ELSEVIER

Journal of Chromatography B, 783 (2003) 61–72

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Purification of dye-labeled oligonucleotides by ion-pair reversed-phase high-performance liquid chromatography

Kenneth J. Fountain, Martin Gilar*, Yeva Budman, John C. Gebler

Waters Corporation, 34 Maple Street, Milford, MA 01757, USA

Received 26 March 2002; received in revised form 8 July 2002; accepted 10 July 2002

Abstract

Singly- and dually-labeled synthetic oligonucleotides were purified by ion-pair reversed-phase high-performance liquid chromatography using a 50×4.6-mm column packed with porous, 2.5 μm C_{18} sorbent. We studied the mechanism of dye-labeled oligonucleotide retention in order to improve the quality of purification. By-products of oligonucleotide synthesis were characterized by liquid chromatography with mass spectrometry detection (LC–MS). We purified oligonucleotides labeled with 6-carboxyfluorescein (6FAM), hexachlorofluorescein (HEX), tetrachlorofluorescein (TET), carboxytetramethylrhodamine (TAMRA) and indodicarboxycyanine (Cy™3) dyes, as well as dually-labeled TaqMan™ probes. Purification of a 0.1-μmole oligonucleotide synthesis in a single injection was demonstrated.

Keywords: Genotyping; Oligonucleotides

1. Introduction

The production of synthetic oligonucleotides has increased dramatically in recent years due to their use in many biological applications such as DNA sequencing [1,2], genotyping [3,4], and quantitative polymerase chain reaction (PCR) [5,6]. These applications often utilize fluorescent oligonucleotides to improve the sensitivity and throughput of the assay [7,8]. However, oligonucleotide synthesis yields a target product contaminated with truncated by-products, so-called failure sequence oligonucleotides [9–11]. It is challenging to produce dye-labeled, especially dually-labeled, oligonucleotides with desirable purity (>90%). Since the performance of genotyping and quantitative PCR assays is sensitive to the quality of fluorescent probes, it is crucial to purify these oligonucleotides prior to use [10,12].

The current methods of choice for fluorescent oligonucleotide purification are polyacrylamide gel electrophoresis (PAGE) and reversed-phase high-performance liquid chromatography (RP-HPLC). While PAGE is considered to provide the highest grade product (>98%) and is traditionally used for its high separation performance, it is difficult to scale up for typical oligonucleotide syntheses of 0.2–1 μmoles [13]. Reversed-phase HPLC became popular for purification of ''trityl on'' oligonucleotides [9,14–16]. Recently, ion-pair (IP) RP-HPLC has been shown to be an efficient method of purification for detritylated synthetic oligonucleotides [10,12,17–21]. The advantage of IP-RP-HPLC over ion-ex-

*Corresponding author. Tel.: +1-508-482-2000; fax: +1-508-482-3100.

E-mail address: Martin_Gilar@waters.com (M. Gilar).

1570-0232/02/$ – see front matter
PII: S1570-0232(02)00490-7

change chromatography, PAGE and the ''trityl on'' method is that the collected fraction does not require post-purification desalting or oligonucleotide deprotection. Due to the compatibility of the mobile phases with mass spectrometry (MS) detection [22–24], IP-RP-HPLC can be used for the characterization of oligonucleotide impurities [10].

With the development of new methods for single nucleotide polymorphism detection and sensitive diagnostic techniques for vial detection based on fluorescently labeled oligonucleotides, there is an urgent need for high-performance purification techniques of these special oligonucleotide probes. While the purification of unlabeled oligonucleotides has been addressed in literature, only a few reports have dealt with the subject of dye-labeled oligonucleotides [12]. It was shown that the retention of DNA fragments in IP-RP-HPLC depends on the nature of the fluorescent tag [21] rather than the length of oligonucleotide. To the best of our knowledge, no comprehensive study describing the HPLC analysis and purification of labeled oligonucleotides has been published to date.

In this work, we used an LC–MS method for the identification of dye-labeled failure sequences and investigated their retention behavior. The impact of the label on oligonucleotide retention was studied in order to improve the IP-RP-HPLC methods for their purification. The developed IP-RP-HPLC method will be a useful application for semi-preparative purifications of dye-labeled diagnostic oligonucleotides as well as recently introduced dually- and triply-labeled oligonucleotide probes.

2. Experimental

2.1. Chemicals and oligonucleotide samples

Triethylamine (TEA), 99.5%, glacial acetic acid, 99.99%, and HPLC grade acetonitrile were purchased from J.T. Baker (Phillipsburg, NJ, USA). Water (18 MΩ cm) was purified in-house using a Milli-Q system (Millipore, Bedford, MA, USA). Polyethylene glycol (PEG), MW 35 000, was purchased from Fluka (Milwaukee, WI, USA). Tris-[hydroxymethyl]aminomethane (Tris), urea, boric acid, ethylenediaminetetraacetic acid, disodium salt dihydrate (EDTA), and *N,N*-Dimethylbutylamine, 99%, were all purchased from Sigma–Aldrich (St. Louis, MO, USA). Crotalus adamanteus venom phosphodiesterase I (3′-exonuclease), lyophilized powder, was purchased from USB Corporation (Cleveland, OH, USA). Unlabeled oligonucleotides were obtained from Midland Certified Reagent Company (Midland, TX, USA), and singly-labeled oligonucleotides were purchased from Sigma-Genosys (The Woodlands, TX, USA). The sequences are as follows (5′–3′): 25 mer GAC TTA GAC TTA GAC TTA GAC TTA G; 25 mer (HEX)-TTT GAC TTA GAC TTA GAC TTA GTT T; 25 mer (TET)-TTT GAC TTA GAC TTA GAC TTA GTT T; 25 mer (Cy3)-TTT GAC TTA GAC TTA GAC TTA GTT T; 25 mer (6FAM)-TTT GAC TTA GAC TTA GAC TTA GTT T; 25 mer TTT GAC TTA GAC TTA GAC TTA GTT T-(TAMRA); 15 mer (6FAM)-TTT TTT TTT TTT TTT. Dually-labeled oligonucleotides were provided by Genetics Institute (Cambridge, MA, USA); sequences unknown.

2.2. HPLC equipment and conditions

An Alliance™ 2690 Separations Module with a 996 Photodiode Array Detector (Waters Corporation, Milford, MA, USA) was used for analytical and semi-preparative HPLC purifications. The built-in column heater of the Alliance™ 2690 HPLC system was set to 60 °C. Oligonucleotides were separated using a 50×4.6-mm XTerra™ MS C_{18} column packed with 2.5 μm particles, average pore diameter 140 Å. The mobile phases were comprised of 0.1 *M* triethylammonium acetate (TEAA), pH 7, and acetonitrile. The flow-rate was 1 ml/min or 0.5 ml/min. For gradient and other conditions see figure captions. Purified oligonucleotides were evaporated using a centrifugal vacuum concentrator (SpeedVac® SC110A, Savant, Holbrook, NY, USA).

