Automated De Novo Sequencing of Nucleic Acids by Liquid Chromatography-Tandem Mass Spectrometry

Herbert Oberacher
Institute of Legal Medicine, University of Innsbruck, Innsbruck, Austria

Bettina M. Mayr and Christian G. Huber
Instrumental Analysis and Bioanalysis, Saarland University, Saarbrücken, Germany

We present the first global computer-aided sequencing algorithm for the de novo determination of short nucleic acid sequences. The method compares the fragment ion spectra generated by collision-induced dissociation of multiply charged oligodeoxynucleotide-ions to the *m/z* values predicted employing established fragmentation pathways from a known reference sequence. The closeness of matching between the measured spectrum and the predicted set of fragment ions is characterized by the fitness, which takes into account the difference between measured and predicted *m/z* values, the intensity of the fragment ions, the number of fragments assigned, and the number of nucleotide positions not covered by fragment ions in the experimental spectrum. Smaller values for the fitness indicate a closer match between the measured spectrum and predicted *m/z* values. In order to find the sequence most closely matching the experimental spectrum, starting from a given nucleotide composition all possible oligonucleotide sequences are assembled followed by identification of the correct sequence by the lowest fitness value. Using this concept, sequences of 5- to 12-mer oligodeoxynucleotides were successfully de novo determined. High sequence coverage with fragment ions was essential for obtaining unequivocal sequencing results. Moreover, the collision energy was shown to have an impact on the interpretability of tandem mass spectra by the de novo sequencing algorithm. Experiments revealed that the optimal collision energy should be set to a value just sufficient for complete fragmentation of the precursor ion. (J Am Soc Mass Spectrom 2004, 15, 32–42) © 2004 American Society for Mass Spectrometry

The hyphenation of liquid chromatography (LC) to mass spectrometry (MS) is one of the most powerful methods for the characterization of oligonucleotides today [1]. In this context the combination of ion-pair reversed-phase high-performance liquid chromatography and electrospray ionization mass spectrometry (ICEMS) has emerged as a versatile tool for the analysis of single- and double-stranded nucleic acids ranging in size from a few nucleotides (nt) up to several hundred base pairs (bp) [2–4]. Although the occurrence of a sequence variation is predictable on the basis of molecular mass measurements [5–7], detailed sequencing information can only be obtained from selective decomposition of oligonucleotides by tandem mass spectrometry (MSMS).

The principles of gas-phase collision-induced dissociation (CID) reactions of oligonucleotides have been studied extensively [8–17] and the resulting product ion spectra are predictable on the basis of known fragmentation pathways. Consequently, MSMS of oligodeoxynucleotides has been successfully applied to characterize or confirm structures [11], to detect chemical modifications [2, 10, 18], and to identify sequence variations [14, 19–21]. However, the deduction of sequence information is time-consuming, highly technical, lacking strict objectivity due to reliance on human interpretation, and can be performed only in laboratories with extensive experience in MSMS. Hence, automation of procedures for interpretation of fragment ion spectra represents a prerequisite for the applicability of MSMS in the routine sequence analysis of nucleic acids.

In principle, attempts to develop an automated sequencing algorithm can follow either local or global search paradigms. McCloskey and coworkers showed that the former approach was highly efficient for the ab initio determination of unknown oligonucleotide sequences at approximately the 20-mer level and below [13, 22]. The so called simple oligonucleotide sequencer (SOS) works by extending from both the 5′ (a-B-ions) and 3′ (w-ions) ends ion series encoding the complete DNA sequence. Mass ladders are identified by sequen-

Published online November 14, 2003

Address reprint requests to Dr. H. Oberacher, Institute of Legal Medicine, University of Innsbruck, Müllerstrasse 44, Innsbruck 6020, Austria E-mail: Herbert.oberacher@uibk.ac.at

1044-0305/04/$30.00
doi:10.1016/S1044-0305(03)00708-6

Received July 2, 2003
Revised September 12, 2003
Accepted September 12, 2003

tially adding each of the four possible nucleotide masses and searching the spectrum for the best match of expected ions.

A global strategy was applied by a recently introduced comparative sequencing algorithm called COMPAS for the sequence verification as well for the detection and localization of point mutations in 5 to 51-mer oligonucleotides [23]. COMPAS is based on the comparison of the experimental spectrum to *m/z* values predicted by employing established fragmentation pathways from a known reference sequence. The closeness of matching between the measured spectrum and the predicted set of fragment ions is characterized by the fitness, which takes into account the difference between measured and predicted *m/z* values, the intensity of the fragment ions, the number of fragments assigned, and the number of nucleotide positions not covered by fragment ions in the experimental spectrum. Smaller values for the fitness indicate a closer match between measured spectrum and predicted *m/z* values. In order to substantiate the identity of investigated sequence and reference sequence, or to identify point mutations, the reference sequence is systematically permutated by incorporating all four possible nucleotides A, T, G, and C at each position in the sequence followed by identification of the correct sequence by the lowest fitness value. Despite its considerable success for sequence verification, the proposed sequencing strategy is not applicable for de novo sequencing since the bulk of the sequence has to be known for the unequivocal detection and localization of dissimilarities.

Therefore, extending this recent communication, we report here, as far as we know, the first fully automated global de novo sequencing algorithm for nucleic acids. Starting from a known nucleotide composition, all possible nucleic acid sequences and corresponding theoretical tandem mass spectra are generated and compared to the experimental spectrum. The degree of matching is reflected in the fitness value, and the correct sequence is indicated by the lowest fitness value. Using this concept, sequences of 5- to 12-mer oligonucleotides were successfully de novo determined from fragment ion spectra generated by LC-MSMS experiments. Moreover, the influence of the collision energy and unspecific fragmentation on the sequencing results was investigated. The algorithm was applied to the unequivocal identification of a synthetic 9-mer showing abnormal retention behavior in LC.

Experimental

Chemicals and Oligodeoxynucleotides

Acetonitrile (HPLC gradient-grade) was obtained from Merck (Darmstadt, Germany). Triethylamine (p.a.) was purchased from Fluka (Buchs, Switzerland). A stock solution (0.50 M) of triethylammonium bicarbonate (TEAB) was prepared by passing carbon dioxide gas (AGA, Vienna, Austria) through a 0.50 M aqueous solution of triethylamine at 5° C until pH 8.4 was reached. For preparation of all solutions, HPLC-grade water (Merck) was used. The synthetic oligodeoxynucleotides listed in Table 1 were ordered from Microsynth (Balgach, Switzerland) and used without further purification.

