Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 231 (2004) 169–177
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Analysis of the in vitro digestion of modified DNA to oligonucleotides by LC–MS and LC–MS/MS

Christine L. Andrews, Andreas Harsch[1], Paul Vouros*

Department of Chemistry and Chemical Biology, Barnett Institute, Northeastern University, 102 Hurtig Hall, 360 Huntington Avenue, Boston, MA 02115, USA

Received 23 June 2003; accepted 11 October 2003

Abstract

Liquid chromatography–mass spectrometry (LC–MS) techniques are ideally suited to the investigation of DNA adduction. Most current experiments focus on the quantification of monomeric DNA adducts, or the reaction of synthetic oligonucleotides with a specific carcinogen. The methodology presented herein allows for examination of the sequence context of an adducted segment by the enzymatic digestion of DNA modified by the carcinogen *N*-acetoxy-2-acetylaminofluorene to oligonucleotide fragments, and subsequent analysis by LC–MS and LC–MS/MS. Synthetic oligonucleotide mixtures were also analyzed to facilitate method development. LC–MS results were compared to predicted masses of adducted oligonucleotide fragments allowing determination of nucleobase compositions bearing a covalent adduct. Subsequent LC–MS/MS analysis enabled identification of specific sequences containing a covalent modification. Relative differences in the abundances of modified sequences were observed in the comparison of the digestion products to the synthetic standards. This approach demonstrates promise for determining binding preferences to DNA by presenting the entire molecule as a target.

Keywords: Liquid chromatography–tandem mass spectrometry; Oligonucleotide adducts; Enzymatic digestion; *N*-Acetoxy-2-acetylaminofluorene

1. Introduction

The initiation of many genotoxic cancers is thought to involve binding of a xenobiotic to cellular biomolecules, in particular nucleic acids. It is well known that there is base sequence selectivity in the binding of certain carcinogens to DNA. For example, benzo[*a*]pyrene (BaP) is known to bind strongly to guanine when presented as runs of guanine (poly[dG]), as well as specific sequences such as CGC, AGC, and TGG [1–3]. *N*-Acetoxy-2-acetylaminofluorene (AAAF) is a well-characterized carcinogen (reviewed in [4,5]) that has shown preferential binding to certain sequences in pBR322, phiX174 and SV40 plasmid DNA, as demonstrated by enzyme inhibition near the sequences T(C/G)TT(G/C) and T(G/C)CTT(G/C) [6–8]. While there is no consensus regarding direct correspondence between site-specific adduction and mutagenesis [9], it is still widely held that adduction by endogenous and/or exogenous compounds may be the first step to mutation [10,11]. There is evidence, for example, of a strong dependency of mutational frequencies on the sequence context of purine targets. BaP exposure is thought to be involved in guanine to thymine (G → T) transversions in several codons along the *p53* gene [12–14]. Mutations attributed to aminobiphenyl exposure are also observed in the same codons as BaP, as well as in additional regions along the *p53* gene [15]. AAAF modification causes frameshift mutations in the NarI region [16–22]. As more human genomic information is gathered and organized into databases, more understanding may be gained into which genes are active in specific tissues and organs. This may make it possible to understand the link between a specific chemical carcinogen and human cancer. Hence, if a sequence within a gene is shown to have a predilection for mutation, it is important to investigate the sequence's susceptibility to adduction.

Complete structural elucidation of carcinogen modified genomic sequences requires not only determination of the oligonucleotide sequence, but also recognition of the adduction site. Mass spectrometry (MS), especially in com-

* Corresponding author. Tel.: +1-617-373-2840; fax: +1-617-373-8478.
E-mail address: p.vouros@neu.edu (P. Vouros).
[1] Present address: Abbott BioResearch Center, Worcester, MA 01605, USA.

1387-3806/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2003.10.005

bination with separation methods, has the sensitivity and selectivity to accomplish the goal of analyzing carcinogen–oligonucleotide adducts in complex mixtures from in vitro and in vivo sources. In recent years, matrix assisted laser desorption ionization (MALDI) mass spectrometry has begun to play a significant role in the investigation of carcinogen modified DNA constituents, especially for the analysis of oligonucleotide adducts. Picomole [23,24] and even femtomole [25–27] detection limits have been demonstrated by MALDI-MS. Generation of structural information from MALDI mass spectra has also been possible by post source decay (PSD) [28,29]. A combination of MALDI-MS and enzymatic digestion was used to locate nucleobase lesions in oligonucleotides [30]. Currently, a limitation of MALDI-MS is that it is not readily amenable to coupling with separation methods such as high performance liquid chromatography (HPLC) and capillary electrophoresis (CE), although efforts to that effect are promising for the analysis of large molecules [31,32]. The significance of the separation step prior to MS analysis is exemplified by the fact that formation of isomeric adducts is a frequent occurrence. Compositional isomers are a particular concern when one is analyzing oligonucleotides. Due to this fact, electrospray mass spectrometry is well suited for the analysis of carcinogen-modified oligonucleotides because of its now routine coupling to separation methods, especially HPLC. Recently, LC–MS was used to indirectly investigate the binding spectrum of benzo[*a*]pyrene along specific codons of the *p53* gene by digesting isotope-labeled oligonucleotide probes to adducted nucleosides [33].

Investigations to date have focused on the analysis of synthetic oligonucleotides reacted with a carcinogen. These studies are extremely useful, but they do not address the binding differences when a carcinogen is presented with a short oligonucleotide versus a longer strand of DNA. First, it should be noted that in DNA there are multiple targets, i.e., purine bases, whereas the oligonucleotides used in the above model studies usually contain only one target to maintain simplicity of analysis. In recent work, double stranded oligonucleotides with multiple target bases were analyzed by LC–MS to assess the relative binding preferences of benzo[*c*]phenanthrene to a known mutational "hot-spot" [34]. However, while more closely mimicking DNA, even double stranded oligonucleotides are not wound as tightly into a double helix due to their short length and the minimal number of hydrogen bonds. Therefore, there are implicit steric differences in the manner the carcinogen approaches the target base in the DNA macromolecule versus that in an oligonucleotide. With these factors in mind, the current study was undertaken to explore the sequence context of carcinogenic binding to DNA.

