RNA interference (RNAi) has quickly become a well-established laboratory tool for regulating gene expression and is currently being explored for its therapeutic potential. The design and use of double-stranded RNA oligonucleotides as therapeutics to trigger the RNAi mechanism and a greater effort to understand the RNAi pathway itself is driving the development of analytical techniques that can characterize these oligonucleotides. Electrospray (ESI) and MALDI have been used routinely to analyze oligonucleotides and their ability to provide mass and sequence information has made them ideal for this application. Reviewed here is the work done to date on the use of ESI and MALDI for the study of RNAi oligonucleotides as well as the strategies and issues associated with siRNA analysis by mass spectrometry. While there is not a large body of literature on the specific application of mass spectrometry to RNAi, the work done in this area is a good demonstration of the range of experiments that can be conducted and the value that ESI and MALDI can provide to the RNAi field.

**Keywords:** RNAi; siRNA; miRNA; RNA interference; mass spectrometry

**I. INTRODUCTION**

The growing importance of inhibitory RNA (RNAi) in regulating gene expression and its potential as a therapeutic has created a need for techniques that can directly analyze RNA both quantitatively and qualitatively. Unlike molecular biology assays or chromatographic methods, mass spectrometry (MS) provides a direct observation of an intrinsic property of a compound: mass. This measurement yields a specificity not available by other methods and allows quantification and identification of an oligonucleotide analyte, its impurities, and its metabolites. In addition to mass, MS techniques can also sequence the oligonucleotide, thus offering a further level of analyte interrogation. These attributes plus the capacity to couple MS analysis to a chromatographic separation provide an analysis format that allows for a direct observation of an intrinsic property of a compound: mass. This measurement yields a specificity not available by other methods and allows quantification and identification of an oligonucleotide analyte, its impurities, and its metabolites. In addition to mass, MS techniques can also sequence the oligonucleotide, thus offering a further level of analyte interrogation. These attributes plus the capacity to couple MS analysis to a chromatographic separation provide an analysis format that is able to identify oligonucleotides using retention time, mass, and sequence. Lastly, as the triggering mechanism for the RNAi pathway is a short (19–22 bp) double-stranded RNA (dsRNA), the ability of soft ionization techniques such as MALDI and electrospray to preserve these non-covalent dsRNA structures during mass measurement is particularly advantageous.

Both MALDI and ESI have been used to study single- and double-stranded oligonucleotides, and the analysis of oligonucleotides by MS has been covered thoroughly, including reviews specifically on MS-based sequencing, MS of RNA, non-covalent oligonucleotide complexes, and LC–MS of oligonucleotides (Limbach, Crain, & McCloskey, 1995a; Limbach, 1996; Crain & McCloskey, 1998; Hofstadler & Griffey, 2001; Huber & Oberacher, 2001; Banoub et al., 2005; Hofstadler, Sannes-Lowery, & Hannis, 2005; Meng & Limbach, 2006; Thomas & Akoulitchev, 2006; Lin, Li, & Dai, 2007). While this earlier body of work does not cover the dsRNA used in RNAi, the techniques and methodologies described in them are fully applicable to these newer classes of RNA compounds. As such, only applications and details relevant to RNAi will be stressed. Most of the work in this review will focus on how MS is being used to understand the RNAi mechanism itself or to analyze the therapeutic dsRNA oligonucleotides created to trigger that mechanism.

**II. INHIBITORY RNA (miRNA AND siRNA)**

The phenomenon termed RNA interference (RNAi) regulates gene expression and can serve an immune function by preventing protein synthesis from mRNA. First discovered in plants and then in worms and mammals, the gene-silencing mechanism operates using short sections of dsRNA (Jorgensen, 1990; Fire et al., 1998; Elbashir et al., 2001a). Since its discovery, the use of RNAi to study gene function has become a standard laboratory procedure, and its importance to biology and its possible therapeutic applications were recognized by a Nobel Prize in 2006 (Barik, 2004; Jorgensen, 2006; Zamore, 2006). There are many excellent reviews on RNAi and therefore it will only be briefly summarized here (Alisky & Davidson, 2004; Novina & Sharp, 2004; Rana, 2007; Yang & Mattes, 2008).

The RNAi pathway is initiated by dsRNA from two sources; dsRNA introduced into cells from exogenous sources is termed as small interfering or siRNA, whereas endogenous non-coding dsRNA produced in the nucleus is called miRNA. In addition to their place of origin, miRNA and siRNA can also differ in length, form, and function in the RNAi pathway (Mack, 2007). This is an area of ongoing research, and the properties and functions of siRNA and miRNA are still being understood and defined.

Currently, it is believed that dsRNA approximately 19 and 22 bp can be loaded into the RNAi pathway via a protein complex.
termed the RNA-induced silencing complex (RISC). It is thought that dsRNA longer than approximately 19–22 bp is first cleaved to this length by the endonuclease DICER before RISC loading and that dsRNA shorter than this length does not load. It has been observed in *Drosophila* that after DICER cleavage, the resulting dsRNA duplex or siRNA is not a blunt structure but has an overhang of two nucleotides on the 3′ ends of each strand (Elbashir et al., 2001b).

RISC loading occurs in the cytoplasm and the dsRNA is unwound as it is incorporated into RISC so that only a single strand remains bound (Filipowicz, 2005). The loading of a specific single strand into RISC is dictated by the thermodynamic stability of the ends of the duplex. The strand with the less stable 5′ end is preferentially loaded (Khvorova, Reynolds, & Jayasena, 2003). In addition to thermodynamic stability, it has also been shown that the phosphorylation of the 5′ end of the oligonucleotide strand affects its ability to be loaded into RISC (Filipowicz, 2005).

The RNA strand that is loaded into RISC is termed the guide strand (sometimes referred to as antisense), and the free strand is termed the passenger strand (sometimes referred to as sense). The passenger strand is believed to be degraded by nucleases once it is unwound and in a single-stranded form. Once bound, RISC uses the single-stranded oligonucleotide from either the siRNA or miRNA as a template to identify target sequences of mRNA (Ameres, Martinez, & Schroeder, 2007). mRNA target recognition by the RISC–RNA complex is guided by the 5′ region of the RNA (Rana, 2007). In cases of a perfect complementary match between guide strand and mRNA, it is hypothesized that the mRNA is cleaved whereas partial complementarity may result in repression of the mRNA (Mack, 2007). After cleavage, the RISC–RNA complex finds a new mRNA to bind and it is in this manner that viral replication and/or gene function is regulated.

