Reversed-phase ion-pair liquid chromatography analysis and purification of small interfering RNA

Sean M. McCarthy*, Martin Gilar, John Gebler

Waters Corporation, Milford, MA 01757, USA

ARTICLE INFO

Article history:
Received 5 March 2009
Available online 2 April 2009

This work is dedicated to Marianna Kele.

Keywords:
Oligonucleotide
Duplex
Double-stranded
RNA
DNA
Ultra-performance liquid chromatography (UPLC)
HPLC
Mass spectrometry

ABSTRACT

Small interfering RNA (siRNA)-induced gene silencing shows great promise in genomic research and therapeutic applications. siRNA duplexes are typically assembled from complementary synthetic oligonucleotides. High-purity single-stranded species are required for in vivo applications. Methods for separation, characterization, and purification of short RNA strands have been developed based on reversed-phase ion-pair liquid chromatography. The purification strategies were developed for both single-stranded and duplex RNA species. The method of duplex purification uses on-column annealing of complementary RNA strands, followed by separation of the target duplex from truncated duplexes and single-stranded RNA forms. The proposed method significantly reduces the purification time of synthetic siRNA.

© 2009 Elsevier Inc. All rights reserved.

Ribonucleic acid interference (RNAi)1 is a recently discovered mechanism for regulation of gene expression. The RNAi mechanism uses two forms of small RNA molecules: microRNA (miRNA), which prevents protein synthesis, and small interfering RNA (siRNA), which degrades messenger RNA (mRNA). The mechanism was first discovered in plants and later discovered in worms and mammals [1–3]. The RNAi pathway uses short oligonucleotide sequences (typically 21–23 nt) in duplex form referred to as siRNA. The currently accepted mechanism of RNAi describes the incorporation of siRNA into the RNA-induced silencing complex (RISC). The RISC uses one of the single-stranded components of the introduced siRNA to sequester complementary mRNA. Following binding, the mRNA is degraded, thereby preventing protein translation [4].

RNA molecules, including siRNA, for therapeutic use are produced synthetically via stepwise synthesis. Although this process is very efficient, with yields as high as 99.5% per coupling step, a simple 21-mer oligo typically used for RNAi therapeutic purposes would have a maximum purity of 90% without further purification. Most impurities, often referred to as “failure sequences,” arise during synthesis [5–7]. The ease of oligonucleotide synthesis, coupled with its utility in gene silencing, is leading to a dramatic increase in gene silencing experiments in genomic research. siRNA is also emerging as a therapeutic strategy for a wide variety of diseases, including cancer, macular degeneration, and viral infection [8–12]. A major challenge in developing siRNA therapeutics remains assurance of purity to minimize off-target gene silencing.

Gene silencing experiments are often performed with siRNA prepared by hybridization (annealing) of two complementary single-stranded counterparts. Each strand introduces its own set of impurities that can further complicate the siRNA mixture with mismatched sequences and noncomplementary single-stranded sequences. The presence of these impurities in a therapeutic mixture may lead to unwanted, and perhaps detrimental, nontargeted gene silencing. The presence of nonhybridized single-stranded RNA (ssRNA) is often associated with a decrease in therapeutic potency, further highlighting the need for efficient purification strategies for both RNA and siRNA.

There have been a number of methods reported for the analysis and purification of oligonucleotides, including polyacrylamide gel
Electrophoresis (PAGE), capillary gel electrophoresis (CGE) [13, 14], anion-exchange high-performance liquid chromatography (AX HPLC) [15–17], and reversed-phase ion-pair liquid chromatography (RP IP LC) [5, 18, 19].

RP IP LC provides good resolution of oligonucleotides and is compatible with mass spectrometry (MS) detection. Only a few reports have shown LC MS analysis of siRNA using a triethylamine/hexafluoroisopropanol (HFIP) ion-pairing system [20, 21].

The preparation of siRNA typically requires separate purification steps for each complementary RNA strand, followed by annealing of the purified products in the final step, yielding the desired siRNA. This article outlines an alternative, less labor-intensive method for purification of siRNA in duplex form using on-column annealing of crude synthetic complementary RNA strands. The efficient method for LC and LC MS analysis of siRNA are also presented.

