FINAL TECHNICAL REPORT

DOE Grant No. DE-FG02-99ER62789

Northwestern University

Annelise Barron, Principle Investigator
Northwestern University
633 Clark Street
Evanston, IL 60208
Phone: 847-491-2778

a. Specific Aims

The research grant had five aims (please see original grant application for these aims, and an introduction to the project). These aims were not substantially modified. During Years 1-3, we focused our efforts on Aims 1-4, although additional studies have been performed more recently which show progress towards Aim 5.

b. Studies and Results

Pursuant to Aim 1, we developed a family of polymeric end-labels, or "drag-tags", for DNA sequencing by End-Labeled Free-Solution Electrophoresis (ELFSE). Synthetic tags were first designed and tested based on poly-*N*-substituted glycine (polypeptoid) chains with lengths of 20, 40 and 60 monomers. These polypeptoids were used to demonstrate single-base resolution of DNA/drag-tag conjugates in free-solution capillary electrophoresis with 20mer and 21mer DNA oligomers. The friction coefficient (α) was found to increase linearly with the peptoid length, with α = 4, 8, and 13 for the 20mer, 40mer and 60mer, respectively. Approximately 5 monomers of *N*-methoxyethylglycine provided drag equivalent to one base of DNA; to achieve a desired α of 80, a 400 monomer-long chain would have to be synthesized. Because it is only practical to synthesize polypeptoid chains up to 60 monomer units, creation of a linear 400mer would require end-on conjugation of seven 60mer chains, which is a challenging proposition.

Two approaches were taken to design drag-tags with large α values: "bottle-brush"-like polypeptoids (BB) and "protein polymer" polypeptides. At the completion of this grant, synthesis had begun on the BB drag-tags. These BB drag-tags show great promise for providing the drag necessary for long-read sequencing, prompting an alternative approach to be carried out in parallel. Through a novel cloning strategy, a series of large, non-natural, repetitive polypeptides (protein polymers) were constructed by genetic engineering, with properties necessary for ELFSE. Using this "controlled cloning" method, monodisperse protein polymers with various numbers of repeats (6, 12, 24 and 48) of a 21-amino acid segment were expressed in *E. coli* and isolated in high purity. Further work beyond the funding period of this grant (supported by the NIH-National Human Genome Research Institute) has used these protein polymers as drag-tags for DNA separations by ELFSE. The results of those studies indicated that large, uncharged drag-tags will have the best DNA-resolving capability for ELFSE separations; up to 230 DNA bases could theoretically be separated using one of the protein polymer drag-tags we created (337mer), had we been successful in obtaining it in a completely monodisperse preparation. These protein polymer drag-tags are based on denatured (unfolded) polypeptide chains with simple, repetitive amino acid sequences, and impose much greater frictional drag per unit molecular weight than the folded proteins that have been used as drag-tags in past studies (e.g. streptavidin).

Also pursuant to Aims 1 and Aim 4, both the smaller polypeptoid and larger protein polymer drag-tags were further characterized and employed for ELFSE separations. The first series of protein polymers had glutamine residues present as 1 in every 7 amino acids in its repetitive sequence [$(GAGQGSA)_{48}G$], which unfortunately underwent deamidation during the cyanogen bromide cleavage procedure necessary to produce

DISCLAIMER

the desired protein sequence. Because each protein polymer drag-tag comprised many such Gln residues in its repetitive sequence, and therefore was deamidated to a different degree, the protein polymers we created were polydisperse and could not be used as effective drag-tags for high-resolution DNA separations by ELFSE. In present work protein polymers with different sequences – not including glutamine -- are being designed, expressed and purified.

In other work related to Aims 1 and 4, we showed that by conjugating monodisperse, fluorescently labeled DNA oligomers to the deamidated protein polymers, the extent of glutamine deamidation could be determined with an accuracy equal or better to that of mass spectrometry. This analytical technique, which we call Free-Solution Capillary Electrophoresis (FSCE) can be used for the high-resolution characterization of polymers and proteins, if they can be conjugated end-on to DNA molecules, by fractionating the DNA-polymer/protein conjugates by free-solution capillary electrophoresis. When we applied this method of analysis to synthetic polypeptoid chains that had been HPLC-purified, impurities in the polypeptoids not observed with standard HPLC could be separated and detected easily and with high accuracy.