### A. RNAi Structure and Properties

The presence of a hydroxyl group on the 2′ position of the ribose and the absence of a methyl group on the nucleotide thymidine to make uridine are what differentiate RNA from DNA. It is the 2′ hydroxyl group, however, that is responsible for the majority of RNA's defining characteristics. When RNA is in a duplex form, the sugar phosphate backbone that links the ribonucleotides together is stabilized by the 2′ OH in the C3 endo position unlike DNA which is in the C2 endo position. The result is that duplex RNA takes the A helix form versus DNA’s B form (Fig. 2). The A helix conformation is thought to be responsible for a greater hydration shell which results in RNA duplexes having enhanced thermodynamic stability and being more rigid than DNA duplexes (Freier et al., 1986; Lesnik & Freier, 1995; Portmann, Usman, & Egli, 1995). RNA helixes complete a 360° rotation in 11–12 bp versus 10 for DNA and the tighter A-form helix geometry has been proposed to be the determining factor for why dsRNA and not dsDNA triggers RNAi (Rana, 2007).

---

**FIGURE 1.** Cartoon depicting RNAi mechanism. Cleavage of the mRNA target strand by RISC is guided by the template sequence from long double-stranded RNA, siRNA, or miRNA. After cleavage RISC remains active to cleave additional mRNA. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
Like DNA duplexes, the RNA duplex forms with the hydrophobic nucleobases Watson–Crick base pairing internally. The pKₐ of the phosphodiester linkage is approximately 2, so in aqueous solution they exist as polyanions. Single-stranded RNA is highly likely to have tertiary structures such as hairpins and bulges (Draper, 2004). The strong affinity of RNA to cations (a result of the 2’ hydroxyl) such as magnesium...
and potassium aids in the formation of duplex or other structures by shielding the negative charges on the phosphodiester backbone (Draper, 2004).

Chemical modifications are often made to synthetic therapeutic oligonucleotides to prevent nuclease degradation and thereby improve their pharmacokinetic properties (Geary et al., 2001; Manoharan, 2004; Egli et al., 2005; Behlke, 2008). Some of the more common modifications are 2′ fluoro, 2′ O-methyl, 2′ MOE, phosphorothioate, and locked nucleic acids (LNA) (Fig. 3). It has been shown that siRNA with 2′ fluoro and 2′ O-methyl modifications retain their functionality, have increased plasma stability, and in some cases are more inhibitory than RNA alone (Layzer et al., 2004; Morrissey et al., 2005a,b). The LNA modification which joins the 2′ and 4′ position of the ribose via a 2′ O–4′ methylene linkage has also been reported to improve siRNA stability while retaining activity (Elmen et al., 2005, 2008; Mook et al., 2007). This is the same for phosphorothioate-modified siRNAs where oxygen in the phosphodiester linkage is replaced with sulfur (Braasch et al., 2003; Elmen et al., 2005, 2008). All of these common modifications increase the hydrophobicity of the oligonucleotide and the temperature at which the duplex denatures or melts (the $T_m$) into its respective single strands (Lesnik et al., 1993; Freier & Altmann, 1997; Egli et al., 2005).

The nucleobases themselves have different hydrophobicities and follow the order C < G < A < T so the sequence of the oligonucleotide greatly affects its behavior both in reversed-phase chromatography and anion exchange chromatography (Huber, Oefner, & Bonn, 1992, 1993; Gilar et al., 2002; Dickman, 2005; Buncek et al., 2006). The sequence and hydrophobicity of the oligonucleotide and the temperature at which the duplex denatures or melts (the $T_m$) into its respective single strands (Lesnik et al., 1993; Freier & Altmann, 1997; Egli et al., 2005).

B. Current Techniques for RNAi Analysis

A typical IC50 for an siRNA in cell culture experiments is in the low to sub-nanomolar range (0.2–5 nM) (Kretschmer-Kazemi Far & Szakiel, 2003; Overhoff, Wunsche, & Sczakiel, 2004). Because of their potency, the low concentration of RNAi compounds used in experiments has favored sensitive molecular biology techniques for their detection and quantification. The bioanalytical techniques used for the study of antisense compounds are applicable to RNAi compounds and these have recently been reviewed (Yu, Geary, & Levin, 2004).

Molecular biology-based methods for quantification of miRNA and siRNA include ELISA assays and hybridization assays with various detection schemes such as bioluminescence and radiolabeled probes (Overhoff, Wunsche, & Szakiel, 2004; Cissell et al., 2008). In addition, a variety of PCR-based methods such as primer-extension qPCR (Raymond et al., 2005), stem-loop RT-PCR (Chen et al., 2005; Cheng et al., 2009; Stratford et al., 2008), Universal Probe Library RT-PCR (Varkonyi-Gasic et al., 2007; Wu et al., 2007), polyadenylation followed by RT-PCR (Shi & Chiang, 2005), ligation methods (Duncan et al., 2006), “crook” siRNA (Jiang et al., 2005), and splinted ligation (Maroney et al., 2007) have also been shown to be effective in quantifying miRNA and siRNA. The detection limits of these different methods range from the low copy number with PCR to the femtomole and low attomole levels with bioluminescent and radiolabeled hybridization methods.

All molecular biology-based methods for detecting oligonucleotides involve hybridization of a probe or a capture and/or detection strand that is complementary to the target miRNA or oligonucleotide sequences have a greater ESI signal due to their outside position in ESI droplet versus hydrophilic sequences which prefer an internal location. This behavior has also been observed in the ESI spectra of duplex siRNA (Beverly, Hartsough, & Machemer, 2005).

