

Analysis and Purification of Synthetic Oligonucleotides by Reversed-Phase High-Performance Liquid Chromatography with Photodiode Array and Mass Spectrometry Detection

Martin Gilar

Waters Corporation, Milford, Massachusetts 01757

Received March 26, 2001; published online October 10, 2001

Native and modified synthetic oligonucleotides were purified by reversed-phase HPLC using volatile ion-pairing mobile phases. Purification of 10–90 nmol of oligonucleotides in a single injection was demonstrated using a 4.6 × 75-mm HPLC column packed with porous 2.5 μm C18 sorbent. Separation of target products from N-1 failure fragments was achieved for oligonucleotides in the 4- to 60-mer size range. We employed a combination of absorbance and mass spectrometry detection to identify by-products of oligonucleotide synthesis. This method was also employed for analysis and purification of fluorescently labeled oligonucleotides. © 2001 Academic Press

Key Words: oligonucleotide purification; ion-pair reversed-phase HPLC; mass spectrometry; fluorescent primers; genotyping.

The annual production of synthetic oligonucleotides has increased dramatically in recent years in order to meet the demand of the biotechnology market. This demand has been propelled by the development of two major technologies: (i) high-throughput DNA sequencing (1, 2) and (ii) DNA genotyping (3–6). Sample preparation for DNA analysis is almost exclusively based on the polymerase chain reaction (PCR).¹ It is usually performed in two steps. First, the interrogated region of genomic DNA is selectively amplified and then the amplicon is used for subsequent generation of a genotyping sample or DNA sequencing ladder. The quality

of the PCR primers used in both steps determines the success of sequencing as well as genotyping accuracy.

Oligonucleotide synthesis typically yields the target oligonucleotide contaminated with truncated by-products, so-called failure, and mismatch sequence oligonucleotides (7, 8). Using low-quality primers for PCR may result in a failure to amplify a selected sequence, as well as in generation of undesirable nonspecific amplification products. Frequent PCR failure presents a serious problem for the logistics of sample reprocessing and sample history record keeping, especially when working in the high-throughput sample processing mode.

It is known that sequencing performance suffers when poor quality primers are used for generation of a DNA sequencing ladder (9, 10). However, the decrease in sequencing accuracy does not directly correlate with the purity of the primers. This is especially true for primers tagged at their 5' end with a fluorescent label, because the prematurely halted failure products of synthesis are not tagged with the fluorescent group. Therefore, these impurities do not interfere with base calling even though they are present in a sequencing mixture.

Genotyping accuracy is one of the important criteria for choosing a genotyping method. Generally, better than 99% accuracy is required. Reliability of many genotyping techniques, however, strongly depends on the purity of the oligonucleotide probes. Low-quality oligonucleotides may have a detrimental impact on the performance of genotyping assays based on fluorescent oligonucleotide probes such as Invader (11, 12), Taq-Man (13–15), and molecular beacons (13, 16, 17). Background fluorescence from failure sequence oligonucleotide probes contributes to false positive or false negative genotype assignments. One of the main advantages of mass spectrometry genotyping techniques is that genotype call accuracy is in principle not af-

¹ Abbreviation used: SPE, solid-phase extraction; RP, reversed-phase; IP, ion-pair; DMT, dimethoxytrityl; HFIP, hexafluoroisopropanol; TEA, triethylamine; TEAA, triethylammonium acetate; FLSN, fluorescein; TAMRA, carboxytetramethylrhodamine; PCR, polymerase chain reaction; ESI, electrospray ionization; PDA, photodiode array.

ected by primer purity. However, frequent PCR failures compromise the efficiency of high-throughput MS data collection (3, 18).

Purification of synthetic oligonucleotides is a straightforward approach to improve the reliability of genotyping techniques. Three techniques traditionally used for DNA separation are (i) polyacrylamide gel electrophoresis, (ii) chromatography, and (iii) solid-phase extraction (SPE). Although the separation ability of highly cross-linked polyacrylamide gel is superior to that of chromatographic techniques, it is difficult to scale up the oligonucleotide mass load without a loss of resolution. Other problems of slab gel electrophoresis are visualization of separated bands, excision from the gel, isolation of DNA from the gel strip, and desalting of the sample contaminated with electrophoretic buffers. Moreover, slab gel electrophoresis is difficult to automate.

Both reversed-phase (RP) and anion-exchange chromatography are commonly used for oligonucleotide purifications. HPLC permits separation of up to approximately 20- to 30-mer oligonucleotides, but the separation selectivity decreases with the oligonucleotide length. In addition, slow diffusion of high-molecular-weight analytes in the sorbent pores causes peak broadening, which results in low separation efficiency. Nonporous sorbents have been shown to be efficient for HPLC separations of polynucleotides and dsDNA (19–21), due to enhanced stationary phase mass transfer. However, the mass capacity of nonporous sorbents is low, which limits their usefulness for preparative purifications.

Nonporous ion-exchange HPLC sorbents have been successfully used for synthetic oligonucleotide cleanup (22, 23). The disadvantage of ion-exchange HPLC is the contamination of oligonucleotides with inorganic salts. An additional desalting step is required, often employing RP-HPLC.

Reversed-phase HPLC is commonly utilized for the so-called “dimethoxytrityl (DMT) on” oligonucleotide purification method (7). The method is based on the hydrophobic properties of the DMT protection moiety used in DNA synthesis. After the last step of synthesis the cleavage of the DMT group from the full-length product is deliberately omitted. As a result, the target oligonucleotide bearing the trityl group is strongly retained on RP-HPLC sorbents and therefore separated from “DMT off” failure products. SPE purification of oligonucleotides uses the same “DMT on” purification principles. “DMT on” SPE offers the advantage of fast and parallel purification. However, due to the lower separation efficiency, it yields lower purity and oligonucleotide recovery than RP-HPLC.

Although simple and popular, the “DMT on” method has several drawbacks. It is known that synthesis often yields failure “mismatch sequence” oligonucleo-

tides, with one or more nucleotides missing within the chosen DNA sequence (8). Because these mismatch oligonucleotides are elongated to the last step of synthesis (last nucleotide of a mismatch oligonucleotide is identical with the full-length product), they have the DMT group attached to the 5' end (7). As a result, these mismatch failure fragments cannot be easily separated from target oligonucleotides using the “DMT on” method. In some cases the content of “DMT on” impurities may exceed 10% or more, which has a detrimental effect on the “DMT on” method performance.