Table 1. Sequences, nucleotide compositions, and molecular masses M_r of oligodeoxynucleotides (1–6) analyzed in this study

No.	Sequence	Nucleotide composition	Length [nt]	M_r[Da]
1	CTGGC	$C_2G_2T_1$	5	1479.0
2	TTCTGGC	$C_2G_2T_3$	7	2087.4
3	ATTCTGGC	$A_1C_2T_3$	8	2400.6
4	CATTCTGGC	$A_1C_3G_2T_3$	9	2689.7
5	GACATTCTGGC	$A_2C_3G_3T_3$	11	3332.3
6	AGACATTTCTGGC	$A_3C_3G_3T_3$	12	3645.5

Capillary High-Performance Liquid Chromatography coupled to Electrospray Ionization Mass Spectrometry

The HPLC system consisted of a low-pressure gradient micro pump (model Rheos 2000, Flux Instruments, Basel, Switzerland) controlled by a personal computer, a vacuum degasser (Knauer, Berlin, Germany), a column thermostat made from 3.3 mm o.d. copper tubing which was heated by means of a circulating water bath (model K 20 KP, Lauda, Lauda-Königshofen, Germany), and a microinjector (model C4-1004, Valco Instruments Co. Inc., Houston, TX) with a 500 nL internal sample loop. ESI-MS was performed on a Finnigan LCQ quadrupole ion trap mass spectrometer (Thermo Finnigan, San Jose, CA) equipped with an electrospray ion source. The 60 × 0.2 mm i.d. monolithic capillary column was prepared according to the published protocol [2] and connected directly to the spray capillary (fused silica, 90 μm o.d., 20 μm i.d., Polymicro Technologies, Phoenix, AZ) by means of a microtight union (Upchurch Scientific, Oak Harbor, WA). A syringe pump equipped with a 250 μL glass syringe (Unimetrics, Shorewood, IL) was used for the post-column addition of a 3.0 μL/min flow of acetonitrile by means of a tee-piece. For analysis with pneumatically assisted ESI, an electrospray voltage of 3.4 kV and a nitrogen sheath gas flow of 40 arbitrary units were employed. The temperature of the heated capillary was set to 200° C. Total ion chromatograms and mass spectra were recorded on a personal computer with the Xcalibur software version 1.1 (Finnigan).

Mass calibration and coarse tuning were performed in the positive ion mode by direct infusion of a solution of caffeine (Sigma, St. Louis, MO), methionyl-arginyl-phenylalanyl-alanine (Finnigan), and Ultramark 1621 (Finnigan). Fine tuning for ESI-MS of oligodeoxynucleotides in the negative ion mode was performed according to reference [24]. For MSMS experiments, the isolation width and the relative collision energy were set to 4.0 mass units and 15–100%, respectively. Zero to 100%

relative collision energy corresponds to a high frequency alternating voltage for resonance excitation from 0 to 5 V maximum-to-maximum. Helium, present in the ion trap at a pressure of 0.1 Pa, served as collision gas.

Computer-Aided Data Interpretation

All calculations were performed on a personal computer under Windows XP operating system (AMD Athlon 1800 MHz, 256 MB RAM). Measured MSMS spectra of oligodeoxynucleotides were exported from the Xcalibur Software (Finnigan) as ASCII files. Automated comparison of measured and predicted spectra was performed with a program written in ActivePerl 5.6.1 (Active State Corporation, Vancouver, BC, Canada). A copy of the program for academic use is available upon request from the corresponding author. All data were collected, evaluated, and prepared for final output in Microsoft Excel for Windows XP.

Results and Discussion

De Novo Sequencing Strategy

The starting point of the sequencing strategy relies on the fact that high accurate molecular mass measurements enable the determination of the nucleotide composition of a DNA molecule [11, 25, 26]. In this context, Russel and coworkers demonstrated that a mass accuracy of better than ±1 ppm, which can be routinely achieved on time-of-flight (TOF) and Fourier transform ion cyclotron resonance (FTICR) instruments, allows the unequivocal assignment of the composition of 20-mers [26]. However, the knowledge of the composition alone is usually not sufficient. In many circumstances detailed sequence information is of even more importance.

The golden standard for the determination of DNA sequences is the fully automated Sanger sequencing method employing multiplexed capillary electrophoretic analysis [27]. Although Sanger sequencing is reliable and employs read lengths up to 1000 nt, this method can not be applied for the sequencing of small oligonucleotides due to the use of sequencing primers, which need to have a length of at least 15 to 20 nt. MSMS could close this gap as this method was shown to be highly suitable for the characterization of small nucleic acids. For example, the characterization of metabolites of oligonucleotides used as therapeutic agents for the treatment of cancer was only possible via tandem mass spectrometric sequencing [28, 29]. However, the deduction of sequence information from an MSMS spectrum is time-consuming, highly technical, lacking strict objectivity due to reliance on human interpretation, and can be performed only in laboratories with extensive experience in MSMS. Hence, automation of procedures for interpretation of fragment ion spectra represents a prerequisite for the applicability of MSMS in the routine sequence analysis of nucleic acids.

On account of this, we developed a fully automated global de novo sequencing algorithm, which involves the comparison of a measured MSMS spectrum to a set of fragment ion *m/z* values predicted from a reference sequence. The closeness of matching between the measured spectrum and the predicted set of ions is characterized by a value FS for the fitness, which takes into account the difference Δ between measured and predicted *m/z* values, the relative intensity I % of the fragment ions, the number K of fragments assigned, and the number M of nucleotide positions not covered by fragment ions in the experimental spectrum. The smaller the value for FS, the closer the match between measured spectrum and predicted *m/z* values. In order to find a sequence most closely matching the experimental spectrum, all possible sequences are assembled from a given nucleotide composition and the corresponding FS values are determined. The correct sequence is identified by the reference sequence having the lowest FS value.

Development of the Sequencing Algorithm

Figure 1 outlines the procedure for the computer-aided spectrum interpretation of MSMS spectra of oligodeoxynucleotides using the de novo sequencing algorithm. The input parameters are the oligonucleotide composition together with the charge state of the precursor ion on one side and a list of *m/z* values and relative intensities I % of the fragments in the experimental spectrum on the other side. Subsequently, the computer program generates a sequence i using the given composition information (Step 1, Figure 1) and calculates from the Sequence i a list of monoisotopic *m/z* values for the a-B- and w-ion series including all possible charge states from 1- up to the charge state of the precursor ion (Step 2, Figure 1). Then, the predicted *m/z* values and those obtained from the experimental spectrum are compared (Step 3, Figure 1). The comparison yields the number of fragment ions that could be assigned to predicted *m/z* values [K(i)], and the sum of non-covered a-B- and w-positions [M(i)]. Non-covered a-B-positions are those in the sequence for which no a-B-fragment of any possible charge state could be identified in the spectrum, whereas the non-covered w-positions are positions in the sequence lacking w-fragments. In order to be assigned or not the *m/z* values must or must not fall within a tolerance of ± mass deviation (Δ). Values for K(i) and M(i) are determined for a Δ value equivalent to 1.0 mass unit.

The next step of the algorithm involves the calculation of a match factor MF(i) representing the quality of the assigned *m/z* signals in terms of mass deviation and intensity (Step 4, Figure 1). The match factor is defined as the sum of all quotients of mass deviation Δ and intensity I (since the input value for intensity is given in percent, the value is multiplied by 100), averaged over