2.3. Capillary electrophoresis

Capillary gel electrophoresis (CGE) was performed on a P/ACE™ MDQ Molecular Characterization System (Beckman–Coulter, Fullerton, CA, USA) using a BioCAP™ DNA Analysis Capillary (Bio-Rad Laboratories, Hercules, CA, USA), 375/75 μm O.D./I.D. (effective capillary length to detector

was 20.7 cm). The capillary was filled with sieving polymer matrix prior to each run. Replaceable sieving matrix was 1.8 g of polyethylene glycol (MW 35 000) dissolved in 7.2 ml of 0.1 *M* Tris–Boric acid, 1.25 m*M* EDTA buffer (0.1 *M* TBE) with 1 *M* urea. The 0.1 *M* TBE, 1 *M* urea buffer was prepared fresh prior to use by dissolving 1.815 g urea in 3 ml of 1 *M* TBE buffer, pH 8.3, and diluting the solution with deionized water up to a final volume of 30 ml. Separation was performed at 15 kV for 40 min. Separation temperature was 30 °C. Samples were injected into the capillary after drop dialysis on a 47-mm, 0.025 μm nitrocellulose membrane (Millipore Corporation, Bedford, MA, USA) by applying a potential of 5 kV for 3–5 s. Sample data was acquired and processed using 32 Karat™ software, version 4.0 (Beckman–Coulter, Fullerton, CA, USA).

2.4. HPLC mobile phase preparation

The 0.1 *M* TEAA buffer, pH 7, was prepared by mixing appropriate volumes of TEA and acetic acid in water. Because of the limited solubility of TEA in water, we used the following protocol (1 l of 0.1 *M* buffer): 5.6 ml of glacial acetic acid was placed in ~950 ml of water. While mixing, 13.86 ml of TEA were slowly added; the pH of the resulting solution was typically between 5 and 9. The pH was carefully adjusted to 7 by the addition of either TEA or acetic acid. Since the desirable pH 7 is more than 1 pH unit apart from the pK_a's of TEA (10.72) and acetic acid (4.71), the amount of acid or base needed to adjust the pH is very small. The volume was adjusted to 1 l with water bringing the final concentration of TEA and acetic acid to ~0.1 *M*.

N,N-Dimethylbutylamine acetate was used as the ion-pairing buffer for liquid chromatography–mass spectrometry (LC–MS) analysis. *N,N*-Dimethylbutylamine is a more effective ion-pairing agent than triethylamine [25], therefore, a lower concentration of buffer can be used for oligonucleotide separation, making the mobile phase more compatible with electrospray ionization-mass spectrometry (ESI-MS) detection [25]. The buffer (100 ml total volume) was prepared by first adding 351 μl of dimethylbutylamine to ~80 ml of deionized water (pH of this solution was between 10 and 12) and adjusting the pH to 7 with glacial acetic acid. The final buffer volume was adjusted with water to 100 ml, resulting in a concentration of 25 m*M* *N,N*-dimethylbutylamine.

2.5. LC–MS analysis

A capillary HPLC system (CapLC™, Waters Corporation, Milford, MA, USA) equipped with a photodiode array detector was connected to an electrospray time of flight (ESI-TOF) mass spectrometer (LCT™, Micromass, Manchester, UK) using a 30-cm fused-silica capillary, 50 μm I.D. A 50×1.0-mm XTerra™ MS C_{18} column packed with 2.5 μm particles (average pore diameter 140 Å) was used for LC–MS experiments. The built-in column heater of the CapLC™ system was set to 50 °C; mobile phase flow-rate was 23.6 μl per min. The LC–MS system was operated by MassLynx™ software, version 3.5 (Micromass, Manchester, UK). For gradient conditions see figure captions. LC–MS chromatograms were acquired in negative ion mode using an ESI-MS capillary voltage of 2.5 kV, a sample cone voltage of 20 V, an extraction cone voltage of 1 V, and an MCP detector voltage of 2700 V. The desolvation gas flow-rate was maintained at 410 l p.h. The cone gas flow-rate was set to 20 l p.h. The desolvation temperature and source temperature were set to 120 and 100 °C, respectively. The scanning range was 350–2500 m/z; the 1.1 s scanning cycle consisted of a 1.0-s scan and a 0.1-s interscan time. Typically, 5–27 individual spectra were acquired to generate a summed spectrum.

3. Results and discussion

3.1. Purification of singly-labeled oligonucleotides

In a previous paper, we described the method of IP-RP-HPLC analysis and purification for unlabeled oligonucleotides [26]. The purification of labeled oligonucleotides, described here, was found to be more challenging, partly because of the more complicated pattern of labeled and unlabeled failure products from synthesis [10]. In addition, little is known about the retention mechanism of dye-labeled oligonucleotides. One published study [21] suggests

that fluorescent dyes notably affect oligonucleotide retention due to the hydrophobicity of the labels. This was found to be the case with both single and double stranded DNA fragments [21].

Fig. 1 shows a complex chromatogram of the IP-RP-HPLC purification of a crude synthetic 25 mer oligonucleotide labeled with 5′ HEX dye. The chromatogram was obtained by loading a 0.1 μmole crude synthesis on a 50×4.6-mm column and monitoring the column output at two wavelengths. The 260 nm trace (Fig. 1A) represents the signal of DNA, while the 539 nm trace (Fig. 1B) records the signal of the HEX dye-labeled oligonucleotides (HEX maximum absorbance is ~539 nm). Comparing the traces, we conclude that the group of peaks eluting between 2 and 6 min (marked as D) are unlabeled failure sequences. Because of the missing dye, these fragments do not absorb at 539 nm. They are also retained less than dye-labeled oligonucleotides. Three later-eluting fractions (F, T and B) were collected; collection intervals are indicated on the chromatogram (dotted lines). The target oligonucleotide (fraction T) and the earlier-eluting impurities (F) are presumably labeled with the dye, since they absorb at both 260 and 539 nm. Interestingly, we observed a strong signal of later-eluting impurities (B) exhibiting absorbance at both wavelengths. The identity of those peaks was further investigated by LC–MS (see Section 3.2).

The purity of the target oligonucleotide (T in Fig. 1) was evaluated by CGE. The purity was found to be 93.5% based on peak area at 254 nm. The crude oligonucleotide and HPLC fraction B (Fig. 1) were also analyzed; electropherograms are shown in Fig. 2.