An alternative to the “DMT on” RP-HPLC method is the ion-pair reversed-phase HPLC (IP-RP-HPLC) separation of detritylated oligonucleotides. IP-RP-HPLC was proposed for purification of synthetic oligonucleotides using volatile buffers typically containing triethylamine (TEA) as the ion-pairing agent (24, 25). The usefulness of ion-pair RP-HPLC was limited to purification of rather short 10- to 26-mer oligonucleotides (24). Better separation performance can be obtained using nonporous polymeric or silica-based sorbents, pellicular sorbents (26), or monolithic columns (27, 28). The drawback of IP-RP-HPLC, however, is that retention of oligonucleotides is driven not only by charge-to-charge interaction of phosphate internucleotide linkages with the ion-pairing agent adsorbed on the stationary phase, but also by hydrophobic interactions of the nucleobases with the reversed-phase sorbent (23, 27). It was found that oligodeoxythymidines are more retained than oligodeoxyadenosines and oligodeoxycytidines, respectively (21). This behavior complicates routine oligonucleotide purification. More efficient ion-pairing agents were utilized to eliminate sequence dependent retention of oligonucleotides. Tetraalkylammonium cations (e.g., tetrabutylammonium ions) are more strongly adsorbed on RP surfaces than triethylammonium ions, which results in stronger retention and predominantly charge-dependent oligonucleotide separations (29). However, the tetraalkylammonium ion-pairing agents are not volatile and cannot be removed from purified oligonucleotides by lyophilization.

Recently, LC-MS methods were developed for oligonucleotide separation and identification using aqueous triethylammonium–hexafluoroisopropylalcohol (TEA–HFIP) buffers compatible with MS detection (30). However, the separation did not fully match the separation achieved with the triethylammonium acetate buffer (31). Alternatively, a triethylammonium bicarbonate mobile phase was used for oligonucleotide separation with postcolumn acetonitrile addition to the eluent. The ion-pairing buffer was chosen to give the best MS detection sensitivity, but the separation performance was not optimal (32).

In this paper we describe a method for analytical and semipreparative purification of synthetic oligonucleotides using RP-HPLC columns packed with a 2.5- μ m

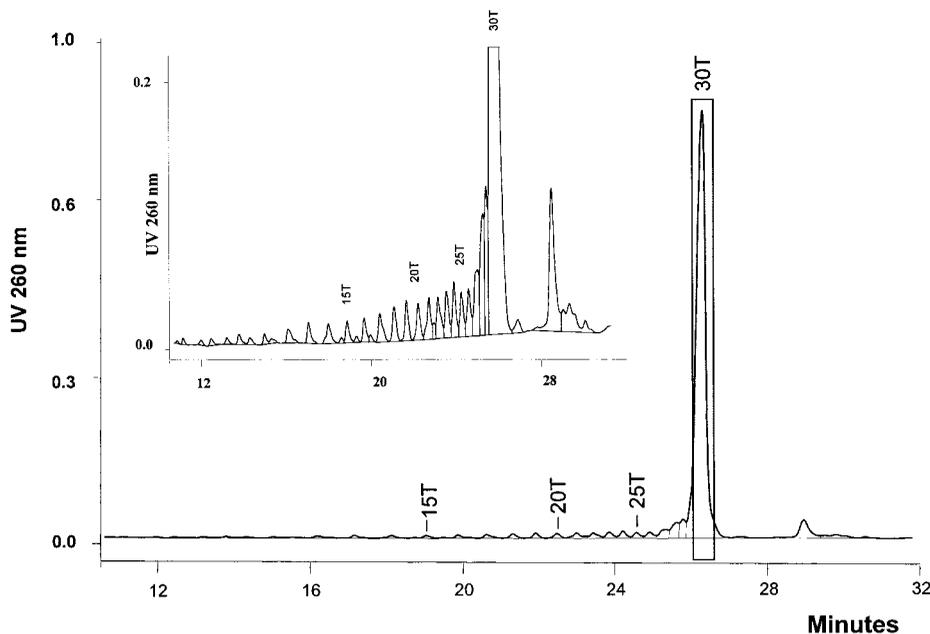


FIG. 1. Ion-pair RP HPLC purification of 1 nmol of crude synthetic 30-mer oligodeoxythymidine. Target fraction was collected in indicated interval. XTerra MS C18 75×4.6 mm, $2.5 \mu\text{m}$. Mobile phase A, 0.1 M TEAA, pH 7; B, ACN. Gradient from 10 to 14.5% B in 36 min, 0.5 ml/min, 51.5°C . (Inset) The enlarged part of the same chromatogram; 30-mer oligodeoxythymidine purity measured as a peak area at 260 nm was 78%.

RESULTS AND DISCUSSION

IP-RP-HPLC Analysis and Purification of Oligonucleotides

As reported earlier (7), we have studied the performance of several RP-HPLC columns for the separation of oligonucleotides. A 75×4.6 -mm column packed with $2.5\text{-}\mu\text{m}$ XTerra MS C18 sorbent was found to be most suitable for separating an 8- to 30-mer oligodeoxythymidine mixture. This column and separation conditions were in this work employed for the purification of crude 30-mer oligodeoxythymidine. Crude oligonucleotide synthesized at the $0.2\text{-}\mu\text{mol}$ scale was obtained from a vendor and dissolved in $100 \mu\text{l}$ of 0.1 M triethylammonium acetate buffer, pH 7. The chromatogram shown in Fig. 1 represents the purification of 2 nmol ($1\text{-}\mu\text{l}$ injection) of oligonucleotide. The target fraction was manually collected within the window indicated in Fig. 1. Separation of the 30-mer full-length product from the N-1 failure sequence oligonucleotide was achieved. The crude 30-mer purity estimated by HPLC analysis was 78% (Fig. 1); the collected fraction purity was better than 99%.

Oligonucleotide synthesis is usually performed on the 0.05- to $1\text{-}\mu\text{mol}$ synthesis scale (the most common scale is probably $0.2 \mu\text{mol}$). Therefore, it is desirable to design an HPLC method suitable for purifying relatively large oligonucleotide mass loads. We studied the influence of mass load on the quality of the oligonucleotide purification performance using the $2.5\text{-}\mu\text{m}$

XTerra MS C18 analytical column, 75×4.6 mm. Figure 2 shows three chromatograms resulting from injection of 0.01, 0.03, and $0.09 \mu\text{mol}$ of 30-mer oligodeoxythymidine. The peak of the 30-mer product noticeably

