

1999

Electrospray tandem mass spectrometry of oligonucleotides and ligand-oligonucleotide adducts

Paula Iannitti
University of Wollongong

Recommended Citation

Iannitti, Paula, Electrospray tandem mass spectrometry of oligonucleotides and ligand-oligonucleotide adducts, Doctor of Philosophy thesis, Department of Chemistry, University of Wollongong, 1999. <http://ro.uow.edu.au/theses/1151>

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

***Electrospray Tandem Mass Spectrometry of
Oligonucleotides and Ligand-Oligonucleotide
Adducts***

Thesis submitted by

Paula Iannitti, BSc(Hons)

*in partial fulfilment of the requirements
for the degree of*

Doctor of Philosophy

Department of Chemistry
University of Wollongong
November, 1999

DECLARATION

I declare that this thesis is my own work, and, except where due acknowledgment has been given in the text of this thesis, does not to my knowledge contain work which has been published or written by another person, or that has been submitted or accepted for the award of a degree or diploma at any institute of higher learning.

Paula Iannitti

ACKNOWLEDGMENTS

I would like to acknowledge the following people for their help and support during the course of this work.

- *My supervisor Associate Professor Margaret Sheil for her continual support, guidance and perseverance throughout the duration of my PhD, and particularly in the final stages.*
- *My supervisor Dr Geoffrey Wickham, to whom I am especially grateful for his enthusiasm and guidance and for providing a constant flow of ideas in this research.*
- *Dr Allan Weimann for driving the research on oligonucleotide fragmentation and for his help with instrumentation, especially the Autospec.*
- *Mr Larry Hick and Mr Roger Kanitz for all their help in the lab throughout this work.*
- *Dr Trevor Lewis for his assistance in the early stages of trouble shooting HPLC problems.*
- *The numerous past and present members of the mass spectrometry lab I came to know during my years at Wollongong, Dr Keiryn Bennett, Dr Susan Hunt, Dr Greg Kilby, Mr Glen Stutchbury, Ms Jacqui Boschenok, Ms Jenny Vazquez and Mr Santiago Vazquez, for their friendship and camaraderie.*
- *My present supervisor, Professor Carol Robinson, for providing me with the opportunity to work at Oxford and for her patience and understanding whilst I have been completing this work in my spare time.*
- *My husband, Mr Mark Tito for his support, encouragement and understanding.*

Finally, I would especially like to thank my parents, Bruno and Louisa Iannitti, for giving me every opportunity to pursue my education and for always believing in me.

PUBLICATIONS

Parts of the work outlined in this thesis have appeared in the following publications:

- 1999** A. Weimann, P. Iannitti-Tito and M.M. Sheil
Characterisation of product ions in high energy tandem mass spectra of protonated oligonucleotides formed by electrospray ionisation
International Journal of Mass Spectrometry (1999) submitted
- 1997** P.Iannitti, G.Wickham, M.M.Sheil
High sensitivity and fragmentation specificity in the analysis of drug-DNA adducts by electrospray mass spectrometry
Journal of the American Chemical Society (1997) 119, 6, 1490-1491

In addition, other work in the University undertaken by the candidate has resulted in the following publications:

- 1995** G.Wickham, P.Iannitti, J.Boschenok and M.M.Sheil
Electrospray ionisation mass spectrometry of covalent ligand-oligonucleotide adducts: evidence for specific duplex ion formation
Journal of Mass Spectrometry (1995) 30.S197-S203
- 1995** G.Wickham, P.Iannitti, J.Boschenok and M.M.Sheil
The observation of a hedamycin-d(CACGTG)₂ covalent adduct by electrospray mass spectrometry
FEBS Letters (1995) 360, 231-234
- 1995** S.F. Ralph, P.Iannitti, R.Kanitz, M.M.Sheil
Determination of the relative stabilities of metal-antibiotic and metal-cryptand complexes by electrospray ionisation mass spectrometry
European Mass Spectrometry (1995) 2, 173-9

ABSTRACT

Electrospray ionisation tandem mass spectrometry (ESI-MS/MS) has been used to examine the fragmentation of a range of unmodified oligonucleotides (from tetramers to octamers) and alkylated oligonucleotide adducts on a hybrid sector-time-of-flight mass spectrometer. The detailed characterisation of all the product ions observed in the MS/MS spectra of two hexanucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' was undertaken using a combination of high resolution MS and MS/MS of product ions generated by in-source collisional activation of oligonucleotides in both positive and negative ion modes. These experiments provide significant new information concerning the identities of the major product ions and their relative propensities for subsequent fragmentation. A comprehensive study of the influence of base composition, base location and precursor ion charge (in both positive and negative ion modes) on the resulting MS/MS spectra of oligonucleotides, showed the major fragmentation pathways *i.e.* the formation of w and a-B sequence ions, and ions arising from loss of charged and neutral bases, to be strongly influenced by the most probable location of the charges and the location and identity of the bases. The effect of base composition was particular evident in the observed lack of preference for loss of neutral thymine in both modes.

In the second part of this work, the utility of ESI-MS/MS for studying the sequence selectivity of alkylating ligands is demonstrated with three alkylating agents, hedamycin, DC-92B and N-bromohexylphenanthridinium bromide. The MS/MS spectra of the resulting adducts show highly specific fragmentation pathways which allow the facile identification of the site(s) of binding of the ligands on the oligonucleotides. Differences in sequence selectivity exhibited by the three ligands are also demonstrated. Overall this work demonstrates that ESI-MS/MS is a powerful technique for the characterisation of oligonucleotides and oligonucleotide adducts.

TABLE OF CONTENTS

<i>Declaration</i>	<i>ii</i>
<i>Acknowledgments</i>	<i>iii</i>
<i>Publications</i>	<i>iv</i>
<i>Abstract</i>	<i>v</i>
 <i>Chapter 1</i>	
<i>INTRODUCTION</i>	<i>1</i>
<i>1.1 NUCLEIC ACIDS</i>	<i>3</i>
<i>1.1.1 Structure and Nomenclature</i>	<i>3</i>
<i>1.1.2 Physical Properties</i>	<i>6</i>
<i>1.1.3 Double Helical DNA</i>	<i>6</i>
<i>1.2 LIGAND-DNA INTERACTIONS</i>	<i>10</i>
<i>1.2.1 Covalent Interactions</i>	<i>10</i>
<i>1.2.2 Noncovalent Interactions</i>	<i>11</i>
<i>1.2.3 Ligands used in this Study</i>	<i>12</i>
<i>(a) Hedamycin and DC92-B</i>	<i>13</i>
<i>(b) N-bromohexylphenanthridinium bromide</i>	<i>16</i>
<i>1.3 STRUCTURAL CHARACTERISATION</i>	<i>17</i>
<i>1.3.1 Mass Spectrometry</i>	<i>19</i>
<i>1.3.2 Mass Spectrometry of DNA and Ligand/DNA Complexes - Past Applications</i>	<i>20</i>
<i>(a) Plasma Desorption (PD)</i>	<i>20</i>
<i>(b) Fast Atom Bombardment (FAB)</i>	<i>20</i>
<i>(c) Matrix Assisted Laser Desorption Ionisation (MALDI)</i>	<i>22</i>
<i>(d) Electrospray Ionisation (ESI)</i>	<i>24</i>
<i>1.4 ELECTROSPRAY IONISATION MASS SPECTROMETRY</i>	<i>24</i>
<i>1.4.1 Origins and Historical Development</i>	<i>24</i>
<i>1.4.2 Principles and Instrumentation</i>	<i>26</i>
<i>1.4.3 The Process of Ionisation</i>	<i>28</i>
<i>1.4.4 Characteristics of ESI Mass Spectra</i>	<i>32</i>
<i>1.4.5 Tandem Mass Spectrometry</i>	<i>34</i>
<i>1.4.6 Electrospray Ionisation Mass Analysers</i>	<i>35</i>
<i>1.5 ELECTROSPRAY MASS SPECTROMETRY OF OLIGONUCLEOTIDES</i>	<i>36</i>
<i>1.5.1 Sample Preparation and Solvent Optimisation</i>	<i>38</i>
<i>1.5.2 Charge State Reduction Studies</i>	<i>40</i>
<i>1.5.3 Molecular Mass Measurement</i>	<i>42</i>
<i>1.5.4 Sequence Analysis and Applications</i>	<i>44</i>
<i>1.5.5 Noncovalent Complexes</i>	<i>45</i>
<i>1.6 TANDEM MASS SPECTROMETRY OF OLIGONUCLEOTIDES</i>	<i>47</i>
<i>1.6.1 Fast Atom Bombardment Mass Spectrometry Studies</i>	<i>48</i>
<i>1.6.2 Studies of Oligonucleotide Fragmentation by Electrospray Tandem Mass Spectrometry</i>	<i>54</i>
<i>(a) Negative Ions</i>	<i>54</i>

	<i>(b) Positive Ions</i>	63
1.6.3	<i>Characterisation of Oligonucleotide Modifications</i>	66
1.7	SUMMARY	67
1.8	OUTLINE OF THIS WORK	69
	REFERENCES	70

Chapter 2

	EXPERIMENTAL	77
2.1	OLIGONUCLEOTIDES	77
2.1.1	<i>Preparation and Purification</i>	77
2.1.2	<i>Determination of Concentration</i>	80
2.2	LIGAND-OLIGONUCLEOTIDE ADDUCTS	80
2.2.1	<i>Hedamycin and DC92-B</i>	80
2.2.2	<i>N-bromohexylphenanthridinium bromide</i>	81
2.3	ELECTROSPRAY MASS SPECTROMETRY	82
2.3.1	<i>Autospec™ oa-TOF Magnetic Sector Mass Spectrometer</i>	82
	<i>(a) Instrumentation</i>	82
	<i>(b) Tandem Mass Spectrometry</i>	85
	<i>(c) Tandem Mass Spectrometry of Source-Generated Product Ions</i>	85
	<i>(d) Instrument Tuning and Calibration</i>	86
2.3.2	<i>Quattro™ Triple Quadrupole Mass Spectrometer</i>	86
	<i>(a) Instrumentation</i>	86
	<i>(b) Tandem Mass Spectrometry</i>	87
	<i>(c) Instrument Tuning and Calibration</i>	88
2.4	SAMPLE ANALYSIS	89
2.4.1	<i>Oligonucleotides</i>	89
2.4.2	<i>Ligand-Oligonucleotide Adducts</i>	91
	REFERENCES	92

Chapter 3

	ELECTROSPRAY IONISATION MASS SPECTROMETRY OF UNMODIFIED OLIGONUCLEOTIDES	93
3.1	INTRODUCTION	93
3.2	OPTIMISATION OF ESI-MS OF OLIGONUCLEOTIDES	93
3.2.1	<i>Solvent Composition</i>	93
3.2.2	<i>Sample Concentration and Flow Rate</i>	96
3.3	FACTORS INFLUENCING THE CHARGE STATE DISTRIBUTION OF OLIGONUCLEOTIDE ANIONS	97
3.3.1	<i>Electrospray Ionisation Source Parameters</i>	97
3.3.2	<i>Addition of Acids and Bases</i>	102
3.4	SUMMARY	107
	REFERENCES	108

Chapter 4**ELECTROSPRAY TANDEM MASS SPECTROMETRY OF OLIGONUCLEOTIDES:
Characterisation of product ions generated by negative and positive ions MS/MS on hybrid
sector-TOF instrumentation.****109**

4.1	INTRODUCTION	109
4.2	OLIGONUCLEOTIDE FRAGMENTATION NOMENCLATURE	110
4.3	COMPARISON OF HYBRID-TOF AND ION TRAP TANDEM MASS SPECTRA OF OLIGONUCLEOTIDE ANIONS	112
4.4	CHARACTERISATION OF NEGATIVE ION ELECTROSPRAY TANDEM MASS SPECTRA OF OLIGONUCLEOTIDES	116
4.4.1	Product Ion Types Observed	117
	(a) Backbone and Nucleotide Fragments	124
	(b) Mononucleotide Ion Types	129
	(c) Polynucleotide Ion Types – Sequence Ions and Internal Ions	132
4.4.2	Analysis of Product Ion Fragmentation Pathways	139
	(a) 5'-d(CACGTG)-3'	140
	(b) 5'-d(CGTACG)-3'	150
	(c) Summary	163
4.5	CHARACTERISATION OF POSITIVE ION ELECTROSPRAY TANDEM MASS SPECTRA OF OLIGONUCLEOTIDES	165
4.5.1	Product Ion Types Observed	165
	(a) Backbone and Nucleotide Fragments	171
	(b) Polynucleotide Ion Types	175
4.5.2	Analysis of Product Ion Fragmentation Pathways	178
	(a) 5'-d(CACGTG)-3'	178
	(b) 5'-d(CGTACG)-3'	184
4.6	COMPARISON OF POSITIVE ION AND NEGATIVE ION FRAGMENTATION OF OLIGONUCLEOTIDES	191
4.6.1	Fragmentation Pathways of Singly Charged Oligonucleotide Anions	191
4.6.2	Fragmentation Pathways of Singly Charged Oligonucleotide Cations	210
4.6.3	Comparison of Positive and Negative Ion Fragmentation Pathways of Oligonucleotides	217
	REFERENCES	223

Chapter 5**FACTORS INFLUENCING THE FRAGMENTATION OF OLIGONUCLEOTIDE ANIONS
IN ELECTROSPRAY TANDEM MASS SPECTRA****224**

5.1	INTRODUCTION	224
5.2	THE EFFECTS OF OLIGONUCLEOTIDE SEQUENCE	225
5.2.1	Fragmentation of Unibase Oligonucleotides	226
5.2.2	Fragmentation of Oligonucleotides of Heterogeneous Base Composition	232
5.2.3	Summary	249
5.3	THE EFFECTS OF PRECURSOR ION CHARGE STATE	253
5.3.1	Nucleobase Loss	260
	(a) Neutral Base Loss	264
	(b) Base Anion Loss	268
5.3.2	Formation of w-type and (a-B)-type ions	281
5.3.3	Summary	297
	REFERENCES	299

Chapter 6**SEQUENCE ANALYSIS OF DNA ALKYLATING AGENTS BY ELECTROSPRAY
TANDEM MASS SPECTROMETRY 300**

6.1	INTRODUCTION	300
6.2	DNA ALKYLATING AGENTS USED IN THIS STUDY	301
6.3	HEDAMYCIN BINDING TO DNA - RESULTS AND DISCUSSION	302
6.3.1	Hedamycin binding to the hexanucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(GCCGGC)-3', and 5'-d(CGGCCG)-3'	302
	(a) Hedamycin binding to 5'-d(CACGTG)-3'	302
	(b) Hedamycin binding to 5'-d(CGTACG)-3'	310
	(c) Hedamycin binding to 5'-d(CACGTG)-3' versus 5'-d(CGTACG)-3': Discussion	317
	(d) Hedamycin binding to 5'-d(GCCGGC)-3'	318
	(e) Hedamycin binding to 5'-d(CGGCCG)-3'	321
	(f) Hedamycin binding to 5'-d(GCCGGC)-3' versus 5'-d(CGGCCG)-3': Discussion	327
6.3.2	Hedamycin binding to the octanucleotides 5'-d(GCGTACGC)-3', 5'-d(ACGTACGT)-3', and 5'-d(TCGTACGA)-3'	328
	(a) Hedamycin binding to 5'-d(GCGTACGC)-3'	328
	(b) Hedamycin binding to 5'-d(ACGTACGT)-3'	333
	(c) Hedamycin binding to 5'-d(TCGTACGA)-3'	336
	(d) Hedamycin binding to octanucleotides: Discussion	338
6.4	DC92-B BINDING TO DNA - RESULTS AND DISCUSSION	341
6.4.1	DC92-B binding to 5'-d(CACGTG)-3',	342
6.4.2	DC92-B binding to 5'-d(CGTACG)-3',	352
6.4.3	DC92-B binding to 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3': Discussion	354
6.5	N-BROMOHEXYLPHSNANTHRIDINIUM BROMIDE BINDING TO DNA - RESULTS AND DISCUSSION	355
6.5.1	PhenC ₆ Br binding to 5'-d(CACGTG)-3',	356
6.5.2	PhenC ₆ Br binding to 5'-d(CGTACG)-3',	361
6.5.3	PhenC ₆ Br binding to 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3': Discussion	366
6.5	SUMMARY OF RESULTS FOR ALKYLATING AGENTS	367
	REFERENCES	371

Chapter 7**CONCLUSIONS 372**

REFERENCES	386
-------------------	-----

APPENDICES

Chapter 1

INTRODUCTION

Deoxyribonucleic acid (DNA), encodes the genetic information fundamental to life. The proper functioning of cells, and thus the maintenance of life, is critically dependent on the precise transcription and translation of the DNA code¹. In the process of DNA transcription, each strand of the DNA double helix acts as a template, transmitting this genetic code to messenger RNA (mRNA), which, in turn, is translated into the synthesis of proteins. The flow of genetic information from mRNA to proteins occurs via the triplet DNA code - the linear sequence of three nucleobases of a single DNA strand specifies an amino acid, and the sequence of base triplets is a linear code of the amino acid sequence of the protein². These processes, however, can be affected when DNA becomes damaged or mutated, for example by substitution of bases, deletion or insertion of base pairs, errors in base pairing, loss of a base, covalent binding of two bases, or covalent modification of bases³. In many cases, damaged DNA can be corrected by repair enzymes which work by detecting and removing the damaged section and then regenerating the correct sequence using the unaffected strand as a template for synthesis. Although these repair mechanisms produce high fidelity DNA replication, they are not always capable of detecting or repairing damaged DNA, so that in some instances modification of DNA can severely affect or impair cell growth and division, and can even lead to cell death⁴. There is now a general recognition that the development of disease states resulting in tumour growth or carcinogenesis is inevitably preceded by some form of DNA modification⁵. Consequently, many of these studies aimed at investigating the mechanism of action of chemical mutagens and carcinogens has been directed towards understanding how such molecules interact with, and modify, DNA.

The critical role DNA plays in cell reproduction also makes it a potential target for chemotherapeutic agents which aim to prevent the reproduction of diseased cells.

Many of the most effective drugs currently employed in chemotherapy are cytotoxic compounds that bind covalently to nucleophilic sites in DNA, or its associated proteins⁶. Examples of these include the platinating agents cisplatin and carboplatin⁷, and alkylating agents such as nitrogen mustards⁸, polyhydric dibromides⁹, 2-haloethylnitrosoureas¹⁰ and cyclophosphamide¹¹. The majority of these anticancer drugs, however have damaging side-effects, since poor selectivity for tumour cells can result in significant damage to normal cells. Furthermore, tumour cells can develop a resistance to drug treatment through either enhanced functioning of enzymatic DNA repair mechanisms, and/or increases in the levels of low molecular weight thiols (such as glutathione) that react with the drug resulting in the loss of its ability to bind to DNA¹². Hence there is on-going interest in the development of improved antitumour drugs. Recently, this has been in part driven by the discovery of new natural products displaying novel modes of interaction with DNA or possessing interesting reactive functionality. The antitumour activity of many of these compounds is thought to relate to the sequence-selective manner by which they interact with DNA. Hence, the identification of the structural features of these molecules that are responsible for the observed DNA sequence selectivity is of great interest to synthetic chemists who are aiming to use rational drug design methods to develop antitumour agents that can be targeted to specific sequences in DNA¹³⁻¹⁵. If structural information can be further correlated with antitumour activity, this can ultimately enhance the process of designing effective analogues. Thus, it is crucial for antitumour drug design to determine the precise nature of the chemical interaction between the ligand or drug and its cellular receptor, DNA.

Hence, the study of DNA modification by synthetic or naturally occurring ligands is an area of considerable biological significance. Efforts are constantly being made to enhance further the understanding of the structural basis of ligand-DNA interactions at the molecular level. The progress in the structural analysis of ligand-DNA adducts to date has predominantly involved the high resolution molecular characterisation methods of X-Ray crystallography and Nuclear magnetic resonance

(NMR) spectroscopy¹⁶. These methods, although providing detailed structural information, lack sensitivity. A number of other biochemical and chemical methods have also been employed for probing structural modifications of DNA which provide much greater sensitivity, albeit lack the structural detail provided by crystallography and NMR, and in some cases may be quite time-consuming. Therefore, in recent times, mass spectrometry has emerged as a potentially powerful tool for the rapid identification and characterisation of ligand-DNA interactions.

1.1 NUCLEIC ACIDS

1.1.1 Structure and Nomenclature

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are linear polymers composed of monomeric nucleotide units. Each nucleotide consists of three substituents - a sugar, a nitrogen heterocyclic base and a phosphate group. In nucleic acids, the 3'-hydroxyl of one nucleoside is connected to the 5'-hydroxyl of the next nucleoside unit by a phosphodiester linkage thus forming a sugar-phosphate backbone (shown in figure 1.1). The sugar in DNA is a 2'-deoxy-D-ribose in its cyclic furanoside form, whereas in RNA it is a D-ribose. The nucleobases in nucleic acids are either derivatives of purine or pyrimidine. The purine bases are adenine (Ade or A) and guanine (Gua or G), and the pyrimidine bases are thymine (Thy or T), uracil (Ura or U) and cytosine (Cyt or C). Of these, thymine is only found in DNA, and uracil is only found in RNA. In nucleosides, the C-1 atom of the sugar is bonded to one of the nitrogenous bases (N-1 of pyrimidine or N-9 of purine) via a β -glycosidic linkage. The nucleosides in DNA (deoxyribonucleosides) are deoxyadenosine, deoxyguanosine, deoxythymidine, and deoxycytidine, and are abbreviated to dN where N is one of the nucleobases. In RNA, the ribonucleosides are adenosine, guanosine, cytidine, and uridine, and are abbreviated to rN. Figure 1.2 shows the structures of the nucleosides in DNA and RNA, and the numbering given to the atoms of the sugar ring and heterocyclic bases.

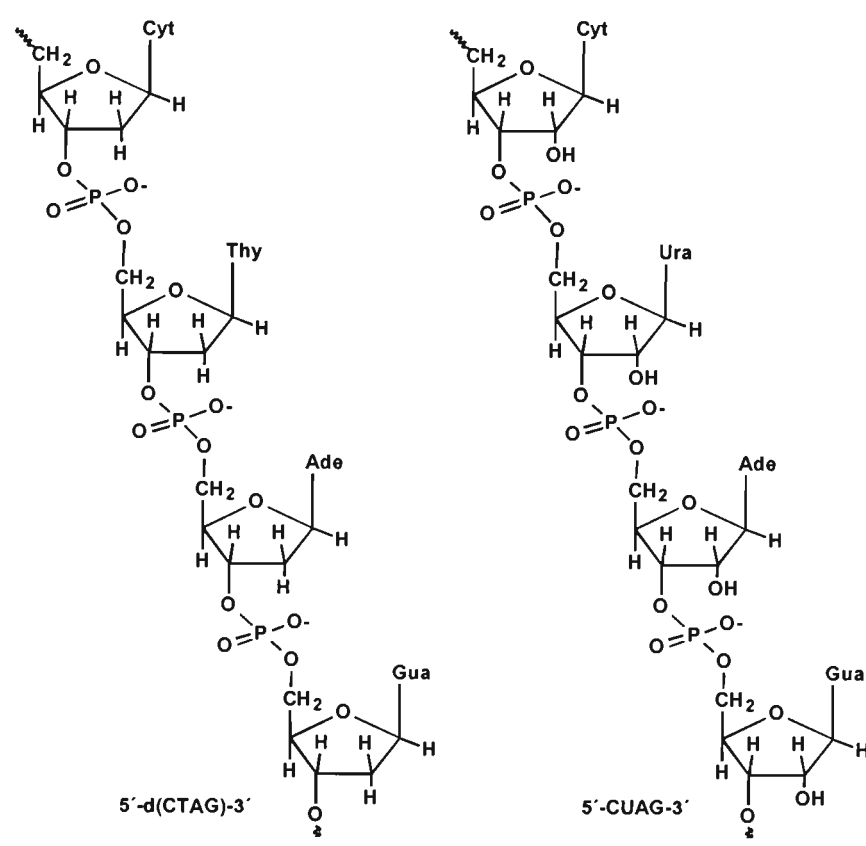


Figure 1.1: Linear structures of oligodeoxyribonucleotides and oligoribonucleotides.

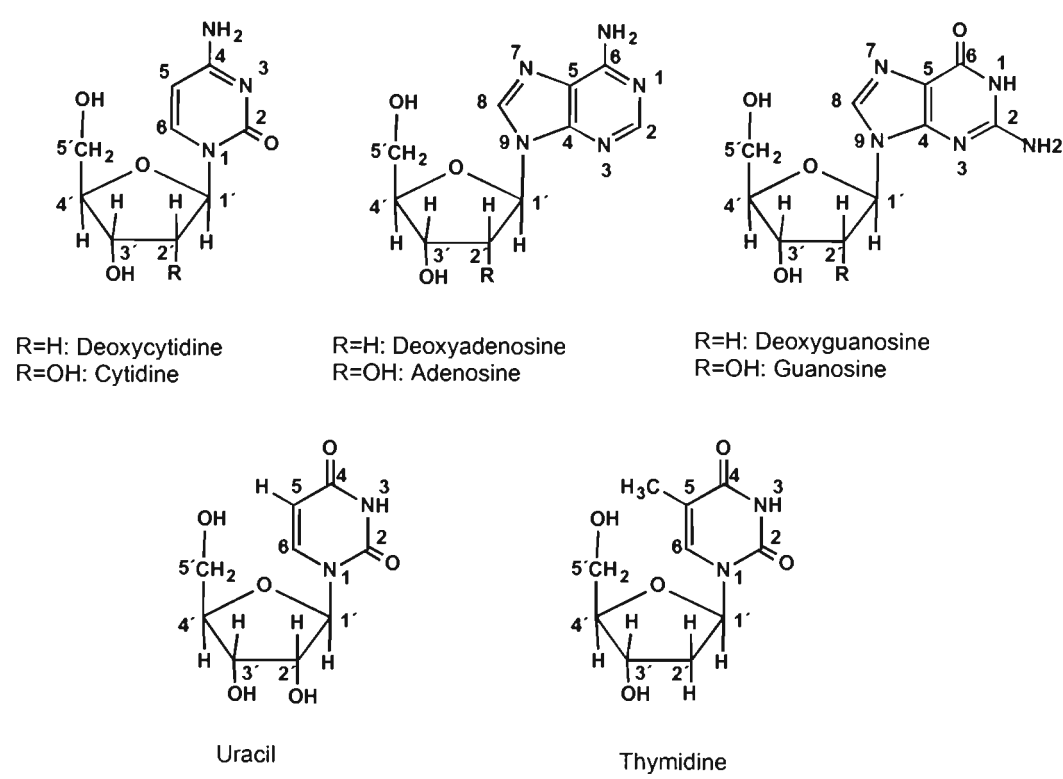


Figure 1.2: The structures of the nucleosides in DNA (R=H) and RNA (R=OH).

Nucleotides are nucleosides which are esterified by a phosphate group at one of the hydroxyl groups of the sugar. Esterification of the phosphate group can be either at the hydroxyl group attached to the C-3 (a 2'-deoxyribonucleoside 3'-monophosphate) or C-5 (a 2'-deoxyribonucleoside 5'-monophosphate) atom of the sugar, although the common site of esterification for natural nucleotides is at the hydroxyl group attached to the C-5 atom. The ribonucleoside 5'-monophosphates are also referred to as adenylate, cytidylate, guanylate, and uridylate (for A, C, G, U respectively), and similarly, the deoxyribonucleoside 5'-monophosphates are named deoxyadenylate, deoxycytidylate, deoxyguanylate, and deoxythymidylate (for A, C, G, T respectively). The nucleotides are designated dNMP when they are derived from DNA (which indicates a 2'-deoxyribonucleoside monophosphate), and rNMP or NMP (which indicates a ribonucleoside monophosphate) in RNA. The site of attachment of the phosphate group at the C-3' or C-5' hydroxyl group is represented by a 3'- or 5'- before dNMP and NMP. When a nucleotide is represented only as dNMP or NMP, without designation as to the location of the phosphate group, it is generally assumed that attachment is to the C-5' hydroxyl group. It should be noted that in common usage a more abbreviated notation, although ambiguous, is generally employed for nucleosides and nucleotides, whereby the letter 'p' is used to represent a monophosphate group rather than 'MP'. For example, dNp designates a 2'-deoxyribonucleoside 3'-monophosphate, d(pN) is a 2'-deoxyribonucleoside 5'-monophosphate, pN is a nucleoside 5'-monophosphate, and Np is a nucleoside 3'-monophosphate.

The specific sequence of an oligonucleotide is defined by the order of bases from the 5'-terminus to the 3'-terminus of the strand. The notation used to represent a DNA or RNA chain is the consecutive sequence of its nucleobases from the 5'-hydroxyl terminus to the 3'-hydroxyl terminus, e.g. 5'-CACGTG-3'. For DNA sequences, the letter 'd' is normally placed before the sequence, so as to stipulate that the sugar groups are deoxyribose, e.g. 5'-d(CACGTG)-3'. As is the case with nucleotide abbreviations, phosphate groups are indicated by the letter 'p' and are sometimes written into the short-hand notation of an oligonucleotide, e.g. 5'-d(CpApCpGpTpG)-3'.

1.1.2 Physical Properties of Nucleic Acids

In solution under physiological conditions (pH 7), nucleic acids exist as polyanionic species which arises from the negative charges on the phosphate backbone. (the phosphodiester groups deprotonate at *ca.* pH 1). For monoester phosphate groups in nucleotides, or terminal monoester phosphate groups in nucleic acids, the second proton is lost at pH 6-7. The sugar groups in both DNA and RNA remain neutral in solution up to a pH of *ca.* 12. At higher pH (>12), the terminal 3'-hydroxyl groups of the sugars in DNA, and the 2'-hydroxyl groups of each ribose as well as the terminal 3'-hydroxyl in RNA are deprotonated. The capacity of nucleobases, nucleosides or nucleotides to act as a proton donator or a proton acceptor in solution is essentially governed by its acid-base properties. Table 1.1 lists the pK_a values for the protonation and deprotonation of the five nucleobases, and their corresponding nucleosides, and nucleotides.

Table 1.1: pK_a values for the protonation and deprotonation of the five nucleobases, and for these nucleobases in nucleosides, and nucleotides recorded at approx. 20°C and at an ionic strength, μ=0. The site of protonation/deprotonation is indicated in brackets under the base name. This data was extracted from Blackburn¹⁷.

	pKa for Protonation			pKa for Deprotonation		
	Adenine (N-1)	Cytosine (N-3)	Guanine (N-7)	Guanine (N-1)	Thymine (N-3)	Uracil (N-3)
Nucleobase	4.18	4.68			10.00	9.52
Nucleoside	3.52	4.17	3.3	9.42	9.93	9.38
3'-Nucleotide	3.70	4.43	3.5	9.84	-	9.96
5'-Nucleotide	3.88	4.56	3.6	10.00	10.47	10.06

The data in the table shows that nucleobase units (whether present as nucleobases or as constituents of nucleosides and nucleotides) are neutral within the range pH 5-9. Below pH 5, protonation of the N-3 of cytosine, N-1 of adenine and N-7 of guanine occurs. It should be noted that other basic nitrogens are available for protonation in the major tautomeric forms (i.e. the amino-and keto- tautomeric forms) of adenine (N-3 and N-7) and guanine (N-3), albeit it has been found that protonation of these sites is much less favourable (Kochetkov *et. al*¹⁸ referenced within¹⁹). From the pK_a values it can be seen

that cytosine is the most basic of the nucleobases, followed by adenine and then guanine. Both thymine and uracil lack an available protonation site, and thus the capacity to behave as bases. These bases, however, can display slight acidic properties at higher pH. In the keto-tautomeric forms of thymine and uracil, the nitrogen at the N-3 position has very weak acidic properties which enables deprotonation of these nucleobases at pH 9-11. Likewise, the N-1 of guanine in its keto-tautomeric form is also acidic in character, with deprotonation occurring in the pH range 9-10.

1.1.3 Double Helical DNA

The genetic information encoded in a nucleic acid is contained in the exact sequence of its bases. An integral aspect of the flow of this genetic information lies in the hydrogen-bonding interactions which occur between purine and pyrimidine bases. Hydrogen bonds form between "Watson Crick" base pairs as shown in figure 1.3 where guanine is paired with cytosine (G-C) and adenine is paired with thymine (A-T), or uracil (A-U).

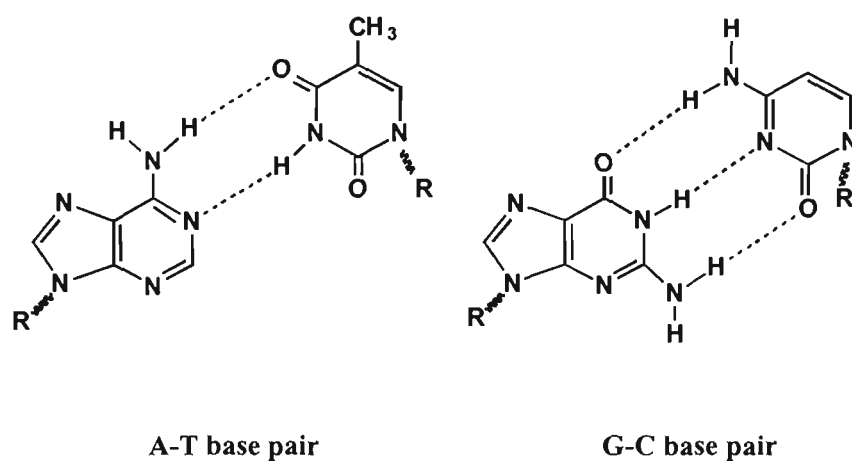


Figure 1.3: "Watson Crick" base pairs.