![FIGURE 3. Common chemical modifications made to oligonucleotides to improve resistance to nuclease degradation.](image-url)
siRNA. The result is that the duplex is never analyzed intact, only the separate strands are detected, and the performance of these assays is always dependent on the sequence and sequence-related factors such as $T_m$ and secondary structure. In addition, due to the inherent probability that a mismatched or truncated target will hybridize, these assays can have problems with specificity (Stratford et al., 2008). This becomes a concern when examining RNAi compounds that have been exposed to the nucleases that are ubiquitous in body fluids and tissues. Nuclease cleavage can occur in a variety of locations along the oligonucleotide sequence and can produce a single nucleotide or $N - 1$ truncation where $N$ is the full-length oligonucleotide and the $-1$ signifies a loss of one nucleotide. Alterations of the oligonucleotide such as truncations, phosphorylation, oxidation, depyrimidation, and depurination that are a result of metabolic processes or that arise during the oligonucleotide’s synthesis may not be differentiated with molecular biology-based methods. Such cross-reactivity of an assay to metabolic truncations or alterations would therefore result in an overestimation of concentration. As strand loading into RISC and mRNA target recognition are both sensitive to the 5' end of the oligonucleotide, truncated strands could have a different activity than full-length strands. It is the ability to identify these various oligonucleotide species and to do so in a manner that is independent of sequence that separates MS from other analytical approaches.

C. Challenges of siRNA Analysis by Mass Spectrometry

While the MS techniques used for siRNA analysis are the same as those for other classes of oligonucleotides, the double-stranded nature of siRNA creates some unique analytical challenges. If analysis of the intact duplex is the goal, preserving the dsRNA duplex intact during sample extraction, a chromatographic step (in the case of LC–MS) and mass spectral analysis can be difficult. Conversely, denaturing the duplex and analyzing the two individual strands have their own difficulties such as mass redundancy and/or co-elution of the two strands and their respective impurities. Complicating matters is the fact that dsRNA compounds can contain twice as many possible analyte species as single-stranded oligonucleotides (Fig. 4). These can consist of synthesis impurities or can be products of metabolism and degradation. In the case of therapeutic RNAi compounds, the single strands are synthesized separately and independently, and then annealed to form the duplex. As a result, synthesis impurities in the single strands can be carried over into the duplex. For example, synthesis impurities in the original oligonucleotides or nuclease degradation of the full-length duplex could result in a duplex where one of the two strands is full length and the other is a truncation. Chromatographic resolution of this truncated duplex species from the full-length could be difficult but has been demonstrated (McCarthy, Gilar, & Gebler, 2009). The mass differences of these truncated duplexes can be detected but even with adequate chromatographic resolution, mass analysis is not without its own set of problems.

A mass redundancy or a 1 Da or less mass difference could exist between the many possible species shown in Figure 4. The mass difference between C and U is 1 Da as is the mass difference between AT and GC base pairs. And when examining the mass of the intact duplex, A to T or C to G substitutions do not cause any mass shift. High mass accuracy and on-line sequencing via MS/MS would help to clarify these ambiguous situations and are two of the advantages that MS brings to oligonucleotide analysis.

Another challenge for MS-based methods is providing the sensitivity needed for bioanalytical studies of RNAi oligonucleotides. This has been a problem for all therapeutic oligonucleotides regardless of their class and adequate sensitivity for PK/PD studies, particularly during the elimination phase, has been difficult. For example, quantitative LC–MS and LC–MS/MS methods have been established for antisense oligonucleo-

FIGURE 4. Synthetic dsRNA could exist in several forms after synthesis, formulation, and metabolism. In addition to the desired duplex, there could be an excess of one strand and impurities or truncations on either strand due to synthesis failures or nuclease cleavage.
tides in tissue and plasma and these have a linear range from the low nanogram to low microgram per milliliter (Yu, Geary, & Levin, 2004; Dai et al., 2005; Murphy et al., 2005; Lin, Li, & Dai, 2007; Zhang et al., 2007). Poor extraction efficiencies from plasma or tissue, strong cation addition, a wide charge state envelope in the case of ESI, and a large molecular weight (usually between 7 and 8 kDa) all contribute to reducing MS sensitivity.

Finally, the use of chemical modifications to the 2’ ribose may pose a challenge to MS-based sequencing efforts. It has been mentioned previously that chemical modification of the 2’ ribose position is a common strategy to add nuclease resistance and thereby improve the PK properties of therapeutic oligonucleotides. Studies on the gas-phase sequencing of RNA and DNA indicate that 2’ ribose chemical modifications hinder the fragmentation pathways that produce sequence information from MALDI and electrospray (Ono, Scaife, & Smith, 1997; Tang et al., 1997; Wang et al., 1997; Sannes-Lowery & Hofstadler, 2003; Tromp & Schurch, 2005). Separate observations by Sannes-Lowery using infrared multiphoton dissociation and Bing with MALDI have indicated that while the intensity of sequence defining fragment ions is greatly diminished with 2’ MOE and 2’ O-methyl modifications, there is enough information to determine the sequence (Wang et al., 1997; Sannes-Lowery & Hofstadler, 2003). Also, work from Hakansson indicates that fragmentation by electron detachment dissociation (EDD) might yield sequence information from 2’ modified oligonucleotides as EDD fragmentation appears to cleave the phosphate backbone without involving the 2’ hydroxyl.

III. MS ANALYSIS OF RNAi OLIGONUCLEOTIDES

A. Desalting

The polyanionic nature of oligonucleotides results in significant cation adduction which can degrade MALDI or ESI spectra by dividing the ion signal between the adduct species. The result is a broad peak and a low signal-to-noise ratio. Regardless of which ionization method is used, to generate good quality spectra, steps must be taken to ensure that sodium, potassium, or other cation adducts are removed from the oligo. The desalting strategy has been to replace the non-volatile cation adduct with a volatile one, usually in the form of an ammonium ion (Nordhoff et al., 1992; Pieles et al., 1993). This technique has added benefit of preserving the cations required to stabilize duplex oligonucleotides.