We used the same HPLC system and conditions to purify four other oligonucleotides labeled with different fluorescent dyes (3′-TAMRA, 5′-6FAM, 5′-Cy™3, and 5′-TET). Fraction collection was performed manually. The purity and yield results for all oligonucleotides are summarized in Table 1. The retention pattern of impurities for these other labeled oligonucleotides is similar to the one shown in Fig. 1; unlabeled failure fragments elute first, followed by

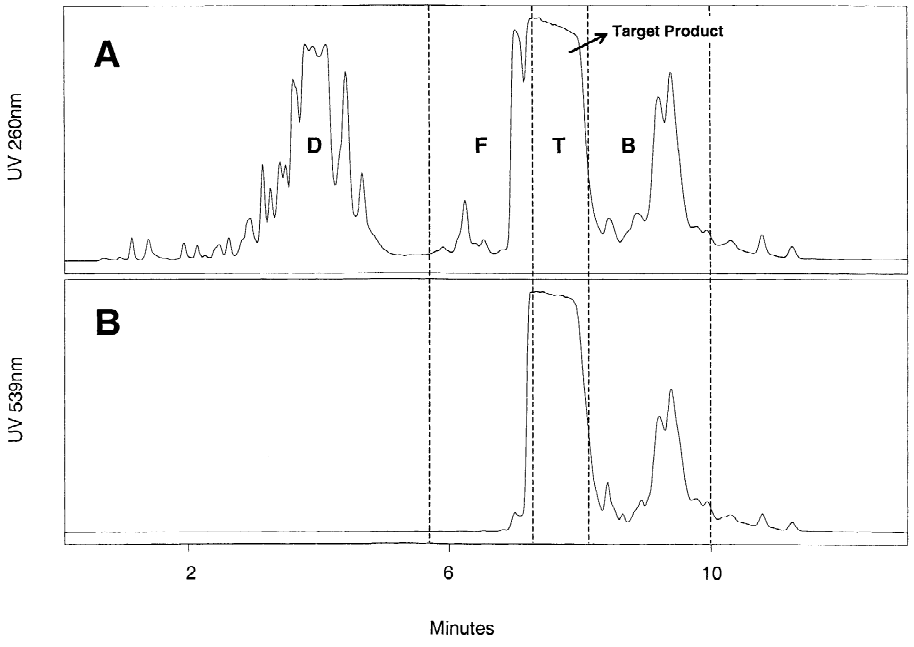

Fig. 1. Single-step, ion-pair RP-HPLC purification of a 0.1-μmole crude synthesis of a 25 mer mixed sequence oligonucleotide with a 5′ HEX fluorescent dye. Fractions were collected in indicated intervals (dotted lines), and monitored at 260 nm (A) and 539 nm (B). XTerra™ MS C_{18} 50×4.6-mm, 2.5 μm column. Mobile phase A: 5% ACN and 95% 0.1 *M* TEAA, pH 7; B: 30% ACN and 70% 0.1 *M* TEAA, pH 7. Gradient from 0% B (5% ACN) to 100% B (30% ACN) in 15 min, 1 ml/min, 60 °C.

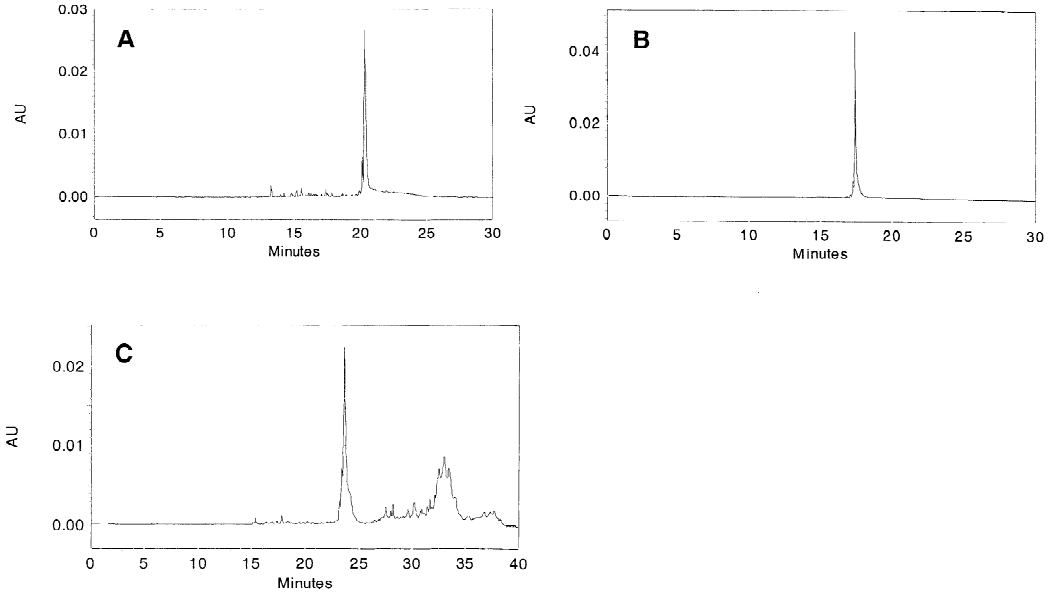

Fig. 2. CGE electropherograms of crude (A), IP-RP-HPLC purified (B), and later eluting contaminants fraction *B* (C) of the 25 mer mixed sequence oligonucleotide labeled with 5′ HEX shown in Fig. 1. Separation conditions: BioCAP™ 75 μm I.D. capillary, 20.7/31.3 cm; 0.1 *M* (1X) PEG sieving matrix in 0.1 *M* (1X) Tris–Boric acid–EDTA-2 Na^+ buffer; UV 254 nm; 15 kV run for 40 min; 3–5 s injection at 5 kV; separation temperature set at 30 °C.

(presumably) truncated dye-labeled oligonucleotides, full-length product, and later-eluting failure sequences.

Surprisingly, we detected a substantial amount of dye-labeled by-products from synthesis eluting after the target oligonucleotide. The nature of these failure products is not obvious. The HPLC and CGE analyses were not sufficient to identify these later-eluting peaks (labeled B in Fig. 1). Capillary electrophoresis analysis performed at a 550-nm detection wavelength revealed that this fraction contains a small amount of the target oligonucleotide, along with shorter dye-labeled fragments migrating before the full-length product (data not shown). However, the majority of dye-labeled failure sequences migrate through the sieving matrix slower than the full-length

Table 1
Purity and yield of labeled and unlabeled oligonucleotides[a]

Oligonuclotide	Sequence (5′–3′)	Yield @ 260 nm (%)	Yield @ dye absorbance (%)	CGE crude purity (%)	CGE purified purity (%)
25 mer no dye	GACTTAGACTTAGACTTAGACTTAG	91.68	–	85.4	91.5
25 mer 5′ TET	(TET)-TTTGACTTAGACTTAGACTTAGTTT	55.17	73.78	58.1	91.1
25 MER 5′ 6 fAM	(6FAM)-TTTGACTTAGACTTAGACTTAGTTT	31.08	45.82	52.3	97.7
25 mer 5′CY3	(CY3)-TTGACTTAGACTTAGACTTAGTTT	59.20	71.82	87.4	96.9
25 mer 5′HEX	(HEX)-TTTGACTTAGACTTAGACTTAGTTT	63.02	85.38	77.3	93.5
25 mer 3′TAMRA	TTTGACTTAGACTTAGACTTAGTTT-(TAMRA)	79.95	81.51	65.5	80.1[b]

[a] Purity was determined as an area percent of target product from CGE. Yield was calculated as an area percent at two different wavelengths; the maximum absorbance of DNA (260 nm) and the maximum UV absorbance of each dye. The latter takes into account labeled oligonucleotide sequences only.