Of particular challenge when dealing with in vitro or in vivo systems is the enzymatic digestion of the DNA into a mixture rich in fragments containing the covalently bound carcinogen. Many different types of enzymes are available to accomplish the generation of oligonucleotides of different lengths. These enzymes include both restriction endonucleases and exonucleases. Some enzymes require specific recognition sequences in order to cut DNA into smaller sequences. Others, such as the exonucleases used in some of the studies mentioned above, cut monomeric units off the DNA strand until they become blocked by the bulky carcinogenic adduct. This hang-up can occur at different intervals from the adduct, generating multiple lengths of oligonucleotides. In addition, exonuclease digestion can generate longer strands containing multiple adduct sites, since there is not an end between the modified bases for the enzyme to attach to and initiate digestion of the strand. A third class of enzyme, non-specific endonucleases, do not require a recognition site, and therefore are the most universal, providing a good basis for analyses of unknown sequences. Benzonase is a random endonuclease which cleaves both single and double stranded DNA and RNA, producing a mixture of oligonucleotides ranging in length from two to eight units [35]. This enzyme has previously been used with alkaline phosphatase to study styrene oxide oligonucleotide adducts generated from an in vitro reaction with DNA by capillary zone electrophoresis–mass spectrometry (CZE–MS) [36–38].

Given the above considerations, it was necessary to initially examine a model system that could address efficiently many of these issues. Accordingly, the model system chosen for this study involved the reaction of calf thymus DNA with *N*-acetoxy-2-acetylaminofluorene followed by digestion with benzonase. AAAF is a reactive metabolite of the aromatic amine carcinogen, acetylaminofluorene [39], which modifies DNA by covalently binding to the N^2 and C8 positions of guanine [19]. The acetoxy group is lost in this modification procedure, producing acetylaminofluorene (AAF)-modified guanine residues, as illustrated in Fig. 1.

By digesting a small amount of DNA to oligonucleotides, multiple peaks are generated not only from differing compositions of modified sequences, but also from isomeric sequences of the same composition. This leads to the need for a high degree of sensitivity in the analysis. The following describes the first attempt to design an LC–MS/MS experiment to analyze modified oligonucleotides generated by the enzymatic digestion of an in vitro reaction of DNA with a known carcinogen.

2. Experimental/materials and methods

All solvents were purchased from Fisher Scientific (Pittsburgh, PA) and were HPLC grade. Deionized water was produced using a Millipore Q-pak filtration system (Bedford, MA).

2.1. Oligonucleotide standards

In order to mimic the results from the benzonase digestion of unmodified calf thymus DNA, reference oligonucleotides were synthesized using standard phosphoramidite

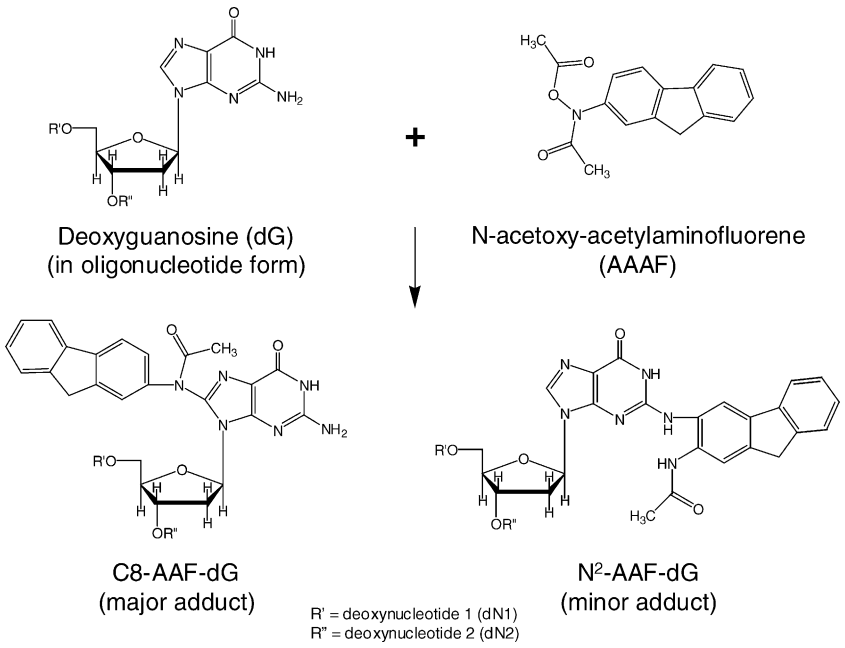

Fig. 1. Modification reaction of deoxyguanosine residues by *N*-acetoxy-2-acetylaminofluorene. This reactions produces two adducts with deoxyguanosine, one at the C8 position, which is the major adduct, and one at the N^2 position of the nucleobase, the minor adduct.

chemistry by Amitof (Boston, MA). Each standard was received lyophilized and was reconstituted in 0.1 ml of deionized water to create stock solutions. Standard 1 (NNN) is a mixture of all possible sequences of trimers, synthesized by introducing all four DNA nucleotides at each stage of the synthetic cycle. Standard 2 (NGN) is a mixture of all possible trimer sequences that contain a guanine base in the middle position. To accomplish this, all four nucleotides are introduced in the first step of the synthesis, only guanine is introduced during the second cycle, and then all four nucleotides are again added simultaneously during the third succession.