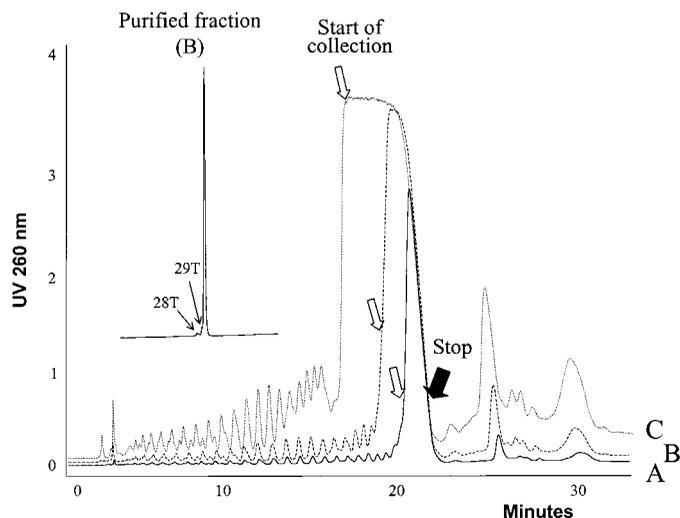


FIG. 2. Semipreparative purification of 30-mer oligodeoxythymidine. Solid line (A) $0.01 \mu\text{mol}$; dashed line (B) $0.03 \mu\text{mol}$; dotted line (C) $0.09 \mu\text{mol}$ injected on column. Target fractions were collected in the indicated intervals. The purity of the collected fraction for injection of $0.03 \mu\text{mol}$ is shown in the inset chromatogram. XTerra MS C18 75×4.6 mm, $2.5 \mu\text{m}$. Mobile phase A, 0.1 M TEAA, pH 7; B, 25% ACN in 0.1 M TEAA, pH 7. Gradient from 40% B to 55% B in 30 min, 0.5 ml/min, 51.5°C .

TABLE 1

Effect of 30-mer Oligodeoxythymidine Mass Load on the Ion-Pair RP-HPLC Purification Efficiency

Injected mass		Injected volume ^a (μ l)	Collected fraction purity ^b (%)	N-1 content ^c (%)	N-2 content ^c (%)
μ mol	mg				
0.01	0.091	5	96.3	1.7	2.0
0.03	0.272	15	96.3	2.2	1.5
0.09	0.816	45	93.2	3.8	3.0

^a 0.2 μ mol synthetic scale oligonucleotide was dissolved in 100 μ l of 0.1 M TEAA buffer, pH 7.

^b Purity was calculated as peak area % at UV 260 nm.

^c No other impurities than N-1 and N-2 were found.

broadens due to mass overloading, which is most apparent for the largest injected mass. Note that the peak shape is also affected by the UV detector nonlinear response at absorbance values greater than 2. The purity of the collected fractions is given in Table 1. Better than 96% purity was obtained for both 0.01- and 0.03- μ mol purification scales. When collected as shown in Fig. 2, the purification of 0.09 μ mol of the oligonucleotide resulted in 93.2% purity. HPLC analysis of the collected fraction confirmed the presence of small amounts of N-1 and N-2 failure sequence contaminants (Table 1). It is evident that the large-scale oligonucleotide purification on a 75×4.6 mm analytical column in overloading conditions may undermine the purity of the target fraction. The purity of the target fraction can be improved by heart cutting of the target peak. Because oligonucleotides are rarely separated with sufficient resolution, heart cutting is a commonly used technique in HPLC purification. However, one should be aware that the purity of the collected fraction is improved at the expense of target oligonucleotide recovery.

Alternative Buffers for Ion-Pair RP-HPLC

Triethylammonium acetate is the most commonly used ion-pairing mobile phase for oligonucleotide separations. However, with this ion-pairing agent it has been observed that retention depends partially on the oligonucleotide sequence. The separation of the ladder of heterogeneous sequence oligonucleotides can be more difficult than the separation of homooligonucleotide ladders. We experienced difficulties in resolving N and N-1 fragments for particular >25 -mer heterogeneous sequence oligonucleotides (data not shown). Using the tetrabutylammonium bromide ion-pairing agent can suppress the oligonucleotide sequence-dependent retention (29). However, the low volatility of this ion-pairing mobile phase makes its use for semi-preparative applications impractical.

An alternative ion-pairing buffer composed of triethylamine and hexafluoroisopropyl alcohol was described recently (30, 31). Apffel *et al.* utilized this buffer system for oligonucleotide analysis with direct ESI-MS detection, because TEAA buffers (33) as well as some other ion-pairing buffers (32) are known to dramatically suppress the ESI-MS detection sensitivity. Huber and Krajete have found that the signal intensity can be correlated to the equivalent conductivity of an acid used as a counter ion for TEA⁺. Although HFIP does not noticeably suppress an oligonucleotide ESI-MS signal, the HPLC separation performance of 2 mM TEA–400 mM HFIP buffer was not optimal (30–32).

One may expect that increased concentration of TEA in the TEA–HFIP buffer system should enhance its separation performance. We prepared and evaluated several TEA–HFIP buffers for oligonucleotide separation (34). Buffers were made by titrating the aqueous solutions of 400 or 100 mM HFIP (weak acid, $pK_a \sim 9.2$; the pH of 400 mM HFIP aqueous solution was 4.9) with TEA (see Materials and Methods). The final concentration of TEA in the 400 mM HFIP buffer was 16.3 mM (pH 7.9), which appears to be the limit of TEA solubility in this particular buffer. The solubility of TEA improves in a solution of 100 mM HFIP, so the two other ion-pairing buffers have been prepared consisting of 8.6 mM TEA–100 mM HFIP, pH 8.25, and 31.4 mM TEA–100 mM HFIP, pH 9.

Apffel *et al.* (30) and others (35) studied the impact of ion-pairing agent concentration in the mobile phase on oligonucleotide separation. Generally, the decrease in pairing ion concentration resulted in a decrease in separation efficiency and selectivity. The TEA concentration in TEA–HFIP buffers is significantly lower than in typical TEAA buffers (100 mM TEA), but the separation efficiency of oligonucleotides has improved (34). The separation performance of the TEA–HFIP ion-pairing buffer does not strictly follow the concentration of TEA in mobile phase. Instead, the resolution improves in line: 8.6 mM TEA–100 mM HFIP < 31.4 mM TEA–100 mM HFIP < 16.3 mM TEA–400 mM HFIP (34). According to the theory, the ion-pairing efficiency is not directly related to the concentration of pairing ion in the mobile phase (36–38). Instead, the concentration of pairing ion (in this case TEA⁺) adsorbed on the stationary phase surface is principally important for the separation. We believe that the limited solubility of the triethylammonium ion in 400 mM HFIP aqueous solution affects the distribution of TEA⁺ between the mobile and stationary phases, which results in more efficient absorption of the triethylammonium cation on the C18 sorbent and improved ion-pairing efficiency.

The example of a heterooligonucleotide ladder separation in IP-RP-HPLC using 16.3 mM TEA–400 mM HFIP buffer with a methanol gradient is shown in Fig.

