
- [76] S. Husale, H. H. J. Persson, and O. Sahin, *Nature*, 2009, 462, 1075–1078.

- [77] H. Šípová, S. Zhang, A. M. Dudley, D. Galas, K. Wang, and J. Homola, *Anal. Chem.*, 2010, 82, 10110–10115.
- [78] G. J. Zhang, J. H. Chua, R. E. Chee, A. Agarwal, and S. M. Wong, *Biosens. Bioelectron.*, 2009, 24, 2504–2508.
- [79] H. Yang, A. Hui, G. Pampalakis, L. Soleymani, F. F. Liu, E. H. Sargent, and S. O. Kelley, *Angew. Chemie - Int. Ed.*, 2009, 48, 8461–8464.
- [80] B. N. Johnson and R. Mutharasan, *Anal. Chem.*, 2012, 84, 10426–10436.
- [81] M. B. Baker, G. Bao, and C. D. Searles, *Nucleic Acids Res.*, 2012, 40, 1–12.
- [82] Y. Kato, *Nucleic Acids Symp. Ser. (Oxf.)*, 2008, 71–72.
- [83] A. Persat and J. G. Santiago, *Anal. Chem.*, 2011, 83, 2310–2316.
- [84] R. B. Schoch, M. Ronaghi, and J. G. Santiago, *Lab Chip*, 2009, 9, 2145–2152.
- [85] L. A. Neely, S. Patel, J. Garver, M. Gallo, M. Hackett, S. McLaughlin, M. Nadel, J. Harris, S. Gullans, and J. Rooke, *Nat Meth*, 2006, 3, 41–46.
- [86] D. W. Wegman, L. T. Cherney, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2013, 85, 6518–6523.
- [87] D. W. Wegman and S. N. Krylov, *Angew. Chemie*, 2011, 123, 10519–10523.
- [88] D. W. Wegman, F. Ghasemi, A. Khorshidi, B. B. Yang, S. K. Liu, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2015, 87, 1404–1410.
- [89] N. Li, A. Nguyen, J. Diedrich, and W. Zhong, *J. Chromatogr. A*, 2008, 1202, 220–223.
- [90] M. V. Berezovski and N. Khan, *Methods Mol. Biol.*, 2013, 1039, 245–259.
- [91] A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh, and J. Wengel, *J. Am. Chem. Soc.*, 1998, 120, 13252–13253.
- [92] P. M. McTigue, R. J. Peterson, and J. D. Kahn, *Biochemistry*, 2004, 43, 5388–5405.
- [93] E. Várallyay, J. Burgyán, and Z. Havelda, *Nat. Protoc.*, 2008, 3, 190–196.
- [94] M. Castoldi, S. Schmidt, V. Benes, M. W. Hentze, and M. U. Muckenthaler, *Nat. Protoc.*, 2008, 3, 321–329.
- [95] W. P. Kloosterman, E. Wienholds, E. de Bruijn, S. Kauppinen, and R. H. A. Plasterk, *Nat. Methods*, 2006, 3, 27–29.

- [96] N. Tolstrup, P. S. Nielsen, J. G. Kolberg, A. M. Frankel, H. Vissing, and S. Kauppinen, *Nucleic Acids Res.*, 2003, 31, 3758–3762.
- [97] R. Owczarzy, Y. You, C. L. Groth, and A. V. Tataurov, *Biochemistry*, 2011, 50, 9352–9367.
- [98] A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil, and D. H. Ly, *J. Am. Chem. Soc.*, 2006, 128, 10258–10267.
- [99] M. Egholm, O. Buchardt, and P. E. Nielsen, *J. Am. Chem. Soc.*, 1992, 114, 1895–1897.
- [100] E. Uhlmann, A. Peyman, G. Breipohl, and D. W. Will, *Angew. Chemie Int. Ed.*, 1998, 37, 2796–2823.
- [101] U. Giesen, W. Kleider, C. Berding, A. Geiger, H. Ørum, and P. E. Nielsen, *Nucleic Acids Res.*, 1998, 26, 5004–5006.
- [102] X. Zhang, T. Ishihara, and D. R. Corey, *Nucleic Acids Res.*, 2000, 28, 3332–3338.
- [103] T. Ishihara and D. R. Corey, *J. Am. Chem. Soc.*, 1999, 121, 2012–2020.
- [104] H. J. Larsen, T. Bentin, and P. E. Nielsen, *Biochim. Biophys. Acta - Gene Struct. Expr.*, 1999, 1489, 159–166.
- [105] S. Ghosh, S. Mishra, T. Banerjee, and R. Mukhopadhyay, *Langmuir*, 2013, 29, 3370–3379.
- [106] M. M. Fabani and M. J. Gait, *RNA*, 2008, 14, 336–346.
- [107] N. Zhang and D. H. Appella, *J. Am. Chem. Soc.*, 2007, 129, 8424–8425.
- [108] S. R. Ryoo, J. Lee, J. Yeo, H. K. Na, Y. K. Kim, H. Jang, J. H. Lee, S. W. Han, Y. Lee, V. N. Kim, and D. H. Min, *ACS Nano*, 2013, 7, 5882–5891.
- [109] S. Karkare and D. Bhatnagar, *Appl. Microbiol. Biotechnol.*, 2006, 71, 575–586.
- [110] B. Hyrup and P. Nielsen, *Bioorg. Med. Chem.*, 1996, 4, 5–23.
- [111] J. P. Vernille, L. C. Kovell, and J. W. Schneider, *Bioconjug. Chem.*, 2004, 15, 1314–1321.
- [112] C. Lau, R. Bitton, H. Bianco-Peled, D. G. Schultz, D. J. Cookson, S. T. Grosser, and J. W. Schneider, *J. Phys. Chem. B*, 2006, 110, 9027–9033.
- [113] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.

- [114] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [115] M. Egholm, O. Buchardt, P. E. Nielsen, and R. H. Berg, *J. Am. Chem. Soc.*, 1992, 114, 1895–1897.
- [116] B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage, and D. H. Ly, *J. Org. Chem.*, 2011, 76, 5614–5627.
- [117] S. Sforza, G. Haaima, R. Marchelli, and P. E. Nielsen, *European J. Org. Chem.*, 1999, 1999, 197–204.
- [118] S. C. Chapin, D. C. Appleyard, D. C. Pregibon, and P. S. Doyle, *Angew. Chemie*, 2011, 123, 2337–2341.
- [119] J. Lu, G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B. L. Ebert, R. H. Mak, A. A. Ferrando, J. R. Downing, T. Jacks, H. R. Horvitz, and T. R. Golub, *Nature*, 2005, 435, 834–838.
- [120] A. N. Glazer and L. Stryer, *Methods Enzymol.*, 1990, 167, 188–194.
- [121] M. J. Waring, J. B. Chaires, and B. A. Armitage, “Cyanine Dye–DNA Interactions: Intercalation, Groove Binding, and Aggregation,” in *DNA Binders and Related Subjects*, 253, Springer Berlin / Heidelberg, 2005, 55–76.
- [122] L. Strekowski and B. Armitage, “Cyanine Dye–Nucleic Acid Interactions,” in *Heterocyclic Polymethine Dyes*, 14, Springer Berlin / Heidelberg, 2008, 11–29.
- [123] A. I. Dragan, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman, and C. D. Geddes, *Biophys. J.*, 2010, 99, 3010–3019.
- [124] A. N. Glazer and H. S. Rye, *Nature*, 1992, 359, 859–861.
- [125] H. S. Rye, S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies, and A. N. Glazer, *Nucleic Acids Res.*, 1992, 20, 2803–2812.
- [126] V. L. Singer, L. J. Jones, S. T. Yue, and R. P. Haugland, *Anal. Biochem.*, 1997, 249, 228–238.
- [127] A. N. Glazer and H. S. Rye, *Nature*, 1992, 359, 859–861.
- [128] M. Reuter and D. T. F. Dryden, *Biochem. Biophys. Res. Commun.*, 2010, 403, 225–229.
- [129] H. Chan, L. Chan, R. N. Wong, and H. Li, *Anal. Chem.*, 2010, 82, 6911–6918.
- [130] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [131] U. Mohanty and N. C. Stellwagen, *Biopolymers*, 1999, 49, 209–214.

- [132] D. N. Heiger, A. S. Cohen, and B. L. Karger, *J. Chromatogr.*, 1990, 516, 33–48.
- [133] H. T. Chang and E. S. Yeung, *J. Chromatogr. B. Biomed. Appl.*, 1995, 669, 113–123.
- [134] S. J. Kim, G. W. Shin, S. J. Choi, H. S. Hwang, G. Y. Jung, and T. K. Seo, *Electrophoresis*, 2010, 31, 1108–1115.
- [135] C. P. Fredlake, D. G. Hert, C.-W. Kan, T. N. Chiesl, B. E. Root, R. E. Forster, and A. E. Barron, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, 105, 476–481.
- [136] A. E. Barron, W. M. Sunada, and H. W. Blanch, *Electrophoresis*, 1996, 17, 744–757.
- [137] P. D. Grossman and D. S. Soane, *Biopolymers*, 1991, 31, 1221–1228.
- [138] J. L. Viovy and T. Duke, *Electrophoresis*, 1993, 14, 322–329.
- [139] A. G. Ogston, *Trans. Faraday Soc.*, 1958, 54, 1754–1757.
- [140] G. W. Slater and J. Noolandi, *Biopolymers*, 1989, 28, 1781–1791.
- [141] G. W. Slater, *Electrophoresis*, 2009, 30.
- [142] P. Mayer, G. W. Slater, and G. Drouin, *Anal. Chem.*, 1994, 66, 1777–1780.
- [143] D. Long and A. Ajdari, *Electrophoresis*, 1996, 17, 1161–1166.
- [144] H. Ren, A. E. Karger, F. Oaks, S. Menchen, G. W. Slater, and G. Drouin, *Electrophoresis*, 1999, 20, 2501–2509.
- [145] L. C. McCormick, G. W. Slater, A. E. Karger, W. N. Vreeland, A. E. Barron, C. Desruisseaux, and G. Drouin, *J. Chromatogr. A*, 2001, 924, 43–52.
- [146] C. Heller, G. W. Slater, P. Mayer, N. Dovichi, D. Pinto, J. L. Viovy, and G. Drouin, *J. Chromatogr. A*, 1998, 806, 113–121.
- [147] W. N. Vreeland, C. Desruisseaux, A. E. Karger, G. Drouin, G. W. Slater, and A. E. Barron, *Anal. Chem.*, 2001, 73, 1795–1803.
- [148] J. Won, R. J. Meagher, and A. E. Barron, *Electrophoresis*, 2005, 26, 2138–2148.
- [149] J. S. Lin, J. C. Albrecht, R. J. Meagher, X. Wang, and A. E. Barron, *Biomacromolecules*, 2011, 12, 2275–2284.
- [150] K. Grass, C. Holm, and G. W. Slater, *Macromolecules*, 2009, 42, 5352–5359.
- [151] J. C. Albrecht, J. S. Lin, and A. E. Barron, *Anal. Chem.*, 2011, 83, 509–515.

- [152] R. J. Meagher, J. I. Won, L. C. McCormick, S. Nedelcu, M. M. Bertrand, J. L. Bertram, G. Drouin, A. E. Barron, and G. W. Slater, *Electrophoresis*, 2005, 26, 331–350.
- [153] R. D. Haynes, R. J. Meagher, J. Won, F. M. Bogdan, and A. E. Barron, *Bioconjug. Chem.*, 2005, 16, 929–938.
- [154] S. Nedelcu and G. W. Slater, *Electrophoresis*, 2005, 26, 4003–4015.
- [155] X. Wang, J. C. Albrecht, J. S. Lin, and A. E. Barron, *Biomacromolecules*, 2012, 13, 117–123.
- [156] J. C. Albrecht, A. Kotani, J. S. Lin, S. A. Soper, and A. E. Barron, *Electrophoresis*, 2013, 34, 590–597.
- [157] H. Song, Y. Wang, C. Garson, and K. Pant, *Microfluid. Nanofluidics*, 2014, 17, 693–699.
- [158] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [159] S. B. Istivan and J. W. Schneider, “Rapid Gel-free Separations of DNA via Transient Attachment to Surfactant Micelles,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2012.
- [160] A. Jones, “Rapid Separations of Kilobase-Sized DNA using Micelle Size-Sampling,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2013.
- [161] H. Wennerström, *Phys. Rep.*, 1979, 52, 1–86.
- [162] A. Patist, J. R. Kanicky, P. K. Shukla, and D. O. Shah, *J. Colloid Interface Sci.*, 2002, 245, 1–15.
- [163] E. A. G. Aniansson and S. N. Wali, *J. Phys. Chem.*, 1973, 78, 1024 – 1030.
- [164] D. G. Hall, *J. Chem. Soc. Faraday Trans. 2*, 1981, 77, 1973.
- [165] M. Kahlweit and M. Teubner, *Adv. Colloid Interface Sci.*, 1980, 13, 1–64.
- [166] S. Harada and H. Sahara, *J. Phys. Chem.*, 1995, 99, 12570–12575.
- [167] C. J. Biaselle and D. B. Millar, *Biophys. Chem.*, 1975, 3, 355–361.
- [168] M. A. Fahrenkopf, B. E. Ydstie, T. Mukherjee, and J. W. Schneider, *Comput. Chem. Eng.*, 2014, 64, 1–8.
- [169] M. E. Cates and S. J. Candau, *J. Phys. Condens. Matter*, 1999, 2, 6869–6892.

- [170] Wei and E. S. Yeung, *Anal. Chem.*, 2001, 73, 1776–1783.
- [171] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [172] T. Kato, N. Taguchi, and D. Nozu, *Prog. Colloid Polym. Sci.*, 1997, 106, 57–60.

# **Chapter 2 – High Affinity $\gamma$ PNA Sandwich Hybridization Assay for Rapid MicroRNA Detection with Single Base Mismatch Discrimination**

## **2.1 – Introduction**

To improve probe-target recognition over background noise, many PCR-less DNA and RNA detection assays employ a sandwich hybridization approach where target sequences are bound between two separate probes. Normally one probe provides target segregation from the bulk solution (the capture probe), while the second probe imparts a measurable signal to the hybridization event (the reporter probe). The sandwich hybridization approach has been successfully implemented in a variety of nucleic acid sensing techniques including fluorescence imaging,<sup>[1]-[4]</sup> electrochemical detection,<sup>[5]-[7]</sup> template-mediated fluorescence activation,<sup>[8], [9]</sup> and surface-enhanced Raman scattering.<sup>[10]</sup> In most cases, the requirement of two orthogonal binding events dramatically reduces background noise. However, sandwich assays are generally not viable for short nucleic acid targets such as microRNA (miRNA),<sup>[11]-[14]</sup> as the binding of probes

to either half of the target is too weak. As such, miRNA detection is normally carried out using enzymatic methods to include polyadenylation,<sup>[15]–[17]</sup> hybridization mediated ligation,<sup>[18], [19]</sup> direct target ligation,<sup>[20], [21]</sup> or RT-PCR,<sup>[22], [23]</sup> which are laborious, time-consuming, and expensive compared with simpler direct, one-step DNA or RNA detection.

To our knowledge, only two other groups have reported direct (enzyme-free) sandwich hybridization of targets less than 25 bases long.<sup>[24], [25]</sup> Cai *et al.* detected targets 15-20 bases long using a cooling protocol ( $T = 20^{\circ}\text{C}$ ) to achieve stability of a DNA sandwich complex.<sup>[24]</sup> Neely *et al.* instead used two locked nucleic acids (LNA), a synthetic DNA analog with greater binding stability than natural DNA or RNA, to detect trace miRNA by fluorescence coincidence detection.<sup>[25]</sup> Although the LNA's higher stability permitted sandwich hybridization, design of LNA probes is complicated by a propensity for LNA oligomers to form hairpins and LNA-LNA duplexes, as well as expensive LNA oligomer synthesis.<sup>[26]–[28]</sup>

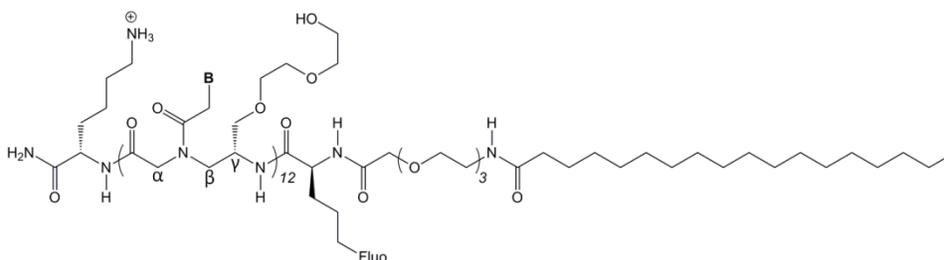
Sahu *et al.* have recently reported a polyethylene glycol (PEG) gamma-substituted peptide nucleic acid ( $\gamma$ PNA) whose nucleobases form highly stable duplexes with both DNA and RNA.<sup>[29]</sup> The high affinity stems from a combination of the  $\gamma$ PNA's reduced electrostatic repulsion between the DNA/RNA backbone and uncharged *N*-(2-aminoethyl)-glycine PNA backbone, and a preorganized B-form like conformation from the  $\gamma$ -carbon modification. Additionally, the PEG substituents on the  $\gamma$ -carbon dramatically increase water solubility of the charge neutral backbone, maintaining its high affinity properties

relative to more hydrophobic substituents such as methyl groups and charged amino-acid residues.<sup>[29]-[31]</sup>

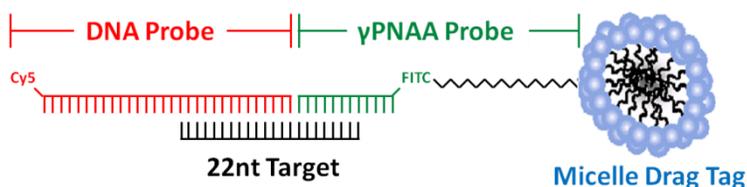
In the case of miRNA detection, the higher affinity of  $\gamma$ PNA may allow for use of inexpensive DNA as a second probe in the binding sandwich, rather than a second  $\gamma$ PNA or LNA probe. This second DNA probe, if long enough, could serve as a highly fluorescent tag using a method introduced by the Armitage group, in which the probe is saturated with fluorescent intercalating dye molecules (about one per four bases of DNA duplex).<sup>[32], [33]</sup> We are developing a miRNA detection method using capillary electrophoresis (CE) based on this idea, and the design of a sandwich that can recognize oligomers 19-23 bases in length, using a natural DNA probe, is critical to that effort.

CE also presents a simple way to measure the yield of sandwich complexes, provided that one of the probes has a much different mobility than the other. In the absence of gels, the mobility of DNA and RNA oligomers is large and approximately constant,<sup>[34]</sup> and that of uncharged  $\gamma$ PNA is negligible. The sandwich will have a slightly lower mobility with  $\gamma$ PNA attached, but the mobility can be drastically shifted by attaching a large non-ionic micelle to the  $\gamma$ PNA terminus. This is done by covalently linking a C18 alkyl group to the  $\gamma$ PNA to form a  $\gamma$ PNA amphiphile ( $\gamma$ PNAA) and adding non-ionic surfactant to the CE running buffer in a manner similar to that reported previously by the Schneider group.<sup>[35], [36]</sup> In this way, bound and unbound components can be easily discriminated in CE using micelle-containing running buffers. The chemical structure of the  $\gamma$ PNAA and a diagram of the resulting sandwich complex are

shown in Figure 2.1. Fluorophores are attached to each probe to aid in identification via coincidence detection.



NH<sub>2</sub>-Lys-**ACTCCATCATCC**-Orn(Fluo)-MP<sub>3</sub>-C<sub>18</sub>



**Figure 2.1** – Structure of  $\gamma$ PNAA (C to N) where **B** = A, T, G, or C used in this study for sandwich hybridization of short nucleic acids and interaction with micelle drag tags.

Using this CE method, we demonstrate successful sandwich hybridization of both probes to 22nt DNA and RNA targets with single base mismatch selectivity. We find that sandwich hybridization is stabilized by the adjacent binding of each probe to the target via coaxial stacking, dramatically increasing the sandwich stability compared to individual binding events. The probes also hybridize under multiple configurations, enabling simple redesign of the sandwich complex without losses in stability. A discussion is provided on the application of the proposed sandwich hybridization strategy for implementation as a miRNA detection method. Although our focus in this report is on the detection of short targets, in principle any nucleic acid target can be quantitatively analyzed using

the  $\gamma$ PNA and DNA probe sandwich hybridization approach in conjunction with capillary electrophoresis techniques.

## **2.2 – Materials and Methods**

### **2.2.1 – Reagents**

Nuclease-free water and hydrochloric acid were purchased from Fisher Scientific (Pittsburgh, PA). Triton X-100 surfactant, sodium hydroxide pellets, DMSO, acetonitrile, and triethylammonium acetate (TEAA) were purchased from Sigma Aldrich (St. Louis, MO). 10x Tris-Borate-EDTA (TBE) buffer was purchased from Promega (Madison, WI). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2m $\Omega$  deionized water, autoclaving, and filtering through a 0.22 $\mu$ m syringe filter. YOYO-1 intercalating dye was purchased from Life Technologies (Carlsbad, CA). DNA and RNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA), with standard desalting purification except the Cy5-labeled DNA probe strand, Duplex strand 1, and Duplex strand 2, which were purchased as HPLC purified. Sequences were designed to avoid undesired non-target complementarity between sandwich components. Dried oligomers were resuspended in nuclease-free water and used without further purification.

### **2.2.2 – $\gamma$ PNA Amphiphile ( $\gamma$ PNA) Synthesis and Purification**

$\gamma$ PNA was manually synthesized as described previously<sup>[29]</sup> where the *N*-terminus was labeled with an ornithine-linked fluorescein dye, a 3 unit mini-polyethylene glycol (MP<sub>3</sub>) spacer, and octadecanoic acid. For ornithine

fluorophore labeling, Boc-*L*-ornithine (Fmoc)-OH was first coupled to the *N*-terminus of the  $\gamma$ PNA on-resin. Fmoc was then removed with 15% piperidine/DMF and the resulting amino side-chain of ornithine was coupled using HATU to 5-carboxyfluorescein. Because of the low solubility of octadecanoic acid in most organic solvents including DMF, which was employed in the coupling step, the reaction was performed three times at 2-hr each. The  $\gamma$ PNAA was cleaved off the resin, ether precipitated, lyophilized, resuspended in nuclease-free water, and purified on a Symmetry C18 300 5 $\mu$ m 4.6mm x 250mm column. A 30 minute linear gradient from 0.1% TEAA in water to acetonitrile at 1mL/min was used for elution. Collected fractions were dried, resuspended in water, and characterized on a PerSeptive Voyager STR MALDI-TOF mass spectrometer (Applied Biosystems, Foster City, CA) (5623.8 obs., 5622.9 theo.).

### **2.2.3 – DNA Target Sandwich Detection with Cy5-labeled DNA Probe**

The  $\gamma$ PNAA probe, 22nt DNA targets, and the Cy5-labeled DNA probe were mixed in 1xTBE, with each component at a concentration of 250nM unless otherwise noted. Solutions were heated to 95°C for at least 5 minutes in a dry bath incubator and allowed to cool to room temperature over the course of 60 minutes. Hybridized complexes not immediately used for separations were stored at -20°C and subjected to minimal freeze-thaws.

CE experiments were performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. Excitation sources were from a 3mW 488nm Argon ion laser and 635nm solid

state laser. LIF detection was performed at 488/520nm and 635/675nm excitation/emission. The capillary was a 50 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and of indicated length in figure captions. Hydrodynamic injections (0.5psi for 10sec) were used to introduce samples into the capillary. Electrophoretic separation was conducted under normal polarity (from cathode to anode) with an applied voltage of 20kV. The capillary was heated to the indicated temperatures using a built in coolant jacket system and hot air convection in the sample chamber. Data collection was performed using 32 Karat software (Beckman Coulter). Peak areas were determined using a MATLAB program where peaks were selected manually and integrated using the trapz function.

Capillaries were prepared for electroosmotic flow (EOF) before each separation using a sequential pressure rinse at 20psi of DI water (5 min), 50/50 methanol/DI water (20min), DI water (5min), 0.1M HCl (20min), DI water (5 min), 0.1M NaOH (20min), and DI water (5min) for a total cycle time of 80min. Once prepared, the capillary could be used for multiple separations with minimal EOF drift. Unhybridized  $\gamma$ PNAA could be used as an EOF marker to determine the electroosmotic velocity of each CE experiment.

#### **2.2.4 – RNA Target Sandwich Detection with YOYO-1 DNA Probe**

The DNA probe for the RNA sandwich complex was a 66 base-pair duplex with a 14 base overhang complementary to the RNA target. Duplex strand 1: 5' ATG CCA TGT GAT CGC AGA TAC GCT TAG ACC TTG TAA CGG CTT GCC

AAG TGA CCT TAC GTG ACT AGT TCA AGC TCA ATG GA 3'. Duplex strand 2: TCC ATT GAG CTT GAA CTA GTC ACG TAA GGT CAC TTG GCA AGC CGT TAC AAG GTC TAA GCG TAT CTG 3'. Preparation of RNA containing solutions was performed in a PCR workstation (Misonix, Creedmorr, NC) equipped with a laminar flow enclosure, HEPA filter, and UV lamp to maintain sterile conditions.

$\gamma$ PNAA probe, 22nt RNA target, Duplex strand 1, and Duplex strand 2 were mixed in 1xTBE, with each component at a concentration of 250nM. Solutions were heated to 95°C for at least 5 minutes in a dry bath incubator and allowed to cool to room temperature over the course of 60 minutes. Hybridized complexes not immediately used for separations were stored at -20°C and subjected to minimal freeze-thaws. YOYO-1 from a DMSO solution was added to the annealed sandwich solution at a concentration of 1 fluorescent dye molecule per every 5 base pairs of DNA duplex. The dye was allowed to intercalate for at least 1 hour in the dark before detection. Stained complexes not immediately used for detection could be stored at -20°C for up to one week.

CE experiments were performed on a P/ACE MDQ equipped for LIF detection. Excitation was from a 3mW 488nm Argon ion laser. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 50 $\mu$ m ID fused-silica capillary and was 20cm to detector and 30cm total length (20/30cm). Hydrodynamic injections (0.5psi for 10sec) were used to introduce samples into the capillary. Electrophoretic separation was conducted under normal polarity (from cathode to anode) with an applied voltage of 20kV. The capillary was

heated to the indicated temperatures using a built in coolant jacket system and hot air convection in the sample chamber. Data collection was performed using 32 Karat software.

Capillaries were prepared for EOF before each separation using a sequential pressure rinse at 20psi of DI water (5 min), 50/50 methanol/DI water (20min), DI water (5min), 0.1M HCl (20min), DI water (5 min), 0.1M NaOH (20min), and DI water (5min) for a total cycle time of 80min. Once prepared, the capillary could be used for multiple separations with minimal EOF drift. Unhybridized  $\gamma$ PNAA could be used as an EOF marker to determine the electroosmotic velocity of each CE experiment.

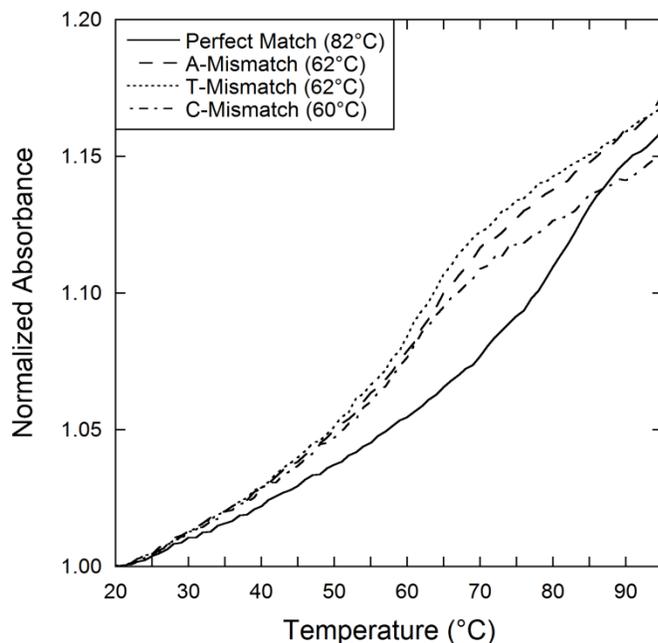
## 2.3 – Results and Discussion

### 2.3.1 – UV Melting Experiments on Full $\gamma$ PNAA-DNA Complements

The melting transition temperature ( $T_m$ ) is used as a measure of DNA duplex stability, where the strands are 50% hybridized and 50% single strands. It can be measured from the hyperchromicity of UV absorbance that occurs during strand denaturation (Figure 2.2) The 12nt  $\gamma$ PNAA hybridized to a fully complementary DNA target yielded a melting transition temperature of 82°C. This value is higher than that predicted using an empirical formula for PNA stability defined by Giesen *et al.*,<sup>[37]</sup> with an additional 2°C per  $\gamma$ -carbon substitution as estimated by Sahu *et al.*,<sup>[29]</sup>  $T_m = 70.1^\circ\text{C}$ ,

$$T_{m \lambda PNA} = 20.79 + 0.83T_{m DNA} - 26.13f_{pyr} + 0.44L + 2.0MP_{units} \quad (2.1)$$

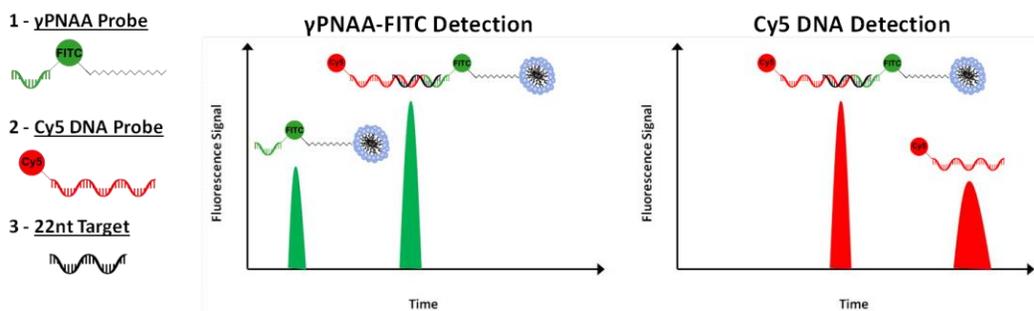
In which  $T_{m\ DNA}$  is the melting temperature calculated using a nearest neighbor model for the corresponding sequence of a DNA/DNA duplex at 1M NaCl concentration as described by SantaLucia et al.,<sup>[38]</sup>  $f_{pyr}$  is the fractional pyrimidine content,  $L$  is the length of the PNA sequence in number of bases, and  $MP_{unit}$  is the number of PNA bases modified with a polyethylene glycol unit on the  $\gamma$ -carbon.



**Figure 2.2** – UV melting profiles of  $\gamma$ PNAA bound to DNA. 5 $\mu$ M Strand concentrations in 1xTBE. Sequences:  $\gamma$ PNAA, C18 - MP<sub>3</sub> - Fluo - CCTACTACCTCA-Lys-NH<sub>2</sub>; DNA, 5' TGA GXT AGT AGG 3' X = A, C, G, T. Both heating and cooling profiles were performed and were identical, only heating curves are shown.

The difference may stem from the empirical nature of the formula compared to well established DNA nearest-neighbor thermodynamic models.<sup>[38]</sup> The  $\gamma$ PNAA  $T_m$  is also much greater than that predicted for a DNA-DNA duplex (38°C), LNA-DNA duplex (65.7°C), and unmodified PNA-DNA duplex (46.1°C) of the same sequence, exemplifying the high affinity properties of the  $\gamma$ PNAA. An alternating LNA-DNA chimera sequence was considered for LNA stability modeling to account for LNA's susceptibility to hairpin formation and LNA-LNA base pairing

when using consecutive LNA modifications.<sup>[27], [28], [39]</sup> The selectivity of the  $\gamma$ PNAA is also maintained at these high stabilities with  $\Delta T_m$ s ranging from -22 to -20°C for single base mismatched off-targets.



**Scheme 2.1** – Detection mechanism for reported sandwich hybridization assay

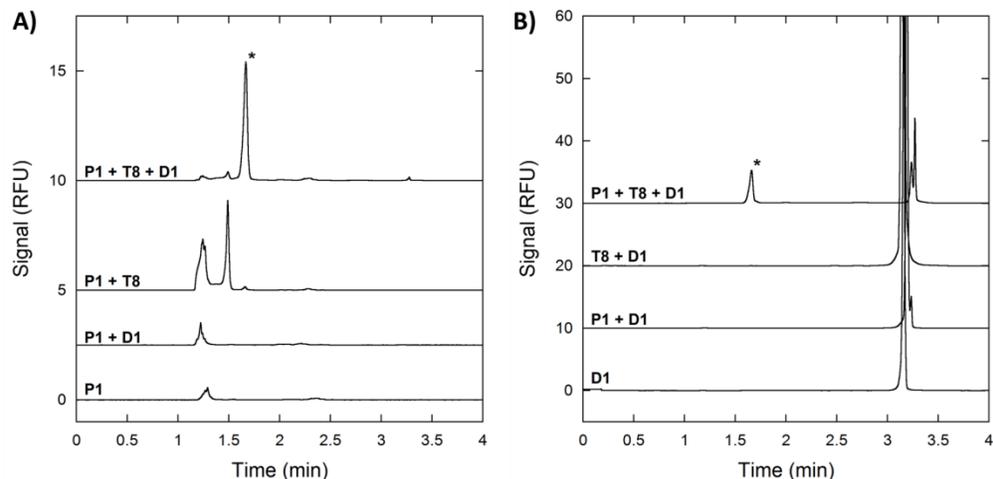
### 2.3.2 – Design of Sandwich Complex

An important aspect of the sandwich design is the number of bases to allocate to each of the probes. To evaluate how many are needed to form a stable complex, we considered a series of  $\gamma$ PNAA each with a different number of bases bound to the DNA target (Table 2.1). Empirical and nearest-neighbor predictions of duplex stabilities suggest that an overlap of only 8 bases is needed to bind a  $\gamma$ PNAA to its DNA target with a  $T_m$  of 41.2°C, while an 8-base overlap would give  $T_m = 28.8^\circ\text{C}$  for LNA,  $T_m = 25.2^\circ\text{C}$  for unmodified PNA, and  $T_m = 16.2^\circ\text{C}$  for DNA.

Sandwich hybridization of the  $\gamma$ PNAA and DNA based probe to the 22nt targets was detected via a mobility shift assay using micelle-containing running buffers in CE.<sup>[35], [36]</sup> Since we operate the CE with a large, constant electro-osmotic counterflow (EOF), the components elute in reverse order of their electrophoretic mobility. Thus, we expect unbound  $\gamma$ PNAA to elute first, followed by the sandwich complex, then the unbound DNA (Scheme 2.1). This is confirmed by

comparing electropherograms collected using detection for the  $\gamma$ PNAA probe (fluorescein) and DNA probe (Cy5), respectively. Figure 2.3 shows CE electropherograms of each component and their stoichiometric mixtures, confirming the expected elution order described above, with a total elution time of less than 3.5 minutes. Note that detection of each component could now be achieved using a single fluorophore, without any quenching moieties, since the expected elution order is confirmed.

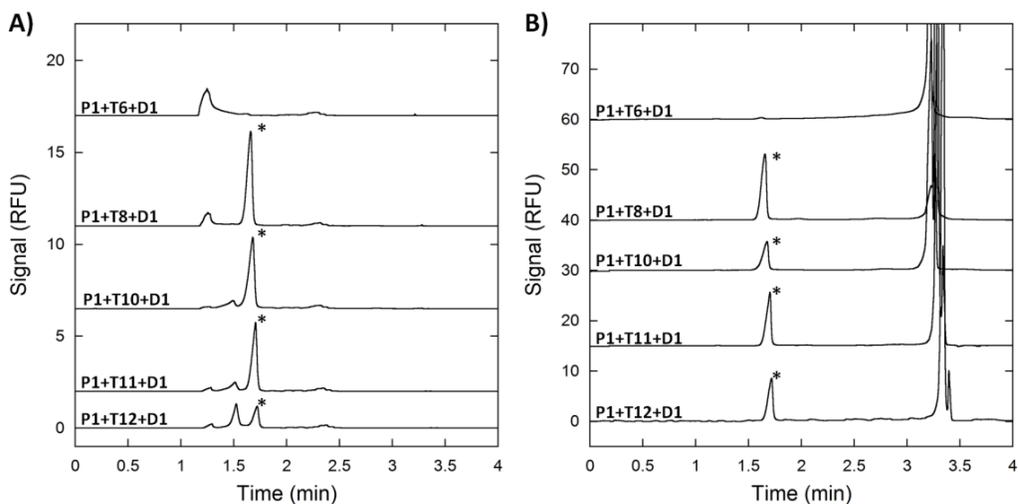
<b>Table 2.1 – <math>\gamma</math>PNAA, DNA Probe, and Target Sequences and <math>T_m</math>s</b>					
Name		Sequence (5'-3')			
<u>Hybridization Probes</u>					
$\gamma$ PNAA (N-C)	P1	C <sub>18</sub> - MP <sub>3</sub> - Fluo – CCTACTACCTCA-Lys-NH <sub>2</sub>			
Cy5-Labeled DNA	D1	ATGCTTCCGAGACCGCATCGACCTTGACCT - Cy5			
<u>Target Configurations</u>					
12 $\gamma$ PNAA- 10 DNA	T12	TCGGAAGCATTGAGGTAGTAGG			
11 $\gamma$ PNAA - 11 DNA	T11	CTCGGAAGCATTGAGGTAGTAG			
10 $\gamma$ PNAA - 12 DNA	T10	TCTCGGAAGCATTGAGGTAGTA			
8 $\gamma$ PNAA - 14 DNA	T8	GGTCTCGGAAGCATTGAGGTAG			
6 $\gamma$ PNAA - 16 DNA	T6	GCGGTCTCGGAAGCATTGAGGT			
Theoretical $T_m$ to Target Bases (°C)					
Probe	DNA <sup>a</sup>	LNA <sup>b</sup>	PNA <sup>c</sup>	$\gamma$ PNA <sup>d</sup>	D1 Probe <sup>a</sup>
P1	38.0	58.5	46.1	70.1	
D1					72.4
T12	38.0	58.5	46.1	70.1	36.9
T11	31.7	46.6	41.2	63.2	45.4
T10	26.2	43.6	37.0	57.0	49.5
T8	16.2	28.8	25.2	41.2	51.9
T6	< 10.0	< 10.0	14.3	26.3	55.4
<i>a</i> – DNA-DNA $T_m$ calculated using nearest-neighbor thermodynamics at 89mM NaCl <sup>[38]</sup>					
<i>b</i> – LNA-DNA $T_m$ calculated using nearest-neighbor thermodynamics for LNA at 89mM NaCl <sup>[39]</sup>					
<i>c</i> – PNA-DNA $T_m$ calculated using Giesen <i>et al.</i> empirical model for PNA stability <sup>[37]</sup>					
<i>d</i> – $\gamma$ PNA-DNA $T_m$ calculated by addition of 2°C per substitution to Giesen model as suggested by Sahu <i>et al.</i> <sup>[29]</sup>					



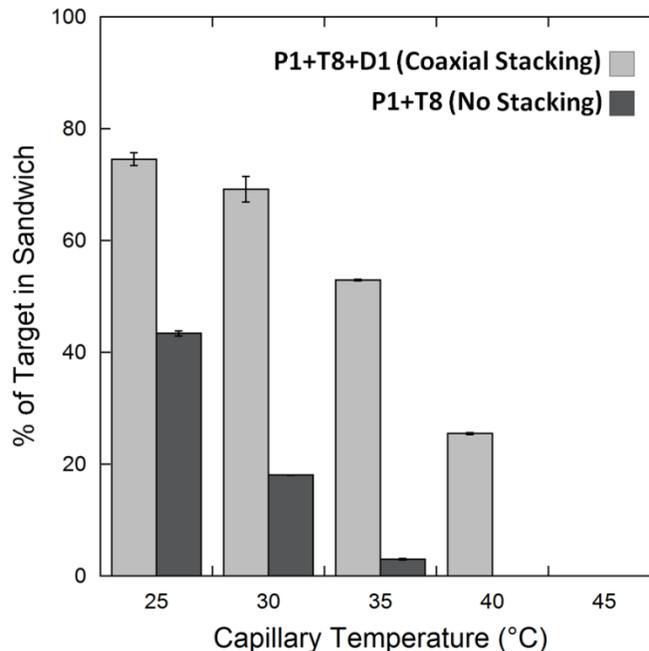
**Figure 2.3** – Micelle CE detection for each component of the sandwich hybridization assay. Run conditions 20kV, 20/30cm capillary, 25°C, 24mM Triton X-100 running buffer in 1xTBE. A) 512nm fluorescence detection of the fluorescein-labeled  $\gamma$ PNAA. B) 675nm fluorescence detection of the Cy5-labeled DNA probe. Curves are labeled with hybridized components using abbreviations according to Table 2.1. Sandwich complex peaks are marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes. RFU = Relative Fluorescence Units

Figure 2.4 shows CE electropherograms for each of the targets in Table 2.1, showing a progressive appearance of the sandwich peak as the number of  $\gamma$ PNAA-overlapping bases is increased. The 6  $\gamma$ PNAA – 16 DNA configuration is the only target where a sandwich complex peak failed to appear. The stability of the  $\gamma$ PNAA to 6 target bases is most likely too low to achieve hybridization at the 25°C detection temperature ( $T_m = 26.3^\circ\text{C}$ ). The 8  $\gamma$ PNAA – 14 DNA configuration, on the other hand, yielded the largest sandwich complex signal. From the stability approximations, one would expect the 10  $\gamma$ PNAA – 12 DNA configuration to have the greatest sandwich complex signal. The difference may stem from the added stability provided by the overhanging  $\gamma$ PNAA bases that was not included in the  $T_m$  models.<sup>[40], [41]</sup> Rather than synthesize a new  $\gamma$ PNAA for each configuration, the hybridization regions were simply shifted down the

$\gamma$ PNAA sequence. Therefore, the 8  $\gamma$ PNAA – 14 DNA configuration contains four *N*-terminal bases of the  $\gamma$ PNAA that are unpaired and extend beyond the 3'-terminus of the target. The helical pre-organization of the  $\gamma$ PNAA residues could provide significant stabilization through end-stacking with the duplex region. The potential added stability allows a shorter  $\gamma$ PNAA to be used, leaving more bases from the target to be recognized by the Cy5-labeled DNA probe, leading to an increase in sandwich complex signal. Because of its higher signal, we continued with further experimentation using the 8  $\gamma$ PNAA – 14 DNA configuration.



**Figure 2.4** – A) 512nm fluorescence detection and B) 675nm fluorescence detection electropherograms of sandwich complexes under different binding configurations. Curves are labeled with hybridized components using abbreviations according to Table 2.1. Separation conditions the same as Figure 2.3. Sandwich complex peaks are marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.



**Figure 2.5** – Added stability from adjacent probe hybridization for sandwich complex formation due to coaxial stacking. Percentage of target bound in sandwich complex (light gray) or by only the  $\gamma$ PNAAs (dark gray). Average of three measurements.

### 2.3.3 – Coaxial Stacking Stabilization

We observed an increase in the percent of target bound in the sandwich versus hybridization solely by the  $\gamma$ PNAAs. Figure 2.5 shows the percentage of the 8  $\gamma$ PNAAs – 14 DNA target bound by either the  $\gamma$ PNAAs alone or in the sandwich at temperatures ranging from 25°C to 45°C. This percentage was determined as the ratio of sandwich complex peak area to that of the unbound  $\gamma$ PNAAs in the 512nm fluorescence detection channel.

An added benefit of sandwich hybridization assays is that the probes may stabilize each other through coaxial stacking when bound to adjacent segments on the target.<sup>[42]–[44]</sup> We observe that this effect gives a higher yield than expected when considering only the  $\gamma$ PNAAs-DNA interaction. The percent of target bound is much larger for the sandwich complex and persists through 40°C, compared to

barely detectable at 35°C for the  $\gamma$ PNA only case. The coaxial stacking effect may permit the use of shorter hybridization segments and greater selectivity against single base mismatch (SBM) off-targets.

Name	Sequence (5'-3')
Perfect (T8)	GGTCTCGGAAGCATTGAGGTAG
MMT1	GG <u>A</u> CTCGGAAGCATTGAGGTAG
MMT2	GGT <u>G</u> TCGGAAGCATTGAGGTAG
MMT3	GGTCTCGGAT <u>G</u> CATTGAGGTAG
MMT4	GGTCTCGGAAGCA <u>A</u> TGAGGTAG
MMT5	GGTCTCGGAAGCAT <u>A</u> GAGGTAG
MMT6	GGTCTCGGAAGCATTG <u>T</u> GGTAG
MMT7	GGTCTCGGAAGCATTGAG <u>C</u> TAG
MMT8	GGTCTCGGAAGCATTGAGG <u>A</u> AG
MMT9	GGTCTCGGAAGCATTGAGGT <u>T</u> G
MMT10	GGTCTCGGAAGCATTGAGGT <u>A</u> C

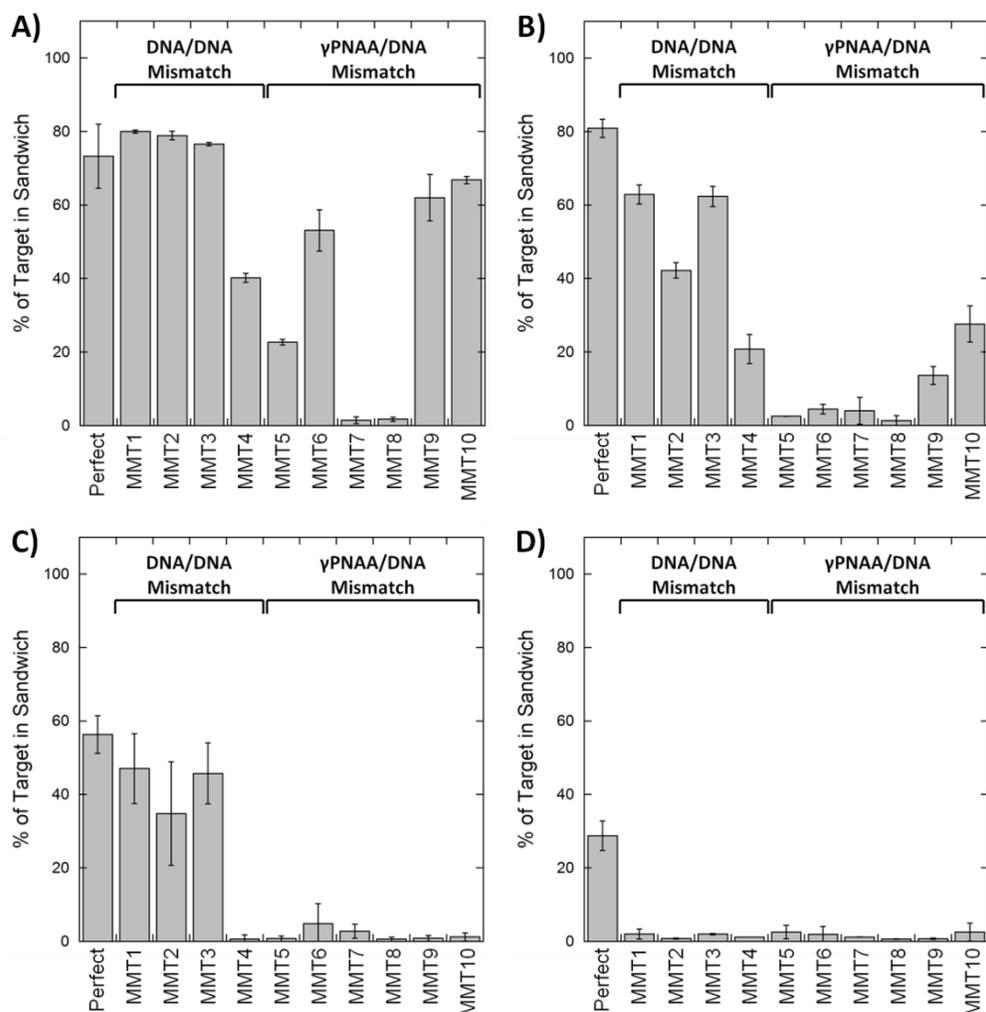
### 2.3.4 – Single Base Mismatch Discrimination

One of the most difficult aspects for direct detection of nucleic acids is single base mismatch (SBM) discrimination between closely related family members. The presented sandwich hybridization assay was tested against 10 different SBM off-targets to determine binding selectivity. The mismatches were located at internal and terminal locations on each probe to test for destabilization effects near probe junctions, as well as target ends. Sequences and names of each mismatch target are listed in Table 2.2. For clarity, the mismatched base on each target sequence is bolded and underlined. Figure 2.6 shows the percent of target bound for each mismatch at temperatures ranging from 25°C to 40°C.

Discrimination against all SBM variants was achieved at 40°C, although the  $\gamma$ PNA achieved complete selectivity at 35°C. The  $\gamma$ PNA hybridizes to 8 bases of the target versus 14 bases recognized by the Cy5-labeled DNA probe. A single

mismatch with the  $\gamma$ PNAA's recognition site should create a larger  $\Delta\Delta G$  upon hybridization as mismatch discrimination is strongly dependent on length for short oligomers.<sup>[45]</sup> The helical preorganization from the  $\gamma$ -carbon substitution might also cause the  $\gamma$ PNAA base opposite a mismatch to remain within the helix. The spatially confined base may create a greater degree of steric hindrance for hybridization compared with a more flexible DNA backbone. This steric hindrance may also explain why mismatches on the  $\gamma$ PNAA positioned closer to the junction with the Cy5-labeled DNA probe strand (MMT5, MMT6, MMT7, and MMT8) are generally more destabilizing than the distal mismatches (MMT9 and MMT10). The conformational freedom of the distal bases due to breathing of the  $\gamma$ PNAA-DNA duplex may be more accommodating than the internal bases to the  $\gamma$ PNAA.

As for the Cy5-labeled probe, only the mismatch at the probe junction (MMT4) is destabilizing at 35°C while the others require 40°C. A single base gap between DNA hybridization probes can result in a destabilizing effect on duplex stability due to reduced coaxial stacking.<sup>[42]</sup> The reduced coaxial stacking combined with the mismatched base may create a higher level of selectivity at the probe junction versus other mismatches. The stability approximations in Table 2.1 also predict a higher  $T_m$  for the Cy5-labeled DNA probe versus the  $\gamma$ PNAA, 51.9°C and 41.2°C, respectively. The lower predicted stability of the  $\gamma$ PNAA may grant it greater selectivity at the 35°C than the Cy5-labeled DNA probe. Through judicious choice of the number of bound bases per probe, these differences can be resolved to normalize stabilities for enhanced selectivity towards SBM off-targets.



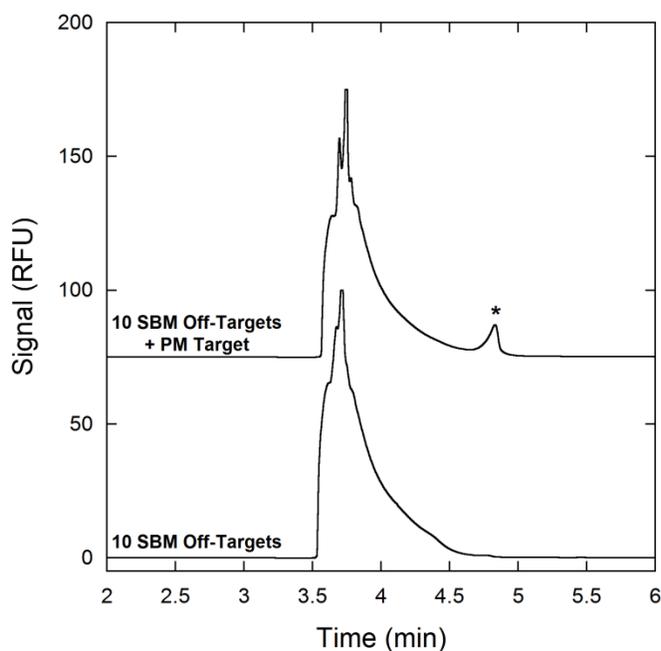
**Figure 2.6** – Percent of target bound by both probes at different capillary temperatures for SBM off-targets. Sequences of each target are seen in Table 2.2. A) 25°C, B) 30°C, C) 35°C, D) 40°C. Average of three measurements.

The selectivity of the probes was also conserved within a mixture of SBM off-targets, Figure 2.7. A solution containing all 10 SBM off-targets failed to yield a sandwich hybridization signal in the presence of excess probes. Only when the perfect match target was spiked into the solution was a sandwich hybridization peak detected. The perfectly match target was also introduced at a 10-fold lower concentration than that of the total off-target concentration, demonstrating the ability of the assay to bind desired targets in the presence of a high concentration

of contaminant sequences. With these SBM results, the sandwich complexes can be easily designed to be highly selective between closely related targets.

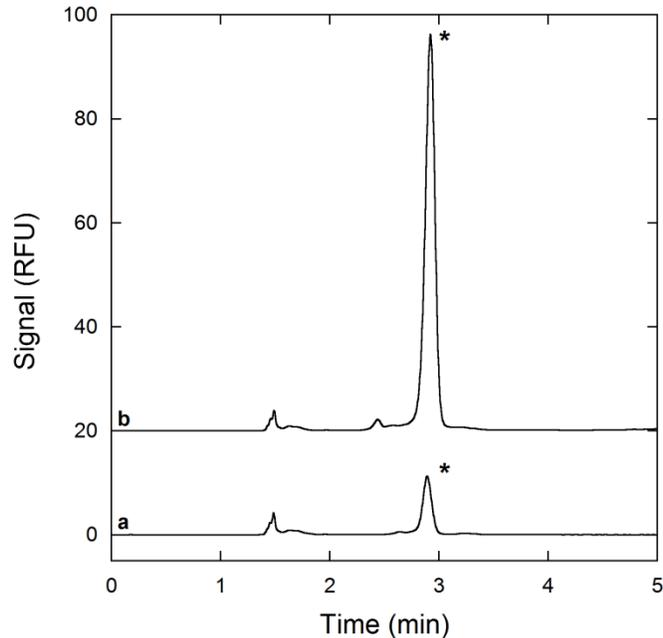
### 2.3.5 – Application Towards miRNA Detection

miRNAs are an attractive biomarker for early stage disease diagnosis due to their highly correlated expression levels with disease onset.<sup>[11], [12]</sup> However, detection of miRNA is complicated by four major challenges; 1) stable hybridization to the 19-23nt sequences, 2) SBM selectivity between/within miRNA families, 3) high-throughput detection for rapid analysis of the thousands of known miRNAs, and 4) sensitive detection limits for extremely low abundance miRNA targets.



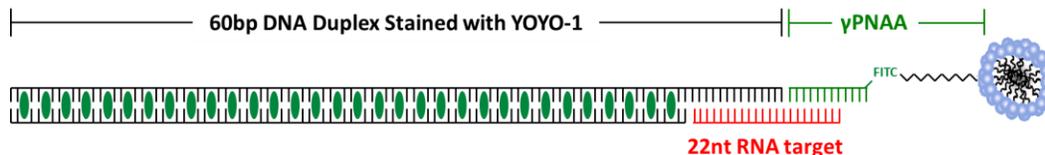
**Figure 2.7** – Detection of perfectly matched target from pool of single base mismatch off-targets. Run conditions 20kV, 40/50cm capillary, 50 $\mu$ m, 40°C, 24mM Triton X-100 running buffer in 1xTBE. Sandwich peak is marked by an asterisk. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.