Ammonium salts such as ammonium citrate and ammonium acetate are commonly added to the MALDI matrix for desalting purposes (Currie & Yates, 1993; Pieles et al., 1993; Fitzgerald & Smith, 1995; Sauer, 2007). Organic bases added to the matrix such as tetramine spermine have also been successful at controlling cation adduction (Asara & Allison, 1999; Distler & Allison, 2001). The most common MALDI matrix for oligonucleotides has been 3-hydroxypicolinic acid (3-HPA) with the additive diammonium hydrogen citrate (DAHC) (Pieles et al., 1993). Other matrices used for oligonucleotide analyses include 6-aza-2-thiothymine (ATT), 2,4,6-trihydroxycetophenone (THAP), and recently diaminobenzophenone (Lecchi, Le, & Pannell, 1995; Lavanant & Lange, 2002; Fu et al., 2006). In addition to cations from the oligonucleotide, the matrix itself can contain sodium and potassium. As a result, cation exchange resins are often mixed with matrix to prevent the matrix from adding salt to the sample (Nordhoff et al., 1992).

The addition of cation chelators to the buffers used for electrospray analysis is also effective at reducing cation adduction. EDTA or organic bases such as imidazole, piperidine, and triethylamine (TEA) have been added to the aqueous sample and/or mobile phase to desalt oligonucleotides prior to ESI (Stults & Marsters, 1991; Greig & Griffee, 1995; Limbach, Crain, & McCloskey, 1995b; Muddiman et al., 1996; Holzl et al., 2005). On-line methods that desalt the oligonucleotide as it elutes into the MS have also been used successfully. These methods include cation exchange and reversed-phase columns as well as microdialysis and size exclusion gels (Liu et al., 1996, 1997; Huber & Buchmeiser, 1998; Xiang et al., 1999; Cavanagh et al., 2003; Fountain, Gilar, & Gebler, 2004; Holzl et al., 2005). It is important to recognize that some of the best solution phase systems for desalting and producing strong ESI signal are not optimal for ion-pair reversed-phase chromatography as the high concentration of organic base will prevent stationary phase interaction and can also interfere with on-line UV analysis.

B. LC–MS of Oligonucleotides

While specific examples of miRNA and siRNA are lacking, a large body of work on the LC–MS analysis of oligonucleotides exists. A variety of chromatographic methods have been used to separate oligonucleotides including anion exchange, reversed-phase, capillary electrophoresis, and size exclusion separation (Baba, 1993; Gilar et al., 1997, 1998; Hernandez-Borges et al., 2004). Because the phosphodiester backbone is negatively charged at most solution pH ranges, anion exchange has traditionally been the most common chromatographic mode used for oligonucleotide separation. However, due to the incompatibility of electrospray ionization with the non-volatile salts and high ionic strength buffers used in anion exchange chromatography, the majority of LC–MS analyses are done with ion-pairing reversed-phase (RP-IP) chromatography. Excellent reviews on the LC–MS analysis of oligonucleotides are given by Huber and Oberacher (2001) and also by Gilar et al. (2003).

To improve reversed-phase performance an ion-pairing agent such as triethylammonium acetate (TEAA) is added to the mobile phase. An ion-pairing agent is an amphiphile and possesses a charged or ionic group on one end and a hydrophobic tail on the other (Bartha & Stahlberg, 1994). For oligonucleotide analysis the ionic group is a cationic amine and common ion-pairing agents are TEAA, tributylammonium acetate, and other trialkylammonium salts. Separation of truncation species that differ from the full-length oligonucleotide by one nucleobase (termed N – 1) is routinely achieved by RP-IP chromatography at the 18–21 bp length of RNAi species (Gilar, 2001).

TEAA, the most common ion-pairing agent used for the reversed-phase separation of oligonucleotides, can suppress electrospray ionization and is not ideal for LC–MS (Bleicher & Bayer, 1994; Apffel et al., 1997a,b; Gilar et al., 2002). In response to the incompatibility between electrospray ionization and
TEAA buffers, an ion-pairing buffer system employing hexafluoroisopropanol (HFIP) and TEA was developed by Apffel et al. (1997a,b). This buffer system provided chromatographic performance equal to TEAA but had greatly improved electrospray behavior. The addition of TEA to the HFIP buffer system has the benefit of increasing the pH to favor negative ionization as well as controlling cation adduction. Since its introduction, the HFIP/TEA buffer has become commonly used for LC–MS analysis of oligonucleotides (Huber & Oberacher, 2001). Gilar et al. have studied the TEAA and TEA/HFIP ion-pairing buffer in detail and have shown that in addition to its beneficial characteristics for ESI, the TEA/HFIP buffer also appeared to yield better chromatographic performance than TEAA (Gilar, Belenky, & Wang, 2001; Gilar et al., 2002, 2003; Fountain, Gilar, & Gebler, 2003). Consistent with Apffel et al., Gilar found optimal separation of oligonucleotides occurred at a concentration of 400 mM HFIP with 16.3 mM TEA as the organic solvent. Along with these empirical studies, equations have been developed that predict the behavior of single-stranded oligonucleotides in ion-pair reversed-phase chromatography based on their sequence and ability to form secondary structures (Gilar et al., 2002; Gilar & Neue, 2007; Sturm et al., 2007).

TEA/HFIP is not the only useful ion-pairing buffer for the LC–MS analysis of oligonucleotides. Huber and Oberacher (2001), Oberacher, Niederstatter, and Parson (2005), and Premstaller, Oberacher, and Huber (2000) have successfully used butyldimethylammonium bicarbonate (BDMAB) and triethylammonium bicarbonate (TEAB) with monolithic capillary columns to separate and mass analyze a variety of single- and double-stranded oligonucleotides. The review by Huber and Oberacher (2001) summarizes the performance characteristics of this system. Monolithic columns were reported to have better separation efficiency than traditional bead-based columns for single- and double-stranded DNA; however, the bead sized that was compared was not the newer sub-2 μm particles which may have better performance (Premstaller, Oberacher, & Huber, 2000; Oberacher et al., 2004c). Holzli et al. (2005) used a BDMAB buffer with a monolithic capillary LC–MS system to analyze a 120mer RNA oligonucleotide in the 50–80 fmol range as well as perform MS/MS on-line sequencing of a 32mer RNA at low picomole levels. Oberacher, Niederstatter, and Parson (2005) evaluated the performance of a similar LC–MS system coupled to a time-of-flight (TOF) mass spectrometer to characterize various lengths of single-stranded DNA. Using this system with BDMAB buffer, a full scan (800–2,400 m/z) lower detection limit of 6.5 fmol/μL of a dT 18mer was achieved.