Pursuant to Aim 3, we explored 1-color sequencing reactions with dye-labeled primers and a polypeptoid drag-tag pre-conjugated to each primer. We were able to see some evidence of chain extension by the Sanger reaction, although meaningful sequencing data was not derived due to the relatively small drag of the polypeptoid drag-tags we used in this feasibility work.

Also pursuant to Aims 3 and 4, small polypeptoid drag-tags comprising 10, 20 and 30 monomer units were applied to enable ELFSE separations of products from a multiplexed single-base extension (SBE) reaction for single nucleotide polymorphism (SNP) genotyping. Three unique SBE oligonucleotide primers, which probed for mutations of clinical importance in the human p53 gene, were covalently conjugated to three unique polypeptoid frictional end-labels and mixed together. This primer-polypeptoid conjugate cocktail was then used in a multiplexed SBE reaction followed by free-solution separation in a 96-capillary array electrophoresis (CAE) instrument. Simultaneous and accurate genotyping of three p53 loci was achieved on five different DNA templates in a single reaction set and single CAE analysis. Given that parallel separation systems like CAE can be operated in a high-throughput mode, we believe this novel SBE-ELFSE technique will be ideally suited for eventual use in molecular diagnostic clinics, where multiple loci in many samples can be screened for clinically relevant SNP genotypes.

Pursuant to Aim 2, the development of a self-assembled monolayer of hydrophilic polymer to coat microchannels was underway at the conclusion of this grant. Tri(ethylene glycol)-terminated alkyltricholosilane was used to create a very thin, protein-resistant coating on the inner surface of a fused-silica capillary. The EG_3-silane compound created a useful coating for rapid separation of cationic proteins near neutral pH. Sufficient electroosmotic flow was retained to separate both cationic and anionic proteins simultaneously in some cases, and the initial results were quite comparable to those of other thin, charge-neutral coatings. Because no buffer additives were required to maintain the coating's stability, it can potentially be used in a range of applications, including capillary electrophoresis-mass spectrometry (CE-MS) and microchannel coatings of glass microfluidic devices. We are following on this work by investigating more efficient methods of tethering the coating to the glass capillary or microchannel surface.

Pursuant to Aim 5, methods for free-solution electrophoresis in microfluidic devices are under development. Due to the lack of a viscous separating matrix, the ELFSE separations are expected to be quite easy.

c. Significance

The further development of End-Labeled Free-Solution Electrophoresis will greatly simplify DNA separation and sequencing on microfluidic devices. The development and optimization of drag-tags is

critical to the success of this research. The neutral 337mer protein polymer drag-tags devised provide enough friction, in theory, for a read length of about 230 bases. The sequencing performance of these first-generation drag-tags was limited by their polydispersity (including glutamine deamidation), so the next generation of neutral drag-tags without glutamine in their sequence is particularly promising. The further development of BB drag-tags that can be chemically derivatized to form branched, monodisperse polymer structures also promises to drastically increase the diversity of available drag-tags, increasing the α value and the read length while also increasing drag-tag solubility and reducing non-specific interactions with the microchannel surface. The arms of the BB drag-tags can be derivatized with various side chains, including polypeptides and monodisperse commercially-available polymers like poly(ethylene glycol) (PEG). Smaller polypeptoid drag-tags are also proving to be useful for exploring experimental parameters in ELFSE separations, and are directly useful for potential genotyping applications.

d. Present Research

The ELFSE project has continued energetically after the completion of the period of funding of our DOE grant, and as mentioned above it is currently funded by NIH through NHGRI (funding period: 2003-2007). The development of monodisperse drag-tags with high a values is still a considerable obstacle. However, our ability to carry out 4-color sequencing by ELFSE with long reads continue to improve. A 127mer protein polymer drag-tag has already provided a longer sequencing read length (about 160 bases) with sharper, cleaner peaks than that obtained in 1999 using streptavidin, a much larger protein (~ 585 amino acids), as the drag-tag. In collaboration with Dr. Gary Slater's research group at the University of Ottawa, the theory of ELFSE is continuing to develop. Theoretical studies have suggested more efficient designs for drag-tags and also clever ways to enhance DNA separations, by performing ELFSE with drag-tags on both ends of the DNA molecules to be analyzed.