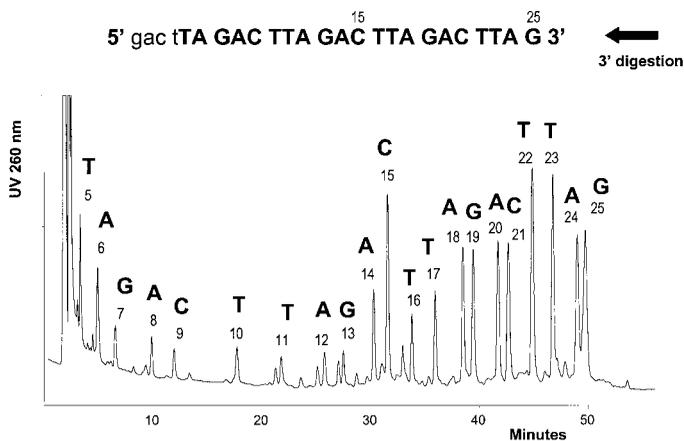


FIG. 3. Ion-pair RP-HPLC separation of an oligonucleotide ladder prepared by partial digestion of 15- and 25-mer oligonucleotides (for sequences see Materials and Methods). Peaks labeled 5–25 were identified by ESI-MS. 3'-terminal nucleotide is indicated at the top of the peak. Unlabeled minor peaks originate from impurities present in crude 15- and 25-mer before digestion. A CapLC system with PDA and ESI-MS detection was equipped with an XTerra MS C18 column, 50×1 mm, 2.5 μ m. Mobile phase A, aqueous TEA-HFIP buffer (16.3 and 400 mM), pH 7.9; B, MeOH. Gradient from 11% B to 20% B in 30 min and then 20% B to 24.5% B from 30 to 60 min, 23.6 μ l/min, 51°C. For ESI-MS conditions see Materials and Methods.

3. Minor peaks in Fig. 3 eluting closely to the 11-, 12-, and 13-mer originated from the contaminants present in the 15-mer peak before digestion. Figure 4 shows the separation of a ladder of partially digested 35-mer and 60-mer oligonucleotides. We achieved a good resolution up to 60-mer, including the separation of the 59-/60-mer peaks.

Impact of Sequence on the Oligonucleotide Separation

As mentioned earlier, the separation of heterooligonucleotides may present a challenge for ion-pair RP-HPLC using TEAA buffers. The TEA-HFIP buffer gave us good separation of a mixed sequence oligonucleotide ladders (Figs. 3 and 4). Figure 3 shows the separation of a 5- to 25-mer oligonucleotide ladder prepared by partial digestion of 15- and 25-mer oligonucleotides with a 3'-exonuclease. One can notice uneven spacing of some peaks, but all N, N-1, N-2 . . . products were well separated. A 100 mM TEAA buffer system (in the pH range 7–8.5) did not permit a complete separation of this mixture; close eluting fragments such as the 24 and 25-mers coeluted (data not shown).

In Figs. 3 and 4, we indicate the fragment length and the 3'-terminal nucleotide at the top of the peaks. A careful look at the chromatogram reveals a pattern. A loss of the more hydrophobic A or T mononucleotides from the 3' end results in a more pronounced decrease in peak retention than loss of rather hydrophilic C and G mononucleotides. The contribution of hydrophobicity

of oligonucleotides to the retention in ion-pair RP-HPLC was noticed earlier (21, 29). The retention of homooligonucleotides is known to follow the trend $p(dT)_n > p(dA)_n \gg p(dC)_n$. We believe that the hydrophobicity contribution to the retention has an impact on the separation of heterooligonucleotide ladder. We studied the retention behavior of homooligonucleotide ladders (digested homooligonucleotide) with a single heterogeneous nucleotide incorporated at the 3' oligonucleotide end using TEAA and TEA-HFIP buffers (data not shown). Based on our observation, we deduced a few rules for prediction of N and N-1 fragments separation success.

(i) Contribution of A and T mononucleotides hydrophobicity to the oligonucleotide retention is similar.

(ii) C and G mononucleotides contributes to the oligonucleotide retention similarly.

(iii) When the oligonucleotide sequence at the 3' end is a combination of A and T nucleotides (or C and G), the peak spacing is rather even as it is common for homooligonucleotide sequences. For example, the separation of 16- to 18-mer, 22- to 24-mer, and 58- to 60-mer (Figs. 3 and 4).

(iv) The separation seems to be enhanced when A and T mononucleotides are cleaved from the 3' end for sequences CA, GA, CT, and GT (5'–3') (see 9-/10-mer, 13-/14-mer, 19-/20-mer, and 31-/32-mer; Figs. 3 and 4).

(v) The separation selectivity decreases when C or G is cleaved from the 3' end for sequences AC, AG, TC, and TG (5'–3') (see 12-/13-mer, 14-/15-mer, 18-/19-mer, 24-/25-mer, and 32-/33-mer; Figs. 3 and 4).

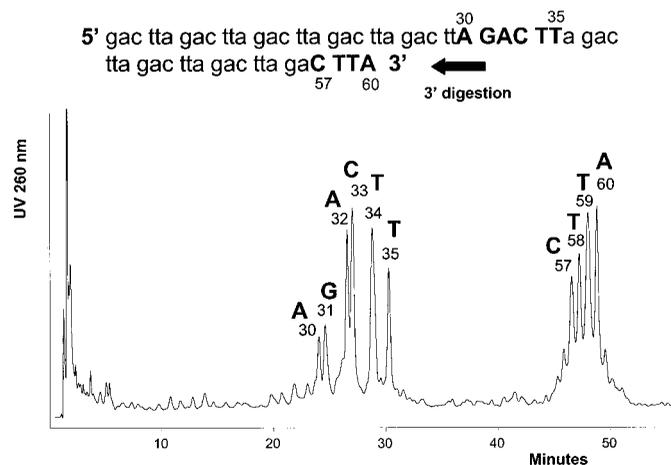


FIG. 4. Ion-pair RP-HPLC separation of an oligonucleotide ladder prepared by partial digestion of 35- and 60-mer oligonucleotides (for sequences see Materials and Methods). The identity of peaks labeled 30–35 and 57–60 was confirmed by ESI-MS. 3'-terminal nucleotide is indicated at the top of the peaks. Conditions: CapLC system with PDA and ESI-MS detection. XTerra MS C18 column, 50×1 mm, 2.5 μ m. Mobile phase A, aqueous TEA-HFIP buffer (16.3 and 400 mM), pH 7.9; B, MeOH. Gradient from 21% B to 27% B in 60 min, 23.6 μ l/min, 51°C. For ESI-MS conditions see Materials and Methods.