[b] This value represents the main peak found by CGE. Some impurities resolved by CGE are also full-length products (Table 2), resulting in a higher purity than is shown. This separation is caused by isomers.

target product (see also Fig. 2C). This behavior was described in literature earlier [27,28]. Short oligonucleotides are separated according to their m/z ratio rather than by a sieving mechanism. Therefore, oligonucleotides shorter than ~6 mer migrate slower than longer ones in CGE. In addition, the interaction of the hydrophobic dye with the sieving polymer matrix was observed to have a substantial impact on the migration of shorter oligonucleotides. The nature of these later-migrating peaks was examined by LC–MS analysis.

3.2. LC–MS analysis of oligonucleotide impurities

We employed LC–MS to identify the components of crude synthetic oligonucleotides. The IP-RP-HPLC purified target oligonucleotides (T fraction, see Fig. 1) as well as later-eluting impurities (B fraction) were collected, concentrated by centrifugal evaporation, and injected into LC–MS. Table 2 shows the observed molecular masses of the purified oligonucleotides (for sequences see Table 1). Interestingly, MS analysis shows several major components with molecular masses close to the expected value. The accuracy of the mass spectrometer for molecular masses in the range of 8000–10 000 is expected to be ± 1 Da (molecular mass was deconvoluted by MaxEnt1™ software). The difference of <2 Da is due to a drift in calibration within a day (variations in lab temperature). However, in some cases, we observed a larger discrepancy between the observed molecular mass and the mass given by the manufacturer (Table 2). This indicates an unknown variation in chemical structure of the linker, or more likely the fluorescent dye itself. These oligonucleotides are presumably full-length product, since the loss of one nucleotide in the case of an $N-1$ failure oligonucleotide results in a mass difference of 313.2, 289.2, 329.2, or 304.2 for an A, C, G or T nucleotide, respectively. Fluorescent dyes used for labeling of oligonucleotides are known to contain isomers (with identical mass) as well as derivatives (amino group, succinimidyl linker, etc.). Due to the combined mass addition or loss of functional groups, it is difficult to identify the nature of modifications from the detected mass differences (~10, 42 and 52 Da). In the target fraction of the oligonucleotide labeled with 5′ TET, we even detected two approximately equimolar components differing in molecular mass by 42 Da (Table 2). Arguably, the possible source of heterogeneity is a partial chemical degradation of the dye during oligonucleotide synthesis (deprotection, purification) or incomplete post-synthesis removal of the protecting groups.

We analyzed later-eluting impurities collected for several different oligonucleotides in similar fashion as suggested in Fig. 1 (fraction B). This fraction from the purification of the 25 mer oligonucleotide labeled with 3′ TAMRA showed a particularly complex pattern of impurities. The LC–MS analysis is shown in Fig. 3. The mass spectra of peaks 1–10 (marked at peak apex on chromatogram) were analyzed, and the molecular mass values of these components are listed in Table 3. In summary, fraction B contains several pronounced peaks strongly absorbing at 556 nm (TAMRA UV maximum). Mass spectrometry suggests that peaks 1–6 are mainly very short oligonucleotides in the range of

Table 2
Mass analysis of HPLC-purified labeled oligonucleotides

Oligonucleotide	MW indicated by manufacturer[a]	Observed MW (major component)[b]	Other components[c]	Δ = observed–indicated MW
25 mer 5′ TET	8327.0	8336.7, 8378.6	8257.4, 8272.3, 8419.6, 8434.0	9.7, 51.6
25 mer 5′ 6FAM	8189.3	8191.0	8227.0, 8289.1	1.7
25 mer 5′ Cy3	8159.4	8201.5	8159.5, 8241.8	42.1
25 mer 5′ HEX	8395.9	8397.7	8318.7, 8438.1	1.8
25 mer 3′ TAMRA	8275.3	8276.0	8311.8	0.7

[a] Molecular weight is calculated from the theoretical additions of dye, linker and oligonucleotide sequence.
[b] All mass values were deconvoluted using the MaxEnt1™ option of MassLynx™ software, v. 3.5. The two components of similar signal intensity were found for 25 mer 5′ TET.
[c] Molecular weight of other components present in sample (5–20% of main component MS signal height).

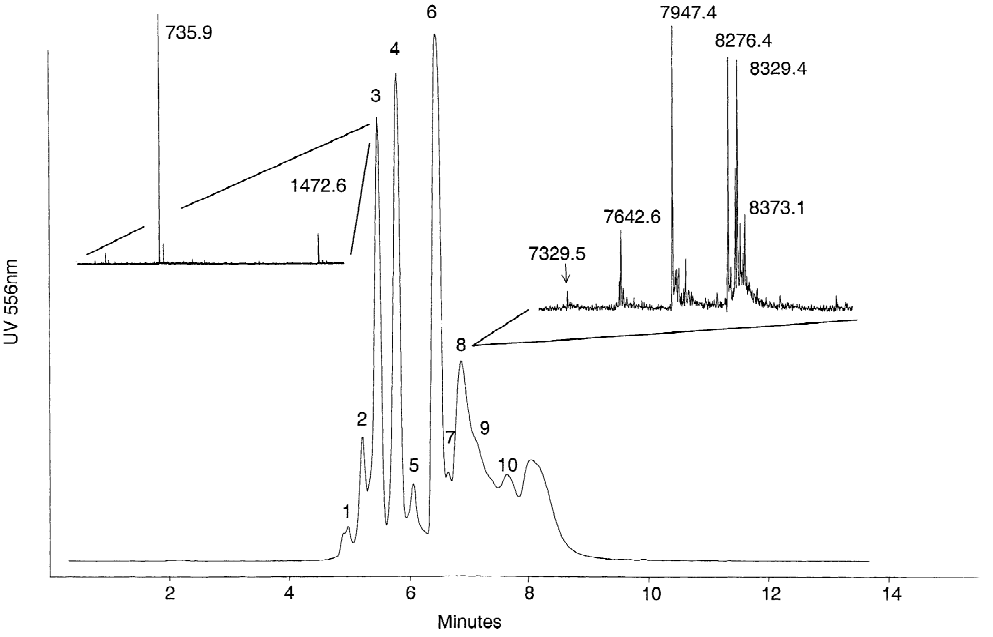

Fig. 3. Capillary LC–MS analysis of the tail fraction of a 25 mer oligonucleotide labeled at the 3′ end with TAMRA. Mass spectra insets are for peaks 3 and 8. The mass spectrum for peak 8 was deconvoluted from an original spectra in the 350–2500 m/z range. Mass values for other peaks can be found in Table 3. Tail fraction collected on a 0.1-μmole scale. UV absorbance at 556 nm. XTerra™ MS C_{18} 50×1.0 mm, 2.5 μm column. Mobile phase A: 25 m*M* dimethylbutylamine, pH 7; B: 100% ACN. Gradient from 15.5% B to 35.5% B in 12.5 min, 23.6 μl/min, 50 °C. For ESI-MS conditions, see Section 2.