As a preliminary feasibility test, we used the CE-sandwich method to detect a 22nt RNA target using the highly fluorescent intercalating dye DNA tags described above. Figure 2.8 is a CE electropherogram for the separation of a sandwich complex formed with a 66 base-pair DNA probe intercalated with YOYO-1 and a 22nt RNA target. Figure 2.9 shows a cartoon representation of the RNA-YOYO-1 DNA sandwich complex. For comparison, we include an electropherogram detected using only the conventional fluorescein dye located on the  $\gamma$ PNA. DNA and RNA have very similar hybridization and electrophoretic properties, and therefore the RNA target behaves comparably to the previously demonstrated DNA targets. However, the elution time of the RNA target sandwich complex peak has shifted to 2.9 minutes, compared to 1.6 minutes for the DNA targets seen in Figure 2.3. The longer DNA probe used for detection of the RNA target delays elution time due to the higher electrophoretic mobility of the sandwich in the presence of a dominant EOF. The ability to tune elution time based on DNA probe length is currently being investigated for multiplexed detection applications.



**Figure 2.8** – 512nm fluorescence detection of the 22nt RNA target (5' CGA UCA CAU GGC AUU GAG GUA G 3') by sandwich complex formation between the  $\gamma$ PNAA and a 66 base-pair duplex DNA probe (sequences in experimental section). a) Detection of fluorescein-labeled  $\gamma$ PNAA only, b) Detection of the fluorescein-labeled  $\gamma$ PNAA and YOYO-1 stained DNA probe. Sandwich complex peaks are marked by an asterisk. Same run conditions as Figure 2.3. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.

Figure 2.8 shows the detection of the RNA target sandwich complex stained with YOYO-1. The fluorescent signal of the YOYO-1 stained peak is 7 times higher than that of the fluorescein-labeled  $\gamma$ PNAA alone due to the addition of approximately 13 YOYO-1 dye molecules loaded into the 66 base-pair DNA probe. The incorporation of multiple fluorophores per sandwich binding event offers a straightforward way to dramatically increase the detection sensitivity of the system for direct miRNA detection. By simply extending the length of the DNA probe duplex, the DNA probe can be loaded with 100s to 1,000s of dye molecules for sensitive detection of low abundance miRNA targets.

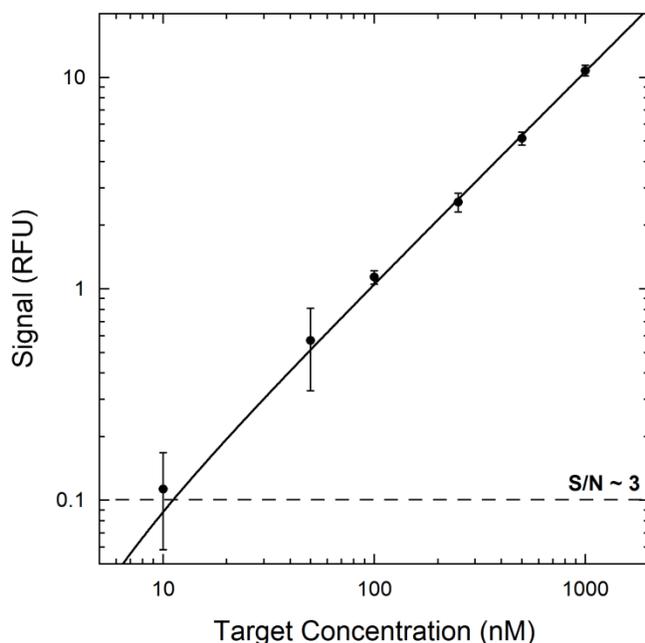


**Figure 2.9** – Cartoon representation of  $\gamma$ PNAA, 22nt RNA target, and 66 base-pair YOYO-1 stained DNA duplex sandwich complex.

The presented sandwich hybridization assay offers many advantages towards implementation as a miRNA detection method. We have demonstrated stable hybridization with SBM sensitivity along both hybridization probes. Sandwich hybridized targets are also detected in as little as 3.5 minutes for rapid analysis of miRNA targets. Additionally, target hybridization and detection occurs in free solution compared to commonly implemented surface based approaches.<sup>[46]–[48]</sup> Surface based hybridization often requires overnight incubations due to high probe surface densities that reduce target diffusion to surface immobilized probes.<sup>[47], [49]</sup> Solution based hybridization can be accomplished in as little as 15 minutes with many common annealing protocols, dramatically reducing the collection to detection analysis time. Furthermore, the free solution nature of the detection mechanism offers rapid turnaround for consecutive detection. By simply rinsing the capillary with fresh micelle-containing running buffer after each run, different samples can be analyzed sequentially for high-throughput analysis.

The remaining challenge that has not been addressed is the detection sensitivity of the system towards low abundance targets. The detection sensitivity of the system was determined as 10nM for the Cy5-labeled DNA probe via titration experiments, as seen in Figure 2.10. CE separations offer the advantage of  $\mu$ L sample requirements compared to mLs for cuvette and surface based techniques. The small sample volumes mean the 10nM limit of detection (LOD) correlates to

a sensitivity of  $2 \times 10^{11}$  targets in the  $30 \mu\text{L}$  sample vial. Rare miRNAs may have expression levels down to tens of copies per cell and small cell biopsies may yield as little as 1,000 cancerous cells for analysis.<sup>[22]</sup> Therefore, the theoretically ideal LOD for direct detection of low abundance miRNA targets would be on the order of  $10^4$  targets, 7 magnitudes lower than our current limits.



**Figure 2.10** – Observed fluorescent signal of sandwich complex at different target concentrations. Constant probe concentration of  $1 \mu\text{M}$ . Solid line represents linear fit with  $R^2 = 0.999$ , averaged over three measurements.

In addition to use of the highly fluorescent YOYO-1 DNA probes, an electrokinetic or micelle focusing step may be implemented to address the theoretically required LOD of  $10^4$  targets in the proposed system. These focusing steps can concentrate the low abundance miRNAs prior to detection, where thousand to million fold increases in concentration have been realized for nucleic acids.<sup>[50]–[53]</sup> The combination of highly fluorescent YOYO-1 DNA probes and focusing may dramatically decrease the LOD to a level more appropriate for

direct detection of low abundance miRNA targets. Custom CE modules using high objective microscope lens and single photon counting detectors have also been built in order to achieve LODs from 1 to 1,000 targets.<sup>[25], [54]</sup> Although these custom modules offer the advantage of sensitive detection, commercially available systems, such as the CE unit used in this study, are more favorable when considering feasible clinical implementation.

## **2.4 – Conclusion**

We have demonstrated the stability and selectivity of a sandwich hybridization assay for rapid detection of short 22 nucleotide DNA and RNA targets. Stable sandwich hybridization is dependent on the high affinity of the recently developed PEG substituted  $\gamma$ PNA. The high affinity  $\gamma$ PNA enables recognition of a relatively short 8 base segment of the target, allowing stable hybridization of a fluorescently labeled DNA reporter probe to the target's remaining 14 bases. A mobility shift assay was utilized to confirm sandwich hybridization, where bound and unbound components are separated in 3.5 minutes using a micelle-containing running buffer in CE.

Sandwich hybridization was found to be stable under multiple binding configurations and stabilized by coaxial stacking upon hybridization of the adjacent probe. The assay is also capable of distinguishing single base mismatches at many locations along the target by simply increasing the capillary temperature to 40°C. We believe the assay offers many advantages towards implementation as a miRNA detection method with its high stability and selectivity towards short targets. Upon future investigation for increasing

detection sensitivity, it should be straightforward to adapt for high-throughput profiling of miRNA expression levels or any other biologically relevant, nucleic acid.

## 2.5 – References

- [1] X. Zhao, R. Tapeç-Dytioco, and W. Tan, *J. Am. Chem. Soc.*, 2003, 125, 11474–11475.
- [2] L. Chen, S. Lee, M. Lee, C. Lim, J. Choo, J. Y. Park, S. Lee, S. W. Joo, K. H. Lee, and Y.-W. Choi, *Biosens. Bioelectron.*, 2008, 23, 1878–1882.
- [3] S. Li, H. Liu, L. Liu, L. Tian, and N. He, *Anal. Biochem.*, 2010, 405, 141–143.
- [4] C. Y. Zhang and J. Hu, *Anal. Chem.*, 2010, 82, 1921–1927.
- [5] M. Dequaire and A. Heller, *Anal. Chem.*, 2002, 74, 4370–4377.
- [6] E. E. Ferapontova, M. N. Hansen, A. M. Saunders, S. Shipovskov, D. S. Sutherland, and K. V. Gothelf, *Chem. Commun. (Cambridge, United Kingdom)*, 2010, 46, 1836–1838.
- [7] J. Wang, G. Liu, and A. Merkoçi, *J. Am. Chem. Soc.*, 2003, 125, 3214–3215.
- [8] R. M. Franzini and E. T. Kool, *Chem. Eur. J.*, 2011, 17, 2168–2175.
- [9] D. J. Kleinbaum and E. T. Kool, *Chem. Commun. (Cambridge, United Kingdom)*, 2010, 46, 8154–8156.
- [10] Y. C. Cao, R. Jin, and C. A. Mirkin, *Science*, 2002, 297, 1536–1540.
- [11] G. A. Calin and C. M. Croce, *Nat. Rev. Cancer*, 2006, 6, 857–866.
- [12] A. Esquela-Kerscher and F. J. Slack, *Nat. Rev. Cancer*, 2006, 6, 259–269.
- [13] K. Cissell and S. Deo, *Anal. Bioanal. Chem.*, 2009, 394, 1109–1116.
- [14] A. W. Wark, H. J. Lee, and R. M. Corn, *Angew. Chemie Int. Ed.*, 2008, 47, 644–652.
- [15] J. Chen, J. Lozach, E. W. Garcia, B. Barnes, S. Luo, I. Mikoulitch, L. Zhou, G. Schroth, and J. B. Fan, *Nucleic Acids Res.*, 2008, 36, e87.
- [16] R. Shi and V. L. Chiang, *Biotechniques*, 2005, 39, 519–525.

- [17] J. Shingara, K. Keiger, J. Shelton, W. Laosinchai-Wolf, P. Powers, R. Conrad, D. Brown, and E. Labourier, *RNA*, 2005, 11, 1461–1470.
- [18] E. M. Harcourt and E. T. Kool, *Nucleic Acids Res.*, 2012, 40, e65.
- [19] Y. Cheng, X. Zhang, Z. Li, X. Jiao, Y. Wang, and Y. Zhang, *Angew. Chemie - Int. Ed.*, 2009, 48, 3268–3272.
- [20] S. C. Chapin and P. S. Doyle, *Anal. Chem.*, 2011, 83, 7179–7185.
- [21] J. Li, S. Schachermeyer, Y. Wang, Y. Yin, and W. Zhong, *Anal. Chem.*, 2009, 81, 9723–9729.
- [22] C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak, and K. J. Guegler, *Nucleic Acids Res.*, 2005, 33, e179.
- [23] Y. Zhang and C. Zhang, *Anal. Chem.*, 2012, 84, 224–231.
- [24] S. Cai, C. Lau, and J. Lu, *Anal. Chem.*, 2010, 82, 7178–7184.
- [25] L. A. Neely, S. Patel, J. Garver, M. Gallo, M. Hackett, S. McLaughlin, M. Nadel, J. Harris, S. Gullans, and J. Rooke, *Nat Meth*, 2006, 3, 41–46.
- [26] A. A. Koshkin, P. Nielsen, M. Meldgaard, V. K. Rajwanshi, S. K. Singh, and J. Wengel, *J. Am. Chem. Soc.*, 1998, 120, 13252–13253.
- [27] N. Tolstrup, P. S. Nielsen, J. G. Kolberg, A. M. Frankel, H. Vissing, and S. Kauppinen, *Nucleic Acids Res.*, 2003, 31, 3758–3762.
- [28] R. Owczarzy, Y. You, C. L. Groth, and A. V. Tataurov, *Biochemistry*, 2011, 50, 9352–9367.
- [29] B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage, and D. H. Ly, *J. Org. Chem.*, 2011, 76, 5614–5627.
- [30] S. Sforza, G. Haaima, R. Marchelli, and P. E. Nielsen, *European J. Org. Chem.*, 1999, 1999, 197–204.
- [31] J. P. Vernille, L. C. Kovell, and J. W. Schneider, *Bioconjug. Chem.*, 2004, 15, 1314–1321.
- [32] A. L. Benven, Y. Creeger, G. W. Fisher, B. Ballou, A. S. Waggoner, and B. A. Armitage, *J. Am. Chem. Soc.*, 2007, 129, 2025–2034.
- [33] H. Özhacı-Ünal and B. Armitage, *ACS Nano*, 2009, 3, 425–433.
- [34] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.

- [35] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [36] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.
- [37] U. Giesen, W. Kleider, C. Berding, A. Geiger, H. Ørum, and P. E. Nielsen, *Nucleic Acids Res.*, 1998, 26, 5004–5006.
- [38] J. SantaLucia Jr., *Proc. Natl. Acad. Sci. U. S. A.*, 1998, 95, 1460–1465.
- [39] P. M. McTigue, R. J. Peterson, and J. D. Kahn, *Biochemistry*, 2004, 43, 5388–5405.
- [40] S. Bommarito, N. Peyret, and J. SantaLucia Jr., *Nucleic Acids Res.*, 2000, 28, 1929–1934.
- [41] B. Datta and B. A. Armitage, *J. Am. Chem. Soc.*, 2001, 123, 9612–9619.
- [42] V. A. Vasiliskov, D. V. Prokopenko, and A. D. Mirzabekov, *Nucleic Acids Res.*, 2001, 29, 2303–2313.
- [43] M. Petersheim and D. H. Turner, *Biochemistry*, 1983, 22, 256–263.
- [44] P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii, *Nucleic Acids Res.*, 2006, 34, 564–574.
- [45] Y. You, B. G. Moreira, M. A. Behlke, and R. Owczarzy, *Nucleic Acids Res.*, 2006, 34, e60.
- [46] J. Shingara, K. Keiger, J. Shelton, W. Laosinchai-Wolf, P. Powers, R. Conrad, D. Brown, and E. Labourier, *RNA*, 2005, 11, 1461–1470.
- [47] C. G. Liu, G. A. Calin, B. Meloon, N. Gamliel, C. Sevigani, M. Ferracin, C. D. Dumitru, M. Shimizu, S. Zupo, M. Dono, H. Alder, F. Bullrich, M. Negrini, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 9740–9744.
- [48] M. Castoldi, S. Schmidt, V. Benes, M. Noerholm, A. E. Kulozik, M. W. Hentze, and M. U. Muckenthaler, *RNA*, 2006, 12, 913–920.
- [49] R. Levicky and A. Horgan, *Trends Biotechnol.*, 2005, 23, 143–149.
- [50] R. L. Chien, *Electrophoresis*, 2003, 24, 486–497.
- [51] B. Jung, R. Bharadwaj, and J. G. Santiago, *Anal. Chem.*, 2006, 78, 2319–2327.
- [52] J. P. Quirino and S. Terabe, *Science*, 1998, 282, 465–468.

- [53] A. Aranas, A. Guidote Jr., and J. Quirino, *Anal. Bioanal. Chem.*, 2009, 394, 175–185.
- [54] B. J. Dodgson, A. Mazouchi, D. W. Wegman, C. C. Gradinaru, and S. N. Krylov, *Anal. Chem.*, 2012, 84, 5470–5474.

# **Chapter 3 – Practical Improvements to MicroRNA $\gamma$ PNA Sandwich Hybridization Assay: Multiplexing, PicoGreen Intercalating Dye, and Protein Contaminant Tolerance**

## **3.1 – Introduction**

MicroRNAs (miRNAs) are a class of ~22 nucleotide, non-protein coding RNAs that play an important role in post-transcriptional gene regulation. Abnormal expression of miRNAs has been associated with many diseases, including most cancers, cognitive disorders, and cardiovascular disease.<sup>[1]-[4]</sup> To date, there are 2,588 mature human miRNAs listed in the miRNA database miRBase 21 (<http://www.mirbase.org/>, July 2014) and thousands of reported miRNA-disease associations, making them a promising class of diagnostic biomarkers.<sup>[5]-[8]</sup> Additionally, identification of tissue and condition specific diseases is typically associated with a unique miRNA signature composed of 2-20 different miRNAs.<sup>[9]-[12]</sup> Therefore, the ability to detect multiple miRNA expression levels in a single, high-throughput assay is crucial towards implementing miRNA as biomarkers.

The detection of miRNA is rather challenging because of their short length, low abundance, and high sequence homology between miRNA family members. Any

viable diagnostic method should be able to overcome these difficulties by detecting multiple miRNAs with low limits of detection (LOD), high specificity, and rapid throughput. The most common miRNA detection methods include northern blotting<sup>[13], [14]</sup>, quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)<sup>[15], [16]</sup>, solid-phase microarrays,<sup>[17], [18]</sup> and next-generation sequencing<sup>[19], [20]</sup>. Although these methods have been crucial in identification of miRNA-disease associations, they typically require extensive sample preparation (miRNA isolation and enzymatic amplification/labeling) that limits their application towards quantifiable diagnostics of miRNA expression levels. These indirect techniques not only introduce complexity, time, and cost to the assay, but also introduce potential enzyme sequence-specific biases that reduce reliability.<sup>[21]–[23]</sup>

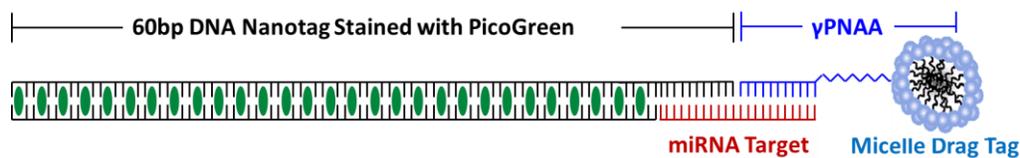
An emerging class of direct (enzyme-free) miRNA detection assays have been established, including bioluminescence<sup>[24]</sup>, atomic force microscopy (AFM)<sup>[25]</sup>, surface plasmon resonance<sup>[26]</sup>, electro-chemical biosensors<sup>[27], [28]</sup>, electro-mechanical devices<sup>[29]</sup>, molecular beacons<sup>[30], [31]</sup>, isotachopheresis<sup>[32], [33]</sup>, two-color fluorescence coincident detection<sup>[34]</sup>, and capillary electrophoresis (CE)<sup>[35]–[39]</sup>. Many of these methods demonstrate exceptional detection sensitivity in the aM-fM range, but are only capable of detecting one miRNA at a time. Only the CE methods have demonstrated promise for multiplexed miRNA detection, albeit with less than desirable nM-pM sensitivity. Additionally, many clinical labs already possess CE instruments, which should make adaptation of new miRNA detection methods by CE much more amenable for new users. New CE methods

that are capable of fM detection limits coupled with high-throughput, multiplexed detection stand to make a dramatic impact on the application of miRNA as disease biomarkers.

We have previously demonstrated detection of single 22nt DNA or RNA targets in a sandwich hybridization format.<sup>[40]</sup> Stable hybridization of two probes to the short targets is achieved using a high affinity *n*-alkylated polyethylene glycol  $\gamma$ -carbon modified peptide nucleic acid amphiphile ( $\gamma$ PNAA). The  $\gamma$ PNAA's exceptionally high affinity enables stable hybridization of a second, fluorescently labeled DNA-based probe to the remaining bases of the short target. Upon hybridization of both probes, an electrophoretic mobility shift is measured via interaction of the *n*-alkane modification on the  $\gamma$ PNA with capillary electrophoresis running buffer containing non-ionic surfactant micelles. Figure 3.1 shows a cartoon representation of our miRNA sandwich hybridization complex.

Here, we demonstrate the extension of this method towards multiplexing detection of 6 *let-7* miRNA family members in a single 4 minute separation. Unlike previous CE methods using multiple drag-inducing polymers to achieve multiplexed miRNA detection, we achieve differences in elution time by simply extending the length of the DNA probe hybridized to the target.<sup>[35], [41]–[43]</sup> We present a theoretical model that predicts our method can achieve multiplexed detection of 26 different miRNA targets in the same 4 minute separation with access to DNA hybridization probes up to 1,000 base pairs (bps) long. Additionally, the DNA probe is saturated with fluorescent intercalating dye

molecules (approximately one per four base pairs of DNA duplex).<sup>[40], [44]–[46]</sup> These DNA “nanotags” increase detection sensitivity by condensing many fluorophores onto a single miRNA binding event. We report an LOD of 100pM using a 70 base pair DNA nanotag stained with fluorescent dye, a 100-fold improvement to our single Cy5 fluorophore system. The fluorescent signal of the sandwich complex is proportional to the length of the DNA nanotag, since longer probes can accommodate more dye. This offers a simple method to achieve fM detection limits by extending the length of DNA nanotag hybridized to the miRNA targets. We also demonstrate excellent single base mismatch discrimination between the closely related *let-7* miRNAs, and the ability to detect miRNA in a high concentration of contaminating proteins from goat serum.



**Figure 3.1** – Cartoon representation of miRNA sandwich hybridization complex.

## 3.2 – Materials and Methods

### 3.2.1 – Reagents

Nuclease-free water was purchased from Fisher Scientific (Pittsburgh, PA). Triton X-100 surfactant and 10x Tris-Borate-EDTA (TBE) buffer were purchased from Sigma Aldrich (St. Louis, MO). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2mΩ deionized water, autoclaving, and filtering through a 0.22μm syringe filter.

PicoGreen dye, BCA Protein Assay Kit, goat serum, and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). DNA and RNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA), with standard desalting purification except for oligomers longer than 60 bases, which were purchased as HPLC purified. Sequences are seen in Table 3.1 and were designed to avoid undesired non-target complementarity between sandwich components. The “capture” DNA probes hybridize directly to the miRNA target. The “complement” probe binds the capture probe to create a double stranded DNA nanotag for intercalating dye staining. Dried oligomers were resuspended in nuclease-free water and used without further purification.

### 3.2.2 – $\gamma$ PNA Amphiphile ( $\gamma$ PNAA) Synthesis and Purification

$\gamma$ PNAAAs were manually synthesized as described previously, where the *N*-terminus was labeled with a 3 unit mini-polyethylene glycol (MP<sub>3</sub>) spacer and octadecanoic acid.<sup>[40], [47]</sup> Because of the low solubility of octadecanoic acid in most organic solvents including DMF, which was employed in the coupling step, the reaction was performed three times at 2-hr each.  $\gamma$ PNAAAs were purified using reverse-phase high performance liquid chromatography (HPLC) and characterized on a PerSeptive Voyager STR MALDI-TOF mass spectrometer. The following masses were obtained: *let7-a,e, f*  $\gamma$ PNAA C18–MP<sub>3</sub>-AACTATACA-Lys-NH<sub>2</sub>, (4323.1 obs., 4320.7 theo.); *let7-b*  $\gamma$ PNAA C18–MP<sub>3</sub>-AACCACAC-Lys-NH<sub>2</sub>, (N/A obs., 3897.5 theo.); *let7-c*  $\gamma$ PNAA C18–MP<sub>3</sub>-AACCATAC-Lys-NH<sub>2</sub>, (3915.3 obs., 3912.5 theo.); *let7-d*  $\gamma$ PNAA C18–MP<sub>3</sub>-AACTATGC-K-NH<sub>2</sub>, (N/A obs., 3943.5 theo.).

<b>Table 3.1 – List of miRNA Sequences, <math>\gamma</math>PNAAs, and DNA probes</b>	
<b><i>let-7</i> miRNAs</b>	<b>Sequence (5' to 3')</b>
<i>let-7a</i>	UGAGGUAGUAGGUUGUAUAGUU
<i>let-7b</i>	UGAGGUAGUAGGUUGUGUGGUU
<i>let-7c</i>	UGAGGUAGUAGGUUGUAUGGUU
<i>let-7d</i>	AGAGGUAGUAGGUUGCAUAGUU
<i>let-7e</i>	UGAGGUAGGAGGUUGUAUAGUU
<i>let-7f</i>	UGAGGUAGUAGAUUGUAUAGUU
<b><math>\gamma</math>PNAAs</b>	<b>Sequence (N to C terminus)</b>
<i>let7-a,e, f</i> $\gamma$ PNAA	C18–MP <sub>3</sub> -AACTATACA-Lys-NH <sub>2</sub>
<i>let7-b</i> $\gamma$ PNAA	C18–MP <sub>3</sub> -AACCCACAC-Lys-NH <sub>2</sub>
<i>let7-c</i> $\gamma$ PNAA	C18–MP <sub>3</sub> -AACCATAC-Lys-NH <sub>2</sub>
<i>let7-d</i> $\gamma$ PNAA	C18–MP <sub>3</sub> -AACTATGC-Lys-NH <sub>2</sub>
<b>DNA Probes</b>	<b>Sequence (5' to 3')</b>
<i>let-7a</i> Capture	ACCTACTACCTCACAGATACGCTTAGACCTAGTAACG GCTTGCCAAGTGACCTTACGTGACTACTTCAAGCTC
<i>let-7a</i> Complement	GAGCTTGAAGTAGTCACGTAAGGTCACCTTGCCAAGC CGTACTAGGTCTAAGCGTATCTG
<i>let-7b</i> Capture	AACCTACTACCTACTAACGAGGAAAGCACGTTATA CGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTAG CGGCGCATTAA
<i>let-7b</i> Complement	TTAATGCGCCGCTACAGGGCGGCTACTATGGTTGCTT TGACGAGCACGTATAACGTGCTTTCCTCGTTAG
<i>let-7c</i> Capture	AACCTACTACCTCAGCATCGACCTTGACCTTAGCCGA GACTTAACCGTCTAGCT
<i>let-7c</i> Complement	AGCTAGACGGTTAAGTCTCGGCTAAGGTCAAGGTGCG ATGC
<i>let-7d</i> Capture	AACCTACTACCTCTCGGAAAGCTGGCTGGAGTGCGAT CTTCCTGAGGCCGATACTGTTCGTCGT C
<i>let-7d</i> Complement	GACGACGACAGTATCGGCCTCAGGAAGATCGCACTC CAGCCAGCTTTCGG
<i>let-7e</i> Capture	ACCTCCTACCTCACGCCAGGGTTGTCCCAGTCA
<i>let-7e</i> Complement	TGACTGGGACAACCCTGGCG
<i>let-7f</i> Capture	ATCTACTACCTCAGTATAATGAGCCAGTTCTTAAAT CGCATA
<i>let-7f</i> Complement	TATGCGATTTTAAGAAGTGGCTCATTATAC

### **3.2.3 – Sandwich Hybridization Conditions**

Sandwich complex hybridization was carried out by mixing 25nM of each *let-7* miRNA target with 250nM of their respective  $\gamma$ PNA and DNA probes in 1xTBE, unless otherwise noted. Samples with multiple miRNA binding complexes were mixed in a single solution. Solutions were heated to 95°C for at least 5 minutes in a dry bath incubator and allowed to cool to room temperature over the course of 60 minutes. Hybridized complexes not immediately used for separations were stored at -20°C and subjected to minimal freeze-thaws.

### **3.2.4 – Capillary Electrophoresis**

Capillary electrophoresis (CE) experiments were performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 50 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and a length of 20cm to the detector and 30cm total. Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer 1xTBE solution and maintained by adding 0.1v/v% POP-6 to all CE buffers. After EOF suppression, the capillary could be used for many sequential separations without observing any detrimental effects.

A 1/400<sup>th</sup> dilution of stock PicoGreen was added to the running buffer to ensure the DNA was stained during electrophoresis. We have found that the addition of PicoGreen dye does not greatly affect the electrophoretic migration of DNA. Before each separation, the capillary was rinsed with fresh 24mM TX-100 1xTBE separation buffer. Hydrodynamic injections (0.5psi for 5sec, ~0.39cm, ~8nL)

were used to introduce samples into the capillary that were incubated for at least 5 minutes with a 1/400<sup>th</sup> dilution of stock PicoGreen. Electrophoretic separation was conducted under reverse polarity (from anode to cathode) with an applied voltage of 20kV. The capillary was heated to the indicated temperatures using a built in coolant jacket system. Data collection was performed using 32 Karat software (Beckman Coulter). Sandwich complex elution peaks were fit using OriginPro 9.0 Peak Analyzer software (Northampton, MA).

### **3.2.5 – Goat Serum Spike-in Experiments**

The concentration of total protein in goat serum was determined following the BCA protein assay protocol. We found a linear protein concentration response from 0.5 – 2.0 mg/mL for 1250 – 4,500x dilutions of the goat serum stock, indicating an approximate 2 g/mL total protein concentration in the goat serum stock. Spike-in samples were prepared as described above by adding the indicated concentration of goat serum to the samples prior to sandwich complex hybridization. Samples were then separated as described above.

## **3.3 – Results and Discussion**

### **3.3.1 – Multiplexed *let-7* miRNA Detection**

The sandwich hybridization approach creates a separation mechanism where unbound probes are inherently shifted from bound miRNAs. In free-solution, the mobility of DNA is high and independent of molecular weight, and that of the uncharged  $\gamma$ PNAA is negligible.<sup>[48]</sup> When both probes bind to a miRNA target, the complex will have a slightly lower mobility due to the uncharged  $\gamma$ PNAA.

The mobility can be drastically shifted by including non-ionic Triton X-100 surfactant micelles that interact with the C18 alkane group on the  $\gamma$ PNAA. The micelles will transiently attach to the alkyl group, acting as drag-tags that enable length based CE separation of DNA.<sup>[40], [49], [50]</sup>

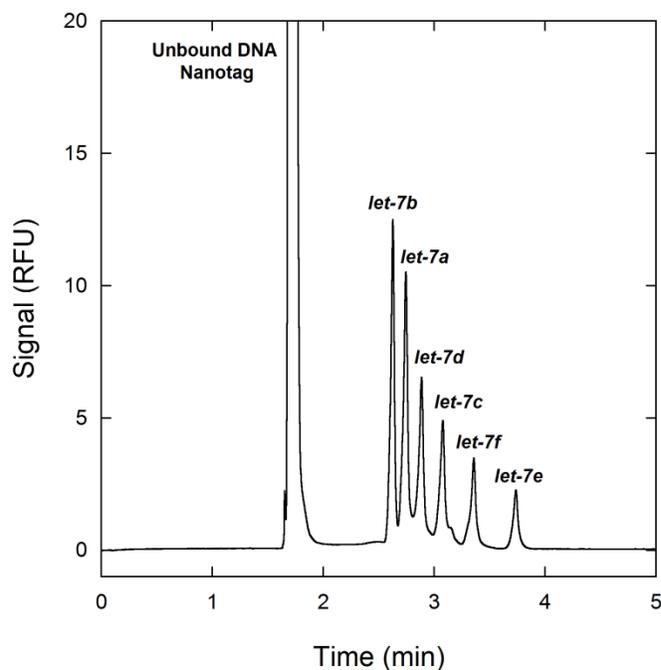
The elution time ( $t_{elute}$ ) of the sandwich complex is predicted by Equation 3.1, where  $l_d$  is the capillary length to the detector,  $l_t$  is the total capillary length,  $V$  is the applied potential,  $\mu_0$  and  $t_0$  are DNA's length invariant free-solution electrophoretic mobility and elution time, respectively,  $L$  is the total length of the DNA nanotag miRNA sandwich complex, and  $\alpha$  is the additional drag of the non-ionic micelle in terms of hydrodynamically equivalent uncharged DNA bases.

$$t_{elute} = \frac{l_d l_t (L + \alpha)}{\mu_0 V L} = t_0 \frac{(L + \alpha)}{L} \quad (3.1)$$

Previous attempts to detect multiple miRNAs using a similar CE approach varied the size of a polymeric drag-tag attached to the DNA hybridization probes, changing  $\alpha$  for each miRNA target.<sup>[35], [37]</sup> Predictions of  $\alpha$  are non-trivial without detailed knowledge of the polymer's chain statistics, such as persistence length, monomer length, and conformation.<sup>[43], [51]</sup> Polymer synthesis, purification, and attachment to DNA probes also adds complex steps to the relatively simple CE based detection approach.

Our sandwich hybridization assay allows simple prediction and separation of multiple miRNA complexes by changing the length of the DNA nanotag hybridized to the miRNA target. Longer lengths of DNA are less hindered by the

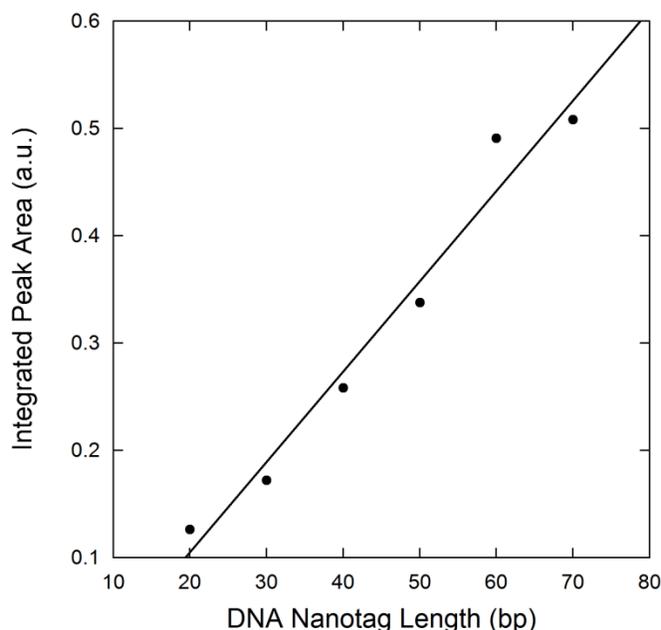
additional drag of the TX-100 micelle attached to the complex, and therefore migrate with a higher electrophoretic velocity. Since interaction with the micelle is achieved through the alkane modification on the  $\gamma$ PNAA, DNA nanotags that are not bound are inherently shifted away from miRNA complex detection. Conversely, when the miRNA is only bound by the  $\gamma$ PNAA, it has no significant DNA length to contribute charge to the complex, and migrates much slower than sandwich complex bound miRNAs.



**Figure 3.2** – Separation of 6 *let-7* miRNA sandwich complexes by CE with end-attached non-ionic surfactant micelles,  $\gamma$ PNAA, and PicoGreen stained DNA nanotags.

We designed DNA nanotags of lengths 20, 30, 40, 50, 60, and 70 bases long corresponding to *let-7* miRNAs *e*, *f*, *c*, *d*, *a*, and *b*, respectively. Figure 3.2 shows the rapid separation of the 6 *let-7* miRNA sandwich complexes in less than 4 minutes. The longest miRNA sandwich complex (*let-7b* = 70 bases) elutes first, followed by sandwich complexes with decreasing DNA nanotag length. It is

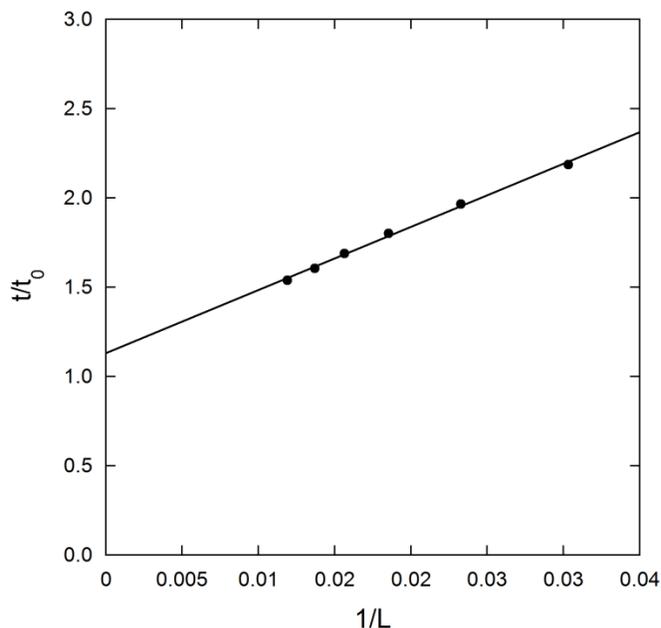
important to note that all 6 complexes were annealed in a one-pot batch, where individual *let-7* targets, DNA nanotags, and  $\gamma$ PNAAs were mixed and annealed in the presence of one another. All sandwich complexes are highly resolved such that integrated area from the peaks can be accurately calculated enabling target concentration quantification.



**Figure 3.3** – Integrated peak areas for the 6 *let-7* miRNAs complexes separated in Figure 3.2. DNA nanotag lengths of 20, 30, 40, 50, 60, and 70 bases corresponding to *let-7* miRNAs *e, f, c, d, a, and b*, respectively.

Figure 3.2 shows a key feature of our sandwich complex design. The complex peaks decrease in signal with increasing elution time, not due to a broadening effect, but rather the decrease in DNA nanotag length. The shorter DNA nanotag incorporates a lesser amount of PicoGreen dye, and therefore has lower fluorescence intensity. The integrated peak areas for each complex are seen in Figure 3.3 plotted against DNA nanotag length. The increase in integrated peak area correlates linearly with the length of DNA nanotags. We view the ability to

lengthen the DNA nanoag hybridized to the miRNA target as a simple means to increase detection sensitivity via incorporation of more PicoGreen dye.



**Figure 3.4** – Inverse sandwich complex length ( $1/L$ ) versus scaled elution time ( $t/t_0$ ) for peaks in Figure 3.2. According to Equation 3.1, the slope can be used to calculate  $\alpha$  for the TX-100 ( $\alpha = 35$ ).

Figure 3.4 shows a plot used to predict the effective drag of the TX-100 micelle. The inverse complex length is plotted against scaled elution time ( $t/t_0$ ) according to Equation 3.1.  $t_0$  is measured by the elution time of unbound DNA nanotags, eluting first in Figure 3.2. From this plot, the effective drag of the TX-100 micelle,  $\alpha$ , can be calculated directly from the slope of the linear fit to the data. We find an  $\alpha$  value of 35 for the TX-100 micelles in this system, agreeing well with the value of 33.5 previously reported for end-alkylated double stranded PCR products.<sup>[50]</sup> This plot also predicts an intercept of 1, where our data shows an intercept of 1.13. This slight discrepancy may be from the small length dependence  $\mu_0$  has for double stranded DNA (dsDNA) less than 150 bases.<sup>[48]</sup>

### 3.3.2 – Theoretical Multiplexing Estimates

There are large gaps in elution time between the unbound DNA nanotags and the first *let-7b* complex, as well as over resolved separation of the shortest *let-7e* and *let-7f* complexes. These gaps represent potential elution times where additional miRNAs may be detected by sandwich complex hybridization in the same 4 minute separation. We apply a theoretical model to predict the lengths of DNA nanotag required to achieve baseline resolved separation of these additional miRNA complexes.

We can estimate the shape of the elution peak as Gaussian and express the spatial baseline width,  $w_b$ , as function of the total spatial peak variance ( $\sigma_x^2$ ),  $w_b = 4\sigma_x^2$ .  $\sigma_x^2$  includes several individual sources of variance, including finite injection/detection width ( $A$ ), diffusion ( $D$ ), wall adsorption ( $W$ ), tag polydispersity ( $B$ ) and Joule heating ( $J$ ).<sup>[52]–[54]</sup> Since each of these terms is independent of each other,  $\sigma_x^2$  can be written as a sum of these uncorrelated sources.<sup>[55]</sup>

$$\sigma_x^2 = A + \frac{2Dl_d}{v} + (B + W)v l_d + \frac{JE^4 l_d v}{D} \quad (3.2)$$

Where  $A$ ,  $B$ ,  $W$ , and  $J$  represents constants for their variance terms,  $D$  is the DNA – micelle complex diffusion coefficient,  $E$  is the electric field strength, and  $v$  is the velocity of the complex.

The variance due to injection,  $A$ , can be estimated from the approximate injection length,  $l_{inj}$ , using the Hagen-Poiseuille equation.<sup>[56]</sup> We find a  $l_{inj}$  of  $\sim 0.39$  cm for our hydrodynamic injections, yielding  $A = 0.0127$  cm<sup>2</sup>.

$$A = \frac{l_{inj}^2}{12} \quad (3.3)$$

The diffusion coefficient for double stranded DNA decreases as  $D = D_0 L^{-0.67}$ , where  $D_0$  is the diffusion coefficient of a single base pair reported as  $8.0 \times 10^{-6}$  cm<sup>2</sup>/s.<sup>[57], [58]</sup> We account for impact of the attached micelle on the dsDNA diffusion coefficient by treating the micelle-DNA conjugate as a double stranded DNA with total length  $(L + \alpha)$ .<sup>[59], [60]</sup>

$$D = D_0 (L + \alpha)^{-0.67} \quad (3.4)$$

We ignore the effects of Joule heating due to the small capillary diameters used and precise temperature control from the built in coolant jacket system.<sup>[54]</sup> This leaves the polydispersity and wall adsorption term  $(B+W)$  as the only fitting parameter in our system. This term represents the effective rate of wall adsorption and micelle sampling effects.<sup>[61]</sup> Smaller times represent a faster interaction and more effective separation.

The total peak variance does not give information regarding successful separation from other miRNA complexes. We use the resolution factor,  $R$ , to calculate the required length spacing between miRNA complexes to achieve baseline resolution. For example, an  $R$  of 2 means the complexes need to be

designed with a 2 base pair difference in DNA nanotag length. We define  $R$  as  $w_b$  divided by the peak spacing,  $dt/dL$ .

$$R = \frac{w_b}{\left| \frac{dt}{dL} \right|} \quad (3.5)$$

We can solve for  $dt/dL$  by taking the derivative of Equation 3.1.

$$\left| \frac{dt}{dL} \right| = \frac{l_d l_t \alpha}{\mu_0 V L^2} \quad (3.6)$$

Because detection takes place at a fixed point down the separation channel, what is actually measured is the temporal variance of the peak  $\sigma_t^2$ . A simple transformation from  $\sigma_x^2$  to  $\sigma_t^2$  can be made via the velocity  $v$  of the peak,  $\sigma_t^2 = \sigma_x^2/v^2$ . The full temporal variance model becomes

$$\sigma_t^2 = \frac{\sigma_x^2 t^2}{l_d^2} = \frac{A t^2}{l_d^2} + \frac{2 D t^3}{l_d^2} + (B + W) t + \frac{J E^4 t}{D} \quad (3.7)$$

By substituting Equation 3.1, 3.6, and 3.7 into 3.5, we can define  $R$  as

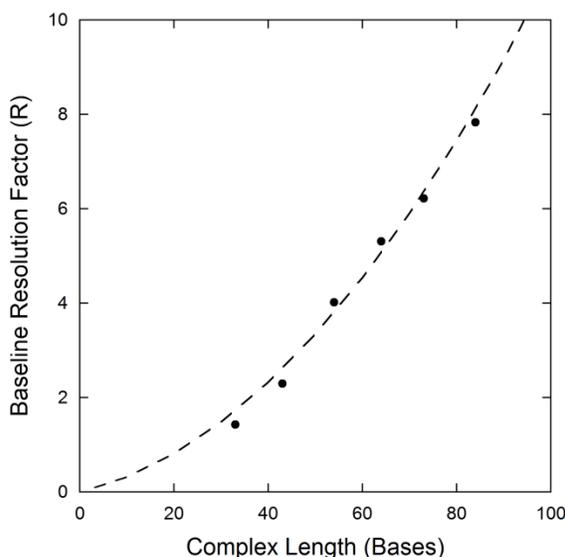
$$R = 4 \left[ \frac{A \mu_0 V l_t L^2 (L + \alpha)^2 + 2 D_0 (L + \alpha)^{7/3} L l_d l_t^2 + (B + W) L^3 \mu_0^2 V^2 l_d}{\mu_0 V l_t l_d^2 \alpha^2} \right]^{1/2} \quad (3.8)$$

We can experimentally measure the baseline resolution using the full-width at half maximum ( $FWHM$ ) and the actual  $dt/dL$  of the elution peaks as

$$R = \frac{4 FWHM}{\sqrt{8 \ln(2)} \left| \frac{dt}{dL} \right|} \quad (3.9)$$

Figure 3.5 shows the excellent agreement between experimental and theoretical  $R$  versus sandwich complex length in our system. The single  $(B+W)$  fitting parameter was found to be 2.8 ms. From our calculations, the injection variance accounts for > 88% of the total variance with < 10% contribution from  $(B+W)$ . Future improvements to multiplexed detection may be achieved through electrokinetic injection focusing techniques that minimize injection variance.<sup>[62]</sup>

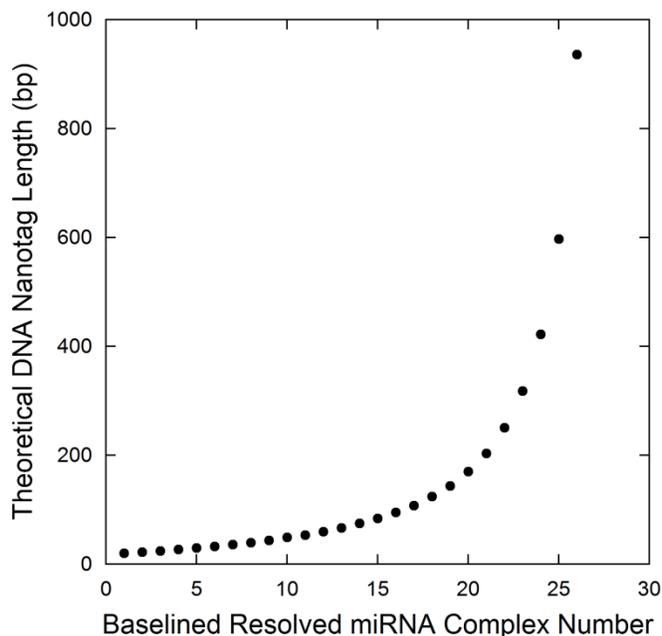
[64]



**Figure 3.5** – Experimental baseline resolution factors from peaks in Figure 3.2 (black circles) vs. theoretical model Equation 3.8 (dotted line).

Given access to DNA nanotags up to 1,000 bps, our theoretical model predicts that 26 different miRNA targets could be baseline resolved in a single 4 minute separation. Considering the low viscosity buffers used, the capillary can be easily rinsed with fresh buffer for sequential separations of different sets of miRNA targets. Many clinical labs also possess capillary array electrophoresis units. Using a 96 capillary array bundle, almost every known miRNA (2,496 out of 2,588) can be detected in a 5 single separation, offering tremendous advantages in

throughput compared to traditional solid-phase microarray methods that require days to analysis.<sup>[65]</sup>

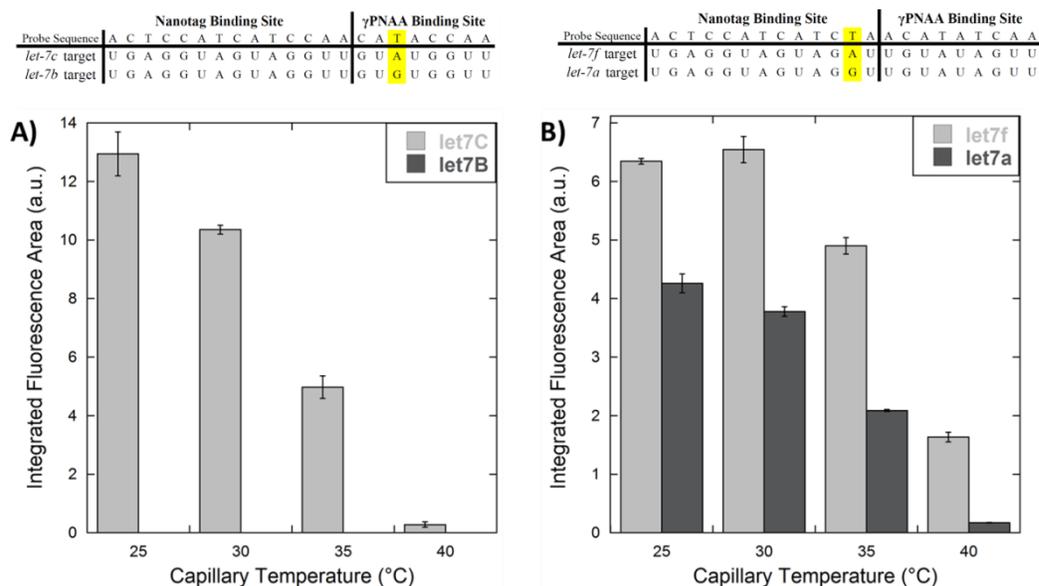


**Figure 3.6** – Theoretical DNA nanotag lengths required to achieve baseline resolution between sequential miRNA complexes.

The length of DNA nanotag for each theoretical complex is shown in Figure 3.6. As the DNA nanotag length increases, the drag from the TX-100 micelle becomes negligible. Sandwich complex spacing over 100 bases is required for baseline resolved resolution for DNA probes above 300 bps. By increasing the size and effective drag of the micelle interacting with the complex, a greater number of miRNAs may be detected by our method. We will investigate these types of buffers in the later chapters of this thesis.

### 3.3.3 – Single Base Mismatch Discrimination

The *let-7* miRNA family is typically used as the most stringent test for discrimination between closely related miRNA family members. We have previously demonstrated selectivity towards single base mismatches across both the DNA nanotag and  $\gamma$ PNAA using DNA targets.<sup>[40]</sup> To ensure that the selectivity is maintained with miRNA targets, the selectivity between *let-7b* and *let-7c* (single base mismatch on the  $\gamma$ PNAA) and *let-7a* and *let-7f* (single base mismatch on the DNA nanotag) was tested. These mismatches express pseudo-binding G-T wobble base-pair mismatches that result in minimal drops in free energy upon hybridization, making them extremely difficult to discriminate.<sup>[66]</sup>



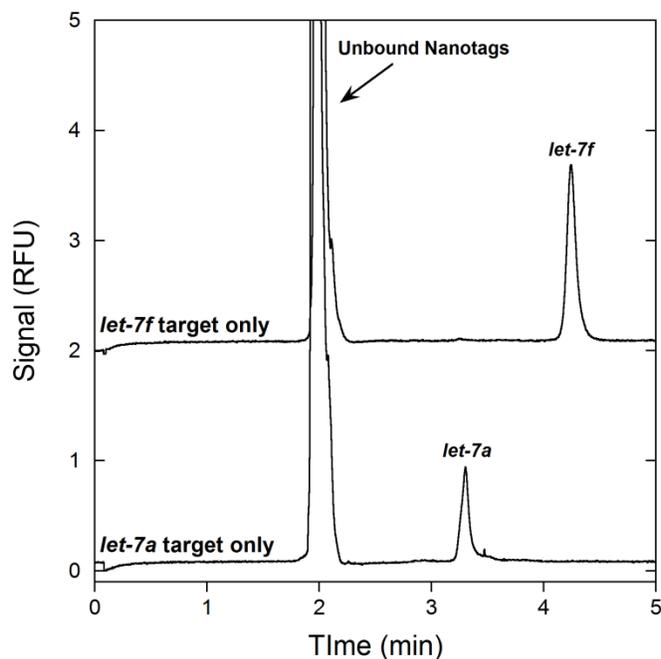
**Figure 3.7** – Integrated peak areas for single base mismatch complexes for different run temperatures. A) *let-7c* or *let-7b* incubated with *let-7c* hybridization probes. B) *let-7f* or *let-7a* incubated with *let-7f* hybridization probes. 250nM miRNA in 2.5 $\mu$ M of respective probe. miRNA,  $\gamma$ PNAA, and nanotag sequences shown above each figure. Error bars represent standard deviation from  $n = 3$  measurements.

The integrated area of perfect match and single base mismatch sandwich complex elution peaks are presented in Figure 3.7. We can see that the G-T

mismatch between *let-7b* and *let-7c* on the  $\gamma$ PNAA binding site is easily discriminated at 25°C. The G-T mismatch between *let-7a* and *let-7f* on the DNA nanotag binding site is not as easily recognized. A temperature of 40°C is required to achieve discrimination, similar to our previous work.<sup>[40]</sup> This could be due to the difference in the number of bases hybridized by each probe, 13 for the DNA nanotag and 9 for the  $\gamma$ PNAA. The shorter recognition site of the  $\gamma$ PNAA should create a larger  $\Delta G$  upon hybridization, as mismatch discrimination is strongly dependent on length for short oligomers.<sup>[67]</sup> The helical preorganization of the  $\gamma$ PNAA may also spatially confine a mismatched base within the helix, creating a greater degree of steric hindrance for mismatched bases compared to the more flexible DNA backbone.

Figure 3.7 also shows that elevating the run temperature results in a dramatic decrease in integrated peak area. The signal decrease is due to the added thermal pressures on miRNA hybridization, reducing the overall amount of miRNA detected. Although our method detects miRNA at 40°C, our detection sensitivity would be greatly hindered by the elevated temperatures. However, in the multiplexed miRNA detection platform, each different miRNA would be in solution with its perfect match hybridization probes, as well as those from single base mismatches. Since the perfect match hybridization probes have a higher thermal stability, the miRNA targets should bind perfect match probes prior to single base mismatch probes as the temperatures slowly drops in our annealing protocol. Thus, any miRNA present in solution will have a propensity to bind its

perfect match hybridization probes prior to any interaction with the single base mismatch probes.

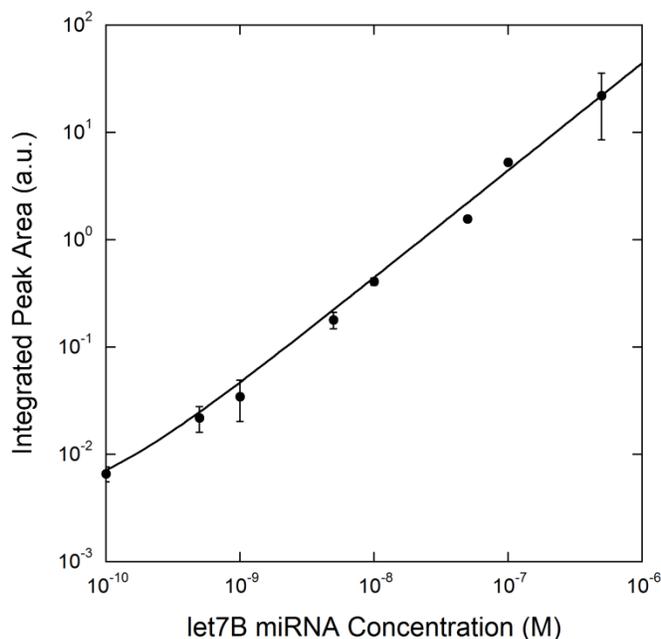


**Figure 3.8** – Representative electropherograms of competitive binding for single base mismatch target discrimination in solutions of both perfect match and mismatched hybridization probes. Separations performed at 25°C with 25nM of indicated miRNA and 250nM of *let-7a* and *let-7f* hybridization probes.

This competitive binding between perfect match and mismatched probes is clearly shown in the electropherograms separated at 25°C in Figure 3.8. Either *let-7a* or *let-f* miRNA targets were incubated in a solution of both *let-7a* and *let-7f* hybridization probes. For *let-7a* miRNA, we observe a single peak eluting around 3.25 minutes, corresponding to the longer 60bp *let-7a* DNA nanotag. For *let-7f* miRNA, we observe a single peak eluting around 4.25 minutes, corresponding to the shorter 30bp *let-7f* DNA nanotag. No off-target hybridization was observed for either miRNA, even though the individual incubations in Figure 3.7 required 40°C for discrimination. Through competitive binding, single base mismatch

discrimination is preserved at low run temperatures while maintaining high detection sensitivity.

### 3.3.4 – Detection Limits

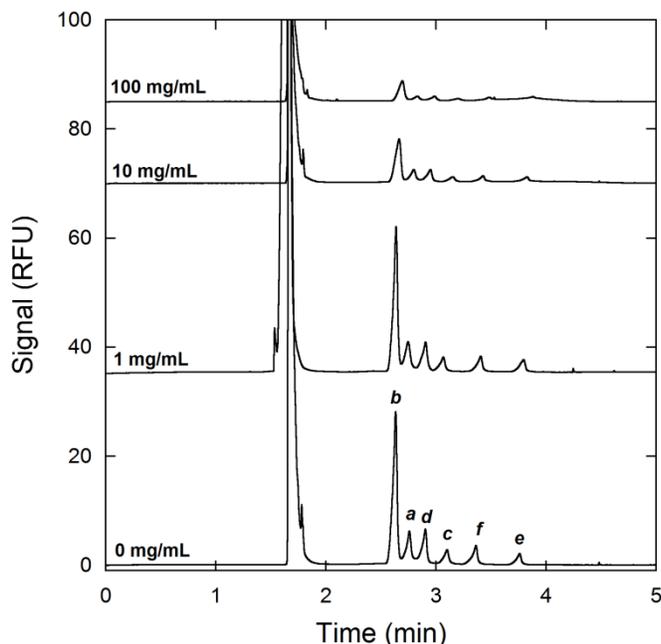


**Figure 3.9** – Integrated peak area for different concentrations of *let-7b* miRNA target in constant excess of sandwich hybridization probes (2.5 $\mu$ M probes). Solid line represents linear fit. Error bars represent standard deviation from  $n = 3$  measurements.