Materials and methods

Materials

Two molar triethylammonium acetate (TEAA, cat. no. 5204-74-0), hexylamine (HA, cat. no. 203-851-8), acetic acid (cat. no. 200-580-7), and phosphodiesterase II (PDII) were purchased from Sigma–Aldrich (St. Louis, MO, USA) and used without further purification. Acetonitrile (ACN, Optima LC/MS grade, cat. no. 75-05-8) was purchased from Fisher Scientific (Fair Lawn, NJ, USA). Synthetic RNA oligonucleotide sequences were purchased from Integrated DNA Technologies (IDT, Coralville, IA, USA). The sequences were as follows and are labeled arbitrarily for the purpose of clarity as upper and lower strands; however, the reader should understand that these distinctions may have important meaning elsewhere. The sequences were 19-mer RNA with DNA TT overhangs and are listed as complementary sequences: RNAi upper 5'-UGG UCA AGC GAU UAC AAG GTT-3', RNAi lower 5'-CCU UGG AAU CGG UUG AGC ATT-3', FAS upper 5'-CCU UGA GAU CCC AGG GCC GTT-3', and FAS lower 5'-CAG GCC UGG GAU CUC AGG GTT-3'. Sequences used for the duplex melting study were as follows: low melt duplex (47.5 °C) upper 5'-TTG TCA TGT TGT AAT C-3' and lower 5'-GAT TAC AAC ACA GTG TGCA A-3'; medium melt duplex (57.0 °C) upper 5'-TTG TCG TCT CCG ACT C-3' and lower 5'-GAT TAC AAC ACA GTG TGCA A-3'; and high melt duplex (66.0 °C) upper 5'-TTG TCG TCG CGG AGC C-3' and lower 5'-GAT TAC AAC ACA GTG TGCA A-3'. A Milli-Q system (Millipore, Bedford, MA, USA) was used to prepare deionized water (18 MΩ cm) for HPLC mobile phases.

RNA and siRNA purification with HPLC

Crude synthetic oligonucleotides were reconstituted to yield a solution of approximately 3.5 nmol/µl in 0.1 M TEAA. The resulting mixture was purified using a Waters Alliance Bio 2696 separation module with a 996 photodiode array detector (Waters, Milford, MA, USA) scanning from 220 to 350 nm. Oligonucleotides were separated using a 4.6 × 50-mm XBridge C18 column packed with 2.5-µm particles (average pore diameter 35 Å, Waters). The mobile phases were as follows: 0.1 M TEAA (pH 7.0) (A) and 80/20% (v/v) A/ACN (B). The mobile phases were degassed via sonication for 10 min. The flow rate was 1 ml/min.

The gradient was from 30% to 52.5% B in 30 min (0.15% ACN/min) for ssRNA and from 25% to 75% B in 30 min (0.33% ACN/min) for duplex separations. Single-stranded separations were performed at 60 °C, and duplex separations were performed at 20 °C. Fractions were collected manually from the peak apex to approximately 30% of the peak height in all cases. Collected fractions were evaporated using a CentriVap (Labconco, Kansas City, MO, USA) and reserved for further analysis and use.

UPLC analysis of single-stranded RNA

The single- and double-stranded nucleic acids were analyzed with an ACQUITY ultra-performance liquid chromatography (UPLC) system using a Waters ACQUITY OST C18 column packed with 1.7-µm particles (average pore size 130 Å). The mobile phase was 100 mM TEAA as mobile phase A and 20% ACN in A with a flow rate of 0.2 ml/min and a gradient of 35–65% B in 10 min. A photodiode array detector was used for detection at 260 nm (all instruments were obtained from Waters). On-column loading of each sample was 7.5–10 pmol. The temperature was varied between 20 and 70 °C in 5 ° increments (see figure captions).

Following purification, ssRNA, double-stranded DNA (dsDNA), and siRNA were analyzed using a UPLC and single quadrupole mass spectrometer. dsDNA was prepared from complementary oligonucleotides by heating the sample to 90 °C and cooling slowly to room temperature over 10 min. The LC MS mobile phases were composed of 25 mM hexylammonium acetate (HA, pH 7.0) as mobile phase A and 100% ACN as mobile phase B. The gradient varied depending on the sample and mobile phase used as described in the figure legends. The MS instrument was used in electrospray ionization (ESI) negative ion mode, capillary voltage was 3 kV, cone voltage was 28 V, extractor voltage was 3 V, source temperature was 150 °C, desolvation temperature was 350 °C, cone gas flow was 31 L/h, and desolvation gas flow was 700 L/h. MS data were deconvoluted using the maximum entropy software MaxEnt 1.