The specificity of these base-pairing interactions is the essence of accurate transcription and translation of genetic information in DNA and RNA. In DNA, base-pairing occurs

between two complementary strands lined up antiparallel to each other which results in the formation of a double helical structure (figure 1.4), which is further stabilised by interactions resulting from the parallel stacking of bases. The DNA double helix shown in figure 1.4 is the B-type conformation (*B-DNA*), and has a right-handed twist. The DNA double helix can exist in other conformations (*A-DNA* and *Z-DNA*), however *B-DNA* is the predominant form which exists at pH 7. The DNA double helix can be denatured through heating or by increasing or decreasing the pH to cause ionisation of the bases³.

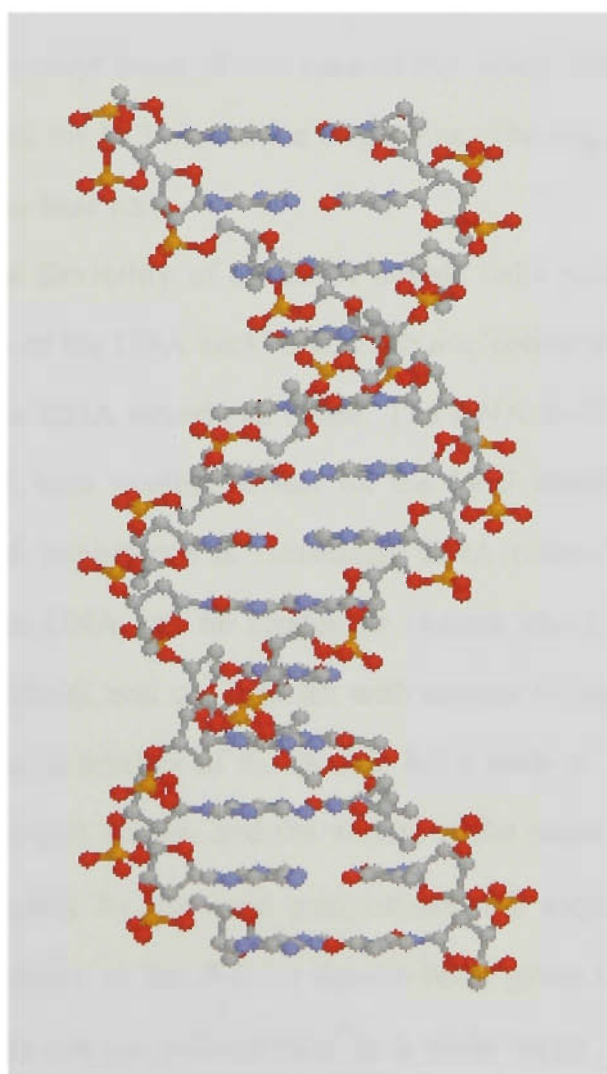


Figure 1.4: The DNA double helix.

The melting temperature of a DNA double helix, T_m (the temperature at which half of the double helix unwinds), is very much dependent of the bases present in its sequence since G-C base pairs have greater stability than A-T base pairs. The interactions

between A-T base pairs involve two hydrogen bonds whereas those of G-C base pairs which involve three hydrogen bonds. The parallel stacking of neighbouring G-C base pairs is also greater than that between neighbouring A-T base-pairs.

The *B-DNA* form of the double helix is a conformationally diverse molecule. An important attribute of this double helical structure is the presence of major and minor grooves along the exterior of the DNA double helix. These grooves exist since, in each base-pair, the glycosidic bonds between the nucleobases and sugar groups are both located on the same side of the base-pair, and also because the base pairs are not centred on the axis of the helix. The minor groove is the region on the side of the base pair, extending from the glycosyl bond of one base to the other, which contains the O-2 of cytosine or thymine and the N-3 of adenine or guanine. The major groove comprises the remaining region of the base pair.

The structural flexibility of the DNA double helix stems predominantly from the rotational freedom of the DNA backbone. Each nucleotide unit of DNA contains six bonds about which the DNA strand can rotate. The DNA double helix can be kinked, bent, and supercoiled with minimal effect on the local double helix structure. This attribute is of integral importance in facilitating DNA interactions with a variety of proteins. The bases in DNA can be rotated or twisted about the N-glycosidic bond (termed *propeller twisting*), and can also tilt with respect to adjacent bases in a strand (called *base roll*). Characteristics of the double helix such as helical twist, base roll, kinking, bending, base-pair angles, and the widths of the major and minor grooves are determined predominantly by the base composition and sequence of DNA¹⁷. These structural features inherent in the *B-DNA* double helix gives rise to the diversity this macromolecule displays in its participation in a wide range of biological functions. Furthermore, the conformation of the DNA double helix, which is to a large extent sequence-dependent²⁰, is an important aspect in the sequence-selective recognition of DNA by both biological macromolecules such as proteins and other DNA strands, as well as small DNA-interactive molecules.

1.2 **LIGAND-DNA INTERACTIONS**

The diverse array of structural features inherent in the composition and conformation of the DNA double helix presents a number of potential sites for both covalent and noncovalent molecular interactions. Of the whole host of known DNA-interactive agents, there are a larger number of different classes of compounds in existence which are of particular interest, either because of their biological effects, or the novel DNA-binding characteristics they display. Some important examples of these include mutagens, carcinogens, antibiotics, dyes, and metal complexes¹². The mutagenic, carcinogenic and/or antitumour properties exhibited by many of these molecules is thought to involve their ability to covalently modify DNA. Therefore, knowledge of their modes of interaction is of crucial importance to understanding their observed activity. There is also significant interest in the noncovalent binding modes exhibited by a range of recently-discovered natural products, many of which have been found to possess antitumour activity. The novel binding features of these molecules have also become the subject of extensive research for their potential applications in the development of new and improved antitumour agents.

1.2.1 **Covalent Interactions**

A range of small molecules can interact covalently, or coordinatively with the heterocyclic bases of DNA either by nucleophilic attack (mainly at the C-6 and C-4 of pyrimidines), or by electrophilic attack (at most nitrogen and oxygen atoms of the bases). These compounds include known carcinogens, such as vinyl chloride, dimethyl sulphate, and methyl methanesulphonate which are alkylating agents that interact directly with DNA forming covalent bonds principally to the nitrogen and oxygen atoms of purine bases. Alternatively, some carcinogenic compounds are metabolically converted *in vivo* to alkylating agents that bind covalently to DNA. For instance, arylamines are metabolised to aryl hydroxylamine derivatives which bind preferentially to the C-8 of deoxyguanosines. Many polycyclic aromatic hydrocarbons (PAH's) are

metabolised to diol epoxides through oxidation by cytochrome P-450 oxidases and reaction with epoxide hydroxylases¹². The epoxides then bind readily to the exocyclic amines on the nucleobases in DNA.

DNA alkylation is also a key cellular event in the mechanism of action of many widely used clinical anticancer agents². A number of naturally occurring antitumour antibiotics have been found to bind covalently to DNA following some initial noncovalent interaction which is responsible for the selective toxicity these compounds display. There are several mechanisms by which covalent binding of antibiotics can occur. For example, the antibiotics daunomycin, adriamycin, actinomycin D and mitomycin C require initial metabolic reduction, whereas for the pyrrolo[1,4]benzodiazepine antibiotics (such as anthramycin and tomamycin), loss of methanol or water precedes covalent attachment¹². The binding of the antitumour antibiotics hedamycin and DC92-B, which were examined in this thesis, is discussed in detail in section 1.2.3(i).

1.2.2 *Noncovalent Interactions*

Ligands that reversibly bind to DNA (i.e. via noncovalent interactions) encompass a wide variety of compounds including metal complexes, antitumour drugs, antibiotics and carcinogens. There are three modes by which ligands can interact noncovalently with nucleic acids. First, binding can occur via electrostatic interactions between the molecule and the exterior of the DNA helix. Such interactions are generally non-specific with respect to the polar bases, sugars and phosphate groups. Examples include condensation type reactions between cations and nucleic acids, and outside stacking of cations along the anionic DNA phosphate-sugar backbone²¹.

Reversible groove-binding to DNA can also occur at the major and minor grooves via van der Waal's forces and hydrogen bonds between the ligand and the edges of the base pairs in the grooves. In general, large molecules such as oligonucleotides and proteins interact with the major groove, while the majority of the small DNA-interactive ligands exhibit binding in the minor grooves of DNA. A common structural feature of

many small minor groove-binding molecules is the presence of several aromatic rings which are bonded together in such a way that free rotation is possible about the rings²¹. This enables the molecule to adopt a helical conformation that aligns with the conformation of the minor groove. To a large extent, minor groove binding molecules preferentially bind to the narrower A-T base pair regions rather than the wider G-C regions since the bond between the guanine 2'-amino group and the carboxyl group of cytosine makes binding in the major grooves sterically unfavourable³. Examples of minor groove binding molecules include netropsin, distamycin and Hoescht 33258²¹.

The third type of reversible binding to DNA is intercalation. Ligands that undergo interactions of this type generally contain a planar aromatic ring structure which can insert between the planar aromatic rings of adjacent base pairs of the DNA double helix²¹. Simple intercalators, such as ethidium bromide and proflavin, contain few side chains and have a positive charge located on the chromophore. More complex intercalators, such as daunomycin and actinomycin, contain neutral chromophores and are structurally more complex owing to the presence of side-chains which do not intercalate. Intercalation generally results in distortions of the DNA backbone such as lengthening of DNA, bending, and unwinding of the double helix²².

1.2.3 *Ligands used in this Study*

The present sequencing study involved detailed examination of the binding characteristics of three alkylating agents: hedamycin, DC92-B, and *n*-bromohexylphenanthridinium bromide (phenC₆). These all comprise an alkylating functional group linked to a DNA-intercalating chromophore and can be further divided into two structural classes. The compounds hedamycin and DC92-B are anthraquinone-based natural products that contain bisepoxide alkylating side-chains, and phenC₆ is a synthetic derivative of phenanthridine which contains an alkylating moiety linked to a quaternised phenanthridine ring .

(a) *Hedamycin and DC92-B*

Hedamycin and its structural analogue DC92-B (shown in figure 1.5), are naturally occurring antitumour antibiotics belonging to the pluramycin family, which are a class of antibiotics comprising highly substituted 4*H*-anthra[1,2,*b*]pyran-4,7,12-triones²³.

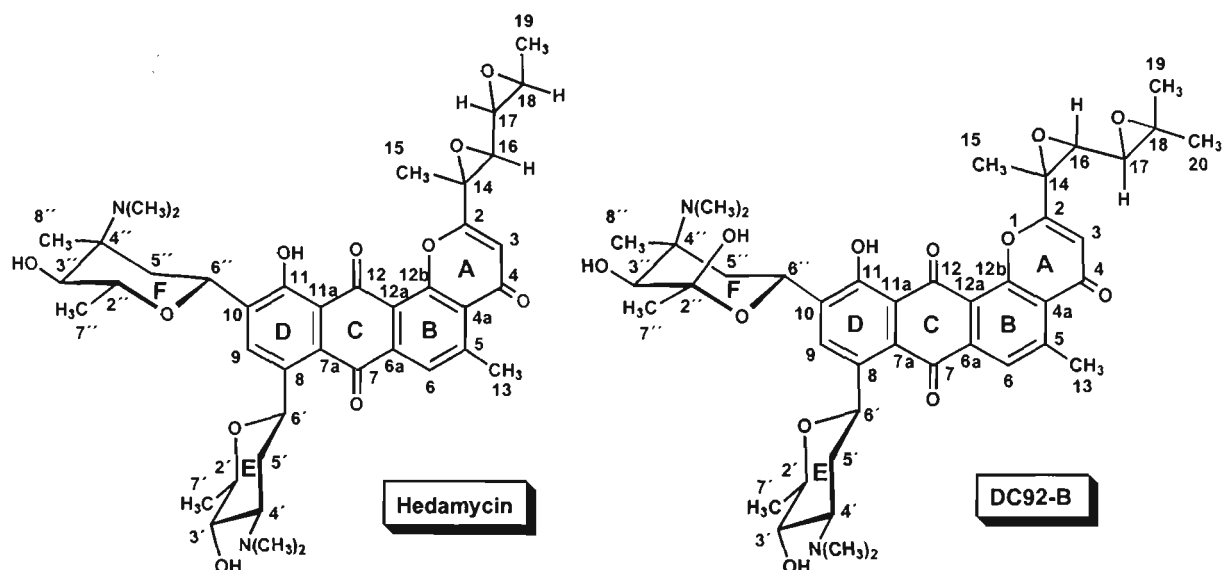


Figure 1.5: Structures of hedamycin and DC92-B.

The isolation of hedamycin from *Streptomyces griseoruber* (strain C-1150) was reported initially by Schmitz *et al.*²⁴ after the discovery of its potent antibacterial activity during the 1960s. Subsequent studies demonstrated the ability of hedamycin to inhibit the growth of HeLa cells in tissue culture, and certain transplanted rodent tumours²⁵, and RNA and DNA polymerases²⁶. White and White²⁷ showed that upon binding of hedamycin to double-stranded DNA, strong complexes are formed which significantly increases the melting temperature (T_m) of the DNA double helix. This was suggested to be a result of hedamycin interacting with DNA via both covalent and noncovalent modes of binding. Subsequently it was proposed by Jernigan *et al.*²⁸ that hedamycin binds to DNA first by rapid intercalation followed by covalent binding to the bases at a much slower rate. Sequin *et al.* undertook a series of chemical and physical studies directed towards the elucidation of the chemical structure of hedamycin, and as a result

of extensive spectroscopic and X-ray crystallographic work, a report of the chemical structure of hedamycin was presented^{23,29,30}.

Hedamycin contains the following structural features²³:

- (i) an anthraquinone structural motif labelled rings B, C, and D and the pyrone ring labelled ring A a 4*H*-anthra[1,2-*b*]pyran-4,7,12-trione chromophore;
- (ii) two amino sugar substituents, an anglosamine which is a 2,3,6-trideoxy-3-(dimethylamino)-*D*-arabino-hexose (ring E) bound at the C-8 position and an *N,N*-dimethylvancosamine which is a 2,3,6-trideoxy-3-(dimethylamino)-*C*-methyl-*L*-lyxo-hexose (ring F) bound at the C-10 position; and
- (iii) and a bis(epoxide)-containing sidechain attached to the C-2 position of the pyrone ring (ring A)

The structurally-related compound DC92-B was first reported in the literature in the late 1980s by Takahashi *et al.*³¹ following its isolation from *Actinomadura*. The structure of DC92-B was then determined by Yasuyawa *et al.*³² from chemical and spectroscopic studies. It was found that the characteristic features of DC92-B that distinguish this compound from hedamycin are the presence of a hydroxyl group attached to the C-2'' position of the ring F aminosugar, and the bis(epoxide) sidechain which contains a *cis*-epoxide and terminal dimethyl groups (as opposed to a *trans*-epoxide and a terminal methyl group in hedamycin).

Studies of the sequence selectivity of hedamycin were undertaken by Bennet *et al.* using piperidine-induced strand breaking methods and high resolution gel electrophoresis. It was found that DNA alkylation by hedamycin occurred solely at guanine bases with a preference exhibited for guanines in 5'-TG sequences. Prakash *et al.*¹³ later studied the sequence selectivity of both hedamycin and DC92-B with both a *Taq* DNA polymerase linear extension method and the piperidine-induced strand breaking procedure used by Bennet *et al.* Their results clearly demonstrated the high sequence selectivity of hedamycin for alkylation of isolated guanines in 5'-pyGT (where py is one of the pyrimidine bases C or T), with 5'-CGT sites being preferred to 5'-TGT sites (these findings have been supported by subsequent studies conducted by

Sun *et al.*³³). Selectivity was observed also for alkylation of guanines in 5'G(G)_n. Their corresponding results with DC92-B showed this compound has the same sequence selectivity as hedamycin, even though the substitution and stereochemistry of the bis(epoxide) sidechain is very different to that of hedamycin. These workers also showed that hedamycin and DC92-B exhibit the same sequence selectivity with intact human cells³⁴.

Very recently, Pavlopoulos *et al.*³⁵ conducted a 2D-NMR study on the complex formed between hedamycin and double-stranded oligodeoxyribonucleotide d(CACGTG)₂. These workers proposed a 3-dimensional model for the interaction between hedamycin and the hexanucleotide DNA duplex on the basis of distances predicted from observed intermolecular NOEs in their NMR data. This model is shown in figure 1.6.

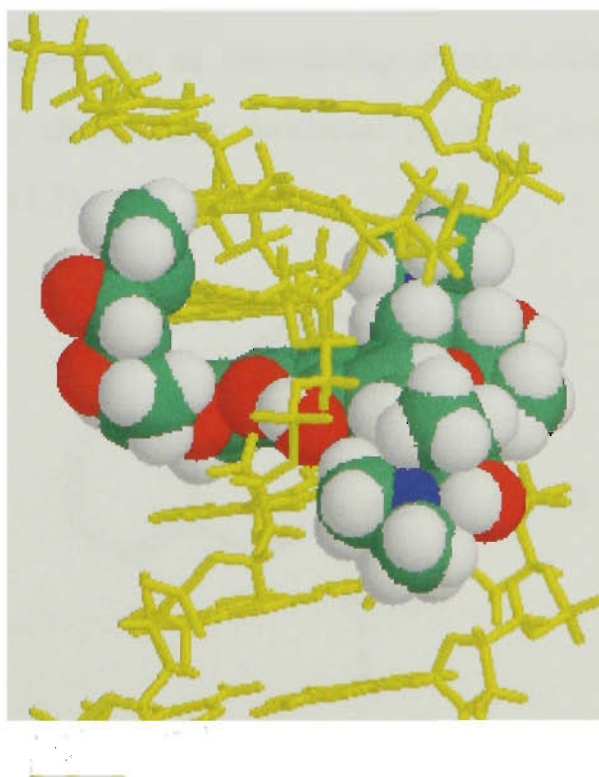


Figure 1.6: NMR-derived model of hedamycin/d(CACGTG)₂.

The planar anthraquinone chromophore is stacked between the 5'-CG-3' base pairs and threads the DNA double helix placing the two aminosugar groups in the minor groove, and the alkylating bis(epoxide) sidechain in the major groove. The ring E anglosamine sugar group lies to the 3'-side to the intercalated 5'-CG-3' base pairs and the ring F N,N-dimethylvancosamine sugar group lies to the 5' side of the base pairs. Hedamycin alkylation was proposed to occur via covalent attachment of the bis(epoxide) sidechain at the N-7 of the central guanine. This was inferred from the resulting increase in lability of the C-8 hydrogen of guanine upon hedamycin binding. These workers suggested that the hedamycin selectivity for 5'-CGT-3' sites arises from the both hydrophobic interactions between the bis(epoxide) sidechain and the methyl group of the 3'-thymine, and the favourable manner in which the bis(epoxide) and sugar substituents fit into the major and minor grooves respectively.

(b) *N*-bromohexylphenanthridinium bromide

The *n*-bromoalkylphenanthridinium bromides (phenC_nBr) are a class of synthetic ligands consisting of an intercalating phenanthridine chromophore that is tethered to a primary alkylbromide functional group by a polymethylene chain of variable length (figure 1.7).

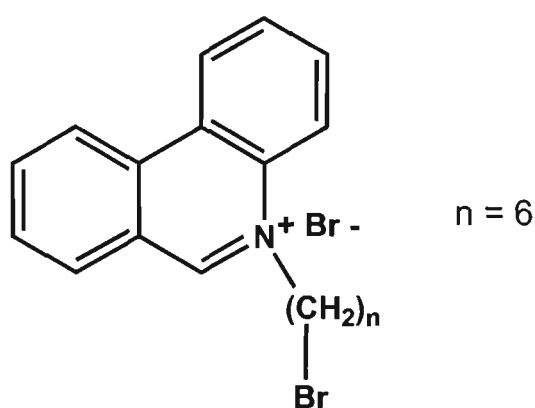


Figure 1.7: Structure of *n*-bromoalkylphenanthridinium bromide, $n=6$.

Interest in these synthetic intercalating-alkylating compounds arose through their use as precursors to platinum(II) diammine complexes linked to a phenanthridinium group and

Wickham *et al.*³⁶ have investigated their DNA-binding properties and sequence selectivity. These compounds were not found to display significant antitumour activity, and their reactivity was found to be dependent on the length of their polymethylene linker chains with the phenC₁₀Br compound showing significantly lower levels of alkylation compared with phenC₆Br. From gel electrophoretic sequencing studies, phenC₆Br was shown to preferentially alkylate guanines (most likely at the N-7 position in the major groove) in 5'-GT-3' sequences and in runs of guanines [(Gp)_n], whereas the phen₁₀Br compound exhibited a high degree of selectivity for 5'-GT-3' sequences and also showed a very slight preference for alkylation of adenines³⁶.

In subsequent studies conducted with plasmid DNA and intact human cells, Murray *et al.*³⁷ employed *Taq* DNA polymerase extension methods to examine the effect of linker chain length of *n*-bromoalkylphenanthridine bromides on the sequence selectivity of these compounds. The compounds with linker chain lengths of *n*=4, 6, and 8 were found to preferentially alkylated guanines, and, in particular, in runs of consecutive guanines. The *n*=10 compound was found to alkylate both guanines and adenines with a preference for alkylation in consecutive G sequences or consecutive A sequences. In the case of alkylation of isolated guanines, 5'-GA-3' sequences were favoured for the *n*=4 compound whereas for the *n*=6, 8, and 10 compounds, a preference for alkylation at 5'-GT-3' sequences was observed. The results obtained for the *n*=6 and 10 compounds were consistent with the previous observations of Wickham *et al.*³⁶.

1.3 **STRUCTURAL CHARACTERISATION OF LIGAND-DNA COMPLEXES**

Knowledge of the precise nature of ligand binding to DNA is fundamental to an enhanced understanding of the structural basis for ligand-DNA interactions. The information gained from the structural characterisation of ligand-DNA adducts can potentially provide an insight into the mechanism of action of mutagens and carcinogens, and can also shed some light on the structural basis for the antitumour activity exhibited by a variety of synthetic ligands and natural products. The pressing

need to discern the fundamental relationship which exists between the structure of a ligand and its biological activity is important for gaining an understanding of the mechanism of carcinogenesis, and, moreover, lies at the very heart of research into the design of improved antitumour agents. Studies of ligand-DNA interactions generally involve a combination of both chemical and biological techniques and are aimed at both obtaining detailed structural information on ligand-DNA complexes and assessing the extent and nature of specific types of ligand binding to DNA.

Two of the commonly used high resolution structural methods for the characterisation of ligand-DNA complexes are X-Ray Crystallography and Nuclear magnetic resonance spectroscopy (NMR). X-Ray Crystallography is a highly specific technique which has been employed extensively for the structural analysis of nucleic acids and for elucidation of the structural implications of DNA modifications¹⁶. A requirement of this technique, however, is that samples must be capable of being crystallised. In some cases this can be difficult to achieve, and quite time-consuming. Furthermore, since samples have to be in solid form for analysis, X-Ray crystallography cannot be applied to the monitoring of DNA structure in solutions comparable to the physiological environment. Nuclear magnetic resonance spectroscopy (NMR) can provide structural information of nucleic acids in solution by measuring proton-proton coupling constants and from through-space nuclear Overhauser effect (NOE) derived distances¹⁶. NMR is a widely used method for the characterisation of intact modified oligonucleotides, since, of all the spectroscopic methods, NMR is one of the most powerful because it yields high resolution structural information concerning the conformation of oligonucleotides in solution. The major disadvantage of NMR, however, is its relatively low sensitivity since μmols or 10s of μmols of sample are required. Mass spectrometry (MS) has recently emerged as a powerful complementary method to these conventional high resolution techniques for the structural characterisation of ligand-DNA complexes owing to its rapid speed of analysis and high sensitivity (with nmol quantities now being amenable to analysis). The following sections give an overview of the previous applications of mass spectrometry in this area

to provide a background for the developments in DNA structural analysis which have taken place since the introduction of electrospray ionisation mass spectrometry.

1.3.1 *Mass Spectrometry*

Mass spectrometry (MS) involves measuring the masses of individual molecular species by converting molecules into gaseous ions and analysing their mass-to-charge ratios. Mass spectrometry is a highly useful technique for deriving both molecular weight and sequence information of DNA and hence has found extensive use as a tool for the characterisation of naturally occurring DNA constituents, synthetic oligonucleotides and modified adducts of nucleobases, nucleosides, nucleotides and oligonucleotides. In the past, the major hurdle in extending mass spectrometry to the analysis of intact nucleic acids was the difficulty in transforming these large, polar and involatile species into gaseous ions. The lability of the glycosidic and internucleotide bonds in nucleic acids limited the applications of conventional ionisation techniques which involve heating compounds for volatilisation (such as electron ionisation (EI), chemical ionisation (CI) and field ionisation (FI)), since these result in thermal decomposition of oligonucleotides. Despite the fact that the routine analysis of intact oligonucleotides was not possible until recently, mass spectrometry has found extensive use in the past for the analysis of more volatile nucleic acid constituents (bases, nucleotides and nucleosides) as well as modified adducts of these^{38,39}.

The introduction of softer ionisation techniques such as plasma desorption (PD), and fast atom bombardment (FAB-MS), opened the way for the direct analysis of thermally labile and involatile species⁴⁰. The desorption ionisation process involves the production of ions directly from solids or solution thus avoiding the need for thermal evaporation prior to ionisation. Electrospray ionisation mass spectrometry (ESI-MS) and matrix assisted laser desorption ionisation mass spectrometry (MALDI-MS) are recently developed soft ionisation techniques that have greatly extended the potential of mass spectrometry as a tool for the structural characterisation of biomolecules⁴¹ by allowing for the first time analysis of biomolecules over 150 kDa. In comparison to

previous mass spectrometric ionisation techniques, both ESI and MALDI have the ability to effectively ionise very large, involatile molecules and thus are amenable to the determination of structural and molecular weight properties of oligonucleotide and modified oligonucleotide adducts.

1.3.2 Mass Spectrometry of DNA and Ligand/DNA Complexes – Past Applications

(a) Plasma Desorption (PD)

Plasma desorption (PD-MS) provided the initial success in the sequencing and molecular weight determinations of oligonucleotides by mass spectrometry. The application of PD to the analysis of intact oligonucleotides was first reported by McNeal and co-workers⁴²⁻⁴⁵. Their results showed that positive ion PD-MS spectra of oligonucleotides largely consisted of salt adducts of the molecular ion⁴⁵ whereas corresponding negative ion spectra were characterised by the presence of the molecular ion, and ions owing to decomposition products of the oligonucleotide⁴⁴. These fragment ions were found to comprise losses of bases from the molecular ion, and ions arising from cleavages at the C-5'-O-5' and C-3'-O-3' bonds yielding 5'- and 3'-terminal sequence ions⁴⁴. The largest oligonucleotide species successfully detected by PD-MS was the sodium adduct of the dimer of a 12-mer protected oligonucleotide at m/z 12 637⁴². PD-MS analysis of fully deprotected oligonucleotides, however, was successful only up to the tetramer level owing to their greater polarity⁴².

(b) Fast Atom Bombardment (FAB)

Subsequently, greater progress in the applications of mass spectrometry to the analysis of intact oligonucleotides was made with fast atom bombardment (FAB)⁴⁶⁻⁴⁹. In FAB, the analyte is suspended in a liquid matrix (for oligonucleotides this is usually glycerol) and bombarded with a beam of high energy xenon atoms or cesium ions, which sputters gaseous molecular ions from the surface of the matrix. The first demonstration of FAB-MS for the analysis of intact oligonucleotides was accomplished by Williams *et*

*al.*⁴⁶. Following from this, groundbreaking studies in the applications of FAB-MS to the characterisation and sequence analysis of oligonucleotides were then reported by Grotjahn and co-workers⁴⁷⁻⁴⁹. These results are discussed in more detail in section 1.6 which gives an overview of the fragmentation of oligonucleotides observed with different mass spectrometry techniques. The longest oligomer that was subsequently sequenced by FAB-MS was a 10-mer and this was achieved through the analysis of series of 5'- and 3'- terminal sequence ions (bi-directional sequencing)⁴⁸. The largest oligonucleotide ionised by FAB-MS is a 13-mer⁵⁰, however, Matsuo *et al.* reported the spectrum of a blocked 24-base synthetic ribonucleotide of molecular weight 13726 Da which contained a small peak corresponding to the molecular ion³⁸.

Analysis of intact modified oligonucleotide adducts has also been demonstrated with FAB-MS and tandem mass spectrometry (MS/MS). Examples of the application of FAB-MS for the analysis of carcinogen-DNA adducts⁵¹ include the characterisation of the modified tetramer d(TpG(C8-aminobiphenyl)pCpA), and a modified pentadeoxynucleotide d(5'-pGpTp-(1-N2-(propano)-2'-dGuO). Dino *et al.* used negative ion FAB in combination with tandem mass spectrometry to identify the site of modification in the species obtained from the *in vitro* reaction of the two tetramers, d(pCpTpCpT) and d(ApGpCpT) with benzo[a]pyrene diol-epoxide which is the active form of the environmental carcinogen benzo[a]pyrene⁵¹. Martin *et al.*⁵² employed FAB-MS/MS for the analysis of cisplatin adducts of four tetranucleotides in both positive and negative ion modes. It was found that the positive ion mode predominantly yielded fragments which contained platinum such as platinum bound to two bases. In contrast, negative ion MS/MS spectra largely contained sequence and internal product ions not bound to platinum. From the fragment ions observed in the positive ion mode the cisplatin binding site could be determined. Costello *et al.*⁵³ obtained both positive and negative ion mass spectra of a series of platinum(II)/oligonucleotide adducts and found fragments predominantly arising from the loss of one or two of the amine groups from the molecular and sequence ions. FAB-MS/MS spectra of platinated oligonucleotides were

found to yield two sets of product ions which were indicative of the location of the platinum binding site corresponding to $\text{Pt}(\text{NH}_3)_{0-2}(\text{B}_x \text{ or } y)$ and $\text{Pt}(\text{NH}_3)_{0-2}(\text{B}_x \text{B}_y)$.

The major limitations associated with the use of FAB-MS for sequencing and structural characterisation of oligonucleotides and modified adducts are the relatively large amounts of sample (nmol quantities) required, and the poor upper limit of oligonucleotides amenable to FAB ionisation.

(c) *Matrix Assisted Laser Desorption Ionisation (MALDI)*

Matrix assisted laser desorption ionisation (MALDI) was developed by Karas and Hillenkamp in 1988⁵⁴ and is a variant of laser desorption whereby the sample is dissolved in a UV-absorbing matrix and ionised by the photon beam of a laser beam tuned to the absorbance of the matrix. The UV-absorbing matrix minimises radiation damage to the sample and facilitates desorption of very large polar compounds. MALDI is most commonly coupled with time-of-flight mass analysers which have the advantages of high sensitivity and mass range. Following the initial reports of the utility of MALDI for the analysis of oligonucleotides⁵⁵ and other oligothymidylic acids⁵⁶, a number of applications in the detection and sequence analysis of nucleic acids have been reported^{19,57,58}. Amongst the largest oligonucleotides successfully detected with MALDI to date include a 500-mer single-stranded DNA⁵⁹, a 622-mer double-stranded DNA⁶⁰ and a 461-mer RNA⁶¹. The sensitivity associated with MALDI MS is a major advantage in the case of oligonucleotide analysis, with quantities within the range of 1-10 pmol being normally sufficient to generate a MALDI mass spectrum, although there have been reports of the detection of femtomole quantities in favourable cases⁶².

There are two general approaches to the use of MALDI for the sequencing and structural characterisation of nucleic acids. First, the application of MALDI to the direct sequence analysis of oligonucleotides has been reported using either fragment ions observed in MALDI mass spectra, or more commonly, in combination with post source

decay (PSD)*. Nordhoff *et al.*⁶³ used PSD fragmentation in IR-MALDI-TOF-MS to confirm the sequence of a 21-mer. Using PSD fragmentation, Juhasz *et al.*⁶⁴ analysed an 11-mer DNA in which the entire set of w-type sequence ions (see section 1.6) were observed. MALDI combined with Fourier transform mass spectrometry (FT-MS) has also been shown to yield sequence and molecular weight information for oligonucleotides^{65,66}. Stemmler *et al.*⁶⁷ showed the ability of MALDI-FT-MS to yield complete sequence information for a 6-mer DNA. Applications of this method, are however impeded by limitations in the mass ranges of oligonucleotides amenable to analysis with MALDI-FT-MS in comparison to TOF mass analysers. The second approach to oligonucleotide sequencing using MALDI involves a combination with either enzymatic digestion of the sample or the analysis of the PCR products generated by Sanger sequencing methods. Sequence analysis by MALDI involving the use of exonuclease digestion was originally reported by Pieleles *et al.*⁶⁸ for the analysis of a 13-mer DNA, and more recently has been found to be applicable to the sequence analysis of oligonucleotides up to the 50-mer level^{64,69}.

The utility of MALDI-MS for the sequence analysis of modified oligonucleotides have been demonstrated in a range of studies¹⁹. Some examples of the analysis of synthetically-modified oligonucleotides include the analysis of -methylphosphonate⁷⁰, -phosphorothioate⁷¹, and, -alkylated phosphorothioate oligonucleotides⁷². Recently, Costello *et al.*⁷³ examined MALDI mass spectra of a series of diammino-platinum(II) oligodeoxyribonucleotides using a range of different conditions, including both UV and IR irradiation.

In the area of DNA sequencing, the speed of analysis of MALDI-MS combined with its sensitivity, gives this technique a substantial advantage over the use of gel electrophoresis in conjunction with Sanger and Maxam Gilbert methods. Two major

* Post source decay is a form of tandem mass spectrometry whereby an ion gate or deflector is used to mass select a precursor ion, and product ions are analysed on the basis of differing energies using a reflector mirror.

limitations associated with the MALDI analysis of oligonucleotides have hindered progress in this area. First, problems with metal cation adduction of oligonucleotides and sample impurities can significantly reduce the mass accuracy of molecular weight determinations owing to their effect on both resolution and sensitivity. Another significant limitation is that fragmentation of oligonucleotides (via loss of purine bases) can reduce molecular ion intensity and resolution. Initially the use of matrices optimised for proteins limited both the sensitivity and resolution. More recently, the use of matrices such as 3-hydroxypicolinic acid and the implementation of delayed extraction have resulted in significant improvements in mass resolution of MALDI-TOF-MS analysis of oligonucleotides⁶².