The detection sensitivity of the assay was determined by dilution of *let-7b* miRNA in a constant concentration of *let-7b* sandwich hybridization probes. Since the *let-7b* complex has the largest DNA nanotag length (70bp), we expect it to have the greatest amount of PicoGreen dye associated with its duplex, and hence most sensitive detection limit. Figure 3.9 shows the integrated peak area for *let-7b* elution peaks versus *let-7b* concentration. The dilution tests demonstrate linear sensitivity over 3 orders of magnitude down to 100pM of *let-7b* target. Using PicoGreen stained DNA nanotags offers a 100-fold improvement in detection sensitivity compared to our original, single end-label Cy5 fluorophore

system of 10nM.<sup>[40]</sup> By simply extending the length of DNA nanotag, we expect the detection sensitivity to increase dramatically to possible fM limits typically associated with alternative miRNA detection methods.

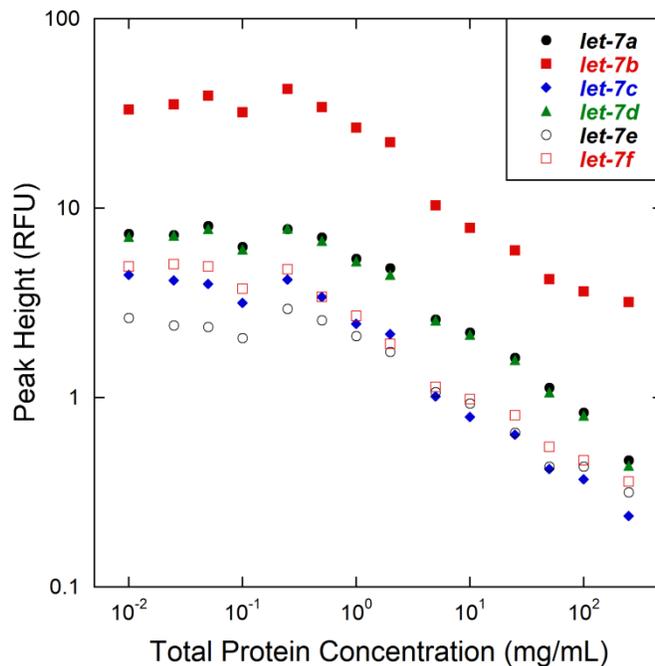
### 3.3.5 – Detection of miRNA in a Biological Matrix



**Figure 3.10** – Detection of *let-7* miRNA complexes in the presence of protein contamination from dilutions of goat serum. Total protein concentration is indicated on each electropherogram trace. Electropherograms are arbitrarily shifted on the y-axis for display purposes.

To test the ability of our sandwich probes to bind targets in a biological protein matrix, we mixed the miRNA sandwich complexes in dilutions of goat serum. Goat serum contains a mixture of different proteins including albumin, immunoglobulin G,  $\alpha_1$ -proteinase inhibitor,  $\alpha_2$ -macroglobulin, and transferrin.<sup>[68]</sup> These proteins may adsorb to the capillary wall, bind miRNA targets and/or hybridization probes, or interact with the hydrophobic alkane modification on the  $\gamma$ PNAAs, all resulting in decreased detection efficiency and reproducibility.<sup>[69]</sup>

Figure 3.10 demonstrates that our assay is tolerant of protein contaminants up to a total protein concentration of 100 mg/mL. The only effect the proteins appear to have on miRNA detection is decreasing signal intensity of the assay with increasing concentration of protein, see Figure 3.11. Peak morphology and separation resolution are unchanged. The lower signal intensity may be an effect of protein binding excess probes or interfering with the PicoGreen dye from staining the DNA. A gradual decrease in peak height begins at 1mg/mL until peaks are unresolved or undetectable at concentrations above 250 mg/mL. We do not expect biological contamination to have a strong impact on our detection assay as 100mg/mL is a high concentration of total protein and atypical for CE separations.



**Figure 3.11** – Effect of total protein concentration in goat serum on the peak height of *let-7* miRNA sandwich complex peaks.

### 3.4 – Conclusion

We have demonstrated the detection of 6 *let-7* miRNA targets in a 4 minute separation using a sandwich hybridization CE assay. Multiplexed detection is achieved via interaction of TX-100 micelles with the C18 alkane modification on a  $\gamma$ PNAA probe, and varying the length of DNA nanotag hybridized to each miRNA target. Theoretical models predict that our method is capable of detecting 26 different miRNAs in the same 4 minute separation by varying the DNA nanotag length up to 1,000 bases. We find a detection limit of 100pM using PicoGreen stained DNA nanotag bound to the miRNA targets, selectivity against G-T wobble single base mismatches, and tolerance to high concentrations (100mg/mL) of total protein contamination. Future work will focus on implementing longer DNA nanotags to achieve high throughput, multiplexed detection, and fM detection limits.

### 3.5 – References

- [1] A. Esquela-Kerscher and F. J. Slack, *Nat. Rev. Cancer*, 2006, 6, 259–269.
- [2] J. P. Cogswell, J. Ward, I. A Taylor, M. Waters, Y. Shi, B. Cannon, K. Kelnar, J. Kemppainen, D. Brown, C. Chen, R. K. Prinjha, J. C. Richardson, A. M. Saunders, A. D. Roses, and C. A Richards, *J. Alzheimers. Dis.*, 2008, 14, 27–41.
- [3] E. Van Rooij, W. S. Marshall, and E. N. Olson, *Circ. Res.*, 2008, 103, 919–928.
- [4] E. E. Creemers, A. J. Tijssen, and Y. M. Pinto, *Circ. Res.*, 2012, 110, 483–495.
- [5] A. Kozomara and S. Griffiths-Jones, *Nucleic Acids Res.*, 2011, 39, D152–D157.
- [6] S. Griffiths-Jones, R. J. Grocock, S. van Dongen, A. Bateman, and A. J. Enright, *Nucleic Acids Res.*, 2006, 34, D140–D144.

- [7] S. Griffiths-Jones, H. K. Saini, S. van Dongen, and A. J. Enright, *Nucleic Acids Res.*, 2008, 36, D154–D158.
- [8] Y. Li, C. Qiu, J. Tu, B. Geng, J. Yang, T. Jiang, and Q. Cui, *Nucleic Acids Res.*, 2014, 42, 1–5.
- [9] S. Volinia, G. A. Calin, C. G. Liu, S. Ambs, A. Cimmino, F. Petrocca, R. Visone, M. Iorio, C. Roldo, M. Ferracin, R. L. Prueitt, N. Yanaihara, G. Lanza, A. Scarpa, A. Vecchione, M. Negrini, C. C. Harris, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 2257–2261.
- [10] M. V. Iorio, M. Ferracin, C. G. Liu, A. Veronese, R. Spizzo, S. Sabbioni, E. Magri, M. Pedriali, M. Fabbri, M. Campiglio, S. Ménard, J. P. Palazzo, A. Rosenberg, P. Musiani, S. Volinia, I. Nenci, G. A. Calin, P. Querzoli, M. Negrini, and C. M. Croce, *Cancer Res.*, 2005, 65, 7065–7070.
- [11] R. Garzon, F. Pichiorri, T. Palumbo, R. Iuliano, A. Cimmino, R. Aqeilan, S. Volinia, D. Bhatt, H. Alder, G. Marcucci, G. A. Calin, C. G. Liu, C. D. Bloomfield, M. Andreeff, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, 103, 5078–5083.
- [12] R. Liu, C. Zhang, Z. Hu, G. Li, C. Wang, C. Yang, D. Huang, X. Chen, H. Zhang, R. Zhuang, T. Deng, H. Liu, J. Yin, S. Wang, K. Zen, Y. Ba, and C. Y. Zhang, *Eur. J. Cancer*, 2011, 47, 784–791.
- [13] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, and T. Tuschl, *Science*, 2001, 294, 853–858.
- [14] E. Várallyay, J. Burgyán, and Z. Havelda, *Nat. Protoc.*, 2008, 3, 190–196.
- [15] T. D. Schmittgen, E. J. Lee, J. Jiang, A. Sarkar, L. Yang, T. S. Elton, and C. Chen, *Methods*, 2008, 44, 31–38.
- [16] C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak, and K. J. Guegler, *Nucleic Acids Res.*, 2005, 33, e179.
- [17] W. Li and K. Ruan, *Anal. Bioanal. Chem.*, 2009, 394, 1117–1124.
- [18] C. G. Liu, G. A. Calin, B. Meloon, N. Gamliel, C. Sevignani, M. Ferracin, C. D. Dumitru, M. Shimizu, S. Zupo, M. Dono, H. Alder, F. Bullrich, M. Negrini, and C. M. Croce, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, 101, 9740–9744.
- [19] R. D. Morin, M. D. O’Connor, M. Griffith, F. Kuchenbauer, A. Delaney, A.-L. Prabhu, Y. Zhao, H. McDonald, T. Zeng, M. Hirst, C. J. Eaves, and M. A. Marra, *Genome Res.*, 2008, 18, 610–621.
- [20] Q. Wu, C. Wang, Z. Lu, L. Guo, and Q. Ge, *Clin. Chim. Acta*, 2012, 413, 1058–1065.

- [21] A. Git, H. Dvinge, M. Salmon-Divon, M. Osborne, C. Kutter, J. Hadfield, P. Bertone, and C. Caldas, *RNA*, 2010, 16, 991–1006.
- [22] U. Bissels, S. Wild, S. Tomiuk, A. Holste, M. Hafner, T. Tuschl, and A. Bosio, *RNA*, 2009, 15, 2375–2384.
- [23] F. Sato, S. Tsuchiya, K. Terasawa, and G. Tsujimoto, *PLoS One*, 2009, 4.
- [24] K. A. Cissell, Y. Rahimi, S. Shrestha, E. A. Hunt, and S. K. Deo, *Anal. Chem.*, 2008, 80, 2319–2325.
- [25] S. Husale, H. H. J. Persson, and O. Sahin, *Nature*, 2009, 462, 1075–1078.
- [26] H. Šípová, S. Zhang, A. M. Dudley, D. Galas, K. Wang, and J. Homola, *Anal. Chem.*, 2010, 82, 10110–10115.
- [27] G. J. Zhang, J. H. Chua, R. E. Chee, A. Agarwal, and S. M. Wong, *Biosens. Bioelectron.*, 2009, 24, 2504–2508.
- [28] H. Yang, A. Hui, G. Pampalakis, L. Soleymani, F. F. Liu, E. H. Sargent, and S. O. Kelley, *Angew. Chemie - Int. Ed.*, 2009, 48, 8461–8464.
- [29] B. N. Johnson and R. Mutharasan, *Anal. Chem.*, 2012, 84, 10426–10436.
- [30] M. B. Baker, G. Bao, and C. D. Searles, *Nucleic Acids Res.*, 2012, 40, 1–12.
- [31] Y. Kato, *Nucleic Acids Symp. Ser. (Oxf.)*, 2008, 71–72.
- [32] A. Persat and J. G. Santiago, *Anal. Chem.*, 2011, 83, 2310–2316.
- [33] R. B. Schoch, M. Ronaghi, and J. G. Santiago, *Lab Chip*, 2009, 9, 2145–2152.
- [34] L. A. Neely, S. Patel, J. Garver, M. Gallo, M. Hackett, S. McLaughlin, M. Nadel, J. Harris, S. Gullans, and J. Rooke, *Nat Meth*, 2006, 3, 41–46.
- [35] D. W. Wegman, L. T. Cherney, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2013, 85, 6518–6523.
- [36] D. W. Wegman and S. N. Krylov, *Angew. Chemie*, 2011, 123, 10519–10523.
- [37] D. W. Wegman, F. Ghasemi, A. Khorshidi, B. B. Yang, S. K. Liu, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2015, 87, 1404–1410.
- [38] N. Li, A. Nguyen, J. Diedrich, and W. Zhong, *J. Chromatogr. A*, 2008, 1202, 220–223.
- [39] M. V. Berezovski and N. Khan, *Methods Mol. Biol.*, 2013, 1039, 245–259.

- [40] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [41] D. W. Wegman and S. N. Krylov, *TrAC - Trends Anal. Chem.*, 2013, 44, 121–130.
- [42] F. Ghasemi, D. W. Wegman, M. Kanoatov, B. B. Yang, S. K. Liu, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2013, 85, 10062–10066.
- [43] L. T. Cherney and S. N. Krylov, *Analyst*, 2013, 138, 553–558.
- [44] H. Özhalıcı-Ünal and B. Armitage, *ACS Nano*, 2009, 3, 425–433.
- [45] A. I. Dragan, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman, and C. D. Geddes, *Biophys. J.*, 2010, 99, 3010–3019.
- [46] V. L. Singer, L. J. Jones, S. T. Yue, and R. P. Haugland, *Anal. Biochem.*, 1997, 249, 228–238.
- [47] B. Sahu, I. Sacui, S. Rapireddy, K. J. Zanotti, R. Bahal, B. A. Armitage, and D. H. Ly, *J. Org. Chem.*, 2011, 76, 5614–5627.
- [48] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [49] J. P. Vernille and J. W. Schneider, *Biotechnol. Prog.*, 2004, 20, 1776–1782.
- [50] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [51] K. Grass, C. Holm, and G. W. Slater, *Macromolecules*, 2009, 42, 5352–5359.
- [52] H. Ren, A. E. Karger, F. Oaks, S. Menchen, G. W. Slater, and G. Drouin, *Electrophoresis*, 1999, 20, 2501–2509.
- [53] P. Mayer, G. W. Slater, and G. Drouin, *Anal. Chem.*, 1994, 66, 1777–1780.
- [54] J. Knox and I. Grant, *Chromatographia*, 1987, 24, 135–143.
- [55] P. D. Grossman, *Capillary Electrophoresis: Theory and Practice*, 1st ed. Academic Press, 1992.
- [56] X. Peng and D. D. Y. Chen, *J. Chromatogr. A*, 1997, 767, 205–216.
- [57] N. C. Stellwagen, S. Magnusdottir, C. Gelfi, and P. G. Righetti, *Biopolymers*, 2001, 58, 390–397.
- [58] S. S. Sorlie and R. Pecora, *Macromolecules*, 1990, 23, 487–497.

- [59] M. A. Fahrenkopf, B. E. Ydstie, T. Mukherjee, and J. W. Schneider, *Comput. Chem. Eng.*, 2014, 64, 1–8.
- [60] L. C. McCormick, G. W. Slater, A. E. Karger, W. N. Vreeland, A. E. Barron, C. Desruisseaux, and G. Drouin, *J. Chromatogr. A*, 2001, 924, 43–52.
- [61] S. B. Istivan and J. W. Schneider, “Rapid Gel-free Separations of DNA via Transient Attachment to Surfactant Micelles,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2012.
- [62] R. Bharadwaj and J. G. Santiago, *J. Fluid Mech.*, 2005, 543, 57–92.
- [63] B. Jung, Y. Zhu, and J. G. Santiago, *Anal. Chem.*, 2007, 79, 345–349.
- [64] Z. Xu, E. Koshimidzu, and T. Hirokawa, *Electrophoresis*, 2009, 30, 3534–3539.
- [65] J. M. Thomson, J. Parker, C. M. Perou, and S. M. Hammond, *Nat. Methods*, 2004, 1, 47–53.
- [66] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner, *J. Mol. Biol.*, 1999, 288, 911–940.
- [67] Y. You, B. G. Moreira, M. A. Behlke, and R. Owczarzy, *Nucleic Acids Res.*, 2006, 34, e60.
- [68] S. Jain and M. N. Gupta, *Appl. Biochem. Biotechnol.*, 2005, 125, 53–62.
- [69] O. Selivanova, “Improved DNA Hybridization Assays Using Surfactant Probes,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2009.

# Chapter 4 – Coupled Transient Isotachophoresis Injections and Micelle ELFSE for Ultrasensitive DNA Detection

## 4.1 – Introduction

We have demonstrated stable and selective binding of microRNA (miRNA) in a sandwich hybridization assay with rapid, multiplexed detection by laser induced fluorescence capillary electrophoresis (CE-LIF).<sup>[1]</sup> The limit of detection (LOD) of the assay was found to be 100pM using double stranded DNA nanotags stained with PicoGreen intercalating dye. However, the natural abundance of miRNA expression levels may be as low as tens of copies per cell.<sup>[2]</sup> Minimally invasive biopsies, such as fine-needle aspiration, typically collect ~1,000 cells for analysis.<sup>[3]</sup> Therefore, the theoretically ideal detection limit for direct (enzyme amplification/ labeling free) detection of low abundance miRNA would be on the order of  $10^4$  targets, or fM LOD for the  $\mu\text{L}$  sample volumes used in the CE-LIF system.

So far, we have only worked with short oligomeric DNA nanotags up to 70 bases long for miRNA detection. We view the extension of these DNA probes to thousands of bases long as a simple way to achieve sensitive detection limits. The fluorescent signal is expected to increase linearly with nanotag length, as a greater amount of PicoGreen is incorporated into the DNA duplex. Therefore, comparing

to our current system, we would need a 7 million base pair DNA nanotag in order to achieve 1fM detection limits. These probe lengths are unrealistic for separation and hybridization. We do not expect lengthening of the nanotag alone to realize direct miRNA detection.

One of the main reasons for low detection sensitivity is the short light path length in the small bore capillaries of CE-LIF systems. To overcome this limitation, many CE-LIF separations take advantage of on-column preconcentration techniques such as field amplified sample stacking (FASS)<sup>[4], [5]</sup> and isotachopheresis (ITP)<sup>[6]-[10]</sup>. FASS and ITP both function under the principle that buffers with different conductivities will generate gradients in the electric field strength ( $E$ ). A low conductivity buffer generates a greater potential drop than the high conductivity buffer, giving it a higher local  $E$ . The velocity ( $v$ ) of a sample with electrophoretic mobility  $\mu$  can be defined as

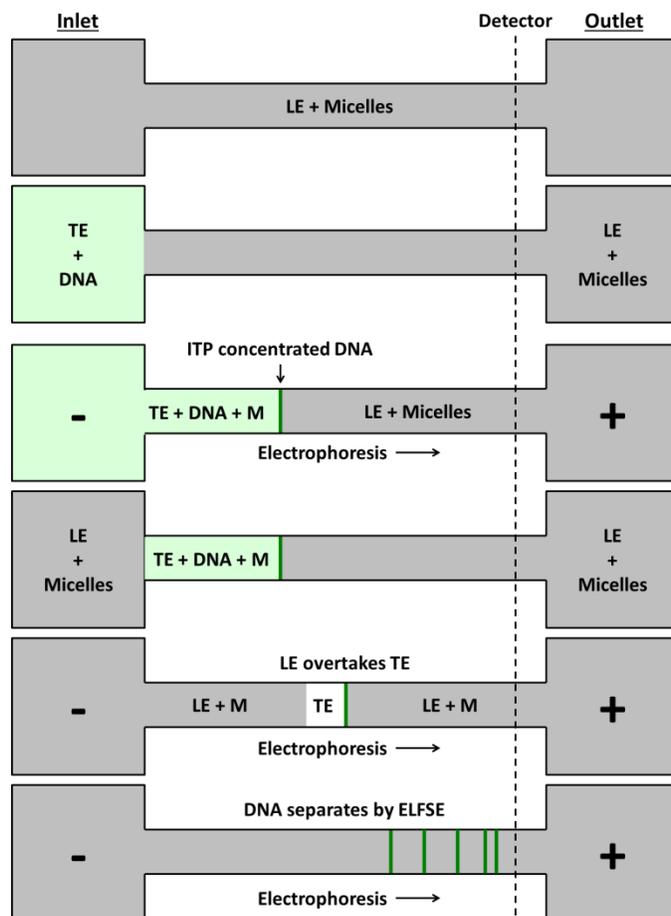
$$v = \mu E \quad (4.1)$$

Hence, as the sample migrates from the low conductivity buffer to the high conductivity buffer, it will slow down and concentrate at the interface.

In FASS, the sample is prepared in a low ionic strength buffer, typically deionized water or highly deionized formamide, to create a low conductivity buffer.<sup>[5], [11]</sup> FASS has the drawback that the low ion concentration sample buffer is a denaturing condition for DNA. Although denaturation is desired for applications such as single stranded DNA sequencing and forensic analysis, when DNA hybridization is required, as is for our miRNA detection method, FASS cannot be easily implemented.

In ITP, the conductivity of the buffers is altered using ions of differing electrophoretic mobility rather than different concentrations. Thus, sample concentration is achieved in high ionic strength buffers so DNA hybridization can be maintained. For ITP concentration of a sample, buffer co-ions (same charge as the sample) are chosen such that the sample has an intermediate mobility between a high mobility leading electrolyte (LE) and a low mobility trailing electrolyte (TE). As the buffer names imply, the sample is introduced between the LE and TE buffers. When an electric field is applied, the ITP buffer system creates sharp, self-correcting electric field gradients and the sample concentrates at the LE-TE boundary.<sup>[7]</sup>

ITP has been used to realize million-fold increases in sample concentration.<sup>[6], [8]</sup> However, when multiple samples are present at trace amounts, as is typically the case with DNA, separation is impossible as all samples will stack at the stable LE-TE interface. In order to trigger separation, the LE-TE interface must be disrupted and samples released.<sup>[12]-[14]</sup> This type of temporary ITP stacking has been implemented for many different samples, including DNA, and is referred to as transient ITP (tITP) or electrokinetic supercharging (EKSC). Typically, the capillary is filled with LE and sample electrokinetically injected from TE. Then, the tITP sample injection is followed by an LE buffer, where the high mobility ions quickly overtake the TE zone and return to a single component separation. This process is illustrated in Figure 4.1



**Figure 4.1** – Illustration of tITP injection process for sensitive micelle ELFSE detection. The capillary is filled with LE buffer containing non-ionic surfactant micelles. A DNA sample prepared in TE buffer is then placed at the inlet of the capillary. A potential is applied to the system, driving electrophoretic motion of LE, TE, and DNA towards the detector. Since the micelles are uncharged, they remain stationary within the capillary. As the TE DNA injection enters the capillary, it becomes a micelle saturated zone. However, due to the ITP conditions within the capillary, the DNA will rapidly concentrate at the interface between the LE and TE buffers without separation. After a substantial injection of DNA in TE, the inlet buffer is replaced by LE buffer containing non-ionic surfactant micelles. A potential is applied and the components migrate towards the detector. The DNA will continue to concentrate at the LE-TE interface until the high mobility LE ions overtake the low mobility TE ions, disrupting the ITP interface. The capillary returns to a single component buffer system and DNA will now separate according to micelle ELFSE.

Here, we discuss our efforts to incorporate tITP injections into our micelle end-labeled free-solution electrophoresis (ELFSE) CE separations of DNA to achieve sensitive detection limits. We demonstrate detection sensitivity down to 10fM of

$\lambda$ DNA using a combination of tITP and YOYO-1 intercalating dyes. The tITP injection grants a 100 fold improvement in sensitivity compared to the single buffer systems. We also demonstrate the effectiveness when coupling large tITP injections with micelle ELFSE separations. We find that larger tITP injections sacrifice capillary length for sample stacking in place of separation. Interestingly, the peak resolution is maintained for the longer tITP injections due to the sharp tITP interface that creates a very small initial DNA plug width. We observe 50 fold increases in signal of PicoGreen stained 1-6kB end-alkylated DNA and a 92 fold increase for a *let-7a* miRNA sandwich complex. We believe that by simply switching to longer capillary systems, we should be able to increase the stacking effect even greater while maintaining efficient separations.

## **4.2 – Materials and Methods**

### **4.2.1 – Reagents**

Triton X-100, nuclease-free water, acetic acid, hydrochloric acid, GeneJET Gel Extraction Kits, and acetone were purchased from Fisher Scientific (Pittsburgh, PA). BODIPY<sup>®</sup> FL C<sub>16</sub> (C16B), low melting temperature agarose, YOYO-1 intercalating dye, PicoGreen intercalating dye, and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). DMSO, 2,2-Dithiodipyridine (DPDS), Lithium perchlorate, Triphenylphosphine (TPP), 4-(Dimethylamino)pyridine (DMAP), Hexadecyltrimethylammoniumbromide (CTAB), Tris base, HEPES, ethidium bromide, acetonitrile, and triethanolamine (TEA) were purchased from Sigma Aldrich (St. Louis, MO). *N*-

dodecylpentaoxyethylene (C<sub>12</sub>E<sub>5</sub>) surfactant was purchased from BaChem (Torrance, CA).  $\lambda$ DNA and LongAmp<sup>TM</sup> Taq DNA polymerase Kit were purchased from New England Biolabs (Ipswich, MA). PCR forward primers and C<sub>6</sub>-amine 5' modified reverse primer were purchased with standard desalting purification from Integrated DNA Technologies (Coralville, IA). Dried oligomers were resuspended in 10mM Tris 0.1mM EDTA buffer and used without further purification.

#### **4.2.2 – DNA Alkylation Reaction**

Alkylation of C<sub>6</sub>-amine 5' modified reverse primer with BODIPY<sup>®</sup> FL C<sub>16</sub> (C16B) was accomplished using a combination of methods outlined by Mukaiyama and Deratani, and closely followed a scheme first reported by Boutorine *et al.*<sup>[15]-[17]</sup>. An aliquot of 10nmol 5'amine modified reverse primer was precipitated in the presence of 750nmol CTAB. The suspension was then dried by vacuum centrifugation under low heat for 30 minutes. Meanwhile, a mixture of 100nmol C16B, 5 $\mu$ mol DMAP, 2.5 $\mu$ mol TPP, 2.5 $\mu$ mol DPDS, and 7.5 $\mu$ mol TEA in DMSO was prepared and allowed to activate at room temperature for 20 minutes. Once activated, the reaction mixture was added to the dried DNA suspension, briefly vortexed and bath sonicated to resuspend the DNA pellet, and allowed to react overnight under continuous agitation in a shaker at room temperature in the dark. The solution was then precipitated with 1mL of 2w/v% lithium perchlorate in acetone and pelleted by centrifugation. The supernatant was removed and the pellet was triple rinsed with acetone. The pellet was resuspended in acetone, centrifuged, and supernatant removed again. The

pellet was then dried by vacuum centrifugation under low heat for 5 minutes and resuspended in 0.1M triethylammoniumacetate pH 7.0 (TEAA) buffer.

Purification of the reaction mixture was performed using a Waters 4.6 mm x 250 mm Symmetry300 C18 HPLC column (Milford, MA) with a 1mL/min flowrate and linear gradient from 0.1M TEAA to 100% acetonitrile over 30 minutes. Any unmodified DNA eluted near 10 minutes while the alkylated DNA eluted near 22 minutes. Product fractions were collected, lyophilized, and resuspended in 10mM Tris 0.1mM EDTA buffer.

### **4.2.3 – PCR Amplification of End-alkylated DNA**

Specific DNA lengths of end-alkylated DNA were amplified using LongAmp<sup>TM</sup> *Taq* DNA polymerase and  $\lambda$ DNA template with the described end-alkylated reverse primer and a length setting forward primer. The end-alkylated DNA ladder included the following lengths of DNA: 9893, 6003, 4111, 2840, 2093, and 1165. Each PCR reaction contained 1ng of  $\lambda$ DNA template, 500nM each primer, 0.3mM each dNTP, 2.5 units of LongAmp<sup>TM</sup> *Taq* DNA polymerase, 1x LongAmp<sup>TM</sup> buffer, and enough water to 25 $\mu$ L. The reaction was denatured for 2min at 95°C and subjected to 35 cycles of denaturation at 95°C for 15s, primer annealing at 50°C for 30 s, and extension at 65°C for 50 seconds per 1000 bases of amplicon length, followed by a final hold for 10min at 65°C.

After thermal cycling, the PCR reaction mixture was directly loaded onto 1wt% agarose gels. Slab gel electrophoresis was performed at 4 V/cm for 2 hours. Gels were imaged using ethidium bromide staining and UV excitation on a BioDoc-It Imaging System from Ultra-Violet Products (Cambridge, UK). Correct PCR

product length bands were excised and purified using GeneJET Gel Extraction Kit spin columns, and eluted using the provided 10mM Tris 0.1mM EDTA buffer.

#### **4.2.4 – Preparation of ITP Buffers**

LE and TE buffers were prepared by first making a solution of 50mM Tris base in 18.2m $\Omega$  deionized water. To make the LE buffer, the 50mM Tris base solution was titrated with HCl to a pH of 8.2. The Cl ion has an electrophoretic mobility of  $-7.91 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . To make the TE buffer, the 50mM Tris base solution was titrated with HEPES to a pH of 7.8. HEPES has an electrophoretic mobility of approximately  $-1.5 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . Titrated LE and TE buffers were filtered through a 0.22 $\mu\text{m}$  filter and could be stored for many months without loss of ITP performance.

#### **4.2.5 – Capillary Electrophoresis**

Capillary electrophoresis (CE) experiments were performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. Excitation sources were from a 3mW 488nm Argon ion. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 50 $\mu\text{m}$  ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and of indicated length in individual methods sections. Samples were electrokinetically injected into the capillary at specific voltages and times as indicated in individual methods sections. Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer and maintained by adding 0.1v/v% POP-6 to all capillary electrophoresis running buffer. The capillary was cooled to 25°C using a built in coolant jacket system. Data collection was performed using 32 Karat

software (Beckman Coulter). Elution peaks were analyzed using OriginPro 9.0 (Northampton, MA). The capillary could be used for many separations without observing any detrimental effects.

Surfactant buffers were prepared by mixing either 6mM C<sub>12</sub>E<sub>5</sub> and 0.75mM Triton X-100 or 24mM TX-100 in LE buffer, briefly vortexing, and shaken for at least 1 hour at room temperature to ensure surfactant solubilization and micelle formation. The mixed surfactant buffer was previously found to provide sufficient separation of long DNA at low voltage.<sup>[18]</sup>

#### **4.2.6 – $\lambda$ DNA Dilutions**

$\lambda$ DNA samples were prepared in LE or TE buffer with YOYO-1 intercalating dye at a ratio of 1 dye to 4 base pairs. The solutions were allowed to rest for 1 hour at room temperature in the dark to ensure the YOYO-1 stained the  $\lambda$ DNA. A capillary length of 40cm to the detector and 50cm total was filled with LE buffer. Samples in either LE or TE were injected into the capillary at 30kV for 1 minute. Electrophoresis was conducted under an applied voltage of 30kV in reverse polarity mode (from anode to cathode) with inlet and outlet buffer in LE.

#### **4.2.7 – tITP End-alkylated DNA Injections and Separations**

The C16B end-alkylated DNAs were prepared by pipetting 1 $\mu$ L of each gel purified PCR product with 30 $\mu$ L of either LE or TE, and a 1/400<sup>th</sup> stock dilution of PicoGreen dye. The capillary length was 20cm to the detector and 30cm total. Samples were electrokinetically injected for total volt-sec indicated in figure captions. Electrophoretic separation was conducted under an applied voltage of

3kV in reverse polarity mode (from anode to cathode). Both sample and capillary buffers contained 1/400<sup>th</sup> dilution of stock PicoGreen dye.

#### **4.2.8 – tITP miRNA Sandwich Hybridization Detection**

Sandwich complex hybridization was carried out by mixing 250nM of *let-7a* miRNA target with 250nM of *let-7a*  $\gamma$ PNA and DNA nanotag in LE or TE buffer. Solutions were heated to 95°C for at least 5 minutes in a dry bath incubator and allowed to cool to room temperature over the course of 60 minutes. The capillary length was 20cm to the detector and 30cm total. Samples were electrokinetically injected into the capillary at 4kV for 30sec. Electrophoretic separation was conducted under an applied voltage of 3kV in reverse polarity mode (from anode to cathode). Both sample and capillary buffers contained 1/400<sup>th</sup> stock dilution of PicoGreen dye.

### **4.3 – ITP Theory**

In a buffer filled capillary, the flux of ions ( $N_i$ ) will have contributions from diffusion and electromigration driven by an electric field.

$$N_i = -D_i \nabla c_i + v_i c_i \quad (4.2)$$

Where  $D_i$  is the  $i$ th ion's molecular diffusivity,  $c_i$  is the concentration of the  $i$ th ion, and  $v_i$  is the velocity of an individual ion. For an EOF suppressed system, there is no bulk fluid flow and  $v_i$  can be expressed as a function of the electrical force on the ion by the local electric field ( $E$ ) multiplied by the  $i$ th ion's mobility ( $\mu_i$ )

$$v_i = Fz_i \mu_i E \quad (4.3)$$

Where  $F$  is Faraday's constant and  $z_i$  is the valence number of the  $i$ th ion. Conservation of mass requires

$$\frac{dc_i}{dt} + \nabla \cdot N_i = 0 \quad (4.4)$$

Substituting Equation 4.2 and 4.3 into 4.4, we recover the advection-diffusion equation for electromigration in the absence of bulk fluid flow.

$$\frac{\partial c_i}{\partial t} = -Fz_i\mu_i \nabla \cdot (C_i E) + D_i \nabla^2 C_i \quad (4.5)$$

We can simplify the equation by assuming the concentration varies only along the axial direction of the capillary ( $x$ ) and ignore diffusion, as stacking dynamics are typically performed under small time scales where diffusion is minimal. This leaves a simplified governing equation

$$\frac{\partial c_i}{\partial t} = -Fz_i\mu_i \frac{\partial (c_i E)}{\partial x} \quad (4.6)$$

When an axial electrical potential is applied, the spatially segregated regions act as a circuit of resistors in series. The series of resistors creates a constant current density,  $j$ , along the axial direction of the capillary. We can express the local  $E$  as a function of the electrical conductivity,  $\kappa$ , at any  $x$ -coordinate in the capillary as

$$E(x) = \frac{j}{\kappa(x)} = \frac{j}{F \sum |z_i| \mu_i c_i(x)} \quad (4.7)$$

Substituting 4.7 into 4.6, summing over all ionic species, and dividing by  $\mu_i$  yields

$$\frac{\partial}{\partial t} \sum_{i=1}^N \frac{c_i}{\mu_i} = -jF \frac{\partial}{\partial x} \frac{\sum_{i=1}^N z_i c_i}{\kappa(x)} \quad (4.8)$$

Since the capillary must maintain electroneutrality, except at the small electrical double layer that we ignore since  $r_{Debye} \ll r_{capillary}$ ,

$$\sum_{i=1}^N z_i c_i = 0 \quad (4.9)$$

Thus, the time derivative of concentration on the left side of the equation must be zero. This must mean that the sum of  $c_i/\mu_i$  itself is not dependent on time and remains constant throughout the electrophoretic process.

$$\sum_{i=1}^N \frac{c_i(x,t)}{\mu_i} = KRF(x) \quad (4.10)$$

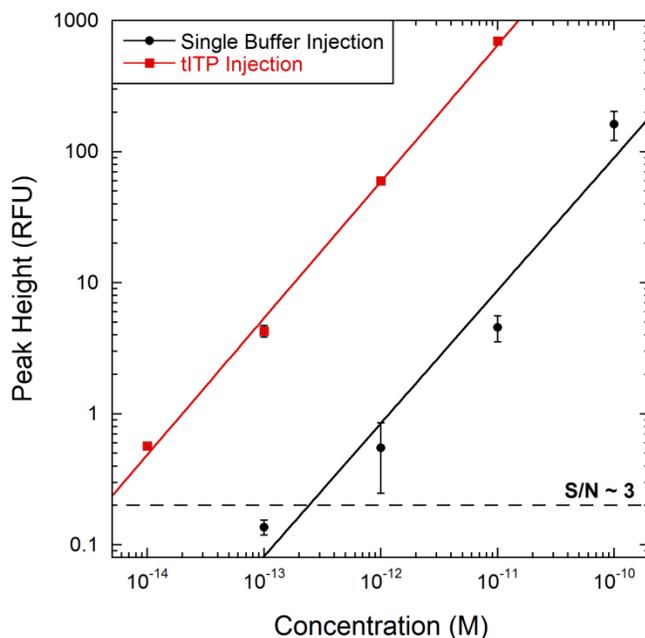
This result is known as the Kohlraush regulating function (*KRF*), as derived by Kohlraush to describe the change in concentration of ionic species as a function of the axial position in a system and electrophoretic mobility.<sup>[19]</sup> Put simply, if a region of high concentration, high mobility ion (LE) is replaced by a low mobility ion (TE), the TE concentration will have to increase in order to maintain the KRF. This is the operating principle of ITP.

## 4.4 – Results and Discussion

### 4.4.1 – tITP $\lambda$ DNA Detection Limits

We tested the effects of tITP injections for ultra-sensitive detection using  $\lambda$ DNA stained with YOYO-1 intercalating dye.  $\lambda$ DNA is a commercially available double stranded DNA 48,502 bases in length. It is typically used for single molecule imaging as the large length allows it to intercalate thousands of YOYO-1 dye molecules. We view these long DNAs as means of ultrasensitive detection

of low abundance microRNAs, where the entire DNA strand becomes a bright fluorescent marker upon hybridization to the target.

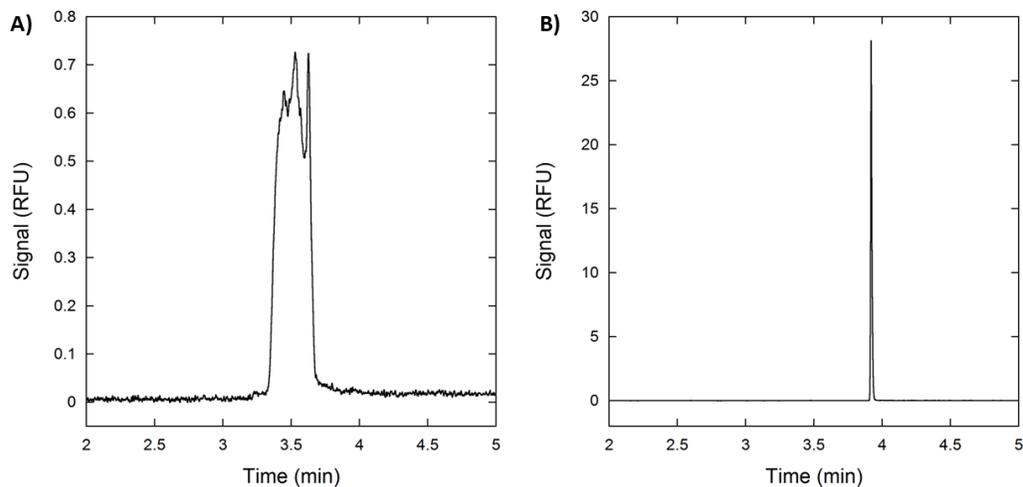


**Figure 4.2** – Elution peak height for titrations of  $\lambda$ DNA stained with YOYO-1 dye for single buffer system injections (black) or tITP injections (red). Error bars are standard deviation for  $n = 3$ .

Figure 4.2 shows the detection limits of YOYO-1 stained  $\lambda$ DNA by conventional single buffer system CE, both sample and capillary containing LE (50mM Tris HCl, pH 8.2). Large plugs of  $\lambda$ DNA were electrokinetically injected into the capillary and electrophoresed past the detector. A detection limit of  $\sim 1\mu\text{M}$  was found for the single buffer system, a 100 fold improvement upon the short, oligomeric DNA nanotags previously used for miRNA detection.

The YOYO-1 stained  $\lambda$ DNA was also prepared in a TE sample buffer (50mM Tris HEPES, pH 7.8). The LE filled capillary creates a tITP injection scheme during electrokinetic sample injection due to the discontinuous buffer system. The detection limits using tITP injections are seen in Figure 4.2. There is a 100 fold improvement in detection compared to the single buffer system, down to  $\sim 10\text{fM}$ .

These fM limits are viable for amplification free detection of very low abundance miRNAs.



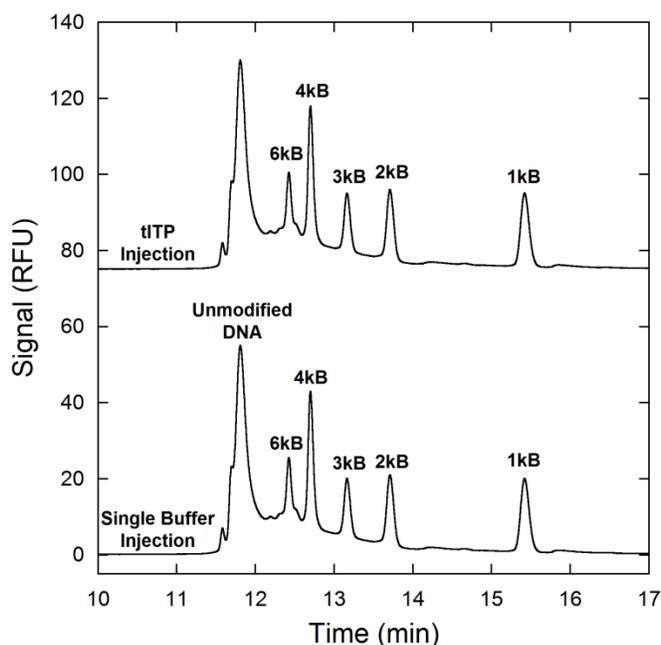
**Figure 4.3** – Representative elution peak morphology for 1pM YOYO-1 stained  $\lambda$ DNA from A) single buffer system injections or B) tITP injections.

The difference between single buffer and tITP injections is clear from the representative electropherograms in Figure 4.3 for 1pM  $\lambda$ DNA. The single buffer injection yields a broad elution peak, leading to poor sensitivity. The tITP injection yields a very sharp elution peak, due to stacking of the  $\lambda$ DNA at the LE-TE interface. Once the ITP interface attains steady state, the interface will not broaden as it migrates down the capillary. This effect is caused by the electric field gradient created during ITP continually focusing the DNA during electrophoresis. The TE zone behind the  $\lambda$ DNA has a higher electric field strength than the LE zone in front it. If the  $\lambda$ DNA diffuses into the TE zone, it will migrate faster than the low mobility TE ions until it reaches the interface. Conversely, if the  $\lambda$ DNA diffuses into the LE zone, it migrates slower than the high mobility LE ions and falls back to the interface. Thus, these interfaces are self-correcting and

have been measured on the order of 10-100 $\mu$ m wide, leading to very sharp stacking injections.

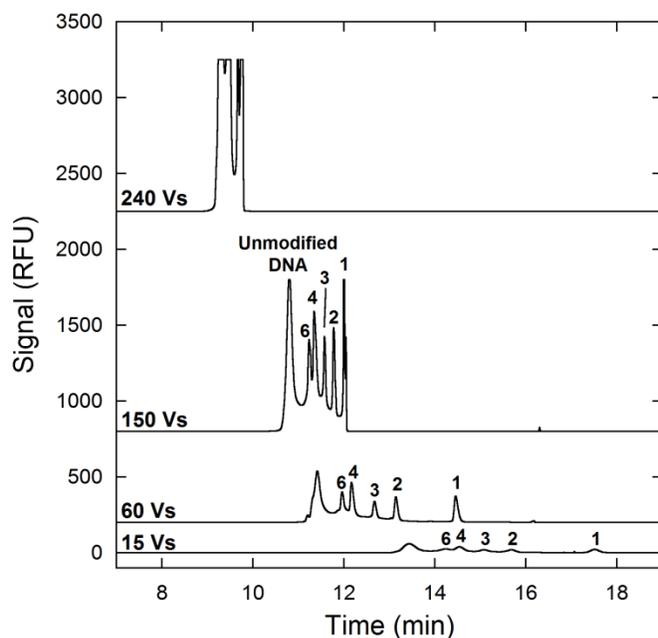
#### 4.4.2 – tITP Injections with Micelle ELFSE

Although tITP injections of  $\lambda$ DNA realized 10fM detection limits, separation of multiple lengths of DNA is made difficult by ITP. The DNA is at too low of a concentration to contribute to local electric field gradients and will all migrate together at the interface, even in sieving matrices or ELFSE systems. Separation must be triggered by disrupting the ITP interface. This can be accomplished by following the tITP injection with LE buffer, where the high mobility LE ions quickly overtake the TE ions in the high electric field zone and the capillary returns to a single component system, see Figure 4.1.



**Figure 4.4** – Representative electropherograms of C16B end-alkylated DNA stained with PicoGreen dye for either single buffer system injections or tITP injections. Injections were 1kV for 15sec. DNA lengths are indicated above each peak. Electropherograms are arbitrarily shifted on the y-axis for display purposes.

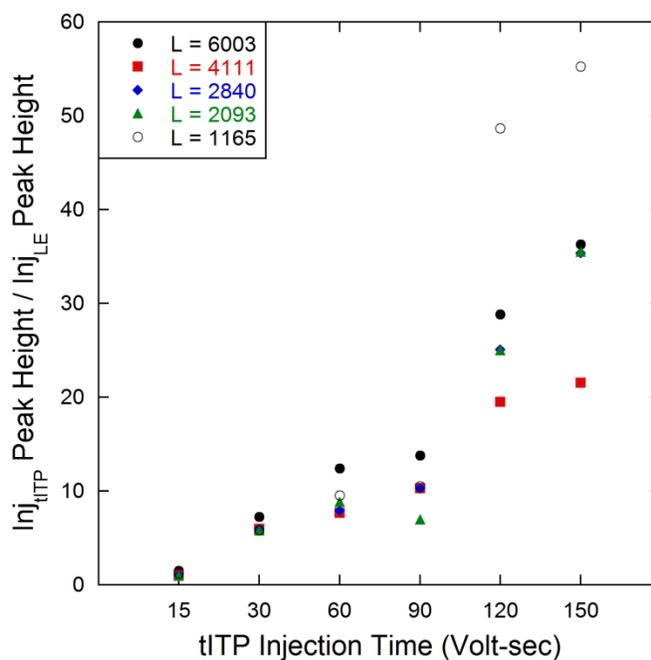
Figure 4.4 shows representative separations of C16B end-alkylated DNA using single buffer systems or tITP injections. The capillary was filled with 6mM C<sub>12</sub>E<sub>5</sub> 0.75mM TX-100 mixed surfactants in LE buffer. The samples were injected from either LE (single buffer system) or TE (tITP). The inlet of the capillary was then placed into a vial containing LE buffer and an electric field applied. We can see that both systems have almost identical separations, demonstrating that tITP and micelle ELFSE are compatible.



**Figure 4.5** – Effect of increasing tITP injection time on peak morphology of long C16B end-alkylated DNA stained with PicoGreen dye. Total injection time is represented as Volt-sec, indicated on each trace. DNA lengths are indicated above each peak. Electropherograms are arbitrarily shifted on the y-axis for display purposes.

The injections used in Figure 4.4 were quite small (15 Vs). Only minor improvements in signal are observed. Longer injections are required in order to have enough time for DNA to migrate to the ITP interface, and allow a greater amount of DNA to enter the system. The effect of increasing tITP injection time is shown in Figure 4.5. As the tITP time increases, we observe dramatic increases

in the height of each DNA length peak. The ratio of tITP injection time peak height versus the single LE buffer system injection peak height for each length is seen in Figure 4.6. 20 – 50 fold improvements are realized for the largest 150 Vs tITP injection that could be resolved. The shortest length, 1165 bases, shows the greatest increase in signal. The 1165 base DNA also has the lowest electrophoretic mobility in micelle ELFSE. The lower mobility means it will have to satisfy the KRF by attaining a higher overall concentration, leading to a higher signal.



**Figure 4.6** – Relative peak height for tITP injected DNA vs. single buffer injections.

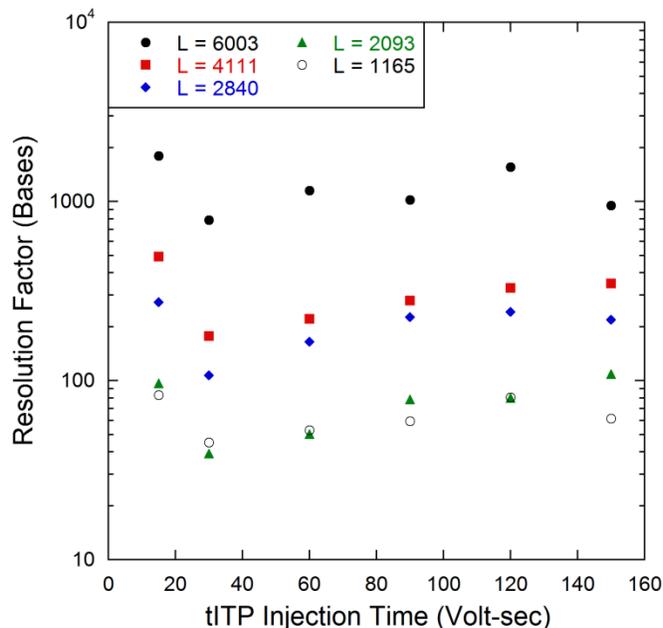
As the tITP injection time increases, less of the capillary is available for separation. Not only does the ITP boundary move further into the capillary, but the time required for the LE ions to overtake the interface also increases. These

effects result in compression of the peaks and earlier elution time due to the effectively shorter capillary, as seen in Figure 4.5.

Interestingly, the DNA lengths are still well resolved with the increased injection times. The resolution factor ( $R$ ) can be calculated from the full-width at half maximum ( $FWHM$ ) and peak spacing ( $dt/dL$ )

$$R = \frac{FWHM}{dt/dL} \quad (4.11)$$

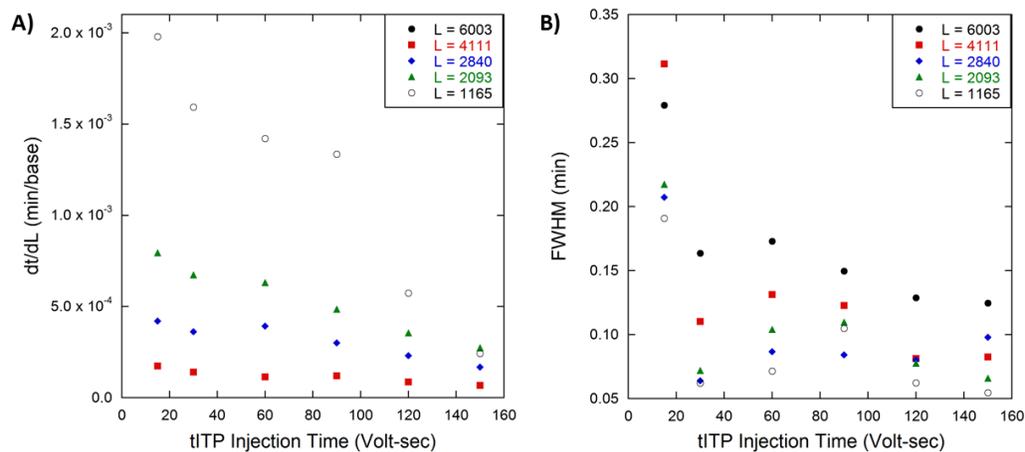
The resolution factors are almost constant vs. injection time, Figure 4.7. Although the peaks become more compressed as injection time increases (smaller  $dt/dL$ ), the sharp ITP interface acts to reduce the  $FWHM$  of the injection and maintain resolution of the peaks, Figure 4.8.



**Figure 4.7** – Resolution factors, calculated using Equation 4.11, for long DNA lengths as a function of tITP injection time.

For the 240Vs injection, the detector is saturated at 1,000 RFU, but there is a dip that appears to represent separation between the DNA lengths. The sacrifice

between sensitive detection and separation is limited by the total length of the capillary. By extending the length of the capillary, we can effectively add more capillary to increase the efficiency of each process. However, increasing the length of the capillary will reduce  $E$  and increase run time. In order to maintain rapid separations, shorter capillaries should be used.



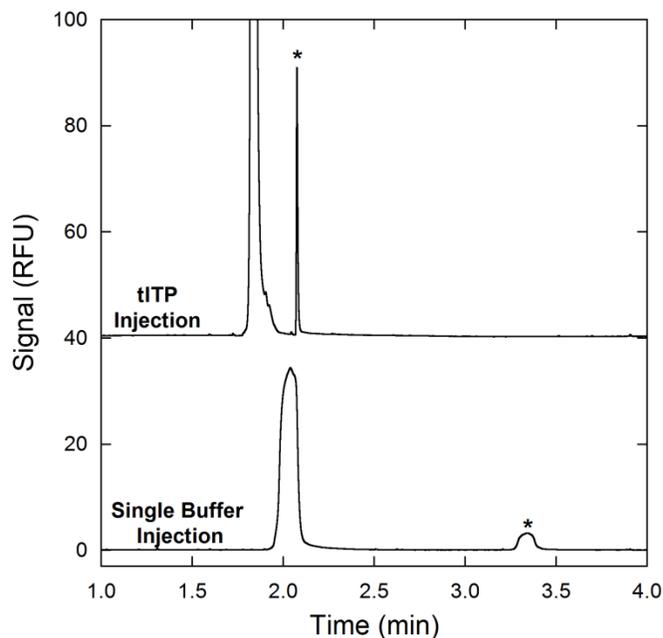
**Figure 4.8** – A) Peak spacing ( $dt/dL$ ) between DNA lengths as a function of tITP injection time. B) Peak broadness ( $FWHM$ ) of elution peaks as a function of tITP injection time.

A very ingenious method to maintain capillary length for separation while utilizing large ITP injections<sup>[10]</sup> was demonstrated by Krylov *et al.*<sup>[10]</sup> In this method, a large plug of DNA in TE was injected from the outlet reservoir into a TE containing capillary. The outlet was switched to a LE buffer and a bulk electroosmotic flow (EOF) was used to carry the ITP interface towards the inlet of the capillary. As the interface migrated towards the inlet, the DNA samples were continually focused at the LE-TE boundary. Once the interface reached the inlet capillary end, the inlet buffer was switched to LE and polarity switched such that the EOF carried the DNA towards the detector at the outlet. Using this method, the entire capillary can be utilized as an ITP focusing step towards the inlet. Then,

by switching the direction of  $E$ , the entire length of the capillary could be used for separation. They were able to achieve a 100 fold increase in detection sensitivity by this method while maintaining efficient ELFSE separations. We plan to investigate this injection procedure for sensitive micelle ELFSE tITP separations.

#### 4.4.3 – tITP miRNA Sandwich Detection

We have attempted to implement tITP focusing for our miRNA sandwich hybridization assay. Figure 4.9 shows the representative electropherograms for the *let-7a* miRNA sandwich complex injected from LE (single buffer injection) and TE (tITP injection). As seen with the long C16B end-alkylated DNAs, there is a dramatic increase in peak height for the tITP injection scheme. In this case, we observe a 92 fold improvement in peak intensity compared to the single buffer injection.



**Figure 4.9** – tITP focusing of *let-7a* miRNA sandwich complex. 250nM *let-7a* miRNA incubated with 250nM of  $\gamma$ PNAA and DNA nanotag in either LE (single buffer injection) or TE (tITP injection). miRNA sandwich peak marked with an asterisk.

Attempts to multiplex detection for the miRNA sandwich complexes with tITP injections have been challenging. We have noticed that the sandwich complexes will elute as one, ITP focused peak separated from the unbound probes, rather than previously attained individual peaks for each miRNA target. This may be a result of the minimal differences in sandwich complex length and the small Triton X-100 micelle used for separation. Since the peaks compress during tITP injections, increased separation performance may be required to achieve multiplexed detection.

