Capillary Electrophoresis of Oligonucleotides

Contents

1. Introduction

Electrophoresis, operationally defined as the migration of analytes within an electrolyte solution under the influence of an electrical field, was first described by Tiselius in the 1930s [1]. Since its introduction, electrophoresis has become a primary tool of biochemical separation and has spawned numerous variants. One of those variants, capillary electrophoresis (CE) began in the 1960s as a technique for understanding ambiguities of the electrophoretic process itself [2]. In the method of zone electrophoresis, products of similar size are cleanly sequestered from those of other sizes; these areas of product are known as zones. Unfortunately, this method suffered from the phenomenon known as zone broadening in which the spaces between the product zones became blurred and the areas of product could not be cleanly resolved. The introduction of solid and semi-solid matrices for electrophoretic separation, such as paper and starch, relieved this problem to some extent. A major advance in the 1970s by Mikkers et al. was the use of narrow tubes made of chemically and electrically inert materials [3]. These tubes not only avoided the zone broadening problem but had the added advantage of compressing the distances between clearly defined zones well beyond the capabilities of other extant methods.

Jorgenson and Lukacs improved on the use of inert tubes by pioneering open-ended glass capillaries [4]. Borrowing the concept of theoretical plates from chromatography [5], Jorgenson and Lukacs showed that separation efficiency in their capillaries could be expressed in terms of the number of theoretical plates, N, in the form,

$$N = \mu V/2D \qquad\qquad (1)$$

where μ is electrophoretic mobility, V is the applied voltage, and D is the molecular diffusion coefficient of the solute in the zone [6]. Since both μ and D are intrinsic properties of the solute, the implication of the result is that efficiencies depend entirely on voltage. Moreover, they found that, apart from run time, the length of the capillary had no effect on efficiency. Thus,

the best separation efficiencies were achieved at very high voltages in capillaries sufficiently long to dissipate the heat generated by the voltage [6].

Once this foundation had been established, the method of capillary electrophoresis became a powerful and popular method for separating all types of biopolymers including proteins, peptides, and oligonucleotides [7]. The resolving capabilities of capillaries that reduced heating and allowed for extremely high electrical fields at once made the technique attractive for DNA sequencing [8]. Aiding this application was the development of flexible, fused silica capillaries and an improved detector system [9, 10]. In 1993, a non-cross-linking polyacrylamide was developed that allowed for reuse of capillaries and the first high-resolution DNA sequences from a multi-use capillary/polymer system were reported in 1995 [11, 12].