(d) Electrospray Ionisation (ESI)

In recent years, electrospray ionisation mass spectrometry (ESI-MS) has been shown to be a powerful method for generating intact ions from large, complex and fragile molecules such as proteins and oligonucleotides^{74,75}. The characteristic multiple charging of gas phase ions (more below) typically yields molecular ions for large biomolecules with m/z ratios between 500-2000 which are suitable for analysis using conventional mass analysers such as quadrupoles⁷⁶. Other advantages of ESI include excellent mass accuracy, high sensitivity (typically in the picomole-femtomole range), minimal fragmentation of the molecular species, and speed and efficiency of analysis⁷⁶. As ESI-MS was used for the studies undertaken in this work, a detailed discussion of this technique is given in the following section, while applications of ESI-MS and ESI-MS/MS of oligonucleotides is described in sections 1.5 and 1.6.

1.4 ELECTROSPRAY IONISATION MASS SPECTROMETRY

1.4.1 Origins and Historical Development

Although the widespread use of ESI-MS is a quite a recent development, the idea of electrospray had its conception almost two hundred years ago in the work of

Bose⁷⁷. This was followed by the studies of Zeleny at the beginning of this century⁷⁸. It was not until the pioneering work of Dole and colleagues towards the end of the 1960s^{79,80} that the potential of electrospray ionisation as a method for mass spectrometry began to be realised. The extensive experiments conducted by these workers into the electrospray process essentially paved the way for the breakthrough that came more than 15 years later, with the simultaneous reports by Yamashita and Fenn⁸¹, and Aleksandrov^{82,83} demonstrating for the first time the application of electrospray as an ionisation method for mass spectrometry. Their results firmly established electrospray as a revolutionary technique for the analysis of biological molecules, and, consequently, propelled electrospray ionisation into the forefront of mass spectrometric techniques for biological research.

The initial electrospray experiments conducted by Dole and co-workers⁷⁹ involved attempts at producing gas-phase macroions from dilute solutions of polystyrene. By applying a high electric field to a small flow of the analyte solution dispersed at atmospheric pressure with a bath gas of nitrogen, these workers were able to generate a fine dispersion of charged droplets. They rationalised the formation of ions on the basis of droplet fission at the Rayleigh limit (more below in section 1.4.3). Although Dole was able to demonstrate the production of gas-phase macroions, the inability of available mass analysers to detect macroions of this size posed considerable difficulties for their experiments. Their subsequent use of ion mobility and ion retardation for the detection and mass measurement of these ions produced inconsistencies and confusion with the interpretation of their results. Nevertheless, these early studies made significant inroads into the establishment of the instrumentation for electrospray ionisation, and, as such, essentially set the scene for the revolutionary developments which were to eventuate.

Interest in electrospray ionisation laid dormant for more than a decade following the studies of Dole until 1984 with the ground-breaking work of both Yamashita and Fenn⁸¹, and Aleksandrov⁸². Fenn and co-workers investigated ion formation of a range of small molecules employing a quadrupole mass analyser.

Additionally they demonstrated the capability to generate and analyse both positive and negative ions with ESI⁸⁴. They found that the major ions generated were due to the solute molecule, or its adducts with cations or anions (where the solute was nonionic), with little or no fragmentation of the molecular species^{81,84,85}. In research conducted concurrently and yet separate to that of Fenn, Aleksandrov and co-workers demonstrated the application of electrospray as an interface for liquid chromatography-mass spectrometry (LC-MS)⁸³. These workers employed a magnetic sector instrument rather than a quadrupole, and applied ESI-MS to the mass analysis of peptides up to a molecular weight of 1500 Da⁸⁶, the sequencing of peptides in conjunction with chemical digestion⁸⁷, and to the analysis of oligosaccharide-containing reaction mixtures⁸⁸.

A remarkable feature of the ESI-MS technique was later demonstrated by Fenn and colleagues⁸⁹ in their mass analysis of poly(ethylene glycols) with molecular weights of 17500 Da made possible by multiple-charging of the ions (more below). In this example the molecules carried up to 23 positive charges which effectively extended the mass range of ions that could be detected, since mass analysis is on the basis of mass-to-charge (m/z) ratio. The enormous potential this technique posed in the analysis of high molecular weight molecules was, however, more emphatically displayed by Fenn and colleagues in 1988, with their demonstration of the successful analysis of intact proteins up to 76000 Da^{90,91}. These results had an enormous impact on biological mass spectrometry^{74,75}, particularly in the analysis of proteins and oligonucleotides. Further, the combination of electrospray and tandem mass spectrometric methods (MS/MS), has found widespread applications, in particular for the sequencing of proteins and, to a lesser extent, oligonucleotides⁴⁰. Interest in the development of the electrospray mass spectrometry technique and instrumentation continues to expand, and is being driven by its potential applications to a challenging array of biological problems.

1.4.2 Principles and Instrumentation

Electrospray involves the production of gas-phase ions directly from solution by spraying a solution of electrolytes in the presence of a strong electric field to form

charged droplets which, through evaporation and desolvation in a bath gas at atmospheric pressure, generate a dispersion of gas-phase ions. A schematic diagram of an electrospray ionisation source is shown in figure 1.8.

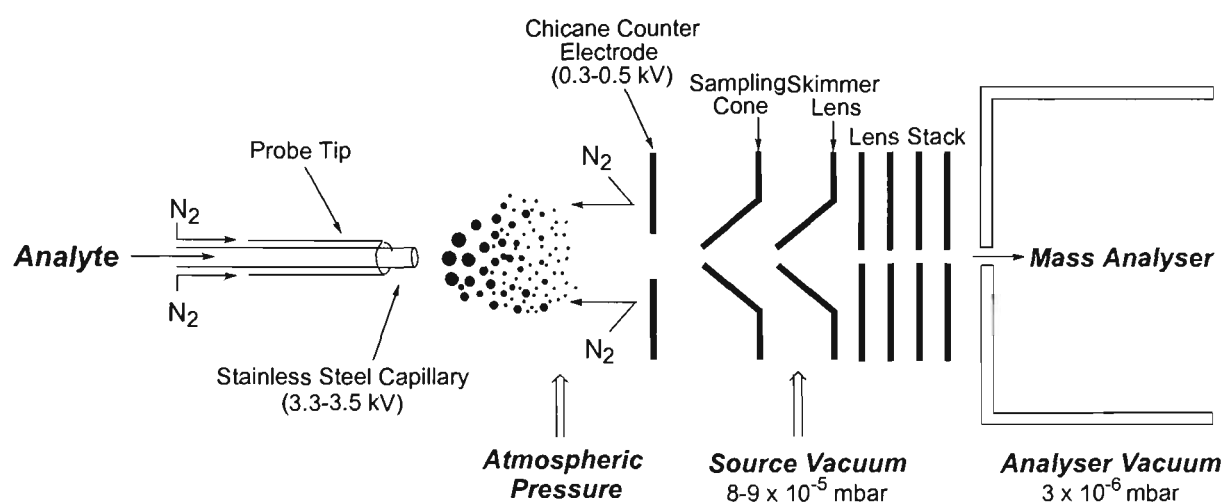


Figure 1.8: Schematic representation of an electrospray ionisation source.

A solution of the analyte is introduced through a metal (such as stainless steel), glass, or fused-silica capillary into the ionisation region which is maintained at atmospheric pressure⁸⁵. In conventional electrospray, the flow rates used typically range between 1-40 $\mu\text{L}/\text{min}$, however, recently developed low-flow electrospray techniques (microflow^[Emmet, 1994 #191] and nanoflow⁹²⁾ use substantially lower flow rates (from 5 nL/min to 1 $\mu\text{L}/\text{min}$) and thus has a much lower sample consumption. The application of a potential difference (of 3 kV to 6 kV) between the end of the capillary and the counter electrode provides a high electrostatic field at the tip of the capillary. As the liquid emerges from the capillary, the presence of the electric field leads to charge aggregation on the surface of the liquid. The instability this creates at the liquid surface causes dispersion of the liquid into a fine spray of highly charged droplets (the capillary bias determines whether positive or negative ions are formed). A flow of warm nitrogen bath gas countercurrent to the flow of sample assists in evaporating solvent from the droplets. When this charged aerosol enters into the region under vacuum, expansion of

the gas produces a supersonic jet which is then passed through a skimmer to form a beam of ions which are focused through the ion optics into the mass analyser.

A variety of different solutions have been shown to be amenable to study by electrospray, however, the commonly used solvents are aqueous solutions containing a percentage of organic solvent such as methanol, acetonitrile, or isopropanol to assist spray formation. In some cases, a small percentage of acid such as acetic or formic acid (in the positive ion mode) or base such as ammonia or triethylamine (in the negative ion mode) is added to the analyte solution in order to promote ionisation. Electrospray of aqueous solvents can also be enhanced through slight heating of the ionisation source which can alter viscosity and surface tension properties of the solution. The suitability of different solvents for the electrospray process is largely dependent on the conductivity and electrolyte concentration of the solution⁷⁵. Generally, solutions with a higher conductivity result in higher '*ESI currents*' when the ions migrate towards the tip of the capillary. This, in turn, creates a higher surface tension in the emerging liquid, such that the greater the surface tension, the greater the potential that needs to be applied at the needle tip in order to disperse the ions in a stable spray. Solutions with a high conductivity are thus typically introduced at much lower flow rates to reduce the amount of droplets formed. The use of high flow rates generally poses difficulties for producing a stable dispersion of droplets due to the larger droplet size. When higher flow rates are required, for example with on-line coupling techniques such as liquid chromatography-mass spectrometry (LC-MS), heating and gas nebulisation have been found to be useful in assisting spray formation, although this can decrease droplet charging and ionisation efficiency⁷⁵.

1.4.3 *The Process of Ionisation*

The exact process underlying the formation of gas-phase ions from charged droplets in electrospray ionisation is an issue which has endured considerable contention since Dole's initial experiments thirty years ago. Subsequently, two schools of thought have evolved regarding the mechanism of ion formation in electrospray albeit, strong

experimental evidence in favour of either mechanism is lacking. Dole and co-workers⁷⁹ proposed a mechanism of ion formation whereby charged droplets evaporate until the charge density at their surface exceeds the Rayleigh limit (i.e. the point at which the electrostatic forces between the charges at the droplet surface are greater than the surface tension)⁹³. This results in 'Coulomb explosion' of the charged droplets into a dispersion of smaller droplets which likewise undergo evaporation and explosions into smaller droplets until there are only charge-bearing solute molecules. Evaporation of the solvent from these droplets with the charge retained by the analyte molecules then generates molecular solute ions. This theory of ion formation, known as the "Charged Residue" mechanism, is schematically illustrated in figure 1.9.

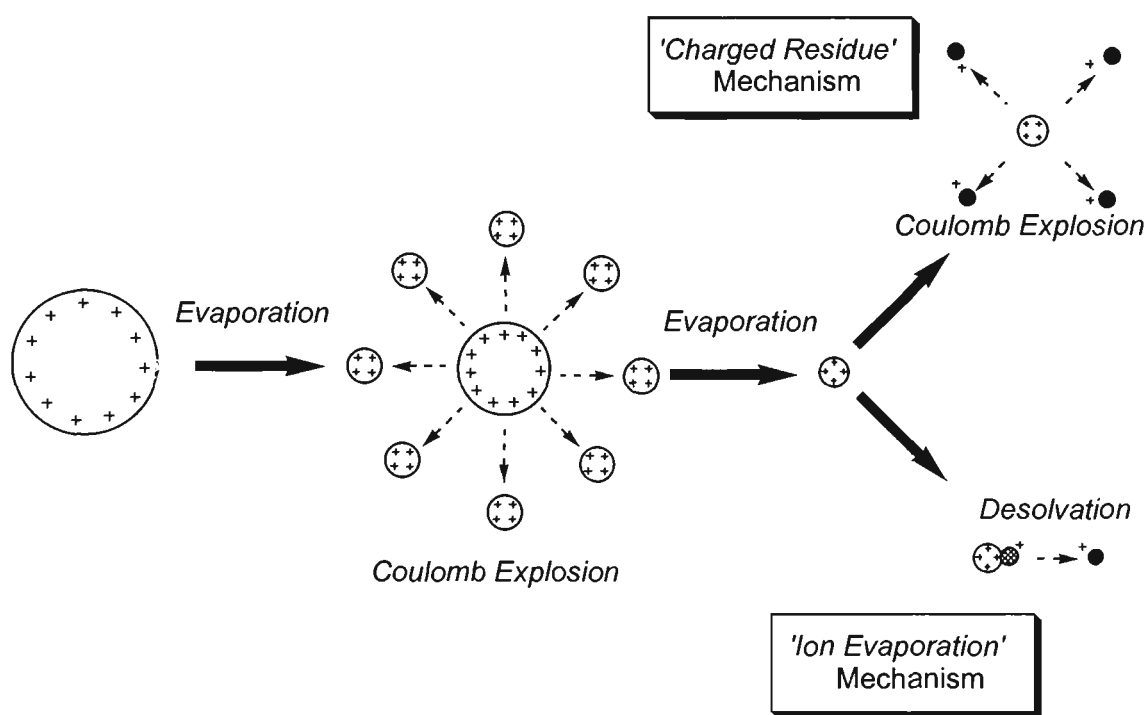


Figure 1.9: Proposed mechanisms of ion formation from charged droplets in electrospray ionisation

Some time later, Iribarne and Thomson⁹⁴, conducted studies on the process of ion formation occurring upon atomisation of charged droplets using an atmospheric pressure ion evaporation (APIE) source. Although this method differs from the electrospray process in which the charging of droplets is responsible for nebulisation

rather than atomisation of the solute, the results these researchers obtained were found to compare favourably with results obtained by Fenn *et al.*⁸¹. Mobility measurements of both positive and negative ions generated upon nebulisation of solutions of different solutes, provided experimental evidence that solute ions could evaporate or "desorb" from a charged, evaporating droplet. These workers calculated the free enthalpy, or Gibb's free energy, of solvation of the solute ions to estimate the energy, or field strength, required for a small solute ion to evaporate from a charged droplet. They then compared this to the Rayleigh limit for droplet stability and found that for droplets of a particular size (in the range of 0.01 mm), the energy required for evaporation of a solute ion could be attained without exceeding the Rayleigh instability limit. Further evidence for the ion desorption process was provided by the same group in a subsequent study which utilised mass spectrometry for the analysis of a variety of solute ions^{95,96}.

Support for the "Ion Evaporation" theory proposed by Iribarne and Thomson was offered by Fenn and co-workers⁸¹ in their initial studies into the vaporisation of small solute species by electrospray. These workers found that the resulting gas-phase ions detected by mass spectrometry corresponded well to the composition of ions present in solution, with no fragmentation being observed for solute or solvent molecules upon vaporisation. Fenn's description of ion formation in the electrospray process was that as the liquid emerges from the capillary tip, the presence of the electric field causes the liquid surface to become charged, resulting in nebulisation of the liquid into a fine aerosol of charged droplets. As these charged droplets migrate in the electric field through a countercurrent of bath gas towards the counter electrode, evaporation of the droplets causes an increase in the charge density at the droplet surface. When the Rayleigh instability limit is reached, Coulombic explosion occurs producing smaller droplets which undergo the same process of evaporation and explosion. When the size of the resulting droplets is sufficiently small, the strength of the electric field at the droplet surface becomes adequate to desorb solute ions from the droplet⁷⁴. A schematic illustration of this ion evaporation mechanism is also shown in figure 1.9. In the process of accepting this model, Fenn *et al.* showed strong misgivings concerning the "Charged

Residue" mechanism initially proposed by Dole, and rationalised that Dole arrived at some erroneous conclusions from the "mistaken" belief that they were producing singly charged solute ions in their electrospray experiments when in fact the droplets being generated consisted of an aggregation of macroions with numerous charges, resulting in the formation of multiply charged solute cluster ions⁸⁵. Another oversight in Dole's experiments was that they failed to acknowledge that heavy molecules travelling in supersonic free jets suffer from "slip effects" resulting in a retardation of their velocity.

In contrast, support for Dole's mechanism of ion formation was provided by Rollgen *et al.*⁹⁷ who asserted that the Rayleigh instability limit was almost always smaller than the strength of the electric field required at a droplet surface to produce ion desorption from the droplet. The extensive work done in later years by Nohmi and Fenn⁹⁸ on the mechanism of ion desorption using poly(ethyleneglycol) (PEG) showed some evidence that Dole's charged residue mechanism may be applicable in the case of ion formation for large oligomeric ions. Siu *et al.*⁹⁹ also investigated the formation of protein ions by electrospray, and proposed that solvent evaporation from charged droplets was not necessary for ion desorption to occur, and, moreover, the ions they observed in their mass spectra were desorbed from solutions which had experienced very little evaporation.

The theories regarding the mechanism of ion formation in electrospray are presently still divided into those which find their basis in Dole's "Charged Residue" mechanism, and those which derive from the "Ion Evaporation" model of Iribarne and Thomson¹⁰⁰. The fundamental difference between these models concerns the way in which the final droplet is generated - as a result of Coulombic explosion or from ion desorption. Although there is much stronger support for the "Ion Evaporation" mechanism, there is still no conclusive experimental evidence for either mechanism of ion formation, and thus the process by which gas phase ions are formed from liquid droplets in the electrospray process remains uncertain.

1.4.4 Characteristics of Electrospray Ionisation Mass Spectra

Perhaps the most distinctive characteristics of electrospray mass spectra of large molecules is that a series of multiply charged molecular ions is typically generated, and that little or no fragmentation occurs during ionisation. The series of multiply charged molecular ion species normally have a Gaussian-shaped distribution. It is also commonly found that the multiply charged molecular ions shift to higher average charge states with increasing molecular mass⁴⁰. The extent of charging that can be observed for a molecular species in ESI mass spectra is mainly dependent on the total available sites of protonation or deprotonation present in the molecule. Thus the structural characteristics and chemical properties of a biomolecule are an important determinant of the number of charges that it can accommodate. The ionisation environment, and the nature of the analyte solution are also significant factors. Experimental parameters of the electrospray source such as the potential applied to the needle tip, the sampling cone voltage, the flow of bath gas and the liquid flow rate have also been demonstrated to affect the observed ion distribution¹⁰¹. Other important influential factors include solution conditions such as the concentration of analyte, solvent composition, solution pH, and conductivity.

The molecular weight of a species, M_r , can be calculated from the m/z ratios of any two adjacent molecular ion peaks which differ by one charge where the charge arises from protonation in positive ions and deprotonation in negative ions^{75,102}. The relationship between m/z and M_r is given by the following simultaneous equations for negative ions (1 and 2) and positive ions (3 and 4)^{40,103}:

For negative ions :

$$\frac{m/z_a}{z_a} = \frac{M_r - z_a H}{z_a} \dots\dots\dots (1)$$

$$\frac{m/z_{a+1}}{z_{a+1}} = \frac{M_r - z_{a+1} H}{z_{a+1}} \dots\dots\dots (2)$$

For positive ions :

$$\frac{m}{z_a} = \frac{M_r + z_a H}{z_a} \dots\dots\dots (4)$$

$$\frac{m}{z_{a+1}} = \frac{M_r + z_{a+1} H}{z_{a+1}} \dots\dots\dots (5)$$

where : m/z_a , and m/z_{a+1} are the mass-to-charge ratios of two adjacent peaks
 M_r is the relative molecular mass of the molecule
 z is the number of charges taken to the nearest integer value
 H is the mass of a proton (1.007825 Da)

By solving these two simultaneous equations, the charge states of the ions at m/z_a and m/z_{a+1} can be determined, and thus the molecular mass of the molecule can then be calculated as follows:

For negative ions: $\therefore M_r = z(m/z + H) \dots\dots\dots (5)$

For positive ions: $\therefore M_r = z(m/z - H) \dots\dots\dots (6)$

A number of distinct advantages arise from this feature of multiple charging^{41,74,104}. First, a mass analyser with a large mass-to-charge range is not required since multiply charged species of increasing charge states will appear at successively lower mass-to-charge ratios. Secondly, the molecular mass can be determined more accurately since there are a series of ions from which the mass can be calculated. The presence of a series of multiply charged molecular species can, however, introduce complications in mass spectral interpretation, especially when analysing complex mixtures owing to the probability of there being overlapping peaks from different species in the spectrum. Another significant disadvantage is encountered in tandem mass spectrometry experiments when precursor ions of higher charge states are selected for collision-induced dissociation. Assigning the correct charge states of product ions can be extremely difficult if the second mass analyser does not have sufficient resolving power and consequently can lead to wrong identifications. This is an important consideration in the analysis of biological molecules such as proteins and

oligonucleotides when there are a large number of possible fragment ions or when the original species or mixture being analysed is of an unknown composition.

1.4.5 Tandem Mass Spectrometry

The gentle nature of the ESI process allows for the ionisation of large molecular species with minimal fragmentation, therefore enabling molecular weight information for these species to be readily discerned from resulting ESI mass spectra. Although the ability to determine the accurate molecular mass can provide confirmation of the elemental composition of a molecule, only minimal structural information can be derived from ESI-MS analysis. Alternatively, the application of two stages of mass analysis or tandem mass spectrometry (MS/MS), is a powerful means of characterising molecular structure. In this process, an ion of interest (the "precursor" ion) is selected in the first stage of mass analysis from the ions generated in the ESI source. The precursor ion is then allowed to undergo collision with a neutral, inert gas which transfers energy to the ion, and, in turn, causes it to fragment. This is known as either collision-induced dissociation (CID) or collision-activated dissociation (CAD) and is represented by the following equation where z represents the charge state of the ion (where m_p^z is the precursor ion, m_d^z is the product ion, and m_n is the neutral fragment formed in the dissociation reaction)¹⁰⁵:

$$m_p^z = m_d^z + m_n \quad \dots\dots\dots(7)$$

The fragment ions generated (the "product" ions) are then subjected to mass analysis with a second mass analyser to yield the product ion (or MS/MS) spectrum. Since all fragment ions appearing in the product ion spectrum originate from the selected precursor ion, no contribution from impurities is observed in the spectrum. Thus MS/MS analysis is essentially a method of separation on the basis of mass-to-charge ratios thereby obviating the need for preliminary isolation and purification of components in mixtures¹⁰⁶.

1.4.6 Electrospray Ionisation Mass Analysers

The phenomenon of multiple charging in electrospray means that mass analysers used in combination with electrospray do not require extensive m/z ranges to analyse large molecules. One of the most common electrospray instruments is the quadrupole mass filter. Quadrupole mass analysers are relatively inexpensive, have the capability to analyse ions over a relatively wide m/z range (typically up to m/z 4000), and can operate up to a large pressure region (10⁻⁵ torr), and have very good mass accuracy^{41,107}.

The successful coupling of electrospray to double-focusing sector instruments¹⁰⁸⁻¹¹⁰ has provided much higher resolution than available with quadrupole instruments, as well as enabling mass analysis over a greater mass range. Magnetic sector instruments have an advantage in the case of tandem mass spectrometry applications by enabling ions to attain higher collision energies in collisionally-induced dissociation experiments¹¹¹. This provides a greater scope in MS/MS capabilities, enabling fragmentation of a wider range of molecules, and applications to some systems which cannot achieve facile decomposition.

The combination of electrospray and Fourier transform (FT) was successfully demonstrated in groundbreaking developments undertaken by McLafferty *et al.*¹¹². This instrumentation provides for very high resolving power enabling accurate mass determinations of large macromolecules^{113,114}. The utility to conduct multiple-stage mass spectrometry experiments on stored ions (MSⁿ) and the added high resolution capacity of determining product ion charge states from resolved isotopic distributions, has been demonstrated to be a powerful combination for the sequencing and fragmentation analysis of biomolecules¹¹⁵⁻¹¹⁷. A major drawback of Fourier Transform instrumentation, however, is the extremely high vacuum required in order to conduct high resolution analyses and the cost and complexity of these instruments compared with triple quadrupoles and ion traps.

Recent developments with ion trap instrumentation coupled to electrospray have shown very promising results^{118,119}. In addition to its compact size, quadrupole ion

traps have very good sensitivity and have high efficiency and sensitivity for MS/MS analysis (which can be significantly higher than that of triple quadrupole instruments in the MS/MS mode). As with FT-MS, multiple-stage mass spectrometry can be conducted on the stored ions¹²⁰. The ion trap has, in many ways, been viewed as ideally suited for electrospray applications since the ability to store ions in the ion trap has numerous applications in experiments with gas-phase chemistry^{121,122}. The major deficiencies of the ion trap are the low dynamic mass range and limits on the mass range in MS/MS mode.

Very recently, results obtained with electrospray time-of-flight (TOF) instruments have shown this to be a potentially powerful combination for the analysis of biological macromolecules, particularly as a consequence of the essentially unlimited m/z range associated with TOF detection in linear instruments¹²³. Added to this advantage is the high resolution that can be achieved with TOF analyses, which enables facile determination of the charge state and hence molecular mass of an ion. Utilisation of orthogonal acceleration with TOF¹²⁴ is a major factor in the high sensitivity and resolution of these instruments, since the presence of an orthogonal acceleration chamber provides the means of accumulating and focusing the ions before they are accelerated into the TOF analyser. It should be noted, however, that unlike linear TOF instruments, orthogonal acceleration instruments have some limitations in mass range.

1.5 ELECTROSPRAY MASS SPECTROMETRY OF OLIGONUCLEOTIDES

Following initial studies by Fenn and co-workers in 1988, it was soon demonstrated that electrospray ionisation mass spectrometry could be applied to the analysis of a wide range of intact, high molecular weight biological molecules, particularly proteins and oligonucleotides⁷⁵. Once the capacity of ESI-MS to effectively ionise intact, polar oligonucleotides was demonstrated¹⁰², the floodgates were opened to a host of ESI-MS applications involving both the sequence analysis and structural characterisation of oligonucleotides as well as their modified adducts.

The first demonstration of the ability of ESI-MS to produce molecular ions from an intact oligonucleotide was by Covey *et al.* in 1988¹⁰². From the analysis of picomole amounts of oligonucleotides by ESI-MS using a quadrupole mass analyser, these workers were able to obtain accurate molecular weight measurements for an oligonucleotide of 14 bases in length. Since that time, rapid progress has been made in the upper molecular mass limit of oligonucleotides capable of being analysed by ESI-MS. Within the next two years, ESI with quadrupole instrumentation had been successfully used to obtain mass spectra for oligonucleotides up to a 61-mer for DNA (M_r 18701.4) , and the 76 nucleotide (nt) long transfer RNA's, tRNA^{phe} (a phenyl isoacceptor transfer RNA from brewers yeast, M_r 24 926.5) and tRNA^{f-met} (and the methionine initiator transfer RNA from *E. coli* , M_r 24 926.3)⁷⁵. The quality of these ESI mass spectra was, however, severely affected by the presence of ions resulting from extensive cation adduction to the polyphosphate backbone of molecular ion species, particularly in the case of ions with less charges. For larger oligonucleotides, which comprise a greater number of phosphate groups, the substitution of H⁺ for Na⁺ was significantly more pronounced. Consequently, the overlapping of sodium adduct peaks and multiply charged molecular ions effectively prevented the resolution of molecular ion charge states above M_r 10 000 Da⁷⁵. In addition, metal cation adduction resulted in decreased signal-to-noise ratios of molecular ion peaks due to the high contribution of multiple sodium adduct peaks to the total ion current. These factors were found to significantly reduce the accuracy of molecular weight measurements of large oligonucleotides with quadrupole mass analysis (a mass accuracy of 0.01% could only be achieved up to M_r 10000 Da), and precluded the applications of ESI-MS to analysis of oligonucleotides of higher mass⁷⁵. Consequently, for many years the molecular weight of oligonucleotides amenable to analysis by ESI-MS with quadrupole instrumentation remained significantly lower than that for proteins (which were successfully analysed up to 150kDa and higher).

Hence, much of the initial work in the analysis of oligonucleotides by ESI-MS involved efforts to overcome the problems of cation adduction by using different sample

preparation techniques, and solvent compositions described in section 1.5.1. A number of investigations have also been undertaken into minimising the extent of charging observed in the ESI mass spectra of oligonucleotide anions ("charge-state reduction" experiments described in section 1.5.2). This work has mainly been directed towards making ESI-MS more amenable to the analysis of complex DNA mixtures. There have also been an increasing number of studies into developing methods for extending the molecular mass of oligonucleotides capable of analysis by ESI-MS (described in section 1.5.3). Much of this work has been driven by recent developments in ESI-MS instrumentation such as the coupling of ESI to mass analysers of extremely high resolving power (FT-ICR), and TOF mass analysers with much higher m/z ranges than quadrupoles.

These improvements in the ESI-MS analysis of oligonucleotides, and developments in ESI instrumentation, have subsequently paved the way for a variety of different applications (described in section 1.5.4 and 1.5.5) including sequencing and the study of ligand-DNA interactions.

1.5.1 Sample Preparation and Solvent Optimisation

Efforts to remove metal cations as a source of interference in the ESI mass spectra of oligonucleotides were first reported by Stults and Marsters¹²⁵. These workers found that the replacement of sodium ions with ammonium ions through precipitation of purified oligonucleotides from a solution of 10 mM ammonium acetate, resulted in a marked improvement in the ESI mass spectra. In the mass spectra obtained for both a 30-mer and 48-mer precipitated from ammonium acetate, the most abundant peak contained no sodium ions, and in the mass spectrum obtained for a 77-mer (M_r 24 039 Da) the most intense peaks observed contained only one sodium ion attached. The absence of any adducts arising from ammonium ion attachment was rationalised on the basis that the transfer of protons from bound ammonium counter ions to their respective phosphate anions resulted in the subsequent loss of neutral gaseous ammonia upon ionisation. Potier *et al.*¹²⁶ later showed that sodium adduction could be

reduced significantly by the addition of ammonia or 1% (v/v) of the bases trimethylamine (TMA), triethylamine (TEA) and diisopropylethylamine to the oligonucleotide sample solution (the most effective base in the removal of sodium cations was shown to be TEA). Successful analysis of a 132-mer oligonucleotide acquired on quadrupole instrumentation was demonstrated using this technique. Bleicher and Bayer¹²⁷ also investigated the effects of addition of TEA to oligonucleotide solutions and found that the signal intensity obtained in the ESI-MS spectrum a 20-mer oligonucleotide increased by a factor of fifty following the addition of TEA. Limbach *et al.*¹²⁸ investigated various methods for reducing the cation adduction from monovalent (Na^+ , K^+) and divalent (Mg^{2+}) metal ions observed in ESI-MS spectra of the transfer RNA, *E. coli* tRNA^{Val1} (76 nt, M_r 24680.96 Da) employing a quadrupole. They examined the effects of precipitation with ammonium acetate, and addition of chelating agents which exhibit a high affinity for divalent metal ions. Their results showed that multiple precipitations of the *E. coli* tRNA^{Val1} with ammonium acetate produced a marked enhancement in signal intensity but did not reduce metal ion adduction sufficiently to enable accurate mass determination. Optimum results for cation removal were obtained through the addition of both EDTA and TEA to oligonucleotide solutions precipitated once from ammonium acetate.

A comparison of the efficiency of a range of organic bases of differing pK_b values (from 5.5 to 11.5) for suppressing alkali cation adduction in ESI-MS spectra has been undertaken by Grieg and Griffey¹²⁹. The most impressive results were obtained through the addition of both imidazole with either piperidine or TEA. The ESI mass spectra generated for each oligomer showed almost complete absence of alkali adducts, improved signal intensity, and a shift in ion distribution to higher charge states. Muddiman *et al.*¹³⁰ also found that the addition of both imidazole and piperidine effectively reduced sodium adduction of oligonucleotides. They suggested two possible processes to account for this. First, imidazole or piperidine can compete with the phosphate groups for adduct ions. Alternatively, the bases may compete with the adduct ions for binding to phosphate groups. In the latter case, imizadole and piperidine would

be lost as neutral gas-phase species upon ionisation, thus causing no interference in the ESI mass spectra. This study also showed that addition of acetic acid and the use of a high percentage of organic solvent (80% acetonitrile) improved sensitivity for oligonucleotide analysis.

The application of reversed-phase high performance liquid chromatography (RP-HPLC) to the purification of oligonucleotide samples has been shown to be an excellent method for removing Na^+ and K^{+131} . In the ESI-FTMS spectra acquired of a sample of tRNA^{phe} (M_r 24 950.3), HPLC desalting was shown to yield a substantial increase in the signal-to-noise ratio and thus resolving power.

1.5.2 Charge State Reduction Studies

The observation of a distribution of multiply charged molecular ion species in ESI-MS spectra of oligonucleotides has significant advantages^{74,75,85,103} including an improvement in the accuracy of mass measurement, and extending the effective mass range of analysis. With the analysis of oligonucleotides in complex mixtures, multiple charging of each component can produce a large number of ions which poses significant problems for spectral interpretation, and can affect sensitivity owing to the spreading of ion current amongst a large number of species. Therefore, various methods have been investigated for limiting the number of charge states of oligonucleotide ions present in ESI mass spectra.