Both oligonucleotide standards were mixed with the carcinogen *N*-acetoxy-2-acetylaminofluorene to mimic the oligonucleotides produced by the enzymatic digestion. The AAAF stock (1.0 mg/ml) solution was prepared in acetonitrile, and mixed in molar equivalence to each of the oligonucleotide standards. The reaction mixtures were incubated at 37 °C for 18 h, and each sample was diluted 1:100 with deionized water with no further sample clean up prior to analysis.

2.2. DNA sample preparation

Calf thymus DNA was purchased from Sigma Aldrich (St. Louis, MO) and reconstituted in 10 mM citrate buffer (pH 6.0) to a concentration of 1.0 mg/ml. A 250 μl aliquot of this solution (250 μg of DNA) was mixed with the AAAF stock solution in 1:1 molar equivalence with the estimated guanine monophosphate content of the DNA. The mixture was incubated overnight (18 h) on a heat block at 37 °C and unreacted AAAF was extracted three times with ethyl acetate. The aqueous fraction was then dried down using a vacuum centrifuge at 1100 rpm (Savant Speed Vac Plus, Farmingdale, NY). This modification procedure was conducted as illustrated previously in our laboratory, and is estimated to yield nearly 1% modification [40].

2.3. Digestion procedure

Samples were reconstituted in 0.469 ml 5 mM Tris/10 mM magnesium chloride buffer (pH 8.0) and enzymes were added as follows: 6 μl of a 1:10 diluted solution of benzonase (EM Science, Gibbstown, NJ (now EMD Chemicals, Inc.)) and 0.94 units alkaline phosphatase (Sigma Aldrich, St. Louis, MO). Samples were incubated for 24 h at 37 °C. In the modification check experiments, the samples were reconstituted and enzymes were added according to a procedure previously used in our laboratory [40] to produce modified monomers. The following amounts of enzymes were added: 46.9 μg (4.69 μl) DNAse I (Sigma Aldrich, St. Louis, MO), 0.047 units (13.1 μl) phosphodiesterase (Amersham Biosciences, Piscataway, NJ), and 0.94 units (4.34 μl) alkaline phosphatase to digest the DNA to nucleosides. This allowed the modification procedure to be isolated from the oligonucleotide digestion procedure.

2.4. Solid phase extraction clean-up

Waters Oasis 1 cm^3 HLB columns (Milford, MA) were used to separate the modified oligonucleotides from the bulk unmodified sequences. Unmodified oligonucleotides were eluted with 5% (v/v) methanol and AAF-modified oligonucleotides were eluted with 50% (v/v) methanol. All fractions

were dried down and reconstituted in 0.1 ml of deionized water for analysis.

2.5. Instrumentation

An Agilent Technologies 1100 liquid chromatograph (Wilmington, DE) with a diode-array detector was coupled on-line to a ThermoFinnigan LCQ Classic mass spectrometer (San Jose, CA) controlled by Navigator 1.2 software. A 1 mm × 150 mm polystyrene divinylbenzene (PSDVB) polymer column (PRP-1) was purchased from Hamilton Chromatography (Reno, NV) run at a flow rate of 0.06 ml/min. This flow rate was introduced without splitting into the mass spectrometer. A solvent system was chosen that had previously demonstrated favorable results with oligonucleotide ladders [41]. 400 mM 1,1,1,3,3,3-hexafluoroisopropanol (Sigma Aldrich, St. Louis, MO) was prepared in 100% DI water (Solvent A) and in 50:50 water/methanol (Solvent B), and a gradient was run from 0 to 100% B over 30 min. For the modification check experiments, a 1 mm × 150 mm Symmetry C18 column from Waters was used with a 0.5% (v/v) acetic acid, water/methanol gradient (0–100% methanol in 10 min). For the detection of nucleosides, the LCQ was operated in the positive mode.

To analyze the oligonucleotide samples, the mass spectrometer was operated in full scan, negative ion detection mode to assess all masses present. The scan range for these experiments was *m/z* 600–1500. The samples were subsequently analyzed by LC–MS/MS, producing fragmentation data for the modified oligonucleotides. The scan ranges for the MS/MS spectra were *m/z* 150–2000. The LC–MS and LC–MS/MS interface conditions were held constant, and are as follows: spray voltage: 3.80 kV, capillary voltage: −4 V, capillary temperature: 200 °C, sheath gas: 25 arbitrary units, tube lens voltage: −25 V. The LC–MS/MS experiments were conducted at a collision width of 1.5 Da and relative collision energy of 30%.

A spreadsheet was constructed in Excel 2000 (Microsoft Office) to calculate the masses of oligonucleotides produced by the predicted enzymatic digestion. Sequences with the same composition were combined so that only unique masses were considered. The masses computed in the spreadsheet were used to create extracted mass chromatograms for the AAF-modified oligonucleotides, and the number of the peaks in each trace was compared to the number of isomers of each composition. Trimers with three unique bases in them would be expected to produce six separate peaks.

3. Results and discussion

There are two distinct aspects to the development of methodology to probe flanking base effects in covalent carcinogen binding to DNA. They involve: sample preparation (i.e., modification and enzymatic digestion of the DNA) and analytical procedures (i.e., solid phase extraction, liquid chromatography, and mass spectrometric detection and characterization). Initially, synthetic oligonucleotides were analyzed to verify the efficiency of the analytical method. To investigate the modification efficiency, a well-established enzymatic digestion was subsequently used to digest the DNA to nucleosides. As discussed in Section 1, benzonase is a random endonuclease, which cleaves DNA in a non-directional manner [35] into oligonucleotides of varying lengths. This enzyme may be sterically hindered by a bulky adduction to a DNA residue, rendering it unable to cut adjacent to the modified base. Due to the unknown effect of modification on the benzonase digestion, it was important to check for the possibility of both modified and unmodified dimers, trimers, and tetramers. The masses of possible digestion products were calculated using spreadsheets, and their fragmentation patterns were established using an Excel macro program.

