Published in final edited form as:

Electrophoresis. 2010 July ; 31(14): 2461–2464. doi:10.1002/elps.201000200.

Capillary sieving electrophoresis and micellar electrokinetic capillary chromatography produce highly correlated separation of tryptic digests

Jane A. Dickerson and Norman J. Dovichi

Department of Chemistry, University of Washington, Seattle WA 98195-1700

Abstract

We perform two-dimensional capillary electrophoresis on fluorescently labeled proteins and peptides. Capillary sieving electrophoresis was performed in the first dimension and micellar electrokinetic capillary chromatography was performed in the second. A cellular homogenate was labeled with the fluorogenic reagent FQ and separated using the system. This homogenate generated a pair of ridges; the first had essentially constant migration time in the CSE dimension, while the second had essentially constant migration time in the MEKC dimension. In addition a few spots were scattered through the electropherogram. The same homogenate was digested using trypsin, and then labeled and subjected to the two dimensional separation. In this case, the two ridges observed from the original two-dimensional separation disappeared, and were replaced by a set of spots that fell along the diagonal. Those spots were identified using a local-maximum algorithm and each was fit using a two-dimensional Gaussian surface by an unsupervised nonlinear least squares regression algorithm. The migration times of the tryptic digest components were highly correlated ($r = 0.862$). When the slowest migrating components were eliminated from the analysis, the correlation coefficient improved to $r = 0.956$.

Sieving electrophoresis separates proteins based on their size. In this separation method, proteins are treated with an anionic surfactant, such as sodium dodecyl sulfate (SDS). At low concentrations of surfactant, the surfactant is incorporated at the rate of ~0.5 g surfactant/g of protein.[1] At higher concentrations, which approach the critical micelle concentration, the surfactant is incorporated at a saturated rate of 1.2 – 1.5 g surfactant / g of protein. The incorporated surfactant produces a more-or-less constant charge-to-size ratio in the protein. Like DNA electrophoresis, the surfactant-treated protein undergoes a size-base separation in the presence of a sieving matrix, such as polyacrylamide, and is the basis of SDS polyacrylamide gel electrophoresis (SDS/PAGE).

Micellar electrokinetic capillary electrophoresis (MEKC) performs separation in the presence of a surfactant at a concentration greater than the CMC but in the absence of a sieving matrix. For low molecular weight analyte, the separation is commonly modeled in terms of the partitioning of analyte between micelles and the solution. The mechanism for separation of proteins is less clear. As noted above, the uniform incorporation of anionic surfactant into a protein will lead to a size-based separation in the presence of a sieving matrix. However, it is known that estimates of protein size determined by SDS/PAGE are often in error.[2] These errors like arise from incorporation of SDS at an anomalous rate due to protein composition and conformation.

In this paper we perform two-dimensional capillary electrophoresis using capillary sieving electrophoresis in the first dimension and MEKC in the second. Two samples are analyzed. The first is a fluorescently-labeled cellular homogenate produced from a Barrett's esophageal cell line (CP18821). The second is a fluorescently-labeled tryptic digest.

Unless stated, all reagents were purchased from Sigma-Aldrich (St. Louis, MO USA). Solutions were made with distilled, deionized water (Barnstead Nanopure, Boston, MA USA) and vacuum filtered through a 0.22 μm filter (Millipore, Billerica, MA USA). Ultra Trol LN and EoTrol LN dynamic coatings were purchased from Target Discovery (Palo Alto, CA USA). FQ was purchased from Molecular Probes (Eugene, OR USA). FQ was dissolved in methanol and divided into 10 μL aliquots of 100 nmol. The aliquots were lyophilized using a SpeedVac. Uncoated fused-silica capillaries were purchased from Polymicro Technologies (Phoenix, AZ USA). Cell media, Keratinocyte-serum free medium supplemented with Bovine Pituitary Extract (BPE) and human recombinant Epidermal Growth Factor (rEGF) was purchased from Invitrogen (Carlsbad, CA USA).

To label proteins with FQ, 5 μL of protein solution was added to 5 μL of 5 mM KCN in 10 mM borate (pH 9) in a tube with 100 nmol of lyophilized FQ. The mixture was heated at 65°C for 5 min, after which the sample was sufficiently labeled and a color change from clear to orange was observed. The sample was quenched with 90 μL of 100-100-3.5 mM Tris CHES SDS buffer. The sample was stored on ice until analysis.