2–6 mer. Peaks 3, 4 and 5 are longer failure oligonucleotides shorter than the main product with molecular masses between 3000 and 7400 Da. Analysis of peak 8 reveals a presence of 22, 23 and 24 mer dye-labeled oligonucleotides. Peaks 9 and 10 consist of multiple components, some with molecular masses close to the target oligonucleotide. These are presumably full-length oligonucleotides labeled with dye derivatives or oligonucleotides modified in the deprotection step (a difference of 53 Da indicates the nucleobase alkylation that occurs during removal of the cyanoethyl protection group) [29,30]. Apparently, these full-length oligonucleotides are more retained in an IP-RP-HPLC system than the target oligonucleotide. It is surprising that some shorter labeled failure sequences were present in the later-eluting fraction B. The possible retention mechanism will be discussed in the Section 3.4.

From Figs. 1–3, we can conclude that IP-RP-HPLC is capable of the efficient separation and purification of dye-labeled oligonucleotides. The results indicate that it is important to avoid collection of the front and ''tail'' fractions (F and B in Fig. 1) of the target product, as these are the major source of dye-labeled failure sequences. The fact that mass spectrometry accuracy is sufficient to distinguish failure sequences as well as indicate the dye derivatives is important for the characterization of labeled oligonucleotides. Since the dye structure has a crucial effect on its fluorescence, and therefore on the quality of a bioassay, LC–MS is a useful method for the quality control of diagnostic probes.

3.3. Purification of dually-labeled oligonucleotides

Dually-labeled oligonucleotides are typically synthesized in a two-step process. A singly-labeled oligonucleotide with a linker is purified by HPLC before conjugation with a second dye. The final product is HPLC-purified in order to remove excess dye reagent and remaining failure sequences. This process is time-consuming and adds to the cost of oligonucleotide probes. The simpler approach is to perform synthesis in one step. However, it results in

Table 3
Mass analysis of later-eluting contaminants (refer to Fig. 3)[a]

Peak	MW	Oligo length	Description
1	596.3, 728.9, 748.3	–	dye, dye derivative, or dye+linker
1	1458.5	<5 mer	dye-labeled failure sequence
2	561.4, 576.8, 740.4, 892.4	–	dye, dye derivative, or dye+linker
2	1154.6	<4 mer	dye-labeled failure sequence
3	2950.1, 4424.1, 58 967.7, 7371.7	<23 mers	dye-labeled failure sequence
4	2927.9, 3511.9, 4095.9, 4683.5, 5855.6, 7020.6	<22 mers	dye labeled failure sequence
4	8201.8	25 mer	Δ=74.5 from main product, dye derivative?
5	4249.7, 5251.2	<17 mer	dye-labeled failure sequences
6	654.4, 864.5, 962.5,	–	dye, dye derivative, or dye+linker
6	1519.8, 1729.6	<6 mer	dye-labeled failure sequence
7	3967.9, 607.9, 79342	<24 mers	dye labeled failure sequence
7	8315.0	25 mer	Δ=+38.6 from main product, dye deriative?
8	7329.6	22 mer	Δ=loss adenine (313.0)
8	7642.6	23 mer	Δ=loss of thymine (304.8)
8	7947.4	24 mer	Δ=loss of guanine
8	8276.4	25 mer-full length product	–
8	8318.0	25 mer	Δ=+41.6 from main product, dye derivative?
8	8329.3	25 mer	Δ=+52.9 from main product (cyanoethyl group)
8	8373.1	25 mer	Δ=+98.2 from main product, dye derivative?
9	7946.6	24 mer	Δ=loss of guaine(329.8)
9	8329.6	25 mer	Δ=+53.2 from main product (cyanoethyl group)
9	8374.6	25 mer	Δ=+98.2 from main product, dye derivative?
10	5986.6, 6100.1, 6404.0, 7046.8, 7689.1	<23 mers	dye-labeled failure sequence
10	8329.0	25 mer	Δ=+52.6 from main product (cyanoethyl group)
10	8375.3	25 mer	Δ=+98.9 from main product, dye derivative?

[a] Mass values greater than 2500 Da were deconvoluted using the MaxEnt1™ option of MassLynx™ software, version 3.5.

a complicated mixture consisting of the target oligonucleotide, oligonucleotide fragments labeled with one of the two dyes, unlabeled failure fragments, and dyes only.

We evaluated the use of IP-RP-HPLC for the purification of these ''one-pot'' synthetic oligonucleotides. Fig. 4 shows the purification of a 36 mer TaqMan™ probe. The Taqman™ oligonucleotide was labeled with fluorescein at the 5′ end and with TAMRA at the 3′ end. Due to the complex pattern of failure products from crude synthesis, the choice of detection wavelength becomes even more crucial than for singly-labeled oligonucleotides. Typically, the separation is simultaneously monitored at 260 nm (DNA absorbance), and at the absorbance maxima of the two dyes; in this case 500 nm (fluorescein absorbance) and 556 nm (TAMRA absorbance). However, due to the presence of numerous failure sequence peaks and the partial spectral overlap of TAMRA and fluorescein dyes, it may be difficult to correctly identify the target fraction. Therefore, we used the detection wavelength of 260 nm in combination with real-time monitoring of the full UV–Vis spectra for the entire duration of analysis (Fig. 4). The target product was collected manually when the characteristic spectra of all three UV–Vis maxima (260, 500 and 556 nm) were detected.

As indicated by the UV–Vis spectra (Fig. 4), the first-eluting peaks are unlabeled oligonucleotides, followed by fluorescein-labeled failure fragments, TAMRA-labeled failure fragments, and the target product. The collection time window is indicated on the chromatogram. Later-eluting peaks are presumably short failure sequences labeled with TAMRA or unconjugated dye. We employed LC–MS for a more detailed characterization of impurities. Mass spectrometry analysis confirmed the retention order of peaks obtained by PDA. These results were published elsewhere [10].