3.1. Modification with N-acetoxy-2-acetylaminofluorene and enzymatic digestions

N-Acetoxy-2-acetylaminofluorene has been utilized extensively in our laboratory, to modify both synthetic oligonucleotides and calf thymus DNA [40,42–44]. The availability of this methodology made AAAF even more attractive because it provided a means to check the reaction yield, and isolate the reaction step from the digestion to oligonucleotides. In this experiment, the digestion to nucleosides verified the presence of guanosine nucleosides adducted with acetylaminofluorene (AAF-dG), and no adduction of the other three nucleosides was detected (data not shown). When benzonase was used to digest unmodified DNA, the result was a mixture of dimers and trimers. This complete digestion to such short fragments is controlled by the time in which the reaction is allowed to progress.

3.2. LC–MS methodology

In order to check the LC–MS methodology separately from the reaction and sample preparation steps, the synthetic standard mixture (NNN) served as a library by presenting all possible trimer sequences. This standard was analyzed before modification, after modification, and after modification and solid phase extraction in order to test each step independently. Since this mixture contained all theoretical sequences, it also was effective in determining whether the chromatographic separation was sufficient to resolve isobaric constituents. When developing the solid phase extraction method, the modified standard was used to determine which fraction contained the adducted sequences, and to assure separation from the bulk unmodified material.

A polystyrene divinylbenzene (PSDVB) polymer column was chosen since it had previously been successful in separating longer modified oligonucleotides [34]. The solvent system chosen, 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) in

water and methanol, had demonstrated baseline resolution with unmodified oligonucleotides and is appropriate for the analysis of hydrophobic analytes [41]. The combination of the PSDVB polymer column with the HFIP gradient produced satisfactory separation and detection of the modified sequences.

The summed full scan spectra from the LC–MS analyses verified the trimer masses present in the mixture, both unmodified and modified (data not shown). Spectra from the NNN standard and the digested calf thymus DNA were identical. These spectra were examined for modified dimer and tetramer masses as well; however, these masses were not detected. The modified trimer ions that were detected in the mixture all contained guanine residues, suggesting that guanine was the preferred target for the AAAF, and that no other bases were modified in high yield. This is substantiated by the results from the samples digested to nucleosides, where monoadducted AAF-dG was the only modified residue detected. To confirm the absence of modified dimers and tetramers, as well as trimer oligonucleotides that contained an AAF modification on a residue other than guanine, all the masses generated in the Excel spreadsheets were used to extract mass chromatograms. No other modified fragments were detected during this process.

Fig. 2 shows the LC–MS extracted mass chromatograms for the modified NNN standard (panel A) and the digested modified calf thymus DNA (panel B). Each trace represents a mass that corresponds to a trimer consisting of a modified guanine and two normal residues. The chromatograms from the digestion sample show distinct differences in both the number and relative intensities of the extracted mass peaks when compared to the standard trimer library sample. While the isomers are not baseline resolved, there is sufficient separation to conduct MS/MS experiments in order to determine their sequences. The retention times for the NNN standard and the digestion product do not correspond exactly to each other. However, these variations in retention time are not significant due to the mass spectrometric detection and characterization. Even with retention time shifts, the digestion peaks can be unequivocally identified using LC–MS/MS.

3.3. LC–MS/MS experiments

Once the compositions of the modified digestion products had been established, the next stage was to explore their sequences. Both the NNN standard and the digested sample were re-analyzed by LC–MS/MS in order to investigate the sequences present for each detected mass. Collision-induced dissociation (CID) experiments were conducted on all extracted masses that produced significant LC–MS peaks. An example is given in Fig. 3 for *m/z* 1080.6, which corresponds to a composition of a cytosine, a thymine, and an