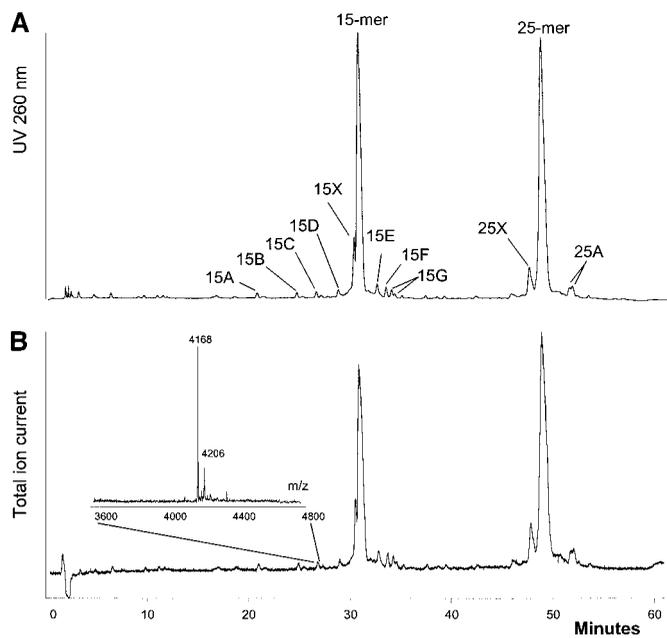


FIG. 5. LC-MS analysis of SPE purified 15- and 25-mer oligonucleotides (for sequences see Table 2). (A) Chromatogram at 260 nm; (B) total ion current MS chromatogram. (Inset) The deconvoluted mass spectrum of peak 15C. See Table 2 for peak identification. LC-MS conditions are the same as in Fig. 3.

We speculate that the rules of hydrophobicity contribution to oligonucleotide retention and separation can be extended to the 5' end sequences, as well as generalized to the whole oligonucleotide. If the influence of oligonucleotide secondary structure on retention is negligible (under the separation conditions), the loss of a nucleotide affects the retention independently whether it occurs at the terminus or in the middle of the oligonucleotide sequence. Recent (39) and earlier published reports (40, 41) support this hypothesis.

Figure 4 shows the retention behavior of partially digested 35- and 60-mer oligonucleotides. Using a 400 mM HFIP-TEA buffer, pH 7.9, the separation of fragments differing by a single nucleotide was achieved. The 59-mer and 60-mer oligonucleotides were clearly separated using a shallow gradient of methanol. One may notice that the separation of the 32-/33-mer oligonucleotides is more difficult than the 57- to 60-mer fragments (identity of all main peaks was confirmed by MS detection). This behavior is caused by the impact of oligonucleotide sequence on its retention.

LC-MS Analysis of Oligonucleotides

Taking advantage of the compatibility of the optimized TEA-HFIP buffer with ESI-MS detection, we were able to use a combination of UV detection and ESI-MS for identification of eluting peaks. The outlet UV detector capillary was directly connected to the

mass spectrometer source. The signal-to-noise (S/N) ratio of the MS chromatogram (total ion current mode) was lower than for UV detection (Figs. 5 and 6). However, the MS sensitivity was sufficient for positive identification of all main peaks. The molecular mass from the oligonucleotide multiply charged signal was deconvoluted by MaxEnt1 software.

To demonstrate the capabilities of LC-MS detection, we identified the molecular mass of all major peaks in Figs. 3 and 4. The difference in mass between adjacent peaks was ~289, 304, 313, or 329 for the loss of a C, T, A, or G mononucleotide, respectively. In this fashion we were able to directly read the oligonucleotide sequence from the ladder (Figs. 3 and 4), although this sequencing method is certainly crude and has a limited application.

Ion-pair RP-HPLC with MS detection was used as an effective tool for quality control of both crude and purified synthetic oligonucleotides. The TEA-HFIP mobile phase allows for the separation of failure sequence oligonucleotides from the target oligonucleotide and the identification of impurities based on the acquired molecular mass. We investigated the purity of two oligonucleotides, a 15-mer and 25-mer (for sequences see Table 2) purified by the "DMT on" SPE method described previously (7). The final purity of the oligonucleotides determined by CE was ~90%. We used LC-MS to study the remaining impurities as a part of the SPE method evaluation.

Figure 5 shows the separation of minor contaminants present in a mixture of the 15-mer and 25-mer

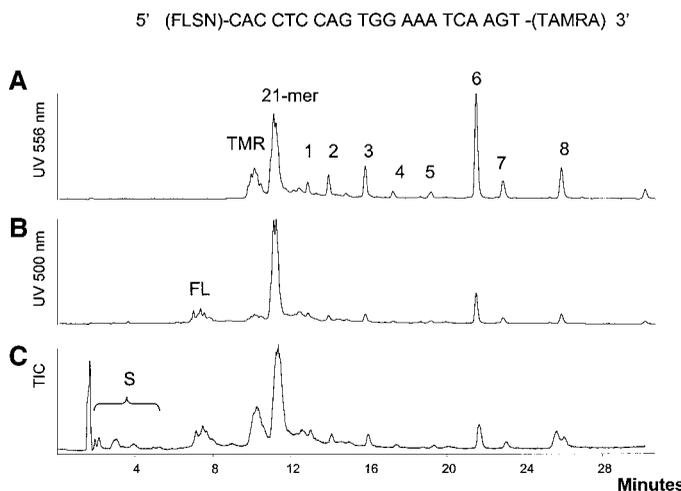


FIG. 6. IP-RP-HPLC analysis of a 21-mer TaqMan probe. (A) Chromatogram at 556 nm, detection of TAMRA; (B) chromatogram at 500 nm, detection of fluorescein; (C) total ion current MS chromatogram. For peak identification see Table 3. Conditions: CapLC system with PDA and ESI-MS detection. XTerra MS C18 column, 50 × 1mm, 2.5 μm. Mobile phase A, aqueous TEA-HFIP buffer (8.6 and 100 mM), pH 8.25; B, MeOH. Gradient from 15% B to 45% B in 30 min, 23.6 μl/min, 50°C. For ESI-MS conditions see Materials and Methods.