We purified two other Taqman™ probes (21 mer and 28 mer) using identical HPLC conditions (Fig. 5). All three crude mixtures show a similar pattern of

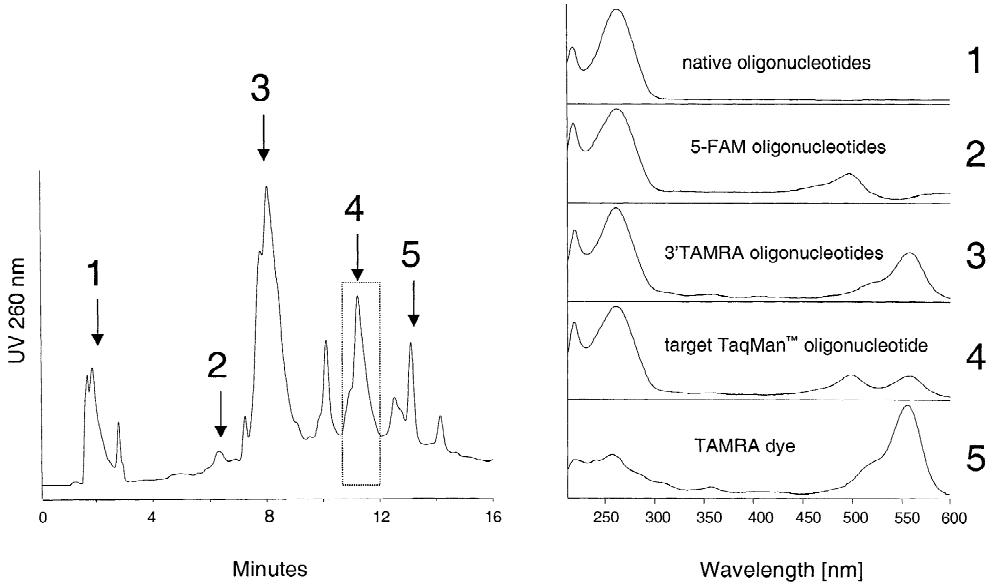

Fig. 4. IP-RP-HPLC separation of a crude synthetic 36 mer Taqman™ oligonucleotide using on-line UV–Vis spectrum monitoring. Individual peak spectra indicated at the peak apex are found to the right of the chromatogram. XTerra™ MS C_{18} 50×4.6-mm, 2.5 μm column. Mobile phase A: 5% ACN and 95% 0.1 *M* TEAA, pH 7; B: 40% ACN and 60% 0.1 *M* TEAA, pH 7. Gradient from 17.1% B (11% ACN) to 60% B (26% ACN) in 30 min, 0.5 ml/min, 60 °C.

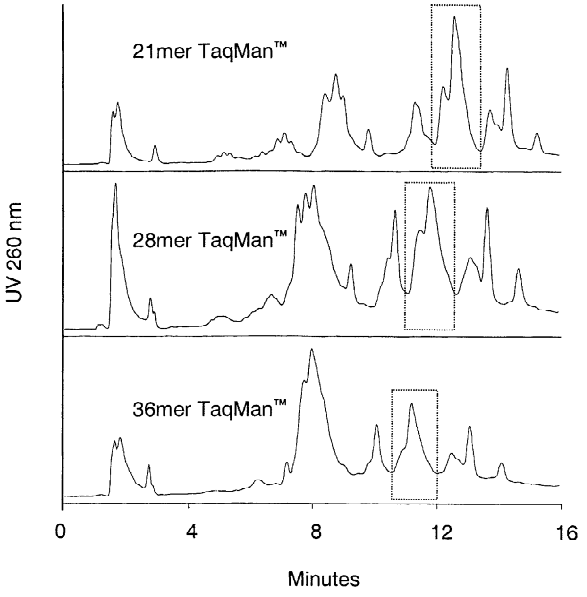

Fig. 5. Overlay of single step, IP-RP-HPLC purification of three Taqman™ oligonucleotides. XTerra™ MS C_{18} 50×4.6 mm, 2.5 μm column. Mobile phase A: 5% ACN and 95% 0.1 *M* TEAA, pH 7; B: 40% ACN and 60% 0.1 *M* TEAA, pH 7. Gradient from 17.1% B (11% ACN) to 60% B (26% ACN) in 30 min, 0.5 ml/min, 60 °C.

impurities, which makes purification routine. The target fractions were collected as suggested in Fig. 5. One may notice that the target fractions appear to be a group of several coeluting peaks. This partial separation of full-length oligonucleotide probes is caused by differences in retention of fluorescent dye isomers. The LC–MS analysis confirmed the presence of only a single molecular mass in the isolated target fraction. Purity of the 36 mer Taqman™ probe was evaluated by CGE (Fig. 6.). The purity values are given on the electropherogram.

3.4. Impact of fluorescent label on oligonucleotide retention and purification

As mentioned in Section 3.1, the retention of unlabeled oligonucleotides is primarily based on an ion-pairing mechanism, with shorter oligonucleotides eluting before longer ones [17,26]. However, the hydrophobicity (base composition) of oligonucleotides was also shown to have an impact on their retention [31,32]. For example, oligonucleotides of the same length with different base composition can be separated in an IP-RP-HPLC system [33].

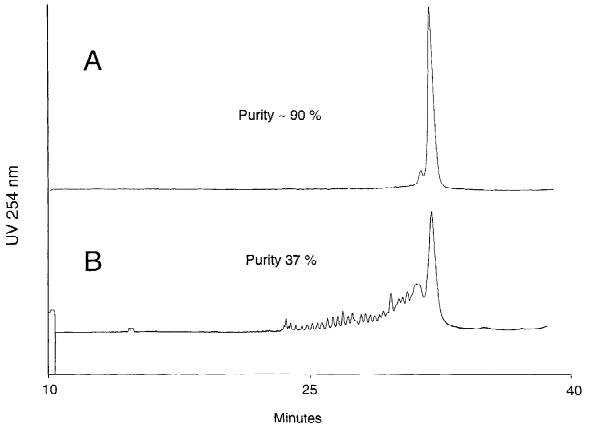

Fig. 6. CGE electropherograms of IP-RP-HPLC purified (A) and crude synthesis (B) of the 36 mer Taqman™ probe shown in Figs. 4 and 5. Purity was based on peak area at 254 nm. Separation conditions: BioCAP™ 75 μm I.D. capillary, 27.5/34.5 cm; PEG sieving matrix with 15% IPA in 0.2 *M* Tris–Boric acid buffer, pH 8.3; UV 254 nm; 15 kV run for 40 min; 5 s injection at 5 kV; capillary cartridge set at 30 °C. Notice that the purity of the 36 mer Taqman™ probe is ~90%. The small peak eluting just before is the remaining 10% of material. LC–MS analysis of this purified oligonucleotide shows that these two peaks are the same molecular mass, indicating that they are isomers, and actual purity is even greater than 90%.