In work carried out by Cheng *et al.*¹³², two different procedures were proposed to minimise the charge states of oligonucleotide ions observed in negative ion ESI mass spectra. First, the effect of addition of acid to the oligonucleotide analyte solutions was investigated with charge state reduction observed with the organic acids formic and acetic. The addition of stronger inorganic acids resulted in marked decreases in the signal intensities of oligonucleotide ions which was explained on the basis that the ionisation of conjugate base anions, and their clusters, would compete for ionisation of phosphate anions and thus decrease the ion current for the oligonucleotide. The second method tested by these workers to promote charge state reduction was the formation of

oligonucleotide diammine adducts in solution, and dissociation of these adducts via collisional activation in the ESI source. These workers found that addition of 1,4-butanediammonium diacetate to solutions of the oligonucleotide d(pT)₁₂ caused a shift in the ions observed at higher m/z (i.e. lower z). This was attributed to the diammonium ions being more effective in interacting with possible charging sites owing to their ability to form stronger complexes with oligonucleotides compared with the ammonium ions since more collision energy was required to completely dissociate the diammine complexes in the ESI interface. A recent study by Muddiman *et al.*¹³⁰ also showed that both charge state reduction and sodium adduct suppression could be achieved by varying the solution composition.

An evaluation of the effect of solution pH on the charge state distribution of unibase hexamers was undertaken by Tong *et al.*¹³³. It was found that upon shifting from pH 7.0 to pH 2.6, a decrease in the maximum charge states were observed for the sequences d(pA)₆, (6⁻ to 4⁻), d(pC)₆ (6⁻ to 3⁻), and d(pG)₆ (5⁻ to 4⁻). No charge-state reduction effect was observed for d(pT)₆. The correlation between these results and the pK_a values of the nucleobases led these workers to conclude that neutralisation of charges within the nucleobases was the factor responsible for charge-state reduction effects observed upon decreasing pH.

The ability to induce charge state reduction of oligonucleotide anions in the gas phase via reactions between ions, or ions and molecules which result in proton transfer to oligonucleotide anions, has been demonstrated with ESI-MS on quadrupole ion trap instrumentation. The utilisation of proton transfer by ion/molecule reactions between stages of mass spectrometry in order to reduce ion charge states and increase resolving power in tandem mass spectra was described by McLuckey *et al.*¹²¹. More impressive results for charge state reduction have been demonstrated with ion/ion gas phase reactions in studies by Herron *et al.*^{134,135} involving reactions between oligonucleotide ions and protonated pyridine cations. Recently, Stephenson *et al.*¹³⁶ showed that

gas-phase electron transfer reactions from multiply charged oligonucleotide anions to ionised xenon cations could effectively reduce the charge states of anions of a 16-mer oligonucleotide.

1.5.3 Molecular Mass Measurement

The coupling of ESI to mass analysers of high resolving power (FTICR-MS) and mass accuracy has seen considerable progress in the size of biomolecules for which accurate molecular mass measurement can be achieved. Earlier studies had demonstrated the capability of analysing poly(ethylene glycols) of megadalton masses on quadrupole instrumentation, however, confirmation of molecular mass could not be achieved owing to the lack of resolving power which prevented the determination of the charge on the ions⁹⁸. Similar problems were encountered by Edmonds *et al.*¹³⁷ in the analysis of plasmid DNA ($M_r=1.946$ MDa) by ESI with a quadrupole MS. Although no accurate mass measurement was possible, nonetheless these results impressively demonstrated the ability to ionise and detect molecular ions for DNA species of megadalton size.

Using the combination of ESI and FTICR, Chen *et al.*¹¹³ later reported the detection of a Coliphage T4 DNA (nominal $M_r=110$ MDa). Multiply charged anions of high charge states were trapped in the ICR cell, detected, and subjected to TOF pulse sequences whereby ions were excited and detected until an appropriate ion population comprising several individual ions were obtained. These ions were then excited and detected to optimise the peak height for an individual ion which reaches a maximum when the ion is excited at an amplitude where it is orbiting at its maximum cyclotron radius. From assuming a value of the ion radius at maximum peak height, the charge on the ion could then be determined, which in turn enabled calculation of the mass of the ion. With their results from one ion, these authors reported a mass measurement of 90.0 (± 9.1) MDa. The limiting factor in the mass accuracy of this method of charge measurement ($\pm 10\%$) is due to uncertainties (estimated to be between 5 and 10%) of measurements of the normalised ion radius. Recently, Cheng *et al.*¹¹⁴ reported an

alternative method for the molecular weight of megadalton size plasmid DNA ($M_r=1.946$ MDa). The analysis of larger sized molecules by FTICR is problematic owing to difficulties in containing larger molecules with a large number of charges in the ICR cell. These workers overcame this problem by using individual multiply charged ions for molecular weight measurements. Gas-phase ion/molecule proton transfer reactions with acetic acid in the ICR cell were used to generate successive shifts in the charge state of individual oligonucleotide ions. The m/z values of the charge state ions derived from the original selected ions were then used to determine the molecular weight of megadalton DNA with the accuracy of 0.2%.

The ionisation, detection, and mass measurement of megadalton DNA has also been demonstrated with ESI-TOFMS instrumentation by employing a method of charge detection (CD) MS. In work conducted by Fuerstenau and Benner¹²³, the masses of DNA samples with nominal molecular weights ranging from 2.8-3.6 MDa were determined by detecting and quantifying the charges of single ions. These workers used a sensitive Faraday cage charge detector to detect the charge on individual ions as they pass through the tube shaped detector. The charge, z , on the ion could be determined since it was proportional to the amplitude of the image-charge signal of the detected ion. Additionally, the m/z values of the ions were determined from measuring the velocity of the ion which was accelerated from a known electrostatic potential. From determining both the charge and m/z of the individual ions, the molecular mass could then be calculated. As with FTICR, the main source of error with mass measurements using CD-TOFMS is associated with uncertainties in measuring the charge state of the ions.

These studies demonstrate that the upper limit of ESI-MS for analysis of large oligonucleotides is limited by the capacity of the currently available mass analysers, rather than the electrospray ionisation process. The specialised nature of the experiments described above, however, limits the development of practical applications based on mass analysis of large intact DNA at the present time.

1.5.4 Sequence Analysis and Applications

A range of methods have been proposed for the sequence analysis of nucleic acids and modified adducts utilising ESI-MS alone rather than in conjunction with MS/MS as described in section 1.6.2. A common aspect of these strategies is that they employ some form of prior sample digestion in order to generate sequence specific oligonucleotides for subsequent ESI-MS analysis. Accurate mass measurement of digested oligonucleotide components has also been applied to determine the base composition of an oligonucleotide sequence¹³⁸. McCloskey and colleagues¹³⁹ have developed a technique for the location of post-transcriptional modifications in RNA which involves enzymatic digestion of nucleic acid samples into smaller oligonucleotide constituents, separation of individual components by HPLC and subsequent accurate mass measurement. In order to characterise modifications in the RNA, complete enzymatic digestion of a portion of the original purified RNA sample into nucleoside components is carried out, followed by LC/MS analysis so that modified nucleosides in the RNA sample can be identified. The accurate masses of the oligonucleotide constituents are compared with information obtained from gene sequence data, and in conjunction with the results obtained with LC/MS for modified nucleosides, the modified sites present in the RNA sample can be located. Following this report, this method was later applied to the characterisation of post-transcriptional modifications in tRNA¹⁴⁰, 5S rRNA¹⁴¹, 16S rRNA¹³⁹ and 23S rRNA¹⁴². The detection of modifications in oligonucleotides has also been carried out using nonspecific endonucleases for oligonucleotide digestion by Janning and co-workers¹⁴³. These workers applied capillary-zone electrophoresis to separate the smaller oligonucleotide fragments of calf thymus DNA prior to ESI-MS analysis for the detection of styrene-oxide/oligomer adducts.

Sequence analysis of nucleic acid samples has also been undertaken using exonucleases to form mass ladders of sequence ions. Modified sites can be identified from the mass shifts in successive mass ladders. A procedure was demonstrated by Limbach *et al.*¹⁴⁴ using exonuclease digestion to sequence a 10-mer oligonucleotide and

for the identification of a modified component in a segment of tRNA. Exonuclease digestion has also been used by Glover *et al.*¹⁴⁵ to sequence a 10-mer oligonucleotide although in this work, each component was separated and isolated by RP-HPLC, and analysed individually by ESI-MS.

ESI-MS has also been employed to analyse the DNA products amplified with polymerase chain reaction (PCR)¹⁴⁶⁻¹⁴⁸. A major difficulty with this method, however, is that contamination from buffers, primers and nucleoside triphosphates can significantly reduce sensitivity in ESI-MS analysis. More success with the analysis of PCR products has been demonstrated with MALDI by Roskey *et al.*¹⁴⁹.

1.5.5 *Noncovalent Complexes*

The formation of double-stranded (or duplex) oligonucleotides and DNA has been shown to involve associations of a noncovalent nature. The duplex and triplex forms of oligonucleotides are important for biological recognition processes and are amongst the most extensively studied noncovalent associations of biopolymers. The applications of ESI-MS to investigate noncovalent associations and interactions in biological systems was initially demonstrated in results which showed structurally specific noncovalent protein association complexes (such as enzyme-substrate, enzyme-product and receptor-ligand associations) being observed^{150,151}. The first report of the detection of duplex forms of oligonucleotides by mass spectrometry was by Light-Wahl *et al.*¹⁵² in 1993. Using conditions which favour preservation of these associations on ionisation (such as low capillary-skimmer interface voltage and low capillary temperature), several duplex ions were observed in the electrospray spectrum of a complementary 20-base oligonucleotide. The presence of salt in the sample (such as 10 mM ammonium acetate) which provides counterions to shield the phosphate anions (and hence minimise repulsions between strands) was shown to be vital for ions corresponding to duplex forms to be observed. In the same year, Ganem *et al.*¹⁵³ also reported detection of duplex forms of three octanucleotides by ESI-MS. The capability

of ESI-MS to detect noncovalent interactions was further demonstrated by Goodlett *et al.*¹⁵⁴ who reported the observation of a quadruplex form of the oligonucleotide, 5'-d(CGCG₄GCG)-3'.

The detection of noncovalent ligand-DNA interactions by ESI-MS has recently become a major focus in this area. Hsieh *et al.*¹⁵⁵ reported the detection of a noncovalent complex formed between the antitumour drug, actinomycin D, and single-stranded DNA. From the results they obtained by ESI-MS, they found that the stability of the complex was enhanced when the oligonucleotide contained neighbouring G and C bases. The successful observation of a duplex DNA/drug noncovalent complex by ESI-MS was reported by Gale *et al.*¹⁵⁶. In this study, ESI-MS was used to monitor the formation of a noncovalent complex between the self-complementary oligonucleotide, 5'-d(CCGCAAATTTGCG)-3', and the naturally occurring antitumour antibiotic, distamycin, which binds to the minor grooves in DNA. Their results showed that ESI-MS is a powerful tool for the observation and analysis of specific drug/oligonucleotide duplex noncovalent associations. Cheng *et al.*¹⁵⁷ have further demonstrated the ability to monitor DNA/protein noncovalent interactions. Gao *et al.*¹⁵⁸ used ESI-MS to characterise the specificity of the interaction between a bulged 22-mer DNA hairpin and a post-activated neocarzinostatin chromophore. In a recent report by Pocsfalvi *et al.*¹⁵⁹, ESI-MS was used to investigate the noncovalent interactions between beauvericin and oligonucleotides (6-mer and 14-mer). The first observation of an covalently-bound/duplex DNA adduct was reported by Wickham *et al.*¹⁶⁰ in 1995. In this study, ESI-MS was used to monitor the formation of an adduct between the antitumour antibiotic, hedamycin, and the double-stranded hexanucleotide, d(CACGTG)₂, upon titration of the ligand into a solution of the double-stranded oligonucleotide.

Hence, ESI-MS has been shown to have considerable promise in applications for the detection and monitoring of noncovalent associations involving oligonucleotides owing to the gentle nature of this ionisation process combined with the inherent stability of the DNA double-helical forms. An important consideration remains in studies of

noncovalent associations by ESI-MS regarding the problem of distinguishing specific and non-specific associations, and more studies are required to establish fully the utility of ESI-MS for the analysis of complexes in which the binding of the ligand to DNA is not already well characterised.

1.6 *TANDEM MASS SPECTROMETRY OF OLIGONUCLEOTIDES*

Tandem mass spectrometry (MS/MS) combined with ESI is a powerful structural tool for the characterisation of oligonucleotides owing to its sensitivity, specificity, and speed of analysis. Furthermore, MS/MS is inherently well suited to the sequence analysis of DNA owing to the wide variety of product ions which can result from fragmentation along the chain of sugar and phosphodiester units which constitute the DNA backbone. Oligonucleotide sequence elucidation by MS/MS has some potential advantages over conventional techniques such as Maxim-Gilbert and Sanger sequencing, since there is no need for prior digestion and separation of individual oligonucleotide fragments. Another important application of MS/MS is in the structural characterisation of DNA modifications. Identification of the exact mode of binding of antitumour agents to DNA is essential to effective antitumour drug design since it is an important step towards understanding how the particular binding characteristics of these DNA interactive agents may give rise to observed antitumour activity. The specificity afforded by MS/MS analyses can yield some detailed structural information on ligand binding to DNA, which is not possible using highly sensitive biological-based methods of detection such as fluorescence detection, ^{32}P postlabelling and 2-D gel electrophoresis.

The effectiveness, and hence utility of ESI-MS/MS for oligonucleotide sequencing and the structural characterisation of modified oligonucleotide adducts ultimately relies on some prior knowledge of the processes governing oligonucleotide fragmentation in MS/MS experiments. Thus, a number of the initial studies of MS/MS of oligonucleotides have focused on the investigation of oligonucleotide fragmentation

patterns, as well as factors influencing dissociation pathways in order to discern the predominant mechanisms involved upon fragmentation of oligonucleotide anions^{120-122,161-163}. Before the advent of ESI and MALDI, these studies were largely confined to dinucleotide derivatives owing to the inherent problems of transferring larger biomolecules into the gas phase intact. However, it was in work conducted with FAB-MS that it was found that certain fragmentation patterns could be discerned from the MS and MS/MS spectra of oligonucleotides^{48,49,164-167}. These earlier studies thus played an important role in planting the conceptual seed for the expansion of applications which was to follow upon the development of ESI and MALDI ionisation techniques.

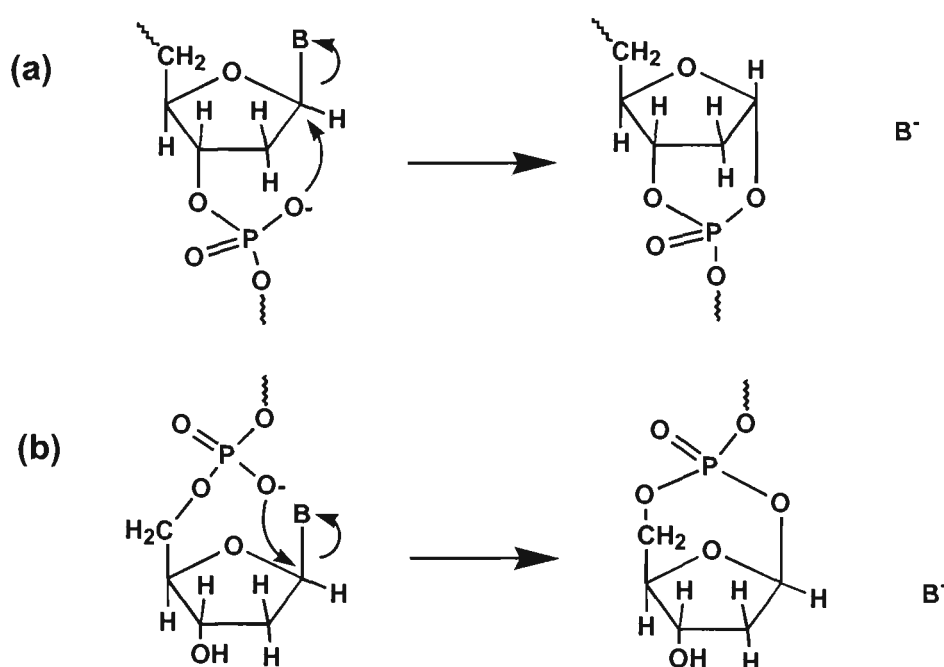
1.6.1 Fast Atom Bombardment Mass Spectrometry Studies

Much of the earlier pioneering work in mass spectrometric sequencing of oligonucleotides prior to the introduction ESI and MALDI, was undertaken by Grotjahn and colleagues in studies employing FAB-MS, and later on, FAB-MS/MS. In initial studies with negative ion FAB-MS, it was found that oligonucleotide sequence information could be obtained from the fragment ions present in resulting mass spectra⁴⁷. Two main fragment ion series were observed owing to cleavages of phosphodiester bonds between C-5' and O-5', and between C-3' and O-3'. These fragment ions were termed 5' and 3' sequence ions in order to indicate the presence of a phosphate group at the 5' and 3' end of the fragment respectively (to avoid confusion in further discussion, references made to 5' and 3' sequence ions indicate ions which contain the 5'-terminus and 3'-terminus respectively of the oligonucleotide). This fragmentation was consistent with observations made by McNeal *et al.*⁴⁴ in a previous report of oligonucleotide sequencing which employed plasma desorption (PD-MS). In a later study, Grotjahn and co-workers⁴⁸ were able to show that the sequence of a decamer could be discerned from the presence of the full set of 5'- and 3'-phosphate sequence ions in the FAB-MS spectrum. If entire sets of both sequence ions are produced, step-wise sequencing of the oligonucleotide is possible from both the 3' and 5' ends. This

capability to confirm the sequence of the oligonucleotide from both the 5' and 3' direction was termed "bidirectional sequencing". These workers analysed in the vicinity of 100 octamers and found that the 3'-phosphate sequence ions were usually of much lower intensity compared with the 5'-phosphate sequence ions. This was rationalised on the basis that the bond between C-3' and O-3' was less labile than the bond between C-5' and O-5'. Furthermore, other ions were observed in their spectra which corresponded to shifts of 18 Da down from sequence and molecular ions assigned as water loss, and shifts of 80 Da up from sequence ions which were attributed to the presence of HPO_3 groups.

In a subsequent extensive investigation into the fragmentation of dinucleotides, Cerny *et al.*¹⁶⁴ utilised FAB combined with tandem mass spectrometry (MS/MS) employing high energy collision-induced dissociation. Base loss from the 5'-terminus was shown to produce a more intense ion compared with the 3'-terminus, and a preference was found for cytosine to be lost as the neutral base. Further, it was observed that although the 3'-base anion was typically of a low intensity, it was significantly increased when a 3'-phosphate group was present. On this basis, a mechanism was postulated for loss of the base anion which involved $\text{S}_{\text{N}}2$ backside nucleophilic attack by the negatively charged phosphate oxygen on the 1'-carbon of the sugar (scheme 1.1(a)). This mechanism involves the formation of a six-membered ring phosphate transition state, which was proposed to be more stable than the seven-membered transition state which would result from cleavage of the 3'-base by $\text{S}_{\text{N}}2$ nucleophilic attack by the negatively charged phosphate oxygen on the 1'-carbon (scheme 1.1(b)). Mechanisms for cleavage of the ribose 3'-C-O bond were also proposed involving nucleophilic attack of the phosphate oxygen on either the 4'- or 2'-hydrogen generating a six-membered ring intermediate. From the intensity of B^- and $[\text{M}-\text{BH}-\text{H}]^-$ ions in the spectra it was further proposed that the order of base ion basicity was $\text{C}^- > \text{T}^-$, $\text{A}^- > \text{G}^-$. These workers later conducted FAB-MS/MS studies on small oligonucleotides (of 3 to 6 bases in length) and found that the propensity for base loss at the 5' terminus was dependent on sequence length, and decreases with longer oligonucleotides¹⁶⁵. Upon increasing oligonucleotide

length, fragment ions resulting from cleavage at the 3'-phosphate termini (5'-sequence ions) became more prominent than those from the 5'-phosphate termini (3'-sequence ions). Furthermore, larger sequences exhibited a greater number of fragment ions which were more intense than the sequence ions produced from cleavages at the 3'-phosphate termini. These competitive fragmentation pathways resulted in MS/MS spectra becoming increasingly complex, and made the elucidation of sequence information more difficult.



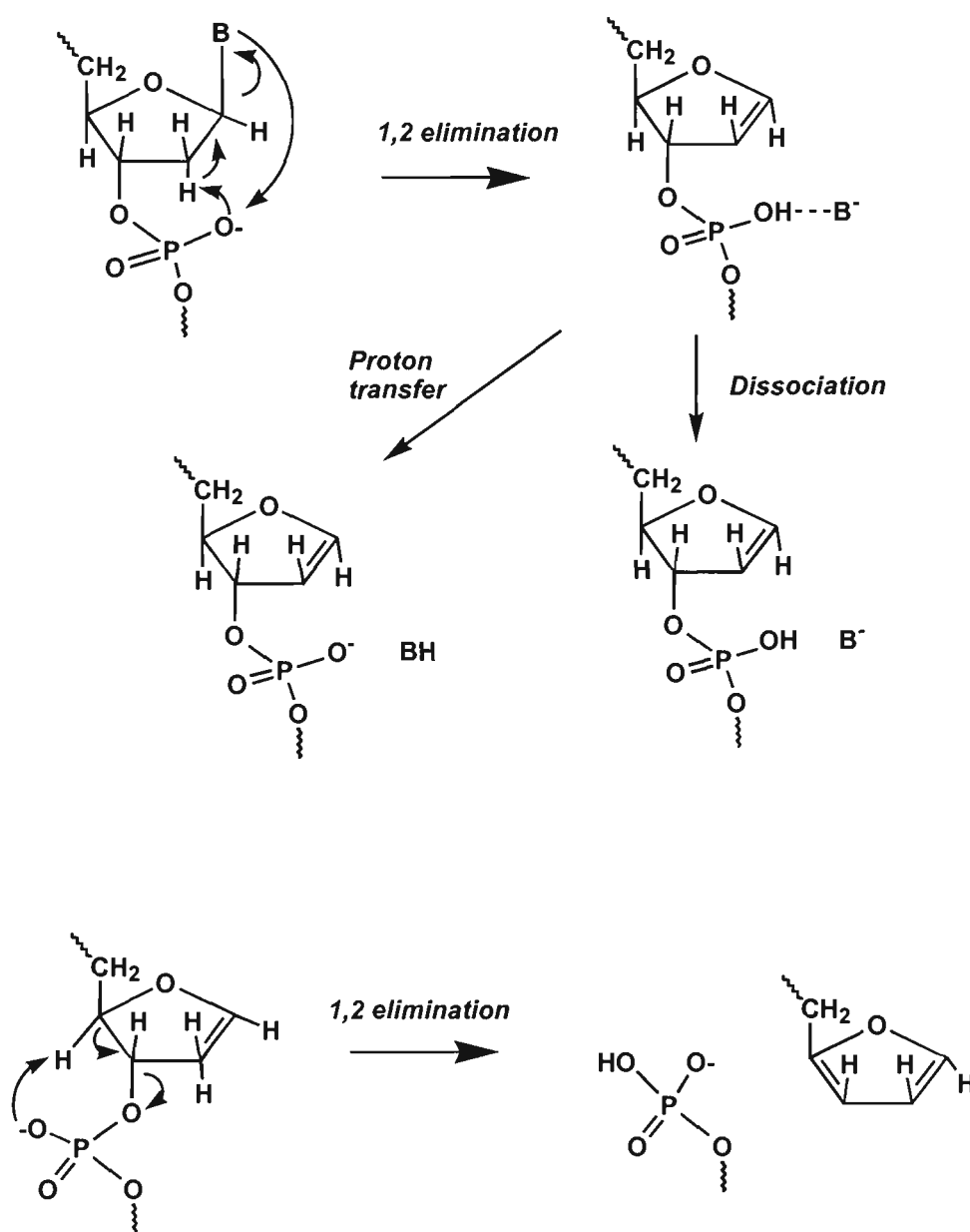
Scheme 1.1: Mechanism proposed by Cerny *et al.*¹⁶⁵ for base loss via S_N2 backside nucleophilic attack on the 1'-carbon of the sugar by a negatively charged phosphate oxygen (a) to the 3' side of the ribose, and (b) to the 5' side of the ribose from which the base is lost.

In order to facilitate understanding of the CID fragmentation pathways of larger oligonucleotides for ultimate applications in MS/MS sequencing methodologies, a significant amount of work has since been conducted on the fragmentation of smaller nucleotide units. Following the studies of Cerny *et al.* on the FAB-MS/MS of dinucleotides, Phillips and McCloskey¹⁶⁶ investigated the low energy collision-induced

dissociation of both protonated and deprotonated species of all possible heterogeneous-base sequences of 2'-deoxyribonucleotide, and ribonucleotides. These workers derived gas-phase structures of both protonated and deprotonated dinucleotide ions indicating the factors contributing to stabilisation of the tertiary structure such as intramolecular hydrogen bonding, and charge delocalisation effects. These structures were proposed as models to rationalise the observed dissociation products. In the fragmentation of protonated oligonucleotides, the tertiary dinucleotide structure involved the bases sharing the proton derived from protonation. From their results, the product ion for the 3'-base was found to be more intense than the 5'-base (except when thymine is the 3'-base). This was attributed to interaction between the phosphate and the 3'-base which induces a partial positive charge on the 3'-base, and enhances glycosidic bond cleavage. For negatively charged dinucleotide ions, the charge was proposed to be delocalised through the oxygens of the phosphate groups. Interactions between phosphate oxygens and the 3'-base hydrogens, or the 2'-hydroxyl group, stabilise the tertiary structure. In all cases, MS/MS spectra showed the 5'-terminal base was of greater intensity compared with the 3'-terminal base. These findings were in contrast to those of positive ions and also those which utilised high energy collisional activation. The rationale given for this finding was in support of the mechanism for base loss which had previously been proposed by Cerny *et al.* From their proposed structural model, the ability for interaction between deprotonated phosphate oxygens and the 5'-sugar group promotes charged base loss of the 5'-nucleotide via nucleophilic S_N2 attack of the phosphate oxygen on C-1' of the sugar.

The low energy collision-induced dissociation of all unibase and heterogeneous deprotonated dinucleotides was subsequently investigated by Rodgers *et al.* by FAB-FTICR-MS/MS¹⁶⁷. The predominant fragment ions present in resulting MS/MS spectra were found to be the 5'-terminal base anion, the fragment arising from loss of the neutral 5'-terminal base, and the phosphate anion PO₃⁻. As was found in previous studies by Cerny *et al.*¹⁶⁴ and Phillips and McCloskey¹⁶⁶, MS/MS spectra of all heterogeneous dinucleotides showed charged base loss from the 5'-terminus to yield

more intense ions than that from the 3'-terminus. These workers proposed a mechanism for base loss involving formation of a proton-bound intermediate complex (shown in scheme 1.2).



Scheme 1.2: Mechanism proposed by Rodgers *et al.*¹⁶⁷ for base loss and ribose 3'-C-O cleavage via base induced elimination resulting from proton transfer to the negatively charged phosphate.

The mechanism proceeds via a base-induced elimination process which is caused by proton transfer from the C-2' of the sugar to the negatively charged phosphate group. An intermediate complex is formed comprising the cleaved base and the protonated phosphate. Charged or neutral base loss then results depending on whether

the proton is transferred to the base. Subsequent proton transfer from the 2'-ribose hydrogen to the negatively charged phosphate results in cleavage of the 3'-phosphodiester bond and reduces Coulombic strain by generating a 3'-terminal sequence ion and a stable furan. It was suggested that the tendency for charged base loss at the 3'-terminus is strongly influenced by the strength of the stabilising interactions between the phosphate group and 3'-terminal base, since strong phosphate interactions will yield more facile base loss.

By comparing the ion abundances of the 3'-terminal base anion in sequences with the same 5'-terminal base, it was proposed that the overall order of the strength of these stabilising interactions between the 3'-terminal base and the phosphate group was $G > T > C > A$. Moreover, by comparing the intensities of the 5'-terminal base anion in sequences containing the same 3'-terminal base, these workers derived an order for the relative acidities of the nucleobases as $A > T > G > C$. These results were found to differ from the order previously reported by Cerny *et al.*¹⁶⁴ in studies employing high-energy CID.

Following this work, Rodgers and co-workers later undertook similar studies on protonated dinucleotides¹⁶⁸. Their results were found to be in good agreement to those of Phillips and McCloskey in their studies of protonated dinucleotides¹⁶⁶. Loss of the protonated 3'-terminal base was observed to be the predominant fragmentation in MS/MS spectra, except in sequences where thymine was the 3'-terminal base in which case the 5'-terminal base was observed to be the most abundant ion. Protonated base loss was proposed to occur via a multi-centered elimination mechanism in which the site of protonation on the base directs elimination of the protonated base as a stable tautomer. Upon base elimination, an intermediate hydrogen-bound complex is formed between the base and the complementary fragment. Competitive dissociation of this complex can generate either the protonated base or the dinucleotide fragment ion resulting from neutral base loss. The observed effect of thymine loss being disfavoured was rationalised first on the basis of its substantially lower proton affinity relative to the

other bases. This agreed with what was expected from estimates of the relative proton affinities of the bases made by Rodgers and co-workers which followed the order $C > G > A \gg T$. Secondly, the fact that the most stable tautomeric form of thymine is much less stable than the corresponding tautomeric forms of the other bases effectively means that thymine is comparatively a much less stable leaving group. From their results, these workers subsequently proposed an order for facile leaving group ability as $C > G > A \gg T$.

The ability to direct fragmentation pathways from charge localisation caused by the binding of a metal cation was demonstrated in a recent study by Rodgers *et al.*¹⁶⁹. This study involved an investigation of the effect of lithium ion attachment in the low energy CID fragmentation pathways observed for dinucleotides in FAB-MS/MS spectra. Lithium attachment to the phosphate group was found to preferentially direct cleavage of the ribose 3'-C-O phosphodiester bond generating a 3'-terminal sequence ion. The mechanism proposed for this pathway was similar to that previously reported by these workers in their studies on deprotonated dinucleotides (described above).

1.6.2 *Studies of Oligonucleotide Fragmentation by Electrospray Tandem Mass Spectrometry*

(a) *Negative Ions*

A great deal of what is known and understood about the collision-induced dissociation pathways of oligonucleotide anions generated by electrospray ionisation to date can be attributed to the extensive work carried out by McLuckey *et al.* using a quadrupole ion trap^{120,122,161-163}. These workers proposed the now widely accepted nomenclature for oligonucleotide fragmentation which is represented in figure 1.10 (note the variation used here whereby the oligonucleotide is shown as uncharged). The 5'-terminal ions are indicated as a, b, c, d and their complementary 3' ions as z, y, x, w respectively. Cleavages at the same position of consecutive phosphodiester units can generate a series of ions and the mass difference between these ions can be used to

identify the bases present at each position along the chain, thereby providing sequence information.

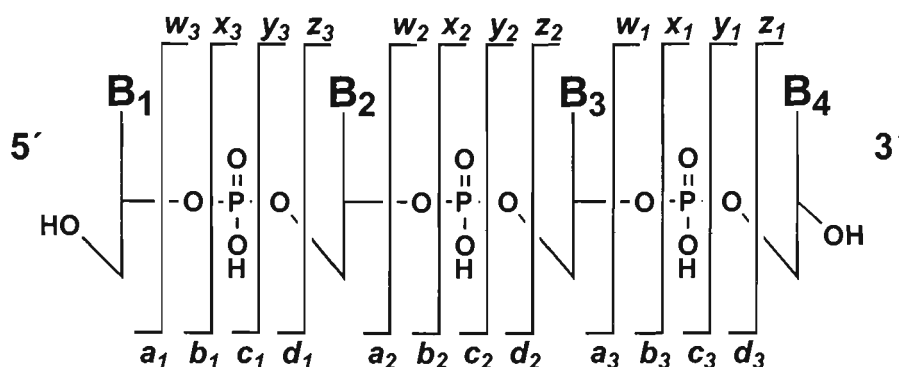


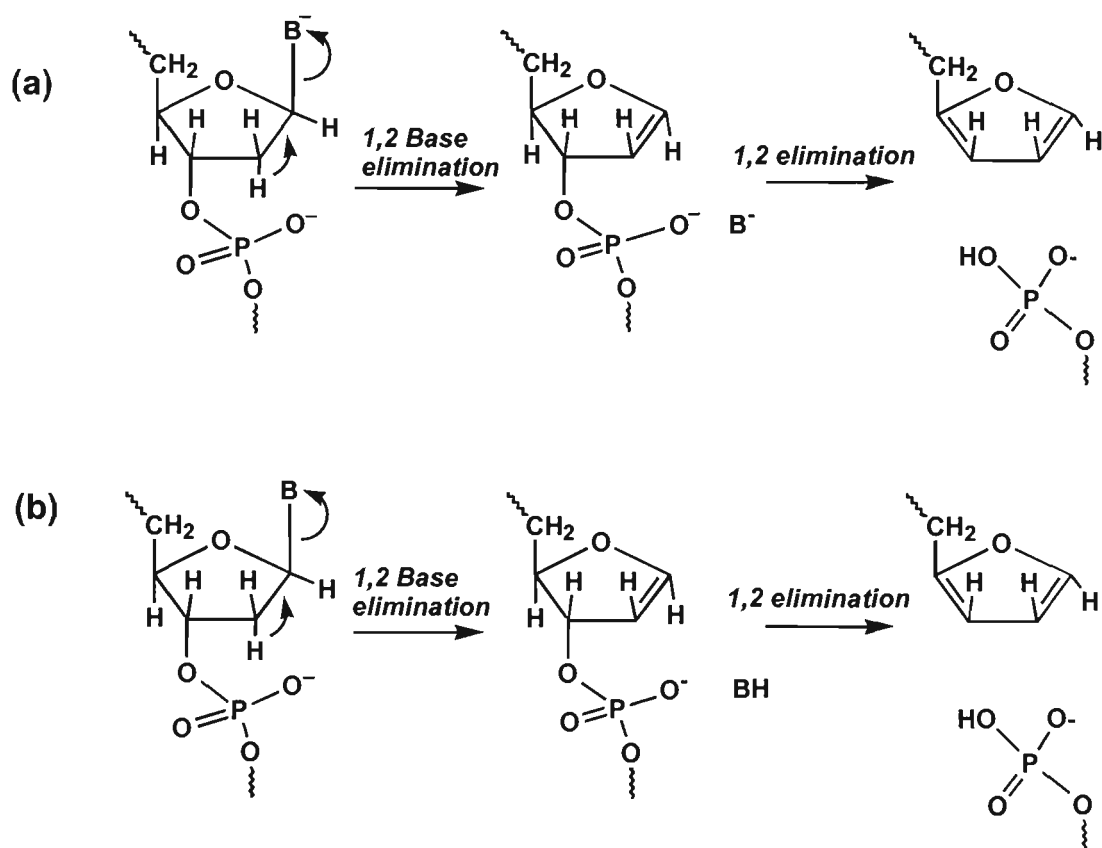
Figure 1.10: Nomenclature proposed by McLuckey et al.¹⁶¹ for oligonucleotide backbone fragmentation.