## **4.5 – Conclusion**

We have demonstrated the effective coupling of tITP preconcentration injections with efficient micelle ELFSE separations. Using tITP injections and YOYO-1 stained  $\lambda$ DNA, we have achieved 10fM detection limits in our capillary. We have demonstrated increased detection sensitivity of PicoGreen stained end-alkylated kB DNA by tITP injections. Long tITP injections yielded 20-50 fold improvements compared to the single buffer system injections. Resolution of the micelle ELFSE separation was maintained by the tight tITP injections that focus DNA at the LE-TE interface. This focusing counters the sacrifice of capillary length for separation for longer tITP injections. We have also demonstrated the application of tITP injection to single miRNA sandwich complex detection, with a 92 fold improvement in sensitivity. Efforts to multiplex miRNA detection failed, believed to be caused by the reduced separation length under tITP conditions. Further investigation of longer capillary lengths and alternative ITP injection

schemes should yield intense, multiplexed miRNA detection by sandwich complex hybridization.

## 4.6 – References

- [1] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [2] C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak, and K. J. Guegler, *Nucleic Acids Res.*, 2005, 33, e179.
- [3] B. J. Dodgson, A. Mazouchi, D. W. Wegman, C. C. Gradinaru, and S. N. Krylov, *Anal. Chem.*, 2012, 84, 5470–5474.
- [4] H. W. Liao, S. W. Lin, U. I. Wu, and C. H. Kuo, *J. Chromatogr. A*, 2012, 1226, 48–54.
- [5] R. Bharadwaj and J. G. Santiago, *J. Fluid Mech.*, 2005, 543, 57–92.
- [6] B. Jung, R. Bharadwaj, and J. G. Santiago, *Anal. Chem.*, 2006, 78, 2319–2327.
- [7] T. K. Khurana and J. G. Santiago, *Anal. Chem.*, 2008, 80, 6300–6307.
- [8] Z. Xu, K. Nakamura, A. R. Timerbaev, and T. Hirokawa, *Anal. Chem.*, 2011, 83, 398–401.
- [9] Z. Xu, T. Ando, T. Nishine, A. Arai, and T. Hirokawa, *Electrophoresis*, 2003, 24, 3821–3827.
- [10] D. W. Wegman, F. Ghasemi, A. Khorshidi, B. B. Yang, S. K. Liu, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2015, 87, 1404–1410.
- [11] L. Gross and E. S. Yeung, *J. Chromatogr.*, 1989, 480, 169–178.
- [12] S. S. Bahga, R. D. Chambers, and J. G. Santiago, *Anal. Chem.*, 2011, 83, 6154–6162.
- [13] G. Garcia-Schwarz and J. G. Santiago, *Anal. Chem.*, 2012, 84, 6366–6369.
- [14] A. Persat and J. G. Santiago, *Anal. Chem.*, 2011, 83, 2310–2316.
- [15] T. Mukaiyama, *Angew. Chemie Int. Ed. English*, 1976, 15, 94–103.
- [16] A. Deratani and T. Maraldo, *React. Polym.*, 1988, 9, 19–28.

- [17] T. Da Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, and B. Cacciari, *Curr. Med. Chem.*, 2005, 12, 71–88.
- [18] A. Jones, “Rapid Separations of Kilobase-Sized DNA using Micelle Size-Sampling,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2013.
- [19] V. Hruška and B. Gaš, *Electrophoresis*, 2007, 28, 3–14.



# Chapter 5 – Unmodified DNA Separation in Entangled Wormlike Micelle Networks

## 5.1 – Introduction

Capillary electrophoresis (CE) is one of the most popular and powerful methods for separating DNA by size. Since DNA has a length invariant charge to friction ratio, they cannot be separated in free solution and require a separation matrix that relies on size dependent interactions. Initial CE attempts of DNA separations utilized traditional cross-linked polyacrylamide slab gel matrices.<sup>[1], [2]</sup> Although functional, difficulties present themselves when polymerizing gels within capillaries. Gel systems were soon replaced by physically entangled (uncross linked), high molecular weight polymer solutions including linear polyacrylamide<sup>[3]</sup>, PEO<sup>[4], [5]</sup>, pDMA<sup>[6]</sup>, and cellulose based polymers<sup>[7], [8]</sup>. The entangled networks form a mesh that the DNA must migrate through by Ogston sieving or reptation mechanisms, similar to the pores of the gel. However, introducing the high viscosity polymer solutions into the small bore capillaries is difficult and dramatically increases the assay time.

Barron *et al.* first observed the separation of DNA in dilute, low viscosity solutions of hydroxyethyl cellulose (HEC) polymers that were not entangled.<sup>[9]</sup> A new transient entanglement coupling (TEC) mechanism was proposed where the DNA collides with individual polymers and is forced to drag the polymer through solution. TEC enables the separation of large kB – MB DNAs at high electric

fields (typically  $> 100$  V/cm). Unfortunately, the resolution of TEC separations is rather poor compared to sieving or reptation in entangled polymer solutions and does not have a unified, well understood separation mechanism. The observation of TEC has led many to develop microfluidic post-array devices that function under similar principles as TEC.<sup>[10]-[12]</sup> However, these systems are difficult to fabricate and are not among the common equipment available to many clinical labs.

Non-ionic *n*-alkyl polyoxyethylene ether surfactants ( $C_iE_j$ s) form large, wormlike micelle structures in aqueous solutions. As the concentration of  $C_iE_j$  surfactant increases, the wormlike micelles grow in length and concentration, leading to entanglement with one another.<sup>[13], [14]</sup> They begin to resemble solutions of physically entangled polymers, with the important distinction that they are not covalently linked monomer chains.  $C_{16}E_6$  and  $C_{16}E_8$  surfactant solutions have been previously demonstrated to act as dynamic polymer matrices for DNA separations.<sup>[15]</sup> Yeung *et al.* proposed a reversible gel model, developed by Duke and Viovy to explain the DNA separation mechanism in these surfactant networks.<sup>[15], [16]</sup> Although the model's assumptions appear valid, no attempts were made to verify it, and design of surfactant buffers was made largely empirically.

Here, we attempt to understand how  $C_{12}E_5$  surfactant solutions affect DNA separation by directly applying Ogston sieving, reptation, and TEC theories to our data. We use  $C_{12}E_5$  as it is the major surfactant used by the Schneider lab for most of our end-labeled free solution electrophoresis (ELFSE) separations. Our goal is

to identify migration methods of unmodified DNA to better understand the behavior of end-alkylated DNA in the same solutions. We find that the observed separation behavior is best characterized by the Ogston sieving theory, where the free volume available to a DNA of length  $L$  for electrophoresis in the matrix is the only factor that sets the mobility of the DNA fragments.

## **5.2 – Materials and Methods**

### **5.2.1 – Reagents**

PicoGreen dye and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). Benzyl alcohol and 10x Tris-Borate-EDTA (TBE) were purchased from Sigma Aldrich (St. Louis, MO).  $C_{12}E_5$  surfactant was purchased from BaChem (Torrance, CA). 2-Log DNA ladder was purchased from New England Biolabs (Ipswich, MA). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2m $\Omega$  deionized water and filtering through a 0.22 $\mu$ m filter.

### **5.2.2 – Lower Critical Solution Temperature (LCST)**

#### **Determination**

The LCST was determined using a Cary300 UV-Vis (Agilent Technologies, Santa Clara, CA). Each surfactant concentration was heated in a cuvette at 1°C/min, and its 488nm absorbance measured every 0.5°C. A sharp increase in absorbance indicated a more turbid sample, and therefore the LCST. The LCST of

C<sub>12</sub>E<sub>5</sub> in 1xTBE was found to be 34°C. This agrees well with a literature value of 33°C.<sup>[14], [17]</sup> The LCST was largely independent of surfactant concentration.

### **5.2.3 – Viscosity Measurements**

Viscosities of C<sub>12</sub>E<sub>5</sub> surfactant solutions ranging from 0.0 wt% to 3.25 wt% were measured via capillary viscometry using a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for UV detection. Higher wt%s resulted in viscosities that were too high (>10cP) to accurately measure via capillary viscometry. The capillary was a 50µm ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) with a length of 20cm to the detector and 30cm total. A small plug of concentrated, UV absorbing benzyl alcohol was injected into the capillary filled with surfactant solution. A pressure of 1 psi was used to hydrodynamically push the plug past the detector. The viscosity of the surfactant solution was determined using the elution time of the benzyl alcohol plug and assuming Hagen-Poiseuille plug flow.

### **5.2.4 – Capillary Electrophoresis**

Capillary electrophoresis (CE) experiments were performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. The excitation source was a 3mW 488nm Argon ion. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 30µm ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and a length of 20cm to the detector and 30cm total. 4kV for 30sec electrokinetic injections were used to introduce samples into the capillary. The capillary was cooled to 20°C using a built in coolant jacket system. At 25°C, the separation failed, believed to be due to

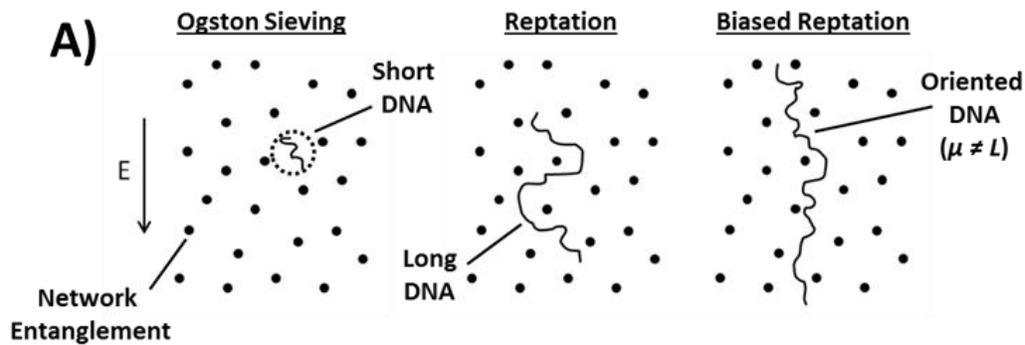
operating close to the LCST of C<sub>12</sub>E<sub>5</sub>. Data collection was performed using 32 Karat software (Beckman Coulter). Elution peaks were analyzed using OriginPro 9.0 (Northampton, MA).

The 2-Log DNA ladder was prepared for separation by diluting to 10µg/mL in 1xTBE, and 1/400<sup>th</sup> dilution of stock PicoGreen dye. Electrophoretic separation was conducted under an applied voltage of 1kV in reverse polarity mode (from anode to cathode,  $E = 33 \text{ V/cm}$ ). A 1/400<sup>th</sup> dilution of stock PicoGreen was added to the running buffer to ensure the DNA was stained during electrophoresis. We have found that the addition of PicoGreen dye does not greatly affect the electrophoretic migration of DNA.

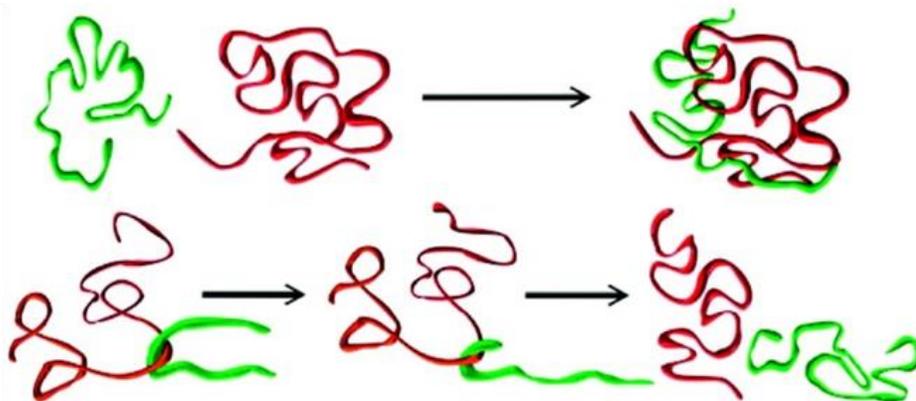
Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer 1xTBE solution and maintained by adding 0.1v/v% POP-6 to all capillary electrophoresis buffers. Capillary electrophoresis buffers were prepared daily by mixing C<sub>12</sub>E<sub>5</sub> in 1xTBE, briefly vortexing, and shaken for at least 1 hour at room temperature to ensure surfactant solubilization and micelle formation. The capillary could be used for many separations without observing any detrimental effects.

### **5.3 – Separation Theory**

The three main separation mechanisms of DNA electrophoresis in physically entangled polymer solutions are depicted in Figure 5.1. This section will detail the underlying theory behind each mechanism.



## B) Transient Entanglement Coupling



**Figure 5.1** – Schematic representations of gel electrophoresis DNA separations. A) Ogston sieving, reptation, and biased reptation through network entanglements. B) TEC of DNA (green) colliding with an individual polymer (red), adapted from [18]

### 5.3.1 – Ogston Sieving

When the size of a migrating DNA length  $L$  is smaller than the average pore size ( $\zeta_p$ ) within a gel, the DNA is thought to separate by an Ogston sieving mechanism. The DNA can traverse the pore space without changing from its globular conformation. This separation mechanism operates under the principle that the pores reduce the free volume available ( $V_A$ ) for DNA electrophoresis within the total volume of the capillary ( $V_T$ ). The electrophoretic mobility ( $\mu$ ) of DNA is equal to the free solution mobility of DNA ( $\mu_0$ ) in  $V_A$ , and is 0 in gel

occupied volume,  $V_T - V_A$ . Thus,  $\mu$  is solely dependent on the volume fraction of gel ( $f$ ) that the DNA can enter.

$$\mu = \mu_0 \frac{V_A}{V_T} = \mu_0 f \quad (5.1)$$

Ogston derived a pore distribution model for a random array of long, thin polymer fibers that calculates the volume fraction of pores that can admit a sphere with radius  $R$ .<sup>[19]</sup> Ogston's model was extended by Rodbard and Chramback to include other molecular conformations rather than just spheres.<sup>[20]</sup> In the case of DNA, the DNA fragment is assumed to maintain an unperturbed globular conformation with effective radius  $R_g$ .<sup>[19], [21]</sup>

$$f = \exp\left(-\pi\nu l(r + R_g)^2\right) \quad (5.2)$$

Where  $\nu$  is the average number of polymer fibers per unit volume of gel,  $l$  is the average length of fiber, and  $r$  is the fiber radius. Realizing that  $\nu l$  is proportional to the concentration of polymer  $C$  and substituting Equation (5.2) into (5.1), a simple relationship for  $\mu$  can be made.

$$\ln(\mu) = \ln(\mu_0) - K_R C \quad (5.3)$$

$K_R$  is known as the retardation coefficient and is a constant based on the shape and size of the migrating DNA fragment. Ferguson first noted the logarithmic dependence of protein mobility vs. the concentration of starch gels following Equation 5.3.<sup>[22]</sup> For this, Ferguson plots ( $\log \mu$  vs.  $C$ ) gained their name, yielding linear fits with an intercept of  $\mu_0$  and slope of  $K_R$ . The Ferguson plot can be used

to calculate  $\xi_p$  by extrapolating the linear fits to the point where  $\mu$  is half of  $\mu_0$ .<sup>[23]</sup>

<sup>[24]</sup> When  $R_g = \xi_p$ , the DNA length  $L$  should be able to access half the available volume in the capillary. Therefore, at  $\mu/\mu_0 = 0.5$  the  $\xi_p$  should be equal to the  $R_g$  of the migrating DNA.

### 5.3.2 – Reptation

When the DNA fragment is too large to fit through a pore of size  $\xi_p$ , the DNA will uncoil in order to enter the pore space. The connectivity of pores creates a tube that topologically constrains the DNA and it is forced to wriggle through head first. The DNA will move through the tube following a reptation mechanism like that proposed by de Gennes for diffusion of long polymers in entangled solutions.<sup>[25]</sup> The major differences between de Gennes reptation and DNA reptation is that electrical forces cause biased DNA motion in the direction of the electric field. The electrophoretic mobility depends on the projection of DNA orientated in the field direction, taken as a ratio of the DNA's total length  $L$  and mean squared end-to-end distance of DNA confined in the tube,  $\langle h_x^2 \rangle$ .<sup>[26], [27]</sup> For weak fields, the DNA maintains Gaussian chain statistics, such that  $\langle h_x^2 \rangle \sim L$ , and the reptation model predicts that the mobility is inversely proportional to the DNA length.

$$\mu = \mu_0 \frac{\langle h_x^2 \rangle}{L^2} \cong \frac{1}{L} \quad (5.4)$$

This ratio manifests itself from the tortuous path the tube creates where the true velocity of DNA within the tube ( $v_{tube}$ ) is greater than the observed velocity through the entire capillary.<sup>[28]</sup>

For strong fields,  $\langle h_x^2 \rangle$  becomes oriented in the direction of the electric field and separation is no longer possible. A scaled electric field,  $\varepsilon$ , is used to define the balance of electrical energy orientating the segment within a pore versus its thermal energy.

$$\varepsilon = \frac{qE\xi_p}{2k_bT} \quad (5.5)$$

Where  $q$  is the charge per DNA segment in a pore, and  $E$  is the electric field strength. A biased reptation model (BRM) has been defined that accounts for orientation as

$$\mu = \mu_0 \left( \frac{1}{3L} + \frac{\varepsilon^2}{9} \right) \quad (5.6)$$

The BRM predicts two regimes of DNA separation.

$$\frac{\mu}{\mu_0} \cong \begin{cases} \frac{1}{3L} & , \quad L < \varepsilon^{-2} \\ \varepsilon^2 & , \quad L > \varepsilon^{-2} \end{cases} \quad (5.7)$$

For short lengths or low  $E$ , electrical forces are not strong enough to orientate the DNA within the tube and the expected inverse length behavior is observed. For large  $L$  or high  $E$ , the electric forces orientate DNA's conformation within the tube. This creates a situation where  $h_x = L$  and the mobility is no longer a function

of length, just the constant  $\varepsilon^2$ . This biased reptation with orientation regime is the major limitation of DNA reptation as size separations are limited to small DNA lengths and low  $E$ .

An important time scale presents itself as the time it takes a DNA of length  $L$  to reptate a distance  $\langle h_x^2 \rangle$  in the direction of the electric field, known as the tube renewal time  $\tau_{rep}$ .

$$\tau_{rep} = \frac{L}{v_{tube} \langle h_x^2 \rangle} \quad (5.8)$$

Duke and Viovy first described the effect of “living polymers” on the constraint of reptating DNA.<sup>[16]</sup> In this model, the polymer chains that compose the network are in constant thermal motion and individual chains diffuse by reptation. The topological constraints on electrophoresing DNA are constantly changing as polymer entanglements are relieved and new ones are formed in the temporary gel, on a timescale known as the dissociation rate  $w$ . When  $w \gg \tau_{rep}^{-1}$ , the entanglements dissociate faster than DNA reptation through the tube. The DNA is not confined by the entanglements and the DNA moves through as if it were in free solution ( $\mu = \mu_0$ ). When  $w \ll \tau_{rep}^{-1}$ , the entanglements persist longer than the DNA reptation time scale and the polymer network acts like a permeant gel ( $\mu \propto L^{-1}$ ).

An interesting limit arises when  $w$  is slow enough to constrain the DNA as it reptates through the tube, but still rapid enough to interfere with the molecular orientation timescale of DNA,  $w\tau_{rep} > \varepsilon^2$ . Longer lengths of DNA may be

separated in these temporary gels versus permanent gels as the biased orientation limit can be extended to higher lengths by eliminating orientation.<sup>[16]</sup>

$$\frac{\mu}{\mu_0} \cong \begin{cases} 1 & , \quad 1 < w\tau_{rep} \\ L^{-1} & , \quad \varepsilon^2 < w\tau_{rep} < 1 \\ w\tau_{rep} & , \quad w\tau_{rep} < \varepsilon^2 \end{cases} \quad (5.9)$$

The  $w\tau_{rep}$  for the physical entangled polymer system is defined as,

$$w\tau_{rep} \sim \left( \frac{c}{c^*} \right)^{-15/4} \frac{\xi_p}{b_k} \quad (5.10)$$

Where  $c$  is the polymer concentration,  $c^*$  is the polymer overlap concentration,  $\xi_p$  is the average distance between entanglements or screening length, and  $b_k$  is the Kuhn length of DNA. This expression represents the limit of biased orientation within the dynamic polymer system.

An important distinction between the polymer chains and wormlike micelles is that the micelles are dynamic as they break and reform reversibly. Cates *et al.* defined the kinetic process of chain breakage with rate constant  $k_M$  into two segments with a mean length  $L_{av}$  as  $\tau_{break}$ .<sup>[13]</sup>

$$\tau_{break} = \frac{1}{k_M L_{av}} \quad (5.11)$$

Accounting for both micelle reptation and breakage, they define a lumped time parameter that relates micelle entanglement lifetimes,  $\tau_M$ , as

$$\tau_M \cong (\tau_{rep} \tau_{break})^{1/2} \quad (5.12)$$

Yeung *et al.* applied Cates theory to Duke and Viovy's temporary gel model to define the upper limit for size separation of DNA in micelle entangled networks.<sup>[15]</sup> For  $w = \tau_M^{-1}$ , the new limit of temporary constant becomes

$$w\tau \sim \left(\frac{c}{c^*}\right)^{-15/4} \frac{\xi_p^{8/3}}{b} \left(\frac{\eta k_M L_{av}}{kT}\right)^{1/2} \quad (5.13)$$

What is important to note is that regardless of the network properties, from static, polymer reptation, or micelle breakage, DNA reptation is always subjected to an upper limit for biased orientation within the polymer constraints.

### 5.3.3 – Transient Entanglement Coupling / Post Collisions

The third and most recent mechanism proposed for electrophoretic DNA separation is transient entanglement coupling (TEC). TEC was first observed in dilute systems of polymer ( $c < c^*$ ) and arises from DNA colliding with individual polymer molecules.<sup>[9], [29], [30]</sup> The DNA drags the polymer through solution, thus increasing the frictional force acting on the DNA. The DNA-polymer conjugate remain coupled until the two molecules slide off one another.

TEC operates in a similar manner to end-labeled free-solution electrophoresis (ELFSE).<sup>[31]</sup> Both systems utilize interaction with uncharged molecules to increase drag on the electrophoresing DNA. However, ELFSE is limited to specific interaction with a single molecule while the number of polymer-DNA interactions is length dependent in TEC. This length based interaction leads to a TEC elution profile where the shorter lengths move faster than longer ones.

The average velocity of DNA undergoing TEC ( $\bar{v}_{TEC}$ ) is given by a similar expression to ELFSE

$$\bar{v}_{TEC} = \frac{q_{eff} LE}{\varepsilon_{DNA} L + \varepsilon_P \bar{P}} \quad (5.13)$$

Where  $q_{eff}$  is the charge of a DNA base,  $\varepsilon_{DNA}$  is the friction coefficient of a DNA base,  $\varepsilon_P$  is the friction coefficient of an entangled polymer, and  $\bar{P}$  is the average number of polymers being dragged.<sup>[30]</sup>  $\bar{P}$  is a function of the time constant for DNA-polymer interaction,  $\tau_C$ , total separation time,  $t$ , and the number of DNA-polymer collisions that occur during the separation,  $N_C$ .<sup>[30]</sup>

$$\bar{P} = \frac{\tau_C N_C}{t} \quad (5.14)$$

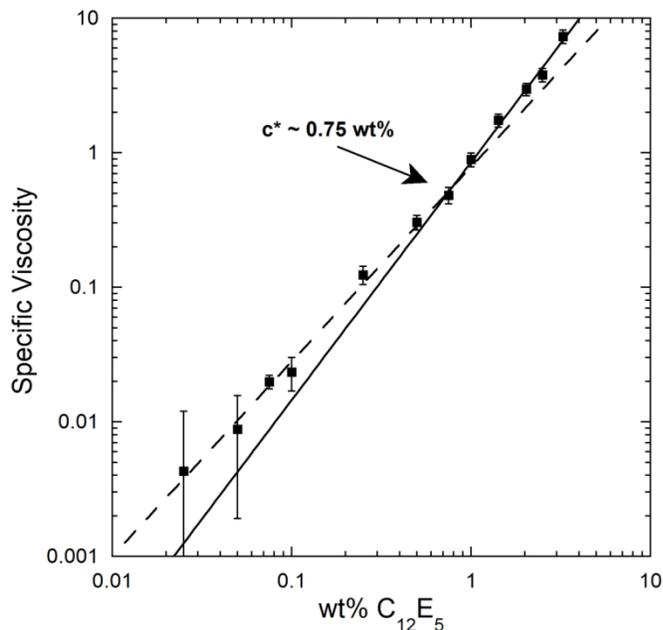
$N_C$  can be determined from collision theory, where the DNA and polymer compose a cylindrical space for interaction based of their respective  $R_g$ s.

$$N_C = C_N \pi (R_{g,DNA} + R_{g,P})^2 l_c \quad (5.15)$$

Where  $C_N$  is the number concentration of polymer and  $l_c$  is the length of the capillary. Note that  $\bar{P}$  contains a velocity term  $l_c/t$ , creating an implicit expression for  $\bar{v}_{TEC}$  that can be solved using the quadratic equation. Thus, the only fitting parameter in the TEC system is  $\tau_C$ .

## 5.4 – Results and Discussion

### 5.4.1 – Determination of $c^*$



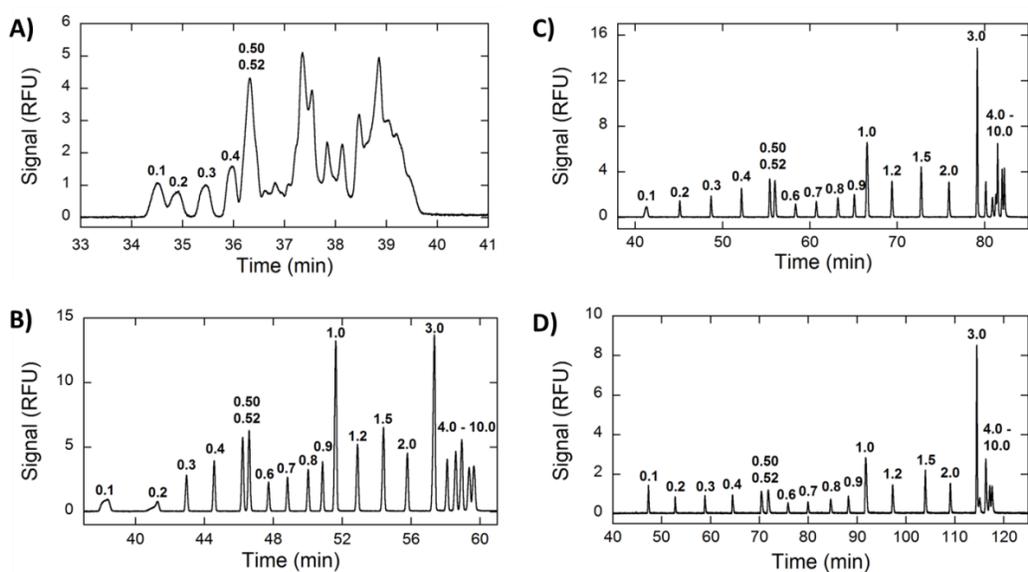
**Figure 5.2** – Specific viscosity of C<sub>12</sub>E<sub>5</sub> surfactant in 1xTBE buffer at 20°C. Dashed line represents linear fit to data below  $c^*$ . Solid line represents linear fit to data above  $c^*$ . Error bars represent standard deviation with  $n = 3$ . The intersection of these two linear fits marks the  $c^*$ .

As the concentration of C<sub>12</sub>E<sub>5</sub> surfactant increases, the wormlike micelles will grow in size and number until individual micelles are no longer hydrodynamically segregated from one another. This concentration is known as the overlap concentration ( $c^*$ ). Above  $c^*$ , the micelle solution is similar to an entangled polymer solution used to separate DNA by conventional gel/polymer matrix electrophoresis.<sup>[15]</sup> The specific viscosity is plotted against C<sub>12</sub>E<sub>5</sub> concentration (wt%) on a log-log scale in Figure 5.2. Two separate linear regimes are observed for the low and high concentrations.  $c^*$  is calculated as the concentration where the two regimes intersect, measured as 0.75 wt%.<sup>[8], [32]</sup> This value agrees well

with that reported by Grosswaser *et al.*, who reported a value of 1.08 wt% using cryo-TEM imaging of  $C_{12}E_5$  wormlike micelles in water at 18°C. [14]

### 5.4.2 – Separation of DNA in $C_{12}E_5$ Solutions

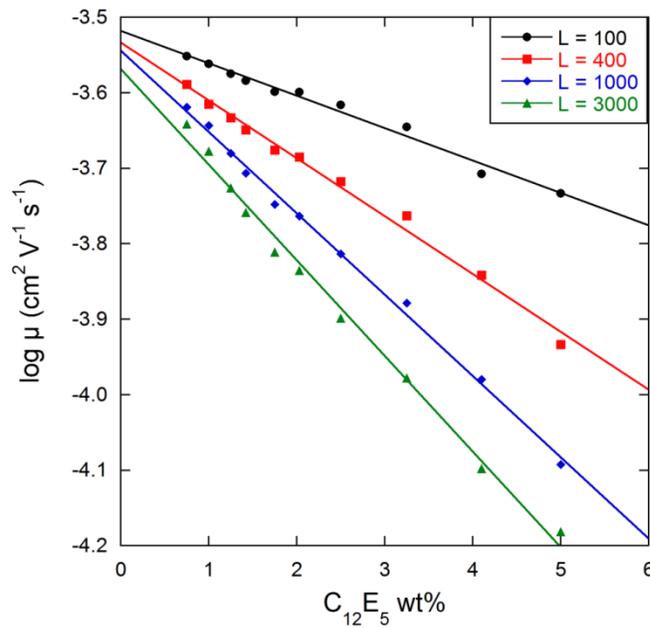
Figure 5.3 shows representative electropherograms of a commercially available DNA ladder separated in different  $C_{12}E_5$   $c/c^*$  buffers. As  $c/c^*$  increases, the wormlike micelles form a more entangled network. The more entangled networks contribute greater amounts of frictional constraint to the migrating DNA molecules and results in better resolved electrophoretic separations with increased elution times. In order to elucidate the separation mechanism of unmodified DNA in entangled networks of  $C_{12}E_5$  wormlike micelles, the three theories of gel electrophoresis presented above were fit to the data.



**Figure 5.3** – Representative electropherograms showing the separation of the unmodified 2-Log DNA ladder at different  $C_{12}E_5$   $c/c^*$  ratios in 1xTBE. DNA lengths are indicated above peaks in kilobases.  $c/c^*$  = A) 0.67, B) 1.90, C) 3.33, D) 6.0.

### 5.4.3 – Ogston Sieving Evaluation

A Ferguson plot ( $\log \mu$  vs.  $C_{12E5}$  concentration) produced from the unmodified DNA separations is seen in Figure 5.4. The Ferguson plot clearly shows a linear dependence of  $\log \mu$  with  $C_{12E5}$  concentration for multiple DNA lengths, demonstrating adherence to Ogston sieving theory. Following Equation 5.3, the free-solution mobility of DNA ( $\mu_0$ ) and retardation coefficient ( $K_r$ ) can be calculated from the intercept and slope of the linear fit for each DNA length, respectively. We find excellent agreement with literature values for  $\mu_0$  in 1xTBE as  $2.84 \times 10^{-4} \pm 9.1 \times 10^{-6} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ .<sup>[33]</sup>



**Figure 5.4** – Ferguson plots for different lengths of DNA based on  $C_{12E5}$  wt%. Lines represent linear regression fits to the data according to Equation 5.3.

Following Ogston’s prediction in Equation 5.2,  $K_R^{1/2}$  should vary linearly with  $R_g$ , as seen in Figure 5.5. The radius of gyration ( $R_g$ ) of DNA is calculated using the Krakty-Porod model for a worm-like chain, Equation 5.16, using the

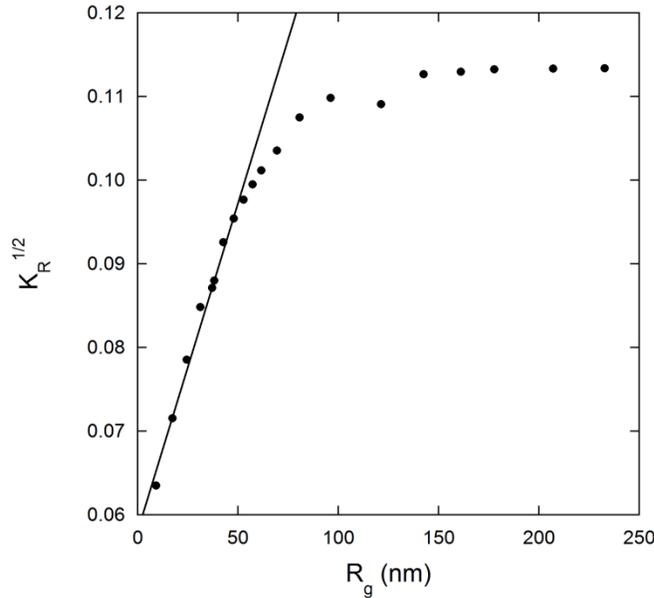
persistence length ( $p$ ) of double stranded DNA as 50nm and monomer length ( $b$ ) of 0.34nm.

$$R_{g,DNA}^2 = \frac{pLb}{3} \left[ 1 - 3 \left( \frac{p}{Lb} \right) + 6 \left( \frac{p}{Lb} \right)^2 - 6 \left( \frac{p}{Lb} \right) \left( 1 - \exp \left( \frac{-Lb}{p} \right) \right) \right] \quad (5.16)$$

As seen in Figure 5.5, smaller DNA fragments agree with Ogston's prediction while larger fragments show significant deviation. This deviation typically represents the transition from Ogston sieving to reptation, when  $R_g \sim 1.4 \xi_p$ .<sup>[23], [34]</sup> From this approximation, we can estimate  $\xi_p$  as 30.5nm for the C<sub>12</sub>E<sub>5</sub> concentrations studied. This is in good agreement with Kato *et al.* DLS results for C<sub>12</sub>E<sub>5</sub> mesh size of ~15nm.<sup>[35]</sup> However, our measurements were made at a finite electric field strength ( $E = 33$  V/cm). Previous studies have shown that decreasing  $E$  to the zero field limit yields more linear fits of  $K_R^{1/2}$  vs.  $R_g$ , rather than transitions to reptation.<sup>[23], [36]</sup> This dependence is believed to be caused by the deformation of larger DNA fragments as they approach pores that would normally be too small for them to pass through. The transition may also indicate the assumption made by Ogston that relates  $K_R^{1/2}$  to  $R_g$  is not the correct functional form of  $f$ . Additional terms may be required to accurately describe the shape and size of a DNA fragment migrating through the wormlike micelle network, as described by Rodbard and Chramaback.<sup>[20]</sup>

The Ferguson plot can be used to calculate  $\xi_p$  by extrapolating the linear fits to the point where  $\mu$  is half of  $\mu_0$ .<sup>[23], [24]</sup> When  $R_g = \xi_p$ , the DNA length  $L$  should be able to access half the available volume in the capillary. Therefore, at  $\mu/\mu_0 = 0.5$  the  $\xi_p$  should be equal to the  $R_g$  of the migrating DNA. A plot of  $\xi_p$  vs. C<sub>12</sub>E<sub>5</sub>  $c/c^*$

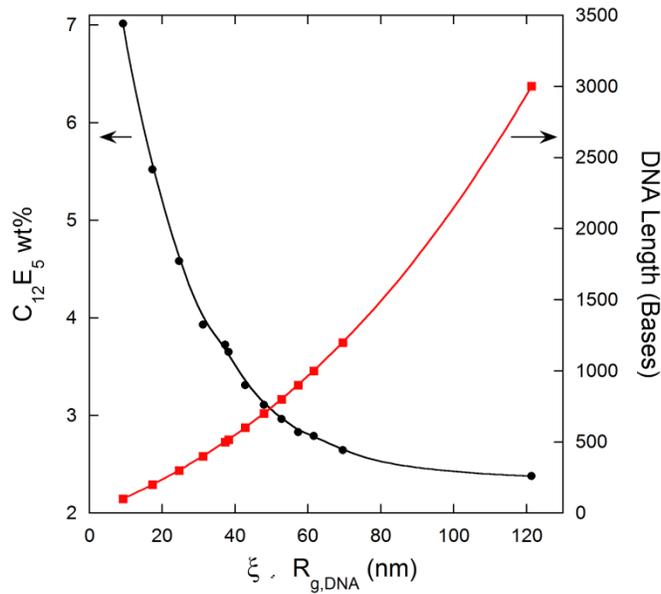
is seen in Figure 5.6. For comparison, the pore sizes for agarose gels are typically on the order of several hundred nms and can effectively separate DNA from 100 base pairs to 10,000 base pairs.<sup>[23], [37]</sup> Polyacrylimide gels are more suited towards separation of 10 – 1,000 base pairs of DNA with reported pore sizes in the tens of nms.<sup>[24]</sup>



**Figure 5.5** – Plot of  $K_R^{1/2}$  vs.  $R_g$  of DNA for data taken at  $E = 33$  V/cm. Deviation from linear fit may indicate a transition from Ogston sieving to reptation.

A sharp decrease in  $\zeta_p$  is observed as  $C_{12}E_5$  wt% increases, due to the greater number of entanglements present in the solution at higher concentrations. The  $\zeta_p$  values calculated in Figure 5.6 are smaller than many of the DNA lengths electrophoresing through the entangled system, violating one of the key assumptions for Ogston sieving. Only DNA < 600 bases has a smaller  $R_g$  than the pore size at 5wt%  $C_{12}E_5$ , the same length when we observe deviation in Figure 5.5. However, the Ferguson plots are linear for DNA lengths up to 10,000 bases. The discrepancy between Ogston sieving theory and our system is most likely

from the dynamic nature of the micelle network. The pores formed by micelle entanglements are most likely short lived (the micellization time of  $C_iE_j$  surfactants is on the order of milliseconds) and may deform in the presence of an electrophoresing DNA fragment.<sup>[38]</sup> The larger DNA lengths may still pass through the pores unperturbed from their globular conformation as the local micelle network “opens up” up for the larger DNAs.

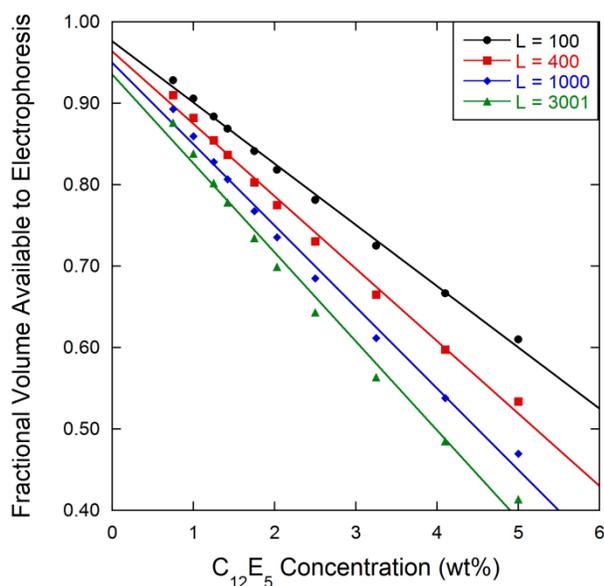


**Figure 5.6** – Effective wormlike micelle network pore size as a function of  $C_{12}E_5$  wt% calculated via extrapolation of Ferguson plot data according to Equations 5.3. Red squares represent DNA  $R_g$  calculated by the WLC model Equation 5.16.

If we are to believe that the DNA migrates under an Ogston sieving method, the only parameter that sets DNA elution time is  $f$  described in Equation 5.1 for a DNA of length  $L$  in a concentration of surfactant  $c$ . In terms of  $K_r$ ,  $f$  follows Equation 5.17.

$$f = \exp(-K_r C) \quad (5.17)$$

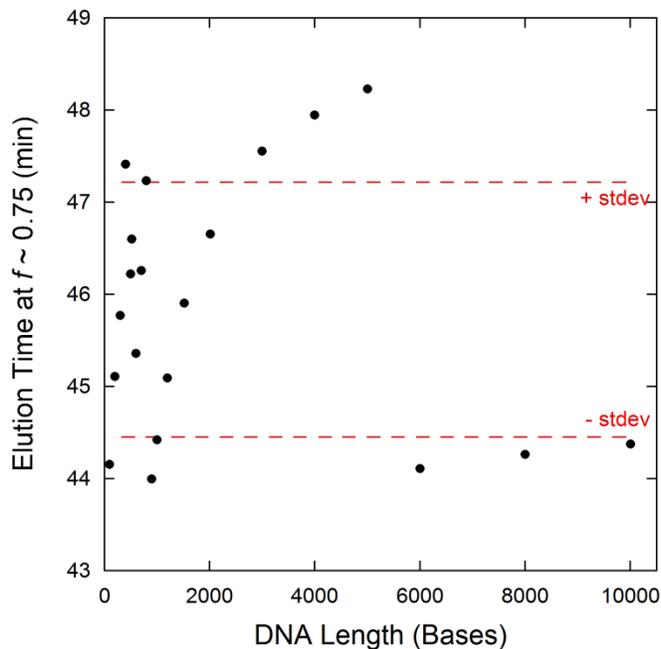
Figure 5.7 shows the trend of  $f$  vs.  $c$  for different lengths of DNA. Longer lengths of DNA occupy larger volumes and have smaller  $f$  for the same surfactant concentration. If two lengths of DNA have the same  $f$  at different surfactant concentrations, they should have the same elution time according to Ogston sieving theory.



**Figure 5.7** – Calculated fractional volume available to electrophoresis vs.  $C_{12}E_5$  concentration.

Figure 5.8 shows the elution time for all lengths of DNA at an  $f$  near 0.75, calculated by Equation 5.17 for different concentrations of  $C_{12}E_5$ . For example, the elution time of the 100 base DNA at  $c = 3\text{wt}\%$  should match that of the 10,000 base DNA at  $c = 0.79\text{wt}\%$ , both at  $f = 0.75$ . We see excellent agreement with the elution times for all lengths at  $f = 0.75$ , eluting at 45.8 min +/- 1.4min. The elution window would appear even narrower if we had access to the exact concentrations that yield  $f = 0.75$ . For DNA lengths 3,000 – 5,000, the real  $f$  was ~

0.73, leading to a slightly longer elution time. For DNA lengths 6,000 – 10,000, the real  $f$  was  $\sim 0.79$ , leading to slightly shorter elution times.

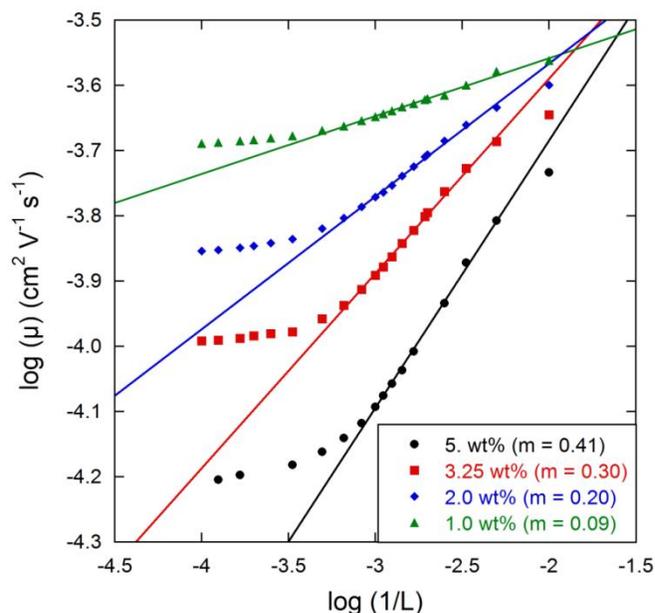


**Figure 5.8** – Elution time for different DNA lengths and concentrations all at  $f = 0.75$

#### 5.4.4 – Reptation Evaluation

Although Ferguson plots and fractional volume arguments appear valid for Ogston sieving, deviations from linear  $K_R^{1/2}$  vs.  $R_g$  plots and estimations of  $\xi_p < R_g$  leave room for doubt. We test the theory of reptation presented in 5.3.2 against our C<sub>12</sub>E<sub>5</sub> separation results. A reptation plot ( $\log(\mu)$  vs  $\log(I/L)$ ) was created where a slope of 1 would indicate reptation without orientation. Figure 5.9 shows that none of the concentrations tested adhere to reptation behavior with a slope of 1, however the slopes do increase towards 1 with increasing concentration. The less than predicted slope may be caused by the greater fraction of time spent migrating in a faster moving globular conformation, consistent with Ogston

sieving.<sup>[8], [36]</sup> The increasing slope may indicate that the DNA molecules have to deform to more oriented conformations with increasing surfactant concentration to pass through pore spaces, but never entering a regime of head first reptation through the network. Additionally, at higher lengths, the data deviates from linear behavior. This deviation is typically believed to be the onset of biased orientation as the mobility becomes length invariant. Since we do not appear to enter a fully reptative regime, orientation appears unlikely. It may represent the point at which longer DNA lengths have to deform more greatly to squeeze through network pores.



**Figure 5.9** – Reptation plots,  $\log \mu$  vs.  $\log (1/L)$ , at different concentration of  $C_{12}E_5$  for unmodified DNA. Straight line represent linear fits to lengths of DNA possibly undergoing reptation, where a slope of 1 would be predicted by  $\mu \propto 1/L$ .

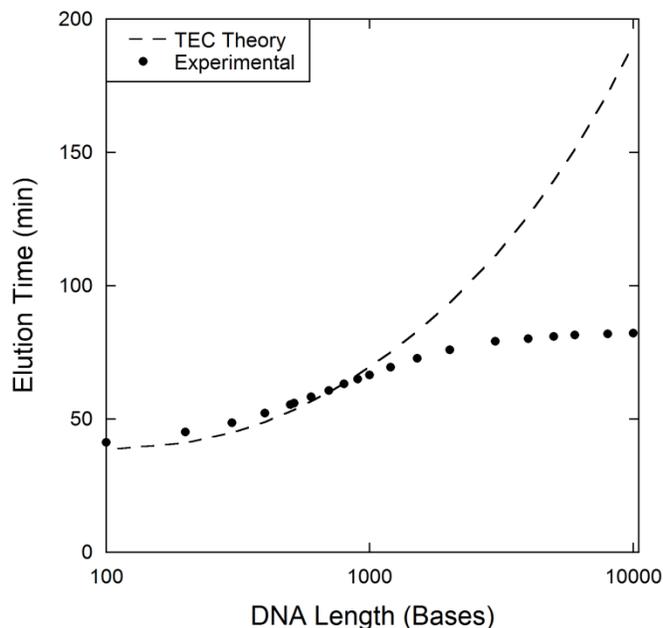
### 5.4.5 – TEC Comparison

We also compare our results to the TEC theory outlined in section 5.3.3. TEC is believed to occur in dilute polymer solutions ( $c < c^*$ ), where an electrophoresing

DNA drags individual polymer chains through solution. We do not observe DNA separations until  $c^*$  is reached, with more resolved separations occurring at higher concentrations. This observation would seem to contradict TEC assumptions. However, Chiesl *et al.* recently used single molecule video microscopy to directly observe the transition from TEC to reptation in increasing concentrations of physically entangled polymer networks.<sup>[18]</sup> Their work demonstrated that TEC is still the main component of DNA separation even at  $c = 10c^*$  for low MW hydroxyethyl cellulose. They demonstrated that only at concentrations well above  $c^*$  do polymers begin to statistically have entanglements with one another, defined as the entanglement concentration  $c_e$ . This physically makes sense because the definition of  $c^*$  is when the polymers just begin sharing occupied volumes, not necessarily interacting with one another. Polymer-polymer interactions are low at  $c^*$  and require much greater concentrations in order to begin to entangle.

Figure 5.10 shows a fit of our experimental separations at 2.5wt% C<sub>12</sub>E<sub>5</sub> to the model given in 5.3.3. For this model, we used the  $MW$  and  $R_H$  for individual C<sub>12</sub>E<sub>5</sub> micelles as determined by SLS and DLS data from Shirai and Einaga to calculate  $C_N$  and  $\varepsilon_p$  by stokes drag, respectively.<sup>[17]</sup> We assumed a  $q_{eff}$  to be  $0.066e$ .<sup>[30], [39]</sup> We calculate  $\varepsilon_{DNA}$  as  $\varepsilon_{DNA} = q_{eff}/\mu_0$  with  $\mu_0 = 2.7 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . Thus,  $\tau_C$  is the lone fitting parameter in the model, which we solve for using the experimental values for average velocities as  $0.82 \mu\text{s}$ . This value agrees well the  $\tau_C$  calculated for cellulose based polymers,  $1.38 - 0.58 \mu\text{s}$ .<sup>[30]</sup>

We can see good agreement with shorter lengths and the TEC model, however significant deviations occur at longer DNA lengths. The TEC model predicts a much greater retardation of the longer DNA's velocity than observed, mainly due to the dramatic increase in collision theory with increasing  $R_g$ . TEC does not appear to predict unmodified DNA behavior in the  $C_{12}E_5$  networks.



**Figure 5.10** – Predicted elution time by TEC theory (dotted line) and actual elution time (black circles) vs. DNA length for 2.5wt%  $C_{12}E_5$  separation buffer.

## 5.5 – Conclusion

We have described three DNA gel separation mechanism theories, Ogston sieving, reptation, and TEC. We have demonstrated the successful separation of unmodified DNA in solutions of  $C_{12}E_5$  surfactant above its overlap concentration. After fitting each separation mechanism model to our data sets, it appears that the DNA separation in  $C_{12}E_5$  network is best characterized by the Ogston sieving mechanism. The entangled micelle network limits the fractional volume available to the DNA for electrophoresis. Longer lengths are more retarded as they access a

smaller fraction of the capillary. Using this fractional volume approach, we will attempt to describe the migration of end-alkylated DNA through these micelle networks undergoing a hybrid sieving and ELFSE separation mechanism.

## 5.6 – References

- [1] H. Drossman, J. A. Luckey, A. J. Kostichka, J. D’Cunha, and L. M. Smith, *Anal. Chem.*, 1990, 62, 900–903.
- [2] H. Swerdlow and R. Gesteland, *Nucleic Acids Res.*, 1990, 18, 1415–1419.
- [3] D. N. Heiger, A. S. Cohen, and B. L. Karger, *J. Chromatogr.*, 1990, 516, 33–48.
- [4] H. T. Chang and E. S. Yeung, *J. Chromatogr. B. Biomed. Appl.*, 1995, 669, 113–123.
- [5] S. J. Kim, G. W. Shin, S. J. Choi, H. S. Hwang, G. Y. Jung, and T. K. Seo, *Electrophoresis*, 2010, 31, 1108–1115.
- [6] C. P. Fredlake, D. G. Hert, C. W. Kan, T. N. Chiesl, B. E. Root, R. E. Forster, and A. E. Barron, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, 105, 476–481.
- [7] A. E. Barron, W. M. Sunada, and H. W. Blanch, 1996.
- [8] P. D. Grossman and D. S. Soane, *Biopolymers*, 1991, 31, 1221–1228.
- [9] A. E. Barron, H. W. Blanch, and D. S. Soane, *Electrophoresis*, 1994, 15, 597–615.
- [10] S. Park, D. W. Olson, and K. D. Dorfman, *Lab Chip*, 2012, 12, 1463.
- [11] K. D. Dorfman, *Rev. Mod. Phys.*, 2010, 82, 2903–2947.
- [12] G. C. Randall and P. S. Doyle, *Macromolecules*, 2006, 39, 7734–7745.
- [13] M. E. Cates, *Macromolecules*, 1987, 20, 2289–96.
- [14] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [15] Wei and E. S. Yeung, *Anal. Chem.*, 2001, 73, 1776–1783.
- [16] T. Duke and J. L. Viovy, *Phys. Rev. E*, 1994, 49, 2408–2416.

- [17] S. Shirai and Y. Einaga, *Polym. J.*, 2005, 37, 913–924.
- [18] T. N. Chiesl, R. E. Forster, B. E. Root, M. Larkin, and A. E. Barron, *Anal. Chem.*, 2007, 79, 7740–7747.
- [19] A. G. Ogston, *Trans. Faraday Soc.*, 1958, 54, 1754–1757.
- [20] D. Rodbard and A. Chrambach, *Proc. Natl. Acad. Sci. U. S. A.*, 1970, 65, 970–977.
- [21] G. W. Slater and J. R. Treurniet, *J. Chromatogr. A*, 1997, 772, 39–48.
- [22] K. A. Ferguson, *Metabolism.*, 1964, 13, SUPPL:985–1002.
- [23] G. W. Slater, J. Rousseau, J. Noolandi, C. Turmel, and M. Lalanda, *Biopolymers*, 1988, 27, 509–524.
- [24] D. L. Holmes and N. C. Stellwagen, *Electrophoresis*, 1991, 12, 253–263.
- [25] P. G. de Gennes, *J. Chem. Phys.*, 1971, 55, 572.
- [26] O. J. Lumpkin, *Biopolymers*, 1982, 21, 2315–2316.
- [27] G. W. Slater, *Electrophoresis*, 2009, 30.
- [28] G. B. Salieb-Beugelaar, K. D. Dorfman, a van den Berg, and J. C. T. Eijkel, *Lab Chip*, 2009, 9, 2508–2523.
- [29] S. J. Hubert, G. W. Slater, and J. L. Viovy, *Macromolecules*, 1996, 29, 1006–1009.
- [30] W. M. Sunada and H. W. Blanch, *Electrophoresis*, 1998, 19, 3128–3136.
- [31] R. J. Meagher, J. I. Won, L. C. McCormick, S. Nedelcu, M. M. Bertrand, J. L. Bertram, G. Drouin, A. E. Barron, and G. W. Slater, *Electrophoresis*, 2005, 26, 331–350.
- [32] D. A. Hill and D. S. Soane, *J. Polym. Sci. Part B Polym. Phys.*, 1989, 27, 2295–2320.
- [33] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [34] G. W. Slater and J. Noolandi, *Biopolymers*, 1989, 28, 1781–1791.
- [35] T. Kato, S. Anzai, and T. Seimiya, *J. Phys. Chem.*, 1990, 94, 7255–7259.
- [36] D. L. Holmes and N. C. Stellwagen, *Electrophoresis*, 1990, 11, 5–15.

- [37] P. G. Righetti, B. C. Brost, and R. S. Snyder, *J. Biochem. Biophys. Methods*, 1981, 4, 347–363.
- [38] A. Patist, J. R. Kanicky, P. K. Shukla, and D. O. Shah, *J. Colloid Interface Sci.*, 2002, 245, 1–15.
- [39] J. M. Schurr and S. B. Smith, *Biopolymers*, 1990, 29, 1161–1165.



# **Chapter 6 – Free-solution Electrophoresis of End-alkylated DNA in Entangled Wormlike Micelle Networks**

## **6.1 – Introduction**

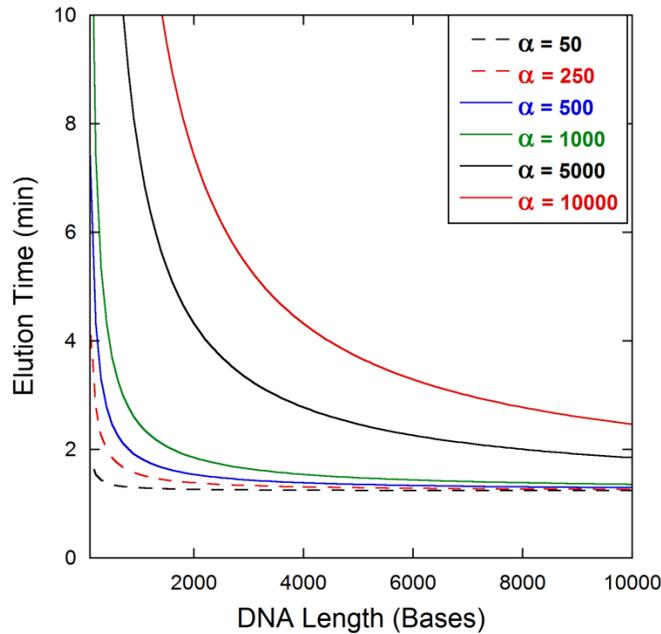
We have demonstrated a miRNA limit of detection (LOD) of 100pM by transitioning to oligomeric DNA nanotags stained with PicoGreen intercalating dyes. We view the extension of these DNA nanotags to thousands of bases long as a simple way to achieve fM detection limits in CE-LIF. The longer strands of DNA can accommodate a greater amount of intercalating dye, dramatically increasing fluorescence intensity. We have demonstrated an LOD of 10fM using  $\lambda$ DNA (48,502 base pairs) with YOYO-1 intercalating dye and optimized CE injections. Thus, separation of miRNA hybridized to long, kilobase (kB) DNA fragments is critical towards our effort for fM detection sensitivity.