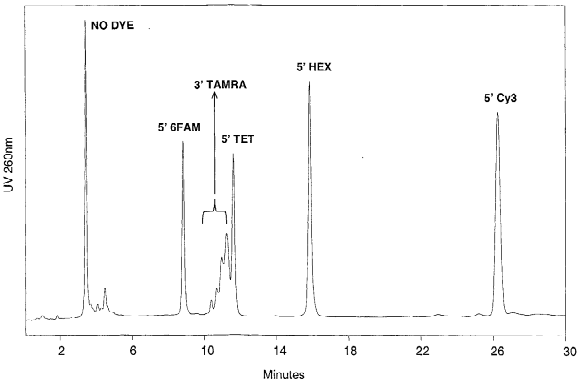

Fig. 7. Ion-pair RP-HPLC separation of 6 oligonucleotides, five with a fluorescent tag and one without, purified by IP-RP-HPLC in a single step. The fluorescent dye is indicated at the peak apex on the chromatogram. UV absorbance at 260 nm. XTerra™ MS C_{18} 50×4.6-mm, 2.5 μm column. Mobile phase A: 5% ACN and 95% 0.1 *M* TEAA, pH 7; B: 30% ACN and 70% 0.1 *M* TEAA, pH 7. Gradient from 12% B (8% ACN) to 92% B (28% ACN) in 40 min, 1 ml/min, 60 °C.

The retention mechanism of fluorescent oligonucleotides was studied to a much lesser degree. As discussed in Section 3.1, by purifying labeled oligonucleotides, we observed a significant amount of failure products eluting after the full-length target oligonucleotide. LC–MS analysis (Section 3.2, 3.3, and Table 3) suggests that some of these fragments are shorter ($N-1$, $N-2$, etc.) failure sequences. Apparently, at least some of the truncated by-products of synthesis are more retained than the full-length dye-labeled oligonucleotide.

To investigate the influence of the dye on oligonucleotide retention, we prepared a mixture of five 25 mer oligonucleotides (of identical sequence) labeled with different fluorescent dyes. An unlabeled 25 mer oligonucleotide was also included. The chromatogram of this mixture is shown in Fig. 7. We found that the label has a pronounced effect on retention. The retention of the labeled 25 mer oligonucleotides increase in order: no dye<5′-6FAM<TAMRA-3′<5′-TET<5′-HEX<5′-Cy™3. The retention is dictated by the nature of fluorescent dye. This finding is consistent with a previous report [21]. Similar to Fig. 7, Fig. 4 shows the impact of dye hydrophobicity on the retention of dually-labeled oligonucleotides. The retention order increases as follows: unlabeled oligonucleotides<5′-fluorescein labeled oligonucleotide(s)<3′-TAMRA labeled oligonucleotide(s)<dually-labeled Taqman™ probe.

It should be noted in Fig. 7 that the 25 mer TAMRA-labeled oligonucleotide elutes in several peaks. These peaks are all full-length products, as determined by LC–MS (Table 2). The partial HPLC separation is due to differences in the retention of the target oligonucleotide labeled with positional isomers of TAMRA dye. Different synthetic lots of this 25 mer labeled with 3′ TAMRA supplied by the manufacturer showed a different ratio of peaks, indicating that either the synthesis or post-synthesis treatment (TAMRA is labile at high pH) generates a heterogeneous product.

We observed similar retention behavior for 25 mer oligonucleotides labeled with fluorescein. Six peaks were separated by IP-RP-HPLC and analyzed by mass spectrometry (data not shown). Some of the peaks have identical molecular mass (positional or optical isomers). Others differ in mass by 15 Da (amino derivative of the dye). This is the reason why

we used an oligonucleotide labeled with a single isomer of fluorescein dye (6-FAM) in this study.

Mass spectrometry analysis of the purified oligonucleotides (Table 2) detected contaminants differing in molecular mass from the target product by 35.8, 40.4, 42.1 and 98.1 Da. These components are either 25 mer oligonucleotides labeled with dye derivatives, or incompletely deprotected oligonucleotides. Interestingly, IP-RP-HPLC does not separate these derivatives from the main component under the conditions used in Fig. 7.

The presence of shorter failure sequences in the later-eluting fraction (B in Fig. 1) suggests a somewhat surprising conclusion that at least some of the truncated oligonucleotides are retained more than the full-length product. We propose two possible explanations for this retention behavior of labeled oligonucleotides. First, the impact of dye derivatives on retention is stronger than oligonucleotide length. That means that $N - x$ failure fragments labeled with dye may be retained more strongly than a 25 mer labeled with a dye isomer. The results from CGE, HPLC and LC–MS analyses suggest the presence of isomers (and derivatives), not only in TAMRA and fluorescein, but also in other dye-labeled oligonucleotides (Table 2).

Second, short labeled oligonucleotides are retained by a hydrophobic interaction of the dye with the reversed-phase sorbent rather than by an ion-pairing mechanism. As a result, the 2–5 mer oligonucleotides labeled with dye exhibit significantly stronger retention than the full-length product (Table 3). We used 2, 4 and 8 mer oligonucleotides labeled with the dimethoxytrityl group (DMT) to demonstrate this retention mechanism (Fig. 8). While unlabeled oligonucleotides follow the expected order of increasing retention (2 mer$<$4 mer$<$8 mer), the presence of the highly hydrophobic DMT group apparently reverses the increasing retention order of the oligonucleotides to 8 mer$<$4 mer$<$2 mer. The reversed retention order of labeled fragments can be explained by the concept of the average molecule hydrophobicity. The most strongly retained is the trityl label only, and the addition of oligonucleotide(s) to the trityl group in fact decreases the molecule's hydrophobicity. In other words, the average hydrophobicity (retention) decreases with the length of the oligonucleotide attached to the trityl group.

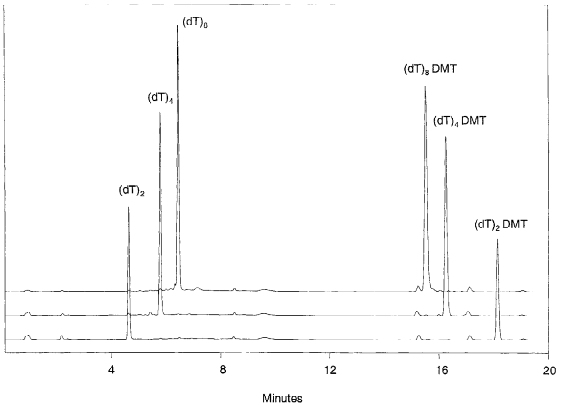

Fig. 8. IP-RP-HPLC separation of oligonucleotides with and without the dimethoxytrityl (DMT) protecting group attached. Oligonucleotide length and DMT label (if applicable) are indicated at peak apex on chromatogram. XTerra™ MS C_{18} 50×4.6-mm, 2.5 μm column. Mobile phase A: 0.1 *M* TEAA, pH 7; B: 100% ACN. Gradient from 0% B to 50% B in 25 min; 1.0 ml/min.; 60 °C.