By utilising the gentle nature of collisional activation associated with an ion trap, McLuckey and co-workers were able to observe the lowest energy fragmentation pathways. From their initial results on the ESI-MS/MS of small multiply charged oligonucleotide anions^{120,161}, these workers extracted a number of fragmentation tendencies, and further postulated some rules describing oligonucleotide decomposition which can be summarised as follows:

- The primary fragmentation pathway observed is loss of a nucleobase either as a neutral or charged species. The identity of the base and the charge state of the precursor ion are the major factors influencing base loss. As the charge state on the precursor ion increases, there is a greater propensity for charged, as opposed to neutral, base loss. The order of preference for charged base loss in the MS/MS of highly charged oligonucleotide anions was $A^- > T^- > G^-$, C^- . No particular order of preference was found for losses of neutral bases in the case of precursor ions with lower charge states.
- Loss of the base at the 3' terminus was disfavoured. Apart from this, the location of the base was not found to be an influential factor determining the order of base loss preference.

- The primary fragmentation pathway of base loss is typically followed by cleavage of the C-3'-O-3' bond of the sugar from which the base was lost. This fragmentation produces a pair of complementary ions: w-type and (a-B)-type ions. Apart from these major fragmentation pathways, other cleavage processes were not found to be significant in MS/MS spectra obtained on the ion trap. Provided that the ion trapping conditions are sufficiently gentle, complementary pairs of fragment ions will not undergo further dissociation and thus can be detected. Identification of complementary ion pairs simplifies spectral interpretation by facilitating assignment of product ion charge states and hence masses.
- Precursor ions of phosphorylated oligonucleotides exhibit loss of a PO_3^- group which can compete with base loss. This fragmentation is also observed for
- w-type precursor ions (which contain a phosphate group at their 5' terminus).

In this early study, McLuckey *et al.* further presented possible fragmentation mechanisms for charged and neutral base loss which are illustrated in scheme 1.3¹²⁰.



Scheme 1.3: Mechanism proposed by McLuckey *et al.*¹²⁰ for (a) charged base loss, and (b) neutral base loss in the ESI-MS/MS fragmentation of oligonucleotide anions.

In each case, base loss was proposed to occur via 1,2 elimination of the base. The subsequent facile cleavage of the ribose 3'-C-O bond after base loss was proposed to arise from 1,2 elimination of the 3'-phosphate group since it results in the formation of a stable furan product.

The discovery of the dependence of charged base loss on precursor ion charge state, base identity and base location was more thoroughly investigated by McLuckey and co-workers in subsequent studies with ESI-quadrupole ion trap MS/MS^{122,163}. To investigate further the dependence of charged versus neutral base loss as a function of base identity, Habibi-Gourdarzi and McLuckey¹⁷⁰ later probed the fragmentation behaviour of deprotonated deoxymononucleoside and deoxydinucleoside monophosphates. The MS/MS spectra of the 2'-deoxyribonucleoside 3'-monophosphates and the 2'-deoxyribonucleoside 5'-monophosphates showed a preference for neutral over charged base loss, with loss of a base being more dependent on the phosphate group located at the 3'-terminus rather than the 5'-terminus. Their MS/MS results of deoxydinucleoside monophosphates confirmed results from previous studies showing that base loss from the 5'-terminus was more facile than loss from the 3'-terminus. Moreover, it was found that there was a preference for neutral over charged base loss at the 3'-terminus (which has also been observed in the CID of 2'-deoxyribonucleoside monophosphates). These workers estimated the relative gas-phase acidities of the nucleobases, and their relative basicities to complementary fragments, and it was found that the propensity for the 5'-terminal base to be lost as an anion versus a neutral species was governed by the order $A > G > T, C$. The preference for charged base loss from the 5'-terminus was additionally found to depend on the identity of the 3'-terminal base in the order $A > C > T > G$. In this study, base loss was assumed to occur by the process previously postulated by Rodgers *et al.*¹⁶⁷ (discussed above) which involves the formation and dissociation of an intermediate proton-bound complex.

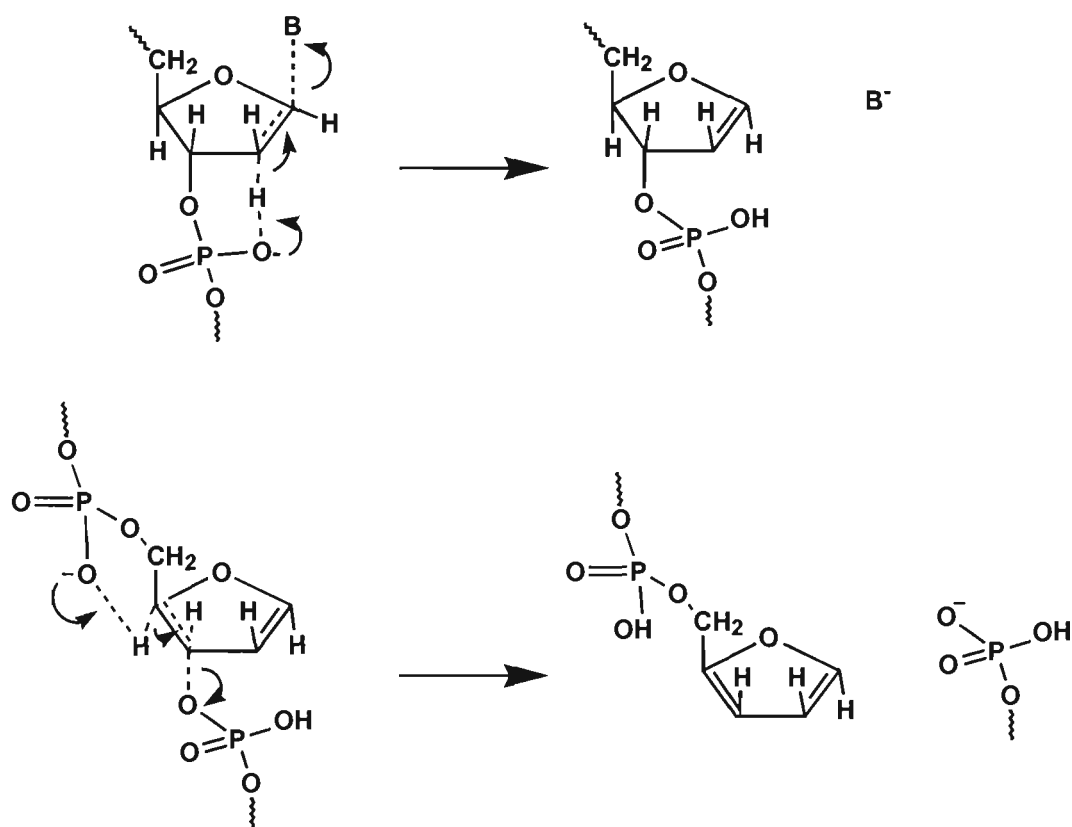
In order to probe the effects of precursor ion charge state on charged versus neutral base loss, McLuckey *et al.*¹⁶³ acquired the MS/MS spectra of a range of multiply

charged unibase adenine oligonucleotide anions (from 3 to 6 bases in length) as a function of charge state. Upon increasing the charge state of the precursor ion, the extent to which adenine was lost as an anion versus a neutral species was greatly increased. These workers concluded that the competition for loss of a charged base over a neutral base is strongly dependent on the internal Coulombic repulsion of the oligonucleotide anion, which in turn results from the charge state of the precursor ion. They rationalised the observed relationship between precursor ion charge state and charged base loss on the basis of a proposed mechanism of base loss fragmentation which involves formation of a proton-bound intermediate between the base and deprotonated phosphate group prior to dissociation which is analogous to that described by Rodgers *et al.*¹⁶⁷ and Habibi-Gourdarzi and McLuckey¹⁷⁰ in earlier work (discussed above). It was suggested that for higher charge states, the Coulomb barrier to neutral base loss increases relative to the barrier for charged base loss, thus making the latter process more facile. In a very recent report, McLuckey *et al.*¹²² investigated the MS/MS fragmentation behaviour of a range of small unibase oligonucleotide anions of varying lengths (from 4-mers to 8-mers) comprising each of the four bases A, G, C, T. These workers examined the effect on oligonucleotide dissociation of a more extensive variety of factors, including precursor ion charge state, sequence length, and base identity, and location. For all oligonucleotide anions analysed, the charge state of the precursor ion was found to be a major determining factor in observed fragmentation pathways owing to its effect on the energy thresholds of various dissociation processes. Base identity and charge site location were also demonstrated to be important factors directing certain fragmentation processes. In the case of the doubly charged tetranucleotides 5'-d(A)₄-3' and 5'-d(C)₄-3', very low ion abundances were observed for the w₂ sequence ions. This was found to confirm the locations of the negative charges on the first and third phosphate groups. The absence of a negative charge on the second phosphate to facilitate base loss and subsequent ribose 3'-C-O bond cleavage of the corresponding sugar moiety resulted in the fragmentation process of w₂ ion formation being disfavoured. These trends were not observed in MS/MS spectra of 5'-d(T)₄-3' and 5'-d(G)₄-3'. Loss of 5'-terminal bases

followed by loss of the sugar groups was found to be the major pathway in MS/MS spectra of unibase T sequences in comparison to the other unibase sequences of corresponding charge states. In contrast, a distinct lack of preference was observed for loss of 5'-terminal bases from unibase G sequences. Upon increasing length and charge state of unibase A sequences, it was found that dissociation processes involving consecutive losses of adenine (either as a charged or neutral species) dominated over phosphate cleavages generating w-type and (a-B)-type sequence ions. This was in contrast to the fragmentation behaviour of other unibase sequences and was attributed to the relatively weak glycosidic bond of adenine.

The characteristic fragmentation pathways of oligonucleotides following CID under the gentle collision conditions in an ion trap show significant differences to other MS/MS instruments. A number of MS/MS fragmentation studies of oligonucleotides have since been undertaken on triple quadrupole instruments, where both the type and extent of dissociation is markedly different to that observed by McLuckey and co-workers^{120,122,161-163}. Barry *et al.*¹⁷¹ utilised a triple quadrupole mass spectrometer to investigate the MS/MS fragmentation of a series of modified hexamers of the general sequence 5'-d(CACGXG)-3' where X was a modified pyrimidine base. As with previous ion-trap studies, base loss followed by cleavage of the ribose 3'-C-O bond was demonstrated to be a primary fragmentation pathway upon CID of oligonucleotide anions. MS/MS analysis of lower charge states such as singly and doubly charged precursor ions produced less informative spectra owing to the greater abundance of neutral fragments and their greater stability compared with higher charge state anions which consequently require higher collision energies to induce dissociation. On the other hand, highly charged precursor ions were found to fragment extensively, producing a large number of lower mass fragments. The tendency for base loss was said to be greatly influenced by the basicity of the base and the electron-withdrawing effects of its substituents, and these workers found that where the modification to the nucleobase is highly electron-withdrawing, loss of the modified base becomes a much more facile process, and effectively dominates resulting MS/MS spectra of the modified

oligonucleotide anions. These observations were rationalised in terms of an internal E2 fragmentation mechanism base catalysed by the phosphate groups which is shown in scheme 1.4. Barry *et al.* postulated that elimination of the charged base occurs from abstraction of the 2'-ribose hydrogen by the negatively charged 3'-phosphate oxygen (although this could also occur from another phosphate group or the 2'-hydroxyl). The loss of the base anion is favourable since it reduces internal strain from Coulombic repulsion forces in the oligonucleotide. Further, there is a low energy barrier to this pathway since it involves formation of a six-membered ring as the intermediate transition state. It was suggested that subsequent cleavage of the 3'-ribose-phosphate group to form a stable furan product could be base catalysed by abstraction of the 5'-ribose hydrogen from the negatively charged 5'-phosphate oxygen. This pathway proceeds favourably since it results in the formation of a stable furan product, and the relief in Coulombic strain from the elimination of the w anion.



Scheme 1.4: Mechanism proposed by Barry *et al.*¹⁷¹ for base loss and subsequent ribose 3'-C-O bond cleavage via an internal base catalysed E2 mechanism.

The differences apparent in the fragmentation behaviour of oligonucleotides in experiments conducted on quadrupole instrumentation as opposed to an ion trap were clearly demonstrated in the results of Gentil and Banoub¹⁷². In the former studies, a triple quadrupole mass spectrometer was used to study the MS/MS spectra of triply and quadruply charged anions of a series of self-complementary DNA hexamers. The extensive fragmentation observed in the MS/MS spectra of these oligonucleotides yielded confirmation of the sequences of each hexamer through the presence of sequence ions which provided the capability for complete bidirectional sequencing. The more abundant product ions observed by these workers in their MS/MS spectra in comparison to those of McLuckey *et al.*^{120,161-163} was attributed to the harsher CID conditions in the quadrupole-hexapole-quadrupole instrument compared with that of an ion trap where relatively more efficient, low energy collisions were employed.

The inherent differences in oligonucleotide dissociation generated from triple quadrupole instruments in comparison to ion trap instruments was also an issue in work undertaken by Bartlett *et al.*¹⁷³ using a triple quadrupole instrument to investigate the role played by phosphate backbone charge in influencing the fragmentation pathways of oligonucleotide anions. An important aspect of the mechanism for base loss previously proposed by McLuckey *et al.*^{120,163} was the effect of the negatively charged phosphate group in driving elimination of the adjacent 5' base. In these experiments, Bartlett *et al.* substituted uncharged methylphosphonate linkages for phosphodiester linkages in a range of oligonucleotides to observe whether the lack of a charge site on phosphate groups adjacent to cleavage sites prevented CID fragmentation pathways generating sequence ions from these positions. The MS/MS spectra obtained showed the presence of complete sets of w-type and (a-B)-type sequence ions regardless of whether the adjacent phosphate linkage could accommodate a negative charge. Thus, these results provided clear experimental evidence that these cleavage processes could occur when adjacent phosphate groups were uncharged. It was suggested that sequence ion formation from sites where adjacent phosphate groups were neutral may be quite a feasible process in light of the results obtained with MALDI demonstrating that these

product ions could be generated from low charge state precursor ions. Furthermore, these workers postulated that with the higher energies and multiple collisions associated with the triple quadrupole collision cell, formation of w-type and a-type ions could proceed without prior base loss, or the location of an adjacent charged phosphate. It was proposed that this was possible from transfer of a proton to a remote phosphodiester anion, or alternatively, via transfer of a hydrogen to the uncharged phosphate oxygen at the site of cleavage.

Another study of the effect of oligonucleotide backbone modifications on CID mass spectra acquired on a triple quadrupole was reported by Ni *et al.*¹⁷⁴. In this particular work, a comparison was made between the CID MS/MS spectra of two oligonucleotides and their corresponding structures in which the phosphodiester linkages were substituted for phosphorothioate linkages. Of interest was whether the substitution would influence the fragmentation pathways involved in the formation of sequence ions. It was found that MS/MS of the phosphorothiorate oligonucleotides yielded complete sequence information thus enabling confirmation of these structures.

The efficient dissociation of oligonucleotide anions in quadrupole collision cells has been demonstrated to have advantages for oligonucleotide sequence analysis owing to the extensive number of sequence ions generated in CID mass spectra. In work conducted by Ni *et al.*¹⁷⁵, a sequencing algorithm was formulated from oligonucleotide dissociation pathways typically observed in MS/MS spectra acquired on a triple quadrupole to aid in the oligonucleotide sequence analysis of unknown structures up to the 15-mer level. The key features involved in the sequencing process were: firstly identification of the 3'-terminal and 5'-terminal residues from which to build mass ladders; secondly attempts were then made to line up overlapping nucleotide units which had been individually constructed; and thirdly the rejection of incorrect sequences on the basis of molecular mass measurements. This method was demonstrated to have applications in locating simple modifications to base and sugar moieties, and also to the direct sequencing of components present in mixtures. In a recent study from the same group, Pomerantz *et al.*¹⁷⁶ demonstrated the potential of

ESI-MS/MS as a means of deconvoluting combinatorial libraries of oligonucleotides and thus enabling identification of individual components of oligonucleotides in complex mixtures.

With instruments of high resolving capabilities, the charge state and hence mass of fragment ions can be determined from isotopically resolved ionic species. This facilitates interpretation of product ion spectra, and effectively enables extensive structural information to be derived from ESI-MS analysis which does not involve prior precursor ion selection. A number of studies have been made utilising the high resolving power of an ESI-FTICR-MS instrument to sequence oligonucleotides, which involve generating fragmentation by either nozzle-skimmer dissociation or through infrared multiphoton dissociation (IRMPD)^{116,131,177,178}. Little *et al.*¹¹⁶ demonstrated the ability to gain complete sequence information for oligonucleotides up to a 50-mer utilising the high resolution and mass accuracy associated with ESI-FTICR-MS experiments employing nozzle-skimmer dissociation and IRMPD. Additional structural information concerning the oligonucleotides analysed was obtained from the presence of internal ions and from MS/MS of fragment ions generated in the ESI source. It was found that nozzle-skimmer dissociation yields similar fragments to those observed with CID-MS/MS, although for larger oligonucleotides, photodissociation was required to generate complete sequence information. This method was also shown to effectively provide sequence verification on 'large' oligonucleotides up to 108 bases long.

(b) Positive Ions

Very recently, several investigators have reported the CID fragmentation of protonated oligonucleotides generated by ESI. The fragmentation behaviour of oligonucleotides in the positive ion mode is of great interest since the complementary structural information provided from MS/MS analysis of oligonucleotides in both positive and negative ion modes can potentially expand oligonucleotide sequencing capabilities by MS/MS. The ability to efficiently generate positive oligonucleotide ions from ESI was demonstrated by Loo *et al.*¹⁷⁹ and in later published work, results were

presented for the positive ion ESI-MS spectra acquired of both DNA and RNA (up to the size of a 77-mer tRNA)¹⁸⁰. It was suggested that protonated oligonucleotides could be generated from gas-phase dissociation of oligonucleotide complexes formed from ammonium ion attachment to phosphate groups and bases. Owing to the higher proton affinities of the nucleobases and phosphate groups, CID of these complexes releases neutral NH₃, resulting in protonation of the bases and neutralisation of the phosphate groups. The results obtained by these workers showed that the production of higher charge state ions for larger protonated oligonucleotides was unfavourable, and, in addition, the average and maximum charge states observed in ESI-MS spectra did not appear to be influenced by the proton affinities of the bases or base identity. These characteristics of oligonucleotide charging in positive ion ESI-MS were attributed to the highly charged species undergoing decomposition in the ESI-source region owing to the higher collision energies associated with acceleration of species with higher charge using the same applied potential. This process was enhanced by the use of harsh ESI collision conditions in the present experiments which was necessary for the dissociation of oligonucleotide-nitrogen-base complexes which generated positive ions.

Preliminary investigations into the fragmentation characteristics exhibited in the CID mass spectra of protonated oligonucleotides (up to 19 bases in length) generated by ESI were undertaken by Ni *et al.*¹⁸¹. These workers first investigated the effect of solvent, specifically ammonium acetate concentration, and pH (ranging from pH 5-10), on positive ion ESI-MS spectra of oligonucleotides. These factors were found to have minimal effect on the charge state distribution of protonated oligonucleotide molecular ions. Similarly, the base composition of the oligonucleotide was not observed to play a major role in influencing charge state which was rationalised by proton sharing or phosphate protonation. Furthermore, consistent with observations by Sannes-Lowery *et al.*, the average charge states observed in positive ions was much lower than that of corresponding negative ion MS spectra which suggested that protonation was occurring at the nucleobases. The CID mass spectra of a range of doubly charged tetramers were examined and a number of important observations were made

- the primary fragmentation pathway observed for protonated oligonucleotides was similar to that observed for deprotonated oligonucleotides - charged base loss followed by cleavage of the C-3'-O-3' phosphodiester bond.
- the ion abundances for loss of protonated bases corresponded with measured proton affinities of the bases ($C > G > A \gg T$) with protonation of thymine being strongly disfavoured
- the intensities of sequence ions resulting from ribose C-3'-O-3' bond cleavage often correlated with the proton affinities of the base to the 5' side of the cleavage site.
- sequence ions were absent for C-3'-O-3' phosphodiester bond cleavage where the base attached to the ribose at the 5' side of the cleavage was a thymine.
- loss of terminal bases, in particular the 3'-terminal base, was favoured over loss of internal bases.

In a separate and almost simultaneous study, Wang and co-workers¹⁸² investigated the CID fragmentation of protonated and deprotonated oligonucleotides up to 20 bases in length. Base composition was found to influence the charge state distribution of protonated oligonucleotide molecular ions in ESI mass spectra with T-containing sequences producing lower average charge states. Consistent with the results of Ni *et al.*, the CID mass spectra acquired from both ionisation modes yielded similar fragmentation characteristics with w-type and (a-B)-type sequence ion formation in the positive ion mode enabling sequence determination. These workers also observed the absence of w-type and (a-B)-type sequence ions in the MS/MS spectrum of a protonated unibase oligonucleotide comprising the base T. The presence of relatively abundant series of x-type and z-type sequence ions, and the lack of protonated thymine was proposed to arise from location of the site of protonation being at the 5'-phosphodiester oxygen rather than on the nucleobases.

1.6.3 Characterisation of Oligonucleotide Modifications

There have been several reports on the sequencing of modified oligonucleotides by ESI-MS/MS such as those with methylphosphonate and phosphothiorate backbones, or modified nucleobases. Most of this work, however, has been directed towards gaining an understanding of oligonucleotide fragmentation behaviour and the rationalisation of mechanisms for oligonucleotide dissociation pathways. Prior to the commencement of this study, there were no reported applications of ESI-MS/MS for the structural characterisation of covalent ligand/oligonucleotide adducts. Since that time McLuckey *et al.* have used ESI on an ion trap to derive sequence information on small oligonucleotides with a modified base¹⁶². The main fragmentation pathway observed was base loss followed by cleavage of the ribose C-O bond to the 3' side of the ribose from which the base was lost. ESI combined with a triple quadrupole mass analyser was later used by Barry *et al.*¹⁷¹ to analyse the fragmentation of hexamers modified with an electron-withdrawing substituent. It was found that the electron-withdrawing substituent resulted in an enhanced loss of the nucleobase to which it was attached.

The use of electrospray mass spectrometry for the characterisation of adducts formed between DNA and carcinogenic and antitumour agents was also reported. Analysis of products formed upon reaction between 4,4'-vinylenedipyridine bis[2,2':6',2''-terpyridine platinum(II)] and the self-complementary oligonucleotide 5'-d(CGTACG)-3' was investigated by Lowe *et al.*¹⁸³ by electrospray ionisation tandem mass spectrometry. Results were consistent with attachment of the terpyridine platinum(II) on the 3'-terminal deoxyguanosine. In a preliminary report based on work presented in this thesis, Iannitti *et al.*¹⁸⁴ showed that the highly specific nature of the fragmentation generated from MS/MS of alkylated DNA adducts formed between the alkylating agent, hedamycin and the hexanucleotides 5'-d(CACGTG)-3', and 5'-d(CGTACG)-3' enabled facile, unambiguous identification of the site of alkylation. Very recently, Marzilli *et al.*¹⁸⁵ examined various adducts formed between the carcinogen, aflatoxin B₁, and DNA adducts using electrospray mass spectrometry on an

ion trap. The fragmentation pattern observed in the spectra obtained for these adducts enabled identification of the modified site in each oligonucleotide.

1.7 SUMMARY

The ability to understand the nature of the interactions between carcinogenic and antitumour agents and their cellular target, DNA, has wide reaching applications in drug discovery and the development of chemotherapeutic agents with enhanced selectivity of action. Of the many methods currently employed for the structural analysis of DNA modifications, mass spectrometry is ideally placed as a method for characterising the nature of such interactions in its ability to provide both high sensitivity and structural specificity. The development of more gentle ionisation techniques capable of analysing large biological macromolecules has impacted greatly on the role of mass spectrometry in this area, making it a rapid and sensitive technique for the analysis of biological systems.

The use of electrospray mass spectrometry in the analysis of intact oligonucleotides has seen a marked advancement since its inception in the late 1980's, with a variety of applications since reported utilising this technique for the accurate mass determination, and fragmentation analysis of oligonucleotides and modified adducts. For the ultimate goal of applying ESI-MS and ESI-MS/MS to elucidating the nature of the structural interactions between DNA and DNA-modifying agents, gaining an understanding of the gas-phase fragmentation processes governing the decomposition of unmodified oligonucleotides is an important preliminary step. From the studies which have been conducted on the fragmentation of oligonucleotide anions, an important conclusion which can be drawn is that the fragmentation behaviour of oligonucleotide anions is influenced by a variety of factors, including precursor ion charge state, base composition, and base identity. Furthermore, instrument type and experimental parameters such as ionisation mode and collision conditions have been shown to greatly impact on oligonucleotide dissociation pathways. The role played by each of these

factors is yet another variable owing to the interdependency of their exhibited effects, for example base loss propensity is influenced by factors such as instrumental collisional conditions and precursor ion charge state. Indeed, the difficulty in formulating general guidelines to the oligonucleotide fragmentation pathways upon collision-induced dissociation stems from the wide variability in both the type of factors which can influence fragmentation, and the extent of their effect.

An appreciation of the MS/MS fragmentation of unmodified oligonucleotides is important in applying MS/MS to the structural characterisation of modified oligonucleotide adducts since it provides a foundation, and a basis for comparison for understanding the fragmentation behaviour of oligonucleotides which have undergone some modification. The ability to determine detailed structural information of the site of DNA modifications provides critical information regarding the manner in which DNA interactions occur at the molecular level upon binding of antitumour agents, mutagenic and carcinogenic agents. Such information can enable insight into the relationship between the structure of these DNA-interactive agents and their biological effect, thereby constituting a crucial first stage in the ultimate goal of improved antitumour drug design.

1.8 **OUTLINE OF THIS WORK**

The work undertaken for this project aimed to utilise ESI-MS/MS on a hybrid magnetic-sector time-of-flight (mag-TOF) instrument for the characterisation of unmodified oligonucleotides, and covalent ligand-DNA adducts. Since no previous studies of the ESI-MS/MS of oligonucleotides on a hybrid mag-TOF instrument had been carried out before the commencement of this work, an initial detailed investigation was undertaken in order to discern trends evident in the fragmentation behaviour of oligonucleotides on the instrument in our laboratory, and the factors affecting observed fragmentation patterns. In the first stage of the study of oligonucleotide fragmentation described in Chapter 4, an initial detailed characterisation of product ions appearing in the ESI-MS/MS spectra of both deprotonated and protonated oligonucleotides was undertaken using multiple stage mass spectrometry experiments. The second part of this study, which is described in Chapter 5, involved an investigation of a variety of factors found to influence the fragmentation behaviour of oligonucleotide anions. In particular, the function and significance of precursor ion charge state on the degree and type of fragmentation observed in MS/MS spectra of oligonucleotide anions was considered. An overall assessment was then made of the interdependence of these various factors and the extent of the role played by each in influencing the fragmentation pattern observed in ESI-MS/MS of oligonucleotides.

The applications of ESI-MS/MS to the structural characterisation of ligand-DNA was then examined, and is described in Chapter 6. The ligands used for this study were the naturally occurring antitumour antibiotics hedamycin and its structural analogue DC92-B, and the synthetic compound phenC₆Br. The aim of this work was to investigate the sequence selective binding of these intercalating-alkylating agents with various defined sequences of DNA using tandem mass spectrometry.

REFERENCES

- (1) Dickerson, R. E. *Sci. Am.* **1983**, 249, 94-111.
- (2) Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; IRL Press at Oxford University Press: Oxford, 1990.
- (3) Stryer, L. *Biochemistry*; Third Edition ed.; W.H. Freeman and Company: New York, 1988.
- (4) Radman, M.; Wagner, R. *Sci. Am.* **1988**, 40-46.
- (5) Waring, M. J. *Annual Reviews in Biochemistry* **1981**, 50, 159-192.
- (6) Montgomery, J. A. *Agents that React with DNA*; Montgomery, J. A., Ed.; American Chemical Society: Washington, DC, 1995, pp 111-204.
- (7) Bruhn, S. L.; Toney, J. H.; Lippard, S. J. *Biological Processing of DNA Modified by Platinum Compounds*; Bruhn, S. L.; Toney, J. H.; Lippard, S. J., Ed.; Wiley: New York, 1990, pp 477-516.
- (8) Wilman, D. E. V.; Connors, T. A. *Molecular Aspects of Anti-cancer Drug Action*; Macmillan: London, 1983.
- (9) Preston, R. K.; Peck, R. M.; Breuninger, E. R.; Miller, A. J.; Creech, H. J. *J. Med. Chem.* **1964**, 7, 471.
- (10) Kohn, K. W. *Cancer Res.* **1979**, 37, 1450-1454.
- (11) Walker, R. T.; Gait, M. J. *Biosynthesis of Nucleotides*; Walker, R. T.; Gait, M. J., Ed.; IRL Press: Oxford, 1990, pp 150-155.
- (12) Blackburn, G. M. *Covalent Interactions of Nucleic Acids with Small Molecules*; Blackburn, G. M., Ed.; IRL Press: Oxford, 1990, pp 259-293.
- (13) Prakash, A. S.; Moore, A. G.; Murray, V.; Matias, C.; McFadyen, W. D.; Wickham, G. *Chemico-Biological Interactions* **1995**, 95, 17-28.
- (14) Nielsen, P. E. *Bioconjug. Chem.* **1991**, 2.
- (15) Boger, D. L. *Chemtracts-Organic Chemistry* **1991**, 4, 329-349.
- (16) Neidle, S. *DNA Structure and Recognition*; Oxford University Press: New York, 1994.
- (17) Blackburn, G. M.; Gait, M. J. *DNA and RNA Structure*; Blackburn, G. M.; Gait, M. J., Ed.; IRL Press: Oxford, 1990, pp 17-70.
- (18) Kochetkov, N. K.; Budovskij, E. I. *Organic Chemistry of Nucleic Acids: Part A*; Plenum Press: London, 1971.
- (19) Nordhoff, E.; Kirpekar, F.; Roepstorff, P. *Mass Spectrom. Rev.* **1996**, 15, 67-138.
- (20) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, 149, 761-786.
- (21) Wilson, W. D. *Reversible Interactions of Nucleic Acids with Small Molecules*; Wilson, W. D., Ed.; IRL Press: Oxford, 1990, pp 295-382.
- (22) Kuroda, R.; Neidle, S. *Drug-Nucleic Acid Interactions at the Molecular Level*; Kuroda, R.; Neidle, S., Ed.
- (23) Sequin, U. *Fortschr. Chem. Org. Naturst.* **1986**, 50, 57-122.
- (24) Schmitz, H.; Crook, K. E.; Bush, J. A. *Antimicrobial Agents and Chemotherapy* **1967**, 1966, 606-612.
- (25) Bradner, W. T.; Heinemann, B.; Gourevitch, A. *Antimicrobial Agents and Chemotherapy* **1967**, 1966, 613-618.
- (26) Joel, P. B.; Goldberg, I. H. *Biochim. Biophys. Acta* **1970**, 224, 361-370.