However, separation of kB DNA by our assay is rather difficult. The separation of bound and unbound hybridization probes is achieved using non-ionic surfactant micelles as hydrodynamic “drag-tags”, a technique known as end-labeled free-solution electrophoresis (ELFSE).<sup>[1]-[3]</sup> As the DNA nanotag length increases, the additional friction from the attached micelle becomes negligible, and the sandwich miRNA complex returns to the inseparable free-solution DNA mobility limit.<sup>[4]</sup> This effect is illustrated in Figure 6.1, where the elution time for different

lengths of DNA nanotag is plotted for different micelle sizes. The elution time ( $t$ ) of a miRNA sandwich complex-micelle conjugate can be calculated by Equation 6.1

$$t = \frac{l_d l_t}{\mu_0 V} \left( 1 + \frac{\alpha}{L} \right) = t_0 \left( 1 + \frac{\alpha}{L} \right) \quad (6.1)$$

Where  $l_d$  is the capillary length to the detector,  $l_t$  is the total capillary length,  $\mu_0$  is the free-solution mobility of DNA,  $t_0$  is the free-solution elution time of DNA,  $V$  is the applied potential,  $L$  is the length of DNA hybridized to the miRNA, and  $\alpha$  is the additional drag of the micelle in terms of hydrodynamically equivalent DNA bases. It is clear from Figure 6.1 that in order to achieve appreciable differences in elution time between kB long DNA hybridization probes, an  $\alpha$  of at least 1,000 is required.



**Figure 6.1** – Effect of increasing micelle size ( $\alpha$ ) on elution time of miRNA complexes for different DNA nanotag lengths. Theoretical run conditions:  $l_d = 20\text{cm}$ ,  $l_t = 30\text{cm}$ ,  $V = 20\text{kV}$ ,  $\mu_0 = 2.7 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ .

A few groups have expanded the molecular conformations and chemistries of covalently bound ELFSE drag-tags in order to separate longer lengths of DNA.<sup>[5]-</sup><sup>[8]</sup> A general trend has seen the transition from globular streptavidin labels to linear polypeptides.<sup>[8]-[10]</sup> The linear conformation of the polypeptide sweeps a larger volume than the globular streptavidin, creating a greater amount of hydrodynamic drag on the electrophoresing DNA. We have previously demonstrated the use of non-ionic Triton X-100 surfactant globular micelles as successful ELFSE drag-tags.<sup>[1]-[3], [11]</sup> With the  $R_h = 2.5\text{nm}$  globular Triton X-100 micelles, we have achieved an  $\alpha = 58$ , clearly too small for separation of kB long DNA hybridization probes.<sup>[11]</sup>

Following the ideals of globular to linear protein drag-tags, transitions to rod-like non-ionic surfactant micelles should benefit from similar increases in drag-tag size. *n*-alkyl polyoxyethylene ether surfactants ( $C_iE_j$ s) form very large micelle structures in aqueous solutions. Increasing the *i/j* ratio (more hydrophobic surfactant monomers) leads to one-dimensional growth of the wormlike micelles with  $\mu\text{m}$  contour lengths and  $\sim 3\text{nm}$  diameters.<sup>[12]-[15]</sup> They appear as ideal candidates for micelle ELFSE drag-tags for kB DNA separations. However, as the concentration of  $C_iE_j$  surfactant increases past the overlap concentration ( $c^*$ ), the micelles will begin to interact with one another leading to an entangled network.<sup>[16]-[19]</sup> These micelle networks mimic polymeric sieving matrices that have been used to electrophoretically separate DNA in a manner similar to gel networks.<sup>[19], [20]</sup> Wormlike micelle network formulation should present an upper limit for their application in micelle ELFSE separations, as the presence of

sieving matrices has been shown to be detrimental for ELFSE separation due to their countering separation mechanisms.<sup>[21]</sup>

Surprisingly, we find that ELFSE-like separations (longest DNA eluting first) of end-alkylated DNA up to 2,093 bases long are observed in  $C_iE_j$  surfactant wormlike micelle buffers well above  $c^*$ . We have demonstrated in Chapter 5 that these buffers form an entangled network that separates unmodified DNA based on free-volume arguments according to Ogston sieving. Here, we demonstrate an observed  $\alpha$  near 1,800 using the same entangled  $C_{12}E_5$  non-ionic surfactant buffers, a magnitude larger than any other  $\alpha$  ever reported for ELFSE separations.<sup>[6]–[9], [22]–[24]</sup> We find that ELFSE theory predicts separation performance up to DNA ~800 bases long. At higher lengths, peaks begin to compress due to the increased sieving pressures from the entangled micelle network. We present a modified ELFSE-sieving model based on DNA sieving theory that accurately captures the deviations from ELFSE. Although our modified theory suggests separation of DNA above 2kB are difficult in entangled micelle networks, our understanding of these large  $\alpha$  buffers has increased tremendously and insights into further optimization are offered.

## **6. 2 – Materials and Methods**

### **6.2.1 – Reagents**

Nuclease-free water, hydrochloric acid, acetic acid, GeneJET Gel Extraction Kits, and acetone were purchased from Fisher Scientific (Pittsburgh, PA). BODIPY<sup>®</sup> FL  $C_{16}$  (C16B), low melting temperature agarose, and POP-6 polymer

were purchased from Life Technologies (Grand Island, NY). Benzyl alcohol, DMSO, 2,2-Dithiodipyridine (DPDS), Lithium perchlorate, Triphenylphosphine (TPP), 4-(Dimethylamino)pyridine (DMAP), cetyltrimethylammonium bromide (CTAB), ethidium bromide, acetonitrile, urea, 10x Tris-Borate-EDTA (TBE), and triethanolamine (TEA) were purchased from Sigma Aldrich (St. Louis, MO). *N*-dodecylpentaerythylene (C<sub>12</sub>E<sub>5</sub>) surfactant was purchased from BaChem (Torrance, CA).  $\lambda$ DNA and LongAmp<sup>TM</sup> *Taq* DNA polymerase Kit were purchased from New England Biolabs (Ipswich, MA). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2m $\Omega$  deionized water and filtering through a 0.22 $\mu$ m filter. PCR forward primers and C<sub>6</sub>-amine 5' modified reverse primer were purchased with standard desalting purification from Integrated DNA Technologies (Coralville, IA). The 22 base Cy5-labeled oligomer was purchased with HPLC purification. Dried oligomers were resuspended in 10mM Tris 0.1mM EDTA buffer and used without further purification.

## 6.2.2 – DNA Alkylation Reaction

Alkylation of C<sub>6</sub>-amine 5' modified reverse primer with C16B was accomplished using a combination of methods outlined by Mukaiyama and Deratani, and closely followed a scheme first reported by Boutorine *et al.*<sup>[25]–[27]</sup>. The chemical structure of the end-alkylated DNA is seen in Figure 6.2. An aliquot of 10nmol 5'amine modified reverse primer was precipitated in the presence of 750nmol CTAB. The suspension was then dried by vacuum centrifugation under low heat for 30 minutes. Meanwhile, a mixture of 100nmol C16B, 5 $\mu$ mol DMAP,



### **6.2.3 – Lower Critical Solution Temperature (LCST)**

#### **Determination**

The LCST was determined using a Cary300 UV-Vis (Agilent Technologies, Santa Clara, CA). Each surfactant concentration was heated in a cuvette at 1°C/min, and its 488nm absorbance measured every 0.5°C. A sharp increase in absorbance indicated a more turbid sample, and therefore the LCST. The LCST of C<sub>12</sub>E<sub>5</sub> in 1xTBE was found to be 34°C. This agrees well with a literature value of 33°C.<sup>[16], [18], [28]</sup> The LCST was largely independent of surfactant concentrations tested.

### **6.2.4 – PCR Amplification of End-alkylated DNA Ladder**

Specific DNA lengths of end-alkylated DNA were amplified using LongAmp<sup>TM</sup> Taq DNA polymerase and λDNA template with the described end-alkylated reverse primer and a length setting forward primer. The end-alkylated DNA ladder included the following lengths of DNA: 86, 133, 189, 214, 265, 375, 421, 488, 551, 603, 647, 707, 787, 808, 847, 909, 986, 1165, and 2093. Each PCR reaction contained 1ng of λDNA template, 500nM each primer, 0.3mM each dNTP, 2.5 units of LongAmp<sup>TM</sup> Taq DNA polymerase, 1x LongAmp<sup>TM</sup> buffer, and enough water to 25μL. The reaction was denatured for 2min at 95°C and subjected to 35 cycles of denaturation at 95°C for 15s, primer annealing at 50°C for 30 s, and extension at 65°C for 50 seconds per 1000 bases of amplicon length, followed by a final hold for 10min at 65°C.

After thermal cycling, the PCR reaction mixture was directly loaded onto 1wt% agarose gels. Slab gel electrophoresis was performed at 4 V/cm for 2 hours. Gels were imaged using ethidium bromide staining and UV excitation on a BioDoc-It

Imaging System from Ultra-Violet Products (Cambridge, UK). Correct PCR product length bands were excised and purified using GeneJET Gel Extraction Kit spin columns, and eluted using the provided 10mM Tris 0.1mM EDTA buffer.

### **6.2.5 – Capillary Electrophoresis**

Capillary electrophoresis (CE) experiments were performed on a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. Excitation sources were from a 3mW 488nm Argon ion and 635nm solid state lasers. LIF detection was performed at 488/520nm and 635/675nm excitation/emission. The capillary was a 30 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and a length of 20cm to the detector and 30cm total. The end-alkylated DNA ladders were prepared by pipetting 2 $\mu$ L of each purified PCR product with 30 $\mu$ L of formamide, heating to 95 $^{\circ}$ C for 2 minutes to ensure denatured single stranded DNA products, and immediately cooled on ice. 4kV for 30sec electrokinetic injections were used to introduce samples into the capillary. Electrophoretic separation was conducted under an applied voltage of 30kV in reverse polarity mode (from anode to cathode). The capillary was cooled to 20 $^{\circ}$ C using a built in coolant jacket system. At 25 $^{\circ}$ C, the separation failed, believed to be due to operating close to the LCST of C<sub>12</sub>E<sub>5</sub>. Data collection was performed using 32 Karat software (Beckman Coulter). Peaks were analyzed using OriginPro 9.0 peak fitting software (Northampton, MA).

Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer and maintained by adding 0.1v/v% POP-6 to all capillary electrophoresis buffers. Surfactant buffers were prepared by mixing the indicated

concentration of C<sub>12</sub>E<sub>5</sub> in 1xTBE, briefly vortexing, and shaken for at least 1 hour at room temperature to ensure surfactant solubilization and micelle formation. The capillary could be used for many separations without observing any detrimental effects.

### **6.2.6 – Viscosity Measurements**

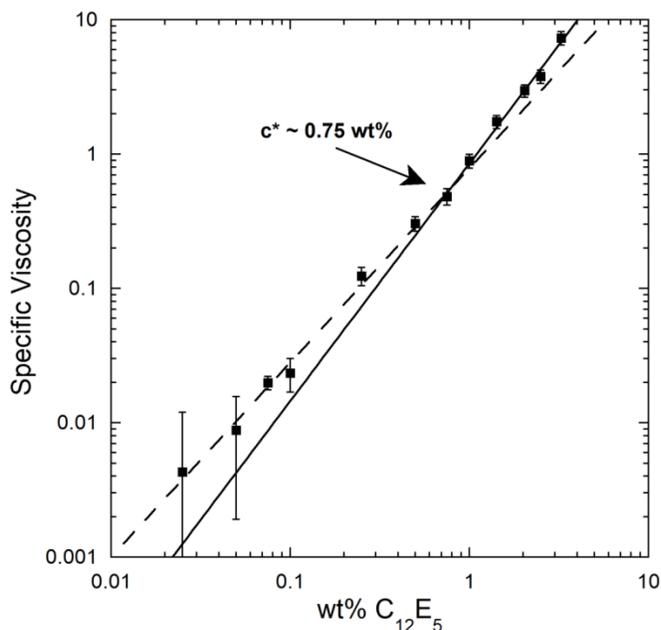
Viscosities of C<sub>12</sub>E<sub>5</sub> surfactant solutions ranging from 0.0 wt% to 3.25 wt% were measured via capillary viscometry using a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for UV detection. The capillary was a 50µm ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) with a length of 20cm to the detector and 30cm total. A small plug of concentrated, UV absorbing benzyl alcohol was injected into the capillary filled with surfactant solution. A pressure of 1 psi was used to hydrodynamically push the plug past the detector. The viscosity of the surfactant solution was determined using the elution time of the benzyl alcohol plug and assuming Hagen-Poiseuille plug flow.

## **6.3 – Results and Discussion**

### **6.3.1 – Determination of $c^*$**

As the concentration of C<sub>12</sub>E<sub>5</sub> surfactant increases, the wormlike micelles will grow in size and number until individual micelles are no longer hydrodynamically segregated from one another. This concentration is known as the overlap concentration ( $c^*$ ). Above  $c^*$ , the micelle solution is similar to an entangled polymer solution used to separate DNA by conventional gel/polymer matrix electrophoresis.<sup>[19]</sup> The specific viscosity is plotted against C<sub>12</sub>E<sub>5</sub> concentration (wt%) on a log-log scale in Figure 6.3. Two separate linear regimes are observed

for the low and high concentrations.  $c^*$  is calculated as the concentration where the two regimes intersect, measured as 0.75 wt%.<sup>[29], [30]</sup> This value agrees well with that reported by Grosswaser *et al.*, who reported a value of 1.08 wt% using cryo-TEM imaging of  $C_{12}E_5$  wormlike micelles in water at 18°C.<sup>[16]</sup>



**Figure 6.3** – Specific viscosity of  $C_{12}E_5$  surfactant in 1xTBE buffer at 20°C. Dashed line represents linear fit to data below  $c^*$ . Solid line represents linear fit to data above  $c^*$ . Error bars represent standard deviation with  $n = 3$ . The intersection of these two linear fits marks the  $c^*$ .

### 6.3.2 – Separation of End-alkylated DNA

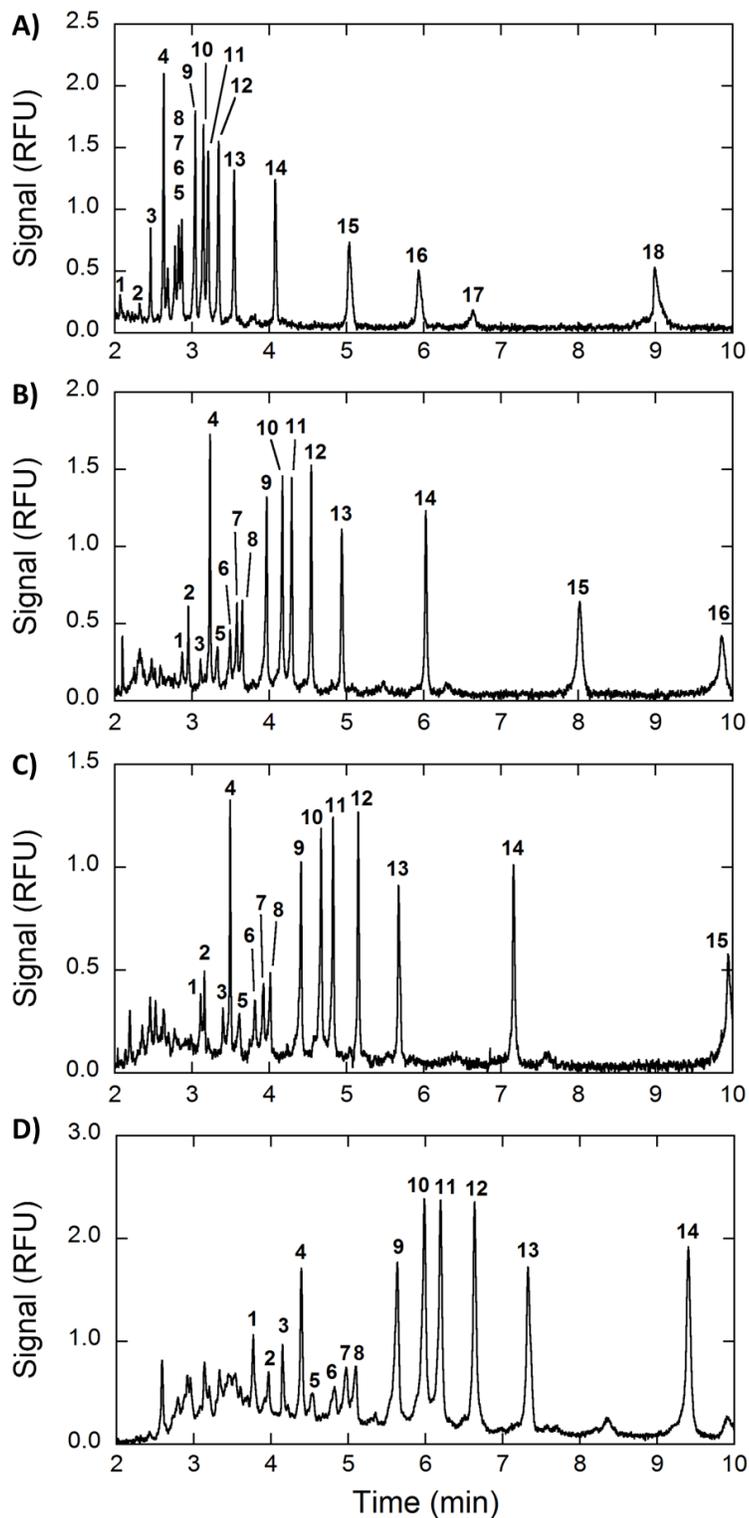
End-alkylated DNA ladders were produced using BODIPY FL<sub>16</sub> (C16B) modified PCR reverse primer, depicted in Figure 6.1. The hydrophobic modification allows the end-alkylated DNA to specifically interact with the hydrophobic core of the  $C_{12}E_5$  wormlike micelles. It is only through this interaction that non-ionic micelles can be used as ELFSE drag-tags. Figure 6.4 shows representative electropherograms of the C16B end-alkylated DNA ladders separated in different  $C_{12}E_5$   $c/c^*$  buffers. As the  $C_{12}E_5$  concentration increases, the

length of the wormlike micelles also increases, as does the hydrodynamic drag it exerts on the electrophoresing DNA. The increased drag results in greater retardation of the DNA's elution time. It is important to note that the longest end-alkylated DNAs are observed to elute first at all  $c/c^*$ s.

### 6.3.3 – End-alkylated DNA ELFSE Parameters

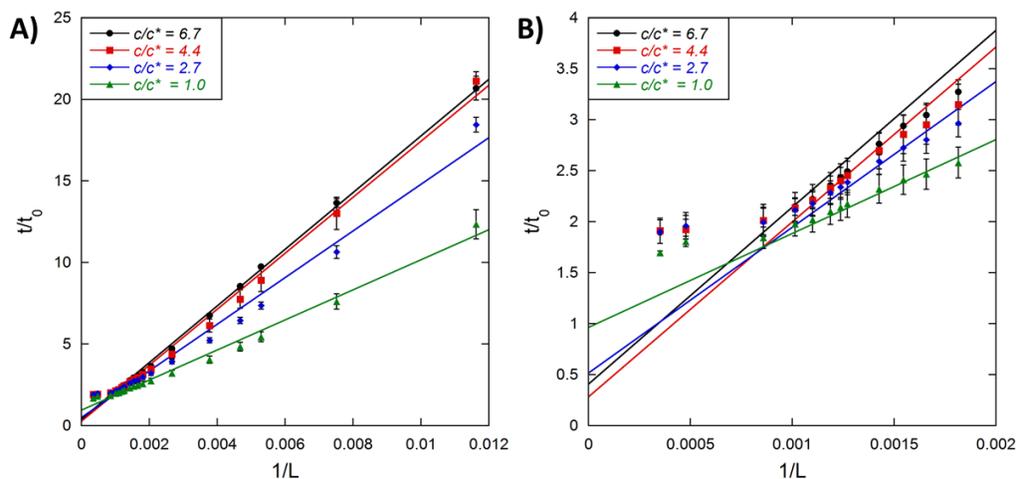
The effective drag of the micelles ( $\alpha$ ) can be calculated from a plot of  $t/t_0$  vs.  $l/L$ , as seen in Figure 6.5. Equation 6.1 predicts that these plots should exhibit linear behavior if the separation mechanism follows ELFSE, with the slope of the line equal to  $\alpha$ . For each separation,  $t_0$  was measured independently using a 22 base Cy5-labeled oligomer. For ssDNA,  $\mu_0$  is constant for DNA lengths greater than a single persistence length, approximately 18 bases, therefore the 22 base oligomer can be used to measure  $t_0$ .<sup>[31], [32]</sup>

For all C<sub>12</sub>E<sub>5</sub>  $c/c^*$ s, the C16B end-alkylated DNA exhibits linear behavior, as seen in Figure 6.5. From these fits, an intercept of 1 would be expected. As the concentration of C<sub>12</sub>E<sub>5</sub> increases, the intercepts begin to decrease below 1, Figure 6.5B. These deviations are believed to be caused by increased sieving pressures from the C<sub>12</sub>E<sub>5</sub> micelle entanglements at higher  $c/c^*$ s. The sieving pressures are evident for the longer DNA lengths (low  $l/L$ ) that show significant deviation from the linear behavior predicted by ELFSE. An entangled micelle network would present an opposing separation mechanism where longer lengths are more retarded than the smaller lengths. The assumption that all lengths are migrating at the  $\mu_0$  may be invalidated in an entangled network.



**Figure 6.4** – Representative electropherograms showing the separation of C16B end-alkylated DNA ladders at different  $C_{12}E_5$   $c/c^*$  ratios in 1xTBE.  $c/c^*$  = A) 1.0, B) 2.7, C) 4.4, D) 6.7. Numbers above peaks correspond DNA lengths (1) 2,840, (2) 2,093, (3) 1165, (4) 986, (5) 909, (6) 842, (7) 808, (8) 787, (9) 701, (10) 647, (11) 603, (12) 551, (13) 488, (14) 421, (15) 375, (16) 265.

The network appears to shift the DNA to longer elution times, but for all concentrations studied the ELFSE elution order is observed with longest DNA eluting first.

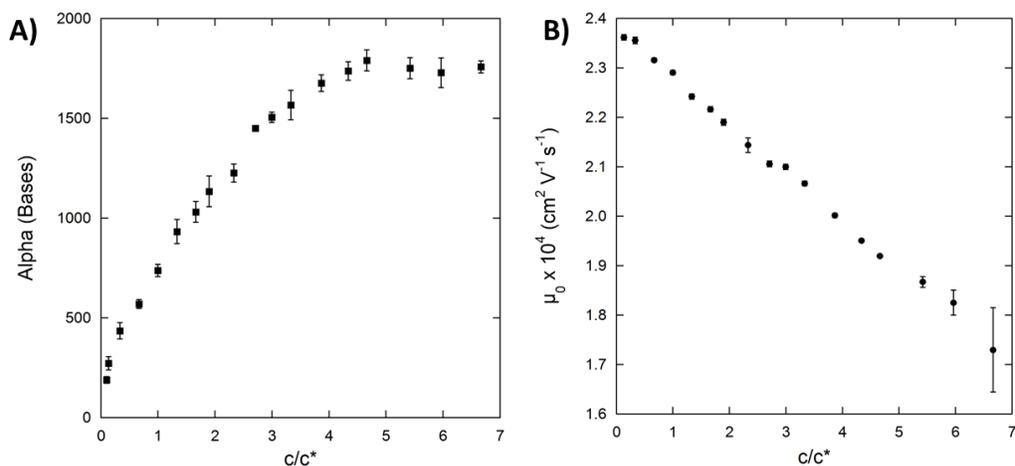


**Figure 6.5** –  $t/t_0$  vs  $1/L$  for the C16B DNA ladder separations at different  $C_{12}E_5$   $c/c^*$  ratios. Linear fits demonstrate behavior according to ELFSE theory, Equation 6.1. Error bars represent standard deviation with  $n = 3$ . A) Full data set. B) Zoom in on long DNA length deviations.

The effect of  $c/c^*$  on  $\mu_0$  and  $\alpha$  are seen in Figure 6.6. A gradual decrease in  $\mu_0$  with increasing  $c/c^*$  is observed, most likely due to an overall higher viscosity at higher  $c/c^*$ s. Sieving of the 22 base Cy5 oligomer is unlikely as very small pores are required to separate oligomeric DNA, such as highly cross-linked polyacrylamide gels with pore sizes on the order of ten nanometers.<sup>[33]</sup>

A gradual increase in  $\alpha$  is observed with  $c/c^*$  for the C16B end-alkylated DNA ladder. These increases reflect the increased drag from the  $C_{12}E_5$  wormlike micelles, but it is unclear whether the DNA is interacting with a single wormlike micelle of large size or micelles that may be entangling with one another.  $c^*$  represents the concentration when the radius of gyration ( $R_g$ ) of the micelles just begin to overlap in solution, not necessarily begin to entangle. Cryo-TEM

imaging has shown that C<sub>12</sub>E<sub>5</sub> micelles begin to exhibit entangled, branched structures at 1.08wt% at 18°C, similar to our  $c^* = 0.75\text{wt\%}$  measured by viscosity.<sup>[16]</sup> Additionally, light scattering has shown that at roughly 1-2wt%, increased contributions from second order virial coefficients in SLS and minima in diffusion coefficients in DLS are observed.<sup>[18], [28]</sup> These measurements all indicate that individual micelles do not exist at concentrations slightly greater than  $c^*$ .



**Figure 6.6** – ELFSE parameters from linear fits on Figure 6.5 according to Equation 6.2 at different C<sub>12</sub>E<sub>5</sub>  $c/c^*$  ratios. A) Effective micelle drag-tag size,  $\alpha$ . B) Measured  $\mu_0$ . Error bars represent standard deviation with  $n = 3$ .

The plateau in Figure 6.6 suggests that a  $c/c^*$  of 4.5 (3.25 wt%) is required for the micelles to reach their maximum size. However, from the literature results, this increase past  $c^*$  most likely represents the interaction with an entangled network acting as a collaborative drag force. Past a  $c/c^*$  of 4.5, additional entanglements have a minimal effect on increasing the collaborative drag from the interacting micelles.

Regardless of the interaction, these 1,800  $\alpha$  values are the largest ever reported for ELFSE separations. These huge drag-tags should enable separation of longer

lengths of DNA and higher lengths of read for DNA sequencing. Covalently attached drag-tags like streptavidin and protein polymers have reported  $\alpha$  values of  $\sim 50$ .<sup>[8], [34]</sup> The size of covalent drag-tags is limited by the requirement that the drag-tags are perfectly monodisperse to maintain efficient separations, as well as attachment of large molecules to PCR primers negatively affecting polymerase enzymes from functioning properly.<sup>[24]</sup> The non-ionic surfactant micelles have a transient interaction with the end-alkylation modification on DNA. This interaction is believed to last on the order of micelle lifetimes ( $ms$ ), leading to a statistical averaging of the polydisperse micelle population throughout the course of the separation.<sup>[11]</sup> This averaging process remedies the requirement for monodisperse drag-tags, granting a tremendous benefit compared to the covalently bound drag-tags. In addition, the small C16B end-alkylation is the only modification required in order to encourage interaction. Clearly this modification does not interfere with the polymerase function and end-alkylated DNA fragments can be easily made by PCR techniques. Changing drag-tags is as simple as changing the surfactant buffer in the capillary rather than synthesizing a whole new DNA.

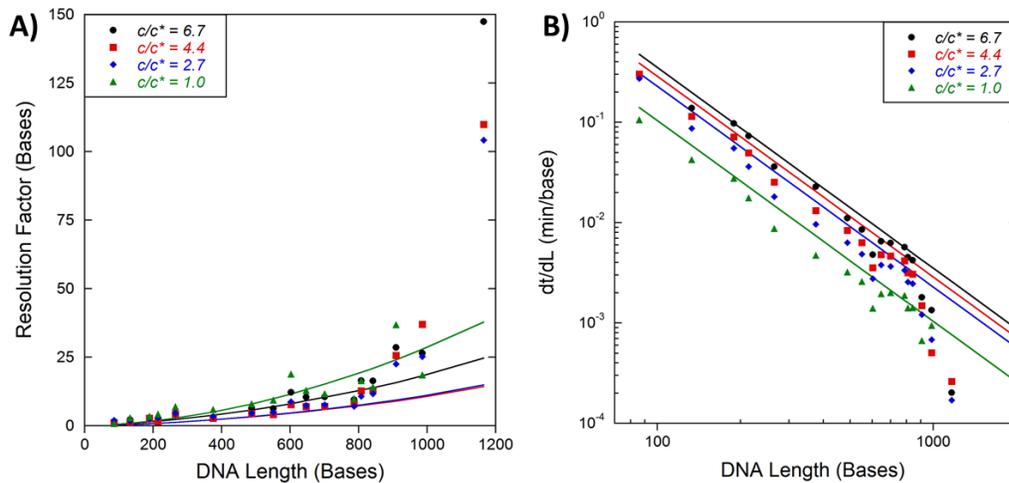
### **6.3.4 – Resolution of End-alkylated DNA**

The effectiveness of a separation can be judged by the resolution factor ( $R$ ), the ratio of peak's full width at half max ( $FWHM$ ) to the spacing between the adjacent DNA length elution peak ( $dt/dL$ ). A smaller value for  $R$  means a more efficient separation, where an  $R \leq 1.5$  is the limit for single base resolution.

$$R = \frac{FWHM}{dt/dL} \quad (6.2)$$

The instantaneous  $dt/dL$  can be calculated for all DNA lengths by taking the derivative of Equation 6.1, as seen in Equation 6.3. Figure 6.7A shows the experimental resolution factors and predicted resolution factors using the FWHM of each elution peak and the predicted  $dt/dL$  from Equation 6.3 and  $\alpha$  parameter in Figure 6.5.

$$\frac{dt}{dL} = \frac{l_d l_t \alpha}{\mu_0 V L^2} \quad (6.3)$$



**Figure 6.7** – A) Experimental (symbols) and theoretical (solid lines) resolution factors based on the FWHM measured for peaks in Figure 6.4. Theoretical resolution factors were found using  $dt/dL$  calculated using Equation 6.4 from the ELFSE parameters found in Figure 6.6. B) A comparison of actual  $dt/dL$  vs. the calculated  $dt/dL$  using Equation 6.3.

There are roughly 10 base resolution factors and excellent agreement with theoretical predictions for DNA lengths < 800 for all  $C_{12}E_5$  buffers. Separation of long DNA lengths may not require single base resolution, such as applications in forensic short tandem repeat analysis ( $R \sim 4$ ) or bacterial genotyping ( $R \sim 100$ ). However, for DNA longer than 800 bases, the theoretical predictions begin to

deviate from the experimentally observed values. The predicted resolution factors are much less, indicating a much more efficient separation than what is experimentally realized.

This difference becomes evident in the  $dt/dL$  predictions at DNA lengths  $>800$  bases seen in Figure 6.7B. The standard ELFSE theory presented in Equation 6.3 predicts a slope of -2 on a log-log plot of  $L$  vs.  $dt/dL$ . The shorter lengths adhere to the standard ELFSE prediction. For  $L > 800$ , we observe smaller  $dt/dL$ s in the C<sub>12</sub>E<sub>5</sub> wormlike micelle buffer than predicted by Equation 6.3, explaining the discrepancies between actual and predicted resolution factors in Figure 6.7A. The decrease in  $dt/dL$  is most likely due to the compaction of elution peaks from the entangled micelle network acting like a sieving matrix. The sieving mechanism of DNA is a counter elution order than ELFSE, where longer lengths of DNA are retarded more than shorter lengths. Since sieving matrices affect longer lengths more than shorter ones, one would expect to see deviations in the long lengths of DNA we are trying to separate by ELFSE.

### **6.3.5 – Modified ELFSE-Sieving Model**

To account for the presence of a sieving matrix, we present a modified ELFSE-sieving model. From our work with unmodified dsDNA, we believe that the micelle network behaves most closely with Ogston sieving theory.<sup>[35], [36]</sup> This theory states that when DNA is traveling through a porous network of entangled polymer fibers, the only parameter that sets the elution time is the fractional volume of the capillary available to the DNA for electrophoresis. This fraction,  $f$ , is a function of DNA length and polymer concentration

$$\mu = \mu_0 f(c, L) \quad (6.4)$$

Ogston's theory is based on those originally observed in gel permeation chromatography (GPC).<sup>[37]</sup> In GPC, a partitioning coefficient is defined that describes the ability of a molecule to partition from the solvent mobile phase into the stationary phase of void spaces in a porous media.<sup>[38]</sup> Larger molecules are excluded from the stationary phase and elute first off the column. Ogston sieving is similar, with the exception that the DNA will only migrate in the axial direction of the capillary when it can partition into the pore space of the network. The mobile phase now becomes the pore spaces within the network and the stationary phase is the excluded volume outside the pores.

Ogston described the partition coefficient as the probability a sphere with radius  $R$  can fit within the pore size of radius  $r$  within a network of long, thin fibers as<sup>[35]</sup>

$$K_{Ogston} = 2\pi\nu LD \exp\left[-\pi\nu L(a + R)^2\right] \quad (6.6)$$

Where  $\nu$  is the average number of polymer fibers per unit volume of gel,  $l$  is the average length of fiber, and  $a$  is the fiber radius.<sup>[35]</sup> This model makes a few key assumptions that may not be exactly valid for wormlike micelle systems. It assumes that the polymer fibers are very long and rigid. Wormlike micelles are transient structures that may exist in multiple conformations.<sup>[16]</sup> This may lead to interactions not predicted by Ogston's model. We have observed deviation from Ogston's model for our unmodified dsDNA separations in Chapter 5.

Instead, we chose to revert back to traditional GPC theory, which DNA sieving partition coefficients were first based on for a spherical particle entering a spherical void space,<sup>[38]</sup> defined as

$$K = \frac{6}{\pi^2} \exp \left[ - \left( \frac{\pi R_g}{r} \right)^2 \right] \quad (6.6)$$

Where  $R_g$  is the radius of gyration of the polymer and  $r$  is the radius of the pore, when the  $r > R_g$ . In this model, all parameters based on the surfactant network are lumped into a single fitting parameter for  $r$ . We use a simplified model of Equation 6.6 to describe the fractional volume available for end-alkylated DNA through the pore spaces.  $f$  is reduced to the exponential ratio of  $(R_g/r)^2$ , which follows the exponential of DNA length  $L$  times a constant  $b$ .

$$f = \exp \left[ - \left( \frac{R_g}{r} \right)^2 \right] \approx \exp(bL) \quad (6.7)$$

We consider a situation where the migration of end-alkylated DNA in free volume is set by the ELFSE elution time defined in Equation 6.1. The free volume is set by Equation 6.7, such that we define a modified ELFSE-sieving model theory for end-alkylated DNA migrating through entangled micelle networks as

$$t = \frac{t_0}{f} \left( 1 + \frac{\alpha}{L} \right) = t_0 \left( 1 + \frac{\alpha}{L} \right) \exp \left[ \left( \frac{R_g}{r} \right)^2 \right] \quad (6.8)$$

We have kept Equation 6.8 in terms of  $R_g$  and  $r$  since it allows us to determine the effective pore size of the network and compare to literature values. We can rearrange Equation 6.8 to yield an expression for the ELFSE contribution to end-alkylated DNA migration in the entangled micelle network.

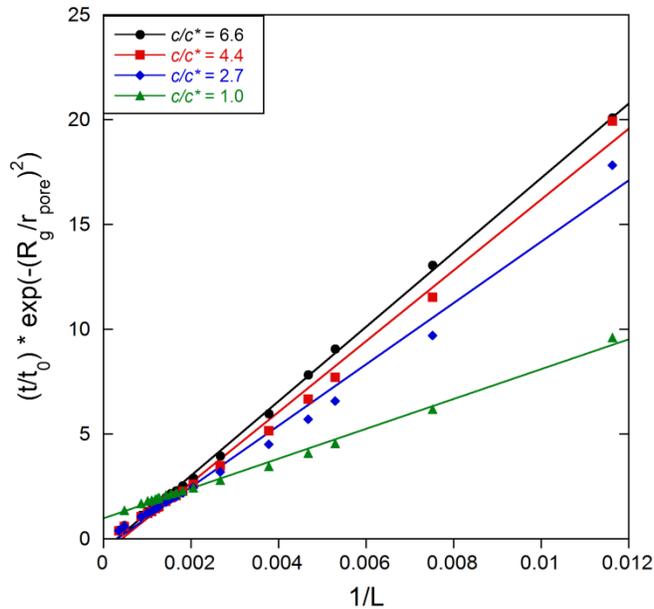
$$\frac{t}{t_0} \exp \left[ - \left( \frac{R_g}{r} \right)^2 \right] = 1 + \frac{\alpha}{L} \quad (6.9)$$

A plot of  $\frac{t}{t_0} \exp\left[-\left(\frac{R_g}{r}\right)^2\right]$  vs.  $1/L$  should yield a linear fit if this model is valid,

as seen in Figure 6.8. We can calculate the  $R_g$  of the end-alkylated DNA using the Krakty-Porod model for a wormlike chain.

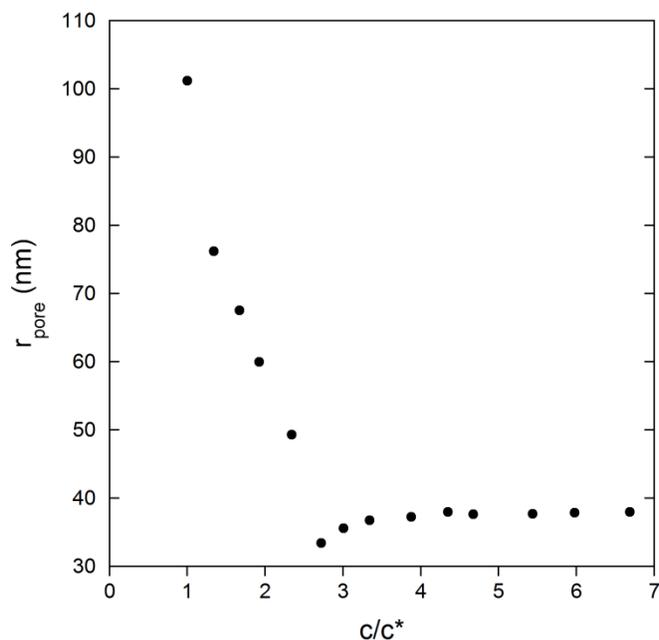
$$R_{g,DNA}^2 = \frac{pLb}{3} \left[ 1 - 3\left(\frac{p}{Lb}\right) + 6\left(\frac{p}{Lb}\right)^2 - 6\left(\frac{p}{Lb}\right) \left(1 - \exp\left(\frac{-Lb}{p}\right)\right) \right] \quad (6.11)$$

Where  $p$  is the persistence length of ssDNA (7nm) and  $b$  is the length of a ssDNA monomer (0.43nm). In order to determine the pore size, we use a non-linear optimization algorithm to maximize the  $R^2$  value for the linear fit in Figure 6.7. We find relatively constant pore sizes of 38nm for  $c/c^* > 3$  from this method, see Figure 6.9. These values agree with pore sizes determined by DLS, roughly 20nm and largely independent of concentration from  $c/c^* = 1 - 4$ .<sup>[18]</sup> They also agree with pore sizes calculated using Ferguson plots in the C<sub>12</sub>E<sub>5</sub> micelle network with unmodified double stranded DNA. Pore sizes were between 20-40nm for  $c/c^* = 2-8$ , with dramatic increases in near  $c^*$ . We believe the increase in pore size near  $c^*$  occurs because the micelles are just beginning to interact. They create large pore spaces and a weak network that has little effect on the end-alkylated DNA migration.



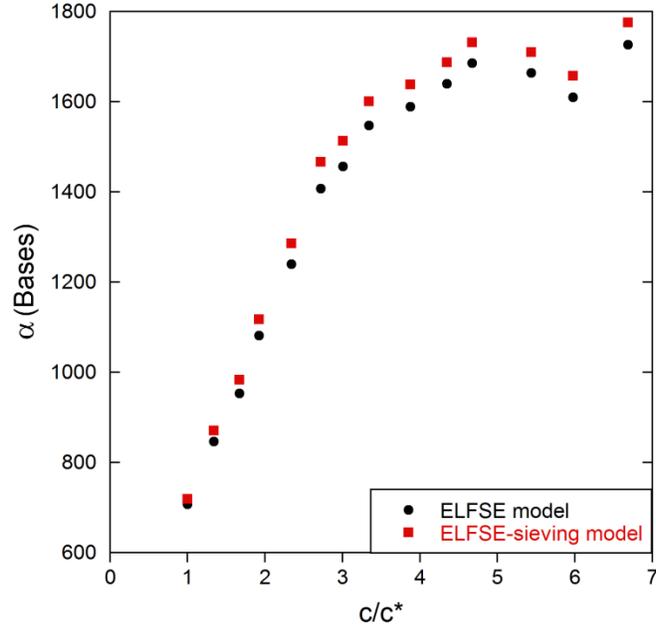
**Figure 6.8** –  $t/t_0$  scaled by fractional volume available to electrophoresis in modified ELFSE equation vs.  $1/L$ . Straight lines represent fits to Equation 6.9.

The inclusion of a fractional volume term in Equation 6.9 results in much better agreement for all DNA lengths. Furthermore, the slopes of these lines yield almost exactly the same  $\alpha$  values for the drag-tags as calculated with the traditional ELFSE model, see Figure 6.10. In these experiments, most of the lengths of DNA are unaffected by the background micelle network. They follow the traditional ELFSE model without sieving quite closely. For future systems with a greater number of longer lengths or fewer data points in the “ELFSE-only” regime, greater deviation may be expected. The modified ELFSE-sieving model should be applied for these systems in order to calculate the correct  $\alpha$ .



**Figure 6.9** –  $C_{12}E_5$  micelle network pore size calculated by Equation 6.9 using a non-linear optimization algorithm to maximize the  $R^2$  value for the linear fits in Figure 6.8

We should also note that the intercepts using this model are much less than the predicted value of 1. The assumption that  $t_0$  represents the free-solution migration of DNA in each system may not be valid. Typical analysis of Ferguson plots for traditional polymer network sieving requires extrapolation to a zero polymer concentration limit in order to accurately quantify  $t_0$ . The same may be true for end-alkylated DNA migrating in the entangled micelle networks, as only the  $c/c^* = 1.0$  separation yields an intercept of 1.



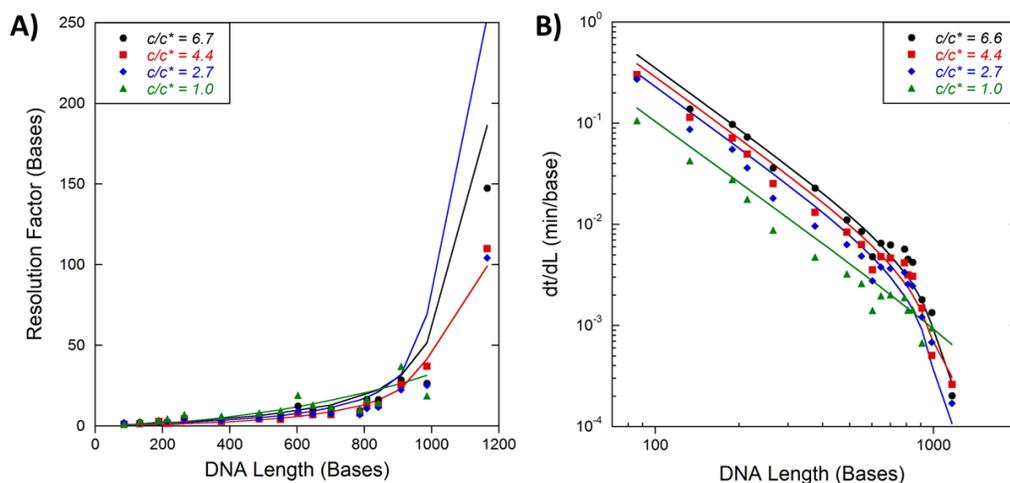
**Figure 6.10** – Effective micelle drag term  $\alpha$  calculated using Equation 6.1, the ELFSE only model, or Equation 6.9, the modified ELFSE-sieving model.

The goal of the modified ELFSE-sieving model is to better predict resolution factors for end-alkylated DNA in entangled micelle systems. We found that the large  $\alpha$  values from an ELFSE only model greatly over predicted the separation performance of the buffer, resulting in much lower  $R$  values. By substituting the length based exponential for  $f$  into Equation 6.8, the modified ELFSE  $dt/dL$  can be solved by Equation 6.11.

$$\left| \frac{dt}{dL} \right| = \left| \frac{t_0 \exp(bL) [\alpha(bL-1) + bL^2]}{L^2} \right| \quad (6.11)$$

Using Equation 6.11, we can compare experimental  $R$  vs. those predicted by the modified ELFSE-sieving model, see Figure 6.11A. Including the additional term for network formation accurately captures the dramatic increase in  $R$  with larger DNA lengths. This becomes evident in the theoretical modified ELFSE-sieving

model  $dt/dL$  in Figure 6.11B. For low lengths, each system follows a power law dependence of -2 on the log-log plot, predicted by an ELFSE only model. For the lengths  $>800$ , peak compression is observed as  $dt/dL$  sharply decreases with length. The modified ELFSE-sieving model accurately predicts this behavior from the additional sieving pressures the micelle network imposes on the longer DNA lengths.



**Figure 6.11** – A) Experimental (symbols) and theoretical (solid lines) resolution factors based on the FWHM measured for peaks in Figure 6.4. Theoretical resolution factors were found using  $dt/dL$  calculated using Equation 6.11 from the ELFSE parameters found in Figure 6.6. B) A comparison of actual  $dt/dL$  vs. the theoretical  $dt/dL$  using Equation 6.11.

Some limitations are present in the proposed model. Equation 6.11 predicts that  $dt/dL$  will begin to increase when  $bL > 1$ . For these lengths, the sequential peak spacing increases due to the entangled network exerting large sieving pressures on the DNA. This would result in an elution order reversal as sieving becomes the dominate separation mode. The  $dt/dL$  terms do not account for elution order reversal and extrapolation of these terms to high length should be used with caution.

Also, calculation of pore size for lower  $c/c^*$  ratios may be problematic. We can see the modified ELFSE-sieving model greatly over predicts  $R$  for  $c/c^* = 2.7$ . From the pore size calculations, there appears to be a discontinuity in behavior at this value. These moderate  $c/c^*$  values may represent points when the surfactant solution is not truly entangled and just begins to follow behavior according to Equation 6.11. Further investigation into micelle network pore size would benefit the application of the modified ELFSE-sieving model.

## 6.4 – Conclusion

We have investigated the use of non-ionic  $C_{12}E_5$  surfactant micelles as large ELFSE drag-tags.  $C_{12}E_5$  forms long wormlike micelles in aqueous solutions, which should provide sufficient hydrodynamic drag on the electrophoresing long end-alkylated DNA strands for separation. However, the wormlike micelles begin to interact with one another at relatively low concentrations, forming an entangled network that opposes ELFSE separation. We have demonstrated that ELFSE separations using  $C_{12}E_5$  micelles above  $c^*$  results in extremely high values for the micelle drag-tag parameter,  $\alpha$ . The highest  $\alpha$  found was 1,800, an order of magnitude larger than any currently published ELFSE drag-tag size. However, the separation efficiency for long end-alkylated DNA strands was greatly over predicted for these systems. We find that the presence of the entangled micelle network compresses elution peaks of long end-alkylated DNA, decreasing separation efficiency. We have proposed a modified ELFSE-sieving model that accounts for the additional sieving pressures arising from the entangled micelle network. We include an additional migration term based on the fractional volume

available to the end-alkylated DNA in the network, following theory derived by Ogston sieving and conventional GPC separations. With the fractional volume term, we are able to accurately predict the resolution factors for all end-alkylated DNA lengths migrating through the entangled micelle system. Although the presence of a micelle network reduces separation efficiency of end-alkylated DNA, an increased understanding of how the network affects migration and resolution enables greater insight into optimization of future micelle drag-tag buffer systems.

## 6.5 – References

- [1] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.
- [2] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [3] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [4] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [5] K. Grass, C. Holm, and G. W. Slater, *Macromolecules*, 2009, 42, 5352–5359.
- [6] X. Wang, J. C. Albrecht, J. S. Lin, and A. E. Barron, *Biomacromolecules*, 2012, 13, 117–123.
- [7] R. D. Haynes, R. J. Meagher, J. Won, F. M. Bogdan, and A. E. Barron, *Bioconjug. Chem.*, 2005, 16, 929–938.
- [8] J. Won, R. J. Meagher, and A. E. Barron, *Electrophoresis*, 2005, 26, 2138–2148.
- [9] D. W. Wegman, L. T. Cherney, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2013, 85, 6518–6523.
- [10] J. S. Lin, J. C. Albrecht, R. J. Meagher, X. Wang, and A. E. Barron, *Biomacromolecules*, 2011, 12, 2275–2284.

- [11] S. B. Istivan and J. W. Schneider, "Rapid Gel-free Separations of DNA via Transient Attachment to Surfactant Micelles," Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2012.
- [12] Y. Einaga, Y. Inaba, and M. Syakado, *Polym. J.*, 2006, 38, 64–72.
- [13] S. Yoshimura, S. Shirai, and Y. Einaga, *Society*, 2004, 15477–15487.
- [14] K. Imanishi and Y. Einaga, *J. Phys. Chem. B*, 2005, 109, 7579–7581.
- [15] Y. Einaga, A. Kusumoto, and A. Noda, *Polym. J.*, 2005, 37, 368–375.
- [16] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [17] K. Kawasaki, B. Lindman, H. Okabayashi, T. Kato, N. Taguchi, and D. Nozu, *Form. Dyn. Self-Organized Struct. Surfactants Polym. Solut.*, 1997, 106, 57–60.
- [18] T. Kato, S. Anzai, and T. Seimiya, *J. Phys. Chem.*, 1990, 94, 7255–7259.
- [19] Wei and E. S. Yeung, *Anal. Chem.*, 2001, 73, 1776–1783.
- [20] J. Lee, H. Li, and E. S. Yeung, *J. Chromatogr. A*, 2004, 1053, 173–179.
- [21] C. Desruisseaux, G. Drouin, and G. W. Slater, *Macromolecules*, 2001, 34, 5280–5286.
- [22] R. J. Meagher, J. I. Won, L. C. McCormick, S. Nedelcu, M. M. Bertrand, J. L. Bertram, G. Drouin, A. E. Barron, and G. W. Slater, *Electrophoresis*, 2005, 26, 331–350.
- [23] R. J. Meagher, J. Won, J. A. Coyne, J. Lin, and A. E. Barron, *Anal. Chem.*, 2008, 1, 2842–2848.
- [24] J. C. Albrecht, J. S. Lin, and A. E. Barron, *Anal. Chem.*, 2011, 83, 509–515.
- [25] T. Mukaiyama, *Angew. Chemie Int. Ed. English*, 1976, 15, 94–103.
- [26] A. Deratani and T. Maraldo, *React. Polym.*, 1988, 9, 19–28.
- [27] T. Da Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, and B. Cacciari, *Curr. Med. Chem.*, 2005, 12, 71–88.
- [28] S. Shirai and Y. Einaga, *Polym. J.*, 2005, 37, 913–924.
- [29] P. D. Grossman and D. S. Soane, *Biopolymers*, 1991, 31, 1221–1228.
- [30] D. A. Hill and D. S. Soane, *J. Polym. Sci. Part B Polym. Phys.*, 1989, 27, 2295–2320.

- [31] N. C. Stellwagen, S. Magnusdottir, C. Gelfi, and P. G. Righetti, *Biopolymers*, 2001, 58, 390–397.
- [32] U. Mohanty and N. C. Stellwagen, *Biopolymers*, 1999, 49, 209–214.
- [33] D. L. Holmes and N. C. Stellwagen, *Electrophoresis*, 1991, 12, 612–619.
- [34] H. Ren, A. E. Karger, F. Oaks, S. Menchen, G. W. Slater, and G. Drouin, *Electrophoresis*, 1999, 20, 2501–2509.
- [35] A. G. Ogston, *Trans. Faraday Soc.*, 1958, 54, 1754–1757.
- [36] J. L. Viovy and T. Duke, *Electrophoresis*, 1993, 14, 322–329.
- [37] D. Rodbard and A. Chrambach, *Proc. Natl. Acad. Sci. U. S. A.*, 1970, 65, 970–977.
- [38] E. Casassa, *J. Polym. Sci. Part B Polym. Lett.*, 1967, 5, 773–778.

# Chapter 7 – Segregation Effects in Rapid Kilobase DNA Separations

## 7.1 – Introduction

Separation of long, kilobase (kB) DNA fragments by micelle end-labeled free-solution electrophoresis (ELFSE) requires large micelle drag-tags to provide sufficient friction on the electrophoresing DNA. We have investigated non-ionic C<sub>12</sub>E<sub>5</sub> *n*-alkyl polyoxyethylene ether surfactant as a candidate large micelle ELFSE drag-tag. The relatively long alkyl chain compared to ethylene oxide head group creates micelles that undergo one-dimensional growth into long, wormlike micelles.<sup>[1], [2]</sup> We found that even though the C<sub>12</sub>E<sub>5</sub> wormlike micelle structures were theoretically large enough to separate kB DNA, micelle entanglements created a DNA sieving-like network that dramatically reduced the separation efficiency.

Our group has previously identified a mixed surfactant buffer of 36mM C<sub>16</sub>E<sub>6</sub>, 24mM C<sub>12</sub>E<sub>5</sub>, 3mM C<sub>10</sub>E<sub>5</sub> as a large micelle buffer for DNA separations.<sup>[3]</sup> Using C<sub>16</sub>E<sub>6</sub> on its own produced very poor separations due its propensity for entanglement at low concentrations.<sup>[4]-[6]</sup> In mixed surfactant systems, smaller *i/j* ratio surfactants decrease the mean wormlike micelle length ( $\bar{L}$ ) of higher *i/j* ratio surfactants, where  $\bar{L}$  can be express as,<sup>[7]-[9]</sup>

$$\bar{L} \cong \phi^{1/2} \exp\left(\frac{E_C}{2kT}\right) \quad (7.1)$$

where  $\Phi^{1/2}$  is the total volume fraction of the surfactant,  $E_C$  is the end-cap energy of the surfactant system,  $k$  is Boltzmann's constant, and  $T$  is the temperature.  $E_C$  is the difference in free energy between adding a surfactant molecule to the cylindrical core compared to the spherical end-cap of a micelle.<sup>[9]</sup> The smaller  $i/j$  ratio surfactants are more easily incorporated into high curvature end-caps of the micelle, reducing  $E_C$  and  $\bar{L}$ . Thus, smaller  $\bar{L}$  should create a lower probability of micelle entanglement to preserve ELFSE separations.

Here, we investigate the use of the ternary C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, and C<sub>10</sub>E<sub>5</sub> system to separate kB long DNA. This buffer was identified using DNA <500 bases long where micelle entanglements are not expected to result in detrimental effects. We once again observe non-linear deviations from standard ELFSE theory for long DNA fragments above the overlap concentration ( $c^*$ ). Interestingly, these deviations appear to result from hydrodynamic segregation of the DNA-micelle conjugates, rather than sieving through the entangled network.<sup>[10]</sup> The separation actually benefits from segregation, due to reduced electrophoretic mobility length dependence. The longer DNA moves slower and is given more time to separate on the capillary.