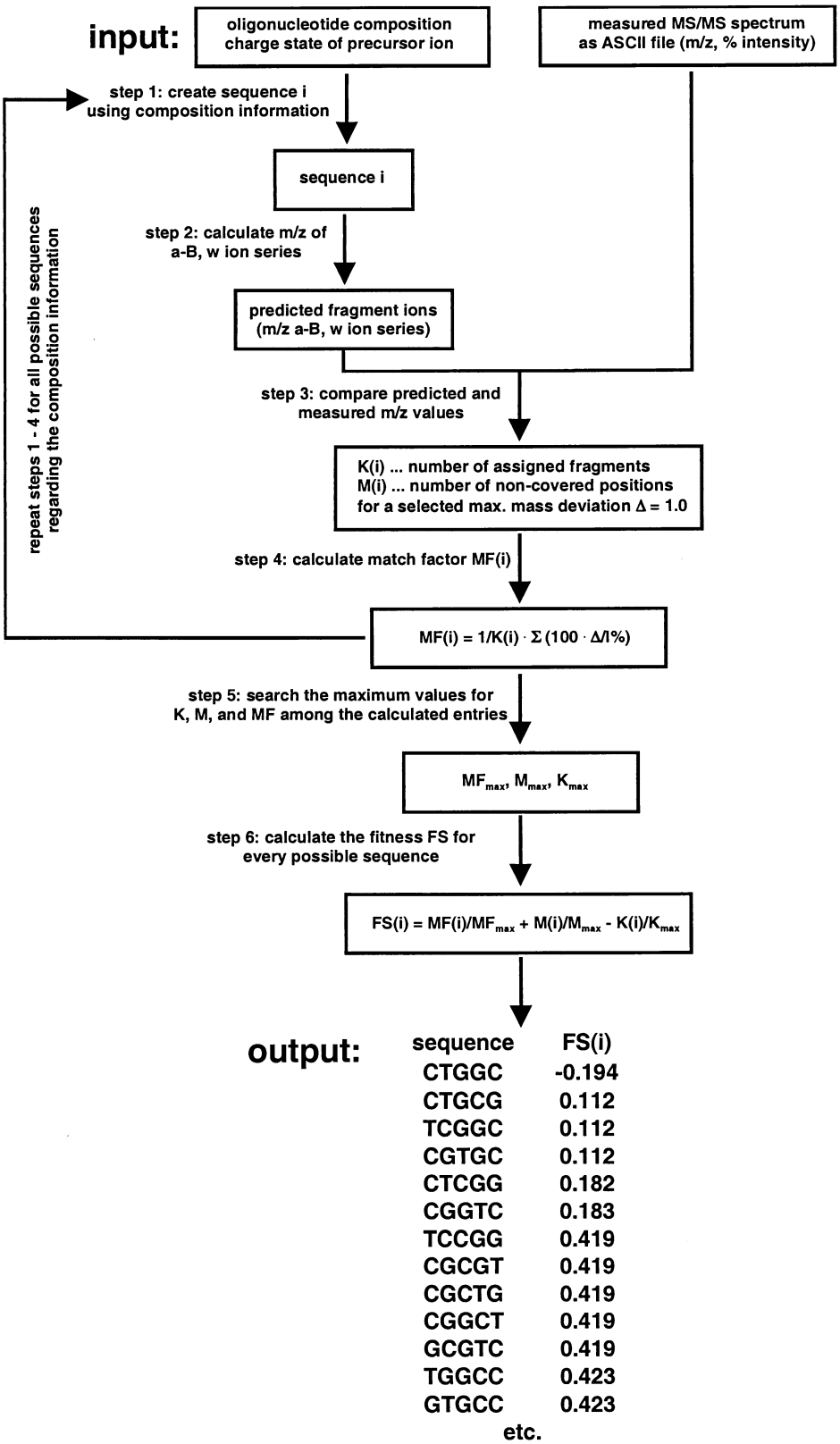

Figure 1. Outline of the steps involved in comparative sequencing by MSMS.

the total number of assignments K(i), using the eq MF(i) = $1/K(i) \cdot \Sigma\ (100 \cdot \Delta/I\ \%)$ (Figure 1). In order to find or exclude alternative sequences showing a better correspondence, the sequence is systematically altered fol-

lowed by matching the new sequence to the experimental spectrum. This is done by sequential permutation of any base with all other nucleobases. Assuming an oligonucleotide composition $A_aC_cG_gT_t$, the number of matches n, which have to be performed, can be calculated using the following eq: $n = (a + c + g + t)!/(a! \cdot c! \cdot g! \cdot t!)$. Obviously, the number of possible sequences grows rapidly with the length of the oligonucleotide.

Then the program automatically searches the maximum values of the three parameters match factor MF_{max}, number of assigned fragments K_{max}, and number of non-covered positions M_{max} among the calculated entries (Step 5, Figure 1). The three coefficients MF_{max}, K_{max}, and M_{max} are used to weigh the terms related to the individual features of the spectrum. Subsequently, these values together with the sequence specific match factor MF(i), number of assigned fragments K(i), and number of non-covered positions M(i) are used to calculate the fitness FS(i) of each sequence according to the eq $FS(i) = MF(i)/MF_{max} + M(i)/M_{max} - K(i)/K_{max}$ (Step 6, Figure 1). It is obvious, that the fitness becomes smaller with smaller mass differences, higher signal intensities, larger number of assigned fragments, and lower number of non-covered positions. The theoretical minimum for the fitness is −1. This value is reached if the sum of the non-covered a-B- and w-positions [M(i)] and the mass deviations (Δ) of all assigned fragments were equal to zero and if the number of assigned fragments K(i) were equal to K_{max}. However, this theoretical minimum cannot be reached experimentally. Even if complete sequence coverage is observed the limited accuracy of mass measurements will be responsible for fitness values larger than −1.

The final output of the algorithm is a list showing the FS(i) values of all permutated sequences (Figure 1). Generally, similar sequences have many fragment ions in common. For example, the exchange of two neighboring bases yields two nearly identical sequences, which have all but four fragment ions in common. However, the slight mass differences of these four specific fragment ions are sufficient to distinguish the two sequences and to achieve an unequivocal and correct sequencing result by applying the sequencing algorithm. Accordingly, as shown in Figure 1 the lowest FS value is usually obtained for the true sequence (sequence 1, Table 1) which provides the best correspondence to the experimental spectrum.

De Novo Sequencing of a 5-Mer Oligodeoxynucleotide

The success of MSMS sequencing largely depends on the purity of the sample that is presented to the mass spectrometer. On-line hyphenation of MSMS to chromatographic separation was chosen as sample preparation method, because it not only efficiently removes adducted cations and other low molecular mass contaminants coming from, e.g., solid-phase oligodeoxynucleotide synthesis or PCR, but also fractionates mixtures of nucleic acids prior to their mass spectrometric investigation. Separation was accomplished by ion-pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC) in a monolithic, 200 μm i.d. poly-(styrene-divinylbenzene) capillary column by application of gradients of acetonitrile in 25 mM triethylammonium bicarbonate [2, 30].

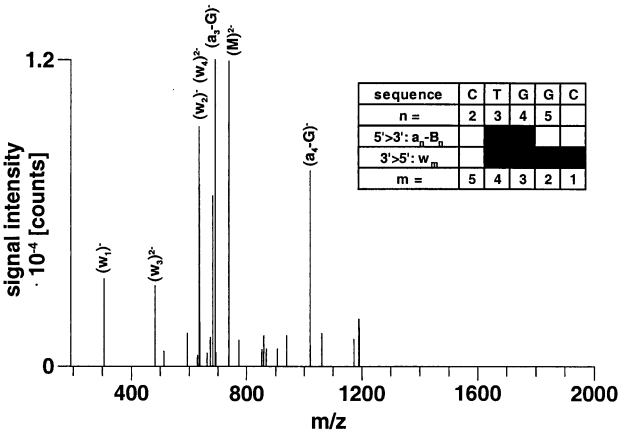

Figure 2. MSMS spectrum and sequence coverage diagram extracted from the LC-MSMS analysis of a 5-mer oligodeoxynucleotide. Column, monolithic PS-DVB, 60 × 0.20 mm i.d.; mobile phase, (**a**) 25 mM TEAB, pH 8.40, (**b**) 25 mM TEAB, pH 8.40, 20% acetonitrile; linear gradient, 5–50% B in 10.0 min; flow-rate, 3.0 μL/min; temperature, 50° C; product ions of the 2-charged species at *m/z* 738.2, 4.0 u isolation width; 27% relative collision energy; scan, 200–2000 u; electrospray voltage, 3.4 kV; sheath gas, 40 units; sheath liquid, 3.0 μl/min acetonitrile; sample, 5-mer (Sequence 1).