Most of the reports in the literature concern single-stranded oligonucleotides and little work has been done to examine the chromatographic behavior of intact double-stranded oligonucleotides at the 19–22mer length of siRNA. Recently, McCarthy et al. (2009) demonstrated that ion-pair reversed-phase chromatography can separate a full-length 21mer siRNA duplex from duplexes where one of the strands contained an N–1 truncation. In addition to highlighting the ability of this system to chromatographically resolve these duplex species, this work also employed an ion-pairing buffer, hexylamine acetate, which was compatible with electrospray ionization.

C. ESI of Duplex Oligonucleotides

Starting with the first report by Light-Wahl et al. (1993) on the ability of ESI to preserve duplex DNA 20mers in the gas phase, a variety of non-covalent oligonucleotide interactions have been examined. This has been a fertile area of research and the reader is directed to some excellent reviews covering these subjects (Sannes-Lowery et al., 2000; Beck et al., 2001; Rusconi, Guillonneau, & Praseuth, 2002; Hofstadler, Sannes-Lowery, & Hannis, 2005). Important to siRNA analysis by MS is the body of work demonstrating that the Watson–Crick base stacking and hydrogen bonding of duplex oligonucleotides in solution are preserved in the gas phase (Schnier et al., 1998; Gabelica & De Pauw, 2002; Gidden, Baker, & Ferzocco, 2004) (Pan, Sun, & Lee, 2006a).

For example, several studies have shown that the collision-induced dissociation energies of duplex DNA and RNA correlate with solution phase Te data (Gabelica et al., 2000; Gabelica & Pauw, 2001; Yang, Thompson, & Hall, 2004). Although unlike the solution phase where base stacking is the major interactive force between strands, it appears that hydrogen bonding between bases is the dominant force in the gas phase (Yang, Thompson, & Hall, 2004; Pan, Sun, & Lee, 2006a,b).

The majority of the ESI studies on oligonucleotide duplexes have been conducted on DNA and using direct infusion of sample via a syringe pump. A volatile salt such as ammonium acetate at 10–50 mM with a low percentage of organic solvent is the most commonly used buffer for examining duplex oligonucleotides (Bayer et al., 1994; Loo, 1997; Hofstadler & Griffey, 2001; Rodrigues Hoyne et al., 2001). As mentioned earlier, while these buffer systems are ideal for preserving non-covalent complexes they are not always amenable to chromatography.

Beverly, Hartsough, and Machemer (2005) used the HFIP/TEA ion-pairing buffer system to separate duplex siRNA from its single-strand components. The 400 mM HFIP/16.3 mM TEA with methanol buffer system enabled the intact siRNA to be preserved in the gas phase during electrospray analysis (Fig. 5A,B). Some denaturation of the duplex was observed but it appears that this was a result of the electrospray process, not the mobile phase conditions, as there were no peaks in the duplex chromatogram corresponding to the single strands alone. In addition to preserving the duplex, the HFIP buffer also produced several high charge states (−10, −11, −12) that were helpful when de-convoluting spectra from a low-resolution instrument. Traditional buffers used for the electrospray of non-covalent complexes do not have a large charge-state envelope so accurate de-convolution back to neutral mass can be difficult without isotopic resolution in the raw spectrum.

D. MALDI of Duplex Oligonucleotides

MALDI observations of intact duplex oligonucleotides have been more difficult due to factors such as the acidic matrix, the crystal formation process, the organic solvent present in the matrix, and the ionization process itself, all of which can contribute to denaturing the non-covalent bonds that exist between strands. MALDI spectra of oligonucleotides can be obtained in either positive or negative ion modes. De-protonation of the phosphodiester backbone produces negative ions, whereas positive ions result from protonation of the nucleobases and
possibly the phosphodiester (Draper, 2004). The strongly acidic matrices that allow positive ion formation have a greater tendency to denature duplex oligonucleotides. As a result, most intact duplex work employs neutral matrices and is carried out in negative mode (Kirpekar, Berkenkamp, & Hillenkamp, 1999). Lecchi, Le, and Pannell (1995) were the first to observe duplex DNA (12mers) with negative mode MALDI using the neutral matrix ATT with ammonium citrate. Importantly, they noted that the duplex retained specific Watson–Crick base pairing in the gas phase. They also noted that duplexes were not observed using the common oligonucleotide matrices of 3-HPA or THAP which are acidic. However, Little et al. (1997) successfully used positive mode MALDI with 3-HPA to observe DNA duplexes of up to 50 bp by employing lower acceleration voltages and decreased sample preparation temperatures. Using glycerol or ATT, and either IR or UV MALDI, Kirpekar, Berkenkamp, and Hillenkamp (1999) were able to analyze DNA duplexes from 12 bp to 1.5 kbp. As with Lecchi, Kirkpekar et al. observed that DNA duplexes smaller than 12mers did not remain intact so it appears that there is a lower limit to the duplex that can remain intact during MALDI. Also, the addition of ammonium acetate or Tris–HCl was found to be required to preserve duplexes. Other additives such as cobalt (III) hexamine with ATT matrix have been used to stabilize DNA duplexes (Distler & Allison, 2002, 2003). Using this approach, Distler and Allison were able to observe duplexes of up to 12mers; however, the ability of more than one additive molecule to bind per duplex created multiple peaks which could complicate the spectra. Sudha and Zenobi (2002) investigated 3-HPA, ATT, and THAP and found that ATT with ammonium citrate best preserved DNA duplexes at a matrix to analyte ratio of 1,500:1 (negative mode).

The majority of MALDI work has examined DNA duplexes. Fortunately, for RNAi analyses, RNA is less prone to MALDI-induced fragmentation than DNA and is also inherently more thermodynamically stable than DNA duplexes (McLuckey, Van Berkel, & Glish, 1992; Kirpekar et al., 1994; Tang et al., 1997). Both of these factors should make MALDI analyses of siRNA easier than DNA duplexes.