The technique of capillary gel electrophoresis (CGE) has also become an industry standard for assessing the purity of biochemical syntheses. Here, the application of CGE as a QC platform in the synthesis of oligonucleotides is presented.
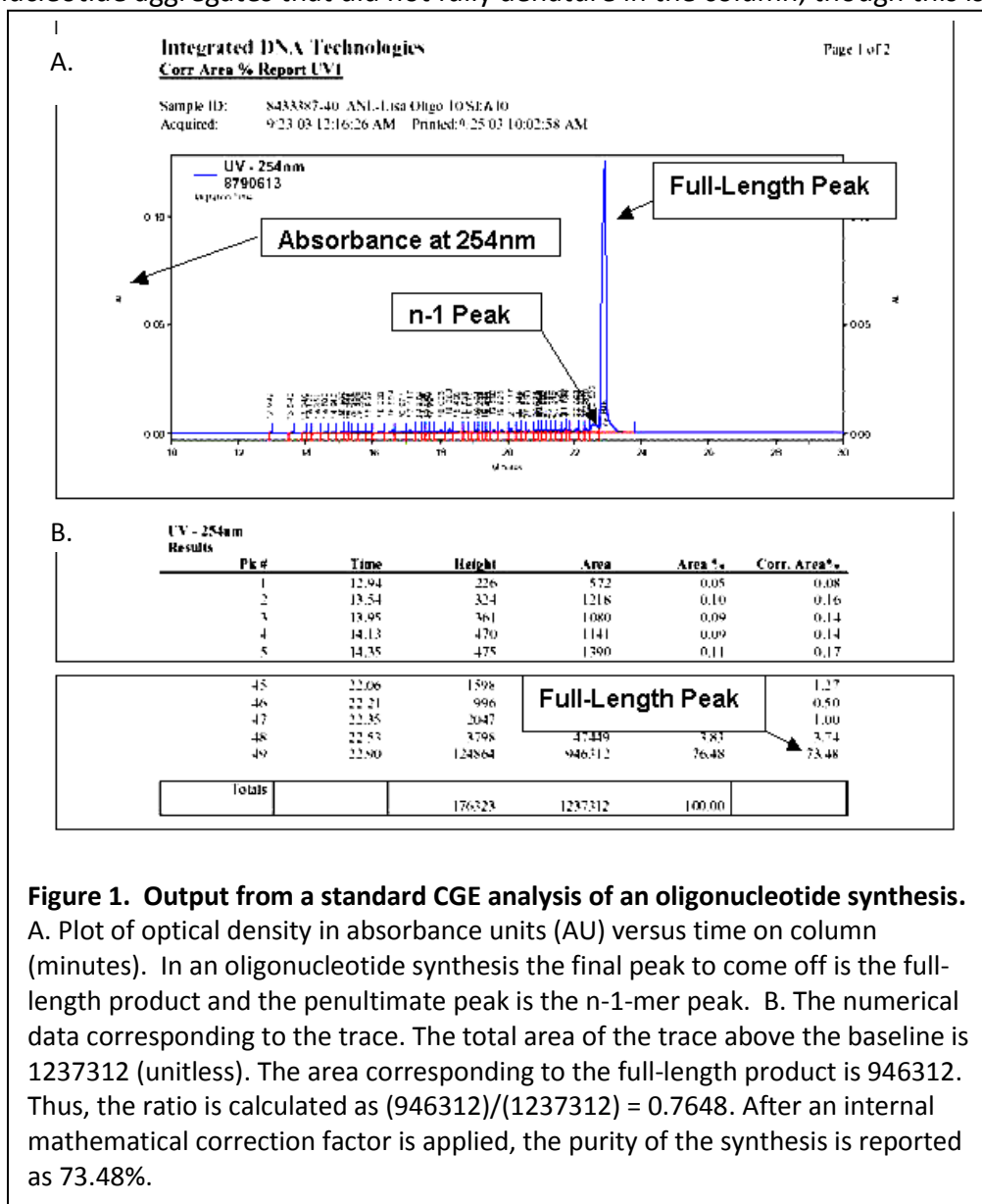
---

2. Capillary Gel Electrophoresis

The essence of CGE as a QC platform in oligonucleotide synthesis is the resolution of polyacrylamide gel electrophoresis (PAGE) in a high throughput capillary format. PAGE is time-consuming, labor intensive, requires relatively large amounts of sample material and, most importantly, does not supply accurate quantitative information about purity. In contrast, CGE is designed to deliver accurate quantitative information in an automated, high throughput environment.
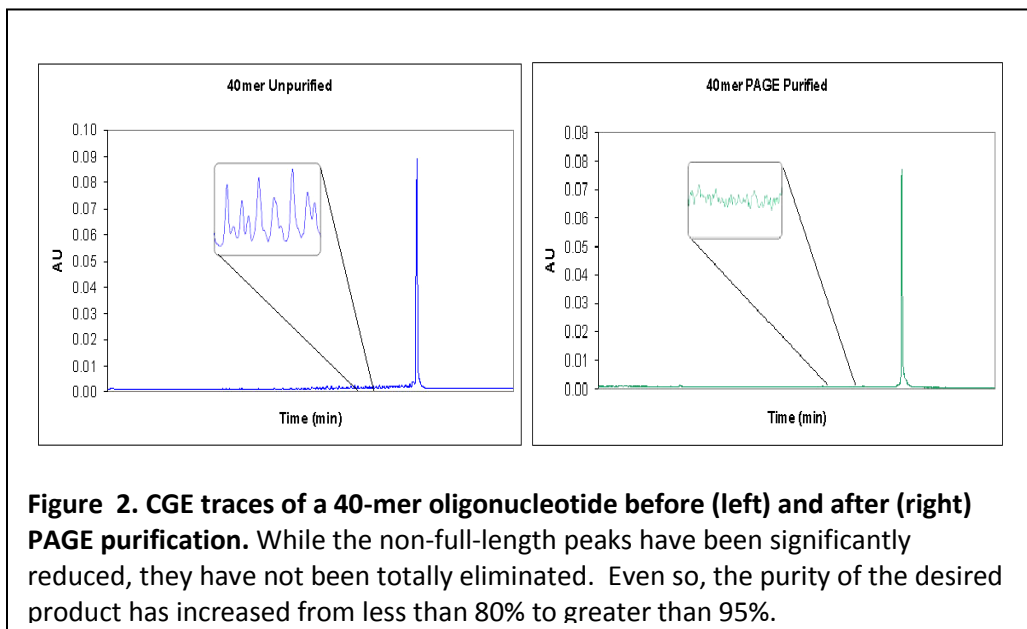
Very small volumes of sample (nanoliter amounts) are introduced into a fused silica capillary, with an inner diameter of 100 μm, filled with a sieving matrix (gel) containing 7M urea. The capillary is heated to 30°C which, along with the urea, denatures the oligonucleotide. In the presence of a high voltage electrical field the oligonucleotides migrate through the gel and the different species present in the injected sample are separated by size. As per equation (1) the resolving power of the capillary is determined by the voltage and, in contemporary systems, there is a potential for hundreds of distinct zones or, in chromatography terms, theoretical plates. An optical window is positioned 20 cm from the injection point and absorbance is measured at 254 nm as the zones migrate past the window. The optical density of each zone is plotted as a function of time in the capillary to create an electropherogram, or trace (Figure 1). Analysis of the CE optical density data begins by integrating the entire area of the trace that lies above the zero absorbance baseline. Oligonucleotide purity is simply the ratio of the area of the peak corresponding to the full-length synthesis product to the total area of the trace.