- (27) White, H. L.; White, J. R. *Biochemistry* **1969**, *8*, 1030-1042.
- (28) Jernigan, H. M.; Irvin, J. L.; White, J. R. *Biochemistry* **1978**, *17*, 4232-4239.
- (29) Sequin, U.; Bedford, C. T.; Chung, S. K.; Scott, I. A. *Helv. Chim. Acta* **1977**, *60*, 897-906.
- (30) Sequin, U.; Furukawa, M. *Tetrahedron* **1978**, *34*, 3623-3629.
- (31) Takahashi, I.; Takahashi, K.; Asano, K.; Kawamoto, I.; Yasuzawa, T.; Ashizawa, T.; Tomita, F.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1151-1153.
- (32) Yasuzawa, T.; Saitoh, Y.; Sano, H. *J. Antibiot.* **1990**, *18*, 485-491.
- (33) Sun, D.; Hansen, M. R.; Hurley, L. H. *J. Am. Chem. Soc.* **1995**, *117*, 2430-2440.
- (34) Murray, V.; Moore, A. G.; Matias, C.; Wickham, G. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1995**, *1261*, 195-200.
- (35) Pavlopoulos, S.; Bicknell, W.; Craik, D. J.; Wickham, G. *Biochemistry* **1996**, *35*, 9314-9324.
- (36) Wickham, G.; Prakash, A. S.; Wakelin, L. P. G.; McFadyen, W. D. *Biochim. Biophys. Acta* **1991**, *1073*, 528-537.
- (37) Murray, V.; Matias, C.; McFadyen, W. D.; Wickham, G. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1996**, *1305*, 79-86.
- (38) Crain, P. F. *Mass Spectrom. Rev.* **1990**, *9*, 505-554.
- (39) McCloskey, J. A.; Crain, P. F. *Int. J. Mass Spectrom. Ion Proc.* **1992**, *118-119*, 593-615.
- (40) Loo, J. A. *Bioconjug. Chem.* **1995**, *6*, 644-665.
- (41) Siuzdak, G. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11290-11297.
- (42) McNeal, C. J.; Macfarlane, R. D. *J. Am. Chem. Soc.* **1981**, *103*, 1609-1610.
- (43) McNeal, C. J.; Ogilvie, K. K.; Theriault, N. Y.; Nemer, M. J. *J. Am. Chem. Soc.* **1982**, *104*, 981-984.
- (44) McNeal, C. J.; Ogilvie, K. K.; Theriault, N. Y.; Nemer, M. J. *J. Am. Chem. Soc.* **1982**, *104*, 976-980.
- (45) McNeal, C. J.; Ogilvie, K. K.; Theriault, N. Y.; Nemer, M. J. *J. Am. Chem. Soc.* **1982**, *104*, 972-975.
- (46) Williams, D. H.; Bradley, C.; Bojesen, G.; Santikarn, S.; Taylor, L. C. E. *J. Am. Chem. Soc.* **1981**, *103*, 5700-5704.
- (47) Grotjahn, L.; Frank, R.; Blocker, H. *Nucleic Acids Res.* **1982**, *10*, 1471-1478.
- (48) Grotjahn, L.; Blocker, H.; Ronald, F. *Biomed. Mass Spectrom.* **1985**, *12*, 514-524.
- (49) Grotjahn, L. *Fast Atom Bombardment Mass Spectrometry of Oligonucleotides*; Grotjahn, L., Ed.; John Wiley & Sons Ltd.: London, 1986, pp 215-234.
- (50) Grotjahn, L.; Taylor, L. C. E. *Org. Mass Spectrom.* **1985**, *20*, 146-152.
- (51) Chiarelli, P.; Lay, J. O. *Mass Spectrom. Rev.* **1992**, *11*, 447-493.
- (52) Martin III, L. B.; Schreiner, A. F.; van Breemen, R. B. *Anal. Biochem.* **1991**, *193*, 6-15.
- (53) Costello, C. E.; Comess, K. M.; Plaziak, A. S.; Bancroft, D. P.; Lippard, S. J. *Int. J. Mass Spectrom. Ion Proc.* **1992**, *122*, 255-279.
- (54) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.
- (55) Spengler, B.; Pan, Y.; Cotter, R. J.; Kan, L. S. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 99-102.
- (56) Karas, M.; Bahr, U. *Trends Anal. Chem.* **1990**, *9*, 321.
- (57) Nordhoff, E. *Trends Anal. Chem.* **1996**, *15*, 240-250.

- (58) Wu, K. J.; Shaler, T. A.; Becker, C. H. *Anal. Chem.* **1994**, *66*, 1637-1645.
- (59) Tang, K.; Taranenko, N. I.; Allman, S. L.; Chang, L. Y.; Chen, C. H. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 272-730.
- (60) Bai, J.; Liu, Y. H.; Lubman, D. M.; Siemieniak, D. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 687-691.
- (61) Kirpekar, F.; Nordhoff, E.; Kristiansen, K.; Roepstorff, P.; Lezius, A.; Hahner, S.; Karas, M.; Hillenkamp, F. *Nucleic Acids Res.* **1994**, *22*, 3866-3870.
- (62) Fitzgerald, M. C.; Parr, G. R.; Smith, L. M. *Anal. Chem.* **1993**, *65*, 3204-3211.
- (63) Nordhoff, E.; Karas, M.; Cramer, R.; Hahner, S.; Hillenkamp, F.; Kirpekar, F.; Lezius, A.; Muth, J.; Meier, C.; Engels, J. W. *J. Mass Spectrom.* **1995**, *30*, 99-112.
- (64) Juhasz, P.; Roskey, M. T.; Smirnov, I. P.; Haff, L. A.; Vestal, M. L.; Martin, S. A. *Anal. Chem.* **1996**, *68*, 941-946.
- (65) Hettich, R. L.; Buchanan, M. V. *Int. J. Mass Spectrom. Ion Proc.* **1991**, *111*, 365-380.
- (66) Hettich, R. L.; Stemmler, E. A. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 321-327.
- (67) Stemmler, E. A.; Hettich, R. L.; Hurst, G. B.; Buchanan, M. V. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 828-836.
- (68) Pieleles, U.; Zurcher, W.; Schar, M.; Moser, H. E. *Nucleic Acids Res.* **1993**, *21*, 3191-3196.
- (69) Smirnov, I. P.; Roskey, M. T.; Juhasz, P.; Takach, E. J.; Martin, S. A.; Haff, L. A. *Anal. Biochem.* **1996**, *238*, 19-25.
- (70) Keough, T.; Baker, T. R.; Dobson, R. L. M.; Lacey, M. P.; Riley, T. A.; Hasselfield, J. A.; Hesselberth, P. E. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 195-200.
- (71) Wang, B. H.; Biemann, K. *Anal. Chem.* **1994**, *66*, 1918-1924.
- (72) Gut, I. G.; Beck, S. *Nucleic Acids Res.* **1995**, *23*, 1367-1373.
- (73) Costello, C. E.; Nordhoff, E.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *132*, 239-249.
- (74) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
- (75) Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Baringa, C. J.; Udseth, H. R. *Anal. Chem.* **1990**, *62*, 882-899.
- (76) Jardine, I. *Nature* **1990**, *345*, 747-748.
- (77) Bose, G. M. *Rescherches sur le cause et sur veritable theorie de l'electricite* Wittenberg, 1745.
- (78) Zeleny, J. *Physics Reviews* **1917**, *10*, 1-6.
- (79) Dole, M.; Mack, L. L.; Hines, R. L. *J. Chem. Phys.* **1968**, *49*, 2240-2249.
- (80) Mack, L. L.; Kralik, P.; Rheude, A.; Dole, M. *J. Chem. Phys.* **1970**, *52*, 4977-4986.
- (81) Yamashita, M.; Fenn, J. B. *J. Chem. Phys.* **1984**, *88*, 4451-4459.
- (82) Aleksandrov, M. L.; Gall, L. N.; Krasnov, V. N.; Nikolaev, V. I.; Pavlenko, V. A.; Shkurov, V. A. *Dokl. Akad. Nauk SSSR* **1984**, *277*, 379-383.
- (83) Aleksandrov, M. L.; Gall, L. N.; Krasnov, V. N.; Nikolaev, V. I.; Pavlenko, V. A.; Shkurov, V. A. *Bioorg. Khim.* **1984**, *10*, 710-712.
- (84) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984**, *88*, 4671-4675.
- (85) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, *9*, 37-70.
- (86) Aleksandrov, M. L.; Baram, G. I.; Gall, L. N.; Krasnov, N. V.; Kusner, Y. S.; Mirgorodskaya, O. A.; Nikolaev, V. I.; Shkurov, V. A. *Bioorg. Khim.* **1985**, *11*, 700-704.

- (87) Aleksandrov, M. L.; Baram, G. I.; Gall, L. N.; Grachev, M. A.; Knorre, V. D.; Krasnov, N. V.; Kusner, Y. S.; Mirgorodskaya, O. A.; Nikolaev, V. I.; Shkurov, V. A. *Bioorg. Khim.* **1985**, *11*, 705-708.
- (88) Aleksandrov, M. L.; Bezukladnikov, P. V.; Grachev, M. A.; Elyakova, L. A.; Zvyagintseva, T. N.; Kondrat'ev, V. M.; Kusner, Y. S.; Mirgorodskaya, O. A.; Fridlyanskii, G. V. *Bioorg. Khim.* **1986**, *12*, 1689-1692.
- (89) Wong, S. F.; Meng, C. K.; Fenn, J. B. *J. Chem. Phys.* **1988**, *92*, 546-550.
- (90) Meng, C. K.; Mann, M.; Fenn, J. B. *Z. Phys. D.: Atoms, Mol. Clusters* **1988**, *10*, 361-368.
- (91) Meng, C. K.; Mann, M.; Fenn, J. B. *Electrospray Ionisation of Some Polypeptides and Small Proteins*; Meng, C. K.; Mann, M.; Fenn, J. B., Ed.: San Francisco, California, USA, 1988.
- (92) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.
- (93) Fenn, J. B. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 524-535.
- (94) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
- (95) Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, *71*, 4451-4463.
- (96) Iribarne, J. B.; Dziedzic, P. J.; Thomson, B. A. *Int. J. Mass Spectrom. Ion Phys.* **1983**, *50*, 331.
- (97) Rollgen, F. W.; Bramer-Weger, E.; Buetfering, L. *J. Phys.* **1987**, *Colloquim C6*, 253-256.
- (98) Nohmi, T.; Fenn, J. B. *J. Am. Chem. Soc.* **1991**, *114*, 3241-3246.
- (99) Siu, K. W. M.; Guevremont, R.; Le Blanc, J. C. Y.; O'Brien, R. T.; Berman, S. S. *Org. Mass Spectrom.* **1993**, *28*, 579-584.
- (100) Kebarle, P.; Tang, K. *Anal. Chem.* **1993**, *65*, 972A-986A.
- (101) Ashton, D. S.; Beddell, C. R.; Cooper, D. J.; Green, B. N.; Oliver, R. W. A. *Org. Mass Spectrom.* **1993**, *28*, 721-728.
- (102) Covey, T. R.; Bonner, R. F.; Shushan, B. I.; Henion, J. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 249-253.
- (103) Smith, R. D.; Loo, J. A.; Ogorzalek Loo, R. R.; Busman, M.; Udseth, H. R. *Mass Spectrom. Rev.* **1991**, *10*, 359-451.
- (104) Mann, M.; Meng, C. K.; Fenn, J. B. *Anal. Chem.* **1989**, *61*, 1702-1708.
- (105) Busch, K. L.; Glish, G. L.; McLuckey, S. A. *Mass Spectrometry/Mass Spectrometry Techniques and Applications of Tandem Mass Spectrometry*; VCH Publishers Inc.: New York, 1988.
- (106) Carr, S. A.; Hemling, M. E.; Bean, M. F.; Roberts, G. D. *Anal. Chem.* **1991**, *63*, 2802-2824.
- (107) Mann, M. *Org. Mass Spectrom.* **1990**, *25*, 575-587.
- (108) Allen, M. H.; Lewis, I. A. S. *Rapid Commun. Mass Spectrom.* **1989**, *3*, 255-258.
- (109) Gallagher, R. T.; Chapman, J. R.; Mann, M. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 369-372.
- (110) Meng, C. K.; McEwen, C. N.; Larsen, B. S. *Rapid Commun. Mass Spectrom.* **1990**, *5*, 400-405.
- (111) Chapman, J. R.; Gallagher, R. T.; Barton, E. C.; Curtis, J. M.; Derrick, P. J. *Org. Mass Spectrom.* **1992**, *27*, 195-203z.
- (112) Henry, K. D.; Quinn, J. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1991**, *113*, 5447-5449.
- (113) Chen, R.; Cheng, X.; Mitchell, D. W.; Hofstadler, S. A.; Wu, Q.; Rockwood, A. L.; Sherman, M. G.; Smith, R. D. *Anal. Chem.* **1995**, *67*, 1159-1163.

- (114) Cheng, X. H.; Camp, D. G.; Wu, Q. Y.; Bakhtiar, R.; Springer, D. L.; Morris, B. J.; Bruce, J. E.; Anderson, G. A.; Edmonds, C. G.; Smith, R. D. *Nucleic Acids Res.* **1996**, *24*, 2183-2189.
- (115) Beu, S. C.; Senko, M. W.; Quinn, J. P.; Wampler, F. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 557-565.
- (116) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.
- (117) Senko, M. W.; Speir, J. P.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2801-2808.
- (118) Schwartz, J. C.; Syka, J. E.; Jardine, I. *J. Am. Soc. Mass Spectrom.* **1991**, *2*, 198-204.
- (119) Van Berkel, G. J.; Glish, G. L.; McLuckey, S. A. *Anal. Chem.* **1990**, *62*, 1284-1295.
- (120) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.
- (121) McLuckey, S. A.; Goeringer, D. E. *Anal. Chem.* **1995**, *67*, 2493-2497.
- (122) McLuckey, S. A.; Vaidyanathan, G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 1-16.
- (123) Fuerstenau, S. D.; Benner, W. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1528-1538.
- (124) Verentchikov, A. N.; Ens, W.; Standing, K. G. *Anal. Chem.* **1994**, *66*, 126-133.
- (125) Stults, J. T.; Marsters, J. C. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 359-363.
- (126) Potier, N.; Van Dorsselaer, A.; Cordier, Y.; Roch, O.; Bischoff, R. *Nucleic Acids Res.* **1994**, *22*, 3895-3903.
- (127) Bleicher, K.; Bayer, E. *Biol. Mass Spectrom.* **1994**, *23*, 320-322.
- (128) Limbach, P. A.; Crain, P. F.; McCloskey, J. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 27-39.
- (129) Greig, M.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 97-102.
- (130) Muddiman, D. C.; Cheng, X.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 697-706.
- (131) Little, D. P.; Thannhauser, T. W.; McLafferty, F. W. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2318-2322.
- (132) Cheng, X.; Gale, D. C.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1995**, *67*, 586-593.
- (133) Tong, X.; Henion, J.; Ganem, B. *J. Mass Spectrom.* **1995**, *30*, 867-871.
- (134) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 529-532.
- (135) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *Anal. Chem.* **1996**, *68*, 257-262.
- (136) Stephenson, J. L.; McLuckey, S. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 575-880.
- (137) Edmonds, C. G.; Springer, D. L.; Morris, B. J.; Thrall, B. D.; Camp II, D. G. *Electrospray Ionization of Plasmid DNA and Detection by Mass Spectrometry*; Edmonds, C. G.; Springer, D. L.; Morris, B. J.; Thrall, B. D.; Camp II, D. G., Ed.: San Francisco, California, **1993**, pp 272a-272b.
- (138) Pomerantz, S. C.; Kowalak, J. A.; McCloskey, J. A. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 204-209.
- (139) Kowalak, J. A.; Pomerantz, S. C.; Crain, P. F.; McCloskey, J. A. *Nucleic Acids Res.* **1993**, *21*, 4577-4585.
- (140) Kowalak, J. A.; Dalluge, J. J.; McCloskey, J. A.; Stetter, K. O. *Biochemistry* **1994**, *33*, 4577-4585.
- (141) Bruenger, E.; Kowalak, J. A.; Kuchino, Y.; McCloskey, J. A.; Mizushima, H.; Stetter, K. O.; Crain, P. F. **1993**, *FASEB Journal*, 196-200.
- (142) Kowalak, J. A.; Bruenger, E. B.; McCloskey, J. A. *J. Biol. Chem.* **1995**, *270*, 17758-17764.

- (143) Janning, P.; Schrader, W.; Linscheld, M. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 1035-1040.
- (144) Limbach, P. A.; McCloskey, J. A.; Crain, P. F. *Nucleic Acids Research Symposium Series* **1994**, *31*, 127-128.
- (145) Glover, R. P.; Sweetman, G. M. A.; Farmer, P. B.; Roberts, G. C. K. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 897-901.
- (146) Naito, Y.; Ishikawa, K.; Koga, Y.; Tsuneyoshi, T.; Ternuma, H.; Arakawa, R. *Rapid Commun. Mass Spectrom.* **1995**, *1995*, 9947-9959.
- (147) Wunschel, D. S.; Fox, K. F.; Fox, A.; Bruce, J. E.; Muddiman, D. C.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 29-34.
- (148) Muddiman, D. A.; Wunschel, D. S.; Liu, C.; Pasa-Tolic, L.; Fox, K. F.; Fox, A.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **1996**, *68*, 3705-3712.
- (149) Roskey, M. T.; Juhasz, P.; Smirnov, I. P.; Takach, E. J.; Martin, S. A.; Haff, L. A. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4724-4729.
- (150) Ganem, B.; Li, Y.-T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 6294-6296.
- (151) Ganem, B.; Li, Y.-T.; Henion, J. D. *J. Am. Chem. Soc.* **1991**, *113*, 7818-7819.
- (152) Light-Wahl, K. J.; Springer, D. L.; Winger, B. E.; Edmonds, C. G.; Camp, D. G.; Thrall III, B. D.; Smith, R. *J. Am. Chem. Soc.* **1993**, *115*, 803-804.
- (153) Ganem, B.; Li, Y.-T.; Henion, J. D. *Tetrahedron Lett.* **1993**, *34*, 1445-1448.
- (154) Goodlett, D. R.; Camp, D. G.; Hardin III, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.
- (155) Hsieh, Y. L.; Li, Y.-T.; Henion, J. D.; Ganem, B. *Biol. Mass Spectrom.* **1994**, *23*, 272-276.
- (156) Gale, D. C.; Goodlett, D. R.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Chem. Soc.* **1994**, *116*, 6027-6028.
- (157) Cheng, X. H.; Morin, P. E.; Harms, A. C.; Bruce, J. E.; Bendavid, Y.; Smith, R. D. *Anal. Biochem.* **1996**, *239*, 35-40.
- (158) Gao, Q.; Cheng, X.; Smith, R. D.; Yang, C. F.; Goldberg, I. H. *J. Mass Spectrom* **1996**, *31*, 31-36.
- (159) Pocsfalvi, G.; Di Landa, G.; Ferranti, P.; Ritieni, A.; Randazzo, G.; Malorni, A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 265-272.
- (160) Wickham, G.; Iannitti, P.; Boschenok, J.; Sheil, M. M. *FEBS Lett.* **1995**, *360*, 231-234.
- (161) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.
- (162) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 740-747.
- (163) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom* **1995**, *30*, 1222-1229.
- (164) Cerny, R. L.; Gross, M. L.; Grotjahn, L. *Anal. Biochem.* **1986**, *156*, 424-435.
- (165) Cerny, R. L.; Tomer, K. B.; Gross, M. L.; Grotjahn, L. *Anal. Biochem.* **1987**, *165*, 175-182.
- (166) Phillips, D. R.; McCloskey, J. A. *Int. J. Mass Spectrom. Ion Proc.* **1993**, *2*, 402-412.
- (167) Rodgers, M. T.; Campbell, S.; Marzluff, M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.
- (168) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1995**, *148*, 1-23.
- (169) Rodgers, M. T.; Campbell, S. A.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *161*, 193-216.

- (170) Habibi-Goudarzi, S.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 102-113.
- (171) Barry, J. P.; Vouros, P.; Van Schepdael, A.; Law, S.-J. *J. Mass Spectrom* **1995**, *30*, 993-1006.
- (172) Gentil, E.; Banoub, J. *J. Mass Spectrom* **1996**, *31*, 83-94.
- (173) Bartlett, M. G.; McCloskey, J. A.; Manalili, S.; Griffey, R. H. *J. Mass Spectrom* **1996**, *31*, 1277-1283.
- (174) Ni, J.; Pomerantz, S. C.; McCloskey, J. A. *Nucleic Acids Symposium Series* **1996**, *35*, 113-114.
- (175) Ni, J. S.; Pomerantz, S. C.; Rozenski, J.; Zhang, Y. H.; McCloskey, J. A. *Anal. Chem.* **1996**, *68*, 1989-1999.
- (176) Pomerantz, S. C.; McCloskey, J. A.; Tarasow, T. M.; Eaton, B. E. *J. Am. Chem. Soc.* **1997**, *119*, 3861-3867.
- (177) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.
- (178) Little, D. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 6783-6784.
- (179) Loo, J. A.; Sannes, K. A.; Hu, P.; Mei, H. Y.; Mack, D. *Studying Non Covalent Protein-RNA Interactions by Electrospray Ionization Mass Spectrometry*; Loo, J. A.; Sannes, K. A.; Hu, P.; Mei, H. Y.; Mack, D., Ed.: Portland, Oregon, USA, 1996, pp 1.
- (180) Sannes-Lowery, K. A.; Mack, D. P.; Hu, P.; Mei, H.-Y.; Loo, J. A. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 90-95.
- (181) Ni, J.; Mathews, M. A. A.; McCloskey, J. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 535-540.
- (182) Wang, P.; Bartlett, M. G.; Martin, L. B. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 846-856.
- (183) Lowe, G.; McCloskey, J. A.; Ni, J.; Vilaivan, T. *Bioorg. Med. Chem.* **1996**, *4*, 1007-1013.
- (184) Iannitti, P.; Sheil, M. M.; Wickham, G. *J. Am. Chem. Soc.* **1997**, *119*, 1490-1491.
- (185) Marzilli, L. A.; Wang, D.; Kobertz, W. R.; Essigmann, J. M.; Vouros, P. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 676-682.

Chapter 2

EXPERIMENTAL

2.1 OLIGONUCLEOTIDES

The experiments conducted for this work involved the use of 21 small oligonucleotides (4 tetramers, 14 hexamers, and 3 octamers). The sequences of these oligonucleotides and the theoretical masses calculated for the neutral monoisotopic species are listed in table 2.1. Oligonucleotide masses were calculated by adding the monoisotopic masses of each nucleotide and then adding a H to the 5'-terminus and subtracting a PO₂ group from the 3'-terminus.

Table 2.1: Oligonucleotides used in this study and their calculated neutral monoisotopic masses, *M_r*.

Oligonucleotide Sequence	<i>M_r</i>
5'-d(CCCC)-3'	1094.2297
5'-d(CCCA)-3'	1118.2410
5'-d(TCCA)-3'	1133.2406
5'-d(ACCT)-3'	1133.2406
5'-d(CCCCCC)-3'	1672.3225
5'-d(AAAAAA)-3'	1816.3894
5'-d(TTTTTT)-3'	1762.3205
5'-d(CGGCCG)-3'	1792.3409
5'-d(GCCGGC)-3'	1792.3409
5'-d(GCCATG)-3'	1791.3457
5'-d(AGGCCT)-3'	1791.3457
5'-d(CACGTG)-3'	1791.3457
5'-d(CGTACG)-3'	1791.3457
5'-d(TACGTA)-3'	1790.3504
5'-d(TCACGA)-3'	1775.3508
5'-d(TCGTGA)-3'	1806.3453
5'-d(ATCGAT)-3'	1790.3504
5'-d(ATGCAT)-3'	1790.3504
5'-d(TCGTACGA)-3'	2408.4493
5'-d(ACGTACGT)-3'	2408.4493
5'-d(GCGTACGC)-3'	2409.4446

2.1.1 Preparation and Purification

The oligonucleotides used in this work were either obtained commercially from Auspep Pty Ltd. (Victoria, Australia) or were prepared on an Applied Biosystems

Model 381A DNA synthesiser. In each case, solid phase β -cyanoethyl phosphoramidite chemistry was employed for oligonucleotide synthesis. Oligonucleotides (both purchased and synthesised) were obtained with a dimethoxytrityl (DMT) protecting group attached to the 5'-terminus (the 'trityl on' form) and still bound to the controlled pore glass bead solid support (CPG). In this form they contained β -cyanoethyl protecting groups on the phosphates, benzoyl protecting groups on adenine and cytosine bases, and isobutyryl protecting groups on guanine bases. These protecting groups are shown in figure 2.1.

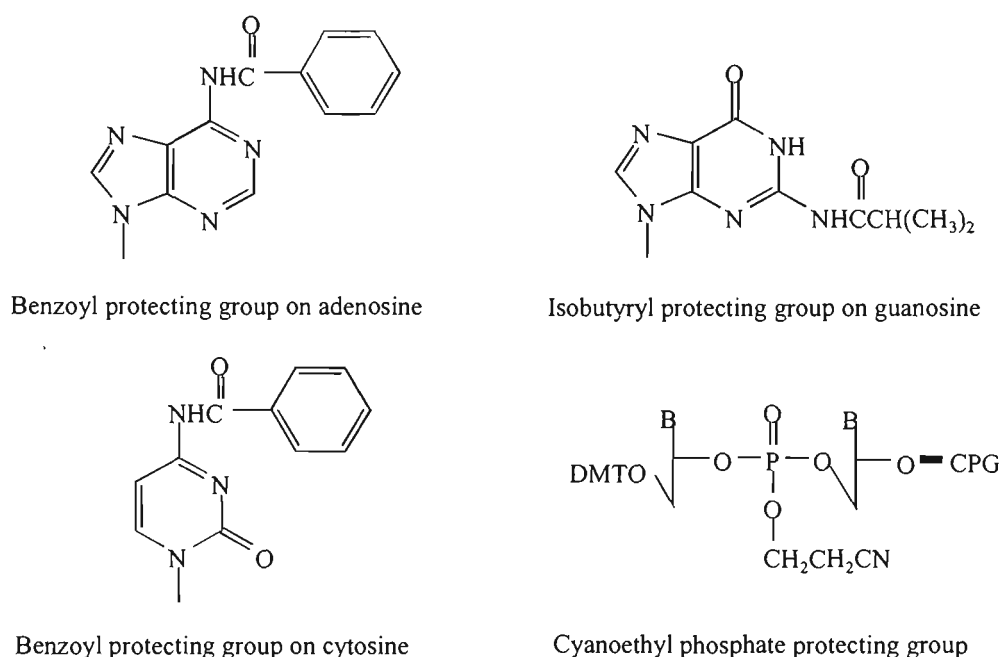


Figure 2.1: Oligonucleotide base and phosphate protecting groups.

Cleavage of oligonucleotides from the CPG solid support was achieved by washing the column with 33% ammonia for 1 hour, three times. Ammonia treatment also removed the β -cyanoethyl phosphate protecting groups. The combined ammonia washings were then heated at 55°C for 15 hours to remove the guanine isobutyryl protecting groups. After complete deprotection, the oligonucleotide solution was cooled, bubbled with nitrogen, and then freeze dried in a Savant Speedvac™ vacuum centrifuge.

Purification of the oligonucleotides was undertaken on a Waters high performance liquid chromatography (HPLC) system which consisted of a U6K injector, two model 6000A pumps, a model 680 solvent programmer and a variable wavelength

UV detector which was set at 260 nm. The separations involved reversed-phase HPLC using a C-18 octadecylsilyl column (an 8 x 100 mm Waters Delta Pak Radial Pak Cartridge). When reversed-phase HPLC is employed to purify oligonucleotides, the hydrophobicity of the DMT protecting group on the oligonucleotide enhances the separation of the desired oligonucleotide product from the capped failure sequences and synthesis impurities which are not tritylated. Hence, all oligonucleotides used in this project were purified initially with the DMT protecting groups left on the 5'-terminus of the oligonucleotide.

The entire dried oligonucleotide product obtained from the synthesis was redissolved in a minimum amount of 10 mM ammonium acetate, pH 7, and filtered through a 0.45 μm cellulose filter to remove insoluble particles prior to injection onto the HPLC. Purification of tritylated oligonucleotides required two stages of HPLC purification. The first stage involved separating the tritylated oligonucleotide from capped failure sequences and other unwanted products of the synthesis. This was achieved by using a linear gradient of 0-60% acetonitrile in 10 mM ammonium acetate, pH 7, over 30 minutes at a flow rate of 1.0 ml min^{-1} . The trityl-on oligonucleotide fraction was then collected, freeze dried and detritylated by treatment with 80% acetic acid at room temperature for 20 minutes. An equal volume of 95% ethanol was then added to precipitate the oligonucleotide out of solution, and the sample was once again freeze dried.

The detritylated oligonucleotide was then purified by HPLC using a linear gradient of 0-60% acetonitrile in 10 mM ammonium acetate over 20 minutes. This second stage of HPLC purification was carried out with the same column and system that was employed to purify the tritylated species. The fraction corresponding to the detritylated oligonucleotide was collected, freeze dried in the Speedivac, and stored at -20°C . Oligonucleotide samples used for electrospray mass spectrometry (ESI-MS) experiments were redissolved in MilliQTM water and desalted by reversed-phase HPLC with a solvent program consisting of pure water for 10 minutes followed by a linear

gradient of 0-60% aqueous acetonitrile over 5 minutes. Samples were once again freeze dried and dissolved in an appropriate solvent for analysis by ESI-MS(/MS).

2.1.2 Determination of Oligonucleotide Concentration

The concentration of the purified oligonucleotides was estimated by measuring the UV absorbance on a model 265 Shimadzu UV Visible recording spectrophotometer. The freeze dried oligonucleotide samples were redissolved in a known volume of distilled, deionised water and the absorbance of the oligonucleotide solution was measured in a 1 cm quartz cell at 260 nm. The concentration of oligonucleotide was calculated from Beer's law ($A=\epsilon lc$), using an approximate extinction coefficient (ϵ) of 9500 M⁻¹cm⁻¹ per base for single-stranded DNA, a 1 cm cell (l), and molar concentrations (c). This compares well with literature values, for example, the sequence 5'-d(AGGCCT)-3' has a reported value of 57 400 M⁻¹cm⁻¹ for the single stranded oligonucleotide¹ compared with 57 000 M⁻¹cm⁻¹ estimated by using this approximation. Oligonucleotide solutions were diluted to obtain absorbance readings (A) in the range of 0.2-1.2 so that errors owing to deviations from Beer's law were minimised.

2.2 LIGAND-OLIGONUCLEOTIDE ADDUCTS

2.2.1 Hedamycin and DC92-B

Hedamycin and DC92-B samples were generously donated by Bristol-Myers Co. (Wallingford, CT, USA) and Kyowa Hakko Kogyo Co. Ltd (Tokyo, Japan), respectively. Stock solutions of hedamycin and DC92-B were prepared by dissolution of each compound in a minimum amount of absolute ethanol and subsequent dilution with 0.1 M ammonium acetate (pH 7) to give a final ligand concentration of 0.5 mM. The solutions were flushed with nitrogen and stored at -20°C, in the dark to prevent photo-oxidation of the ligands in solution.

Hedamycin binding reactions were carried out with the hexanucleotides

5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(GCCGGC)-3', 5'-d(CGGCCG)-3', and the octanucleotides 5'-d(ACGTACGT)-3', 5'-d(TCGTACGA)-3', and 5'-d(GCGTACGC)-3'. DC92-B binding reactions were undertaken with the two hexanucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. Dried oligonucleotide samples were dissolved in 0.1 M ammonium acetate (pH 7) to give a concentration of 0.5 mM in oligonucleotide duplex. An aliquot of the ligand stock solution was added to approximately 200 μL of the 0.5 mM oligonucleotide duplex solution in a molar ratio of 1:1 ligand:oligonucleotide duplex, and then diluted to give a final concentration of 100 $\text{pmol}\mu\text{L}^{-1}$ in ligand-oligonucleotide duplex. Reactions were performed at room temperature, under nitrogen, and in the absence of light.

All ligand-oligonucleotide binding reactions were monitored by reversed-phase HPLC using the same system and column used for oligonucleotide purification. A linear gradient from buffer A (10 mM ammonium acetate, pH 7) to buffer B (60% acetonitrile in 10 mM ammonium acetate, pH 7) over 40 minutes at 1 mLmin^{-1} was employed for separation of reaction products. The final percentage of buffer B was varied (60-100%) in order to achieve optimum separation between adduct peaks. Binding reactions were allowed to proceed for 1 hour, after which the major ligand-oligonucleotide adducts formed were isolated by preparative HPLC, collected and freeze dried by vacuum centrifugation. The samples were then dissolved in water and desalted by reversed-phase HPLC as described previously. The samples were once again freeze dried and dissolved in 50% aqueous methanol for ESI-MS analysis.

2.2.2 *N*-bromohexylphenanthridinium bromide

Synthesis of the *n*-bromohexylphenanthridinium bromide (phenC_6Br) has been reported previously by Wickham *et al.*². This compound was dissolved in a minimum amount of ethanol and diluted with 0.1 M ammonium acetate to a concentration of 5 mM. Binding reactions were performed with the sequences 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' by adding an aliquot of the ligand solution to approximately 0.2 μmol of dried oligonucleotide to give a ligand:oligonucleotide duplex molar ratio of

5:1, and a final oligonucleotide duplex concentration of 1 mM. Reaction mixtures were then incubated at 37°C for 18 hours.

PhenC₆Br binding reactions were monitored by reversed-phase HPLC (using equipment described above) employing a linear gradient from buffer A (10 mM ammonium acetate, pH 7) to 60% buffer B (60% acetonitrile in 10 mM ammonium acetate, pH 7) over 40 minutes at 1 mLmin⁻¹. After the binding reactions had proceeded for 18 hours, the major adducts formed judged by the initial HPLC analyses of the mixtures were isolated by preparative HPLC. These samples were freeze dried by vacuum centrifugation and desalted (using the method described above), freeze dried again and then dissolved in 50% aqueous methanol for ESI-MS(/MS) analysis.

2.3 ELECTROSPRAY MASS SPECTROMETRY

The majority of tandem mass spectrometry experiments undertaken in this work were performed on a Micromass/VG AutospecTM oa-TOF magnetic sector mass spectrometer (Micromass, Wythenshawe, England) equipped with an electrospray ionisation source. A Micromass VG Biotech QuattroTM was also employed for a smaller number of ESI-MS and ESI-MS/MS experiments. The QuattroTM was routinely used for the rapid ESI-MS analysis of HPLC-purified oligonucleotide samples to check sample purity, and for analysis of the products of ligand binding reactions isolated by HPLC in order to identify peaks in the HPLC chromatograms giving rise to ligand-oligonucleotide adducts.

2.3.1 AutospecTM oa-TOF Magnetic Sector Mass Spectrometer

(a) Instrumentation

A schematic illustration of the VG AutospecTM is shown in figure 2.2. This sector hybrid instrument comprises three sectors as MS 1 of *E-B-E* configuration (an electrostatic analyser, ESA 1 (*E*); the magnet (*B*); and a second electrostatic analyser, ESA 2 (*E*)) with a *m/z* limit of 6500, and an orthogonal acceleration time-of-flight

analyser (oa-TOF) as MS 2. The electrospray interface on this instrument comprises a three-stage pumping arrangement in order to transfer ions formed at the tip of the spray probe into the high vacuum region for mass analysis. The intermediate region between the electrospray atmospheric pressure chamber and the high vacuum region comprises two differentially pumped regions. A schematic of this electrospray source is shown in figure 2.3.