Bahr, Aygun, and Karas (2008) have investigated the MALDI conditions needed to analyze siRNA as an intact duplex and used the method to determine the relative quantity of duplex and single strands. Using ATT with DAHC in water as the matrix, dsRNA 21mer duplexes with a variety of GC ratios were examined. A concentration of 100 mM DAHC and 2–10 μM of duplex was found to produce the maximum duplex signal intensity. Significantly, Bahr et al. found that the matrix must be in pure water to allow detection of the intact siRNA (Fig. 6). The addition of acetoni trile to the matrix, which is commonly done to improve ATT solubility, eliminated the signal for the intact duplex. In order for the MALDI approach to provide the relative percentage of intact duplex a single-stranded internal standard was used. This approach was needed due to differences in single- and double-strand MALDI ionization efficiencies, similar to what has been observed when using ESI.
APPLICATIONS OF MASS SPECTROMETRY

FIGURE 5. (Continued)
A linear range of 0.13–10 μM was established for the internal standard, and the relative amount of intact siRNA duplex between 94% and 97% was determined. Also using this method, excess amount of single strand was determined and duplexes constructed from $N-1$ truncation impurities could be observed. The development of a MALDI method to rapidly determine the amount of duplex siRNA will be important to the manufacture and quality control of synthetic siRNAs.

IV. APPLICATION OF MS TO THE STUDY OF siRNA AND miRNA

The therapeutic potential of RNAi has made the study of siRNA metabolism and degradation the most common application of MS. Due to the soft nature of ESI and MALDI, siRNA can be analyzed as denatured single strands or as an intact duplex and there are advantages and disadvantages to each strategy.

Denaturing the duplex is easily achieved by heating the column above the $T_m$ of the duplex or using an acidic MALDI matrix. Chromatographic resolution of single strands is expected to be better due to the greater mass transfer characteristics of the smaller single strands as well as at the higher temperatures used to denature the duplex. The performance of most mass analyzers is also better in the lower molecular weight range of the single strands than for the larger duplex. In addition, elevated temperatures prevent RNA from having secondary structure, which can complicate chromatographic and MS performance. For example, desalting of the single strands was found to be easier than the duplex (Beverly, Hartsough, & Machemer, 2005). Finally, when examining the duplex, the chromatographic and MS systems have to be constrained by a set of conditions (buffers, column temperatures, matrices, ion-optics, ion-source temperatures) that

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{MALDI mass spectra of siRNA double-stranded TN1-GC-ds (a) under denaturing conditions, with HPA as matrix, (b) with ATT as matrix and 20 mM DAHC, and (c) under optimized MALDI conditions using ATT/100 mM DAHC as matrix. Reproduced from Bahr, Aygun, and Karas (2008) with permission from American Chemical Society (Copyright 2008).}
\end{figure}
preserve the duplex but may compromise sensitivity/instrument performance. Analysis of the duplex in its natural state is therefore more difficult for the reasons given above. The advantage, however, to keeping the duplex intact is that it provides direct information as to how various single-strand species are paired up. This information can be important when assessing the amount of duplex formation for quality control purposes or can be used to offer insight into the mechanism of duplex metabolism.

It is worthwhile to recognize that the use of MS as an analytical tool for quality control and impurity identification during the synthesis and manufacturing of oligonucleotides is common and it serves a vital role in this area. The literature contains many examples (far to many to cite here) where MS has been critical to the identification of a synthesis impurity (Ravikumar et al., 2003; Capaldi et al., 2004; Gaus et al., 2005; Oberacher, Niederstatter, & Parson, 2005; Rentel et al., 2005; Kurata et al., 2006) and as the recent work by Bahr et al. demonstrates, MS will continue to be an important analytical tool to assess siRNA quality.

A. Sequencing

The use of MS in sequencing oligonucleotides is a routine practice and can be accomplished by either directly sequencing the oligonucleotide in the gas phase or by examining the products of enzymatic or chemical reactions designed to produce a sequence ladder. There are numerous examples of MS-based sequencing and the reader is directed to an excellent review for more information (Limbach, 1996). While sequencing oligonucleotides is common, the duplex nature of siRNA does make the analysis more complicated as there exists the possibility of mass overlap and confusion of the ladder or gas-phase fragment ions between the two strands. Recently, both MALDI and ESI have been used to analyze the sequence ladder from siRNA single strands.

In two reports by Farand and Beverly (2008) and Farand and Gosselin (2009), the chemical approach combined with accurate mass and MS/MS successfully sequenced oligonucleotides with multiple 2′ fluoro and 2′ O-methyl modifications with one of these siRNAs being sequenced de novo. In this work six different chemical reactions were found to produce acid or base cleavage products from the modified sequences. The pattern of cleavage from each of these reactions was used collectively to reconstruct the sequence. The siRNA was analyzed as separate single strands as there was the potential that digestion of the duplex would produce a complex mixture of reaction products that could not be unambiguously assigned to an individual strand.

Another recent report by Bahr, Aygun, and Karas (2009) demonstrated the use of acid hydrolysis followed by MALDI to analyze the ladder sequence both single- and double-stranded RNA. Bahr was able to sequence both strands of the duplex simultaneously by using the molecular weight of each full-length sequence as a starting point for subsequent mass losses and thereby avoid mixing fragments from the two strands. This method did not cleave RNA with 2′ O-methyl and would not be expected to work for 2′ ribose modifications as the hydrolysis reaction requires a hydroxyl group on the 2′ position. Enzymatic or chemical approaches have not previously been used to examine duplex oligonucleotides directly as the reaction products from each of the strands will mix. Direct gas-phase sequencing of a duplex oligonucleotide without first isolating each strand avoids a possible mix up and can be achieved in several ways. With high temperature LC, the duplex can be denatured and separated on-column where the two strands then elute into the mass spectrometer for gas-phase fragmentation. The ionization process itself, MALDI or ESI, denatures the duplex so that the two strands, now isolated from one another in the mass spectrometer, can then be sequenced using characteristic gas-phase fragmentation patterns.

An example of this direct approach was recently reported by Huang et al. (2008) who sequenced an RNA duplex that was electrospayed intact. During ESI, the intact siRNA duplex was observed and dissociated; however, the resulting CID spectra of the intact duplex were too complicated to be used for sequencing. As a result, the ions corresponding to the dissociated sense and antisense strands were used for sequencing. The use of proton-transfer reactions to decrease the charge states simplified the MS/MS spectra and made spectral interpretation easier. In addition, manipulation of the ion trap collisional activation energy enabled the fragmentation pathway of the siRNA single strand to produce either c/y ions or a-B/w ions that resulted in complete sequencing of the sense and antisense strands.