Results and discussion

Purification of RNA single strands

The RP IP LC method is well suited for purification of nucleic acids [5]. As shown in Fig. 1, useful amounts of RNA can be purified in a single injection using an analytical column. The purification up to 140 nmol of a 21-mer oligonucleotide was performed using TEAA mobile phases. As the column load increased, typical peak broadening is observed with a partial retention shift of failed sequences (selected impurities are labeled with asterisks in Fig. 1). Due to the displacement effect, ’hearth cutting’ is a viable strategy to obtain good purity of target oligonucleotides with only a moderate sacrifice in product recovery [5].

For all mass loads illustrated in Fig. 1, the main peak was collected from its apex to approximately 30% of the peak height, shown in the figure for the largest injection by arrows. Typically, 1–2 ml was collected as a single fraction depending on the column mass load. The yield was between 55% and 70% based on the peak area in the collected range versus the total peak ultraviolet (UV) area.

UPLC analysis of the collected fractions indicated that the isolated oligonucleotides were typically >95% pure (Fig. 1 inset), which is suitable for a variety of molecular biology and therapeutic applications.

Duplex DNA stability in RP IP LC

RNA can be purified in single-stranded or duplex form. The latter approach is intriguing, potentially shortening the process. Nevertheless, several practical problems need to be addressed. The foremost concern is the stability of siRNA duplexes under RP IP LC conditions. It has been well documented that RP IP LC is suitable for analysis of 50- to 1000-bp-long dsDNA restriction fragments [22, 23]. However, unmodified siRNA species are typically 19 bases long complementary region, with melting temperatures (T_m) ranging between 45 and 65 °C. Some on-column melting of siRNA can potentially occur during the analysis or purification, especially at elevated temperatures [22].
To investigate the stability of short duplexes, we used a set of DNA duplexes with various melting points. Three duplexes were prepared from purified single-stranded counterparts with melting points of 47.5 °C (low T\textsubscript{m}), 57.0 °C (medium T\textsubscript{m}), and 66.0 °C (high T\textsubscript{m}). The variation in melting point was accomplished by strategic replacement of T and/or A bases with G and/or C bases that offer greater thermal stability due to increased H bonding. The duplexes were prepared with a 50% excess of upper strand so as to monitor RP IP LC retention of both single-stranded oligonucleotide and dsDNA species.

Figs. 2A–C illustrate the analysis of duplexes at various temperatures. Approaching the melting temperature, dramatic peak broadening occurs, indicating the on-column duplex melting. The dsDNA melting is accompanied by an appearance of complementary oligonucleotides. The T\textsubscript{m} values observed in Fig. 2 roughly correspond to the calculated dsDNA values. When the column temperature reached and exceeded the melting point, the duplexes rapidly melted on injection on column and eluted from column as a pair of sharp and symmetrical single-stranded peaks. Retention times and peak shape of the single-stranded oligonucleotides injected separately are provided for comparison as supplemental materials (see Supplemental Figs. 1–3 in supplementary material).

It is worth pointing out that retention of oligonucleotides is noticeably affected by their sequence (nucleobase composition) [5,24], whereas all dsDNA species elute in nearly identical retention time. The RP IP LC retention of duplexes is driven less by sequence and more by charge-to-charge interaction (dsDNA is more retained than its corresponding oligonucleotides).

**UPLC analysis of siRNA**

Based on the experiment shown in Fig. 2, we selected 20 °C as a generic separation temperature for further siRNA experiments. The duplex siRNA samples were prepared with an excess of either upper or lower strand to demonstrate the resolution of duplex siRNA from its single-stranded counterparts.

As shown in Figs. 3A and B, the RP IP LC clearly separates full-length duplex from single-stranded oligonucleotide “impurities.” No apparent deterioration of duplex peak shape is observed, indicating little or no on-column melting of siRNA duplex.

Fig. 3C shows the analysis of siRNA duplex prepared by annealing nearly equimolar amounts of complementary RNA oligonucleotides. A small excess of the lower RNA strand was present, whereas the peak corresponding to the upper strand was not detected. This supports the conclusion that siRNA duplex is stable at given chromatographic conditions. Minor peaks eluting prior to the target siRNA peak are presumably partially truncated duplex species. The ability of the RP IP LC method to resolve siRNA from its truncated duplex impurities is discussed in more detail in the following section.

**Resolution of partially truncated siRNA species**

“Mismatched” truncated duplexes arise when annealing oligonucleotides are contaminated with shorter failed sequences. Little is known about the performance of RP IP LC for the separation of truncated duplex species. Naturally, the purification of siRNA in duplex form is a viable strategy only if an efficient separation of target duplex from truncated species is achieved.