We have previously reported the instrument and interface for 2D capillary electrophoresis.[3,4] The 2D set up is comprised of two fused silica capillaries (30 μm inner diameter, 147 μm outer diameter, 20–25 cm length) dynamically coated with Ultra Trol LN, joined at an interface, and connected to two high voltage power supplies. Sample was injected hydrodynamically at 5 psi for 3 sec. 2D separations were performed using the same conditions previously reported.[5]

Our laser-induced fluorescence detectors have been described in detail.[6,7] Briefly, we employed a post-column sheath-flow cuvette. Fluorescence was excited by a 473-nm solid-state diode-pumped laser (Lasermate Group, Pionoma, CA USA), collected with an M-PLAN 60x, 0.7 NA microscope objective (Universe Kogaku, Oyster Bay, NY USA), and filtered with a 580LP long-pass filter (Omega Optical, Brattleboro, VT USA). Light was detected by an avalanche photodiode single-photon counting module (EG&G Canada, Vaudreuil, Canada). Voltage programming and fluorescence detection were controlled by home-built LabView software. The signal was corrected to account for the dead-time response of the photodetector.

The data were analyzed using Matlab.[5] Data were first corrected for the nonlinear response of the avalanche photodiodes. They were then passed through a seven point median filter to remove spikes due to light scatter from particles in the solution. A two-dimensional data array was constructed from stacks of successive 13-s long segments of the filtered data. Data were then down-sampled to 10 Hz to generate a more manageable size array. The baseline was estimated through a morphological opening operation across a circular region with six-pixel radius. The baseline was then subtracted to generate the reconstructed electropherogram.

An algorithm was used to identify all local maxima that were roughly four standard deviations above the noise in the baseline, estimated as the standard deviation of a relatively featureless region in the reconstructed electropherogram. Next, an unsupervised nonlinear least squares routine based on the Marquardt algorithm was used to fit a Gaussian surface to each of the local maximum. The fits were performed across a square region that was centered at the local maximum and was five fractions tall in the CSE dimension and 2.5 sec wide in the MEKC dimension. Fits that failed to converge or that generated a spot center that was outside of the fitting region were rejected.

The separation of fluorescently labeled proteins is challenging. Most labeling reagents react with primary amines (the N-terminus and the α -amine of lysine residues). Common labeling

reagents, such as fluorescein isothiocyanate, are anionic, and the labeling reaction converts the cationic lysine residue into a negatively charged product. This change in charge is problematic because the labeling reaction seldom goes to completion. There are 2^n-1 possible reaction products, and the reaction converts each protein into a complex mixture of products, which results in a very complex and essentially useless electropherogram.[8–10] Other labeling reactions, such as 3-(2-furoyl)quinoline-2-carboxaldehyde, convert the cationic lysine residue into a neutral product. In many buffers, this labeled product generates poor separation performance. However, we discovered that the labeled proteins generate outstanding electrophoretic behavior in buffers that contain an anionic surfactant such as SDS.[11] It appears that the SDS ion-pairs with unlabeled lysine residues, neutralizing the charge and introducing a similar hydrodynamic drag as the label itself. We employed FQ-labeled proteins in this paper.

Figure 1 presents the two-dimensional electropherograms generated by a cellular homogenate that had been labeled with FQ. The top electropherogram is of the homogenate itself while the bottom electropherogram is of the tryptic digest of the homogenate that had been labeled with FQ. Note that it is not appropriate to digest the labeled homogenate; the label is likely to interfere with the action of trypsin, which cleaves unlabeled lysine residues.

The two-dimensional electropherogram of the intact homogenate shows a nearly horizontal streak for fractions that migrate early in the CSE dimension (dashed line). Note that the zero time point in the MEKC dimension is arbitrary and can be shifted as appropriate. A second streak extends at roughly a 60 angle for intermediate fractions. This streak falls off the right edge of the reconstructed electropherogram and reappears on the left. The use of a longer MEKC separation window would not have split the components between the two sides of the electropherogram. The supporting material section shows the reconstructed electropherogram generated by pasting a portion of the left side of the electropherogram to the right side, which clearly shows this protein ridge. Finally, a few spots are present between the dashed lines.

The two-dimensional electropherogram of the labeled tryptic digest shows that all of the components along the dashed lines of the top panel have disappeared, thus confirming that they are proteins. Several components are present in both electropherograms, including the relatively intense spot in fraction 225. The protein components are replaced with a series of spots that nearly fall on the diagonal of the electropherogram, dashed curve. These spots must be tryptic peptides because they are not present in the intact homogenate electropherogram.