Figure 2 illustrates the MSMS spectrum of a 5-mer (Sequence 1, Table 1). The spectrum was extracted as the average of 3 scans from the peak eluting at 1.6 min from the monolithic separation column. With the knowledge of the entire sequence of the 5-mer the sequence specific a-B- and w-ions could be easily assigned. We believe that diagrams like that shown in Figure 2 are convenient to illustrate the coverage of a certain sequence by fragment ions. In the first line the complete oligonucleotide sequence is denoted. The presence of fragments that are useful for the identification of the nucleotides is indicated by black rectangles at the corresponding position in the sequence. The one-unit shift to the left in the numbering of the a-B-ions with respect to the oligonucleotide sequence reflects the fact that, due to the loss of the nucleobase at the position of the backbone cleavage, the fragments contain sequence information only about the previous nucleotide in the sequence, and not about the nucleobase present exactly at the cleavage position. An a_3-B_3 fragment, for example, gives information about the nucleotides at positions 1 and 2 but provides no evidence of the nucleotide at position 3.

Using this diagram, gaps in sequence information coming from missing fragment ions of both series are easily spotted. Therefore, full sequence coverage was only obtained for the w-series, whereas just two a-B-

Table 2. Results obtained from the comparison of a fragment ion spectrum to all possible sequences derived from the 5-mer composition $C_2G_2T_1$

Sequence	M(i)	MF(i)	K(i)	$M(i)/M_{max}$	$MF(i)/MF_{max}$	$K(i)/K_{max}$	FS(i)
CTGGC	2	244.63	9	0.50	0.556	1.000	−0.194
CTGCG	3	275.16	8	0.375	0.626	0.889	0.112
TCGGC	3	275.20	8	0.375	0.626	0.889	0.112
CGTGC	3	275.24	8	0.375	0.626	0.889	0.112
CTCGG	5	0.77	4	0.625	0.002	0.444	0.182
CGGTC	5	0.93	4	0.625	0.002	0.444	0.183
TCCGG	6	1.00	3	0.750	0.002	0.333	0.419
CGCGT	6	1.10	3	0.750	0.003	0.333	0.419
CGCTG	6	1.10	3	0.750	0.003	0.333	0.419
CGGCT	6	1.10	3	0.750	0.003	0.333	0.419
GCGTC	6	1.21	3	0.750	0.003	0.333	0.419
TGGCC	6	2.64	3	0.750	0.006	0.333	0.423
GTGCC	6	2.64	3	0.750	0.006	0.333	0.423
GGTCC	6	2.64	3	0.750	0.006	0.333	0.423
TCGCG	4	314.46	7	0.500	0.715	0.778	0.437
CGTCG	4	314.50	7	0.500	0.715	0.778	0.438
GCTGC	4	314.54	7	0.500	0.715	0.778	0.438
CCTGG	7	0.12	2	0.87	0.000	0.222	0.653
CCGTG	7	0.12	2	0.875	0.000	0.222	0.653
CCGGT	7	0.12	2	0.87	0.000	0.222	0.653
GGCTC	7	0.28	2	0.875	0.001	0.222	0.653
GCCGT	7	1.60	2	0.875	0.004	0.222	0.656
GCCTG	7	1.60	2	0.875	0.004	0.222	0.656
GCGCT	7	1.60	2	0.875	0.004	0.222	0.656
GTCGC	5	366.46	6	0.625	0.833	0.667	0.792
TGCGC	5	366.46	6	0.625	0.833	0.667	0.792
GCTCG	5	366.90	6	0.625	0.834	0.667	0.793
GGCCT	8	0.15	1	1.000	0.000	0.111	0.889
TGCCG	6	439.67	5	0.750	1.000	0.556	1.194
GTCCG	6	439.67	5	0.750	1.000	0.556	1.194

fragment ions were present which partially covered the sequence of the 5-mer. However, this information is sufficient for the verification of the complete sequence, but we believe that even experts would need a considerable amount of time to unequivocally deduce the sequence de novo from the MSMS spectrum shown in Figure 2. Sequence information inherited in MSMS spectra derived from multiply charged oligonucleotides is often not obvious from causal inspection due to the fact that product ion charge states can vary from unity up to the parent ion charge. Moreover, with increasing length of the oligonucleotide the number of possible fragment ions and the likelihood of missing fragments in the series of a-B- or w-ions increase, which both complicate and in some cases prevent the successful manual passage through the whole sequence.

To circumvent the laborious, time-consuming, and disaffected manual interpretation, we have used a fully automated de novo sequencing algorithm for the ab initio sequence determination of the 5-mer. The mass spectral data served as input for the de novo sequencing algorithm. The results obtained from the comparison of the fragment ion spectrum to all thirty sequences derived from the oligonucleotide composition $C_2G_2T_1$ are depicted in Table 2. Obviously the lowest FS value was obtained for the true sequence (Sequence 1, Table 1). In this example solely the best matching sequence is characterized by the highest K(i)- and the lowest M(i)-value. Since this is more likely an exception than the rule, and since wrong sequences can yield much lower MF(i)-values than the true sequences (Table 2), in general only the combination of the three terms into a comprehensive match factor allows the unequivocal de novo sequencing of oligonucleotides.

Characterization of 7- to 9-Mer Oligodeoxynucleotides Showing Abnormal Retention Behavior

IP-RP-HPLC is a well established and popular technique for the separation of single- and double-stranded nucleic acids [31, 32]. The phase system comprises a hydrophobic stationary phase and a hydroorganic eluent containing an amphiphilic ion. According to retention models, the positively-charged amphiphilic ions are adsorbed at the interface between the stationary phase and the mobile phase, resulting in the formation of a positive surface potential. Retention is mainly accomplished by the interaction of the surface potential with the negatively-charged DNA molecules and is therefore governed by chain length. However, additional solvophobic interactions between the quite hydrophobic nucleobases and the surface of the stationary

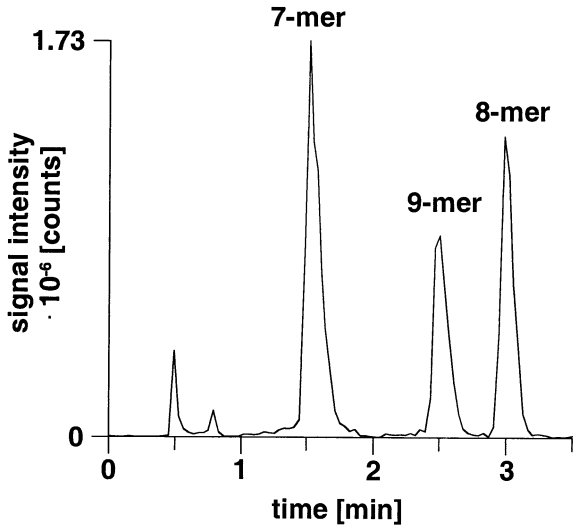

Figure 3. LC-MSMS analysis of a mixture of 3 oligodeoxynucleotides. Column, monolithic PS-DVB, 60 × 0.20 mm i.d.; mobile phase, (**a**) 25 mM TEAB, pH 8.40, (**b**) 25 mM TEAB, pH 8.40, 20% acetonitrile; linear gradient, 5–50% B in 10.0 min; flow-rate, 3.0 μL/min; temperature, 50° C; product ions of the 2-charged species at *m/z* 1042.1, 1198.5, 1343.6; 4.0 u isolation width; 23% relative collision energy; scan, 285–2000 u for the 7-mer, 325–2000 u for the 8-mer, 365–2000 u for the 9-mer; electrospray voltage, 3.4 kV; sheath gas, 40 units; sheath liquid, 3.0 μl/min acetonitrile; sample, 7-mer (Sequence 2), 8-mer (Sequence 3), 9-mer (Sequence 4).

phase are observed, which results in a sequence-dependent retention of oligonucleotides [33]. Although this phenomenon has been known for quite some time, the retention behavior is still unpredictable on the basis of the oligonucleotide sequence. Therefore, there is a strong need for reliable methods like LC-MS or even better LC–MSMS, which unequivocally prove the identity of chromatographically separated oligonucleotides.