e. Products of the Research

Eleven publications have so far resulted from this project which was begun "from scratch" (from concept) with the support of DOE (please see below). In addition, so far three excellent graduate students (Wyatt N. Vreeland, Jong-In Won, and Robert J. Meagher) have earned their Ph.D.'s in the course of this work. Dr. Vreeland is a scientist at NIST in Maryland, Dr. Won is an assistant professor at Hongik University in South Korea, and Dr. Meagher is a postdoctoral research associate at Sandia National Laboraties in Livermore, CA. Both Dr. Vreeland and Dr. Meagher intend to go on to academic positions. Presently, four other graduate students are working on the project: Jennifer Lin, Russell Haynes, Jordan Bertram, and Jennifer Coyne.

e. Publications

(1) W.N. Vreeland, C. Desruisseaux, A.E. Karger, G. Drouin, G.W. Slater, **A.E. Barron,** "Molar mass profiling of synthetic polymers by free-solution capillary electrophoresis of DNA-polymer conjugates", *Analytical Chemistry* 2001, *73*, 1795-1803.

(2) W.N. Vreeland, G.W. Slater, **A.E. Barron,** "Profiling solid-phase synthesis products by free-solution conjugate capillary electrophoresis", *Bioconjugate Chemistry* 2002, *13*, 663-670.

(3) W.N. Vreeland, R.J. Meagher, **A.E. Barron,** "Multiplexed, high-throughput genotyping by single-base extension and end-labeled free-solution electrophoresis", *Analytical Chemistry* 2002, *74*, 4328-4333.

(4) J. Won, **A.E. Barron,** "A new cloning method for the preparation of long repetitive polypeptides without a sequence requirement", *Macromolecules* 2002, *35*, 8281-8287.

(5) E.A.S. Doherty, R.J. Meagher, M.N. Albarghouthi, **A.E. Barron,** "Microchannel wall coatings for protein separations by capillary and chip electrophoresis", *Electrophoresis* 2003, *24*, 34-54 (review article).

(6) W.N. Vreeland, S.J. Williams, **A.E. Barron,** A.P. Sassi, "Tandem isotachophoresis-zone electrophoresis via base-mediated destacking for increased detection sensitivity in microfluidic systems", *Analytical Chemistry* 2003, *75*, 3059-3065.

(7) J. Won, R.J. Meagher, **A.E. Barron,** "Characterization of glutamine deamidation in a long, repetitive protein polymer via bioconjugate capillary electrophoresis", *Biomacromolecules* 2004, *5*, 618-627.

(8) R.J. Meagher, J. Seong, P.E. Laibinis, **A.E. Barron,** "A very thin coating for capillary zone electrophoresis of proteins based on a tri(ethylene glycol)-terminated alkyltrichlorosilane", *Electrophoresis* 2004, *25*, 405-414.

(9) R.J. Meagher, J. Won, L.C. McCormick, S. Nedelcu, M.M. Bertrand, J.L. Bertram, G. Drouin, **A.E. Barron,** G.W. Slater, "End-labeled free solution electrophoresis of DNA", *Electrophoresis* 2005, *26*, 331-350 (review article).

(10) J. Won, R.J. Meagher, **A.E. Barron,** "Protein polymer drag-tags for DNA separations by End-Labeled Free-Solution Electrophoresis", *Electrophoresis* 2005, *26*, 2138-2148.

(11) R.D. Haynes, R.J. Meagher, J. Won, F.M. Bogdan, **A.E. Barron,** "Comb-like, monodisperse polypeptoid drag-tags for DNA separations by End-Labeled Free-Solution Electrophoresis (ELFSE)". *Bioconjugate Chemistry* 2005, *16*, 929-938.