To investigate the ability of our method to afford these separations, we prepared a 5’ truncated ladder of one complementary strand by its partial digestion with the exonuclease PDII. We then annealed this mixture of truncated sequences with a full-length complementary strand. The resulting mixture mimics a crude synthetic mixture that is often contaminated with failure sequences (Scheme 1).

Full-length truncated duplexes and remaining single-stranded species were separated using UPLC with 25 mM HAA mobile phase, which provides suitable chromatographic resolution and adequate MS sensitivity. Data shown in Fig. 4 suggest that the RP IP LC method employed resolved the truncated duplexes.

The retention order of peaks in Fig. 5 was confirmed by MS. The traces of selected ion chromatograms were generated by extracting the –3 or –4 charge state mass for the appropriate oligonucleotides, and the results are shown in Fig. 5A. Interestingly, when inspecting the MS data under detected peaks, we do not observe the expected masses of duplexes. Instead, the MS data reveal the presence of two complementary oligonucleotides corresponding to the original duplexes. Apparently, the siRNA species survive the chromatographic separation, but siRNA duplexes are melted in the ESI source upon ionization.
The MS data for truncated U21/L20 duplex peak is shown in Fig. 5B. The MS spectra arising from multiply charged species were deconvoluted, yielding the expected masses of 21 nt upper strand and 20 nt lower strand oligonucleotide (Fig. 5C). This confirms that the RP IP LC method is capable of resolving closely related truncated duplex species.

We investigated the prospects for MS detection of the intact duplex mass by decreasing both the desolvation and source temperature. Although we were able to produce spectra that contained the duplex along with some single-stranded components, these species were also present with a large percentage of hexylamine adducts (data not shown). We investigated the use of ammonium acetate for intact duplex MS analysis with greater success. We found that the use of ammonium acetate yielded ionization of the intact duplex using normal MS operating conditions, but the chromatographic elution was reversed in this non-ion-pairing buffer (duplex species eluting prior to the single-stranded species) and the separation selectivity suffered.

MS signal of oligonucleotides typically contains alkali ion adductation. Fig. 5C shows a significant iron adductation in the presented case. The source of iron is unclear; we believe that traces of iron are present in the mobile phase.

Closer inspection of Fig. 4 indicates that as the lower strand length decreases from 21- to 16-mer and lower, incomplete duplex formation and/or on-column melting become apparent. We calculated the theoretical melting points for truncated duplexes [25]. The \( T_m \) values decreased from 49 °C for the full-length duplex to 34 °C for the U21/L16 mismatched duplex (only 14 base pairs contribute to annealing). This dramatic decrease in melting point is the likely source of incomplete duplex formation and melting for the shortest duplexes even when using a 20 °C separation temperature.

**Purification of siRNA from crude oligo mixtures**

After demonstrating the stability of siRNA duplexes and their resolution from truncated failure sequence products, we investigated the RP IP LC method utility for duplex purification. Similar to our purification strategy with single-stranded oligonucleotides, we evaluated the ability of our method to yield purified siRNA in a limited number of steps.

First, we annealed crude complementary oligonucleotide mixtures off-line to prepare the duplex siRNA. The expected product impurity profile is more complex than the one obtained by annealing a partially digested single strand with a full-length complementary strand shown in Scheme 1. Nevertheless, the HPLC purification of the resulting mixture was successful, and the method yielded good separation of single-stranded and mismatched duplexes from the desired full-length duplex (data not shown). The full-length duplex was collected by appropriate hearth cutting of the peak, and subsequent analysis via UPLC revealed that the full-length duplex purity was \( \geq 98\% \).

Following our success in duplex purification using a separate annealing step, we investigated whether the method can be further improved and shortened. This was prompted by the observation that when injecting two complementary RNA strands on chromatographic column at conditions below melting temperature, they spontaneously anneal and elute as duplex. This is the case even when including the 8 M urea in the sample as denaturant. It appears that as the urea is washed away by the mobile phase and the complementary strands are focused on the head of column, they spontaneously and effectively anneal to the corresponding duplex.