An example for an oligonucleotide showing abnormal retention behavior is depicted in Figure 3. Three homologous oligonucleotides, obtained by addition of one (A) or two nucleotides (CA) to the 5′ end of a 7-mer sequence (Sequences 2–4, Table 1), were analyzed by IP-RP-HPLC. Surprisingly, the 9-mer oligonucleotide eluted between the 7- and the 8-mer from the chromatographic column. For the investigation of this phenomenon and the identification of the three oligonucleotides, LC-MSMS was applied. The oligonucleotide mixture was separated using a gradient of 1 to 10% acetonitrile in 25 mM triethylammonium bicarbonate in 10 min. The three investigated oligonucleotides eluted from the capillary column at 1.4, 2.4, and 2.9 min, respectively. The corresponding MSMS spectra originated from fragmentation of the 2− charge states at a relative collision energy of 23%. The mass spectral data served as input for the de novo sequencing algorithm. The sequencing results (Table 3) obviously proved that Sequence 2 and the sequence of the investigated 7-mer were identical.

The sequences of the 8-mer and 9-mer could not be unequivocally identified as in both cases the sequences differing in the order of the final two bases matched the fragment ion spectrum identically well. This ambiguous result was caused by the absence of the w_1-ions in the fragment ion spectra (Figure 4). As discussed earlier, due to the loss of the nucleobase at the position of the backbone cleavage, a-B-fragments contain sequence information only about the previous nucleotide in the sequence, and not about the nucleobase present exactly at the cleavage position. Consequently, the order of the final two bases is indefinable using a-B-fragment ions.

The complete nucleotide order can only be determined based on the sequence information stored in the

Table 3. De novo sequencing of three synthetic oligonucleotides: (a) 7-mer, (b) 8-mer, (c) 9-mer

(a) Sequence	FS(i)	(b) Sequence	FS(i)	(c) Sequence	FS(i)
TTCTGGC	−0.585	ATTCTGCG	−0.546	CATTCTGGC	−0.492
TTCTGCG	−0.491	ATTCTGGC	−0.546	CATTCTGCG	−0.492
TCTTGGC	−0.323	TATCTGCG	−0.485	ACTTCTGCG	−0.440
TTTCGGC	−0.247	TATCTGGC	−0.485	ACTTCTGGC	−0.440
CTTTGGC	−0.231	CGTCTGTA	−0.412	CGCTCTGAT	−0.405
TTCGTGC	−0.204	CGTCTGAT	−0.412	CGCTCTGTA	−0.405
TCTTGCG	−0.187	TTACTGCG	−0.361	CCCTGTGAT	−0.338
TTCTCGG	−0.149	TTACTGGC	−0.361	CCCTGTGTA	−0.338
TTGTCGC	−0.148	ATTCGTGC	−0.354	CTATCTGGC	−0.328
TTTCGCG	−0.111	ATTCGTCG	−0.354	CTATCTGCG	−0.328
CTTTGCG	−0.092	ATCTTGCG	−0.301	TACTCTGCG	−0.327
TTCGTCG	−0.071	ATCTTGGC	−0.301	TACTCTGGC	−0.357
TTTGCGC	−0.023	CGCTTGTA	−0.297	ATCTCTGGC	−0.323
TTGTCCG	−0.008	CGCTTGAT	−0.297	ATCTCTGCG	−0.323
TTGCTGC	0.007	TATCGTCG	−0.291	CGTCCTGTA	−0.323
TCTGTGC	0.055	TATCGTGC	−0.291	CGTCCTGAT	−0.323
TTCCTGG	0.068	CGTCTATG	−0.266	TGCCCTGTA	−0.312
TTCGGTC	0.107	CGTCTAGT	−0.289	TGCCCTGAT	−0.312
TTTGCCG	0.119	ATTTCGGC	−0.265	CTCGCTGAT	−0.293
TTCCGTG	0.130	ATTTCGCG	−0.265	CTCGCTGTA	−0.293
etc.		etc.		etc.	

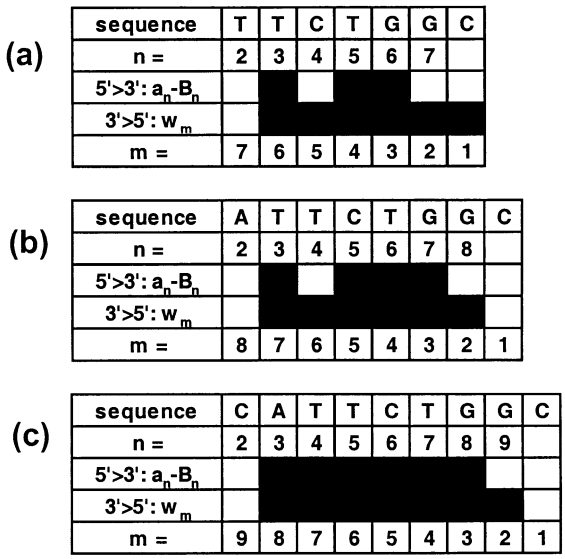

Figure 4. Sequence coverage diagrams extracted from the LC-MSMS analysis of 3 oligodeoxynucleotides. Same conditions as in Figure 3.

w-ion series. Under these circumstances, the assignment of the w_1-ion is of utmost importance to reveal the order of the final two bases. For technical reasons, the settings of the ion trap mass analyzer prevented the detection of that distinct ion during the MSMS analysis of the 8-mer and the 9-mer. The parent ion is translationally excited in the ion trap after mass selection by applying a supplementary radio frequency voltage to the end cap electrodes. The product ions resulting from CID of the excited ions with the helium gas are trapped and scanned. Depending on the mass-to-charge ratio (*m/z*) of the parent ion, fragment ions with *m/z* values lower than a characteristic cutoff are not trapped anymore. Therefore, due to the low-mass cutoffs at 285 for the 7-mer, at 325 for the 8-mer, and at 365 for the 9-mer, respectively, the w_1-ion with *m/z* of 306.2 was only observed for the 7-mer, which allowed the unequivocal sequence determination exclusively for that distinct oligonucleotide. However, despite the existing ambiguity of the nucleotide order at the 3′-end, de novo sequencing testified the conformity of the three sequences and corroborated the hypothesis that the inversion in elution is based on differences in the 8- and 9-mer sequences.