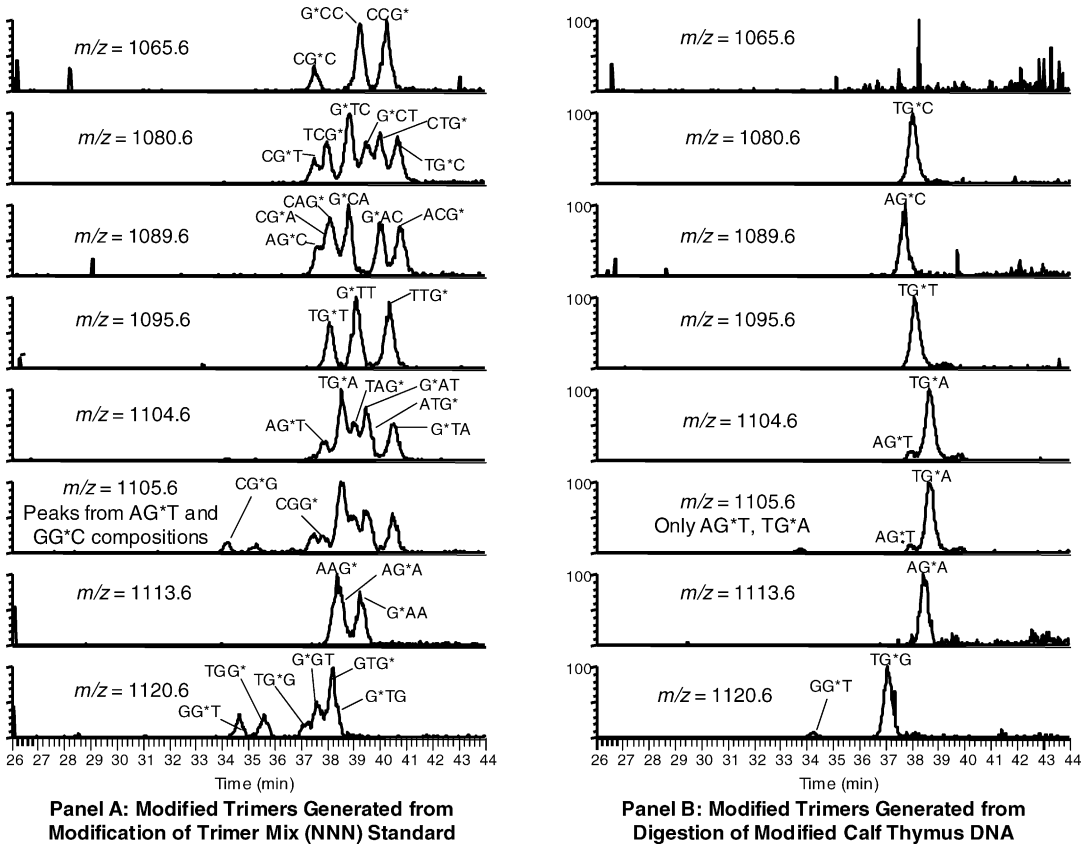

Fig. 2. Extracted LC–MS chromatograms for the standard trimer library (panel A) and the digestion products of the enzymatic cleavage of AAF-modified calf thymus DNA (panel B). Each peak is labeled with its specific sequence, and the site of adduction is indicated by an asterisk (*).

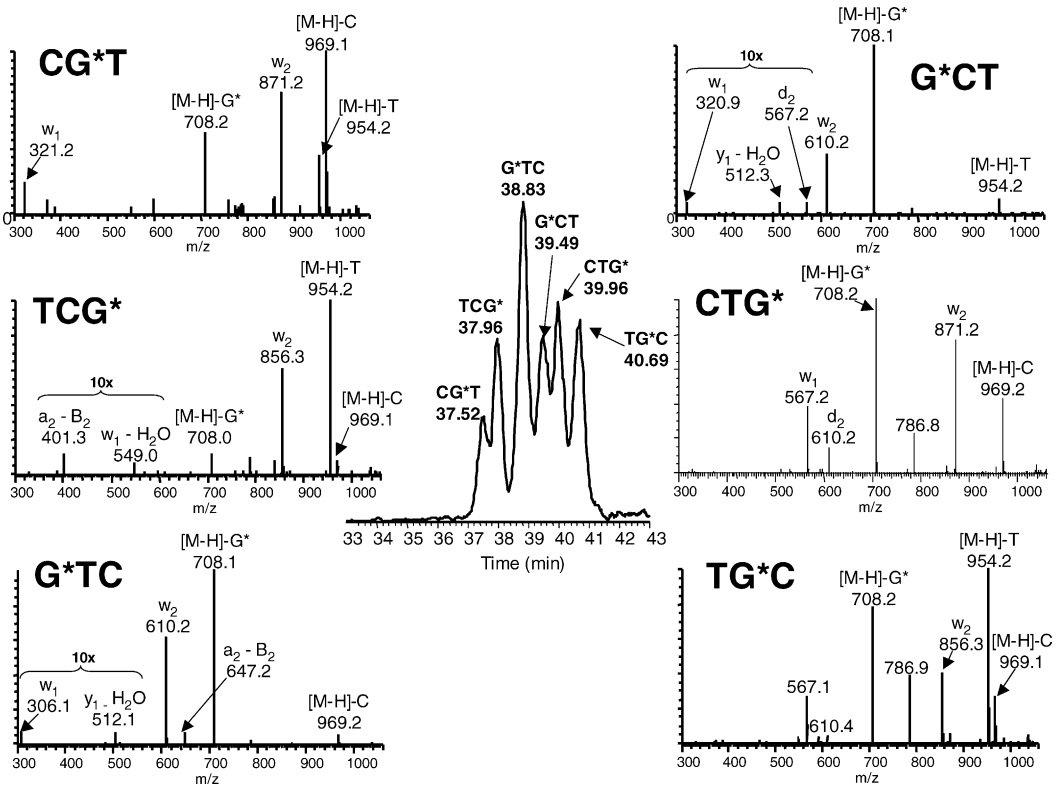

Fig. 3. Extracted mass chromatogram for *m/z* 1080.6 from the analysis of the modified NNN standard (from Fig. 2, panel A), MS/MS spectra for each peak in the chromatogram. Characteristic fragments are labeled using standard nomenclature [45,46] and were used to assign all the sequences.

acetylaminofluorene-adducted guanine residue (CTG*). The chromatogram from the NNN standard shows all possible isomers of the modified mass, accounting for the six peaks. Each of the isomer peaks is represented by a MS/MS spectrum. Characteristic fragmentation patterns were observed in all cases, allowing for sequence determination and peak assignment according to the well-established nomenclature for tandem MS oligonucleotide sequencing [45,46]. These characteristic fragmentation patterns were generated using an Excel macro program, and are illustrated in Fig. 4, with a detailed example shown for the sequence CTG*. The theoretical patterns were compared to the LC–MS/MS spectra of the standard mixture and the digestion sample. This was conducted to determine the presence or absence of characteristic ions that would allow for the assignment of the correct sequence to each chromatographic peak. For example, CG*T and CTG* share the same w_2 ion (*m/z* 871.2), which is usually one of the most intense peaks in the spectrum. However, these two sequences differ in their w_1 fragmentation (*m/z* 321.2 versus *m/z* 567.2), allowing simple characterization. In a more complex comparison, TCG* and TG*C share the same w_2 ion and both have weak w_1 signals, preventing facile sequence differentiation. Nevertheless, by amplifying the signal in certain regions of the spectra, it is possible to detect an a_2–B_2 peak corresponding to an apurinic ion in the spectrum of TCG*. This peak is absent in the other spectrum, and there is an additional characteristic d_2 ion at 871.2, indicative of the sequence TG*C.