We find that segregation is independent of electric field strength or steric limitations of DNA, both opposed to predictions in literature.<sup>[10]-[12]</sup> We believe segregation is caused by the entangled micelle microstructure, where DNA passes through the network unperturbed, but micelles cannot. Using these segregated,

ternary surfactant buffers, we demonstrate excellent separation of 1 – 10kB end-alkylated DNA in 3 minutes with ~100 base resolution.

## 7.2 – Materials and Methods

### 7.2.1 – Reagents

Nuclease-free water, acetic acid, GeneJET Gel Extraction Kits, and acetone were purchased from Fisher Scientific (Pittsburgh, PA). BODIPY<sup>®</sup> FL C<sub>16</sub> (C16B), low melting temperature agarose, Hi-Di formamide, and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). Benzyl alcohol, DMSO, 2,2-Dithiodipyridine (DPDS), Lithium perchlorate, Triphenylphosphine (TPP), 4-(Dimethylamino)pyridine (DMAP), Hexadecyltrimethyl ammoniumbromide (CTAB), *N*-hexadecylhexaoxyethylene (C<sub>16</sub>E<sub>6</sub>), ethidium bromide, acetonitrile, urea, 10x Tris-Borate-EDTA (TBE), stearic acid, and triethanolamine (TEA) were purchased from Sigma Aldrich (St. Louis, MO). *N*-dodecylpentaoxyethylene (C<sub>12</sub>E<sub>5</sub>) and *N*-decylpentaoxyethylene (C<sub>10</sub>E<sub>5</sub>) surfactant were purchased from BaChem (Torrance, CA). λDNA and LongAmp<sup>™</sup> *Taq* DNA polymerase Kit were purchased from New England Biolabs (Ipswich, MA). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2mΩ deionized water and filtering through a 0.22μm filter. PCR forward primers and C<sub>6</sub>-amine 5' modified reverse primers were purchased with standard desalting purification from Integrated DNA Technologies (Coralville, IA). Dried oligomers were resuspended in 10mM Tris 0.1mM EDTA buffer and used without further purification.

## 7.2.2 – DNA Alkylation Reaction

Alkylation of C<sub>6</sub>-amine 5' modified reverse primer with C16B was accomplished using a combination of methods outlined by Mukaiyama and Deratani, and closely followed a scheme first reported by Bourtine *et al.*<sup>[13]–[15]</sup> An aliquot of 10nmol 5'amine modified reverse primer was precipitated in the presence of 750nmol CTAB. The suspension was then dried by vacuum centrifugation under low heat for 30 minutes. Meanwhile, a mixture of 100nmol C16B or C18, 5μmol DMAP, 2.5μmol TPP, 2.5μmol DPDS, and 7.5μmole TEA in DMSO was prepared and allowed to activate at room temperature for 20 minutes. Once activated, the reaction mixture was added to the dried DNA suspension, briefly vortexed and bath sonicated to resuspend the DNA pellet, and allowed to react overnight under continuous agitation in a shaker at room temperature in the dark. The solution was then precipitated with 1mL of 2w/v% lithium perchlorate in acetone and pelleted by centrifugation. The supernatant was removed and the pellet was triple rinsed with acetone. The pellet was resuspended in acetone, centrifuged, and supernatant removed again. The pellet was then dried by vacuum centrifugation under low heat for 5 minutes and resuspended in 0.1M triethylammoniumacetate pH 7.0 (TEAA) buffer.

Purification of the reaction mixture was performed using a Waters 4.6 mm x 250 mm Symmetry300 C18 HPLC column (Milford, MA) with a 1mL/min flowrate and linear gradient from 0.1M TEAA to 100% acetonitrile over 30 minutes. Any unmodified DNA eluted near 10 minutes while the end-alkylated

DNA eluted near 22 minutes. Product fractions were collected, lyophilized, and resuspended in 10mM Tris 0.1mM EDTA buffer.

### **7.2.3 – Lower Critical Solution Temperature (LCST)**

#### **Determination**

The LCST was determined using a Cary300 UV-Vis (Agilent Technologies, Santa Clara, CA). Each surfactant concentration was heated in a cuvette at 1°C/min, and its 488nm absorbance measured every 0.5°C. A sharp increase in absorbance indicated a more turbid sample, and therefore the LCST. The LCST of 12mM C<sub>16</sub>E<sub>6</sub>, 8mM C<sub>12</sub>E<sub>5</sub>, 1mM C<sub>10</sub>E<sub>5</sub> was found to be 35°C.

### **7.2.4 – PCR Amplification of C16B End-alkylated DNA Ladder**

Specific DNA lengths of C16B end-alkylated DNA were amplified using LongAmp<sup>TM</sup> *Taq* DNA polymerase, λDNA template, and a C16B end-alkylated λDNA reverse primer, and a length setting forward primer. The C16B end-alkylated DNA ladder included the following lengths of DNA: 86, 133, 189, 214, 265, 375, 421, 488, 551, 603, 647, 707, 787, 808, 847, 909, 986, 1165, 2093, 2840, 4111, 6003, and 9893. Each PCR reaction contained 1ng of λDNA template, 500nM each primer, 0.3mM each dNTP, 2.5 units of LongAmp<sup>TM</sup> *Taq* DNA polymerase, 1x LongAmp<sup>TM</sup> buffer, and enough nuclease free water to 25μL. The reaction was denatured for 2min at 95°C and subjected to 35 cycles of denaturation at 95°C for 15s, primer annealing at 50°C for 30 s, and extension at 65°C for 50 seconds per 1000 bases of amplicon length, followed by a final hold for 10min at 65°C.

After thermal cycling, the PCR reaction mixture was directly loaded onto 1wt% agarose gels. Slab gel electrophoresis was performed at 4 V/cm for 2 hours. Gels were imaged using ethidium bromide staining and UV excitation on a BioDoc-It Imaging System from Ultra-Violet Products (Cambridge, UK). Correct PCR product length bands were excised and purified using GeneJET Gel Extraction Kit spin columns, and eluted using the provided 10mM Tris 0.1mM EDTA buffer.

### **7.2.5 – Capillary Electrophoresis**

Analysis of C16B end-alkylated DNA ladders was performed on a P/ACE MDQ capillary electrophoresis instrument (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. Excitation sources were from a 3mW 488nm Argon ion. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 50 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and a length of 20cm to the detector and 30cm total. The C16B end-alkylated DNA ladders were prepared by pipetting 2 $\mu$ L of each purified PCR product with 30 $\mu$ L of formamide, heating to 95°C for 2 minutes to ensure denatured single stranded DNA products, and immediately cooled on ice. 4kV for 30sec electrokinetic injections were used to introduce samples into the capillary. Electrophoretic separation was conducted under applied voltages indicated in figure captions in reverse polarity mode (from anode to cathode). The capillary was heated to 25°C using a built in coolant jacket system. At higher temperatures the separation failed, believed to be caused by operating close to the LCST. Data collection was performed using 32 Karat

software (Beckman Coulter). Peaks were analyzed using OriginPro 9.0 peak fitting software (Northampton, MA).

Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer and maintained by adding 0.1v/v% POP-6 to all capillary electrophoresis buffers. Surfactant buffers were prepared by mixing the indicated concentration of  $C_iE_j$  surfactants in 1xTBE, briefly vortexing, and shaken for at least 1 hour at room temperature to ensure surfactant solubilization and micelle formation. The capillary could be used for many separations without observing any detrimental effects.

### **7.2.6 – Viscosity Measurements**

Viscosities of surfactant solutions were measured via capillary viscometry using a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for UV detection. The capillary was a 50 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) with a length of 20cm to the detector and 30cm total. A small plug of concentrated, UV absorbing benzyl alcohol was injected into the capillary filled with surfactant solution. A pressure of 1 psi was used to hydrodynamically push the plug past the detector. The viscosity of the surfactant solution was determined using the elution time of the benzyl alcohol plug and assuming Hagen-Poiseuille plug flow.

## **7.3 – Segregation Theory**

### **7.3.1 – Segregated Electrophoretic Mobility**

The theory currently utilized to describe ELFSE was developed by Long *et al.*, where they considered the electrophoresis of a composite object composed of charged and uncharged segments.<sup>[16]</sup> The electrophoretic mobility of the composite is simply the electrophoretic mobility of each segment ( $\mu_C$ ,  $\mu_U$ ) weighted by their respective frictional coefficients ( $\xi_C$ ,  $\xi_U$ ).

$$\mu = \mu_C \frac{\xi_C}{\xi_C + \xi_U} + \mu_U \frac{\xi_U}{\xi_C + \xi_U} \quad (7.1)$$

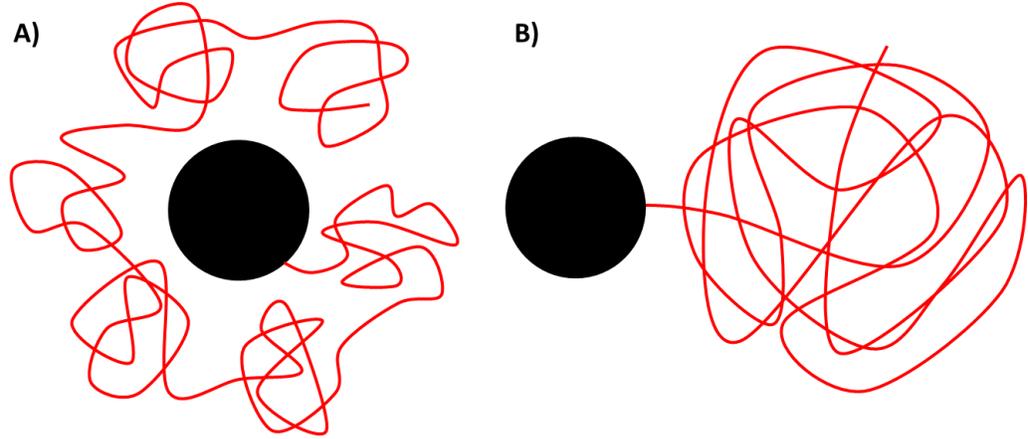
For a DNA-micelle system,  $\mu_C = \mu_0$  and  $\mu_U = 0$ . Standard ELFSE theory is derived from Equation 7.1 assuming that the DNA-micelle conjugate forms an unsegregated, single random coil hydrodynamic unit. The DNA and micelle are considered as blobs having the same hydrodynamic radii, allowing the substitution of charged and uncharged friction coefficients as  $L$  and  $\alpha$ , respectively.<sup>[17]</sup>  $\alpha$  acts as a scaling factor to convert the number of drag-tag monomers into the number of hydrodynamically equivalent uncharged DNA monomers.

$$\mu = \mu_0 \left( \frac{L}{L + \alpha} \right) \quad (7.2)$$

$$\alpha = N \frac{b_m b_{k,m}}{b_D b_{k,D}} \quad (7.3)$$

Where  $N$  is the number of micelle monomers,  $b_m$  is the micelle monomer size,  $b_{k,m}$  is the micelle Kuhn length,  $b_D$  is the DNA monomer size, and  $b_{k,D}$  is the DNA Kuhn length.

As the micelle or DNA grows in size, the micelle may begin to lag behind the electrophoresing DNA, resulting in hydrodynamic segregation, Figure 7.1.<sup>[10], [12]</sup>



**Figure 7.1** – Schematic representation of DNA drag-tag conjugates under A) standard, unsegregated conformation, and B) segregated conformation. DNA represented as solid red line, drag-tag as black circle.

Segregation would result in a system where the additional friction of the micelle can no longer be considered as that of equivalent uncharged DNA bases, but rather its own hydrodynamics. The frictional coefficients of DNA and micelle must be correctly evaluated for use of Equation 7.1. We consider the friction coefficients of the DNA and micelle as a random coil polymer chain, given by the Stokes drag.

$$\xi = 6\pi\eta R_H = 4\pi\eta R_g \quad (7.4)$$

Where  $\eta$  is the solution viscosity,  $R_H$  is the hydrodynamic radius,  $R_g$  is the radius of gyration, and we have applied the Kirkwood-Riseman approximation to express Equation 7.4 in terms of  $R_g$ .<sup>[18]</sup>

$$R_H \cong \frac{2}{3}R_g \quad (7.5)$$

We can approximate  $R_g$  as that of a random Gaussian coil when the total contour length ( $Lb$ ) is much greater than a coil's Kuhn length ( $b_k$ )

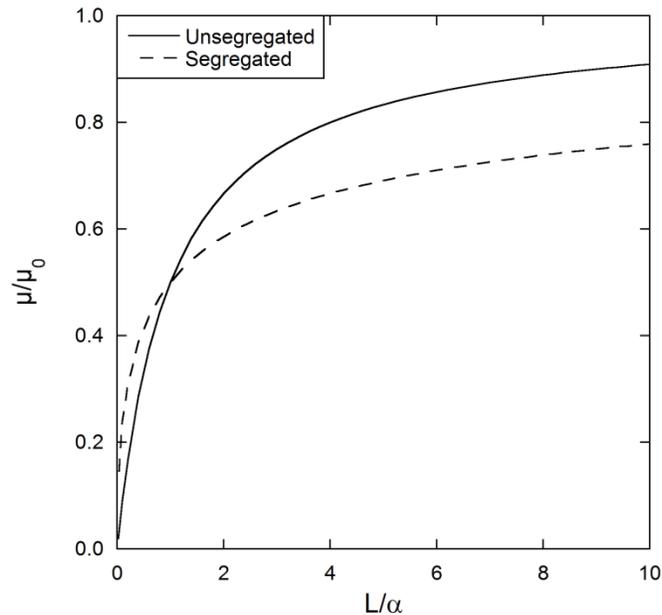
$$R_{g,i} = \sqrt{\frac{L_i b_i b_{k,i}}{6}} \quad (7.6)$$

For single stranded DNA (ssDNA),  $b$  is the length of a single base (0.43nm) and  $b_k \sim 15\text{nm}$ .<sup>[19]</sup> Thus, this model is appropriate for ssDNA greater than 100 bases ( $\sim 3b_k$ ), much smaller than the DNA lengths investigated here. This assumption for wormlike micelles is a bit trickier. Light scattering and cryo-TEM imaging have found that C<sub>1</sub>E<sub>3</sub> wormlike micelles can have contour lengths above 1 $\mu\text{m}$  with diameters roughly  $>5\text{nm}$ .<sup>[2], [20]</sup> Magnetic birefringence techniques on ionic wormlike micelles calculated a persistence length of 20nm.<sup>[21]</sup> One would expect a more flexible micelle for non-ionic micelles from reduced electrostatic repulsion, giving some validity to our assumption that the micelle behaves like a random, Gaussian coil. However, we should point out that changing the conformation of the micelle will only change how the final drag-tag parameter  $\alpha$  is expressed and does not change the fundamental behavior of the DNA upon segregation.

Substituting Equations 7.3, 7.4, and 7.6 into 7.2 yields an equation to describe ELFSE migration of hydrodynamically segregated, random Gaussian coils of DNA and micelle drag-tags.

$$\mu = \mu_0 \frac{\sqrt{L}}{\sqrt{L} + \sqrt{\alpha}} \quad (7.8)$$

The segregated and unsegregated models appear quite similar. The segregated state simply scales with  $L^{0.5}$  rather than  $L$ . As seen in Figure 7.2, the DNA-micelle complex will migrate with a lower mobility than the unsegregated case when  $L/\alpha > 1$ , which is typical for most cases of interest when separating DNA by ELFSE. Since the DNA acts as the polyelectrolyte engine of the DNA-micelle complex, the square root dependence acts to decrease the total effective charge of the complex. The segregated state should aid in resolving long DNAs by allowing more time on the capillary for greater separation.



**Figure 7.2** – Relative mobility of unsegregated and segregated DNA-micelle complexes plotted against  $L/\alpha$

### 7.3.2 – Causes for Segregation

There are two theories on the establishment of ELFSE DNA micelle segregation. The first is thought to occur when the electric and drag forces on the DNA-micelle complex are high.<sup>[10]</sup> These situations could arise from the large

micelles exerting a greater amount of drag onto the DNA, possibly creating a system where the DNA appears effectively anchored at one end as the other end freely electrophoreses. Alternatively, very long DNA strands have a large amount of electrophoretic force ( $F_D$ ) acting on the connection between the DNA and micelle. When  $F_D$  is greater than the entropic forces ( $F_E$ ) resisting segregation at the connection, the DNA and micelle would become hydrodynamically distinct and segregated. These forces have been previously published considering the DNA drag-tag conjugate as a bead spring model, with the final results summarized here as<sup>[3],[10]</sup>

$$F_D = 4\pi\eta\mu_0 E \frac{R_{g,DNA} R_{g,micelle}}{R_{g,DNA} + R_{g,micelle}} \quad (7.10)$$

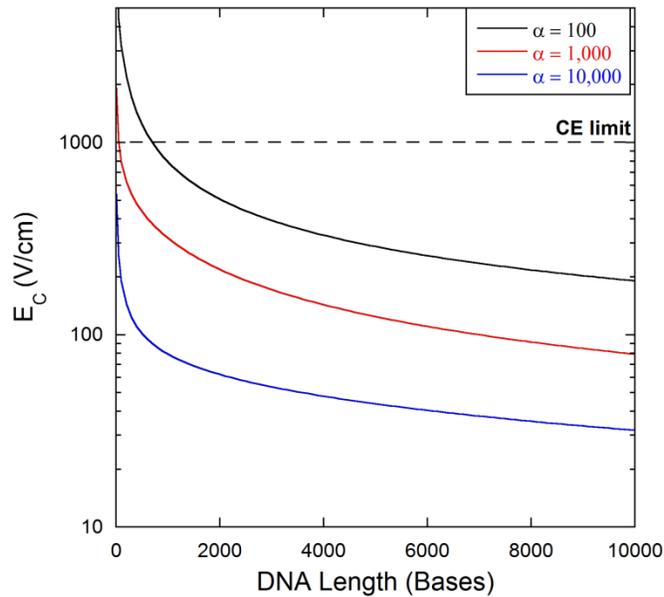
$$F_E = \frac{3kT(R_{g,DNA} + R_{g,micelle})}{\sqrt{6}(R_{g,DNA}^2 + R_{g,micelle}^2)} \quad (7.11)$$

Setting Equation 7.10 and 7.11 equal to each other and solving for  $E$  yields the critical electric field strength ( $E_c$ ) for DNA-micelle segregation. Substituting in  $R_g$  and  $\alpha$  terms leads to the expression for  $E_c$  as

$$E_c = E_0 \frac{\left(1 + \sqrt{\alpha/L}\right)^2}{\sqrt{\alpha L} \left(1 + \alpha/L\right)} \quad (7.12)$$

$$E_0 = \frac{3\sqrt{6}kT}{4\pi\eta\mu_0 b_D b_{k,D}} \quad (7.13)$$

The main result of Equation 7.12 is that  $E_c$  gets monotonically smaller as  $L$  and  $\alpha$  increase (see Figure 7.3). Larger drag-tags and DNAs are easier to segregate via electrical forces. We should be able to probe the space of electrically induced segregation by changing  $E$  for our long DNA separations. Figure 7.3 clearly demonstrates the feasible  $E$  regime we can use to test segregation of our DNA-micelle conjugates.



**Figure 7.3** – Critical segregation electric field strength ( $E_c$ ) calculated using Equation 7.12 for different  $\alpha$  values.

The second theory propose that DNA micelle segregation is caused by steric limitations of the DNA to maintain a random walk conformation around the large micelle drag-tag.<sup>[11], [12]</sup> If the DNA does not contain a sufficient number of Kuhn lengths to encompass the micelle, we would expect steric segregation to occur. Steric segregation of double stranded DNA (dsDNA),  $b_k \sim 300$  bases, 100nm, has been observed with streptavidin drag-tags.<sup>[12]</sup> The extremely stiff dsDNA is unable to curve around the small protein. The much more flexible ssDNA does

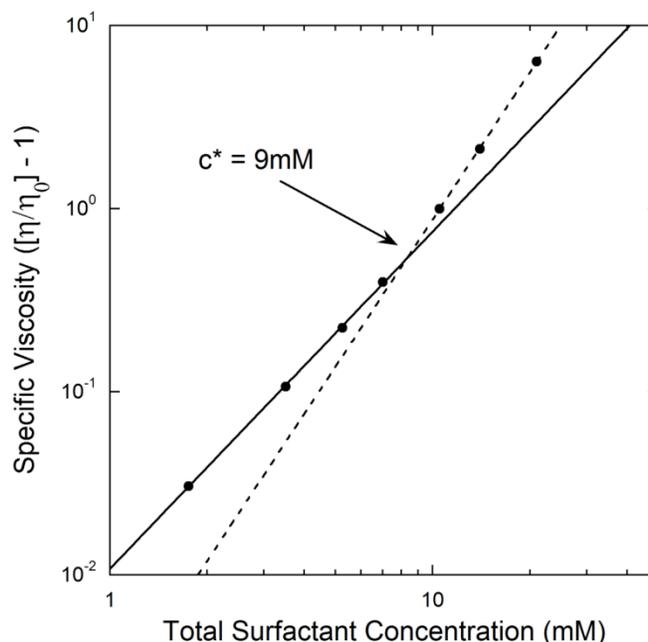
not typically encounter steric limitations until very short lengths are investigated (< 50 bases). However, when the micelles grow in size, longer fragments of ssDNA are required to sufficiently include the large micelle in its random coil.

## 7.4 – Results and Discussion

### 7.4.1 – Determination of $c^*$

The non-ionic  $C_{16}E_6$  surfactant forms  $\mu\text{m}$  long wormlike micelles in aqueous buffers.<sup>[4]</sup> Although the large size of these micelles appears ideal for micelle ELFSE separations, they entangle at low concentrations. The entanglement leads to a sieving type separation that reduces ELFSE separation performance.<sup>[6], [22]</sup> In order to reduce entanglement, smaller  $C_{12}E_5$  and  $C_{10}E_5$  surfactants are added to the  $C_{16}E_6$  solution. The shorter  $C_iE_j$  surfactants act as end-capping molecules that shorten overall micelle length, reducing interaction and entanglement.

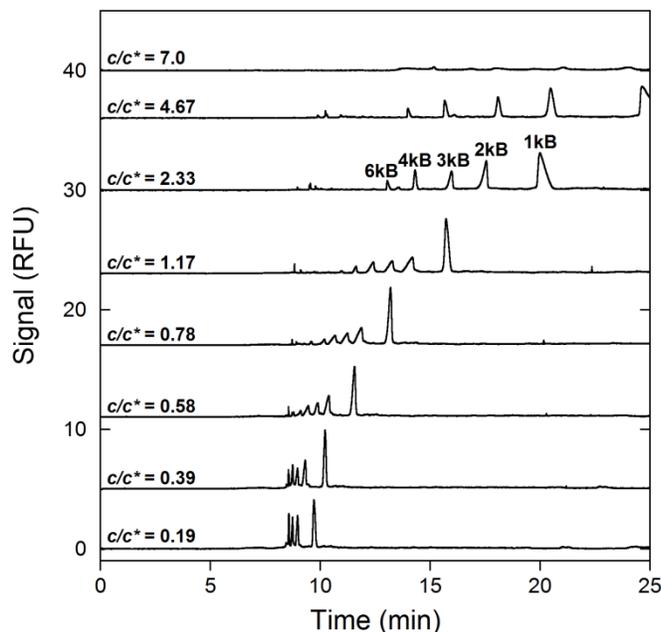
To determine the onset of entanglement, we measured the viscosity of serial dilutions of 36mM  $C_{16}E_6$ , 24mM  $C_{12}E_5$ , 3mM  $C_{10}E_5$ , termed 36/24/3. The overlap concentration ( $c^*$ ) can be measured from a plot of specific viscosity against total surfactant concentration (mM) on a log-log scale, as seen in Figure 7.4. Above  $c^*$ , the micelle solution is similar to an entangled polymer solution used to separate DNA by conventional gel/polymer matrix electrophoresis.<sup>[6]</sup> Two separate linear regimes are observed for the low and high concentrations.  $c^*$  is calculated as the concentration where the two regimes intersect, measured as 9mM total surfactant concentration, or a 5/3.5/0.5 ratio.



**Figure 7.4** – Specific viscosity of mixed C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> surfactant buffers in 1xTBE at 25°C. Solid line represents linear fit to data below  $c^*$ . Dotted line represents linear fit to data above  $c^*$ . Intersection of the fits indicates  $c^*$ .

### 7.4.2 – Long DNA Separations

Figure 7.5 shows representative electropherograms of 1 to 6kB C16B end-alkylated DNA separations as a function of surfactant concentration in the ternary C<sub>16</sub>E<sub>6</sub>, C<sub>12</sub>E<sub>5</sub>, C<sub>10</sub>E<sub>5</sub> buffers. As the concentration increases, the micelles grow in size and number, leading to increased elution times. The 36/24/3 ( $c/c^* = 7.0$ ) buffer previously used for separation of short DNA fragments yielded indistinguishable peaks, believed to be caused by micelle entanglement.



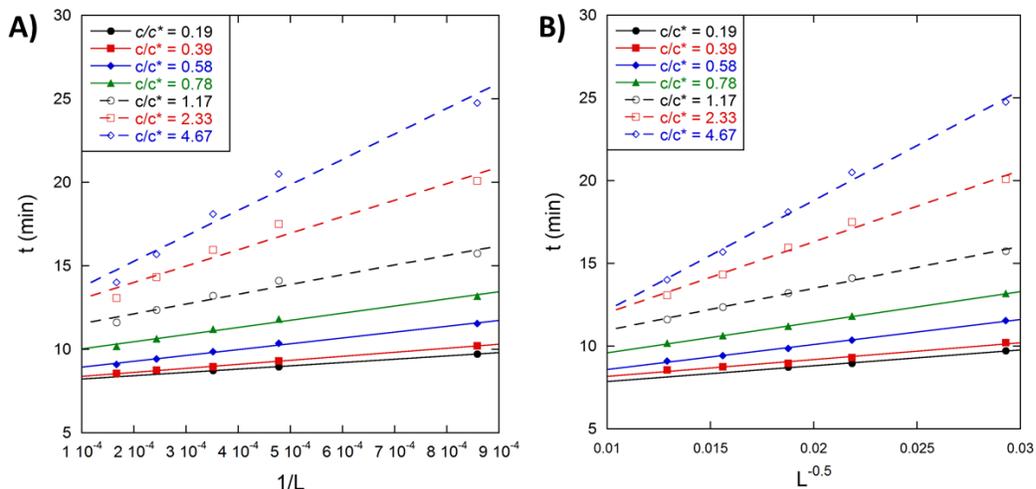
**Figure 7.5** – Representative electropherograms of 1-6kB C16B end-alkylated DNA separated in  $C_{16}E_6$ ,  $C_{12}E_5$ ,  $C_{10}E_5$  mixed surfactant 1xTBE buffers at 25°C. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.  $V = 5\text{kV}$

Plots of elution time vs.  $1/L$  are shown in Figure 7.6A. These plots should exhibit linear behavior according to standard, unsegregated ELFSE theory in Equation 7.2. The elution time ( $t$ ) follows as

$$t = \frac{l_d l_t}{\mu_0 V} \left( 1 + \frac{\alpha}{L} \right) \quad (7.14)$$

Where  $l_d$  is the capillary length to the detector,  $l_t$  is the total capillary length,  $V$  is the applied voltage,  $L$  is the DNA length,  $\mu_0$  is the free-solution mobility of DNA, and  $\alpha$  is the effective size of the micelle drag-tag in terms of hydrodynamically equivalent DNA bases. We have previously seen non-linear deviations from ELFSE theory for higher concentrations of  $C_{12}E_5$  surfactant, caused by the sieving pressures of the entangled micelle network. The entangled network increased

retardation of longer lengths than predicted by ELFSE, resulting in a peak compression effect that reduced separation efficiency.

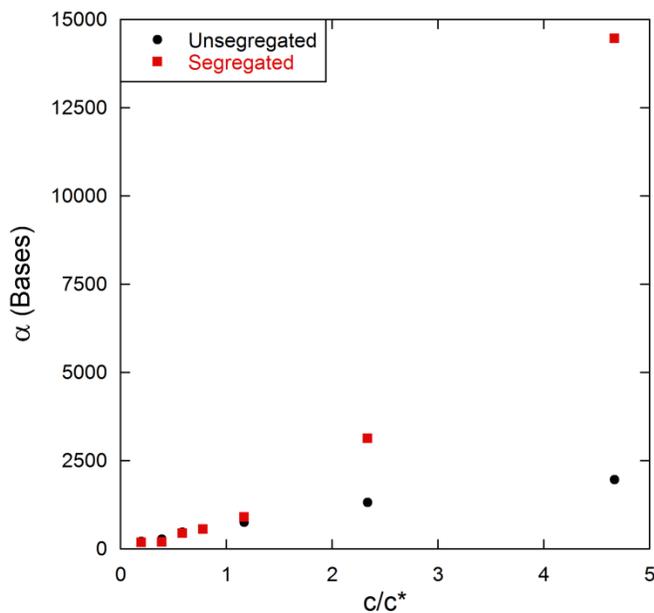


**Figure 7.6** – A)  $t$  vs  $1/L$  for the C16B end-alkylated DNA lengths separated in Figure 7.5. Linear fits demonstrate behavior according to standard, unsegregated ELFSE theory, Equation 7.14. B)  $t$  vs  $L^{-0.5}$  for the C16B end-alkylated DNA lengths separated in Figure 7.5. Linear fits demonstrate behavior according to segregated ELFSE theory, Equation 7.15.

Here, we observe much different non-linear behavior. Below  $c^*$ , the fits appear linear and are predicted by standard ELFSE behavior in Equation 7.14. Above  $c^*$ , an increasingly concave downward behavior is observed. This same concave behavior has been reported for stiff, dsDNA fragments sterically segregated from streptavidin drag-tags.<sup>[12]</sup> The segregated ELFSE model in terms of DNA elution time is found in Equation 7.15 and a representative plot of  $t$  vs.  $L^{-0.5}$  is in Figure 7.6B. We see much better agreement with the segregated model than the unsegregated model for high  $c/c^*$ .

$$t = \frac{l_d l_t}{\mu_0 V} \left( 1 + \frac{\sqrt{\alpha}}{\sqrt{L}} \right) \quad (7.15)$$

We can calculate the effective drag parameter,  $\alpha$ , for each  $c^*$  using the two models, see Figure 7.7. We can see that both models predict similar  $\alpha$  near and below  $c^*$ . However, at higher  $c/c^*$  the segregated model predicts a substantially higher  $\alpha$ . This may mark the transition from an unsegregated to segregated case. We would expect a very large micelle and  $\alpha$  for the highly resolved separations of long DNA in Figure 7.5. However, we should note that the ambiguity regarding application of each model leads to questions about the correct value of  $\alpha$  for high  $c/c^*$  buffers.

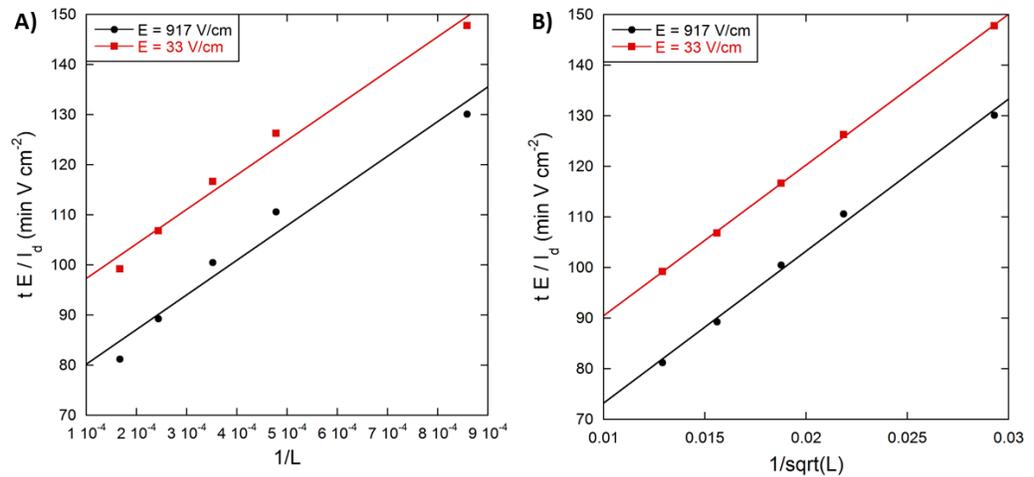


**Figure 7.7** – Effective drag-tag parameters for unsegregated and segregated ELFSE models vs.  $c/c^*$  of ternary  $C_{16}E_6$ ,  $C_{12}E_5$ ,  $C_{10}E_5$  surfactant.

### 7.4.3 – Effect of $E$ on Segregation

To test the effects of  $E$  on segregation, we performed separations at varying  $E$  using the  $c/c^* = 2.33$  buffer. Figure 7.8 shows the ELFSE behavior for  $E = 33$  V/cm, below the predicted  $E_c$  for any  $\alpha$  and  $L$ , and  $E = 917$  V/cm, well above the

predicted  $E_c$  segregation limit. The elution times in Figure 7.8 are scaled by the prefactor term  $l_d/E$  in Equations 7.14 and 7.15 to allow better comparison of  $t$  at different  $E$ . We can see that both systems are better fit by segregated behavior even though the  $E = 33$  V/cm case should not segregate according to Equation 7.12. Segregation of the DNA-micelle complex does not appear to be induced by high electrical forces. In fact, the low  $E$  separation exhibits more linear behavior according to the segregation model than the high  $E$  separation. The slight non-linear behavior in the high  $E$  separation may be caused by stretching of the DNA, resulting in different friction coefficients for the more rod like DNA.<sup>[12], [23]</sup>

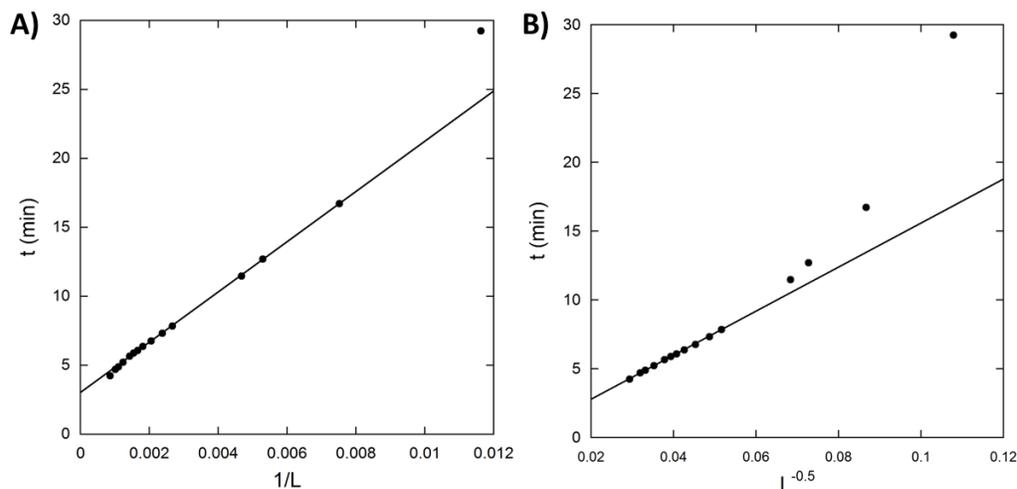


**Figure 7.8** – Scaled elution time vs. A)  $1/L$  (unsegregated model) and B)  $1/\sqrt{L}$  (segregated model) for  $c/c^* = 2.33$  surfactant buffer at different  $E$ .

#### 7.4.4 – Effect of $L$ on Segregation

To explain the non-linear ELFSE behavior, we consider the effects of steric segregation. The main  $C_{16}E_6$  component of our ternary buffer forms  $\mu\text{m}$  long wormlike micelles, compared to the hundreds of nanometers by the previously investigated pure  $C_{12}E_5$  system.<sup>[2], [4], [20]</sup> The segregated prediction of  $\alpha > 3,000$

above  $c^*$  may indicate that the micelles are too large even for kB DNA to maintain a hydrodynamically coupled random walk conformation. If this were true, we would expect segregated behavior for all DNA lengths below 6kB in this system. Figure 7.9 shows segregated and unsegregated ELFSE fits for separations of 86 to 1165 C16B end-alkylated DNA bases long in the  $c/c^* = 2.33$  buffer. The standard, unsegregated ELFSE model fits the DNA lengths from 133 – 551 quite well (Figure 7.9A). Non-linear deviations as concave downward behavior predicted by a segregated model are only observed for the longer lengths of DNA (Figure 7.9B). Thus, steric segregation due to the limited length of DNA available to encompass the micelle does explain the observed segregation effects.



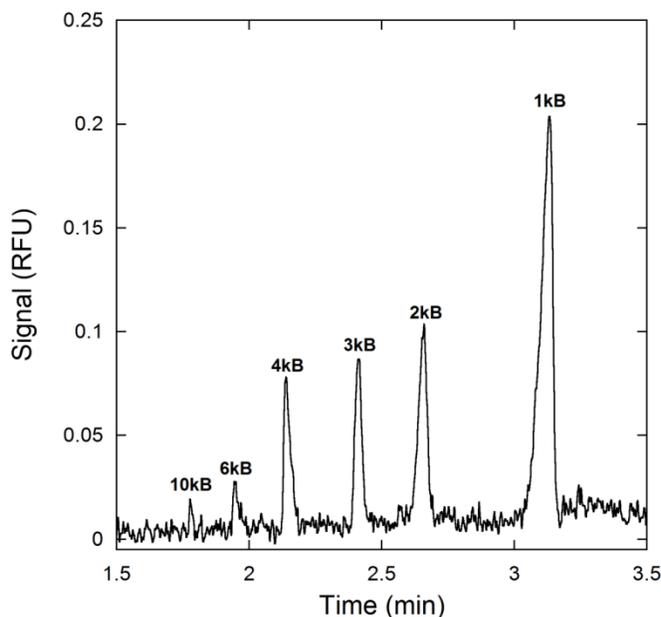
**Figure 7.9** – Comparison of A) unsegregated and B) segregated ELFSE models for separation of 86 – 1165 C16B end-alkylated DNA in ternary  $C_{16}E_6$ ,  $C_{12}E_5$ ,  $C_{10}E_5$  surfactant buffers at  $c/c^* = 2.33$ .

It is possible that the wormlike micelle microstructure is perturbing the system to a segregated state. Non-linear behavior was only observed for separations above  $c^*$ . The tertiary buffer may create pore spaces large enough for a random coil of DNA to pass through relatively unperturbed, minimizing the effects of

sieving on DNA. However, the combination of a large micelle-DNA conjugate may be too large to traverse the pore space. The polyelectrolyte DNA engine would most likely electrophorese through the pore space, with the large micelle lagging behind, leading to segregation.

### 7.4.5 – Rapid Long DNA Separations

The extremely large  $\alpha$  and reduced length dependence from segregation yields highly resolved separation of 1 – 10kB C16B end-alkylated DNA in 3 minutes, Figure 7.10.



**Figure 7.10** – Rapid separation of kB C16B DNA lengths using mixed surfactant system. Run Conditions: 12mM  $C_{16}E_6$ , 8mM  $C_{12}E_5$ , 1mM  $C_{10}E_5$  ( $c/c^* = 2.33$ ) in 1xTBE, 25°C, 25kV,  $l_d = 20$ cm,  $l_t = 30$ cm,  $d = 50\mu$ m, 4kV 30sec injection from formamide.

The resolution factors are seen in Figure 7.11, where the 1 – 4kB peaks demonstrate  $R \sim 100$ . The 6 and 10kB fragments were excluded due to the poor estimates of FWHM from their low signal. Each  $R$  was calculated using Equation

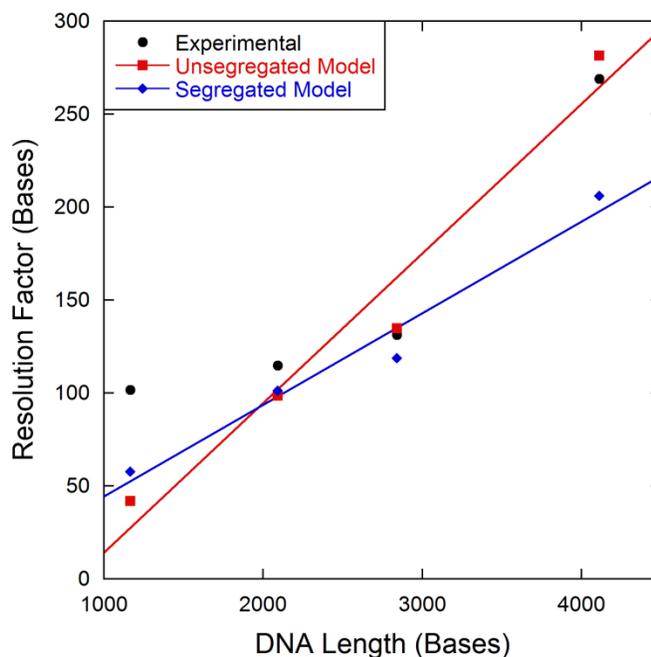
7.16, where  $FWHM$  is the elution peak's full-width at half maximum, and  $dt/dL$  is the peak spacing between sequential lengths of DNA.

$$R = \frac{FWHM}{\left| \frac{dt}{dL} \right|} \quad (7.16)$$

Predictions of  $dt/dL$  from the derivatives of unsegregated and segregated theory are given in Equation 7.17 and 7.18, respectively. Both  $dt/dLs$  were fit using the ELFSE parameters for  $\alpha$  and  $\mu_0$  for their respective models. We can see there is good agreement between experimental  $R$  and both predicted  $R$ s. This may be due to the ambiguity of which  $\alpha$  and  $\mu_0$  is correct. However, an important distinction between the models is evident in Figure 7.11. The segregated model  $dt/dL$  scales with  $L^{-3/2}$ , creating a weaker dependence on  $R$  for  $L$  and extending the resolution factors for longer DNA fragments.

$$\left| \frac{dt}{dL} \right| = \frac{l_d l_t \alpha}{\mu_0 V L^2} \quad (7.17)$$

$$\left| \frac{dt}{dL} \right| = \frac{l_d l_t \sqrt{\alpha}}{2\mu_0 V L^{3/2}} \quad (7.18)$$



**Figure 7.11** – Resolution factors for peaks in Figure 7.10, with predictions from unsegregated  $dt/dL$  in Equation 7.17 (red squares, red line) and segregated  $dt/dL$  in Equation 7.18 (blue diamonds, blue line)

## 7.5 – Conclusion

Mixed surfactant systems offer the potential of minimizing entanglement while maintaining large micelle size. We find that a ternary system of  $C_{16}E_6$ ,  $C_{12}E_5$ ,  $C_{10}E_5$  allowed rapid separation of kB DNA fragments without detrimental sieving effects. We find that the DNA micelle conjugate hydrodynamically segregated, improving the separation of long DNA. Segregation is not apparently caused by high electrical forces or steric factors. We believe it is caused by the wormlike micelle network. These types of buffers should aid in our ability to detect fM concentrations of microRNA by hybridizing kB DNA probes loaded with fluorescent intercalating dyes.<sup>[24]</sup>

## 7.6 – References

- [1] T. Kato, S. Anzai, and T. Seimiya, *J. Phys. Chem.*, 1990, 94, 7255–7259.
- [2] S. Shirai and Y. Einaga, *Polym. J.*, 2005, 37, 913–924.
- [3] A. Jones, “Rapid Separations of Kilobase-Sized DNA using Micelle Size-Sampling,” Ph.D. dissertation, Dept. Chem. Eng, Carnegie Mellon University, Pittsburgh, PA, 2013.
- [4] Y. Einaga, Y. Inaba, and M. Syakado, *Polym. J.*, 2006, 38, 64–72.
- [5] Z. Lin, L. E. Scriven, and H. T. Davis, *Langmuir*, 1992, 8, 2200–2205.
- [6] Wei and E. S. Yeung, *Anal. Chem.*, 2001, 73, 1776–1783.
- [7] Y. Einaga, Y. Kito, and M. Watanabe, *Polym. J.*, 2006, 38, 1267–1277.
- [8] S. Puvvada and D. Blankschtein, *J. Phys. Chem*, 1992, 96, 5579–5592.
- [9] M. E. Cates and S. J. Candau, *J. Phys. Condens. Matter*, 1999, 2, 6869–6892.
- [10] L. C. McCormick and G. W. Slater, *Electrophoresis*, 2007, 28, 674 – 682.
- [11] C. Desruisseaux, D. Long, G. Drouin, and G. W. Slater, *Macromolecules*, 2001, 34, 44–52.
- [12] H. W. Lau and L. A. Archer, *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 2010, 81, 1–8.
- [13] T. Mukaiyama, *Angew. Chemie Int. Ed. English*, 1976, 15, 94–103.
- [14] A. Deratani and T. Maraldo, *React. Polym.*, 1988, 9, 19–28.
- [15] T. Da Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, and B. Cacciari, *Curr. Med. Chem.*, 2005, 12, 71–88.
- [16] D. Long and A. Ajdari, *Electrophoresis*, 1996, 17, 1161–1166.
- [17] L. C. McCormick, G. W. Slater, A. E. Karger, W. N. Vreeland, A. E. Barron, C. Desruisseaux, and G. Drouin, *J. Chromatogr. A*, 2001, 924, 43–52.
- [18] W. Van Saarloos, *Phys. A Stat. Mech. its Appl.*, 1987, 147, 280–296.
- [19] B. Tinland, a Pluen, J. Sturm, and G. Weill, *Macromolecules*, 1997, 9297, 5763–5765.

- [20] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [21] G. Porte, J. Appell, and Y. Poggi, *J. Phys. Chem.*, 1980, 3105–3110.
- [22] C. Desruisseaux, G. Drouin, and G. W. Slater, *Macromolecules*, 2001, 34, 5280–5286.
- [23] L. C. McCormick and G. W. Slater, *Electrophoresis*, 2007, 28, 3837–3844.
- [24] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.



# Chapter 8 – Improved DNA Sequencing by Reducing Micelle Entanglement

## 8.1 – Introduction

The advent of personalized medicine is dependent on fast, cheap whole-genome sequencing.<sup>[1]</sup> Capillary gel electrophoresis (CGE) was vital in completion of the Human Genome Project in 2003.<sup>[2], [3]</sup> However, it is ill suited for personalized medicine due to its long separation times (1,000 bases in 1 hour).<sup>[4]</sup> Alternative non-electrophoretic methods, such as pyro-sequencing and nanopore sequencing, have demonstrated potential to achieve efficient whole-genome sequencing.<sup>[5], [6]</sup> However, the read length of pyro-sequencing is relatively short (100-200 bases) and the error rate of nanopore sequencing is still too high for use as a reliable sequencing method.<sup>[7], [8]</sup>

End-labeled free-solution electrophoresis (ELFSE) has emerged as a much faster alternative to DNA sequencing than traditional CGE.<sup>[9]</sup> In ELFSE, DNA strands are end-labeled with an uncharged, monodisperse molecule, termed a “drag-tag”, to generate additional hydrodynamic drag on the electrophoresing DNA. Longer DNA strands are able to overcome the additional drag from the end-label and electrophoresis with a higher velocity than shorter lengths, creating a separation order opposite of gels.

An ever increasing push in ELFSE is to extend the length of read (*LOR*) for DNA sequencing separations.<sup>[10]-[15]</sup> However, drag-tags have been limited to small sizes to maintain monodispersity requirements for the DNA sequencing single base resolution. To date, the best ELFSE DNA sequencing *LOR* for covalently bound drag-tags was 265 in a 30 minute separation, not much of an improvement compared to gel systems.<sup>[11]</sup>

We have previously demonstrated the use of non-ionic surfactants as transiently attached drag-tags for efficient ELFSE separations.<sup>[16]-[18]</sup> We have achieved DNA sequencing of 110 bases in 4 minutes using the small Triton X-100 micelle, and a *LOR* of 420 bases in 45 minutes using a mixed C<sub>12</sub>E<sub>5</sub> C<sub>10</sub>E<sub>5</sub> 3M urea surfactant system.<sup>[15]</sup> However, DNA sequencing requires the presence of high concentrations of urea (7M) in order to maintain denaturing conditions for the single stranded DNA, as well as to increase fluorescence intensity of the hydrophobic FRET coupled dyes in running buffers.<sup>[19]</sup> Our Triton X-100 and 3M urea buffer were only able to detect a single dye from the BigDye® DNA sequencing kit, making implementation of the 4 color sequencing kits difficult.

The presence of urea acts to disrupt hydrogen bonding between water molecules, increasing both the solvation of the hydrophobic surfactant monomers and solvated head group size.<sup>[20]</sup> This leads to transition from wormlike micelles to globular conformations.<sup>[21]</sup> The spherical micelles occupy a smaller hydrodynamic volume than wormlike micelles, creating a lesser amount of hydrodynamic drag and reducing  $\alpha$  for ELFSE separations.<sup>[14]</sup> Our efforts to

create a high drag 7M urea DNA sequencing micelle buffer have been limited, to a *LOR* of 300 bases in 18 minutes.

Here, we investigate the use of 7M urea with our large, entangled C<sub>12</sub>E<sub>5</sub> buffer system. With an excessively high drag in 0M urea, the presence of 7M urea may still lead to large micelles for use as a high *LOR* DNA sequencing buffer. Additionally, the transition from long wormlike to spherical micelles in 7M urea may reduce entanglements that are detrimental to high resolution DNA sequencing. We find that addition of 7M urea still yields large micelles for separation, but the presence of an entangled micelle network leads to poor DNA sequencing results (*LOR* < 300). We demonstrate an increase in DNA sequencing *LOR* to 500 bases by doping in a small amount of C<sub>10</sub>E<sub>5</sub> surfactant to the 7M urea C<sub>12</sub>E<sub>5</sub> surfactant buffer, in addition to fine adjustments of separation temperature, both to reduce entanglement. Surprisingly, the highest *LOR* is achieved in a slightly entangled network. There is a trade-off in *LOR* between maintaining large micelle size at higher temperatures with reduced network formation at lower temperatures. This trade-off can be easily monitored using the viscosity of surfactant solutions, and offers a much simpler experimental protocol for future surfactant buffer optimization.

## **8.2 – Materials and Methods**

### **8.2.1 – Reagents**

Nuclease-free water, acetic acid, GeneJET Gel Extraction Kits, and acetone were purchased from Fisher Scientific (Pittsburgh, PA). BODIPY<sup>®</sup> FL C<sub>16</sub>

(C16B), low melting temperature agarose, BigDye<sup>®</sup> Terminator v1.1 Cycle Sequencing Kit, Hi-Di formamide, Centri-Sep spin columns, and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). Benzyl alcohol, DMSO, 2,2-Dithiodipyridine (DPDS), Lithium perchlorate, Triphenylphosphine (TPP), 4-(Dimethylamino)pyridine (DMAP), Hexadecyltrimethyl ammoniumbromide (CTAB), ethidium bromide, acetonitrile, urea, 10x Tris-Borate-EDTA (TBE), stearic acid, and triethylamine (TEA) were purchased from Sigma Aldrich (St. Louis, MO). *N*-dodecylpentaoxyethylene (C<sub>12</sub>E<sub>5</sub>) and *N*-decylpentaoxyethylene (C<sub>10</sub>E<sub>5</sub>) surfactant were purchased from BaChem (Torrance, CA). λDNA, m13mp18 single stranded DNA, and LongAmp<sup>™</sup> *Taq* DNA polymerase Kit were purchased from New England Biolabs (Ipswich, MA). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2mΩ deionized water and filtering through a 0.22μm filter. PCR forward primers and C<sub>6</sub>-amine 5' modified reverse primers were purchased with standard desalting purification from Integrated DNA Technologies (Coralville, IA). Dried oligomers were resuspended in 10mM Tris 0.1mM EDTA buffer and used without further purification.

### **8.2.2 – DNA Alkylation Reaction**

Alkylation of C<sub>6</sub>-amine 5' modified reverse primer with C16B or stearic acid (C18) was accomplished using a combination of methods outlined by Mukaiyama and Deratani, and closely followed a scheme first reported by Boutorine *et al.*<sup>[22]</sup>–<sup>[24]</sup>. An aliquot of 10nmol 5'amine modified reverse primer was precipitated in the presence of 750nmol CTAB. The suspension was then dried by vacuum

centrifugation under low heat for 30 minutes. Meanwhile, a mixture of 100nmol C16B or C18, 5 $\mu$ mol DMAP, 2.5 $\mu$ mol TPP, 2.5 $\mu$ mol DPDS, and 7.5 $\mu$ mole TEA in DMSO was prepared and allowed to activate at room temperature for 20 minutes. Once activated, the reaction mixture was added to the dried DNA suspension, briefly vortexed and bath sonicated to resuspend the DNA pellet, and allowed to react overnight under continuous agitation in a shaker at room temperature in the dark. The solution was then precipitated with 1mL of 2w/v% lithium perchlorate in acetone and pelleted by centrifugation. The supernatant was removed and the pellet was triple rinsed with acetone. The pellet was resuspended in acetone, centrifuged, and supernatant removed again. The pellet was then dried by vacuum centrifugation under low heat for 5 minutes and resuspended in 0.1M triethylammoniumacetate pH 7.0 (TEAA) buffer.

Purification of the reaction mixture was performed using a Waters 4.6 mm x 250 mm Symmetry300 C18 HPLC column (Milford, MA) with a 1mL/min flowrate and linear gradient from 0.1M TEAA to 100% acetonitrile over 30 minutes. Any unmodified DNA eluted near 10 minutes while the end-alkylated DNA eluted near 22 minutes. Product fractions were collected, lyophilized, and resuspended in 10mM Tris 0.1mM EDTA buffer.

### **8.2.3 – Lower Critical Solution Temperature (LCST) Determination**

The LCST was determined using a Cary300 UV-Vis (Agilent Technologies, Santa Clara, CA). Each surfactant concentration was heated in a cuvette at 1 $^{\circ}$ C/min, and its 488nm absorbance measured every 0.5 $^{\circ}$ C. A sharp increase in

absorbance indicated a more turbid sample, and therefore the LCST. The LCST of the C<sub>12</sub>E<sub>5</sub> in 7M urea 1xTBE was found to be 48°C and largely independent of concentration. The small addition of C<sub>10</sub>E<sub>5</sub> did not appreciably change the LCST.

#### **8.2.4 – PCR Amplification of C16B End-alkylated DNA Ladder**

Specific DNA lengths of C16B end-alkylated DNA were amplified using LongAmp<sup>TM</sup> Taq DNA polymerase, λDNA template, and a C16B end-alkylated λDNA reverse primer, and a length setting forward primer. The C16B end-alkylated DNA ladder included the following lengths of DNA: 86, 133, 189, 214, 265, 375, 421, 488, 551, 603, 647, 707, 787, 808, 847, 909, 986, and 1165. Each PCR reaction contained 1ng of λDNA template, 500nM each primer, 0.3mM each dNTP, 2.5 units of LongAmp<sup>TM</sup> Taq DNA polymerase, 1x LongAmp<sup>TM</sup> buffer, and enough nuclease free water to 25μL. The reaction was denatured for 2min at 95°C and subjected to 35 cycles of denaturation at 95°C for 15s, primer annealing at 50°C for 30 s, and extension at 65°C for 50 seconds per 1000 bases of amplicon length, followed by a final hold for 10min at 65°C.

After thermal cycling, the PCR reaction mixture was directly loaded onto 1wt% agarose gels. Slab gel electrophoresis was performed at 4 V/cm for 2 hours. Gels were imaged using ethidium bromide staining and UV excitation on a BioDoc-It Imaging System from Ultra-Violet Products (Cambridge, UK). Correct PCR product length bands were excised and purified using GeneJET Gel Extraction Kit spin columns, and eluted using the provided 10mM Tris 0.1mM EDTA buffer.