Dependence of De Novo Sequencing on Fragmentation Energy

In the course of investigations about the applicability of COMPAS to the verification of sequences longer than 20 nucleotides, we found that collision energy had a major impact on the ability to interpret MSMS spectra [23]. Similarly, the influence of the collision energy on the de novo sequencing results of small oligonucleotides was investigated. For this study the 7-mer, 8-mer, and 9-mer oligonucleotide (Sequences 2–4, Table 1), respectively, were fragmented by applying collision energies between 15 and 100%. Afterward, all possible sequences derived from the composition of the three oligonucleotides were matched to the measured fragment ion spectra and ranked according to the magnitude of the calculated match factor. The rank of the correct sequence was used as an indicator for the usefulness of a certain collision energy for the sequencing procedure.

Table 4. Ranks of the correct 7-mer, 8-mer, and 9-mer sequences, respectively, by applying the de novo sequencing algorithm to fragment ion spectra generated at different collision energies

Relative collision energy [%]	Rank of sequence TTCTGGC	ATTCTGGC	CATTCTGGC
15	1 (5)[a]	1 (1)	2 (1)
17	2	1 (1)	1(1)
20	1	1 (1)	2 (1)
23	1	1 (1)	1 (1)
25	1	1 (1)	3 (1)
27	1	4 (1)	13 (1)
30	1	1 (1)	62 (1)
100	1	1 (1)	14 (1)

[a]The numbers of additional sequences yielding the same FS(i)-value like the investigated sequence are given in brackets.

The results summarized in Table 4 clearly revealed that de novo sequencing is possible by applying a broad range of collision energies and in particular the results of the 7-mer and 8-mer suggest that the spectral quality does not change significantly between 15 and 100% relative collision energy (Table 4). However, it is important to note that at a relative collision energy of 15%, 6 different 7-mer sequences matched the fragment ion spectrum equally well. Obviously, in this experiment the collision energy was too low to generate enough sequence specific fragments, which made the unequivocal de novo sequence determination impossible. On the other hand, the failure to interpret the MSMS spectrum of the 9-mer recorded at collision energies higher than 27% is most probably due to secondary and/or alternative fragmentation reactions viable at higher collision energies giving rise to fragments that may be misinterpreted as regular a-B- or w-ions. Thus, the optimum collision energy has to be high enough to generate fragments covering the sequence as completely as possible by a-B- and w-ions but low enough to suppress secondary and/or alternative fragmentation pathways. For the sequence verification of oligonucleotides by COMPAS, this compromise was empirically found to be in the region of the collision energy just sufficient for complete fragmentation of the precursor ion [23]. Obviously, this rule is valid in the same manner for de novo sequencing. However, since MSMS spectra generated at a broad range of collision energies were applicable to the sequencing of small oligonucleotides, the proper setting of that experimental parameter is not so critical. Certainly, the influence of the

Table 5. Rank of different 9-mer sequences by applying the de novo sequencing algorithm to the fragment ion spectra generated from the oligonucleotide CATTCTGGC at different collision energies

Relative collision energy [%]	15	17	20	23	25
Rank of CATTCTGGC	2 (1)[a]	1 (1)	2 (1)	1 (1)	3 (1)
Rank of CCCTGTGAT	1 (1)	5 (1)	1 (1)	4 (1)	33 (1)
Rank of CCCTGGTTA	1 (1)	11 (1)	5 (1)	107 (1)	317 (1)
Rank of CGTCCTGTA	9 (23)	37 (11)	4 (1)	8 (1)	1 (1)

[a]The numbers of additional sequences yielding the same FS(i)-value like the investigated sequence are given in brackets.

collision energy will gain more importance once the algorithm will be applied to the de novo sequencing of nucleic acids at the 15-mer level and beyond.

Nevertheless, even by using appropriate energy settings (15 up to 25%) incorrect 9-mer oligonucleotide sequences matched some fragment ion spectra better than the correct one (Table 4). Anyhow, concerning the complete set of applied collision energies, the correct 9-mer sequence yielded the overall best rankings of all oligonucleotide sequences (Table 5). To illustrate this, the sequence CCCTGTGAT, which received two first ranks at 15 and 20% collision energy, respectively, gained comparably worse rankings than the correct sequence at the remaining energy levels clearly indicating the incorrectness of that sequence. Hence, the treatment of a set of fragment ion spectra generated at different collision energy values is advisable to achieve the desired unequivocal de novo sequencing results.

Alteration of the de Novo Sequencing Result of an 11-Mer by Missing Fragments

A general problem of MSMS sequencing is the occurrence of missing fragments in the a-B- or w-ion series, respectively, which prevents the coverage of the complete sequence with sequence specific ions. According to the literature, fragment formation is site-specific [8, 10, 16, 34]. The first step of the fragmentation process involves the protonation of a nucleobase, which initiates the neutral base loss and thereby leads to the backbone cleavage to form a-B- and w-ions. The preference for losing G, A, C, or T, respectively, follows the order of the proton affinities of the nucleobases. Therefore, in particular the cleavage of the phosphate bond at the 3′-side next to thymine is disfavored leading to low signal intensities or to the complete loss of the corresponding fragment ions.

Figure 5a depicts the sequence coverage diagram of the 11-mer (Sequence 5, Table 1). Obviously, some nucleotide positions were not covered by fragment ions. Therefore, the likelihood was high that other sequences matched the MSMS spectrum equally or even better than the correct sequence. In Table 6a the de novo sequencing results of the 11-mer are summarized. As anticipated, the sequence GACCGTCTGAT yielded a slightly lower FS(i) value indicating a better match than the correct sequence GACATTCTGGC. Interestingly, both sequences can be transferred into each other by exchanging the two bases at position 4 and 5 with the final two bases (AT-GC). As the molecular mass of AT (555.4 Da) differs from GC (556.4 Da) just by 1.0 Da, all fragment ions, which were long enough to contain a complete nucleotide pair, could fall within the *m/z* tolerance of ±1.0, which becomes even more likely considering higher charge states. Therefore, only the a_6-B_6-, the w_1-, and the w_8-ions, respectively, would unequivocally distinguish the two sequences based on molecular mass differences larger than 1.0 Da. Unfortunately, none of these specific fragment ions were observed matching the correct sequence, but the theoretical a_6-G-fragment corresponding to the sequence GACCGTCTGAT was assigned to an unspecific and low abundant signal in the measured fragment ion spectrum (Figure 5b). Consequently, the number of assigned fragments was higher, and the sum of noncovered a- and w-positions was lower, which yielded an overall better FS(i)-value for that distinct sequence. This example clearly proves that a high coverage with sequence specific fragment ions is a prerequisite for successful de novo sequencing. Otherwise incorrect sequences will fit the experimental spectrum equally or even better than the correct sequence due to assignments of hypothetical fragment ions of wrong sequences to unspecific signals in the MSMS spectrum.

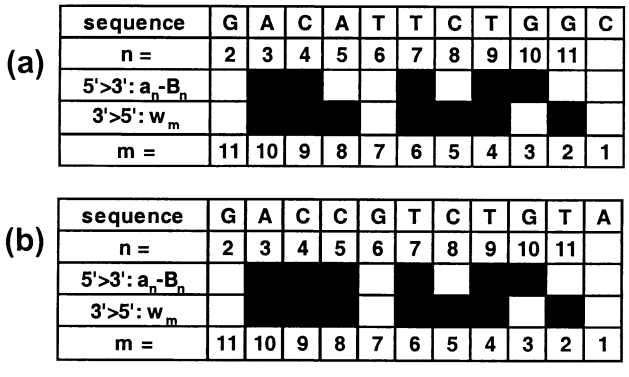

Figure 5. Sequence coverage diagrams for (**a**) the correct and (**b**) the better matching sequence extracted from the LC-MSMS analysis of an 11-mer oligodeoxynucleotide. Column, monolithic PS-DVB, 60 × 0.20 mm i.d.; mobile phase, (**a**) 25 mM TEAB, pH 8.40, (**b**) 25 mM TEAB, pH 8.40, 20% acetonitrile; linear gradient, 5–50% B in 10.0 min; flow-rate, 3.0 μL/min; temperature, 50° C; product ions of the 2-charged species at *m/z* 1664.2; 4.0 u isolation width; 23% relative collision energy; scan, 445–2000 u; electrospray voltage, 3.4 kV; sheath gas, 40 units; sheath liquid, 3.0 μl/min acetonitrile; sample, 11-mer (Sequence 5).