Most of the extracted chromatograms for the digestion sample, shown in Fig. 2, contain a single peak for each mass, while some contain two peaks. This is in contrast to the NNN standard, where each extracted mass is represented by multiple signals, usually three to six peaks per trace. MS/MS experiments on the digestion product show that each peak represents an isomer with the modified guanine residue in the middle. It is interesting to note that there are differences in the relative abundances of certain sequences, and absence of

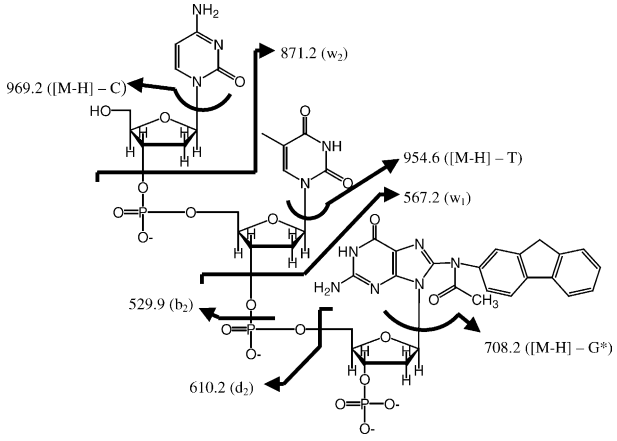

Fig. 4. Characteristic fragmentation patterns generated by the MS/MS sequencing of oligonucleotides [45,46], and specific example of the characteristic cleavages of the sequence CTG*.

some possible digestion products containing a modified G in the middle of the sequence. For example, the digestion trace for *m/z* 1080.6, contains a single peak (TG*C), while there is no peak for the sequence CG*T. Similarly, the *m/z* 1089.6 trace contains a single peak (AG*C) and the *m/z* 1104.6 trace contains one major peak (AG*T) and one minor peak (TG*A). There are two additional traces not shown in Fig. 2, *m/z* 1129.6, corresponding to AG*G, and 1145.6, corresponding to GG*G. The 1129.6 trace revealed all six possible isomers in the standard mixture, but only GG*A and AG*G in the digestion, consistent with the other traces illustrated in Fig. 2. All of these findings seem to point to the possibilities of an uneven base distribution in the DNA sequence or selection in the modification or digestion processes. Chemoselectivity in the modification of the intact DNA may be the most logical conclusion, as there seems to be a preference for modification of a guanine 5′ to a cytosine, whereas there is no such trend observed for modification next to a thymine. In a previous study, AAAF was shown to modify guanines along a plasmid sequence with as much as a 40-fold difference in reactivities [19]. In another related study, restriction enzymes were consistently and preferentially inhibited near the sequences T(C/G)TT(G/C) and T(G/C)CTT(G/C) in three types of DNA, indicative of AAAF modification in these regions [6–8]. These data, which our LC–MS studies are consistent with, point to the importance of the microenvironment of a target base when considering the binding preference of a carcinogen.

The mass chromatogram for the *m/z* 1105.6 trace did not reveal the presence of any trimers of composition GG*C in the digest (Fig. 2, panel B). The MS/MS data determined that the peaks in this trace are due to the sequences TG*A and AG*T, whose *m/z* is 1104.6. The inability to detect any of the GG*C composition in the *m/z* 1105.6 digestion trace demonstrates the need for improved mass resolution and sensitivity. In the present study, the collision width of the mass spectrometer was set to 1.5 Da, therefore including 1104.6 in its range. When the synthetic standard mixture was analyzed, certain sequences with the GG*C composition were detected in low abundance in the *m/z* 1105.6 trace. Even with their weak signals, these sequences were unequivocally identified by their MS/MS fragmentation by extracting the masses of characteristic fragments (w and d series) for each sequence. In the digestion analyses though, no sequences for the GG*C composition are seen, and the chromatographic peak profile perfectly shadows that of the *m/z* 1104.6 trace. Decreasing the width of the isolation window did not further elucidate any signals from the GG*C composition in the *m/z* 1105.6 trace. Although there seems to be a small peak at approximately 34 min in the *m/z* 1105.6 trace of the digestion similar to the one identified as CG*G in the trimer standard trace, MS/MS did not reveal the presence of this composition in the digestion. These findings are both interesting in a biologically context, and demonstrative of the need for high-resolution experiments and MS/MS analyses for confirmation of these results. In these experiments, it was difficult

Table 1
Summary of all AAF-modified sequences detected and characterized by LC–MS/MS

Mass	All sequences	NNN standard	NGN standard	Digestion
1065.6	CG*C	×	×	
	G*CC	×		
	CCG*	×		
1080.6	CG*T	×	×	
	TCG*	×		
	G*TC	×		
	G*CT	×		
	CTG*	×		
	TG*C	×	×	×
1089.6	AG*C	×	×	×
	CAG*	×		
	G*CA	×		
	G*AC	×		
	ACG*	×		
	CG*A	×	×	
1095.6	TG*T	×	×	×
	G*TT	×		
	TTG*	×		
1104.6	TG*A	×	×	× (Major)
	TAG*	×		
	G*AT	×		
	G*TA	×		
	AG*T	×	×	× (Minor)
	ATG*	×		
1105.6	GCG*	×		
	CGG*	×	×	
	G*GC	×	×	
	GG*C	×	×	
	G*CG	×		
	CG*G	×	×	
1113.6	AG*A	×	×	×
	AAG*	×		
	G*AA	×		
1120.6	GG*T	×	×	× (Minor)
	TGG*	×	×	
	TG*G	×	×	× (Major)
	G*GT	×	×	
	GTG*	×		
	G*TG	×		
1129.6	AGG*	×	×	
	G*GA	×	×	
	GG*A	×	×	× (Major)
	GAG*	×		
	G*AG	×		
	AG*G	×	×	× (Minor)
1145.6	GGG*		×	
	G*GG		×	
	GG*G			

NNN standard is library of all trimer sequences, NGN contains all trimer sequences with a guanine residue in the middle position, and the digestion sample was generated by cleaving modified calf thymus DNA with benzonase and alkaline phosphatase.

to distinguish between the two sequences that were a single Dalton apart in full scan. In addition, no multiply charged ions were observed in these experiments, which would have allowed detection of longer digestion segments. This may be due to the inability of the instrument to resolve peaks less than 1 Da apart. The LCQ is not a high-resolution instrument, and this illustrates the importance of LC–MS/MS data in determining sequence assignments.