As shown in Figure 1, the final, main, peak is the full-length synthesis product. The penultimate peak usually represents the population of products in which a single nucleotide is missing. These are, collectively, the n-1-mers. In theory, the missing base can occur anywhere in the

sequence except at the 3'-most base. This is because the 3'-most base is covalently attached to the solid synthesis support (the CPG). Since chemical reactions are never 100% efficient, there will always be a finite, non-zero probability that not all of the available couplings will occur at each step of the synthesis. The majority of these failed couplings are capped in order to prevent additional couplings to an incorrect sequence, but a small population of the growing oligo-nucleotides will remain reactive. The n-1 peak will lie immediately to the left of the main peak. Other peaks to the left are composed primarily of capped oligonucleotides from n-2 to n-(n-1). Occasionally, peaks will appear to the right of the main peak. Such peaks contain oligonucleotide species that were not completely deprotected, and so still contain an extra trityl group, or have extra bases and are (n + 1)-mers. These slower migrating entities may also be oligonucleotide aggregates that did not fully denature in the column, though this is rare.

Figure 1. Output from a standard CGE analysis of an oligonucleotide synthesis. A. Plot of optical density in absorbance units (AU) versus time on column (minutes). In an oligonucleotide synthesis the final peak to come off is the full-length product and the penultimate peak is the n-1-mer peak. B. The numerical data corresponding to the trace. The total area of the trace above the baseline is 1237312 (unitless). The area corresponding to the full-length product is 946312. Thus, the ratio is calculated as (946312)/(1237312) = 0.7648. After an internal mathematical correction factor is applied, the purity of the synthesis is reported as 73.48%.

The sensitivity that can be achieved by current CGE platforms is exemplified in Figure 2. Here, a 40-mer oligonucleotide is assessed before and after polyacrylamide gel electrophoresis (PAGE). Even though the raw synthesis of the 40-mer is good, there are truncation products that require removal. Following PAGE purification, the CE shows that the vast majority of the mass of the (n-1) through n-(n-1) species has been removed.

Figure 2. CGE traces of a 40-mer oligonucleotide before (left) and after (right) PAGE purification. While the non-full-length peaks have been significantly reduced, they have not been totally eliminated. Even so, the purity of the desired product has increased from less than 80% to greater than 95%.

CGE is currently the accepted standard for assessing oligonucleotide purity. However, the data provided by a CGE analysis contains more information than just the ratio of the full-length product peak to the total. CGE trace data can be used to back-calculate average base coupling efficiencies for each synthesis if an unpurified oligonucleotide is analyzed. In addition, the length of an oligonucleotide can be estimated by CE using known size standards. For example, the size standard used by IDT is a purified 40-mer that is used to define a retention time for the capillary. For example, if the 40-mer standard has a retention time of 22 minutes and the average time between nucleotides is 0.2 minutes, then the retention time of a 20-mer will be 22 – (20 x 0.2) = 18 minutes. Similarly, for a 70-mer, retention time is 22 + (30 x 0.2) = 28 minutes. Such calculations are, of course, only rough estimates because retention times of oligonucleotides of the same length will vary by specific nucleotide sequence. Thus, this method cannot be used to precisely determine the length of an oligonucleotide. However, it can be used to perform standardization checks on capillaries, gel matrix lots, etc.