Table 6. De novo sequencing results for two oligonucleotides: (a) 11-mer, (b) 12-mer

(a) Sequence	FS(i)	(b) Sequence	FS(i)
GACCGTGTGAT	−0.349	AGACATTCTGGC	−0.349
GACCGTCTGTA	−0.349	AGACATTCTGCG	−0.349
GACATTCTGGC	−0.346	AGACTATCTGCG	−0.289
GACTATCTGCG	−0.346	AGACTATCTGGC	−0.289
GACTATCTGGC	−0.346	AGACTTACTGGC	−0.289
GACATTCTGCG	−0.346	TATGGTCAGCAC	−0.289
GACGCTCTGTA	−0.286	AGACTTACTGCG	−0.289
GACGCTCTGAT	−0.286	AGACTTCATGCG	−0.286
GACCTGCTGTA	−0.262	AGACTCTATGGC	−0.286
GACCTGCTGAT	−0.262	AGACTCTATGCG	−0.286
GACCGTTCGTA	−0.250	AGACTTCATGGC	−0.286
GACCGTTCGAT	−0.250	AGACTGCCTGTA	−0.266
GACTTACTGCG	−0.245	AGACTCGCTGAT	−0.266
GACTTACTGGC	−0.245	AGACTGCCTGAT	−0.266
GACTACTTGCG	−0.244	AGACTCGCTGTA	−0.266
GACTACTTGGC	−0.244	AGACTTCTAGGC	−0.265
GACATGTTGCG	−0.244	AGACTCTTAGGC	−0.266
GACATCTTGGC	−0.244	AGACTTCTAGCG	−0.266
GAACTTCTGCG	−0.234	AGACTCTTAGCG	−0.265
GAACTTCTGGC	−0.254	AGACTTCTGAGC	−0.263
etc.		etc.	

Limitations of the De Novo Sequencing Algorithm

Previously established sequencing algorithms like SOS and COMPAS were shown to be highly suitable for the sequence verification of nucleic acids up to the 20-mer [13, 22] or 50-mer [23] level, respectively. However, the longest sequence, which could be correctly determined so far using our newly-developed de novo sequencing algorithm, was the sequence of a 12-mer oligonucleotide (Sequence 6, Table 1).

The 12-mer was purified by using a gradient of 1 to 10% acetonitrile in 25 mM triethylammonium bicarbonate in 10 min. The MSMS spectrum originating from fragmentation of the 3- charge state at 20% relative collision energy together with the nucleotide composition $A_3C_3G_3T_3$ served as input for the de novo sequencing program. The results are shown in Table 6b. Once more, apart from the order of the final two bases the complete oligonucleotide sequence was correctly determined, which clearly proves the suitability of our approach for the sequencing of nucleic acids. We believe that by applying the developed matching procedure the sequencing of even longer oligonucleotides is feasible. However, due to the rapid increase of the number of possible sequences with increasing oligonucleotide length the algorithm encountered difficulties at the 12-mer level, which could not be overcome so far. Whereas the permutation of the 5-mer sequence yielded only 30 possible sequences, 369 600 unique combinations were obtained starting from the 12-mer composition (Table 7). Since the operating expense is directly connected to the number of possible sequences, calculation time and occupied disc space increase tremendously with growing oligonucleotide length. Therefore, due to a sequencing time of 46 h and an occupied disc space of 11.9 MB (Microsoft Excel was not able to open the file completely and to show all 369 600 lines of the calculated list) the determination of the 12-mer sequence so far represents the upper limit for the applicability of the sequencing algorithm in its current form. However, the introduction of some kind of a pruning technique, which will limit the combinatorial explosion and therefore reduce calculation time, will extend the size limit of the algorithm enormously in the future.

Table 7. Comparison of the number of possible sequences, calculation time, and occupied disc space, respectively, necessary for the de novo sequencing of 5- to 12-mer oligonucleotides

Correct sequence	Number of possible sequences	Calculated time	Occupied disc space [kB]
CTGGC	30	10 s	<1
CATTCTGGC	5,040	30 min	145
GACATTCTGGC	92,400	10 h	2,748
AGACATTCTGGC	369,600	46 h	11,901

Conclusions

Electrospray ionization tandem mass spectrometry under optimized fragmentation conditions in combination with a computer-aided de novo sequencing algorithm is applicable to the fully automated sequence analysis of small nucleic acids. The sequence determination is based on the comparison of the experimental tandem mass spectrum with a set of fragment ions predicted from a reference sequence created by permutation of a given oligonucleotide composition. Like for any other de novo sequencing algorithm, unequivocal sequence determination is restricted by the absence of sequence-specific fragment ions. At this stage of the study the applicability of the sequencing algorithm is limited to the 12-mer level. Due to an exponential increase of the number of possible sequences with increasing oligonucleotide length, calculation time becomes unacceptably long for larger nucleic acids. Therefore, future investigations will focus on an extension of the size range by the introduction of pruning techniques, which will limit the combinatorial explosion and consequently reduce sequencing time.

References

1. Huber, C. G.; Oberacher, H. Analysis of Nucleic Acids by On-Line Liquid Chromatography-Mass Spectrometry. *Mass Spectrom. Rev.* **2001,** *20,* 310–343.
2. Premstaller, A.; Oberacher, H.; Huber, C. G. High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry of Single- and Double-Stranded Nucleic Acids Using Monolithic Capillary Columns. *Anal. Chem.* **2000,** *72,* 4386–4393.
3. Oberacher, H.; Oefner, P. J.; Parson, W.; Huber, C. G. On-Line Liquid Chromatography-Mass Spectrometry: A Useful Tool for the Detection of DNA Sequence Variation. *Angew. Chem. Int. Ed.* **2001,** *40,* 3828–3830.