#### **8.2.5 – Sanger Sequencing C18 End-alkylated DNA Amplification**

Sanger sequencing products were amplified using a C18 end-alkylated m13mp18 reverse primer and m13mp18 single stranded DNA template. The PCR reaction mixture contained 4 $\mu$ L of BigDye<sup>®</sup> Terminator v1.1 Master Mix, 500ng of m13mp18 template, 320nM C18 end-alkylated primer, and enough nuclease-free water to 10 $\mu$ L. The reaction was denatured for 5min at 95°C and subjected to 35 cycles of denaturation at 95°C for 30s, primer annealing at 50°C for 30 s, and extension at 60°C for 4min, followed by a final hold for 4 min at 60°C.

After thermal cycling, the Sanger sequencing reaction mixture was purified using Centri-Sep spin columns. In order to increase fluorescent signal, 2 reactions were combined and added to a single Centri-Sep column. The purified sample was resuspended in 30 $\mu$ L of formamide, heated to 95°C for 2 min, and snap-cooled on ice to ensure the DNA was properly denatured.

### **8.2.6 – Capillary Electrophoresis of C16B PCR Products**

Analysis of C16B end-alkylated DNA ladders was performed on a P/ACE MDQ capillary electrophoresis instrument (Beckman Coulter, Fullerton, CA) equipped for laser induced fluorescence (LIF) detection. Excitation sources were from a 3mW 488nm Argon ion. LIF detection was performed at 488/520nm excitation/emission. The capillary was a 30 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) and a length of 20cm to the detector and 30cm total. The C16B end-alkylated DNA ladders were prepared by pipetting 2 $\mu$ L of each purified PCR product with 30 $\mu$ L of formamide, heating to 95°C for 2 minutes to ensure denatured single stranded DNA products, and immediately cooled on ice. 4kV for 30sec electrokinetic injections were used to introduce

samples into the capillary. Electrophoretic separation was conducted under an applied voltage of 30kV in reverse polarity mode (from anode to cathode). The capillary was heated to 40°C using a built in coolant jacket system. At 45°C the separation failed, believed to be caused by operating close to the LCST of C<sub>12</sub>E<sub>5</sub> 7M urea. Data collection was performed using 32 Karat software (Beckman Coulter). Peaks were analyzed using OriginPro 9.0 peak fitting software (Northampton, MA).

Electroosmotic flow (EOF) was suppressed by first rinsing the capillary with 10v/v% POP-6 polymer and maintained by adding 0.1v/v% POP-6 to all capillary electrophoresis buffers. Surfactant buffers were prepared by mixing the indicated concentration of C<sub>12</sub>E<sub>5</sub> and C<sub>10</sub>E<sub>5</sub> in 7M urea 1xTBE, briefly vortexing, and shaken for at least 1 hour at room temperature to ensure surfactant solubilization and micelle formation. The capillary could be used for many separations without observing any detrimental effects.

### **8.2.7 – Capillary Electrophoresis of C18 Sanger Sequencing Products**

Analysis of C18 Sanger sequencing products was performed on an ABI 310 capillary electrophoresis instrument (Life Technologies) equipped for LIF detection. The capillary was a 30µm ID fused-silica capillary, 31cm to the detector, and 41cm in total length. The capillary was flushed with 10v/v% POP-6 polymer in 1xTBE to suppress EOF using the built in syringe pump. The capillary was then flushed with the indicated surfactant buffers containing 0.1v/v% POP-6. C18 Sanger sequencing DNA product was then injected into the capillary at

2.5kV for 15sec. Separations were performed at 15kV and of indicated temperatures using a built in hot air convection heating system. Peaks were analyzed using OriginPro 9.0 peak fitting software (Northampton, MA).

### **8.2.8 – Viscosity Measurements**

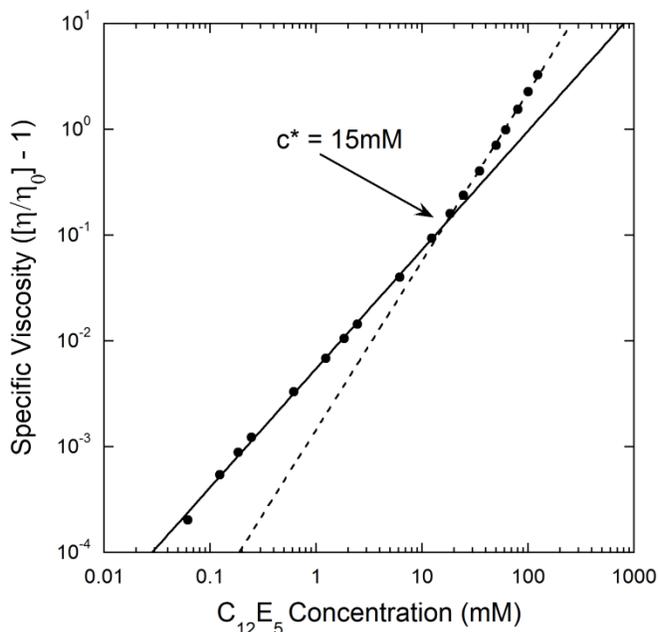
Viscosities of surfactant solutions were measured via capillary viscometry using a P/ACE MDQ (Beckman Coulter, Fullerton, CA) equipped for UV detection. . The capillary was a 50 $\mu$ m ID fused-silica capillary (Polymicro Technologies, Phoenix, AZ) with a length of 20cm to the detector and 30cm total. A small plug of concentrated, UV absorbing benzyl alcohol was injected into the capillary filled with surfactant solution. A pressure of 1 psi was used to hydrodynamically push the plug past the detector. The viscosity of the surfactant solution was determined using the elution time of the benzyl alcohol plug and assuming Hagen-Poiseuille plug flow.

## **8.3 – Results**

### **8.3.1 – Determination of $c^*$**

The specific viscosity of C<sub>12</sub>E<sub>5</sub> in 7M urea is plotted against surfactant concentration (mM) on a log-log scale in Figure 8.1. Two separate linear regimes are observed for the low and high concentrations. The intersection of these regimes indicates micelle interaction at the overlap concentration ( $c^*$ ), measured as 15mM. This corresponds to the same  $c^*$  for 0M urea C<sub>12</sub>E<sub>5</sub>. The similarity may represent the fact that as the micelles decrease in size and form more globular structures, the concentration of micelles increases. More surfactant monomer

would be available to form a greater number of micelles in the high urea solution. In such a case,  $c^*$  would be similar for a greater number of spherical micelles vs. larger wormlike micelles. However, we would expect a weaker entanglement interaction from the transition to smaller micelles.

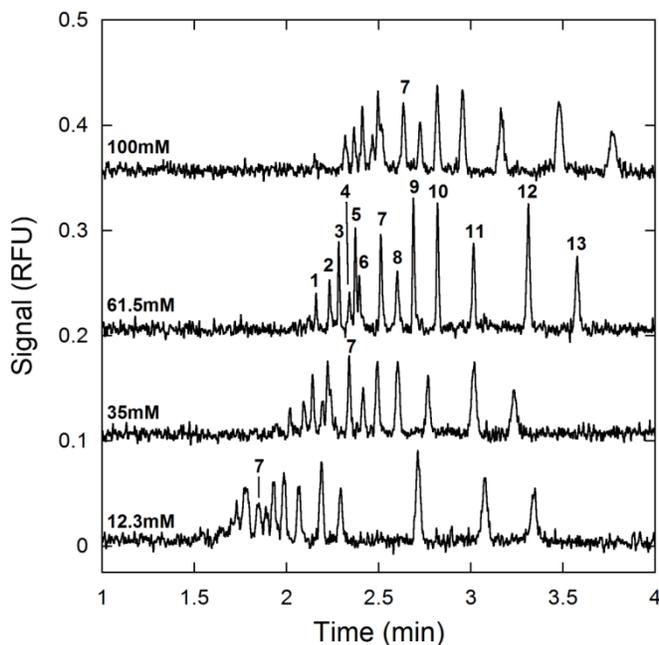


**Figure 8.1** – Specific viscosity of  $C_{12}E_5$  surfactant in 7M urea 1xTBE at 40°C. Solid line represents linear fit to data below  $c^*$ . Dotted line represents linear fit to data above  $c^*$ . Intersection of the fits indicates  $c^*$ .

### 8.3.2 – C16B End-alkylated DNA Separations

Figure 8.2 shows representative electropherograms of  $C_{12}E_5$  7M urea separation of C16B end-alkylated DNA ladders. For all  $C_{12}E_5$  concentrations, the micelles are large enough to resolve the moderate PCR product lengths. As the concentration increases, the elution times become more retarded due to the growth of the micelles. All lengths are resolved for the 61.5mM buffer, however at 100mM, the peaks appear to broaden and the triplet of 842, 808, and 747 is not

resolved. This poor resolution and broadening is due to the increased sieving pressures at higher concentrations as an entangled network is formed.

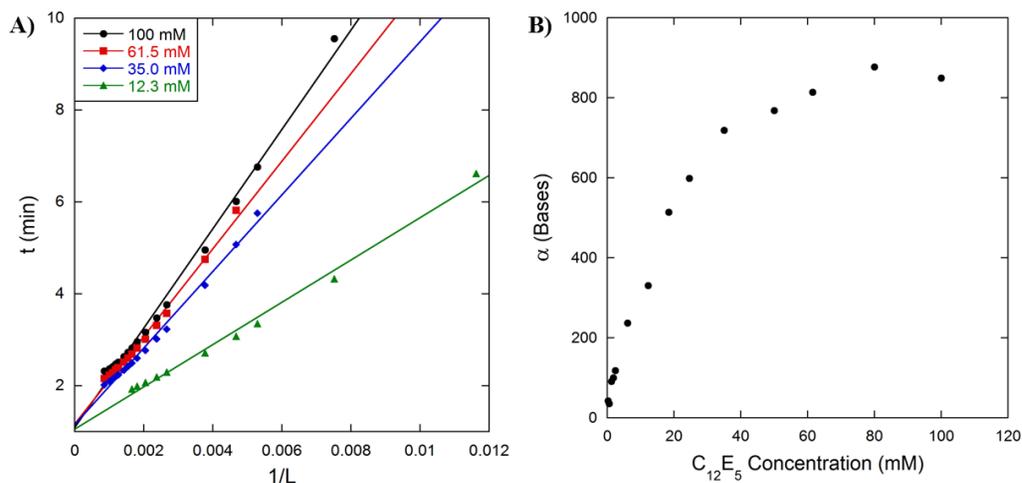


**Figure 8.2** – Representative electropherograms of C16B end-alkylated DNA separated in C<sub>12</sub>E<sub>5</sub> 7M urea 1xTBE buffers at 40°C. Numbers above peaks correspond to lengths (1) 1165, (2) 986, (3) 909, (4) 842, (5) 808, (6) 787, (7) 701, (8) 647, (9) 603, (10) 551, (11) 488, (12) 421, (13) 375. Electropherograms are arbitrarily shifted in the positive y-direction for display purposes.

Figure 8.3A shows the plots of elution time ( $t$ ) vs.  $1/L$  for the C<sub>12</sub>E<sub>5</sub> 7M urea buffers. Linear behavior is fit according to ELFSE theory in Equation 8.1, where  $l_d$  is the capillary length to the detector,  $l_t$  is the total capillary length,  $V$  is the applied voltage,  $L$  is the DNA length,  $\mu_0$  is the free-solution mobility of DNA, and  $\alpha$  is the effective size of the micelle drag-tag in terms of hydrodynamically equivalent DNA bases.

$$t = \frac{l_d l_t}{\mu_0 V} \left( 1 + \frac{\alpha}{L} \right) \quad (8.1)$$

As the surfactant concentration increases, the micelles grow in size and a higher slope, representing the  $\alpha$  drag-tag parameter, is seen. Once again, deviations from linear behavior are observed for higher lengths and higher concentrations of surfactant due to the entangled micelle network. However, the deviations are much less severe than the 0M urea case discussed in Chapter 6. The intercepts for all curves originate at a similar value, whereas previous intercepts for the 0M urea case were seen to significantly change. The smaller deviations are most likely caused by the smaller micelles in 7M urea forming a less entangled network.



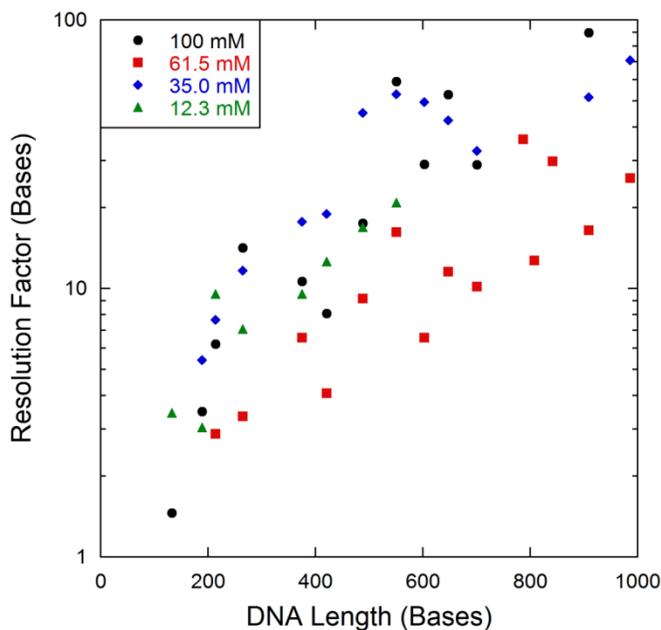
**Figure 8.3** – A) Elution time ( $t$ ) vs  $1/L$  for the C16B end-alkylated DNA lengths separated in  $C_{12}E_5$  7M urea buffers in Figure 8.2. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Effective micelle drag-tag size,  $\alpha$ , vs.  $C_{12}E_5$  surfactant concentration, calculated from the slope and intercept of Figure 8.3A, according to Equation 8.1

A plot of  $\alpha$  vs.  $C_{12}E_5$  concentration is seen in Figure 8.3B. The  $\alpha$  value increases with concentration and plateaus at a value of roughly  $4x c^*$ , just as in the 0M urea case. Previous 7M  $C_iE_j$  surfactant systems achieved an  $\alpha$  of  $\sim 50$  and  $LOR$  of  $\sim 300$  in 18 minutes. With an  $\alpha$  of 850, substantial increases in the  $LOR$  for DNA sequencing applications are possible. We also note that the 7M urea  $C_{12}E_5$

surfactant system has smaller  $\alpha$  values than the 0M urea case, as expected with the addition of urea creating more globular micelle structures.<sup>[20]</sup>

DNA sequencing requires single base resolution ( $R < 1.5$ ), given by Equation 8.2 as the ratio of the full-width half-maximum ( $FWHM$ ) of the elution peak and peak spacing ( $dt/dL$ ).

$$R = \frac{FWHM}{|dt/dL|} \quad (8.2)$$



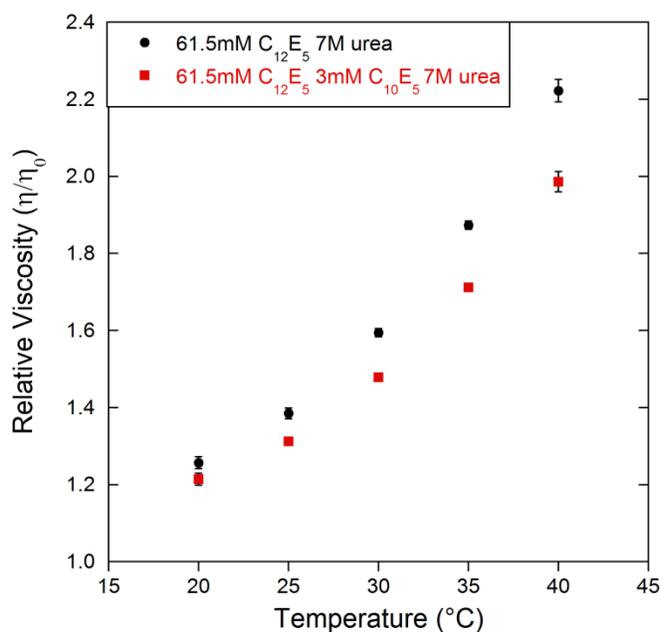
**Figure 8.4** – Resolution factors calculated for different C16B end-alkylated DNA lengths using Equation 8.2 from peaks in Figure 8.2.

The resolution factors for the C16B end-alkylated DNA separations are in Figure 8.4. The 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea buffer had the best separation efficiency with the lowest resolution factors, however they are still well above the  $R = 1.5$  limit for sequencing. The presence of a micelle network appears to be once again

reducing separation efficiency. This is evident in the high resolution factors of the 100mM C<sub>12</sub>E<sub>5</sub> 7M urea buffer

### 8.3.3 – Reduced Temperature DNA Sequencing

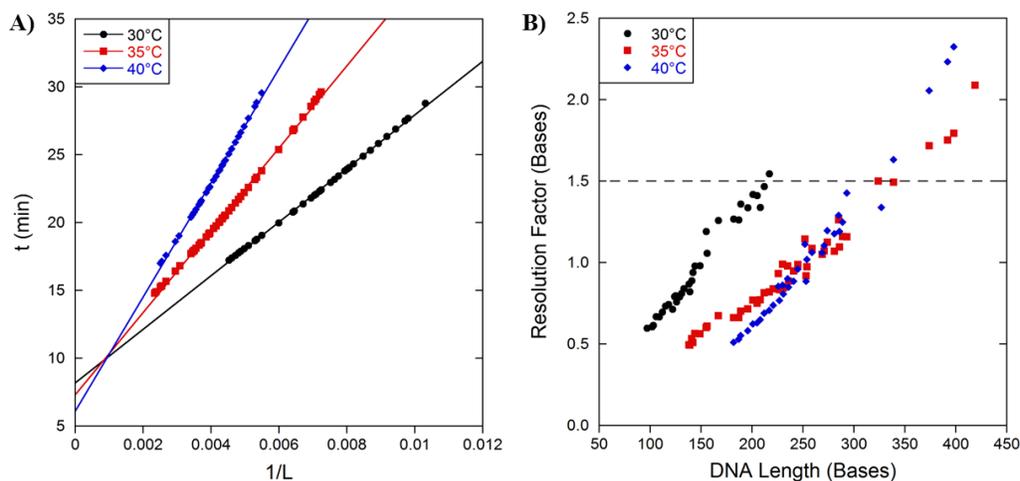
In order to reduce the presence of the entangled micelle network and decrease  $R$  to the  $R = 1.5$  limit for sequencing, we explore the effects of lowering separation temperature. Wormlike surfactant micelles shrink as the solution temperature is lowered, due to the increased solvation of the hydrophilic head group.<sup>[25], [26]</sup> The relative viscosity of 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea decreases with temperature, indicating a less entangled system, see Figure 8.5.



**Figure 8.5** – Relative viscosity of 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea (black circles) and 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> (red squares). A higher relative viscosity represents a greater extent of micelle interaction. Error bars represent standard deviation for  $n = 3$ .

We performed DNA sequencing separations using 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea at 30°C, 35°C, and 40°C to test its effect on sequencing  $LOR$ . Plots of  $t$  vs.  $l/L$  from

the T nucleotide channel sequencing separations are seen in Figure 8.6A. As the temperature decreases, the slope of the linear fits decreases, indicating a smaller  $\alpha$  value for the buffer system. From the fits, we find an  $\alpha$  value of 241, 373, and 664 for 30°C, 35°C, and 40°C, respectively.



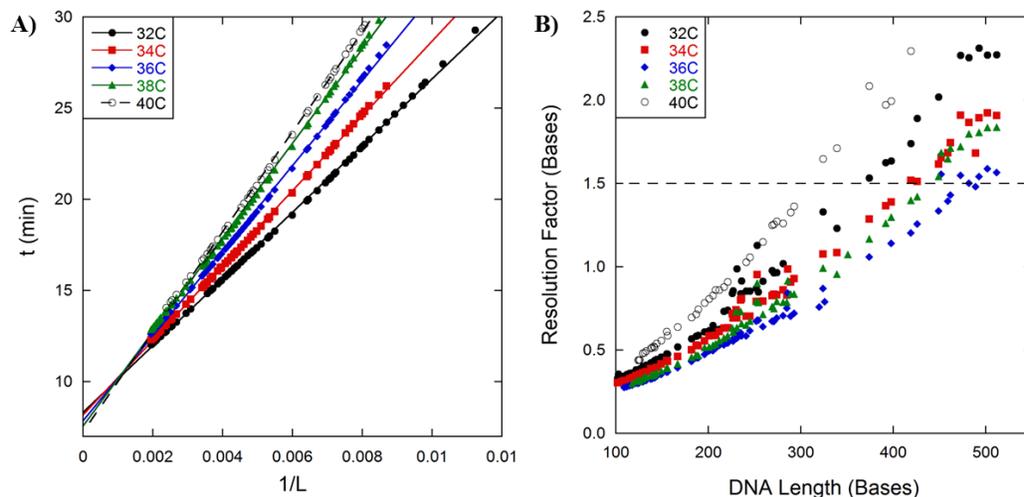
**Figure 8.6** – A) Elution time ( $t$ ) vs  $1/L$  for T channel BigDye<sup>®</sup> C18 end-alkylated DNA sequencing products separated in 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea 1xTBE buffer on ABI310 at different temperatures. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Resolution factors calculated for or T channel BigDye<sup>®</sup> C18 end-alkylated DNA sequencing products using Equation 8.2. Dotted line represents resolution limit for DNA sequencing, corresponding to the *LOR*.

The resolution factors for the T channel DNA sequencing results are seen in Figure 8.6B. The *LOR* for the separation can be found when a specific DNA length meets the  $R = 1.5$  limit. We find a *LOR* of 221, 293, and 293 for 30°C, 35°C, and 40°C, respectively. Reducing temperature was not able to fully alleviate the entangled network from decreasing separation efficiency in the large micelle buffer. The previous 7M urea micelle buffer had an  $\alpha \sim 50$  and *LOR* of  $\sim 300$  under the same run conditions. Even though the  $\alpha$ s here are more than 5 fold larger, the *LOR*s are less. Clearly the presence of an entangled network greatly limits DNA sequencing separation efficiency.

### 8.3.4 – C<sub>10</sub>E<sub>5</sub> End-Capping DNA Sequencing

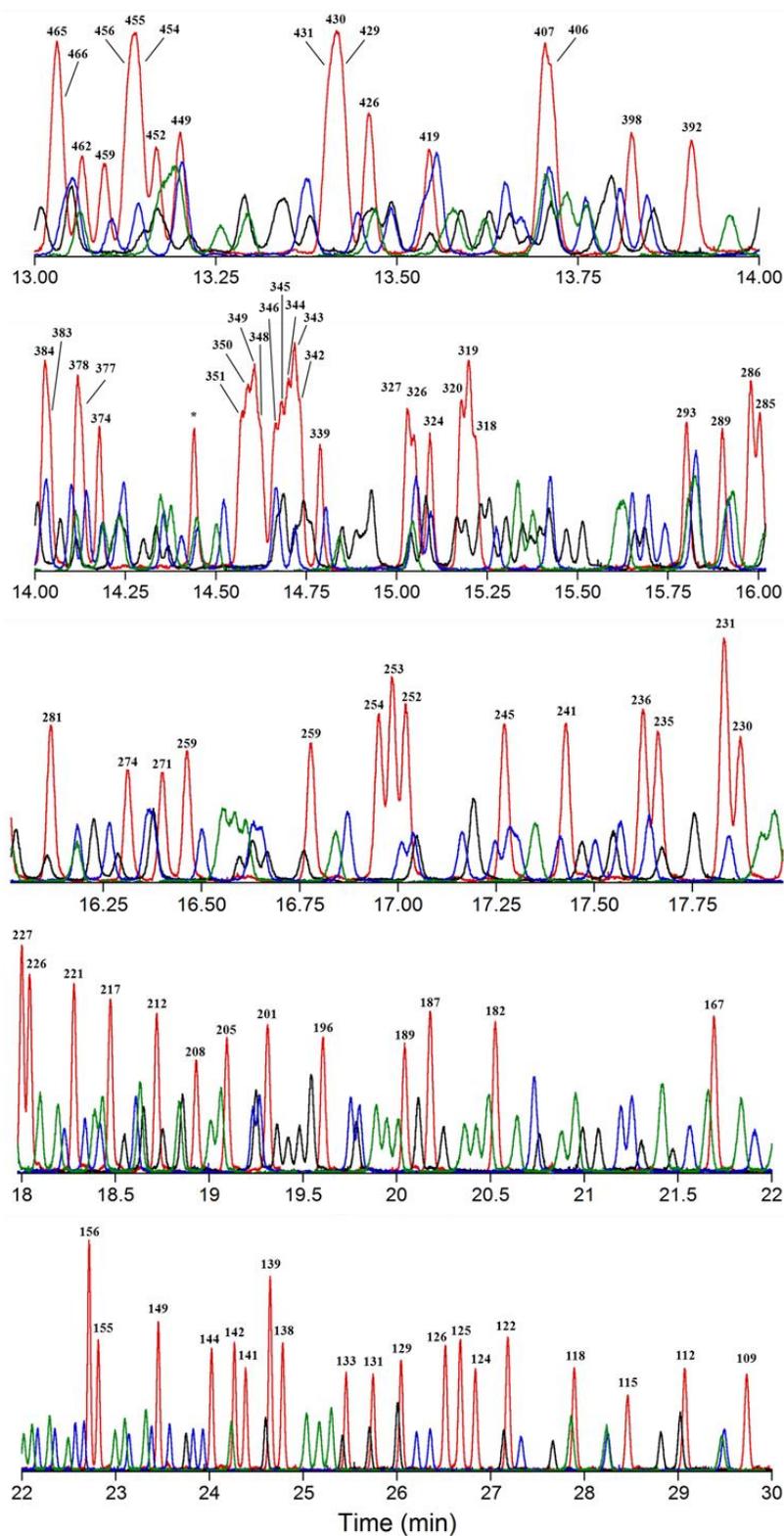
To further reduce the presence of an entangled network, we added a small fraction of C<sub>10</sub>E<sub>5</sub> end-capping surfactant to the system. Figure 8.5 shows the scaled viscosity of 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea buffers vs. temperature. The addition of 3mM C<sub>10</sub>E<sub>5</sub> to the 61.5mM C<sub>12</sub>E<sub>5</sub> 7M urea buffer reduces the relative viscosity compared to pure 61.5mM C<sub>12</sub>E<sub>5</sub>, indicating a slightly less entangled micelle system.

The temperature was also varied for the 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE separations to test for the effect of possible sieving on the sequencing performance. Figure 8.7A shows plots of  $t$  vs.  $l/L$  for the 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea buffers for the T base channel with linear fits according to Equation 8.1. As expected, an increase in slope is observed with increasing temperature. This indicates that the micelles are growing in size with temperature. The  $\alpha$  values are 219, 233, 300, 341, and 377 for temperatures 32°C – 40°C, respectively. The  $\mu_0$  values also increase slightly with temperature from 1.99 to  $2.28 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for 32°C and 40°C, respectively. This increase in  $\mu_0$  is due to a decrease in overall viscosity with increasing temperature.



**Figure 8.7** – A) Elution time ( $t$ ) vs  $1/L$  for T channel BigDye<sup>®</sup> C18 end-alkylated DNA sequencing products separated in 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE buffers on ABI310 at different temperatures. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1. B) Resolution factors calculated for or T channel BigDye<sup>®</sup> C18 end-alkylated DNA sequencing products using Equation 8.2. Dotted line represents resolution limit for DNA sequencing, corresponding to the LOR.

A representative DNA sequencing electropherogram is seen in Figure 8.8 for 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE at 36°C, with the T channel (red curve) labeled with respective DNA lengths. Figure 8.8 demonstrates sequencing of bases 465 – 109 in 30 minutes, increasing the LOR of C<sub>12</sub>E<sub>5</sub> only 7M urea buffers by almost 200 bases. It is important to note that the other nucleotide channels, black curve = G, green curve = A, and blue curve = C, are also present and easily seen. Previous separations using 3M urea only yielded fluorescent signal from the T channel.<sup>[15]</sup>

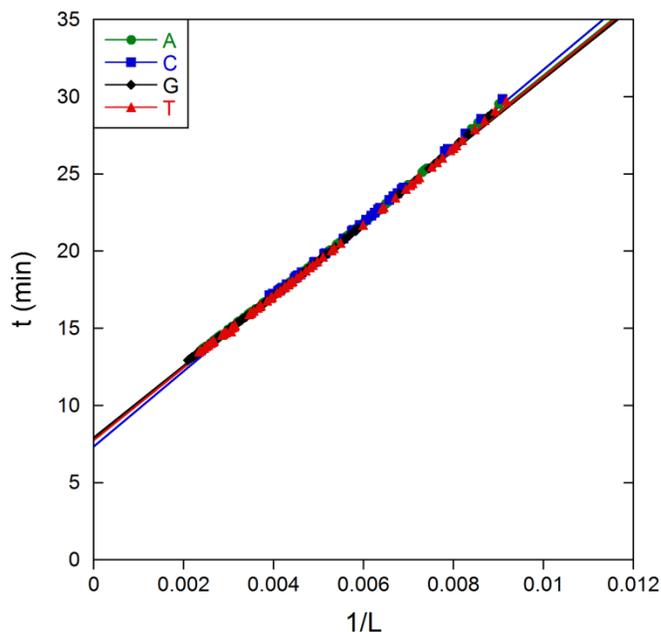


**Figure 8.8** – Four channel DNA sequencing of C18 end-alkylated BigDye® sequencing products in 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea 1xTBE. A = green curve, C = blue curve, G = black curve, T = red curve. Numbers above peaks correspond to T channel DNA lengths. Elution times between different channels have been adjusted using Equation 8.4.

The resolution factors for the T channel sequencing product for 61.5mM C<sub>12</sub>E<sub>5</sub> 3mM C<sub>10</sub>E<sub>5</sub> 7M urea at different temperatures are seen in Figure 8.7B. The dotted horizontal line at  $R = 1.5$  represents the single base resolution limit, where the corresponding length indicates the *LOR*. At 40°C, we have the smallest *LOR* (300) even though  $\alpha$  is the highest (377). This is due to the small amount of background network formation that presents sieving effects on the end-alkylated DNA. As the temperature decreases, the *LOR* increases dramatically up to 475 bases for 36°C. As the temperature decreases further, the *LOR* shrinks due to the smaller micelle size ( $\alpha = 219$ ). Figure 8.7B indicates the very delicate balance of increasing *LOR* with a large micelle, but the necessity of minimal micelle network formation. We need to operate on the edge of network formation because it enables the use of large micelles without the negative broadening from extensive sieving pressures. Since this transition can be monitored using viscosity, optimization of future buffers can be narrowed to a much smaller range of concentrations and temperatures quite easily. This offers a tremendous advantage to experimentalists where simple 5 minute hydrodynamic tests in the capillary can yield a great deal of information vs. the traditional 1 hour electrophoretic separations with analysis.

Taking a close look at the electropherogram in Figure 8.8, there appears to be co-elution of sequencing products between the different nucleotide channels. The different dyes corresponding to each nucleotide are designed to have very similar electrophoretic mobilities in gels; however, due to possible interaction with surfactant or small differences in size, we observe differences in ELFSE behavior.

These differences are presented in Figure 8.9 as a plot of  $t$  vs.  $1/L$  for each nucleotide.



**Figure 8.9** – Elution time ( $t$ ) vs  $1/L$  for all four channels in Figure 8.8. Linear fits demonstrate behavior according to ELFSE theory, Equation 8.1.

Although the curves look to overlap significantly, there are large enough differences in ELFSE parameters to skew each channel's electropherogram. From Figure 8.9, we find an  $\alpha$  of 292, 333, 300, and 301 for the G, C, T, and A channels, respectively. The intercepts also indicate differences in  $\mu_0$ , with 2.09, 2.26, 2.12, and  $2.12 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for the G, C, T, and A channels, respectively.

We can account for these differences using a transformation of the different channels' elution time according to their respective ELFSE parameters. By rearranging Equation 8.1 to solve for DNA length, the effective DNA length at any given elution time for each channel can be solved for.

$$L_N = \alpha_N \left[ \frac{t_N \mu_{0,N} V}{l_d l_t} - 1 \right]^{-1} \quad (8.3)$$

Where  $L_N$  is the DNA length for channel  $N$  corresponding to elution time  $t_N$ ,  $\alpha_N$  is the channel specific  $\alpha$ , and  $\mu_{0,N}$  is the channel specific  $\mu_0$ . Equation 8.3 effectively converts the channel's elution time into a continuous, non-integer theoretical length space using its ELFSE specific parameters. This theoretical length space is then converted back into a time domain using the ELFSE parameters for a single channel, in this case T. We substitute the channel specific length domain from Equation 8.3 into the Equation 8.1 to produce

$$t_c = \frac{l_d l_t}{\mu_{0,T} V} + \frac{\alpha_T \mu_{0,N}}{\alpha_N \mu_{0,T}} \left[ t_N - \frac{l_d l_t}{\mu_{0,N} V} \right] \quad (8.4)$$

Where  $t_c$  is the calibrated time domain scaled to the T channel,  $\alpha_T$  is the T channel  $\alpha$ , and  $\mu_{0,T}$  is the T channel  $\mu_0$ . This transformed time domain now creates a system where equal DNA lengths labeled with different dye terminators results in identical elution times with different  $\alpha$  and  $\mu_0$  parameters.

This transformation has in fact already been applied to Figure 8.8, yet we still see co-elution of DNA lengths for different channels. The reason for this discrepancy is that there is a length dependence of  $\alpha$  and  $\mu_0$  even though the curves appear highly linear in Figure 8.9. For the T channel fits,  $\alpha$  decreases from 361, to 298, to 250 for the lengths 100-200, 200-300, and 300-400, respectively.  $\mu_0$  also decreases in a similar manner, from 2.41 to 2.13 to  $1.95 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . These differences may be caused by the slight network formation of the surfactant

or possible segregation of the DNA from the micelle. We have presented a modified ELFSE-sieving model to account for the contributions of sieving. However, solving for an effective pore size was based on significant deviations from the linear fit. Here,  $R^2$  values are almost 1, making use of our previous model difficult. Additionally, segregation behavior resulted in an inverse square root dependence of elution time with DNA length. Since the traditional ELFSE model seems to fit the data well, further investigation into applying these newer models is required to align the different nucleotide sequencing products. Regardless of the alignment, DNA sequencing close to 500 bases was achieved in a 7M urea buffer by understanding network entanglement and its effect on DNA separation efficiency.

## 8.4 – Conclusion

We have presented the optimization of a 7M urea C<sub>12</sub>E<sub>5</sub> buffer system for increased *LOR* DNA sequencing separations. We were able to design a 7M urea buffer with  $\alpha = 300$  and *LOR* of 465 bases by reducing micelle entanglement through addition of end-capping C<sub>10</sub>E<sub>5</sub> and lowering the separation temperature. We find there is a trade-off between micelle size and entanglement to achieve increased *LOR*. Therefore, future surfactant buffers should be designed such that the micelles are just beginning to entangle, but not forming a fully percolated network in order to achieve optimal separation. This transition can be monitored using simple viscosity measurements, greatly decreasing the experimental strain for future buffer design.

## 8.5 – References

- [1] G. Church, *Science*, 2006, 311, 1544–6.
- [2] J. C. Venter *et al.*, *Science*, 2001, 291, 1304–1351.
- [3] E. S. Lander, *Nature*, 2001, 409, 860–921.
- [4] O. Salas-Solano, E. Carrilho, L. Kotler, A. W. Miller, W. Goetzinger, Z. Sosic, and B. L. Karger, *Anal. Chem.*, 1998, 70, 3996–4003.
- [5] J. M. Rothberg *et al.*, *Nature*, 2011, 475, 348–352.
- [6] D. W. Deamer and M. Akeson, *Trends Biotechnol.*, 2000, 18, 147–151.
- [7] B. Merriman, I. Torrent, and J. M. Rothberg, *Electrophoresis*, 2012, 33, 3397–3417.
- [8] G. F. Schneider and C. Dekker, *Nat. Biotechnol.*, 2012, 30, 326–328.
- [9] H. Ren, A. E. Karger, F. Oaks, S. Menchen, G. W. Slater, and G. Drouin, *Electrophoresis*, 1999, 20, 2501–2509.
- [10] X. Wang, J. C. Albrecht, J. S. Lin, and A. E. Barron, *Biomacromolecules*, 2012, 13, 117–123.
- [11] J. C. Albrecht, J. S. Lin, and A. E. Barron, *Anal. Chem.*, 2011, 83, 509–515.
- [12] P. Mayer, G. W. Slater, and G. Drouin, *Anal. Chem.*, 1994, 66, 1777–1780.
- [13] L. C. McCormick and G. W. Slater, *Electrophoresis*, 2006, 27, 1693–1701.
- [14] K. Grass, C. Holm, and G. W. Slater, *Macromolecules*, 2009, 42, 5352–5359.
- [15] S. B. Istivan and J. W. Schneider, “Rapid Gel-free Separations of DNA via Transient Attachment to Surfactant Micelles,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2012.
- [16] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [17] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.
- [18] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.

- [19] B. B. Rosenblum, L. G. Lee, S. L. Spurgeon, S. H. Khan, S. M. Menchen, C. R. Heiner, and S. M. Chen, *Nucleic Acids Res.*, 1997, 25, 4500–4504.
- [20] G. Briganti, S. Puvvada, and D. Blankschtein, *J. Phys. Chem.*, 1991, 95, 8989–8995.
- [21] C. L. Bianco, C. S. Schneider, M. Santonicola, A. M. Lenhoff, and E. W. Kaler, *Ind. Eng. Chem. Res.*, 2011, 50, 85–96.
- [22] T. Mukaiyama, *Angew. Chemie Int. Ed. English*, 1976, 15, 94–103.
- [23] A. Deratani and T. Maraldo, *React. Polym.*, 1988, 9, 19–28.
- [24] T. Da Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, and B. Cacciari, *Curr. Med. Chem.*, 2005, 12, 71–88.
- [25] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [26] T. Kato, S. Anzai, and T. Seimiya, *J. Phys. Chem.*, 1990, 94, 7255–7259.

# Chapter 9 – Rapid, High Resolution Microchip Micelle ELFSE DNA Separations

## 9.1 – Introduction

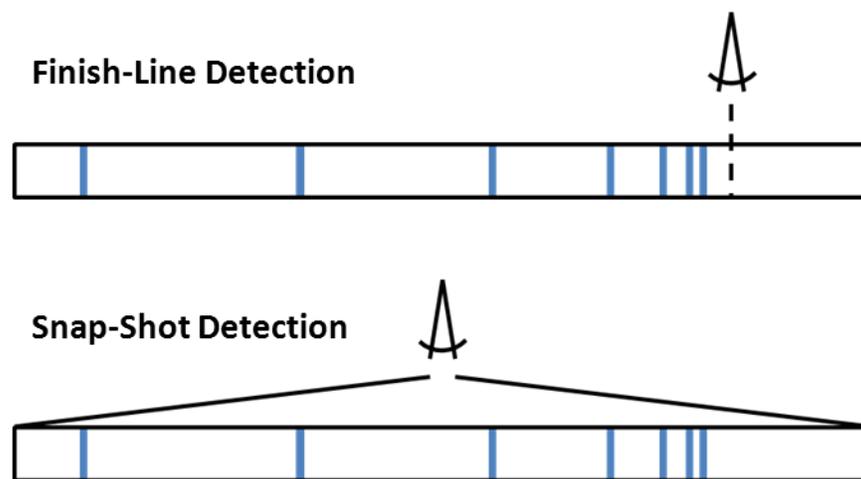
The Human Genome Project could not have been completed in the early 2000s without capillary electrophoresis (CE) instruments.<sup>[1], [2]</sup> The transition from slab gel to CE offered many advantages, such as: i) the ability to apply high electric field strengths (>100 V/cm) in small diameter capillaries without development of Joule heating peak broadening effects. The higher electric fields decreased DNA sequencing run times, sequencing up to 1,000-1,300 bases in 1-2 hours.<sup>[3], [4]</sup> ii) Automation of the CE fluid handling so the polymer sieving matrix could be replaced with fresh polymer after the separation was complete. iii) Simplified data analysis from integrated detection and computer software to process data. iv) Reduced sample and reagent size, on the order of  $\mu\text{Ls}$ , that minimizes both cost and the necessity to collect large sample mass.

The next evolutionary step for DNA electrophoresis and sequencing is the further miniaturization of the CE platform to microfluidic systems. Microchip electrophoresis (ME) was first developed using planar glass chips in 1992.<sup>[5]</sup> However, it is the more recent developments in chip fabrication that have enabled commercialized use of ME. ME offers many advantages compared to the benchtop CE units. ME separation channels are only a few cm long compared to the >20cm capillaries in CE. The shorter channel enables the use of even higher

electric field strengths and short migration distances, completing separations in only seconds to a few minutes. Sample injections in ME systems are defined by the geometry of the chips, with typical cross-T and staggered injections 10-100 $\mu$ ms wide.<sup>[6], [7]</sup> CE injections are on the order of mms due to the end-on introduction of the sample into the inlet of the capillary.<sup>[8], [9]</sup> The smaller injection reduces initial peak broadness, creating a more efficient separation. A microchip can also be highly customized with multiple channels, valves, and reactors to create microscale total analysis systems ( $\mu$ TAS). DNA purification, amplification and labeling, and detection/separation are all performed on a single “lab-on-a-chip” device, reducing sample handling and enabling sample-to-answer diagnostics.<sup>[10]–[12]</sup>

Although the transition from slab gel to CE to ME has progressed over the past two decades, the choice of using DNA sieving matrices has not. Many ME systems still rely on polymer networks to separate DNA. Polymer networks present difficulties when loading the high viscosity solutions into microchannels, and face biased reptation electric field limits that reduce the ability to take full advantage of the short channel lengths in ME for rapid separations.<sup>[6]</sup> End-labeled free-solution electrophoresis (ELFSE) has been developed as an alternative to sieving matrices for rapid DNA separation in free-solution. ELFSE relies on the attachment of uncharged drag-tags to DNA to create length dependent electrophoretic mobilities. Long DNA fragments migrate faster than shorter fragments under ELFSE, opposite of conventional gel electrophoresis. The long fragments are also the most difficult to resolve, and once resolved, the shorter

DNA lengths are also resolved. However, the separation is not complete until the short DNA passes the detector in finish-line detection modes. Thus, ELFSE separations would benefit greatly from the ability to detect all DNA lengths once the longest DNA becomes resolved. The small foot-print of ME systems would enable full chip “snap-shot” detection compared to the conventional finish-line detection mode of CE.<sup>[13], [14]</sup> The difference between finish-line and snap-shot detection is illustrated in Figure 9.1. Snap-shot detection would enable a faster ELFSE readout of the short DNA lengths, dramatically reducing run time even further for ME systems.



**Figure 9.1** – Illustration of finish-line or snap-shot detection modes in microchip ELFSE. All DNA lengths (blue lines) are separated from one another by the time the longest DNA reaches the detector. The shorter lengths have barely migrated down the channel by the time the longest lengths reach the finish-line detector (top cartoon). This creates a situation where most of the run time is spent waiting for the short DNA to pass the detector. In snap-shot mode (bottom cartoon), the entire channel is imaged once the longest DNA lengths are resolved. The separation of all DNA lengths is captured even though the short lengths are still far from the end of the channel.

Implementation of ELFSE into ME devices has not been fully realized. The main difficulty when implementing ELFSE in ME is that the small size of the drag-tag does not yield efficient separation of DNA fragments within the short

separation channel of the microchip.<sup>[7], [15]–[18]</sup> We have previously demonstrated that non-ionic surfactant micelles can be used as large ELFSE drag-tags.<sup>[19]–[23]</sup> The large self-assembled structures should yield large mobility shifts and grant the ability to separate DNA fragments efficiently in the short channel lengths.

Here, we discuss the first ever implementation of micelle ELFSE in a microfluidic format, demonstrating highly resolved and predictive separations of DNA fragments up to 1165 bases long in ~ 1.5 min. A theoretical model is presented that demonstrates our ME separations are almost purely limited by diffusion, indicating operation near the thermodynamic separation limit. With these highly efficient ME separations, we discuss channel length limitations and how an increase from 3cm to 8cm could yield a theoretical DNA sequencing result of 450 bases in 2 minutes under current microchip operating conditions with snap-shot detection.

## **9.2 – Materials and Methods**

### **9.2.1 – Reagents**

Nuclease-free water, Triton X-100 surfactant, acetic acid, and GeneJET Gel Extraction Kits, were purchased from Fisher Scientific (Pittsburgh, PA). BODIPY<sup>®</sup> FL C<sub>16</sub> (C16B), low melting temperature agarose, and POP-6 polymer were purchased from Life Technologies (Grand Island, NY). DMSO, 2,2-Dithiodipyridine (DPDS), Lithium perchlorate, Triphenylphosphine (TPP), 4-(Dimethylamino)pyridine (DMAP), Hexadecyltrimethylammoniumbromide (CTAB), Ethidium bromide, acetonitrile, urea, 10x Tris-Borate-EDTA (TBE), and

triethanolamine (TEA) were purchased from Sigma Aldrich (St. Louis, MO). *N*-dodecylpentaoxyethylene (C<sub>12</sub>E<sub>5</sub>) and *N*-decylpentaoxyethylene surfactant (C<sub>10</sub>E<sub>5</sub>) were purchased from BaChem (Torrance, CA). λDNA and LongAmp<sup>TM</sup> Taq DNA polymerase Kit were purchased from New England Biolabs (Ipswich, MA). 1xTBE buffers (89mM Tris, 89mM Borate, 2mM EDTA) were prepared by diluting the 10x stock with Millipore 18.2mΩ deionized water and filtering through a 0.22μm filter. PCR forward primers and C<sub>6</sub>-amine 5' modified reverse primer were purchased with standard desalting purification from Integrated DNA Technologies (Coralville, IA). Dried oligomers were resuspended in 10mM Tris 0.1mM EDTA buffer and used without further purification.

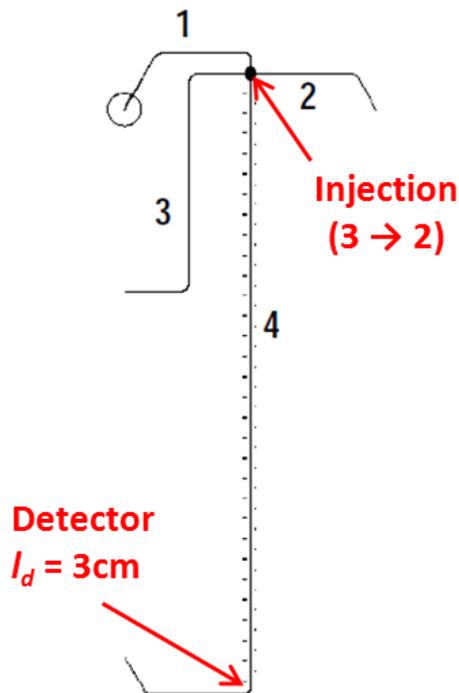
Standard NS114 microfluidic chips were purchased from Perkin-Elmer Microfluidic Chip Foundry (Waltham, MA) provided in plastic caddies. Chip geometry and channel dimensions are shown in Figure 9.2. The plastic caddies enabled the use of ~30μL wells above the channel inlets that could be used for fluid management and electrode placement. The separation channel (channel 1 to 4) yields a 3cm long distance to the detection point prior to the turn. The sample injection channel (channel 3 to 2) yields a perpendicular path for introducing sample into the separation channel.

### **9.2.2 – DNA Alkylation Reaction**

Alkylation of C<sub>6</sub>-amine 5' modified reverse primer with C16B was accomplished using a combination of methods outlined by Mukaiyama and Deratani, and closely followed a scheme first reported by Boutorine *et al.*<sup>[24]–[26]</sup>. An aliquot of 10nmol 5'amine modified reverse primer was precipitated in the

presence of 750nmol CTAB. The suspension was then dried by vacuum centrifugation under low heat for 30 minutes. Meanwhile, a mixture of 100nmol C16B, 5 $\mu$ mol DMAP, 2.5 $\mu$ mol TPP, 2.5 $\mu$ mol DPDS, and 7.5 $\mu$ mole TEA in DMSO was prepared and allowed to activate at room temperature for 20 minutes. Once activated, the reaction mixture was added to the dried DNA suspension, briefly vortexed and bath sonicated to resuspend the DNA pellet, and allowed to react overnight under continuous agitation in a shaker at room temperature in the dark. The solution was then precipitated with 1mL of 2w/v% lithium perchlorate in acetone and pelleted by centrifugation. The supernatant was removed and the pellet was triple rinsed with acetone. The pellet was resuspended in acetone, centrifuged, and supernatant removed again. The pellet was then dried by vacuum centrifugation under low heat for 5 minutes and resuspended in 0.1M triethylammonium acetate pH 7.0 (TEAA) buffer.

Purification of the reaction mixture was performed using a Waters 4.6 mm x 250 mm Symmetry300 C18 HPLC column (Milford, MA) with a 1mL/min flowrate and linear gradient from 0.1M TEAA to 100% acetonitrile over 30 minutes. Any unmodified DNA eluted near 10 minutes while the alkylated DNA eluted near 22 minutes. Product fractions were collected, lyophilized, and resuspended in 10mM Tris 0.1mM EDTA buffer.



		Depth ( $\mu\text{m}$ )	Mask Width ( $\mu\text{m}$ )	Lengths (mm)
channel	1	12	10	7.7
channel	2	12	10	6.2
channel	3	12	10	15.6
channel	4	12	10	36.7

**Figure 9.2** – Schematic of NS114 microfluidic chip from Perkin Elmer with dimensions of each channel.

### 9.2.3 – PCR Amplification of End-alkylated DNA Ladder

Specific lengths of C16B end-alkylated DNA were amplified using LongAmp<sup>TM</sup> Taq DNA polymerase and  $\lambda$ DNA template with the described C16B end-alkylated reverse primer and a length setting forward primer. The end-alkylated DNA ladder included the following lengths of DNA: 86, 133, 189, 214, 265, 375, 421, 488, 551, 603, 647, 707, 787, 808, 847, 909, 986, and 1165. Each PCR reaction contained 1ng of  $\lambda$ DNA template, 500nM each primer, 0.3mM each dNTP, 2.5 units of LongAmp<sup>TM</sup> Taq DNA polymerase, 1x LongAmp<sup>TM</sup> buffer, and enough

water to 25 $\mu$ L. The reaction was denatured for 2min at 95°C and subjected to 35 cycles of denaturation at 95°C for 15s, primer annealing at 50°C for 30 s, and extension at 65°C for 50 seconds per 1000 bases of amplicon length.

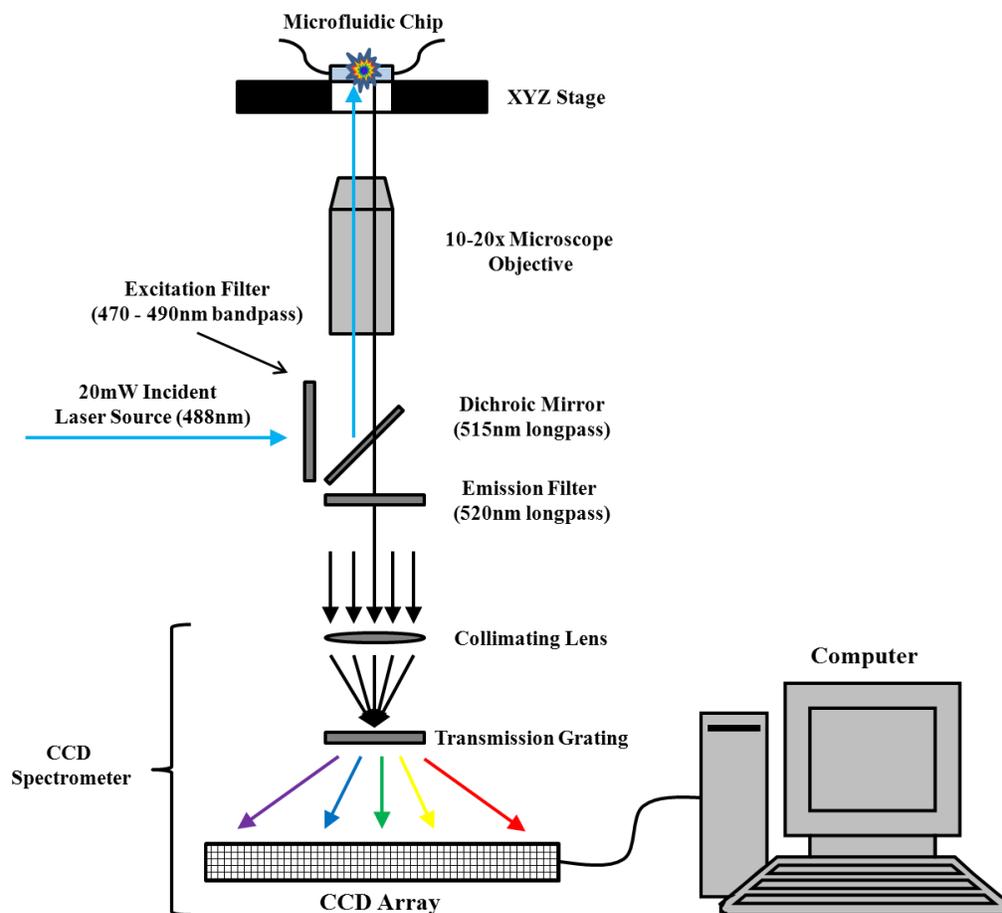
After thermal cycling, the PCR reaction mixture was directly loaded onto 1wt% agarose gels. Slab gel electrophoresis was performed at 4 V/cm for 2 hours. Gels were imaged using ethidium bromide staining and UV excitation on a BioDoc-It Imaging System from Ultra-Violet Products (Cambridge, UK). Correct PCR product length bands were excised, purified using GeneJET Gel Extraction Kit spin columns, and eluted using the provided 10mM Tris 0.1mM EDTA buffer. PCR products were stored in the dark at 4°C for up to one month.

#### **9.2.4 – Microchip Electrophoresis Detection System**

The microchip electrophoresis detection system was a Nikon Eclipse TE2000-U inverted epifluorescence microscope. A schematic of the detection system is seen in Figure 9.3. A 20mW OBIS 488nm LS solid-state laser (Santa Clara, CA) was used as the excitation source for laser induced fluorescence (LIF). The laser was installed into a Nikon D-Eclipse C1-LU3S laser sled control module that provided safety controls for laser light exposure, manual attenuation of laser intensity, and fiber optic light path management to the microscope. The excitation light passed through a FITC filter cube equipped with a 470 – 490 nm bandpass excitation filter, 515nm longpass dichromic mirror, and a 520nm longpass emission filter. The excitation light was focused onto the microchip using a 20x objective ( $NA = 0.45$ ), which created a laser spot size slightly larger than the full channel width, ~50 $\mu$ m spot. The spot was precisely focused onto the channel 3cm from the

sample channel intersection using an MS-2000 XYZ automated stage from Applied Scientific Instrumentation (Eugene, OR) while monitoring the spot on a QuantEM:512SC EMCCD camera from Photometrics (Tucson, AZ) coupled directly to a C-mount on the microscope. The fluorescent excitation light was collected through the same 20x objective and sent to a QE-Pro high-sensitivity CCD spectrometer from Ocean Optics (Dunedin, FL) through a 600 $\mu$ m fiber optic precisely aligned with the excitation light out of another C-mount using an XYZ translation mount with fiber optic attachment from ThorLabs (Newton, NJ). The QE-Pro was equipped with a Hamamatsu S7031-1006 scientific grade, back-thinned, thermoelectric cooled CCD array. Detection using front illuminated CCDs without cooling were not sensitive enough to detect the fluorescent DNA in our system. Full wavelength excitation spectra were captured every 50ms over the course of the separation using the QE-Pro OceanView V1.4 software. Single wavelength electropherograms were reduced from the full wavelength spectra using a custom MATLAB code to import the OceanView text files. Electropherogram peaks were analyzed using OriginPro 9.0 peak fitting software (Northampton, MA).

A broad spectrum EXFO X-Cite 120W metal halide lamp (Quebec, Canada) was used for full field-of-view imaging. The excitation light was passed through the FITC filter cube and focused onto the microchip using a 10x objective ( $NA = 0.45$ ) for full field-of-view imaging. The fluorescent excitation light was collected through the objective and sent to the QuantEM:512SC EMCCD camera. Video capture during pull-back voltage experiment was performed at 50Hz.