To further investigate the peptide electropherogram, we used an unsupervised routine to identify the local maxima in the electropherogram and to fit a Gaussian surface to each spot. [5] The 20 most intense spots are marked in Figure 2. Three components were also present in the protein electropherogram and are marked with a circle.

We performed a weighted linear least squares analysis of the migration time in the two dimensions, figure 3. The weights for the fit were the square root of the sum of the peak variances in both the CSE and MEKC dimensions, as determined by the regression analysis used in figure 2.

The migration times are reasonably well correlated ($r = 0.86$), particularly for the faster migrating components. We did a second fit on the 12 fastest migrating components. The data were very highly correlated ($X^2_{\nu}=4$ for 10 degrees of freedom; an unweighted fit gave $r = 0.96$).

The data of figure 1 provide insights into the separation mechanisms for peptides and proteins using CSE and MEKC. It is clear that MEKC provides mediocre resolution of proteins. Presumably, the constant size-to-charge ratio produced by the uniform incorporation of surfactant into the protein results in consistent mobility for those components. The contrast with the separation of tryptic peptides is striking. Those peptides show very highly correlated separation in the two dimensions, and smaller peptides show the highest correlation. It is tempting to suggest that peptides that are smaller than the micelle are able to partition into the micelle, and that the separation mechanism is similar to classic MEKC. This hypothesis requires that the peptide-SDS complex is smaller than the pore size of the sieving matrix. It will be interesting to revisit this problem using standard peptides of known size.

Acknowledgments

We acknowledge the assistance of Professor James Kraly of Keene State University in setting up this experiment. This research was funded by a grant from the National Institutes of Health (R33CA122900).

References

1. Takagi T, Tsujii K, Shirahama K. J Biochem. 1975; 77:939–947. [PubMed: 1158859]
2. Banker GA, Cotman CW. J Biol Chem. 1972; 247:5856–5861. [PubMed: 4560419]
3. Michels DA, Hu S, Schoenherr RM, Eggertson MJ, Dovichi NJ. Mol Cell Proteomics. 2002; 1:69–74. [PubMed: 12096142]
4. Michels DA, Hu S, Dambrowitz KA, Eggertson MJ, et al. Electrophoresis. 2004; 25:3098–3105. [PubMed: 15472976]
5. Kraly JR, Jones MR, Gomez DG, Dickerson JA, et al. Anal Chem. 2006; 78:5977–5986. [PubMed: 16944874]
6. Wu S, Dovichi NJ. J Chromatogr. 1989; 480:141–155. [PubMed: 12408116]
7. Cheng YF, Dovichi NJ. Science. 1988; 242:562–564. [PubMed: 3140381]
8. Zhao JY, Waldron KC, Miller J, Zhang JZ, Harke H, Dovichi NJ. J Chromatogr. 1992; 608:239–242. [PubMed: 1430027]
9. Richards DP, Stathakis C, Polakowski R, Ahmadzadeh H, Dovichi NJ. J Chromatogr A. 1999; 853:21–25. [PubMed: 10486708]
10. Craig DB, Dovichi N. J Anal Chem. 1998; 70:2493–2494.
11. Pinto DM, Arriaga EA, Craig D, Angelova J, Sharma N, Ahmadzadeh H, Dovichi NJ, Boulet CA. Anal Chem. 1997; 69:3015–3021.

Figure 1.

Separation of a cellular homogenate by two-dimensional capillary electrophoresis. CSE was used for the first dimension separation and MEKC was used for the second. The intact homogenate was used in the top panel while the tryptic digest of the homogenate was used in the bottom. FQ was used to label the intact homogenate in the top panel and to label the tryptic digest in the bottom panel. The color is related to the logarithm of the fluorescence intensity. A false-color intensity map is used, where blue is low fluorescence amplitude and red is high amplitude. The image was slightly overexposed to highlight lower abundance components. The dashed curves are guides to the eye.

Figure 2.

The 20 most intense components in the two-dimensional capillary electropherogram of the tryptic digest of figure 1. Three components were also present in the protein electropherogram; those components are marked with a circle. The others are marked with a plus sign.

Figure 3.

Weighted least squares fit of a line to the migration times determined from the 17 spots identified with the plus sign in figure 2. The weights were determined as the square root of the sum of the peak variances determined by the nonlinear fit of a Gaussian surface to the spots in figure 2. χ^2_ν was 33 for 15 degrees of freedom. An unweighted fit was also performed and generated a correlation coefficient of $r = 0.86$.