4. Walcher, W.; Oberacher, H.; Troiani, S.; Hölzl, G.; Oefner, P.; Zolla, L.; Huber, C. G. Monolithic Capillary Columns for Liquid Chromatography-Electrospray Ionization Mass Spectrometry in Proteomic and Genomic Research. *J. Chromatogr. B* **2002,** *782,* 111–125.
5. Oberacher, H.; Parson, W.; Mühlmann, R.; Huber, C. G. Analysis of Polymerase Chain Reaction Products by On-Line Liquid Chromatography-Mass Spectrometry for Genotyping of Polymorphic Short Tandem Repeat Loci. *Anal. Chem.* **2001,** *73,* 5109–5115.
6. Berger, B.; Hölzl, G.; Oberacher, H.; Niederstätter, H.; Huber, C. G.; Parson, W. Single Nucleotide Polymorphism Genotyping by On-Line Liquid Chromatography-Mass Spectrometry in Forensic Science with the Y-Chromosomal Locus M9. *J. Chromatogr. B* **2002,** *782,* 89–97.
7. Oberacher, H.; Huber, C. G.; Oefner, P. J. Mutation Screening by Ion-Pair Reversed-Phase High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry (ICEMS). *Hum. Mutat.* **2003,** *21,* 86–95.
8. McLuckey, S. A.; Berkel, G. J.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992,** *3,* 60–70.
9. Kirpekar, F.; Krogh, T. N. RNA Fragmentation Studied in a Matrix-Assisted Laser Desorption/Ionization Tandem Quadrupole/Orthogonal Time-of-Flight Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **2001,** *15,* 8–14.
10. McLuckey, S. A.; Habibi-Goudarzi, S. Ion-Trap Tandem Mass-Spectrometry Applied to Small Multiply-Charged Oligonucleotides with a Modified Base. *J. Am. Soc. Mass Spectrom.* **1994,** *5,* 740–747.
11. Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. Rapid Sequencing of Oligonucleotides by High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1994,** *116,* 4893–4897.
12. McLuckey, S. A.; Vaidayanathan, G.; Habibi-Goudarzi, S. Charged Versus Neutral Nucleobase Loss from Multiply Charged Oligonucleotide Anions. *J. Mass Spectrom.* **1995,** *30,* 1222–1229.
13. Ni, J.; Pomerantz, S. C.; Rozenski, J.; Zhang, Y.; McCloskey, J. A. Interpretation of Oligonucleotide Mass Spectra for Determination of Sequence Using Electrospray Ionization and Tandem Mass Spectrometry. *Anal. Chem.* **1996,** *68,* 1989–1999.
14. Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. Sequence Information from 42-108-Mer DNAs (Complete for a 50-Mer) by Tandem Mass Spectrometry. *J. Am. Chem. Soc.* **1996,** *118,* 9352–9359.
15. Vrkic, A. K.; O'Hair, R. A. J.; Foote, S.; Reid, G. E. Fragmentation Reactions of all 64 Protonated Trimer Oligodeoxynucleotides and 16 Mixed Base Tetramer Oligodeoxynucleotides via Tandem Mass Spectrometry in an Ion Trap. *Int. J. Mass Spectrom.* **2000,** *194,* 145–164.
16. Wan, K. X.; Gross, J.; Hillenkamp, F.; Gross, M. L. Fragmentation Mechanisms of Oligodeoxynucleotides Studied by H/D Exchange and Electrospray Ionization Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2001,** *12,* 193–205.
17. Wan, K. X.; Gross, M. L. Fragmentation Mechanisms of Oligodeoxynucleotides: Effects of Replacing Phosphates with Methylphosphonates and Thymines with Other Bases in T-Rich Sequences. *J. Am. Soc. Mass Spectrom.* **2001,** *12,* 580–589.
18. Barry, J. P.; Vouros, P.; Van Schepdael, A.; Lay, S.-J. Mass and Sequence Verification of Modified Oligonucleotides Using Electrospray Tandem Mass Spectrometry. *J. Mass Spectrom.* **1995,** *30,* 993–1006.
19. Genti, E.; Banoub, J. Characterization and Differentiation of Isomeric Self-complementary DNA Oligomers by Electrospray Tandem Mass Spectrometry. *J. Mass Spectrom.* **1996,** *31,* 83–94.
20. Little, D. P.; McLafferty, F. W. Sequencing 50-Mer DNAs using Electrospray Tandem Mass Spectrometry and Complementary Fragmentation Methods. *J. Am. Chem. Soc.* **1995,** *117,* 6783–6784.
21. Krahmer, M. T.; Walters, J. J.; Fox, K. F.; Fox, A.; Creek, K. E.; Pirisi, L.; Wunschel, D. S.; Smith, R. D.; Tabb, D. L.; Yates, J. R. MS for Identification of Single Nucleotide Polymorphisms and MS/MS for Discrimination of Isomeric PCR Products. *Anal. Chem.* **2000,** *72,* 4033–4040.
22. Rozenski, J.; McCloskey, J. A. SOS: A Simple Interactive Program for ab Initio Oligonucleotide Sequencing by Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2002,** *13,* 200–203.
23. Oberacher, H.; Wellenzohn, B.; Huber, C. G. Comparative Sequencing of Nucleic Acids by Liquid Chromatography-Tandem Mass Spectrometry. *Anal. Chem.* **2002,** *74,* 211–218.
24. Oberacher, H.; Walcher, W.; Huber, C. G. Effect of Instrument Tuning on Detectability of Biopolymers in Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2003,** *38,* 108–116.
25. Pomerantz, S. C.; Kowalak, J. A.; McCloskey, J. A. Determination of Oligonucleotide Composition from Mass Spectrometrically Measured Molecular Weight. *J. Am. Soc. Mass Spectrom.* **1993,** *4,* 204–209.
26. Koomen, J. M.; Russell, W. K.; Tichy, S. E.; Russell, D. H. Accurate Mass Measurement of DNA Oligonucleotide Ions Using High-Resolution Time-of-Flight Mass Sspectrometry. *J. Mass Spectrom.* **2002,** *37,* 357–371.
27. Dovici, N. J.; Zhang, J. Wie die Kapillarelektrophorese das menschliche Genom Sequenzierte. *Angew. Chem.* **2000,** *112,* 4635–4640.
28. Griffey, R. H.; Greig, M. J.; Gaus, H. J.; Liu, K.; Monteith, D.; Winniman, M.; Cummins, L. L. Characterization of Oligonucleotide Metabolism in Vivo via Liquid Chromatography/Electrospray Tandem Mass Spectrometry with a Quadrupole Ion Trap Mass Spectrometer. *J. Mass Spectrom.* **1997,** *32,* 305–313.
29. Gaus, H. J.; Owens, S. R.; Winniman, M.; Cooper, S.; Cummins, L. L. On-Line HPLC Electrospray Mass Spectrometry of Phosphorothioate Oligonucleotide Metabolites. *Anal. Chem.* **1997,** *69,* 313–319.
30. Oberacher, H.; Huber, C. G. Capillary Monoliths for the Analysis of Nucleic Acids by High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry. *TRAC* **2002,** *21,* 166–173.
31. Huber, C. G.; Oefner, P. J.; Bonn, G. K. Rapid and Accurate Sizing of DNA Fragments by Ion-Pair Reversed-Phase Chromatography on Alkylated Nonporous Poly(Styrene-Divinylbenzene) Particles. *Anal. Chem.* **1995,** *67,* 578–585.
32. Huber, C. G.; Oefner, P. J.; Preuss, E.; Bonn, G. K. High-Resolution Liquid Chromatography of DNA Fragments an Non-Porous Poly(Styrene-Divinylbenzene) Particles. *Nucleic Acids Res.* **1993,** *1061,* 1061–1066.
33. Huber, C. G.; Oefner, P. J.; Bonn, G. K. High-Resolution Liquid Chromatography of Oligonucleotides on Highly Crosslinked Poly(Styrene-Divinylbenzene) Particles. *Anal. Biochem.* **1993,** *212,* 351–358.
34. Wang, Z.; Wan, K. X.; Ramanathan, R.; Taylor, J. S.; Gross, M. L. Structure and Fragmentation Mechanisms of Isomeric T-Rich Oligodesoxynucleotides: A Comparison of Four Tandem Mass Spectrometry Methods. *J. Am. Soc. Mass Spectrom.* **1998,** *9,* 683–691.