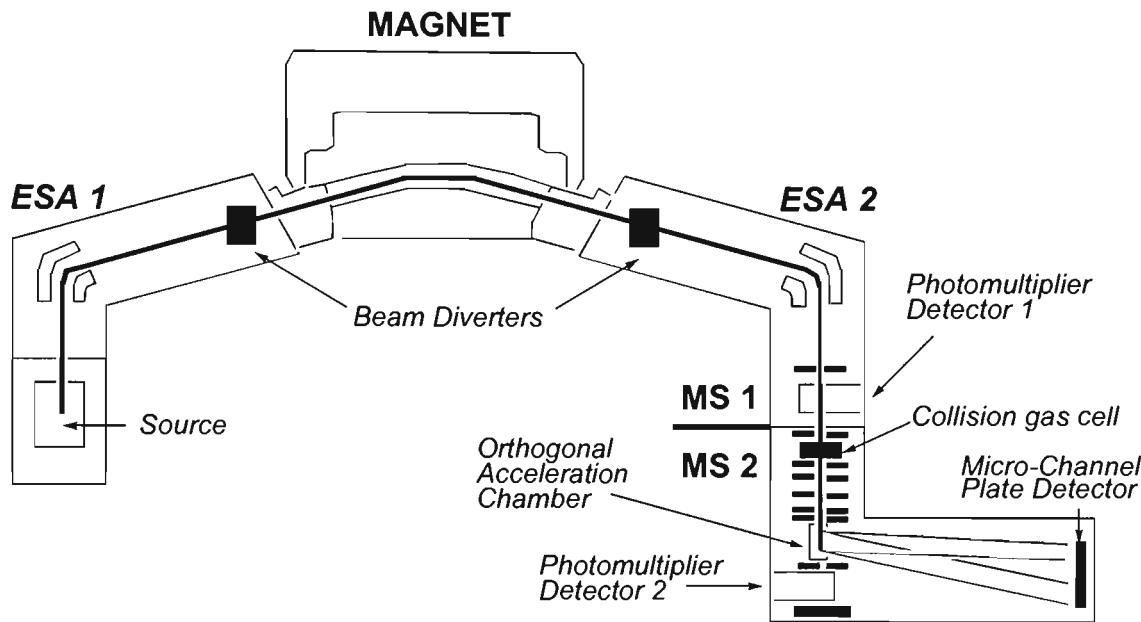


Figure 2.2: Schematic of the VG Autospec™ magnetic sector mass spectrometer.

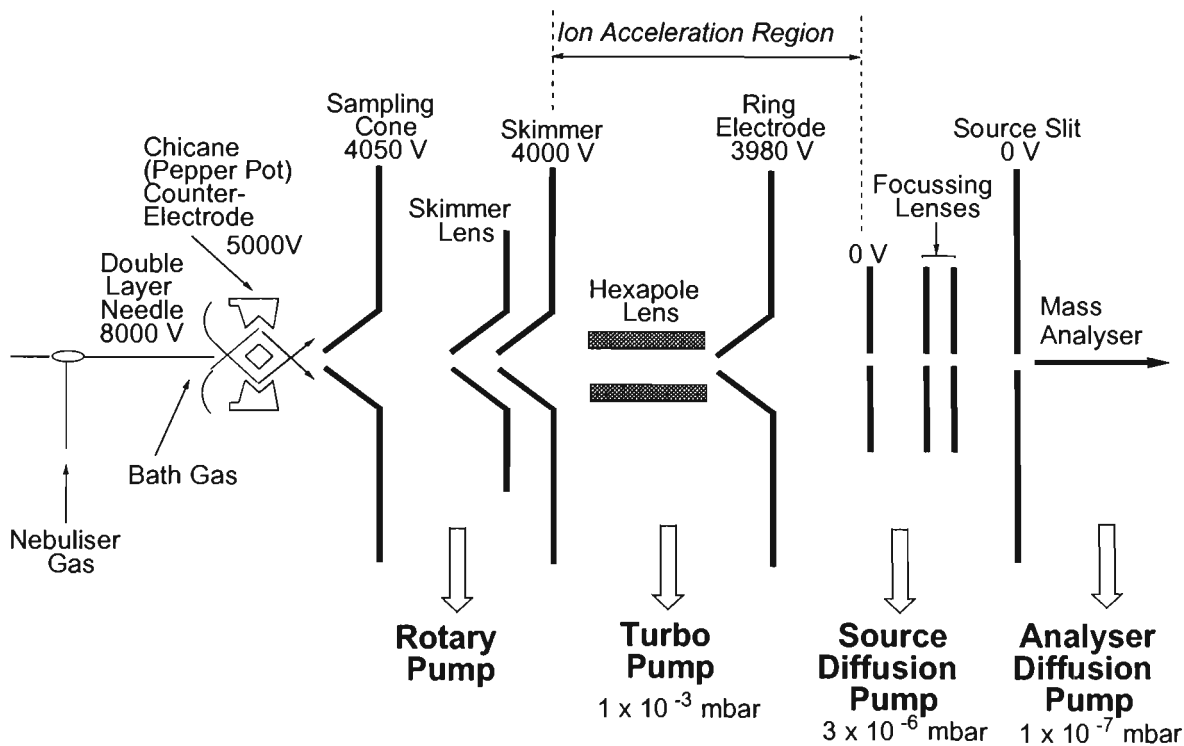


Figure 2.3: Schematic representation of the electrospray ionisation source on the VG Autospec™ mass spectrometer.

Analyte solutions were delivered into the electrospray source through a stainless steel needle at flow rates between 5-20 μLmin^{-1} . A Harvard Apparatus (South Natick, MA, USA) Model 22 syringe pump was employed for solvent delivery and continuous infusion of analyte solutions. A flow of nitrogen nebulising gas concurrent to the stainless steel capillary at a flow rate of 8-12 Lhour^{-1} (7 bar) was used to assist formation of a stable spray. Evaporation of the droplets formed was promoted by a dry, warm, counter-current flow of nitrogen bath gas at a flow rate of $\sim 250 \text{ Lhour}^{-1}$ (7 bar). Ions passing through the sampling cone enter into the first of the two differentially pumped intermediate regions between the probe tip at atmospheric pressure and the high vacuum region after the ring electrode. A hexapole lens in this region served to remove neutral species and optimise ion transmission. Mass spectra were acquired in both the positive and negative ion modes. For negative ions, the electrospray probe tip potential was -8 kV with -5 kV on the counter electrode. The potentials applied to the sampling cone and skimmer lens were typically about 50 V and 60 V respectively below the accelerating voltage, $V_{\text{acc.}}$ of -4 kV. For positive ions, the electrospray probe tip potential was 8 kV with 5 kV on the counter electrode. The potentials applied to the sampling cone and skimmer lens were typically about 50 V and 60 V respectively above the accelerating voltage, $V_{\text{acc.}}$ of 4 kV. Other parameters are shown in table 2.2.

Table 2.2: Typical operating parameters used for the electrospray ionisation source on the Autospec TM.

Source Parameters	+ve ions	-ve ions	Comments
Needle Voltage	$\sim 8 \text{ kV}$	$\sim -8 \text{ kV}$	(variable)
Counter Electrode	$\sim 1050 \text{ kV}$	$\sim -1050 \text{ kV}$	
Sampling Cone	$\sim 50 \text{ V} > V_{\text{acc.}}$	$\sim 50 \text{ V} < V_{\text{acc.}}$	(sample dependent)
Skimmer Lens	$\sim 60 \text{ V} > V_{\text{acc.}}$	$\sim 60 \text{ V} < V_{\text{acc.}}$	(sample dependent)
Skimmer ($V_{\text{acc.}}$)	4 kV	-4 kV	
Hexapole	$\sim 10 \text{ V} < V_{\text{acc.}}$	$\sim 10 \text{ V} > V_{\text{acc.}}$	
Ring Electrode	$\sim 20 \text{ V} < V_{\text{acc.}}$	$\sim 20 \text{ V} > V_{\text{acc.}}$	
Bath Heater	60-80°C	60-80°C	

ESI mass spectra were acquired in multi-channel analysis (MCA) mode scanning at a rate of 1 second per 10 m/z from 0-100 m/z , and 1 second per 100 m/z

above 100 m/z . Between 5-50 scans were averaged to obtain representative data. The resolution was normally set to obtain isotopic separation of the molecular species at 5% valley. Operation and data processing was controlled by VG Opus™ software.

(b) Tandem Mass Spectrometry

For negative and positive ion MS/MS experiments carried out on the VG Autospec™, a collision cell voltage (E_{lab}) of 400 V was used for singly-charged precursor ions and 200 V for all multiply-charged precursor ions. For typical MS/MS analyses, the pressure of the collision gas (methane) in the cell was adjusted to correspond to approximately 30% transmission of the incident ion beam. All MS/MS spectra were acquired in the continuum mode over the m/z range of 1-3500 Da at a rate of typically 20 seconds per acquisition. Spectra from between 5-50 scans were then averaged to obtain representative data. In the majority of studies on oligonucleotide fragmentation, the pressure in the cell was altered during the course of the acquisition of MS/MS spectra for each precursor ion depending on the type of fragmentation desired (i.e. low or high mass product ions or both). Where the intensity of the precursor ion was adequate, the resolution in MS 1 was increased sufficiently to enable separation of the molecular species with differing number of ^{13}C isotopes. The ion corresponding to the species that contained only ^{12}C was selected as the precursor ion for MS/MS analysis.

(c) Tandem Mass Spectrometry of Source-Generated Product Ions

A series of experiments were conducted on product ions of interest observed in the MS/MS spectra of molecular ions of oligonucleotides and ligand-oligonucleotide adducts. The product ions were firstly generated in the ion source by optimising the potential of the sampling cone and skimmer lens so as to induce sufficient collisional activation of the molecular species during ion acceleration in the ESI interface. The in-source generated fragment ion was then selected as the precursor ion in MS 1 (the magnetic sector), passed through to the collision cell in MS 2 (the oa-TOF) where they were subjected to CID after which the MS/MS spectrum of the fragment or product ion

was acquired using the conditions as described in the section above for precursor ions derived from intact oligonucleotides. Where the intensity of the source-generated product ion was adequate, the resolution in the magnetic sector analyser was increased sufficiently to enable separation of the different isotopic species, and the ion corresponding to the species containing only ^{12}C was selected as the precursor ion for MS/MS experiments.

(d) Instrument Tuning and Calibration

For all ESI-MS, MS/MS, multiple stage MS and high resolution MS experiments undertaken on the AutospecTM, the instrument was tuned and calibrated prior to analysis. For optimisation of tuning parameters, a solution of 30 pmol μL^{-1} sodium dodecyl sulphate (SDS) in 50% aqueous isopropanol was used for tuning in the negative ion mode. A solution of 30 pmol μL^{-1} of leucine enkephalin was used for positive ion tuning. Solutions were continuously infused into the mass spectrometer at a flow rate of 10 μLmin^{-1} . Calibration of the mass range of the instrument for all positive and negative ion experiments described in this work was carried out with a solution of aqueous sodium iodide, 1 mgmL $^{-1}$. The calibration solution was injected via a loop injector (Rheodyne 7125) and delivered into the instrument in 50% aqueous methanol at a flow rate of 5 μLmin^{-1} . Approximately 20 μL of sample was injected for calibration of the mass range of MS 1 and 40 μL of sample was used for calibration of the mass range of MS 2.

2.3.2 QuattroTM Triple Quadrupole Mass Spectrometer

(a) Instrumentation

A VG QuattroTM mass spectrometer equipped with an electrospray ionisation source (VG Biotech (Micromass), Altrincham, UK) was used for both MS and MS/MS analysis. A schematic of this instrument is shown in figure 2.4. This instrument has a quadrupole/hexapole/quadrupole mass analyser configuration and a mass range of

m/z 0-4000. The solvent was delivered through fused silica capillary tubing by an ISCO (Lincoln, NE, USA) SFC-500 microflow syringe pump. Flow rates of between 5-15 μLmin^{-1} were typically employed. The formation of a stable spray was assisted by a flow of nitrogen nebulising gas concurrent to the stainless steel capillary in the probe at a flow rate of $\sim 10 \text{ Lhr}^{-1}$. To enhance droplet evaporation, a flow of dry, warm, nitrogen bath gas was applied counter-current to the solvent flow at a rate of $\sim 350 \text{ Lhr}^{-1}$. The potential on the electrospray probe tip was typically -3.5 kV with -0.5 kV on the chicane counter electrode (for negative ions). The photomultiplier was set to 650 V.

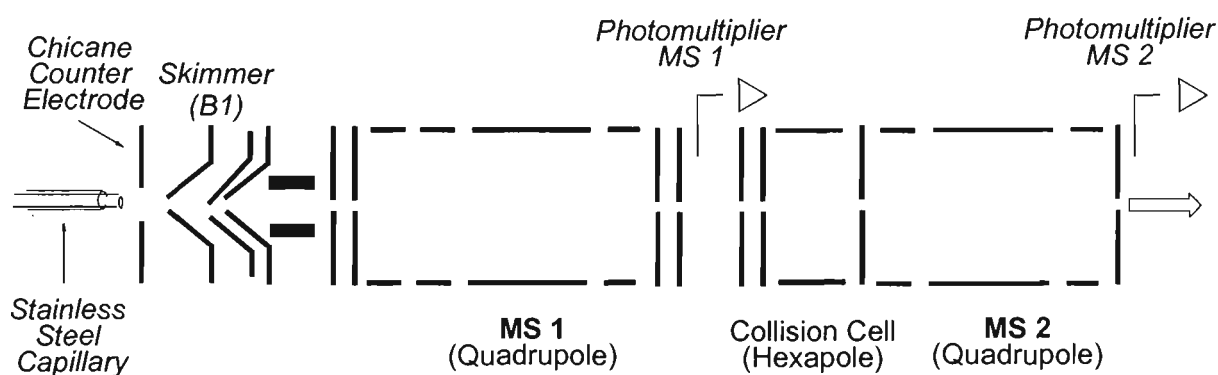


Figure 2.4: Schematic representation of the VG Quattro™ triple quadrupole mass spectrometer.

All MS spectra were acquired in the negative ion mode (ES^-) by multi-channel analysis (MCA) at a scan rate of 1 second per 100 m/z units and between 10-30 scans were summed to obtain a representative spectrum. For ESI-MS analysis, the low mass (LM) and high mass (HM) resolution parameters were set to 12.5 and 10.0 respectively (arbitrary units) corresponding to a resolution of 500 at 50% valley and spectra were typically acquired over the range of m/z 400-1800.

(b) Tandem Mass Spectrometry

In MS/MS experiments performed on the VG Quattro™, the resolution in the first quadrupole was set to a minimum to allow for maximum transmission of ions through to the second quadrupole. A collision energy of 80 eV typically was used in the

analysis of all charge states of unmodified oligonucleotides. The collision gas (argon) pressure was varied as indicated in the text to generate the degree of fragmentation appropriate for the experiment. Typical parameters used for both MS and MS/MS analysis are shown in table 2.3.

Table 2.3: Typical operating parameters used for MS and MS/MS experiments conducted on the Quattro™

Source Parameters	
Capillary	3.0-3.5 kV
HV Lens (Chicane Counter Electrode)	0.3-0.5 kV
Cone (B1 Skimmer lens)	30-100V
Lens 1	50-80 V
Lens 2	1-10 V
Lens 3	35-50 V
Lens 4	85-115 V
Source Temperature	70°C
MS 1	
LM Resolution	12.5
HM Resolution	11.0
Energy Filter	6.0
Ion Energy	2.0
Ion Energy Ramp	0.0
Lens 5	147 V
Lens 6	0-10 V
MS 1 Photomultiplier	650 V
MS 2	
LM Resolution	12.5
HM Resolution	9.0
Collision	0 V
Ion Energy	4.0
Ion Energy Ramp	0.0
Lens 7	300-350 V
Lens 8	300-350 V
Lens 9	60-80 V
MS 1 Photomultiplier	650 V

(c) Instrument Tuning and Calibration

Tuning of the instrument parameters in the negative ion mode was normally achieved by optimising the intensity of the doubly charged ion of the oligonucleotide 5'-d(CACGTG)-3'. A stock solution of the oligonucleotide sample (prepared in water at

a concentration of $100 \text{ pmol}\mu\text{L}^{-1}$), was injected via a loop injector (Rheodyne 7125) and delivered to the instrument by a solvent flow of 50% aqueous methanol at a flow rate of $10 \text{ }\mu\text{Lmin}^{-1}$. The mass range of the instrument for both MS and MS/MS analysis was calibrated with a sugar mixture composed of 0.5 mg/ml of each raffinose, maltotetraose, dextrose, and maltooligosaccharide (corn syrup) in 50% aqueous acetonitrile.

2.4 SAMPLE ANALYSIS

2.4.1 Oligonucleotides

For initial studies involving the optimisation of ESI-MS of oligonucleotides on the Autospec™, the oligonucleotide 5'-d(CACGTG)-3' was analysed in solvent compositions of 50% aqueous methanol, 50% aqueous acetonitrile, and 50% aqueous isopropanol. In each case, oligonucleotide solutions were prepared by dissolving HPLC-purified and desalted samples in water and this solution was diluted 1:1 with the organic solvent to give a final concentration of $50 \text{ pmol}\mu\text{L}^{-1}$ of the single stranded oligonucleotide. The effect of organic solvent on oligonucleotide signal intensity was examined by acquiring ESI mass spectra of 5'-d(CACGTG)-3' in solvents with varying percentages of isopropanol (the organic solvent found to yield optimum oligonucleotide signal intensity) in MilliQ™ water over the range 0-90% isopropanol. The optimisation of sample concentration and flow rate was carried out by measuring the total ion signal of the oligonucleotide 5'-d(CACGTG)-3' generated in ESI-MS spectra obtained at concentrations over the range of $5\text{-}100 \text{ pmol}\mu\text{L}^{-1}$ analysed at flow rates of $5 \text{ }\mu\text{Lmin}^{-1}$, $10 \text{ }\mu\text{Lmin}^{-1}$ and $20 \text{ }\mu\text{Lmin}^{-1}$ for each concentration of oligonucleotide. For the optimisation of individual charge state ions in the ESI-MS spectra of oligonucleotides, the effects of adjusting sampling cone and skimmer lens voltages on the intensities of charge state ions of 5'-d(CACGTG)-3' were examined. To investigate further the influence of solvent composition on the charge states of oligonucleotide ions observed, ESI mass spectra of 5'-d(CACGTG)-3' were acquired in solvents containing either 1% acetic acid or 1% ammonium hydroxide in 50% isopropanol.

In order to examine the effects of instrument type on the fragmentation pathways observed for oligonucleotide anions, MS/MS analysis was conducted on the series of tetramers 5'-d(CCCC)-3', 5'-d(CCCA)-3', 5'-d(ACCT)-3', and 5'-d(TCCA)-3' since these oligonucleotides had been the subject of previous detailed MS/MS studies on ion trap instrumentation. MS/MS spectra of these oligonucleotides were acquired in the negative ion mode on both the Quattro™ and Autospec™ mass spectrometers for all precursor ion charge states. For the purposes of characterising the product ions of oligonucleotide fragmentation in both negative and positive ion MS/MS, a detailed examination of the product ions observed in the MS/MS spectra of the $[M-H]^-$ and $[M+H]^+$ ions of each of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' was carried out using multiple stage mass spectrometry experiments.

For the investigation of the effects of precursor ion charge state on the fragmentation behaviour of oligonucleotide anions, MS/MS spectra were acquired in the negative ion mode for all charge state anions (1^- to 5^-) of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3', and 5'-d(GCCGGC)-3'. For studies conducted into the effects of base identity and sequence on the negative ion MS/MS fragmentation of oligonucleotides, MS/MS spectra were acquired of the $[M-2H]^{2-}$ ions of all hexanucleotides listed in table 2.1.

2.4.2 *Ligand-Oligonucleotide Adducts*

All ligand-oligonucleotide adducts of hedamycin, DC92-B and phenC₆Br were prepared for analysis by dissolving the dried, purified adduct in 50% methanol/water to give a concentration of 50 pmol μ L⁻¹. The solvent compositions used for oligonucleotide adduct samples were selected in order to yield optimum signal intensity of adduct samples. For initial molecular weight determination of ligand-oligonucleotide adducts, and examination of sample purity, all adduct samples were analysed by ESI-MS on the Quattro™ mass spectrometer. All subsequent MS/MS experiments of ligand-oligonucleotide adducts were carried out on the Autospec™ mass spectrometer. Ligand-

oligonucleotide adducts were analysed in the negative ion mode, and, in some cases, MS/MS analysis in the positive ion mode was also conducted to obtain additional structural information.

REFERENCES

- (1) Caradonna, J. P.; Lippard, S. J. *Inorg. Chem.* **1988**, 27, 1454-1466.
- (2) Wickham, G.; Prakash, A. S.; Wakelin, L. P. G.; McFadyen, W. D. *Biochim. et Biophys. Acta* **1991**, 1073, 528-537.

Chapter 3

ELECTROSPRAY IONISATION MASS SPECTROMETRY OF OLIGONUCLEOTIDES

3.1 INTRODUCTION

For the studies conducted into the ESI-MS/MS fragmentation of oligonucleotides, an initial optimisation of the conditions for analysis of oligonucleotides was necessary. Furthermore, the need to optimise for the observation of oligonucleotide ions of a range of charge states for subsequent MS/MS experiments required a preliminary investigation of the parameters affecting the charge state distribution of oligonucleotide anions in electrospray mass spectra. The phosphodiester backbone of an oligonucleotide is deprotonated and fully charged in solution at physiological pH, and so the negative ion ESI mass spectra of oligonucleotides are generally characterised by the observation of molecular ion species with different numbers of negative charges. As outlined in chapter 1 (section 1.5), a variety of factors have been cited in the literature as influencing the charge states of oligonucleotide ions observed in ESI mass spectra. These include: changing the solvent composition and pH^{1,2}, adjusting the voltage applied across the ESI interface^{1,3}, altering the source temperature and bath gas^{4,5}, gas-phase proton transfer reactions and gas-phase dissociation of complexes^{1,6-9}. The methods explored in this work were the manipulation of ESI interface conditions, and the addition of acids and bases to the oligonucleotide solutions.

3.2 OPTIMISATION OF ESI-MS OF OLIGONUCLEOTIDES

3.2.1 Solvent Composition

Solvent compositions of methanol, 2-propanol, and acetonitrile were examined in order to find the optimum organic solvent composition for ESI analysis of

oligonucleotides. The oligonucleotides analysed in the present study had been subjected to prior purification and desalting by reversed-phase HPLC which was found to produce excellent results for overcoming problems associated with the adduction of ubiquitous sodium and potassium to the phosphate backbone in ESI analysis of oligonucleotides. Therefore no attempts were made to examine other methods of manipulating solution conditions for removal of cation adduction. In these preliminary studies, the oligonucleotide 5'-d(CACGTG)-3' was dissolved in water and then diluted with 50% of each organic solvent examined yielding a concentration of 50 pmol μ L⁻¹. ESI mass spectra of each sample were recorded at flow rates of 20 μ L/min and at cone voltages which produced optimum signal intensities of the [M-2H]²⁻ ion (m/z 895) and [M-H]⁻ ion (m/z 1790) of the oligonucleotide respectively (the influence of cone voltage on oligonucleotide charge states is discussed in detail below). Figure 3.1(a) shows the intensity of the [M-2H]²⁻ ion using each solvent composition (at the optimum cone voltage). The summed intensities of all the molecular ions of different charges (the total oligonucleotide signal intensity) is also shown in the figure. The corresponding results obtained at the optimum cone voltage for the [M-H]⁻ ion are given in figure 3.1(b). A solvent composition of 50% isopropanol was found to yield superior oligonucleotide signal intensity of the three solvents tested in both cases. This was most evident in spectra where the [M-2H]²⁻ ion was optimised which showed a substantial improvement in both the absolute intensity of the 2⁻ ion and the absolute intensity of all molecular ion species when 50% isopropanol was used. In ESI spectra acquired at the optimum cone voltage for the [M-H]⁻ ion (figure 3.1(b)), a solvent composition of 50% aqueous acetonitrile produced the greatest signal intensity for the [M-H]⁻ ion although the total oligonucleotide ion intensity was slightly less than that associated with the use of 50% isopropanol. Thus isopropanol was selected as the optimum organic solvent for the analysis of oligonucleotides on the magnetic sector orthogonal time-of-flight instrument (Mag oa-TOF) used in this study.

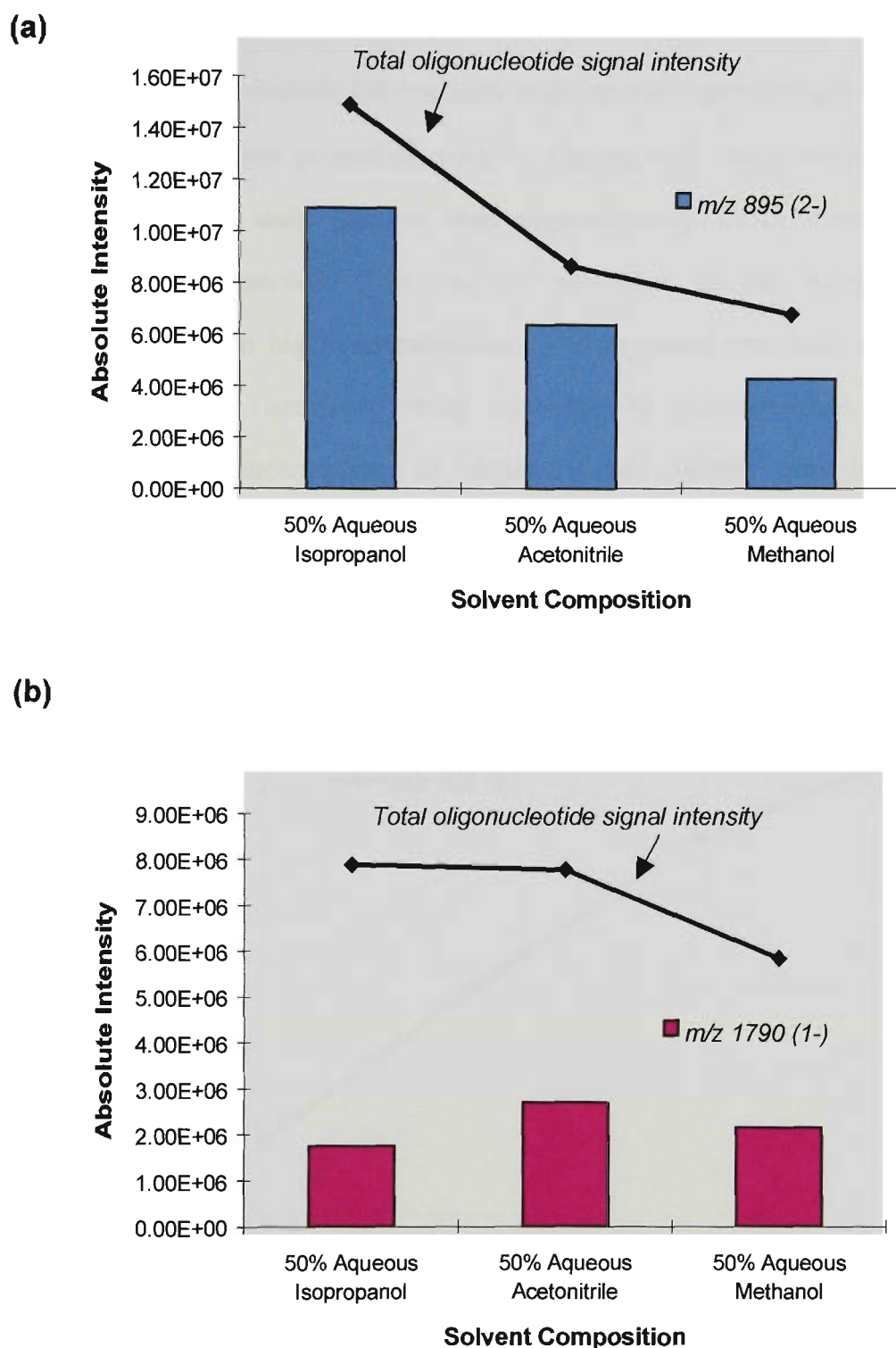


Figure 3.1: Variation of optimum signal intensities of the (a) 2^- charge state ion, and (b) 1^- charge state ion in the ESI mass spectra of the oligonucleotide 5'-d(CACGTG)-3' in different solvent compositions.

The effect of different percentages of organic solvent was also examined and figure 3.2 shows how the absolute intensity of the $[M-2H]^{2-}$ ion (the most intense ion observed in ESI spectra at the cone voltage used) varied with increasing composition of isopropanol. A steady increase in the ion intensity was observed with increasing

percentages of organic solvent up to a concentration of 75%. This observation of increasing oligonucleotide ion intensity with increased percentages of organic solvent is consistent with other published work¹⁰. Alternatively, when higher concentrations of isopropanol were used (greater than approximately 75%), a sharp drop in signal intensity was observed. This can be attributed to the decreased solubility of oligonucleotides in high concentrations of isopropanol (the least polar of the organic solvents tested). Therefore, owing to solubility considerations, and the fact that relatively high concentrations of oligonucleotide sample were being used so as to produce optimum quality MS/MS spectra, a composition of 50% isopropanol was preferred for use in all subsequent analyses.

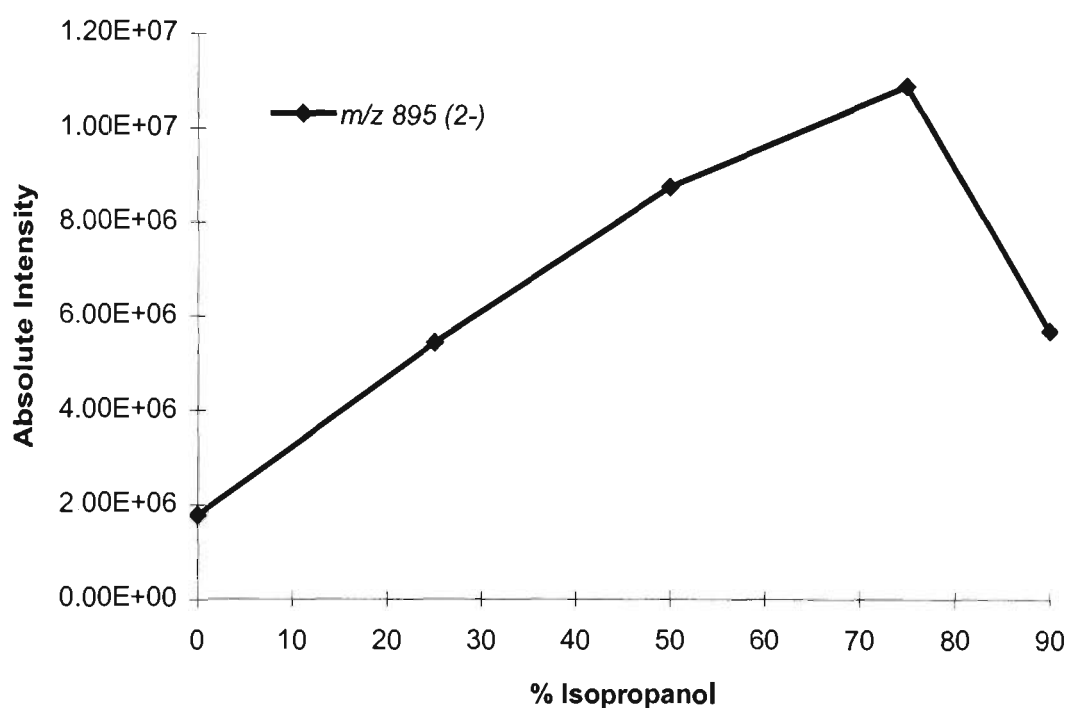


Figure 3.2: The effect of the percentage of isopropanol on the absolute intensity of the $[M-2H]^{2-}$ ion in the ESI mass spectra of 5'-d(CACGTG)-3'.

3.2.2 Sample Concentration and Flow Rate

The total ion signal of the oligonucleotide 5'-d(CACGTG)-3' was measured at concentrations ranging from 5 pmol μL^{-1} to 100 pmol μL^{-1} , and at flow rates of 5 μLmin^{-1} , 10 μLmin^{-1} and 20 μLmin^{-1} for each concentration. Figure 3.3 shows the variation of the summed oligonucleotide ion intensity at each flow rate as a function of

sample concentration. The oligonucleotide signal intensity was observed to increase upon increasing the sample concentration to $100 \text{ pmol}\mu\text{L}^{-1}$, and, for each sample concentration, the intensity was greater at higher flow rates. An increase in the flow rate from $10 \text{ }\mu\text{Lmin}^{-1}$ to $20 \text{ }\mu\text{Lmin}^{-1}$ however, produced a smaller percentage increase in signal intensity compared with increasing the flow rate from $5 \text{ }\mu\text{Lmin}^{-1}$ to $10 \text{ }\mu\text{Lmin}^{-1}$. From the results obtained, the conditions chosen for MS/MS analysis in the current work were a concentration of $50 \text{ pmol}\mu\text{L}^{-1}$ with a flow rate of typically $5 \text{ }\mu\text{Lmin}^{-1}$, but ranging up to $10 \text{ }\mu\text{Lmin}^{-1}$ depending on the sensitivity required for the particular experiment. These conditions were selected as resulting in the best compromise between signal intensity and sample consumption.