TABLE 2
Oligonucleotide Impurities Identified by LC-MS

HPLC peak label ^a	Measured <i>m/z</i> (Da)	<i>m/z</i> difference from parent compound (Da) ^b	Sequence difference ^c (5'–3')	Peak assignment
15 A	3524	–1051	Loss of CGA, A depurination	12-mer mismatch or 12-mer truncated from 3'
15 B	3837	–738	Loss of CG, A depurination	13-mer mismatch
15 C	4167	–408	Loss of C, A depurination	14-mer mismatch
15 D	4271	–304	Loss of T	14-mer mismatch
15 X	4443	–132	G depurination	15-mer depurinated
15-mer	4575	0		15-mer
15 E	4810	235	Branched A	15-mer + A branch
15 F	4629	54	Addition of cyanoethyl group	15-mer + CNEt
15 G	4325	–250	Loss of T, addition of cyanoethyl group	14-mer mismatch + CNEt
25 X	7366	–313	Loss of A	25-mer mismatch
	7375	–304	Loss of T	25-mer mismatch
25-mer	7679	0		25-mer
25 A	7732	53	Addition of cyanoethyl group	25-mer + CNEt

^a See Fig. 5.

^b Loss of dC, –289.187 Da; loss of dT, –304.198 Da; loss of dA, –313.211 Da; loss of dG, –329.211 Da; G depurination, –133 Da; A depurination, –117 Da; acetyl capping, +42 Da; cyanoethyl protection = +53.

^c Oligonucleotide sequence (5'–3'): 15-mer, GAC TTA GAC TTA GAC; 25-mer, GAC TTA GAC TTA GAC GAC TTA GAC T.

purified oligonucleotides. Interestingly, impurities elute both before and after the target oligonucleotide peaks. Two of the more retained peaks were found to be corresponding full-length oligonucleotides with an addition of ~53 Da (peaks 15F and 25A; see Table 2), which indicates the presence of the cyanoethyl moiety used as a protecting group in the oligonucleotide synthesis. Apparently, the cyanoethyl groups were not completely cleaved in the postsynthesis deprotection reaction with ammonium hydroxide. The hydrophobic character of the protecting moiety is responsible for the increased oligonucleotide retention. Peaks in the group labeled 25A have identical molecular mass. These 25-mers were separated probably due to differences in the position of the cyanoethyl group in the oligonucleotide chain.

A summary of the LC-MS analysis is given in Table 2. We did not find the presence of 5'-truncated failure sequence oligonucleotides. Instead we observed the N-1 and N-2 oligonucleotides with the molecular mass suggesting the loss of a different nucleotide than expected (for detail see Table 2). The results strongly suggest that these oligonucleotides are "mismatch" failure sequences. Because mismatch sequences have the trityl group attached after the last step of the synthesis (7), they were not removed from the full-length product in the wash step of the SPE procedure. On the other hand, the absence of 5'-truncated N-1 and N-2 oligonucleotides confirms that the SPE procedure efficiently removes "trityl off" failure sequences.

Among the other impurities we detected were products of depurination and a branched-backbone oligonucleotide (see Table 2). The sensitivity of mass spectrom-

etry was sufficient to measure a molecular mass even for a minor contaminant such as the peak labeled 15C (Fig. 5). Other unlabeled peaks eluting before 20 min and between 35 and 45 min are also oligonucleotide fragments (MS data are not shown).

LC-MS-PDA Analysis of Modified Oligonucleotides

LC-MS permits not only the identification of synthetic impurities, but most importantly the identification of the target oligonucleotide. This is beneficial for difficult syntheses, such as those of special oligonucleotides with modified backbones, DNA–RNA combinations, or oligonucleotides labeled with two different fluorescent tags. In our experience the syntheses of modified backbone and fluorescent oligonucleotides often result in several major products. MS detection of the desired molecular mass permits the choice of the proper collection interval and offers the possibility of fraction collection automation. Figure 6 illustrates the separation of a 21-mer TaqMan oligonucleotide probe labeled at the 5' and 3' ends with fluorescein and TAMRA, respectively. The current approach to purifying TaqMan probes is to collect the peaks with significant absorbance at both 500 nm (fluorescein) and 556 nm (maximum absorbance for TAMRA dye). The goal is to remove nonlabeled failure sequences, as well as the failure oligonucleotides labeled with only one of the fluorescent tags. Since the fluorescence of such products is not quenched, they may undermine the performance of TaqMan-based genotyping assays and quantitative PCR.

TABLE 3
LC-PDA-MS Quality Control of 21-mer TaqMan Oligonucleotide

HPLC peak label ^a	<i>m/z</i> measured (Da) ^b	Peak assignment
S	580, 785, 824, 992, 1137, 1197, 1315, 1424, 1449, 1492, 1530, 1543, 1590, 1601, 1625, 1655, 1663, 1673, 1821, 1823, 1920, 1973, 2016, 2121, 2132, 2135, 2217, 2230	unidentified short oligonucleotides fragments, not labeled
FL	6997 7004 7045 7257 7299	5' FLSN labeled 21-mer 5' FLSN labeled 21-mer + 7 Da (?); 5' FLSN labeled 21-mer + 7 + 42 Da (capping?); 5' FLSN labeled 21-mer + 7 + 42 + 212 Da (FLSN derivative + capping?) 5' FLSN labeled 21-mer + 7 + 42 + 212 + 42 Da (FLSN derivative + 2× capping?)
TMR	3684 4013 4317 4646 4958 5249 5538 5843 6132 6420 6733 7021 3097, 3306, 4354, 5755, 6462, 7066	5'-G AAA TCA AGT-TAMRA-3' 10-mer 5'-GG AAA TCA AGT-TAMRA-3' 11-mer 5'-TGG AAA TCA AGT-TAMRA-3' 12-mer 5'-G TGG AAA TCA AGT-TAMRA-3' 13-mer 5'-AG TGG AAA TCA AGT-TAMRA-3' 14-mer 5'-CAG TGG AAA TCA AGT-TAMRA-3' 15-mer 5'-C CAG TGG AAA TCA AGT-TAMRA-3' 16-mer 5'-TC CAG TGG AAA TCA AGT-TAMRA-3' 17-mer 5'-CTC CAG TGG AAA TCA AGT-TAMRA-3' 18-mer 5'-C CTC CAG TGG AAA TCA AGT-TAMRA-3' 19-mer 5'-AC CTC CAG TGG AAA TCA AGT-TAMRA-3' 20-mer 5'-CAC CTC CAG TGG AAA TCA AGT-TAMRA-3' 21-mer Unidentified mismatch sequences labeled at 3' with TAMRA
21-mer	7619 7629 7670 7684 7711	Target 21-mer, 3'TAMRA, 5' FLSN Target 21-mer, 3'TAMRA, 5' FLSN + 10 Da (derivative?) Target 21-mer, 3'TAMRA, 5' FLSN + 10 + 41 Da (derivative, capping?) Target 21-mer, 3'TAMRA, 5' FLSN + 65 Da (dye derivative?) Target 21-mer, 3'TAMRA, 5' FLSN + 92 Da (dye derivative?)
1, 2, 3, 4, 5, 6, 7, 8	1818, 1506, 1194, 1178, 890, 1729, 1310, 1122	3' TAMRA labeled oligonucleotide fragments (2-4-mer?)