B. Metabolism/Degradation

Oligonucleotides are degraded in vivo by nucleases that cleave the phosphodiester linkage. Cleavage can occur at the ends of the strand and proceed sequentially inward (exonucleases), can occur internally (endonuclease), or can be a combination of both. The metabolism of single-stranded antisense therapeutic oligonucleotides has been studied extensively by LC–MS due to its ability to identify metabolites by their unique mass, retention time, and sequence (Gaus et al., 1997; Gilar et al., 1997, 2003; Lin, Li, & Dai, 2007; Zhang et al., 2007). Understanding the exact location of oligonucleotide degradation is important as sites in the sequence that are susceptible to enzymatic digestion can be identified and then chemically modified to improve resistance. A priori knowledge of the sequence makes identification of nuclease-derived metabolites possible by mass alone as these metabolites often have unique masses. In cases where the masses of metabolites are not unique, identification can be achieved by MS/MS sequencing or high-resolution mass measurement. For example, on-line MS/MS sequencing has been used successfully to further confirm the identity of metabolites from antisense oligonucleotides (Griffey et al., 1997; Dai et al., 2005). Recently, the identification of siRNA metabolites has been reported using both MALDI and LC–MS techniques (Beverly, Hartsough, & Machemer, 2005; Turner et al., 2007; Zou et al., 2008).

Using HFIP/TEA with reversed-phase chromatography and ESI, the metabolism of siRNA in urine and ocular vitreous humor was examined and compared to its single-stranded components (Beverly, Hartsough, & Machemer, 2005). As mentioned earlier, the HFIP/TEA buffer allowed the intact duplex to be chromatographically separated from single strands as well as preserving
the duplex in the gas phase. Degradation of duplex siRNA was slower than single strands confirming that the duplex structure provides protection from nucleases. In addition, the observed metabolites were consistent with the chemical modifications made to prevent nuclease digestion.

Ocular metabolites of a therapeutic siRNA targeting the VEGF receptor were examined by LC–MS with HFIP/TEA at various time points after being dosed into rabbits (Beverly et al., 2006). Samples ranging from 10 to 100 μg/mL of duplex were extracted from ocular vitreous humor and retinal tissue using phenol/chloroform extraction. The siRNA was analyzed as an intact duplex as well as denatured into single strands by heating the reversed-phase column. The chromatographic resolution of the intact duplex was much poorer than the single strands, and the sensitivity of detecting the intact duplex was 5–10 times less than for the single strands. Metabolism products of the denatured duplex as well as the intact duplex revealed a degradation pattern in which the end of the duplex with the weakest interstrand

![Diagram of siRNA sequence and metabolites](image)

**FIGURE 7.** De-convoluted ESI mass spectrum of vitreous extract taken at Day 1 and Day 4. The intact duplex is observed at mass 13,678 and the two single strands (marked S for sense and AS for antisense) are at mass 7,014 and 6,663. Day 1 shows full-length duplex and single strands. The Day 4 spectrum shows two duplexes made up of metabolites at mass 12,886 and 11,587. Also, visible are the single strands that make up the metabolite duplexes at mass 6,222 and 5,364. The siRNA sequence is given above with the arrows indicating the direction of observed nuclease degradation. Reproduced from Beverly et al. (2006) with permission from John Wiley & Sons, Ltd (Copyright 2006).
FIGURE 8. MALDI-TOF mass spectrum of an siRNA duplex (50 μM) incubated with 10% mouse serum for 5 days using THAP and diammonium hydrogen citrate (1:1). Peak identities are listed in the table. Sequence of duplex with the following key: circled U indicates 2',O-methyl modification on ribos, idT indicates thymidine, cP indicates 2',3' cyclic phosphate, and P indicates 2' or 3' phosphate. Reproduced from Turner et al. (2007), with permission from the Royal Society of Chemistry (Copyright 2007).
The binding energy was most susceptible to degradation (Fig. 7). This result suggested that the end of the duplex that is most able to "breath" or temporarily denature into single strands is more likely to degrade.

Using MALDI-TOF, Turner et al. (2007) studied the metabolism patterns of various siRNAs spiked into plasma. Using a matrix of 20 mg/mL THAP diluted 1:1 with 50 or 80 mg/mL DAHC, Turner et al. were able to analyze siRNA from 90% plasma with little interference. siRNA was not observed in the duplex form but was denatured into its single strands. The majority of siRNA cleavages occurred after pyrimidine residues, specifically AU, supporting the hypothesis that degradation of siRNA is via an RNAse A endonuclease (Fig. 8). The degradation patterns observed also suggested that "breathing" at the ends of the duplex may make it more vulnerable to degradation.

High-resolution instruments that provide accurate mass have been used to identify in vitro siRNA metabolites based on prior knowledge of the sequence. Metabolites of a chemically

![Figure 9](image-url)  
**FIGURE 9.** An example of metabolite identification using accurate mass measurements. The inset shows the isotopic spectrum of a metabolite from the test sequence compared with spectra of two possible metabolites. The profile identifies the metabolite as S(N-10)³, but not AS(N-9)⁵ + P. Reproduced from Zou et al. (2008), with permission from John Wiley & Sons, Ltd (Copyright 2008).