It is interesting to note that the retention order of dye-labeled oligonucleotides is not completely reversed. Some of the dye-labeled failure sequences elute before the target product (Fig. 1, Fig. 2). We studied the retention behavior of a 10–15 mer oligodeoxythymidine ladder labeled at the 5′ end with 6FAM dye (Fig. 9). Interestingly, we did not

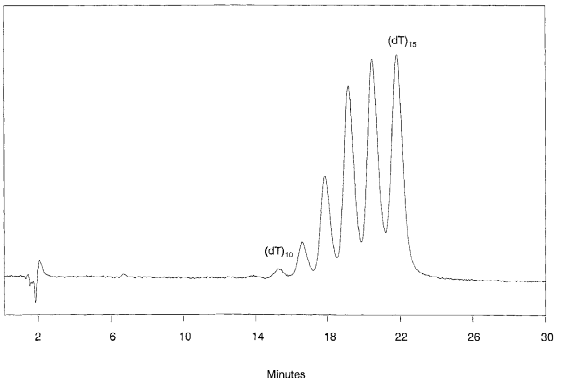

Fig. 9. IP-RP-HPLC separation of a purified 5′ 6FAM labeled oligodeoxythymidine digested with 3′ exonuclease (phosphodiesterase I). UV absorbance at 500 nm. XTerra™ MS C_{18} 50×4.6-μμ, 2.5 μm column. Mobile phase A: 5% ACN and 95% 0.1 *M* TEAA, pH 7; B: 30% ACN and 70% 0.1 *M* TEAA, pH 7. Gradient from 20% B (10% ACN) to 52% B (18% ACN) in 160 min; 0.5 ml/min.; 60 °C.

detect a switch in the retention order of these dye-labeled fragments. The retention time increased with oligonucleotide length, mimicking the retention behavior of unlabeled oligonucleotides. Presumably, the switch in retention order of labeled oligonucleotides depends on their length as well as on the hydrophobicity of the labeling group (the 6FAM label is less hydrophobic than the trityl moiety).

4. Conclusion

An ion-pair reversed-phase HPLC method was developed for the purification of diagnostic oligonucleotide probes. Singly and dually dye-labeled oligonucleotides were purified on a 50×4.6-mm column packed with porous, 2.5 μm C_{18} sorbent. Purification of a 0.1 μmole synthesis was routinely performed in a single injection. A new method was developed for the isolation of dually-labeled oligonucleotides synthesized in the ''one-pot'' approach, which offers substantial advantages over current two-step synthesis/two-step purification processes. We employed IP-RP-HPLC with mass spectrometry detection for the characterization of the target oligonucleotide as well as failure sequences. The dye-labeled failure sequences were found to elute both before and after the target product. Therefore, we investigated the mechanism of dye-labeled oligonucleotide retention in order to maximize purity of the full-length target product. Since we observed a highly variable quality of oligonucleotides supplied by several manufacturers, the IP-RP-HPLC method was successfully utilized for the quality control of synthetic oligonucleotides.

References

[1] R.T. Pon, G.A. Buck, K.M. Hager, C.W. Naeve, R.L. Niece, M. Robertson, A.J. Smith, Biotechniques 21 (1996) 680.
[2] K. Kleparnik, F. Foret, J. Berka, W. Goetzinger, A.W. Miller, B.L. Karger, Electrophoresis 17 (1996) 1860.
[3] S.J. Laken, P.E. Jackson, K.W. Kinzler, B. Vogelstein, P.T. Strickland, J.D. Groopman, M.D. Friesen, Nat. Biotechnol. 16 (1998) 1352.
[4] I. Tapp, L. Malmberg, E. Rennel, M. Wik, A.C. Syvanen, Biotechniques 28 (2000) 732.
[5] C.A. Heid, J. Stevens, K.J. Livak, P.M. Williams, Genome Res. 6 (1996) 986.
[6] U.E. Gibson, C.A. Heid, P.M. Williams, Genome Res. 6 (1996) 995.
[7] S.A. Marras, F.R. Kramer, S. Tyagi, Genet. Anal. 14 (1999) 151.
[8] L.G. Lee, K.J. Livak, B. Mullah, R.J. Graham, R.S. Vinayak, T.M. Woudenberg, Biotechniques 27 (1999) 342.
[9] M. Gilar, E.S.P. Bouvier, J. Chromatogr. A 890 (2000) 167.
[10] M. Gilar, Anal. Biochem. 298 (2001) 196.
[11] J. Temsamani, M. Kubert, S. Agrawal, Nucleic Acids Res. 23 (1995) 1841.
[12] W.A. Rudert, E.R. Braun, S.J. Faas, R. Menon, A. Jaquins-Gerstl, M. Trucco, Biotechniques 22 (1997) 1140.
[13] K.L. Fearon, J.T. Stults, B.J. Bergot, L.M. Christensen, A.M. Raible, Nucleic Acids Res. 23 (1995) 2754.
[14] W.J. Warren, G. Vella, Mol. Biotechnol. 4 (1995) 179.
[15] T.L. Hill, J.W. Mayhew, J. Chromatogr. 512 (1990) 415.
[16] F.W. Hobbs, J.A. Yarem, Biotechniques 14 (1993) 584.
[17] C.G. Huber, P.J. Oefner, G.K. Bonn, Anal. Biochem. 212 (1993) 351.
[18] J.B. Crowther, R. Jones, R.A. Hartwick, J. Chromatogr. 217 (1981) 479.
[19] C.H. Becker, J.W. Efcavitch, C.R. Heiner, N.F. Kaiser, J. Chromatogr. 326 (1985) 293.
[20] J.B. Murray, A.K. Collier, J.R. Arnold, Anal. Biochem. 218 (1994) 177.
[21] P.J. Oefner, C.G. Huber, F. Umlauft, G.N. Berti, E. Stimpf, G.K. Bonn, Anal. Biochem. 223 (1994) 39.
[22] A. Apffel, J.A. Chakel, S. Fischer, K. Lichtenwalter, W.S. Hancock, Anal. Chem. 69 (1997) 1320.
[23] C.G. Huber, A. Krajete, Anal. Chem. 71 (1999) 3730.
[24] H. Oberacher, B. Wellenzohn, C.G. Huber, Anal. Chem. 74 (2002) 211.
[25] H. Oberacher, W. Parson, R. Muhlmann, C.G. Huber, Anal. Chem. 73 (2001) 5109.
[26] M. Gilar, K.J. Fountain, U.D. Neue, Y. Budman, K.R. Yardley, D. Rainville, R.J. Russell II, J.C. Gebler, J. Chromatogr. A 958 (2002) 167.
[27] L.A. DeDionisio, D.H. Lloyd, J. Chromatogr. A 735 (1996) 191.
[28] M. Gilar, A. Belenky, A.S. Cohen, Electrophoresis 21 (2000) 2999.
[29] A.P. Guzaev, M. Manoharan, J. Am. Chem. Soc. 123 (2001) 783.
[30] A. Wilk, A. Grajkowski, R.L. Phillips, S.L. Beaucage, J. Org. Chem. 64 (1999) 7515.
[31] C.G. Huber, E. Stimpf, P.J. Oefner, G.K. Bonn, LC–GC 14 (1996) 114.
[32] R. Haefele, D. Gjerde, Transgenomics Application note 103.
[33] K. Makino, H. Ozaki, H. Imaishi, T. Takeuchi, T. Fukui, Chem. Lett. (1987) 1251.