We scaled up the siRNA purification using the on-column annealing with the following strategy. The first strand was injected at initial gradient conditions, followed by injection of the complementary strand. The gradient elution was started immediately afterward. We observed quantitative formation of the duplex siRNA and its elution at expected elution time. The purification strategy was scalable (Fig. 6), yielding results similar to those found for ssRNA. Expected single-stranded and double-stranded impurities were eluting mostly prior to the target peak. The collection was performed by hearth cutting the main peak from the apex to 30% of the peak height, as indicated in Fig. 6 by the arrows (for the 85-nmol purification scale). Collected fractions were analyzed via UPLC. Fig. 7A shows a typical UPLC chromatogram of a purified du-
plex at 20 °C overlaid with crude upper and lower strands analyzed at the same conditions. These data confirm the purity of the collected duplex. Two small peaks of single-stranded oligonucleotides were detected at 0.6–0.8% levels, suggesting that minor on-column duplex melting occurs on injection at a given low-ionic-strength LC MS mobile phase system (25 mM TEAA, pH 7.0); mobile phase B: 20% ACN in mobile phase A. Gradient: from 35% B (7% ACN) to 85% B (13% ACN) in 10 min, 0.2 ml/min, 20 °C. The eluent was monitored at 260 nm.

Fig. 3. UPLC separation of siRNA duplex from excess of upper chain of ssRNA (A) and excess of single-stranded lower chain (B). (C) Separation of siRNA duplex prepared by annealing nearly stoichiometric ratio of complementary RNA oligonucleotides. Column: Waters OST BEH C18, 2.1 × 50 mm, 1.7 μm. Mobile phase A: 100 mM TEAA (pH 7.0); mobile phase B: 20% ACN in mobile phase A. Gradient: from 35% B (7% ACN) to 85% B (13% ACN) in 10 min, 0.2 ml/min, 20 °C. The eluent was monitored at 260 nm.

Scheme 1. Potential mismatched duplexes arising from annealing of a mixture of full-length upper strand (U21) and complementary truncated lower strand (L21, L20, L19, ..., L1) RNA sequences.

Limitations of the method

We investigated the generality of the RNA purification methods by purifying multiple sets of RNA oligonucleotides. Interestingly, we found that oligonucleotides with strong secondary structure yield broad fronting and tailing peaks and could be difficult to purify in single-stranded form even at 60 °C. This may present a problem for scaling up the method to a preparative scale, where the column is typically kept at ambient temperature. Occasionally, the on-column formation of intramolecular structures may interfere with purification even for oligonucleotides with relatively weak secondary structure.

To overcome this issue, we annealed a complementary pair of oligonucleotides with significant secondary structure off-column and performed purification in duplex form (see Supplemental Fig. 4 in supplementary material). The purification was successful; the secondary structure impact apparently was eliminated by the formation of duplex. On the other hand, the on-column annealing method was less efficient; besides the duplex peak, we detected increased amounts of non-annealed single-stranded oligonucleotides. It appears that secondary structure can, to a certain extent, prevent the on-column formation of duplex. The duplex purification was still possible, but the yield was reduced to approximately half for the given set of oligonucleotides. For oligonucleotides with strong secondary structure, it is advantageous to use a separate off-column annealing step if high purification yields are desired.

Conclusions

The methods presented in this article allow semipreparative scale purification of RNA in both single-stranded and duplex forms. The proposed methods are rapid and gentle while offering impressive resolution of both single-stranded and duplex RNA impurities.
The scalability of the LC purification method allows scaling up to a preparative scale. The proposed methods are fast, efficient, and cost-effective, and they represent a significant advance in siRNA purification strategies. In addition, we developed the method for LC MS analysis of ssRNA and dsRNA using the nondenaturing mobile phase consisting of HAA. This concurrent analysis via UV and MS reduces analysis times considerably compared with other methods. The use of HAA is also cost-effective compared with HFIP-based mobile phases, making it an attractive alternative for routine duplex analysis. The mobile phase system was shown to be nondenaturing for duplex analysis, as evidenced by intact duplex elution.
Fig. 6. RP IP HPLC purification of siRNA duplex formed on column after injecting two complementary synthetic RNA oligonucleotides. Column: XBridge OST BEH C18, 4.6 × 50 mm, 2.5 μm. Mobile phase A: 100 mM TEAA (pH 7.0); mobile phase B: 20% ACN in mobile phase A. Gradient: from 25% B (5% ACN) to 75% B (12% ACN) in 30 min, 1 ml/min, 20 °C. The eluent was monitored at 260 nm.

Fig. 7. UPLC verification of purified siRNA quality. (A) UPLC analysis of purified siRNA duplex and each crude complementary strand at 20 °C. (B) UPLC analysis of purified siRNA and each crude complementary strand at 60 °C. Column: Waters OST BEH C18, 2.1 × 50 mm, 1.7 μm. Mobile phase A: 25 mM HAA (pH 7.0); mobile phase B: 100%. Gradient: from 30% B to 40% B in 10 min, 0.2 ml/min.
Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ab.2009.03.042.

References