![Figure 10](image-url)  
**FIGURE 10.** In vitro assembly of the human RLC. Electrospray MS of the intact RLC complex.  
**Lower:** Mass spectrum of the intact heterotrimeric RLC complex (purple star) confirms the presence of Dicer, Ago2, and TRBP subunits with unit stoichiometry. The charge-state series allows the molecular mass to be measured as 371.129 ± 95 Da, indicating the presence of an additional component with a molecular mass 11 kDa higher than expected. Inset: This larger mass is likely due to an RNA bound to Ago 2, because the Ago 2 alone has a molecular mass of 109.950 ± 53 Da. At m/z 7,500, the Dicer/TRBP (orange circle) heterodimer also was observed (molecular mass of 259.212 ± 62 Da), indicating its formation in solution.  
**Upper:** MS/MS analysis confirms that the RLC complex is composed of TRBP (pink star), an 11-kDa RNA fragment (yellow circle), Dicer, and Ago2 (blue star) subunits. Reproduced from MacRae et al. (2008) with permission from the National Academy of Sciences, USA (Copyright 2008). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
modified siRNA in plasma and liver microsomes were analyzed by LC–MS using an Orbi-trap (Zou et al., 2008). In this report, siRNA was extracted from samples with phenol/chloroform followed by solid-phase extraction. High-resolution MS made it possible to distinguish between two metabolite species differing in mass by <1 Da (Fig. 9). The use of a commercial software package to semi-automate the de-convolution and identification of the nuclelease metabolites helped to reduce the data analysis time. The resulting siRNA metabolism profiles for liver microsomes versus plasma highlighted the different nucleases present in each matrix.

C. RNAi Mechanism

To obtain a clearer picture of how the enzyme DICER cleaves its dsRNA substrate, Rose et al. (2005) used ESI–MS to examine the cleavage patterns of various 25–30mer dsRNA constructs. It was found that asymmetric duplexes with two bases overhangs on the 3’ end produced predictable cleavage patterns. These insights allowed 27mer DICER–substrate RNA duplexes to be constructed that would produce predictable 21mer cleavage products and enable a specific strand to be loaded into RISC. The longer RNA duplexes created in this fashion produced a 100-fold greater knockdown than their 21mer siRNA counterparts (Rose et al., 2005). Later, Kubo et al. (2007) confirmed the DICER processing behavior observed by Rose and then used MALDI to determine that 27mer dsRNAs have an increased serum stability over 21–22mer siRNA.

In efforts to create light-activated siRNA, Shah, Rangarajan, and Friedman (2005) used ESI to examine siRNA modified with various photolabile groups. The effect of modifications made to the 5’ phosphate of the guide strand on siRNA potency was also examined (Shah & Friedman, 2007). ESI was used to identify and quantify various impurities on the 5’ guide strand to account for their role in siRNA potency (Shah & Friedman, 2007, 2008). Their results indicated that the modifications on the 5’ guide strand did not completely block knockdown thus preventing a true binary “on/off” light switching mechanism.

To further understand the RNAi mechanism, the RISC complex itself has been examined by non-denaturing ESI to determine the stoichiometry of the three proteins that make up the active complex (MacRae et al., 2008). The 371 kDa molecular weight complex that was observed showed a 1:1:1 ratio of the three proteins: Dicer, Ago2, and TRBP (Fig. 10). Interestingly, the analysis also revealed that there was an 11 kDa RNA contaminant in the mixture that was bound to the RISC complex.

MS has also been used to characterize miRNAs. Recently, Katoh et al. demonstrated that terminal 3’ adenylation of miR-122 occurs in the cytoplasm and that these additional nucleotides help stabilize miRNA. In these experiments, miR-122 from mouse liver was isolated and concentrated with reciprocal circulating chromatography and then analyzed by nano electro-spray (Katoh et al., 2009). The addition of nucleotides to the 3’ end of miRNA-122 was determined using a combination of accurate mass and MS/MS sequencing.

Similar types of experiments have been conducted for larger RNA sequences where unknown RNA isolated from cell culture has been sequenced on-line via LC–MS (Oberacher et al., 2004b, 2004; Holzl et al., 2005). Other researchers have used MS to identify the methylation of plant miRNA and recently, LC–MS analysis of miRNAs has found that even these small RNAs contain the post-transcriptional modifications previously observed in larger mRNA and rRNAs (Yu et al., 2005; Ohara et al., 2007). It is certain that the MS techniques pioneered by Limbach and McCloskey to identify RNA modifications will be critical to further understanding how miRNA is modified (Limbach, Crain, & McCloskey, 1994; McCloskey et al., 1999; Qiu & McCloskey, 1999; Meng & Limbach, 2006).

V. CONCLUSION

siRNA and miRNA oligonucleotides are being developed as therapeutics, used to provide insight into the RNAi mechanism, and screened to identify their roles in regulating gene expression. The advantages for using MS as the analysis platform for miRNA and siRNA are similar to those described by others for the analysis of therapeutic oligonucleotides, oligonucleotide complexes, and RNA (Beck et al., 2001; Hofstadler & Griffey, 2001; Huber & Oberacher, 2001; Meng & Limbach, 2006). These include the specificity of the analysis by providing both mass and sequence information, the ability to couple to chromatographic separation, and an analysis technique that is for the most part independent of sequence (unlike PCR or hybridization methods). In addition, the use of soft ionization techniques like MALDI and ESI to preserve the intact duplex makes MS especially attractive to the analysis of siRNA. Questions regarding how well the gas-phase observations of dsRNA reflect solution-phase behavior are still being addressed but it is clear that major characteristics such as Watson–Crick base pairing are preserved in the gas phase observations of dsRNA reflect solution-phase behavior are still being addressed but it is clear that major characteristics such as Watson–Crick base pairing are preserved in the gas phase whether the ionization is by MALDI or ESI. New MS techniques such as ion mobility, which can provide information on the secondary and tertiary structure of oligonucleotides, will undoubtedly offer unforeseen benefits to the study of siRNA and miRNA.

REFERENCES


Shah S, Friedman SH. 2007. Tolerance of RNA interference toward
Schnier PD, Klassen JS, Strittmatter EF, Williams ER. 1998. Activation
Sauer S. 2007. The essence of DNA sample preparation for MALDI mass
Sannes-Lowery KA, Drader JJ, Griffey RH, Hofstadler SA. 2000. Fourier
Sannes-Lowery KA, Hofstadler SA. 2003. Sequence confirmation of
Rose SD, Kim DH, Amarzguioui M, Heidel JD, Collingwood MA, Davis ME,
Rodrigues Hoyne P, Benson LM, Moschos SA, Lindsay MA, Gait MJ. 2007. MALDI-
Rentel C, Wang X, Batt M, Kurata C, Oliver J, Gaus H, Krotz AH, McArule
Ravikumar VT, Kumar RK, Capaldi DC, Cole DL. 2003. Synthesis of high