^a See Fig. 6.

^b *m/z* shown only for major MS signals. TAMRA dye + linker, 623 Da; FLSN dye + linker = 598.6 Da. The nonlabeled 21-mer (expected mass 6398 Da) was not detected in crude oligonucleotide mixture.

Figure 6 shows the separation of a 21-mer target product from impurities. The group of peaks labeled S represents nonlabeled failure sequences. The peaks labeled FL were found to be mostly 5'-FLSN-21-mer isomers (see Table 3). A partially resolved group of peaks eluting at approximately 10 min labeled TMR comprises 3'-TAMRA labeled failure sequences truncated from the 5'-end (we identified 10- to 21-mer failure oligonucleotides) along with unidentified mismatch 3'-TAMRA oligonucleotides. The main fraction eluting at 11.5 min contains the target TaqMan probe. The observed peak split is likely due to the partial separation of dye positional isomers. The nonlabeled 21-mer oligonucleotide (expected *m/z* 6398.3) was not detected. However, we identified 21-mer-TAMRA conjugate (*m/z* 7021; TAMRA dye + linker mass is 623) and 21-mer-FLSN (*m/z* 6997; FLSN + linker mass is 598.6) (for details see Table 3).

Several peaks were found to elute after the main fraction (peaks 1–8 eluting between 12.8 and 26.0 min). They show pronounced absorbance at 556 nm

(TAMRA); the signal at 500 nm is caused by a dye spectral overlap (TAMRA has residual absorbance at 500 nm). The MS results indicate that peaks 1–8 are 2- to 4-mer oligonucleotide fragments. However, the *m/z* values do not correspond to those expected for the 5'-truncated oligonucleotides with 3'-TAMRA. Strong retention of these fragments indicates that the retention of short oligonucleotides is governed by the hydrophobicity of fluorescent dye rather than the oligonucleotide length (for a summary of 21-mer TaqMan impurities see Table 3). We are currently evaluating the use of mass directed fraction collection for automated purification of both native and fluorescent oligonucleotides.

CONCLUSION

Ion-pair RP-HPLC is an efficient method for analytical and semipreparative purification of oligonucleotides up to 60-mer length. Using a 75 × 4.6-mm XTerra MS C18 column, 2.5 μm, we purified up to 90 nmol of a

30-mer oligonucleotide in a single injection. We have evaluated mass spectrometry-friendly TEA–HFIP ion-pairing buffers for the separation of native and fluorescent oligonucleotides. TEA–HFIP (16.3 mM TEA, 400 mM HFIP, pH 7.9) was found to be a more efficient ion-pairing buffer for the separation of oligonucleotides than TEAA (100 mM TEA, 100 mM acetic acid, pH 7). We studied the impact of oligonucleotide sequence (nucleobase composition) on the separation of the target product from N-1, N-2 . . . failure sequence fragments.

A 1 × 50 mm XTerra MS C18 column, 2.5 μm was utilized for the LC-MS analysis and quality control of synthetic oligonucleotides. We investigated the remaining impurities in purified oligonucleotide samples that were not removed by the “trityl on” SPE purification method. The impurities were found to be mostly mismatch sequence oligonucleotides originating from incomplete capping in the oligonucleotide synthesis. We also detected impurities eluting after the main product peak. They were identified as full-length oligonucleotides with a cyanoethyl protecting group, which was not cleaved in the postsynthesis deprotection step. Finally, we employed HPLC with PDA and MS detection to identify the main product and impurities in a crude 21-mer TaqMan synthetic mixture.

REFERENCES

- Carrilho, E. (2000) DNA sequencing by capillary array electrophoresis and microfabricated array systems. *Electrophoresis* **21**, 55–65.
- Dovichi, N. J. (1997) DNA sequencing by capillary electrophoresis. *Electrophoresis* **18**, 2393–2399.
- Tang, K., Fu, D. J., Julien, D., Braun, A., Cantor, C. R., and Koster, H. (1999) Chip-based genotyping by mass spectrometry. *Proc. Natl. Acad. Sci. USA* **96**, 10016–10020.
- Little, D. P., Braun, A., O'Donnell, M. J., and Koster, H. (1997) Mass spectrometry from miniaturized arrays for full comparative DNA analysis. *Nature Med.* **3**, 1413–1416.
- Pastinen, T., Raitio, M., Lindroos, K., Tainola, P., Peltonen, L., and Syvanen, A. C. (2000) A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res.* **10**, 1031–1042.
- Morin, P. A., Saiz, R., and Monjabez, A. (1999) High-throughput single nucleotide polymorphism genotyping by fluorescent 5' exonuclease assay. *Biotechniques* **27**, 538–544.
- Gilar, M., and Bouvier, E. S. P. (2000) Purification of crude DNA oligonucleotides by solid-phase extraction and reversed-phase high-performance liquid chromatography. *J. Chromatogr. A* **890**, 167–177.
- Temsamani, J., Kubert, M., and Agrawal, S. (1995) Sequence identity of the n-1 product of a synthetic oligonucleotide. *Nucleic Acids Res.* **23**, 1841–1844.
- Pon, R. T., Buck, G. A., Hager, K. M., Naeve, C. W., Niece, R. L., Robertson, M., and Smith, A. J. (1996) Multi-facility survey of oligonucleotide synthesis and an examination of the performance of unpurified primers in automated DNA sequencing. *Biotechniques* **21**, 680–685.
- Buck, G. A., Fox, J. W., Gunthorpe, M., Hager, K. M., Naeve, C. W., Pon, R. T., Adams, P. S., and Rush, J. (1999) Design strategies and performance of custom DNA sequencing primers. *Biotechniques* **27**, 528–536.
- Mein, C. A., Barratt, B. J., Dunn, M. G., Siegmund, T., Smith, A. N., Esposito, L., Nutland, S., Stevens, H. E., Wilson, A. J., Phillips, M. S., Jarvis, N., Law, S., de Arruda, M., and Todd, J. A. (2000) Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res.* **10**, 330–343.
- Hessner, M. J., Budish, M. A., and Friedman, K. D. (2000) Genotyping of factor V G1691A (Leiden) without the use of PCR by invasive cleavage of oligonucleotide probes. *Clin. Chem.* **46**, 1051–1056.
- Tapp, I., Malmberg, L., Rennel, E., Wik, M., and Syvanen, A. C. (2000) Homogeneous scoring of single-nucleotide polymorphisms: Comparison of the 5'-nuclease TaqMan assay and molecular beacon probes. *Biotechniques* **28**, 732–738.
- Whitcombe, D., Brownie, J., Gillard, H. L., McKechnie, D., Theaker, J., Newton, C. R., and Little, S. (1998) A homogeneous fluorescence assay for PCR amplicons: Its application to real-time, single-tube genotyping. *Clin. Chem.* **44**, 918–923.
- Nasarabadi, S., Milanovich, F., Richards, J., and Belgrader, P. (1999) Simultaneous detection of TaqMan probes containing Fam and Tamra reporter fluorophores. *Biotechniques* **27**, 1116–1118.
- Fang, X., Li, J. J., Perlette, J., Tan, W., and Wang, K. (2000) Molecular beacons: Novel fluorescent probes. *Anal. Chem.* **72**, 747A–753A.
- Tyagi, S., Bratu, D. P., and Kramer, F. R. (1998) Multicolor molecular beacons for allele discrimination. *Nature Biotechnol.* **16**, 49–53.
- Buetow, K. H., Edmonson, M., MacDonald, R., Clifford, R., Yip, P., Kelley, J., Little, D. P., Strausberg, R., Koester, H., Cantor, C. R., and Braun, A. (2001) High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc. Natl. Acad. Sci. USA* **2**, 2.
- Huber, C. G. (1998) Micropellicular stationary phases for high-performance liquid chromatography of double-stranded DNA. *J. Chromatogr. A* **806**, 3–30.
- Huber, C. G., Oefner, P. J., Preuss, E., and Bonn, G. K. (1993) High-resolution liquid chromatography of DNA fragments on non-porous poly(styrene-divinylbenzene) particles. *Nucleic Acids Res.* **21**, 1061–1066.
- Huber, C. G., Oefner, P. J., and Bonn, G. K. (1993) High-resolution liquid chromatography of oligonucleotides on nonporous alkylated styrene-divinylbenzene copolymers. *Anal. Biochem.* **212**, 351–358.
- Warren, W. J., and Vella, G. (1995) Principles and methods for the analysis and purification of synthetic deoxyribonucleotides by high-performance liquid chromatography. *Mol. Biotechnol.* **4**, 179–199.
- Ausserer, W. A., and Biros, M. L. (1995) High-resolution analysis and purification of synthetic oligonucleotides with strong anion-exchange HPLC. *Biotechniques* **19**, 136–139.
- Makino, K., Ozaki, H., Matsumoto, T., Imaishi, H., Takeuchi, T., and Fukui, T. (1987) Reversed-phase ion-pair chromatography of oligodeoxyribonucleotides. *J. Chromatogr.* **400**, 271–277.
- Huber, C. G., Oefner, P. J., and Bonn, G. K. (1992) High-performance liquid chromatographic separation of detritylated oligonucleotides on highly cross-linked poly(styrene-divinylbenzene) particles. *J. Chromatogr.* **599**, 113–118.