In an effort to verify all of the above findings, an additional standard was obtained and modified in the same manner as the original NNN standard. This standard is the NGN mixture discussed in the Section 2, which is a mixture of trimers with guanine as the middle residue. The additional data from this standard confirmed that the sequences found in the digestion all have the modified guanine in the middle, and no enzymatic cleavage adjacent to a modified base was detected. This seems to indicate both a lack of adjacent modifications on the DNA strand, and the inability of benzonase to cleave next to a modified residue. As AAF modification is a fairly bulky adduction, this proposal is reasonable, especially since previous research studies have established that an AAF modification decreases the degree of subsequent modification on an adjacent guanine by other carcinogens [47,48], and that AAF-modified guanines causes many enzymes to be blocked [6–8,17,49–51]. Table 1 summarizes the results of the CID investigations, indicating the sequences identified in each of the three samples.

4. Conclusions

The results presented here illustrate the applicability of LC–MS and LC–MS/MS to the analysis of adducted segments of DNA. In this study, enzymatic digestion products from an in vitro reaction were compared to modified standards, and spreadsheets were utilized to provide an organized approach to the data analysis. This methodology introduces an opportunity to systematically investigate a complex model system. Standard mixtures presented all theoretical sequences, allowing for verification of sufficient chromatographic resolution and characterization by LC–MS/MS. When the extracted mass chromatograms of the digestion were compared to those of the standard, fewer peaks consistent with the masses of the modified oligonucleotides were observed. Additionally, a binding preference for guanines with a cytosine 3′ is demonstrated by AAAF in this experiment.

Comparisons of the CID data indicate the AAF-modified guanine is found predominantly in the middle of the digested trimer sequences. This seems to indicate both an inability of the AAAF to modify neighboring bases and a difficulty in enzymatic cleavage next to modified residues.

There is a general interest in investigating the relationship between the susceptibility of a gene for adduction and its predilection for mutation. On-line coupling of liquid chromatography and mass spectrometry is a facile way to explore whether any relationship exists between these events. Additional studies with carcinogens that exhibit a greater degree of sequence specific binding will be needed to fully validate this methodology. In order to conduct these studies, it may be necessary to adapt the LC–MS techniques to increase their sensitivity. Using capillary LC instead of normal bore columns may provide the increases necessary. In addition, chromatographic advances such as monolithic capillary columns may improve the resolution and sharpness of the LC peaks, thus increasing the sensitivity. Current studies have demonstrated the utility of polymeric monolith columns for separating ladders and plasmid digestions [52,53]. These experiments utilized alternative techniques such as ion pairing chromatography and ion exchange chromatography in conjunction with mass spectrometry. Utilizing an instrument with a larger mass range will allow analysis of larger carcinogens and may also lead to experiments with different enzymes that cut the sequence into longer fragments, providing more complete information about neighboring base effects on adduction. Another manner in which to proceed would be to alter the experimental conditions so as to aid in the detection of multiply charged ions, which may allow longer oligonucleotide segments to be detected. Another important consideration to take into account when proceeding with different enzymes is the fact that longer sequences will have more isomers, thus re-enforcing the need for additional sensitivity and resolution. Since there was difficulty in distinguishing between sequences 1 Da apart, caution would need to be taken when analyzing multiply charged ions that are less than 1 Da apart. Use of a high-resolution instrument, such as a Fourier transform mass spectrometer would facilitate these analyses.

The general approach and methodology presented here are promising for the detection as well as characterization of the sequence context of carcinogen adduction, and offers a pseudo-combinatorial approach of determining binding preferences to DNA by presenting the entire molecule as a target. Further experiments to expand the scope of this research are currently underway in our laboratory.

Acknowledgements

This work was supported by a grant from the National Institutes of Health (1RO1 CA69390).

References

[1] V.V. Lobanenkov, M. Plumb, G.H. Goodwin, P.L. Grover, Carcinogenesis 7 (1986) 1689.
[2] R.L. Rill, G.A. Marsch, Biochemistry 29 (1990) 6050.
[3] D.B. Reardon, C.A. Bigger, J. Strandberg, H. Yagi, D.M. Jerina, A. Dipple, Chem. Res. Toxicol. 2 (1989) 12.
[4] J.A. Miller, E.C. Miller, Environ. Health Perspect. 49 (1983) 3.
[5] L. Verna, J. Whysner, G.M. Williams, Pharmacol. Ther. 71 (1996) 83.