3. Alternative Oligonucleotide QC Methods

Other methods can be used to test oligonucleotide purity. However, it is important to understand the limitations of these methods and how they compare to CGE. For example,

reverse-phase HPLC (RP-HPLC), in which oligonucleotide species are separated on the basis of hydrophobicity, is an excellent method for assessing the efficiency of tagging a hydrophobic ligand, such as a fluorescent dye or biotin group, onto an oligonucleotide. However, this method provides a poor assessment of purity. An oligonucleotide that appears to be 100% pure by RP-HPLC may be observed to be only 80% pure by CGE. Anion-exchange HPLC (IE-HPLC) is another method that can potentially provide information about purity but it does not perform any better than RP-HPLC in this capacity.

Mass spectrometry (MS) is a method that is used to complement CGE. A CGE analysis of an oligonucleotide will provide an estimate of length and purity while the MS analysis provides the exact molecular weight (compound identity). Mass spectrometry will indicate the presence or absence of additional molecular species in the sample but these different species will ionize at different efficiencies. Thus, MS traces are not normally able to provide an accurate assessment of purity. In fact, it has been shown that there is no direct comparability between the size of an n-1 peak on a MALDI-TOF trace and the percentage of (n-1)-mer identified on a CGE trace of the same oligonucleotide sample.

---

4. CGE at IDT

The IDT High Throughput Analysis Department has the capacity to perform thousands of CGE analyses daily. At the present time we provide CGE on all oligonucleotides that have a purity guarantee (typically those shorter than 60 bases). Both CGE and MALDI-TOF QC results are available on the IDT web site.

The minimum sample requirement for CGE analysis is 0.3 ODs of sample. This provides sufficient mass to carry out multiple runs if necessary. The best results are obtained if the oligonucleotide has been desalted because salt ions will compete with the oligonucleotide for entry into the capillary during electrokinetic injection. Also, CGE analyses carried out in salt buffers have characteristically rounded peaks with decreased resolution compared to analyses carried out in water.

Finally, certain oligonucleotides will display anomalous migration in CGE. Secondary structures, such as hairpins, routinely behave anomalously in CGE. In such cases, the traces are forwarded to the IDT Technical Support Group. Members of the Technical Support Group provide assistance in interpreting non-standard CGE results.

---

5. References

1.   Tiselius A. (1959) Introduction. in Electrophoresis: theory, methods, and applications, M. Bier, Editor. Academic Press: New York.

2. Hjerten S. (1967) Free zone electrophoresis. Chromatogr Rev, 9(2): 122−219.
3. Mikkers FEP, Everaerts FM, and Verheggen TPEM. (1979) High-performance zone electrophoresis. J. Chromatogr, 169: 11−20.
4. Jorgenson JW and Lukacs KD. (1981) Zone electrophoresis in open-tubular glass capillaries. Anal. Chem., 53(8): 1298−1302.
5. Martin AJ and Synge RL. (1941) A new form of chromatogram employing two liquid phases: A theory of chromatography. 2. Application to the micro-determination of the higher monoamino-acids in proteins. Biochem J, 35(12): 1358−1368.
6. Jorgenson JW and Lukacs KD. (1983) Capillary zone electrophoresis. Science, 222(4621): 266−272.
7. Kemp G. (1998) Capillary electrophoresis: a versatile family of analytical techniques. Biotechnol Appl Biochem, 27 (Pt 1): 9−17.
8. Viovy JL and Duke T. (1993) DNA electrophoresis in polymer solutions: Ogston sieving, reptation and constraint release. Electrophoresis, 14(4): 322−329.
9. Swerdlow H, Wu SL, et al. (1990) Capillary gel electrophoresis for DNA sequencing. Laser-induced fluorescence detection with the sheath flow cuvette. J Chromatogr, 516(1): 61−67.
10. Swerdlow H, Zhang JZ, et al. (1991) Three DNA sequencing methods using capillary gel electrophoresis and laser-induced fluorescence. Anal Chem, 63(24): 2835−2841.
11. Ruiz-Martinez MC, Berka J, et al. (1993) DNA sequencing by capillary electrophoresis with replaceable linear polyacrylamide and laser-induced fluorescence detection. Anal Chem, 65(20): 2851−2858.
12. Zhang J, Fang Y, et al. (1995) Use of non-cross-linked polyacrylamide for four-color DNA sequencing by capillary electrophoresis separation of fragments up to 640 bases in length in two hours. Anal Chem, 67(24): 4589−93.