26. Kirkland, J. J., Truszkowski, F. A., Dilks, C. H., Jr., and Engel, G. S. (2000) Superficially porous silica microspheres for fast high-performance liquid chromatography of macromolecules. *J. Chromatogr. A* **890**, 3–13.
27. Sykora, D., Svec, F., and Frechet, J. M. (1999) Separation of oligonucleotides on novel monolithic columns with ion-exchange functional surfaces. *J. Chromatogr. A* **852**, 297–304.
28. Premstaller, A., Oberacher, H., and Huber, C. G. (2000) High-performance liquid chromatography-electrospray ionization mass spectrometry of single- and double-stranded nucleic acids using monolithic capillary columns. *Anal. Chem.* **72**, 4386–4393.
29. Haefele, R., and Gjerde, D. Quality control and purification of oligonucleotides on the WAVE(R) nucleic acid fragment analysis system. *Transgenomics*, Application note 103.
30. Apffel, A., Chakel, J. A., Fischer, S., Lichtenwalter, K., and Hancock, W. S. (1997) New procedure for the use of HPLC-ESI MS for the analysis of nucleotides and oligonucleotides. *J. Chromatogr. A* **777**, 3–21.
31. Apffel, A., Chakel, J. A., Fischer, S., Lichtenwalter, K., and Hancock, W. S. (1997) Analysis of oligonucleotides by HPLC-ESI MS. *Anal. Chem.* **69**, 1320–1325.
32. Huber, C. G., and Krajete, A. (2000) Sheath liquid effects in capillary high-performance liquid chromatography-electrospray mass spectrometry of oligonucleotides. *J. Chromatogr. A* **870**, 413–424.
33. Huber, C. G., and Buchmeiser, M. R. (1998) On-line cation exchange for suppression of adduct formation in negative-ion electrospray mass spectrometry of nucleic acids. *Anal. Chem.* **70**, 5288–5295.
34. Gilar, M., Wheat, T. E., Chumsae, C., Budman, E., and Bouvier, E. S. P. (2000) Semi-preparative Purification of Native and Chemically Modified Oligodeoxyribonucleotide Probes by Reversed-Phase HPLC with ESI MS Detection, Poster 1112, HPLC 2000, Seattle, WA.
35. Huber, C. G., and Krajete, A. (1999) Analysis of nucleic acids by capillary ion-pair RP-HPLC coupled to negative-ion electrospray ionization MS. *Anal. Chem.* **71**, 3730–3739.
36. Bidlingmeyer, B. A., Deming, S. N., Price, J. W. P., Sachok, B., and Petrussek, M. (1979) Retention mechanism for reversed-phase ion-pair liquid chromatography. *J. Chromatogr.* **186**, 419–434.
37. Bartha, A., Vigh, G., and Stahlberg, J. (1990) Extension of the electrostatic retention model of reversed-phase ion-pair chromatography to include the simultaneous effects of the organic modifier and the pairing ion. *J. Chromatogr.* **506**, 85–96.
38. Bartha, A., Vigh, G., and Varga-Puchony, Z. (1990) Basis of the rational selection of the hydrophobicity and concentration of the ion-pairing reagent in reversed-phase ion-pair high-performance liquid chromatography. *J. Chromatogr.* **499**, 423–434.
39. Devaney, J. M., Pettit, E. L., Kaler, S. G., Vallone, P. M., Butler, J. M., and Marino, M. A. (2001) Genotyping of two mutations in the HFE gene using single-base extension and high-performance liquid chromatography. *Anal. Chem.* **73**, 620–624.
40. Ikuta, S., Chattopadhyaya, R., and Dickerson, R. E. (1984) Reverse-phase polystyrene column for purification and analysis of DNA oligomers. *Anal. Chem.* **56**, 2253–2256.
41. Haupt, W., and Pingoud, A. (1983) Comparison of several high-performance liquid chromatography techniques for the separation of oligodeoxynucleotides according to their chain lengths. *J. Chromatogr.* **260**, 419–427.