[6] N.J. Combates, S.A. Winkle, J. Biomol. Struct. Dyn. 10 (1992) 63.
[7] M.A. Mallamaci, M.L. Bascoy, J. Brown, N.J. Combates, S.A. Winkle, J. Biomol. Struct. Dyn. 10 (1992) 83.
[8] M.A. Mallamaci, D.P. Reed, S.A. Winkle, J. Biomol. Struct. Dyn. 10 (1992) 73.
[9] S.N. Rodin, A.S. Rodin, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 12244.
[10] E.C. Miller, J.A. Miller, Cancer 47 (1981) 2327.
[11] R.W. Pero, J. Seidegard, D.G. Miller, Prog. Clin. Biol. Res. (1986) 225.
[12] M. Hollstein, B. Shomer, M. Greenblatt, T. Soussi, E. Hovig, R. Montesano, C.C. Harris, Nucleic Acids Res. 24 (1996) 141.
[13] S.P. Hussain, P. Amstad, K. Raja, M. Sawyer, L. Hofseth, P.G. Shields, A. Hewer, D.H. Phillips, D. Ryberg, A. Haugen, C.C. Harris, Cancer Res. 61 (2001) 6350.
[14] A. Puisieux, S. Lim, J. Groopman, M. Ozturk, Cancer Res. 51 (1991) 6185.
[15] Z. Feng, W. Hu, W.N. Rom, F.A. Beland, M.S. Tang, Biochemistry 41 (2002) 6414.
[16] P. Belguise-Valladier, R.P. Fuchs, Biochemistry 30 (1991) 10091.
[17] P. Belguise-Valladier, R.P. Fuchs, J. Mol. Biol. 249 (1995) 903.
[18] D. Burnouf, P. Koehl, R.P. Fuchs, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 4147.
[19] R.P. Fuchs, J. Mol. Biol. 177 (1983) 173.
[20] P. Koehl, P. Valladier, J.F. Lefevre, R.P. Fuchs, Nucleic Acids Res. 17 (1989) 9531.
[21] I.B. Lambert, R.L. Napolitano, R.P. Fuchs, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 1310.
[22] T.H. Broschard, N. Koffel-Schwartz, R.P. Fuchs, J. Mol. Biol. 288 (1999) 191.
[23] E.A. Stemmler, M.V. Buchanan, G.B. Hurst, R.L. Hettich, Anal. Chem. 66 (1994) 1274.
[24] E.A. Stemmler, M.V. Buchanan, G.B. Hurst, R.L. Hettich, Anal. Chem. 67 (1995) 2924.
[25] M. George, J.M.Y. Wellemans, R.L. Cerny, M.L. Gross, K. Li, E.L. Cavalieri, J. Am. Soc. Mass Spectrom. 5 (1994) 1021.
[26] N.P. Christian, S.M. Colby, L. Giver, C.T. Houston, R.J. Arnold, A.D. Ellington, J.P. Reilly, Rapid Commun. Mass Spectrom. 9 (1995) 1061.
[27] P.M. Chiarelli, X. Gu, A.A. Aldridge, H. Wu, Anal. Chim. Acta 368 (1998) 1.
[28] J.K. Gooden, J. Byen, L. Chen, E.G. Rogan, E.L. Cavalieri, M.L. Gross, Intl. J. Mass Spectrom. Ion Process. 169/170 (1997) 241.
[29] Z. Wang, K.X. Wan, R. Ramanathan, J.S. Taylor, M.L. Gross, J. Am. Soc. Mass Spectrom. 9 (1998) 683.
[30] N. Tretyakova, B. Matter, A. Ogdie, J.S. Wishnok, S.R. Tannenbaum, Chem. Res. Toxicol. 14 (2001) 1058.
[31] J. Preisler, F. Foret, B.L. Karger, Anal. Chem. 70 (1998) 5278.
[32] J. Preisler, P. Hu, T. Rejtar, E. Moskovets, B.L. Karger, Anal. Chem. 74 (2002) 17.
[33] N. Tretyakova, B. Matter, R. Jones, A. Shallop, Biochemistry 41 (2002) 9535.
[34] A. Harsch, J.M. Sayer, D.M. Jerina, P. Vouros, Chem. Res. Toxicol. 13 (2000) 1342.
[35] G. Meiss, P. Friedhoff, M. Hahn, O. Gimadutdinow, A. Pingoud, Biochemistry 34 (1995) 11979.
[36] P. Janning, W. Schrader, M. Linscheid, Rapid Commun. Mass Spectrom. 8 (1994) 1035.
[37] W. Schrader, M. Linscheid, Arch. Toxicol. 71 (1997) 588.
[38] W. Schrader, M. Linscheid, J. Chromatogr. A 717 (1995) 117.
[39] E.C. Miller, U. Juhl, J.A. Miller, Science 153 (1966) 1125.
[40] S.M. Wolf, P. Vouros, Chem. Res. Toxicol. 7 (1994) 82.
[41] A. Apffel, J.A. Chakel, S. Fischer, K. Lichtenwalter, W.S. Hancock, Anal. Chem. 69 (1997) 1320.
[42] R.S. Annan, R.W. Giese, P. Vouros, Anal. Biochem. 191 (1990) 86.
[43] A. Harsch, P. Vouros, Anal. Chem. 70 (1998) 3021.
[44] S.M. Wolf, P. Vouros, Anal. Chem. 67 (1995) 891.
[45] S.A. McLuckey, G.J. Van Berkel, G.L. Glish, J. Am. Soc. Mass Spectrom. 3 (1992) 60.
[46] S.A. McLuckey, S. Habibi-Goudarzi, J. Am. Chem. Soc. 115 (1993) 12085.
[47] B. Said, M.K. Ross, T. Salib, R.C. Shank, Carcinogenesis 16 (1995) 3057.
[48] M.K. Ross, B. Said, R.C. Shank, Toxicol. Sci. 53 (2000) 224.
[49] E. Seeberg, R.P. Fuchs, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 191.
[50] P. Belguise-Valladier, H. Maki, M. Sekiguchi, R.P. Fuchs, J. Mol. Biol. 236 (1994) 151.
[51] M.T. Hess, U. Schwitter, M. Petretta, B. Giese, H. Naegeli, Chem. Biol. 3 (1996) 121.
[52] D. Sýkora, F. Svec, J.M.J. Fréchet, J. Chromatogr. A 852 (1999) 297.
[53] H. Oberacher, C.G. Huber, Trends Anal. Chem. 21 (2002) 166.