**Figure 9.3** – Schematic of microchip electrophoresis laser-induced fluorescence detection system, described in Section 9.2.4.

### 9.2.5 – Microchip Preparation

Before each separation, the NS114 microfluidic chip was scanned to ensure no blockages were in any of the channels. Each buffer was filtered through a 0.22 $\mu$ m syringe filter to remove dust particulates that could block the small microchannels. EOF was suppressed by adding 10v/v% POP-6 in 1xTBE to the wells above channels 1, 3, and 4 with vacuum pulled on channel 2 via tubing connected to an aspirator. After a noticeable decrease in fluid volume of the wells (typically 10-20 minutes), the 10v/v% POP-6 was removed from the wells. Wells above channels 1 and 4 were filled with a surfactant separation buffer. The well

above 3 was filled with either 1xTBE or 3M urea 1xTBE to match the background composition of the separation buffer. Vacuum was applied at 2 until a noticeable decrease in fluid volume in the wells (typically 10-20 minutes). The well above channel 3 was replaced with the indicated DNA sample and ready for separation.

A high voltage power supply (HVS488 6000D, Labsmith, Livermore, CA) was used to apply an indicated voltage between -3kV and 3kV to the 4 wells via platinum electrodes. Labsmith Sequence V1.161 software was used to precisely control the voltage set at each well. Multiple steps could be programmed for simple transition from injection to separation.

### **9.2.6 – Pull-back Voltage Imaging**

Pull-back experiments were performed using the EXFO X-Cite 120W excitation system and video monitoring by the QuantEM:512SC EMCCD camera. A 5 $\mu$ M C16B DNA reverse primer in 10mM Tris 0.1mM EDTA was used as the sample. The DNA was electrokinetically introduced into the sample channel using applied voltages of 0V channel 1, 1000V channel 2, 0V channel 3, and 0V channel 4. After fluorescence was observed in the sample channel, the voltage was switched to 0V channel 1, XV channel 2, XV channel 2, and 2220V channel 4, where X indicates the pull-back voltage. A custom MATLAB script was used to analysis the captured movies and images generated from individual frames.

### **9.2.7 – Microchip Electrophoretic Separation**

Micelle ELFSE separations were performed by injecting the C16B DNA ladder from channel 3 well. The DNA ladder was prepared by diluting 2 $\mu$ L of each C16B DNA length into 30 $\mu$ L of formamide, creating an approximately 50% formamide/50% 10mM Tris 0.1mM EDTA buffer, then heated to 95°C for 2 minutes and snap cooled on ice to denature the DNA. The DNA was electrokinetically introduced into the sample channel for 1 minute using applied voltages of 0V channel 1, 1000V channel 2, 0V channel 3, and 0V channel 4. The voltage was then switched to 0V channel 1, 1000V channel 2, 1000V channel 3, and 2220V channel 4, creating a 500V V/cm electric field down the separation with 1kV pull-back voltages.

## **9.3 – Results and Discussion**

### **9.3.1 – Optimization of Pull Back Voltage**

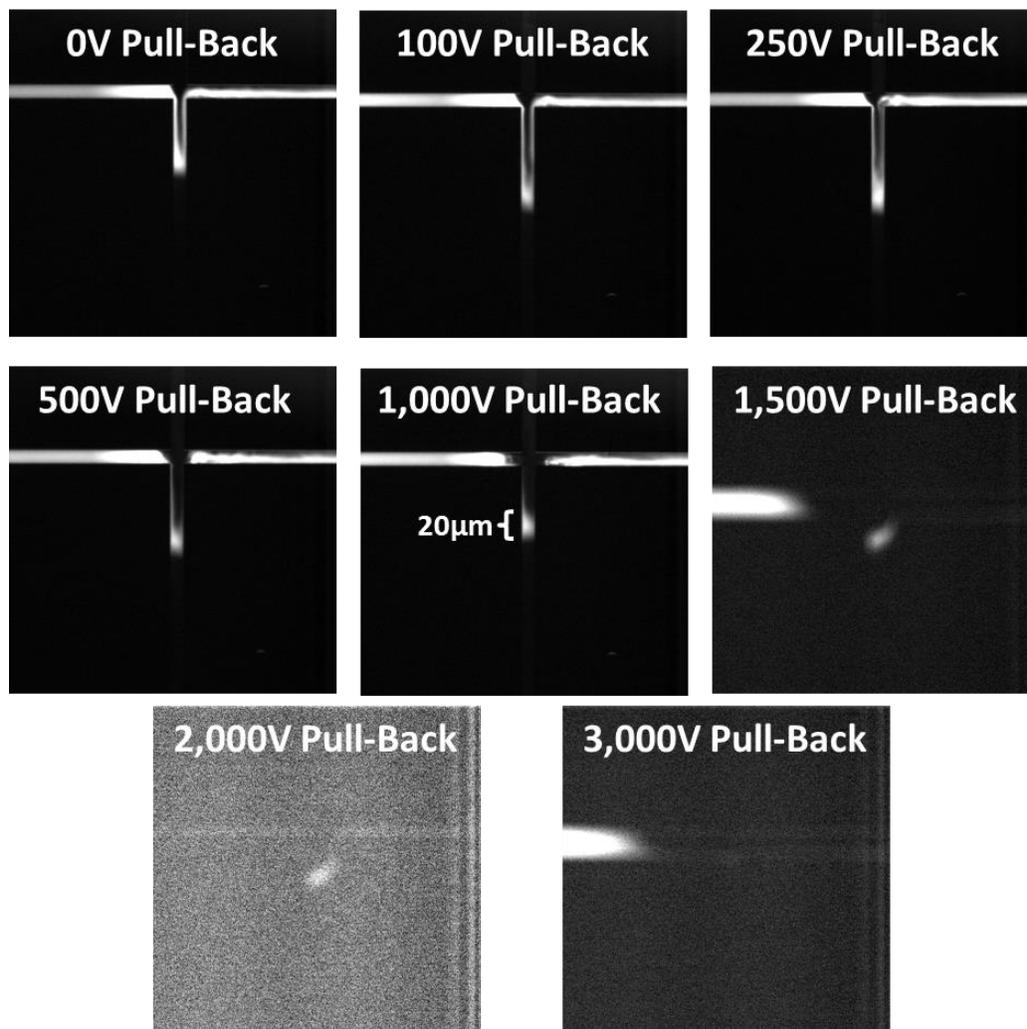
CE injections are performed by submerging the capillary end into the sample solution and applying voltage. The sample is electrokinetically injected into the capillary end, typically on the order of a few mms into the capillary.<sup>[8], [9]</sup> ME injections are performed using a cross channel geometry. The sample is injected in an injection channel perpendicular to the separation channel. When an electric field is applied down the separation channel, the sample at the intersection of the two channels is sent down the separation channel for detection, creating a small injection roughly the width of the channel. However, the injection channel is still filled with sample, which will flow into the separation channel unless a countering potential is applied to keep it from entering the separation channel.

This voltage is known as the pull-back voltage and needs to be optimized to ensure proper injection of the sample without bleed through.

Figure 9.4 shows the effect of increasing the potential of the pull-back voltage along the injection channels when injecting DNA into the separation channel. Each image represents the instantaneous response from changing the injection of DNA (0V channel 1, 1000V channel 2, 0V channel 3, 0V channel 4) to separation (0V channel 1,  $X$ V channel 2,  $X$ V channel 3, 2220V channel 4), where  $X$  represents the pull-back voltage. The effect of sample bleed through is obvious for the lower pull-back voltages of 0, 100, and 250V. This bleed through effect results in an initially larger sample injection plug, as well as constant signal response at the detector from DNA continuously entering the separation channel.

At 500 and 1,000 V, the DNA at the intersection enters the separation channel and DNA sample is pulled back into the injection channel. A tight 20 $\mu$ m injection with a full-width at half-maximum (*FWHM*) was measured for the 1,000V pull-back. Above 1,000V in pull-back voltage, DNA at the intersection is pulled back into the injection channel before entering the separation channel and weak injections are observed. Since the 1,000V pull-back gave us the best injections, we continued with that system as the set of voltages used to apply for separations. Even after long periods of time (>5min), no leakage of DNA from the injection channel was observed. Attempts to apply constant  $E$  in the injections channels led to leakage from the shorter channel 2 into the separation channel. This was most likely caused by the higher potential applied at 3 than 2, creating a net migration

from 2 to 3. For changes in separation voltage, the pull-back voltage should be reconsidered with half the applied  $V$  as a starting point.



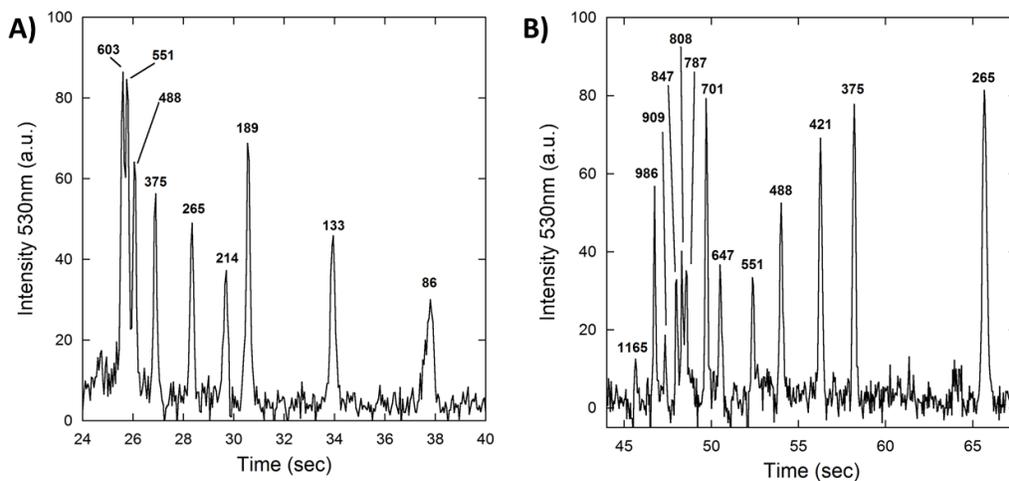
**Figure 9.4** – Effect of pull-back voltage applied to injection channels to prevent leakage of into separation channel. Images are taken at the channel intersection with the separation channel down the page, right after the separation voltage is applied. Pull-back voltages represent potential applied to channel 2 and 3 wells. Channel 4 was grounded ( $V = 0$ ) and channel 3 has a potential of 2,220V.

### 9.3.2 – Microchip Micelle ELFSE Separations

We tested two buffers for microchip micelle ELFSE separations: the commonly used Triton X-100 globular micelle buffer and a  $C_{12}E_5/C_{10}E_5$  3M urea binary

surfactant buffer used for DNA sequencing applications by our group. Representative electropherograms are shown in Figure 9.5. The microchip separations exhibit ELFSE like behavior, with the longest lengths of DNA eluting first and the most difficult to resolve. The 24mM TX-100 buffer yields separation of lengths 603 to 86 bases long in 38 seconds. Separation of longer lengths was not observed on the chip due to the small size of the TX micelle ( $\alpha \sim 50$ ,  $R_h = 2.3\text{nm}$ ).

For the 48mM  $C_{12}E_5$ , 8mM  $C_{10}E_5$ , 3M urea buffer, separation of DNA fragments 1165 to 375 bases long is achieved in 1 minute.  $C_iE_j$  surfactants are known to form large wormlike micelles in aqueous solutions. Due to its larger size than the globular TX-100 micelle, we see resolution of the longer lengths of DNA.



**Figure 9.5** – Representative electropherograms for microchip micelle ELFSE separations of C16B end-alkylated DNA ladders. A) 24mM TX-100 in 1xTBE. B) 48mM  $C_{12}E_5$ , 8mM  $C_{10}E_5$ , 3M urea 1xTBE. Run conditions as described in Section 9.2.7.

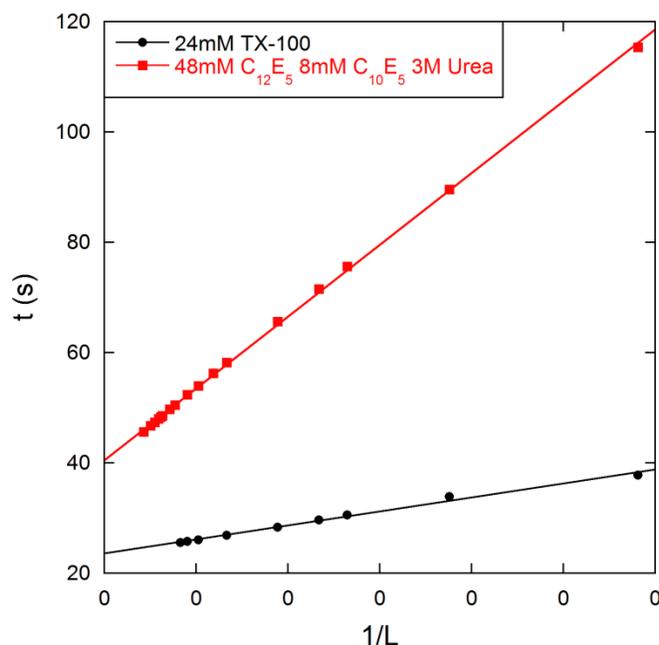
The high quality separations in Figure 9.5 are somewhat unexpected due to the short channel length of the microchip (3cm). Drag-tag polydispersity can lead to significant peak broadness in ELFSE separations. Although surfactant micelles are polydisperse in size, dynamic surfactant exchange during the course of the separation creates an averaging effect that significantly reduces polydispersity.<sup>[22]</sup> With such a short separation time and length, one would expect deleterious peak broadening effects to become problematic. Surprisingly, the separations appear highly resolved for the short channel lengths. This may be in part due to the increased injection efficiency of the microchip compared to the capillary. Although polydispersity broadening may increase as the separation channel is shortened, the decreased initial injection width (from ~1mm to 20 $\mu$ m) appears to maintain small peak widths.

Figure 9.6 shows a plot of elution time ( $t$ ) vs.  $l/L$  for both micelle buffer separations in Figure 9.5. The microchip data exhibits highly linear behavior, demonstrating excellent agreement with standard ELFSE theory in Equation 9.1.

$$t = \frac{l_d l_t}{\mu_0 V} \left( 1 + \frac{\alpha}{L} \right) \quad (9.1)$$

Where  $l_d$  is the channel length to the detector (3cm),  $l_t$  is the total channel length (4.4cm),  $V$  is the applied potential,  $\mu_0$  is DNA's length invariant free-solution electrophoretic mobility,  $L$  is the DNA length, and  $\alpha$  is the additional drag of the non-ionic micelle in terms of hydrodynamically equivalent DNA bases. From Figure 9.6,  $\alpha$  can be calculated from the slope/intercept and  $\mu_0$  from 1/intercept of the linear trend fit to the data, respectively.

We find  $\alpha = 53.6$  and  $\mu_0 = 2.52 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for the 24mM TX-100 ME separations. An  $\alpha$  of 55 was reported for ssDNA using benchtop CE systems, showing agreement in buffer performance between CE and ME.<sup>[22]</sup> The  $\mu_0$  also agrees well with literature values  $2.7 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for ssDNA in 1xTBE buffers.<sup>[27]</sup>



**Figure 9.6** – Elution time ( $t$ ) vs.  $1/L$  from Figure 9.5 C16B end-alkylated DNA ladders separated by ME in 24mM TX-100 (black circles) or 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M urea (red squares). Solid lines are linear fits according to Equation 9.1.

The higher slope for the 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M urea buffer indicates a higher  $\alpha$  parameter,  $\alpha = 161$ . The larger size of the C<sub>i</sub>E<sub>j</sub> wormlike micelles leads to this larger  $\alpha$ , allowing resolution of the longer DNA lengths. The higher intercept indicates a lower  $\mu_0$ ,  $\mu_0 = 1.47 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . The decrease in  $\mu_0$  may be due to the increased viscosity of the urea containing buffer vs. 1xTBE. Typical CE separations with this buffer are performed at 33°C, with ELFSE parameters of

$\alpha = 430$  and  $\mu_0 = 2.2 \times 10^{-4} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ . The higher temperature increases the size of the wormlike micelles, leading to a higher  $\alpha$  and reducing buffer viscosity increasing  $\mu_0$ . Overall, we observe excellent agreement between ELFSE theory and microchip separations of C16B end-alkylated DNA.

### 9.3.3 – ME Resolution Factors and Theory

The effectiveness of a separation can be judged by the resolution factor ( $R$ ), the ratio of a peak's full width at half max ( $FWHM$ ) to the spacing between the adjacent DNA length elution peak ( $dt/dL$ ). A smaller value for  $R$  means a more efficient separation, where an  $R \leq 1.5$  is the limit for single base resolution for DNA sequencing.

$$R = \frac{FWHM}{dt/dL} \quad (9.2)$$

To predict separation performance, we can estimate the shape of the elution peak as Gaussians where the  $FWHM$  can be described in terms of the temporal variance ( $\sigma_t^2$ ) of the peak

$$FWHM = \sqrt{8\ln(2)\sigma_t^2} \quad (9.3)$$

The total variance can be expressed as a combination of individual sources of variance, including finite injection/detection width ( $A$ ), diffusion ( $D$ ), wall adsorption ( $W$ ), and Joule heating ( $J$ ).<sup>[28]–[30]</sup> Since each of these terms is an independent random walk of variance, they can be summed as a contribution of uncorrelated sources of spatial variance ( $\sigma_x^2$ ).<sup>[9]</sup> Because detection takes place at a fixed point down the separation channel, what is actually measured is the

temporal variance of the peak. A simple transformation from  $\sigma_x^2$  to  $\sigma_t^2$  can be made via the velocity  $v$  of the peak,  $\sigma_t^2 = \sigma_x^2/v^2$ . The full temporal variance model becomes

$$\sigma_t^2 = \frac{\sigma_x^2 t^2}{l_d^2} = \frac{At^2}{l_d^2} + \frac{2Dt^3}{l_d^2} + Wt + \frac{JE^4 t}{D} \quad (9.4)$$

Where  $A$ ,  $W$ , and  $J$  represent constants for their variance terms, and  $D$  is the DNA – micelle complex diffusion coefficient. The microchip geometry sets the injection width of DNA, which we measured as  $FWHM = 20\mu\text{m}$ . Due to the small injection width, we can ignore the small amount of variance it contributes in microfluidic systems. Additionally, the Joule heating constant  $J$  has a strong dependence on channel radius ( $J \propto r^6$ ).<sup>[30]</sup> The small channel radius in the microchannel allows us to ignore contributions from Joule heating as well.

The diffusion coefficient for ssDNA decreases as  $D = D_0 L^{-0.5}$ , where  $D_0$  is the diffusion coefficient of a single base pair reported as  $3.2 \times 10^{-6} \text{ cm}^2/\text{s}$ .<sup>[28], [31]</sup> We account for impact of the attached micelle on the ssDNA diffusion coefficient by treating the micelle-DNA conjugate as a ssDNA with length  $(L+\alpha)$ .<sup>[14], [32]</sup>

$$D = D_0(L + \alpha)^{-0.5} \quad (9.5)$$

This leaves the wall adsorption term  $W$  as the only fitting parameter in our system. This term represents the effective rate of DNA interaction with the wall. Smaller times represent a shorter interaction and more effective separation.

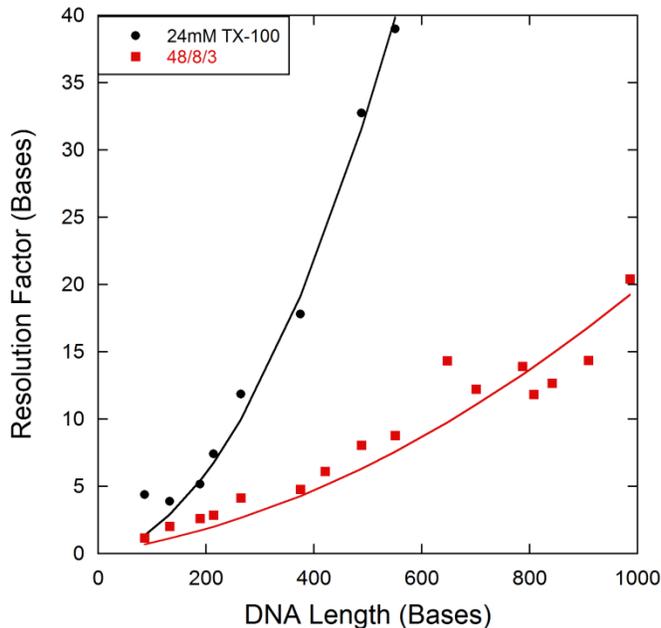
The instantaneous  $dt/dL$  can be calculated for all DNA lengths by taking the derivative of Equation 9.1. Here, we have substituted an expression for  $t$  for the group of separation parameters ( $l_d, l_t, \mu_0$  and  $V$ ) in Equation 9.1.

$$\frac{dt}{dL} = \frac{l_d l_t \alpha}{\mu_0 V L^2} = \frac{t}{\left(1 + \frac{\alpha}{L}\right)} \frac{\alpha}{L^2} \quad (9.6)$$

Substituting Equation 9.3 – 9.6 into Equation 9.2, the theoretical expression for  $R$  with diffusion and wall adsorption broadening terms becomes

$$R = \frac{L^2 \left(1 + \frac{\alpha}{L}\right)}{\alpha} \sqrt{8 \ln(2) \left[ \frac{2D_0 t}{l_d^2 \sqrt{L + \alpha}} + \frac{W}{t} \right]} \quad (9.7)$$

Figure 9.7 shows the resolution factors for the ME separations in Figure 9.5. The solid lines in Figure 9.7 are fits of Equation 9.7 to the resolution data, using  $W$  as the lone fitting parameter. We find  $W$  terms of 174 $\mu$ s and 18 $\mu$ s for the 24mM TX-100 and 48mM C<sub>12</sub>E<sub>5</sub> 8mM C<sub>10</sub>E<sub>5</sub> 3M urea buffers, respectively. The presence of urea is thought to break up hydrogen bond formation between water molecules, stabilizing surfactant monomers in solution.<sup>[33]–[35]</sup> Increasing urea in the separation buffer may stabilize the C16B end-alkyl modification in the aqueous running buffer as well. This would reduce interaction with the capillary wall, creating more efficient separations.<sup>[22]</sup>



**Figure 9.7** – Resolution factors for C16B end-alkylated DNA ladders separated by ME in 24mM TX-100 (black circles) or 48mM C<sub>12</sub>E<sub>5</sub>, 8mM C<sub>10</sub>E<sub>5</sub>, 3M urea (red squares). Solid lines represent fits to theoretical resolution model for ELFSE in Equation 9.7.

The major conclusion from these theoretical models is that our microchip separations are almost completely limited by diffusive peak broadening, corresponding to >98% of the total variance for the 48mM C<sub>12</sub>E<sub>5</sub> 8mM C<sub>10</sub>E<sub>5</sub> 3M urea buffer. Since thermal motion of the DNA cannot be prevented, our ME separations appear to be operating at the thermodynamic separation limit. This result means that faster separations (from increased  $E$  or reduced  $l_d$ ) will only aid in improving resolution when peaks are sufficiently separated. It also explains why our separations are so efficient in the small channels of the microchip.

We have previously demonstrated CE DNA sequencing with 24mM TX-100 1xTBE up to 110 bases in 4 minutes.<sup>[22]</sup> Our DNA lengths are much larger than 100 bases, but our resolution model predicts single base resolution in the 3cm

microchip up to 94 bases with a  $W$  of  $174\mu\text{s}$ , in good agreement with the CE results. The smaller injection widths in the ME may enable the use of shorter channel lengths to achieve very similar separations as the CE.

### **9.3.4 – Theoretical Estimates for Microchip Micelle ELFSE DNA Sequencing**

One of the main ME applications using micelle ELFSE is to have dramatically reduced run times for DNA sequencing applications. Typical CE micelle ELFSE DNA sequencing takes roughly 45 minutes to achieve a length of read ( $LOR$ ) of 420 bases.<sup>[22]</sup> The bulk of the time is not spent separating DNA, but rather waiting for smaller fragments of DNA to elute past the finish-line detection method 20-30cm down the capillary. Since ELFSE operates under the unique condition that the longest DNA elutes first and is the most difficult to resolve, once the first resolved base ( $LOR$ ) crosses the finish-line, all other smaller lengths of DNA are resolved.

Microchips offer the advantage of small front prints that can be analyzed via a snap-shot detection mode, where the entire separation channel is imaged by scanning the detection point back along the channel from the finish-line point.<sup>[13]</sup><sup>[14]</sup> We can estimate the run-time and channel length required to achieve baseline resolution ( $R = 1.5$ ) for a specific  $LOR$  under microchip run conditions with snap-shot detection. Figure 9.8 shows a plot of the theoretical channel length vs. run time to achieve  $R = 1.5$  for a  $LOR$  of 450 bases using an  $\alpha = 500$  micelle buffer, the typical value achieved for the 48mM  $C_{12}E_5$  8mM  $C_{10}E_5$  3M urea buffer heated

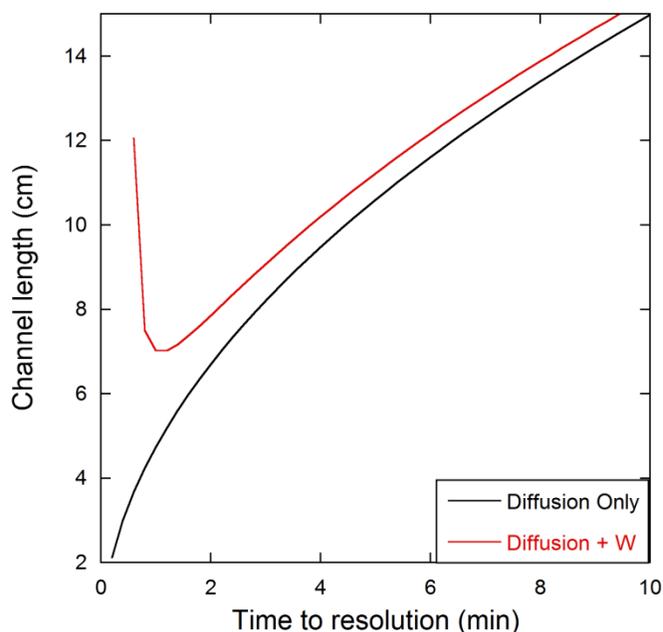
to 33°C. The theoretical channel length was calculated by rearranging Equation 9.7 to solve for the  $l_d$  at  $R = 1.5$ .

$$l_d = \sqrt{\frac{2D_0t}{\sqrt{L+\alpha} \left( \frac{R^2\alpha^2}{8\ln(2)L^4 \left(1+\frac{\alpha}{L}\right)^2} - \frac{W}{t} \right)}} \quad (9.8)$$

We include a completely diffusion limited case ( $W = 0$ ) and a  $W = 18\mu\text{s}$  case for the fitted parameter from our 48mM C<sub>12</sub>E<sub>5</sub> 8mM C<sub>10</sub>E<sub>5</sub> 3M urea buffer. For the purely diffusive broadening case, our theory predicts that the 3cm microchip could meet the 450 base sequencing limits in as little as 0.4 min with snap-shot detection. This is the absolute minimum for micelle ELFSE DNA sequencing on a microchip. When we include the small  $W$  term, we see a dramatic change in behavior of this plot. At low  $t$ , the required channel length sharply decreases from an initially high value. This decrease is caused by the  $W$  term contributing a small amount of peak broadening that cannot be resolved in short times without a long separation channel length. A 12cm channel is required for a run time of 0.6 minutes on the microchip. This would correspond to an electric field strength of 3.25kV/cm, much higher than possible with currently available power supplies, and not to mention the incumbent effects of Joule heating at high  $E$ .

As the run time increases, we observe a minimum in the required channel length where a balance is struck between  $W$  peak broadening and diffusion. This minimum indicates that with wall adsorption, our ME system can sequence 450 bases in 1 minute using a microchannel 7cm long. This corresponds to a much

more reasonable electric field strength of 1kV/cm, just above the limits of our 6kV power supply but below Joule heating extremes. For our current 6kV power supply, we could perform a 2 minute separation with snap-shot detection on an 8cm microfluidic channel ( $E = 750$  V/cm) and achieve our DNA sequencing *LOR* of 450 bases. Our current microfluidic chips are limited to 3cm separation channels, but we are investigating longer channel lengths to achieve these extremely desirable sequencing results.



**Figure 9.8** – Theoretical estimates of microchip channel length and run-time to detect 450 bases with single base resolution and an  $\alpha = 500$  micelle buffer according to Equation 9.8. Diffusion only model (black line) and diffusion + wall adsorption with  $W = 18\mu\text{s}$  (red line).

## 9.4 – Conclusion

We have demonstrated micelle ELFSE in a microchip format for the first time. We find that ME separation of C16B end-alkylated DNA follows ELFSE theory and matches performance on CE benchtop systems. Using arguments for peak

broadening terms during electrophoresis, we find that our ME separations are almost entirely dominated by the thermodynamic limit of diffusive band broadening. Our theoretical model predicts that sequencing with an *LOR* of 450 bases can be achieved in 2 minutes on an 8cm microchip, given an  $\alpha = 500$  micelle buffer and snap-shot detection, where the entire separation channel is imaged rather than conventional finish-line detection. These run times and *LORs* are unparalleled by current microchip DNA sequencing and offer tremendous potential for future DNA sequencing applications such as personalized medicine whole genome sequencing.

## 9.5 – References

- [1] J. C. Venter *et al.*, *Science*, 2001, 291, 1304–1351.
- [2] E. S. Lander, *Nature*, 2011, 470, 187–197.
- [3] H. Zhou, A. W. Miller, Z. Susic, B. Buchholz, A. E. Barron, L. Kotler, and B. L. Karger, *Anal. Chem.*, 2000, 72, 1045–1052.
- [4] O. Salas-Solano, E. Carrilho, L. Kotler, A. W. Miller, W. Goetzinger, Z. Susic, and B. L. Karger, *Anal. Chem.*, 1998, 70, 3996–4003.
- [5] D. J. Harrison, A. Manz, Z. Fan, H. Luedi, and H. M. Widmer, *Anal. Chem.*, 1992, 64, 1926–1932.
- [6] C. P. Fredlake, D. G. Hert, C. W. Kan, T. N. Chiesl, B. E. Root, R. E. Forster, and A. E. Barron, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, 105, 476–481.
- [7] J. C. Albrecht, M. B. Kerby, T. P. Niedringhaus, J. S. Lin, X. Wang, and A. E. Barron, *Electrophoresis*, 2011, 32, 1201–1208.
- [8] J. P. Quirino and S. Terabe, *Science*, 1998, 282, 465–468.
- [9] P. D. Grossman, *Capillary Electrophoresis: Theory and Practice*, 1st ed. Academic Press, 1992.
- [10] P. Liu, X. Li, S. A Greenspoon, J. R. Scherer, and R. A Mathies, *Lab Chip*, 2011, 11, 1041–1048.

- [11] E. T. Lagally, P. C. Simpson, and R. A. Mathies, *Sensors Actuators B Chem.*, 2000, 63, 138–146.
- [12] N. Beyor, L. Yi, T. S. Seo, and R. A. Mathies, *Anal. Chem.*, 2009, 81, 3523–3528.
- [13] M. A. Fahrenkopf, T. Mukherjee, B. E. Ydstie, and J. W. Schneider, *Electrophoresis*, 2014, 35, 3408–3414.
- [14] M. A. Fahrenkopf, B. E. Ydstie, T. Mukherjee, and J. W. Schneider, *Comput. Chem. Eng.*, 2014, 64, 1–8.
- [15] J. C. Albrecht, A. Kotani, J. S. Lin, S. A. Soper, and A. E. Barron, *Electrophoresis*, 2013, 34, 590–597.
- [16] W. Cao, L. Chen, Y. Fu, Z. Tan, and B. Qu, *J. Sep. Sci.*, 2011, 34, 939–946.
- [17] H. Song, Y. Wang, C. Garson, and K. Pant, *Microfluid. Nanofluidics*, 2014, 17, 693–699.
- [18] H. Song, Y. Wang, C. Garson, and K. Pant, *Anal. Methods*, 2015, 7, 1273–1279.
- [19] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [20] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.
- [21] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [22] S. B. Istivan and J. W. Schneider, “Rapid Gel-free Separations of DNA via Transient Attachment to Surfactant Micelles,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2012.
- [23] A. Jones, “Rapid Separations of Kilobase-Sized DNA using Micelle Size-Sampling,” Ph.D. dissertation, Dept. Chem. Eng., Carnegie Mellon University, Pittsburgh, PA, 2013.
- [24] T. Mukaiyama, *Angew. Chemie Int. Ed. English*, 1976, 15, 94–103.
- [25] A. Deratani and T. Maraldo, *React. Polym.*, 1988, 9, 19–28.
- [26] T. Da Ros, G. Spalluto, M. Prato, T. Saison-Behmoaras, A. Boutorine, and B. Cacciari, *Curr. Med. Chem.*, 2005, 12, 71–88.
- [27] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [28] H. Ren, A. E. Karger, F. Oaks, S. Menchen, G. W. Slater, and G. Drouin, *Electrophoresis*, 1999, 20, 2501–2509.

- [29] P. Mayer, G. W. Slater, and G. Drouin, *Anal. Chem.*, 1994, 66, 1777–1780.
- [30] J. Knox and I. Grant, *Chromatographia*, 1987, 24, 135–143.
- [31] B. Tinland, A. Pluen, J. Sturm, and G. Weill, *Macromolecules*, 1997, 9297, 5763–5765.
- [32] L. C. McCormick, G. W. Slater, A. E. Karger, W. N. Vreeland, A. E. Barron, C. Desruisseaux, and G. Drouin, *J. Chromatogr. A*, 2001, 924, 43–52.
- [33] H. Patel, G. Raval, M. Nazari, and H. Heerklotz, *Biophys. Chem.*, 2010, 150, 119–128.
- [34] S. B. Sulthana, S. G. T. Bhat, and A. K. Rakshit, *Colloids Surfaces A Physicochem. Eng. Asp.*, 1996, 111, 57–65.
- [35] G. Briganti, S. Puvvada, and D. Blankschtein, *J. Phys. Chem.*, 1991, 95, 8989–8995.



# Chapter 10 – Conclusions, Future work, and Outlook

## 10.1 – Conclusions

The work presented in this thesis has outlined a new method for direct, high-throughput miRNA detection by sandwich hybridization and micelle end-labeled free-solution electrophoresis (ELFSE) detection. In Chapter 2, we were able to demonstrate stable hybridization of two probes to the short, ~22 nucleotide miRNA targets using a high affinity *n*-alkylated polyethylene glycol  $\gamma$ -carbon modified peptide nucleic acid amphiphile ( $\gamma$ PNAA).<sup>[1]</sup> The  $\gamma$ PNAA's exceptionally high affinity enables stable hybridization of a second, fluorescently labeled DNA-based probe to the remaining bases of the miRNA. Upon hybridization of both probes, an electrophoretic mobility shift is measured via interaction of the *n*-alkane modification on the  $\gamma$ PNA with capillary electrophoresis (CE) running buffer containing non-ionic surfactant micelles.

In Chapter 3, we were able to demonstrate rapid, multiplexed detection of 6 *let-7* miRNAs in a 4 minute separation by the CE sandwich hybridization assay. Importantly, hybridization of each miRNA target to its respective  $\gamma$ PNAA and DNA probes was accomplished in a single solution via a simple heating/cooling protocol. The kinetics of solution based hybridization are much higher than traditional surface based multiplexed detection platforms, such as microarrays, greatly minimizing the total assay time.<sup>[2]</sup> Our theoretical model predicts that our

current separation can multiplex detection of 26 different miRNAs in the same 4 minute separation by simply increasing the length of the DNA probe hybridized to the miRNA target. If we consider the 96 capillary array CE systems present in many clinical labs, our method is theoretically capable of profiling almost every known miRNA (2,496 out of 2,588) in a single 4 minute separation. This high-throughput profiling offers a 200 fold improvement compared to the gold-standard of multiplexed miRNA detection by solid phase microarray technologies.<sup>[3]</sup>

We have also demonstrated excellent single base mismatch discrimination between the closely related *let-7* family of miRNA. These mismatches express pseudo-binding G-T wobble base-pair mismatches that result in minimal drops in free energy upon hybridization, making them extremely difficult to discriminate.<sup>[4]</sup> We find that through competitive binding between perfectly match probes and single base mismatch probes in solution, are system can discriminate these G-T single base mismatches at 25°C. Separations at lower temperatures preserve high degrees of target hybridization and increased signal response of bound miRNA complexes.

We also found miRNA detection limits of 100pM by staining the DNA probe with PicoGreen intercalating dye. We term these dye laden DNA duplex hybridization probes as DNA nanotags. However, the natural abundance of miRNA expression levels may be as low as tens of copies per cell.<sup>[5]</sup> Minimally invasive biopsies, such as fine-needle aspiration, typically collect ~1,000 cells for analysis.<sup>[6]</sup> Therefore, the theoretically ideal detection limit for direct (enzyme

amplification/ labeling free) detection of low abundance miRNA would be on the order of  $10^4$  targets, or fM concentrations for the  $\mu\text{L}$  sample volumes used in the CE system.

To increase our detection limits closer to the theoretically ideal fM range, we investigated sample stacking isotachophoretic injection schemes in Chapter 4 that yielded 20-50 fold improvements in detection sensitivity without losses in resolution.<sup>[7]-[11]</sup> To further increase detection sensitivity, we may simply increase the length of our DNA nanotags to thousands of bases long. The dye associates with the DNA duplex at a ratio of one PicoGreen molecule per four base pairs.<sup>[12]</sup> We have shown that the fluorescent signal of each complex linearly correlates to the length of DNA nanotag hybridized to each miRNA target. As the DNA nanotag increases in length, it should incorporate a greater amount of PicoGreen dye and yield a greater fluorescent signal.

However, separation of kilobase (kB) DNA by ELFSE is rather difficult. The separation of bound and unbound hybridization probes is achieved using non-ionic surfactant micelles as hydrodynamic drag-tags.<sup>[1], [13], [14]</sup> As the DNA nanotag length increases, the additional friction from the attached micelle becomes negligible, and the sandwich miRNA complex returns to the inseparable free-solution DNA mobility limit.<sup>[15]</sup>

We have investigated the use of *n*-alkyl polyoxyethylene ether surfactants ( $\text{C}_i\text{E}_j$ s) that form very large, wormlike micelle structures as large ELFSE drag-tags. As the concentration of  $\text{C}_i\text{E}_j$  surfactant increases past the overlap

concentration ( $c^*$ ), the micelles will begin to interact with one another leading to an entangled network.<sup>[16]–[19]</sup> These micelle networks mimic polymeric sieving matrices that have been used to electrophoretically separate DNA in a manner similar to gel networks.<sup>[19], [20]</sup> Wormlike micelle network formulation should present an upper limit for their application in micelle ELFSE separations, as the presence of sieving matrices has been shown to be detrimental for ELFSE separation due to their countering separation mechanisms.<sup>[21]</sup>

In Chapter 5, we found that separation of unmodified DNA, DNA without an  $n$ -alkane modification for direct interaction with non-ionic surfactant micelles, in  $C_{12}E_5$  buffers above  $c^*$  followed an Ogston sieving like separation mechanism.<sup>[22]</sup><sup>[23]</sup> This theory states that the DNA is separated by length, based on the fractional volume available for each DNA length to access free volume within the pore spaces of the surfactant network.

In Chapter 6, we were able to separate end-alkylated DNA, DNA with the  $n$ -alkane modification for separation by ELFSE, in these same  $C_{12}E_5$  buffers above  $c^*$ . We found that the end-alkylated DNA separated in an ELFSE-like manner with the longest DNA eluting first. However, we observed deviations from ELFSE theory caused by the background micelle network, which reduced separation efficiency of the longer kB DNA fragments. The background network resulted in peak compression of the longer DNA lengths. We were able to develop a modified ELFSE-sieving model that accounted for the reduced free volume for electrophoresis based on GPC theory for longer DNAs in the micelle networks. This modified model accurately reflected the reduced peak spacing of end-

alkylated DNA in the micelle networks. Our conclusion from this work was that although the  $C_{12}E_5$  micelles proved to be large ELFSE drag-tags, micelle network formation significantly reduced our ability to separate the desired kB DNA fragments.

In Chapters 7 and 8, we explored methods to reduce micelle network formation by adding end-capping  $C_iE_j$  surfactants, changing separation temperature, and adding urea to our buffers. We were able to develop a ternary  $C_{16}E_6$ ,  $C_{12}E_5$ ,  $C_{10}E_5$  buffer that rapidly separated 1 – 10kB end-alkylated DNA in 3 minutes. Interestingly, we observed deviations from ELFSE theory once again caused by micelle network formation. However, these deviations were best characterized by considering the DNA-micelle conjugate as two hydrodynamically segregated units, rather than the single random coil assumed for traditional ELFSE theory. The segregated behavior results in a reduced electrophoretic mobility relationship with DNA length, yielding slower moving fragments that are more resolved by ELFSE. We believe segregation is caused by the large size of the DNA-micelle conjugate that cannot traverse the micelle network pore spaces as a single hydrodynamic unit. The two most segregate to migrate through pore spaces, leading to these efficient separations of long DNA. The pores are still large enough to allow DNA to migrate through without negative sieving effects.

In Chapter 8, we applied our knowledge of micelle network formation to  $C_{12}E_5$  surfactant in 7M urea DNA sequencing buffers for increased length of read (*LOR*) separations. We were able to formulate buffers, based on reducing entanglement by temperature and end-capping surfactants, to increase the sequencing *LOR* from

300 to 500 bases. Since  $c^*$  can be measured through simple viscosity measurements, future buffer optimization based on network formation will not require intensive DNA separations. One would simply measure the viscosity of several buffers to judge network formation to dramatically narrow the window for proper DNA sequencing buffers.

Finally, in Chapter 9 we expanded the domain of micelle ELFSE DNA separations from the CE benchtop units to small microchips. Although the path length for microchips is quite small (3cm), we were able to demonstrate highly resolved and predictive separations of DNA fragments up to 1165 bases long in ~ 1.5 min on the chips. We find that these separations are almost purely limited by diffusional peak broadening, mainly due to the small injections into the microchip separation channels. We offered a model for theoretical DNA sequencing by whole chip imaging, known as “snap-shot” detection, where access to an 8cm separation channel would yield DNA sequencing of 450 bases in 2 minutes.

## **10.2 – Future Work**

The work presented in this thesis has laid down a large portion of the groundwork for high-throughput, direct miRNA detection with fM detection limits. We have developed a sandwich hybridization assay for selective miRNA detection by micelle ELFSE. We have developed new surfactant buffers capable of rapid separation of kB DNA for sensitive detection by DNA staining with PicoGreen dye. We have investigated on-column isotachopheresis sample stacking injections to further increase detection sensitivity. However, we have not

brought each of these components together to realize our goal of direct miRNA detection. This section will set out to define some specific aims for future experiments to achieve this goal.

### **10.2.1 – Longer Nanotags for fM Detection Sensitivity**

We have demonstrated a miRNA detection limit of 100pM using oligomeric DNA nanotags. Oligomers are limited to lengths approximately 100 bases long by phosphoramidite DNA synthesis. In order to achieve truly direct, amplification free miRNA detection, fM detection limits are required. A simple way to increase detection sensitivity would be to extend the length of DNA nanotag hybridized to the miRNA targets. The extended DNA duplex would incorporate a greater amount of fluorescent dye, increasing detection sensitivity.

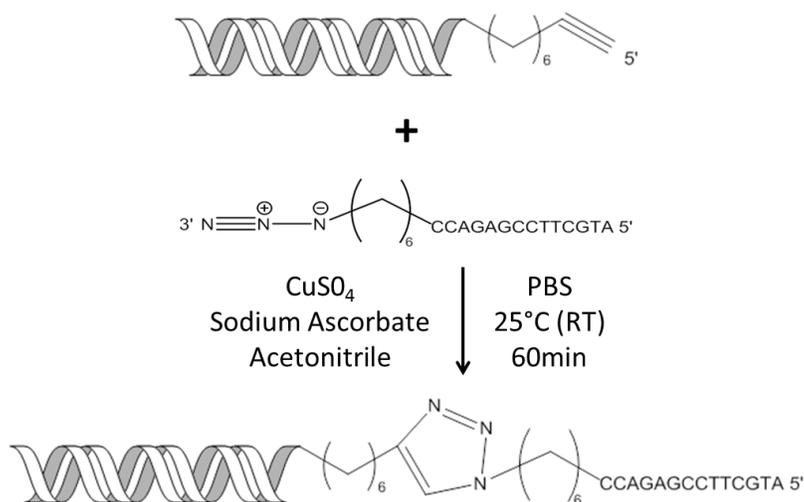
A good portion of this thesis was devoted towards improving surfactant buffers for long DNA micelle ELFSE separations. In order to implement long DNA as nanotags, they would require an overhanging, single stranded complementary sequence (cohesive end) to recognize the miRNA targets. Incorporating this cohesive end onto the long, kB DNA is actually somewhat of a challenge. Enzymatic techniques, such as ligation or PCR, appear to be ideal for cohesive end incorporation. However, these methods are not always suitable for hybridization applications. DNA ligase enzymes have low ligation efficiencies. They simply cannot produce the concentrations of DNA required for hybridization without high costs and time. The polymerase enzyme of PCR would incorporate any overhanging 5' bases on the primer into the amplicon, inhibiting

its ability to recognize miRNA targets. A technique known as sticky-end PCR uses two asymmetric PCR (aPCR) reactions to generate the amplicons with cohesive ends.<sup>[24]</sup> The two strands are amplified separately and annealed after aPCR. We have demonstrated the feasibility of this technique for a 200 base amplicon, but the poor yield of longer DNA by aPCR made it difficult to use for miRNA hybridization.

To overcome the limitations of enzymatic methods, we propose the use of two alternative methods for linking long DNA probes to the short miRNA targets. One solution would utilize click chemistry to link the long DNA strand to a short miRNA specific oligomer. The best example of click chemistry is the copper catalyzed alkyne-azide cycloaddition (CuAAC) reaction (Figure 10.1).<sup>[25]</sup> An alkyne and azide react to form an extremely stable triazole ring, covalently linking the two reactants together. Click chemistry of nucleic acids is quite common and has many advantages. The azide and alkyne functional groups are simple to attach to DNA by phosphoramidite synthesis. These modifications do not greatly affect the DNA binding, especially when incorporated at the distal ends of the oligomers. The reaction itself is quite simple and progresses to near quantitative yields at room temperature in an hour. Finally, it is almost entirely unreactive towards the function groups on biopolymers (-NH<sub>2</sub>, -COOH, -PO<sub>4</sub>, etc.). The bioorthogonal reaction limits potential side products as the two groups only react with each other.

We could simply use an alkyne modified PCR primer to generate long DNA amplicons. After purification, we could then react an azide modified oligomer to

the amplicon to yield long, cohesive end DNA probes. Additionally, the triazole ring may present a barrier for polymerase amplification.<sup>[25]</sup> If we perform the CuAAC reaction prior to PCR, the amplicon may contain the cohesive end by blocking enzyme function at the triazole ring.



**Figure 10.1** – The CuAAC click chemistry reaction between an alkyne modified PCR amplicon and azide modified miRNA specific oligomer. Under simple reaction conditions, the long DNA probes with miRNA specific cohesive ends can be produced.

We could also investigate the attachment of long DNA probes to miRNA using the strand invasion properties of  $\gamma$ PNA.<sup>[26]</sup> This type of system would grant us two advantages for miRNA detection. The  $\gamma$ PNA could capture double stranded DNAs longer than the limits of PCR (10,000 bases), such as plasmids (1-10kB),  $\lambda$ DNA (48.5kB), or T4 DNA templates (166kB). The increased length would give us even more sensitive detection limits. Additionally, the DNA probe would have a  $\gamma$ PNA cohesive end for miRNA hybridization. The miRNA would be sandwiched between two  $\gamma$ PNA probes, taking full advantage of their high affinity and selectivity.

## **10.2.2 – Single Molecule Imaging of DNA in Entangled Micelle Networks**

A good portion of this thesis was focused on understanding the effect of wormlike micelle entanglement on micelle ELFSE DNA separations. We have found that entanglement reduces the efficiency of long DNA separations due to a background sieving effect. Reducing entanglement by lowering separation temperatures or adding small, end-capping surfactants returns the efficient ELFSE separation. We offered a modified ELFSE-sieving model to account for the additional effects of sieving on the separation.

Our understanding of how DNA interacts with the entangled micelle network could benefit from single molecule imaging of DNA, both native and end-alkylated. The DNA could be stained with either YOYO-1 or PicoGreen dye and our microfluidic detection system used for analysis. We can utilize the DNA for pseudo-particle tracking measurements. The diffusion coefficients for native DNA and end-alkylated DNA as a function of surfactant concentration could be compared. We should be able calculate pore sizes of the micelle network from these measurements and compare to our electrophoretically measured pore sizes. The effects of temperature and small-end-capping surfactants can be explored in greater detail as well. We could also monitor the DNA undergoing electrophoresis to observe any changes in conformation. We can confirm our theory assuming DNA is sieving through the micelle network rather than reptation, as well as segregated vs. unsegregated conformations.

## 10.3 – Outlook

Beyond what has already been summarized and presented, here we will discuss the outlook of micelle ELFSE for future scientific contributions. One of the major topics not discussed thus far is sample preparation of the miRNA targets prior to CE analysis. We have demonstrated that our separations are tolerant of goat serum contamination up to a total protein concentration of 100 mg/mL. We should be able to detect miRNA from raw cell lysate with minimal clean-up. However, our current annealing protocol requires ~1 hour of heating/cooling to correctly hybridize miRNA to their respective probes. We may be able to shorten the protocol to ~5 minutes due to the rapid hybridization kinetics in solution. Correct miRNA target hybridization may be accomplished by holding at elevated temperature (~40°C) rather than a slow cool down from 95°C.

Although we may shorten the annealing protocol, this type of hybridization still presents auxiliary steps to achieve detection. Ideally, we would be able to lyse cells, mix and anneal miRNA with hybridization probes, and perform separation with minimal handling. These types of processes are commonly implemented with microchip set-ups, known as micro total analysis systems ( $\mu$ TAS).<sup>[27]</sup> This work has demonstrated successful micelle ELFSE separations on a microchip format. Future work may focus on utilizing  $\mu$ TAS microchips for complete miRNA detection from cells to miRNA profiling in a single chip.

Additionally, isotachopheresis (ITP) has been used to isolate nucleic acids from whole blood and increase hybridization kinetics by concentrating complementary nucleic acids at the ITP interface.<sup>[28], [29]</sup> We may wish to further investigate ITP

beyond the scope of concentration to perform these integrated steps on the microchip for reduced miRNA processing.

Finally, the field of ELFSE has stalled as a whole, mainly due to the inability to generate larger covalently bound drag-tags for separation of longer DNA fragments.<sup>[30]–[33]</sup> The contributions presented in this work towards micelle ELFSE buffer formulation enable a whole new approach for future buffer design. From our increased understanding of micelle network formation and its effect on ELFSE separation, we have demonstrated simple buffer design concepts for long kB DNA separation and increased length of read DNA sequencing. This work should offer a revival of ELFSE for its mainstream application in separation of DNA from 20 to 20kB bases long.

## 10.4 – References

- [1] J. M. Goldman, L. A. Zhang, A. Manna, B. A. Armitage, D. H. Ly, and J. W. Schneider, *Biomacromolecules*, 2013, 14, 2253–2261.
- [2] R. Levicky and A. Horgan, *Trends Biotechnol.*, 2005, 23, 143–149.
- [3] M. Castoldi, S. Schmidt, V. Benes, M. Noerholm, A. E. Kulozik, M. W. Hentze, and M. U. Muckenthaler, *RNA*, 2006, 12, 913–920.
- [4] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner, *J. Mol. Biol.*, 1999, 288, 911–940.
- [5] C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak, and K. J. Guegler, *Nucleic Acids Res.*, 2005, 33, e179.
- [6] B. J. Dodgson, A. Mazouchi, D. W. Wegman, C. C. Gradinaru, and S. N. Krylov, *Anal. Chem.*, 2012, 84, 5470–5474.
- [7] B. Jung, R. Bharadwaj, and J. G. Santiago, *Anal. Chem.*, 2006, 78, 2319–2327.
- [8] T. K. Khurana and J. G. Santiago, *Anal. Chem.*, 2008, 80, 6300–6307.

- [9] Z. Xu, K. Nakamura, A. R. Timerbaev, and T. Hirokawa, *Anal. Chem.*, 2011, 83, 398–401.
- [10] Z. Xu, T. Ando, T. Nishine, A. Arai, and T. Hirokawa, *Electrophoresis*, 2003, 24, 3821–3827.
- [11] D. W. Wegman, F. Ghasemi, A. Khorshidi, B. B. Yang, S. K. Liu, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2015, 87, 1404–1410.
- [12] A. I. Dragan, J. R. Casas-Finet, E. S. Bishop, R. J. Strouse, M. A. Schenerman, and C. D. Geddes, *Biophys. J.*, 2010, 99, 3010–3019.
- [13] J. M. Savard, S. T. Grosser, and J. W. Schneider, *Electrophoresis*, 2008, 29, 2779–2789.
- [14] S. T. Grosser, J. M. Savard, and J. W. Schneider, *Anal. Chem.*, 2007, 79, 9513–9519.
- [15] N. C. Stellwagen, C. Gelfi, and P. G. Righetti, *Biopolymers*, 1997, 42, 687–703.
- [16] A. Bernheim-Groswasser, E. Wachtel, and Y. Talmon, *Langmuir*, 2000, 16, 4131–4140.
- [17] K. Kawasaki, B. Lindman, H. Okabayashi, T. Kato, N. Taguchi, and D. Nozu, *Form. Dyn. Self-Organized Struct. Surfactants Polym. Solut.*, 1997, 106, 57–60.
- [18] T. Kato, S. Anzai, and T. Seimiya, *J. Phys. Chem.*, 1990, 94, 7255–7259.
- [19] Wei and E. S. Yeung, *Anal. Chem.*, 2001, 73, 1776–1783.
- [20] J. Lee, H. Li, and E. S. Yeung, *J. Chromatogr. A*, 2004, 1053, 173–179.
- [21] C. Desruisseaux, G. Drouin, and G. W. Slater, *Macromolecules*, 2001, 34, 5280–5286.
- [22] A. G. Ogston, *Trans. Faraday Soc.*, 1958, 54, 1754–1757.
- [23] J. L. Viovy and T. Duke, *Electrophoresis*, 1993, 14, 322–329.
- [24] G. Zeng, *Biotechniques*, 1998, 25, 206–208.
- [25] A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, 39, 1388–1405.
- [26] G. He, S. Rapireddy, R. Bahal, B. Sahu, and D. H. Ly, *J. Am. Chem. Soc.*, 2009, 131, 12088–12090.
- [27] P. Liu, X. Li, S. A. Greenspoon, J. R. Scherer, and R. A. Mathies, *Lab Chip*, 2011, 11, 1041–1048.

- [28] A. Rogacs, Y. Qu, and J. G. Santiago, *Anal. Chem.*, 2012, 84, 5858–5863.
- [29] M. Bercovici, C. M. Han, J. C. Liao, and J. G. Santiago, *Proc. Natl. Acad. Sci.*, 2012, 109, 11127–11132.
- [30] J. C. Albrecht, A. Kotani, J. S. Lin, S. A. Soper, and A. E. Barron, *Electrophoresis*, 2013, 34, 590–597.
- [31] D. W. Wegman, L. T. Cherney, G. M. Yousef, and S. N. Krylov, *Anal. Chem.*, 2013, 85, 6518–6523.
- [32] D. W. Wegman and S. N. Krylov, *Angew. Chemie*, 2011, 123, 10519–10523.
- [33] H. Song, Y. Wang, C. Garson, and K. Pant, *Microfluid. Nanofluidics*, 2014, 17, 693–699.