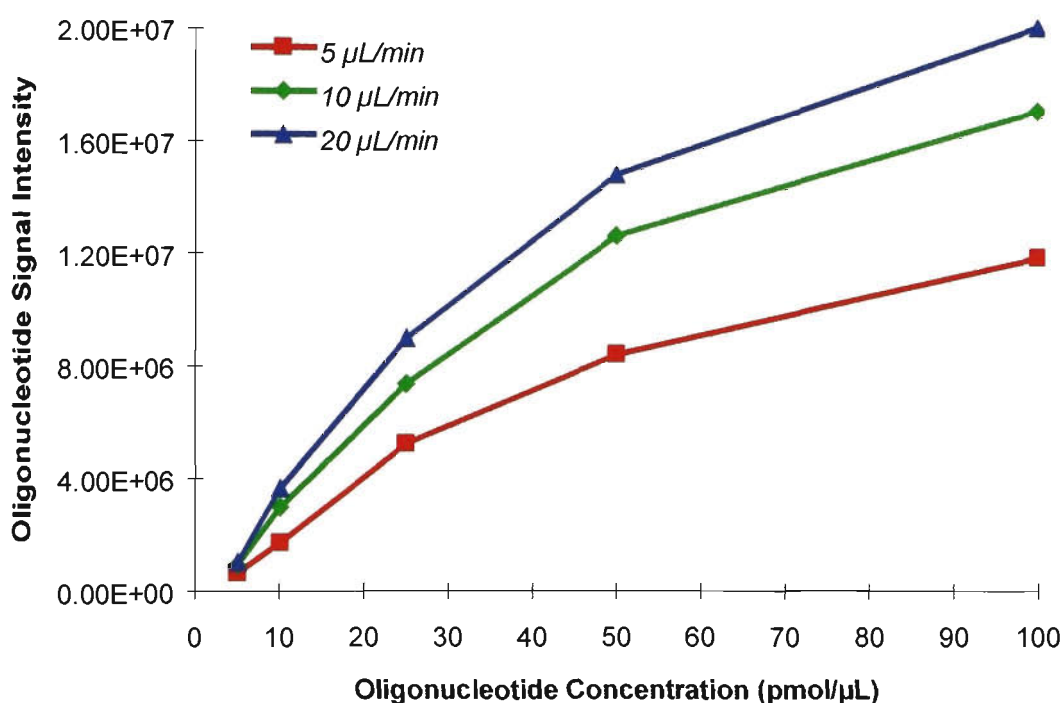


Figure 3.3: The effect of both flow rate and concentration on the total oligonucleotide signal generated in the ESI mass spectra of 5'-d(CACGTG)-3'.

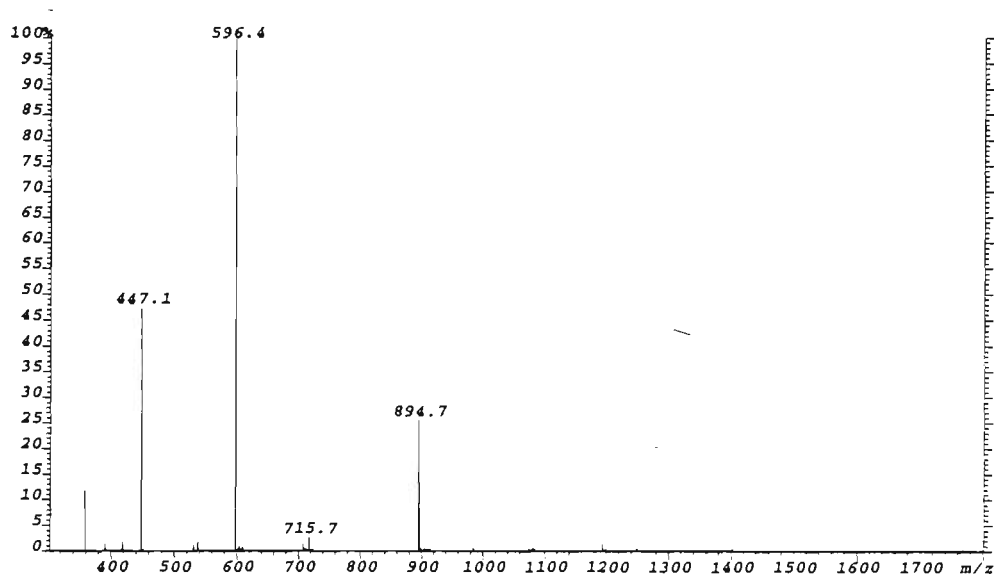
3.3 FACTORS INFLUENCING THE CHARGE STATE DISTRIBUTION OF OLIGONUCLEOTIDE ANIONS

3.3.1 Electrospray Ionisation Source Parameters

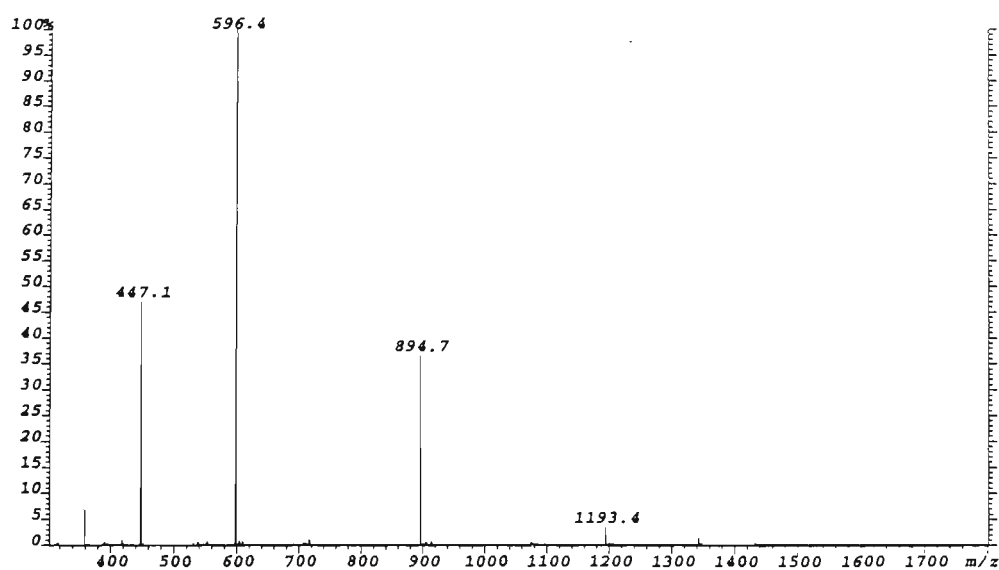
For the optimisation of individual charge states of oligonucleotides, an examination of the effects of sampling cone and skimmer lens voltages on the charge

state distribution and focusing efficiency of oligonucleotide anions in ESI mass spectra were also examined. Experiments were conducted on the oligonucleotide 5'-d(CGTACG)-3' ($M_r=1791.34\text{Da}$) whereby the cone and skimmer voltages were adjusted so as to produce optimum signal intensity for each ion of different charges. Figures 3.4 shows the ESI mass spectra generated at parameters optimising for the $[\text{M}-5\text{H}]^{5-}$ (m/z 357), $[\text{M}-4\text{H}]^{4-}$ (m/z 447), $[\text{M}-3\text{H}]^{3-}$ (m/z 596), $[\text{M}-2\text{H}]^{2-}$ (m/z 895), and $[\text{M}-\text{H}]^{-}$ (m/z 1790) ions respectively.

(a)



(b)



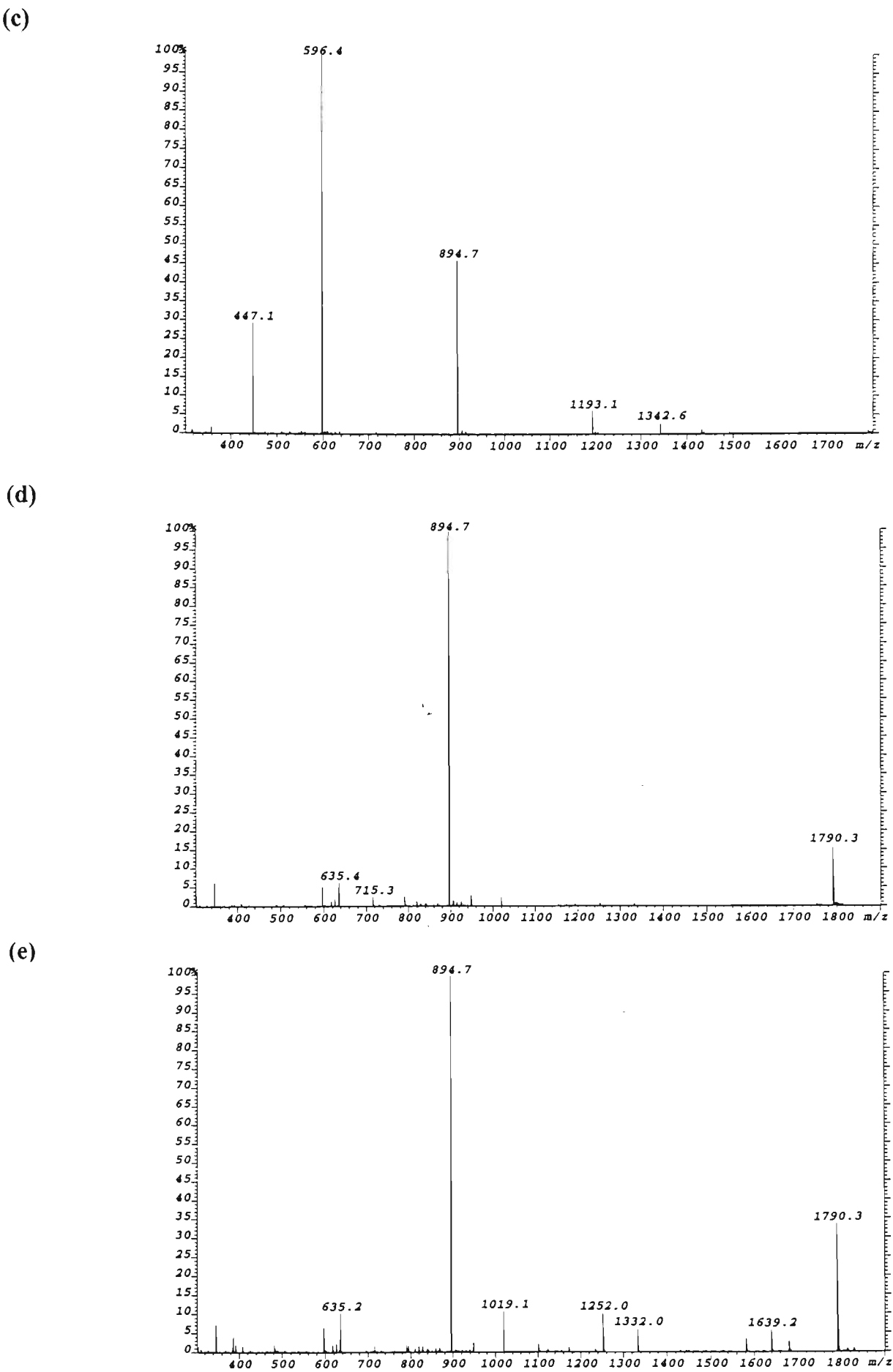


Figure 3.4: ESI mass spectra generated at sampling cone and skimmer voltages optimising for the (a) $[M-5H]^{5-}$ (m/z 357), (b) $[M-4H]^{4-}$ (m/z 447), (c) $[M-3H]^{3-}$ (m/z 596), (d) $[M-2H]^{2-}$ (m/z 895), and (e) $[M-H]^{-}$ (m/z 1790) ions of 5'-d(CGTACG)-3'.

Significant shifts in ion charge state distribution were observed upon altering the cone and skimmer lens voltages, with ions of lower charge state being optimised at successively higher voltages. This effect is illustrated in figure 3.5 which shows how the intensity of each charge state ion varies as a function of the sampling cone voltage (the voltage on the skimmer lens was approximately 10-15V higher than that applied to the cone). Ions of higher charge state, in particular the 5⁻, 4⁻, and 3⁻ ions, were optimised at much lower cone voltages (cone=18.0V; skimmer lens=30.0V) compared with the 2⁻ and 1⁻ ions (cone =40.7V, skimmer lens=54.1V). In ESI mass spectra recorded at cone voltages below 18.0V, the singly charged species was not observed whereas above 40.7V, the 4⁻ and 5⁻ ions were absent. The 3⁻ and 2⁻ species were generated over the entire range of cone voltages used (cone=11.8V-67.4V, skimmer lens=21.5V-82.3V) with the 3⁻ ion being the dominant charge state observed in all ESI mass spectra recorded at low voltages, and the 2⁻ ion dominating spectra recorded at higher voltages.

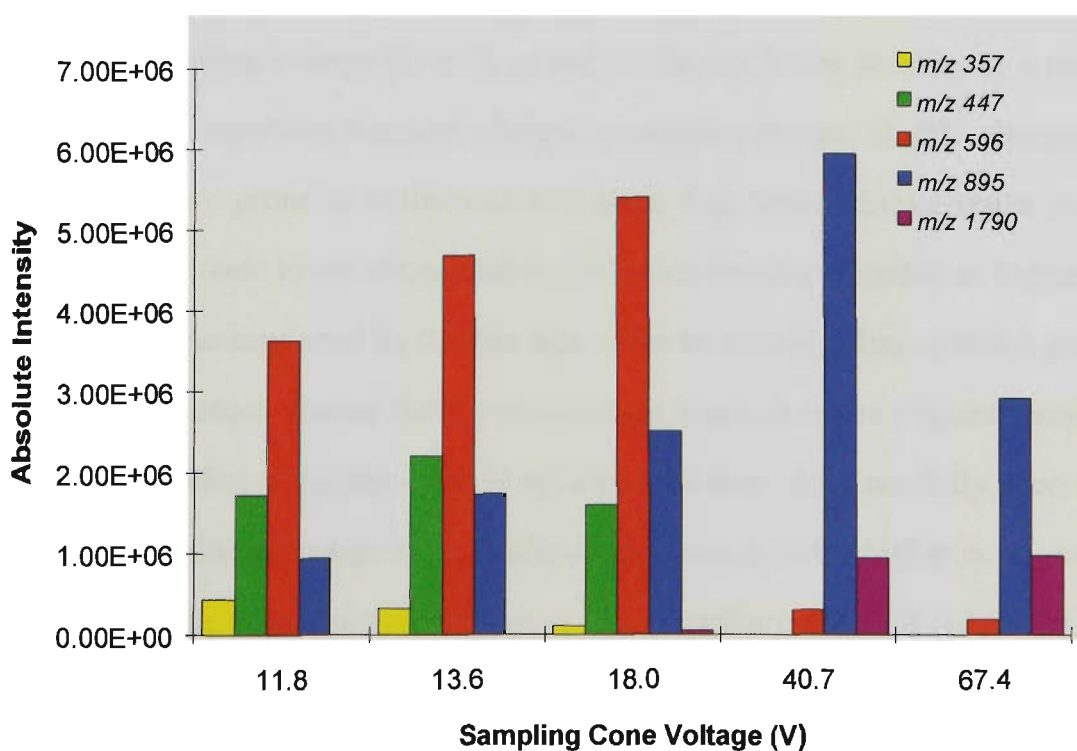


Figure 3.5: Variation of the absolute intensities of each charge state ion as a function of the sampling cone voltage in the ESI mass spectra of 5'-d(CGTACG)-3'.

The effect of increasing cone voltages on shifting ion charge state distributions to lower charge states has been observed in previous studies of oligonucleotides^{11,12}, proteins^{3,13}, and synthetic polymers¹⁴. Although it has been proposed that the major factor influencing the charge state distribution is the ion distribution in solution, numerous studies have demonstrated the ability to bring about charge exchange of ions in the gas-phase^{6-8,15,16}. When the voltage difference in the intermediate atmospheric pressure region is increased, ions accelerating through the greater electrostatic field in the atmospheric pressure region undergo more extensive and energetic collisions with bath gas and/or other background molecules. The results obtained here show that an extensive number of fragment ions are present in the ESI mass spectra recorded at high cone and skimmer lens voltages (figure 3.4(d) and (e)). Dissociation of the oligonucleotide in the ESI interface was found to give rise to ions owing to base cleavage ($[M-CH]^-$ at m/z 1678.9 and $[M-GH]^-$ at m/z 1638.9), and phosphate backbone cleavage: the full range of w-type ions (w_1^- (m/z 346.0), w_2^- (m/z 635.2), w_3^- (m/z 948.2), w_4^- (m/z 1252.0), w_5^- (m/z 1580.9)), and (a-B)-type ions ($[a_2-B_2H-2H]^-$ (m/z 386.1), $[a_3-B_3H-2H]^-$ (m/z 715.2), $[a_4-B_4H-2H]^-$ (m/z 1019.2)) are observed. Compared with lower charge states, ions of higher charge possess a greater amount of internal energy at the applied accelerating voltage ($E=qV$), as well as having lower stability as a result of greater Coulombic repulsion between charged phosphate groups. Highly charged ions would thus be more prone to collisional activation than lower charge states and this could also explain their lower abundance in ESI mass spectra recorded at higher cone voltages. This is also supported by the fact that at the high cone voltages which generate fragmentation, ESI mass spectra show a concomitant decrease in the oligonucleotide ion current. Fragmentation of highly charged species, however, does not fully account for the observed alteration of charge state distribution associated with higher cone voltages, since if this were the case, the shift in charge state distribution would only be relative, and no increase would be observed for the absolute abundances of the lower charge state ions. This is clearly not the case since the ESI spectra recorded at the higher skimmer voltages yielding fragmentation (40.7V) corresponded to optimum absolute intensities

of the singly and doubly charged ions respectively (figure 3.4(d) and (e)). From studies conducted on the effects of cone voltage on polymer and peptide ions it has been proposed that ions of the same mass but differing charge are focussed in proportion to their mass-to-charge ratios under the influence of an electric field with ions of lower m/z having optimum focussing voltages¹⁴. Therefore the results observed for the dependence of oligonucleotide charge state on skimmer voltages can be rationalised as arising from a combination of fragmentation of higher charge state ions in the ESI interface in addition to these focussing effects.

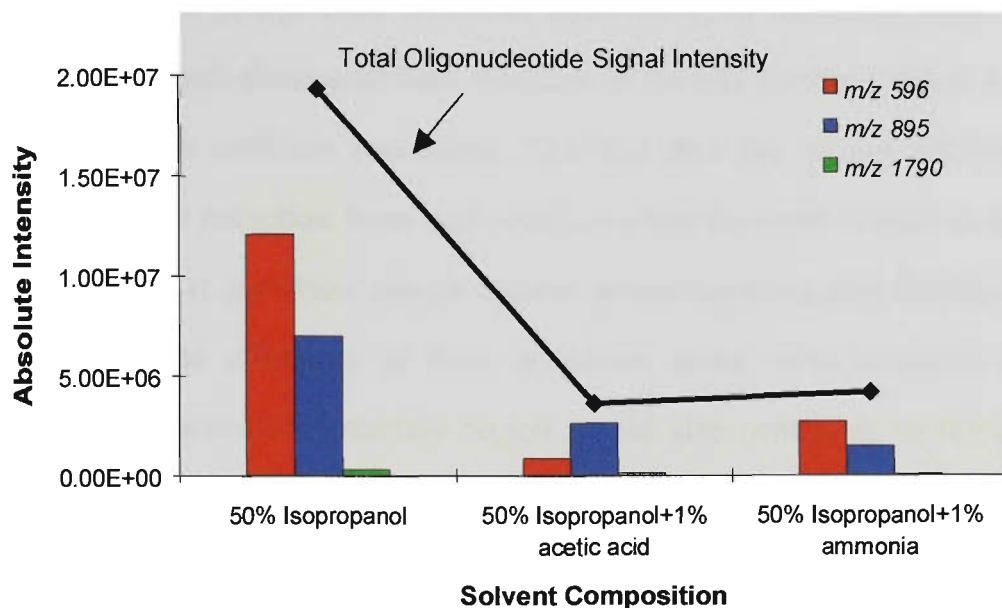
3.3.2 *Addition of Acids and Bases*

Solution pH influences the extent of ion formation in solution, and as such has been found to be an important factor determining the charge state distribution of ions in ESI mass spectra¹⁷. In solution, and under physiological pH, oligonucleotides exist predominantly as deprotonated species since the phosphodiester groups are fully deprotonated above pH~1 and the nucleobases and sugar groups are neutral. Owing to the high acidity of the phosphate backbone, the addition of acids and bases would not be expected to have any significant effect on phosphate ionisation. The nucleobases, however, can be protonated on decreasing the pH (G, A, and C, are protonated in pH ranges of 3.3-3.6, 3.5-4.2 and 4.2-4.6 respectively) and deprotonated on increasing the pH (G and T are deprotonated above pH 9.4-10.5 respectively). In addition to influencing charge formation of oligonucleotides in solution, the addition of acids and bases has also been found to effect the charge states of oligonucleotides via processes occurring in the gas phase (acid-base proton-transfer reactions), and the formation and dissociation of oligonucleotide-base complexes¹. In the present study the effect of addition of 1% acetic acid and 1% ammonium hydroxide on the ESI mass spectra of the oligonucleotide 5'-d(CACGTG)-3' was examined. Since the voltage applied across the intermediate atmospheric pressure region is an important factor driving gas-phase proton-transfer reactions and the dissociation of oligonucleotide complexes, experiments were conducted at low (20.8V), intermediate (44.0V) and high cone

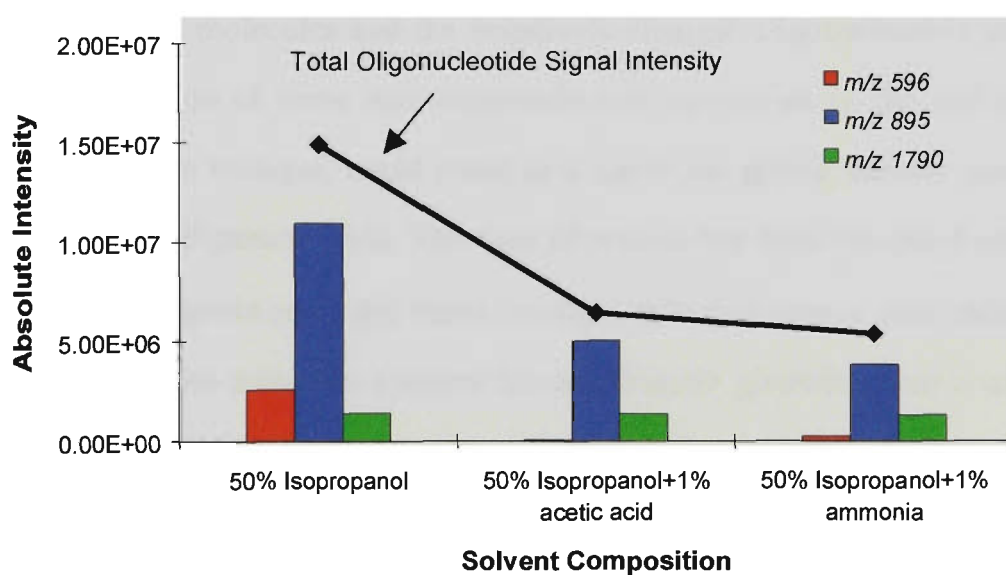
voltages (64.0V) which correspond to the optimisation of the 3⁺, 2⁺ and 1⁺ charge states of this oligonucleotide respectively. In order to assess the effects of addition of acid and base on the observed oligonucleotide charge state distributions at different cone voltages, the ESI spectra obtained at each cone voltage were compared to the ESI spectrum of the oligonucleotide solutions to which no acid and base had been added recorded with the same interface parameters. Figure 3.6 shows the effect of addition of acid and base on the distribution of the 3⁺, 2⁺ and 1⁺ ions, and on the total oligonucleotide ion signal intensity at each respective cone voltage (the ESI spectra obtained are given in Appendix 3.1).

The addition of acid at lower cone voltages was found to result in a large decrease in intensity of the triply charged species with a shift in the relative ion intensities to the doubly charged ion. These effects were also accompanied by significant ion suppression (an 80% decrease in oligonucleotide signal intensity). At a cone voltage optimum for the intensity of the doubly charged ion (figure 3.6(b)), the addition of acid was found to increase the proportion of singly charged ions relative to the total oligonucleotide ion signal from 9% to 21%. The absolute abundance of this ion, however, was approximately the same since the total oligonucleotide ion signal had decreased by 57%. The most efficient charge state reduction from acid addition was observed at a cone voltage of 64.0V which optimised for the singly charged species. In the ESI mass spectrum to which no acid has been added the singly charged species constituted approximately 12% of the total oligonucleotide ion signal. When acid was added to the solution using these interface parameters, an increase in the absolute abundance of the singly charged ion by about 53% was observed with a shift in the relative ion distribution such that the singly charged ion now comprised 50% of the total oligonucleotide ion signal and was the most predominant charge state observed. The decrease in oligonucleotide signal was also much lower than that experienced at lower cone voltages (approximately 30% reduction).

(a)



(b)



(c)

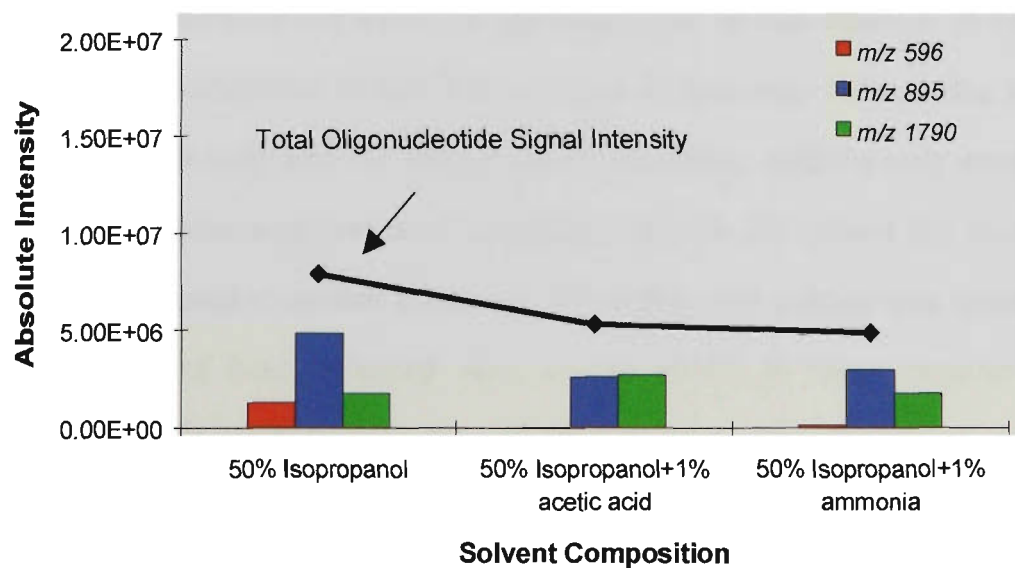


Figure 3.6: The effect of acid and base addition on the distribution of the 3⁺, 2⁺ and 1⁺ charge state ions, and the total oligonucleotide ion signal intensity generated in the ESI mass spectra of 5'-d(CACGTG)-3' acquired at cone voltages of (a) 20.8V, (b) 44.0V, and (c) 64.0V.

The enhanced charge state reduction observed upon increasing cone voltages may be as a result of gas-phase acid-base reactions in the ESI interface which are driven by the higher energy collision conditions. The fact that the results obtained show optimum charge state reduction from acid addition when the cone voltage is increased appears to suggest that gas-phase charge transfer interactions are contributing to these effects. The improved efficiency of these processes under more energetic collision conditions in the intermediate interface region would also contribute to the observed reduction in oligonucleotide ion suppression effects as the cone voltage is increased. Alternatively, these findings could be explained by the formation, and subsequent dissociation, of complexes arising from non-specific hydrogen bonding interactions between neutral acid molecules and the negatively charged oligonucleotide phosphate backbone. Dissociation of these acid-oligonucleotide complexes in the ESI interface induced by high cone voltages, could result in a gas-phase proton transfer mechanism from the acid to the oligonucleotide. This type of process has been described previously in studies utilising organic acids and bases for oligonucleotide charge state reduction^{1,2}. The occurrence of these processes however are questionable given that ions arising from intact acid-oligonucleotide complexes were not found to be present in ESI mass spectra obtained over the range of cone voltages examined.

The addition of 1% ammonium hydroxide to the oligonucleotide solution was also examined to determine the effect of increasing pH on the relative charge state distribution of oligonucleotides in ESI mass spectra. At low cone voltage, the addition of base did not significantly alter the charge state distribution, with the only appreciable effect being a significant suppression of the oligonucleotide ion current (by about 78% of the intensity compared to neutral solutions). When the cone voltage was increased to ~40V, the addition of base produced very similar effects to those observed upon addition of acid. The intensity of the doubly charged ion relative to the total oligonucleotide ion signal remained approximately the same whereas that of the singly charged ion increased from 9% to 24%, and the relative abundance of the triply charged species decreased from 18% to 5%. These effects were accompanied by a decrease in

the absolute intensities of ions from all charge states corresponding to a loss in oligonucleotide ion current of about 64%. At high cone voltages, the relative abundance of the singly charged ion as a fraction of total oligonucleotide ion signal increased from 22% to 37% upon base addition, albeit it can be seen that the absolute intensity of the singly charged ion remained approximately the same and the signal intensities of all other charge state ions decreased.

Overall, the addition of 1% ammonium hydroxide rather than promoting the formation of higher charge state ions, was found to shift the distribution of oligonucleotide ions to lower charges. Although the addition of base enhances deprotonation, the fact that oligonucleotides are already highly charged species in solution at very low pH suggests that this would have minimal effect on the degree of deprotonation of oligonucleotides in solution. The observed charge state reduction could thus be rationalised by the formation and subsequent dissociation of ammonium-phosphate ion pairs on the oligonucleotide backbone. The ability of oligonucleotides to form adducts with ammonium cations has been utilised extensively in past studies of ESI of oligonucleotides as a means of reducing sodium cation adduction. This has been attributed to the ammonium cations displacing sodium ions and forming ion pairs with the phosphate anionic backbone¹². Since ammonium-oligonucleotide adducts are generally absent from ESI mass spectra it has been suggested that dissociation of these ion pairs in the ESI interface results in gas phase proton transfer from ammonium ions to phosphate ions resulting in loss of neutral ammonia¹². This process would result in neutralisation of charged phosphate sites thus causing charge state reduction. The results obtained here show the most efficient charge state reduction following addition of base in ESI mass spectra recorded at the highest cone voltage. The fact there was no increase in the absolute abundance of the singly charged ion compared with the spectrum recorded at the same cone voltage to which no base had been added indicates that ion suppression was also a contributing factor to the observed attenuation of the singly charged ions.

3.4 SUMMARY

The charge state distribution of oligonucleotide anions observed in ESI mass spectra was found to be significantly influenced by the sampling cone and skimmer lens voltages, with low voltages favouring the production of highly charged species, and higher voltages predominantly yielding ions of low charge state. The dependence of oligonucleotide charge state distributions on the cone and skimmer lens voltages, enabled the intensities of individual charge state ions to be readily optimised through adjustment of these parameters, and this proved to be a valuable tool for the MS/MS studies described later in this thesis.

The charge state distribution of oligonucleotide ions observed in ESI mass spectra could also be altered through changing the solution pH by the addition of acid or base. Solution compositions containing either 1% acetic acid or 1% ammonia were both found to shift the relative charge state distribution of oligonucleotide ions to lower charge states, with high cone voltages producing optimum results. A significant decrease in total oligonucleotide ion intensity was observed in all cases following addition of acid and base, and the extent of this ion suppression upon addition of ammonia countered the observed effects of charge state reduction so that no improvement in the ion intensity of the doubly or singly charged ions could be achieved. Although acid addition was also accompanied by significant suppression of the oligonucleotide ion signal, it was found to produce much more efficient charge state reduction at higher cone voltages yielding an increase in the absolute intensity of the singly charged ion. The addition 1% acetic acid was therefore utilised in addition to adjustment of ESI source parameters in order to produce optimum intensity of the singly charged species for subsequent MS/MS analyses.

REFERENCES

- (1) Cheng, X.; Gale, D. C.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1995**, *67*, 586-593.
- (2) Muddiman, D. C.; Cheng, X.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 697-706.
- (3) Ashton, D. S.; Beddell, C. R.; Cooper, D. J.; Green, B. N.; Oliver, R. W. A. *Org. Mass Spectrom.* **1993**, *28*, 721-728.
- (4) Mirza, U. A.; Chait, B. T. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 173-181.
- (5) Leblanc, J. C. Y.; Wang, J. Y.; Guevremont, R.; Siu, K. W. U. *Org. Mass Spectrom.* **1994**, *29*, 587-593.
- (6) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 529-532.
- (7) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *J. Am. Chem. Soc.* **1995**, *117*, 11555-11562.
- (8) Herron, W. J.; Goeringer, D. E.; McLuckey, S. A. *Anal. Chem.* **1996**, *68*, 257-262.
- (9) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom.* **1995**, *30*, 1222-1229.
- (10) Bleicher, K.; Bayer, E. *Biol. Mass Spectrom.* **1994**, *23*, 320-322.
- (11) Cheng, X.; Chen, R.; Bruce, J. E.; Schwartz, B. L.; Anderson, G. A.; Hofstadler, S. A.; Gale, D. C.; Smith, R. D. *J. Am. Chem. Soc.* **1995**, *117*, 8859-8860.
- (12) Stults, J. T.; Marsters, J. C. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 359-363.
- (13) Loo, J. A.; Udseth, H. R.; Smith, R. D. *Rapid Commun. Mass Spectrom.* **1990**, *4*, 207.
- (14) Hunt, S. M. *Electrospray Mass Spectrometry of Polymers*; Hunt, S. M., Ed.; University of Wollongong: Wollongong, **1997**.
- (15) Winger, B. E.; Hofstadler, S. A.; Bruce, J. E.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 566-577.
- (16) McLuckey, S. A.; Goeringer, D. E. *Anal. Chem.* **1995**, *67*, 2493-2497.
- (17) Guevremont, R.; Siu, K. W. M.; Le Blanc, J. C. Y.; Berman, S. S. *J. Am. Soc. Mass Spectrom.* **1992**, *2*.

Chapter 4

ELECTROSPRAY TANDEM MASS SPECTROMETRY OF OLIGONUCLEOTIDES:***Characterisation of product ions generated by negative and positive ion MS/MS on hybrid sector-TOF instrumentation.***

4.1 INTRODUCTION

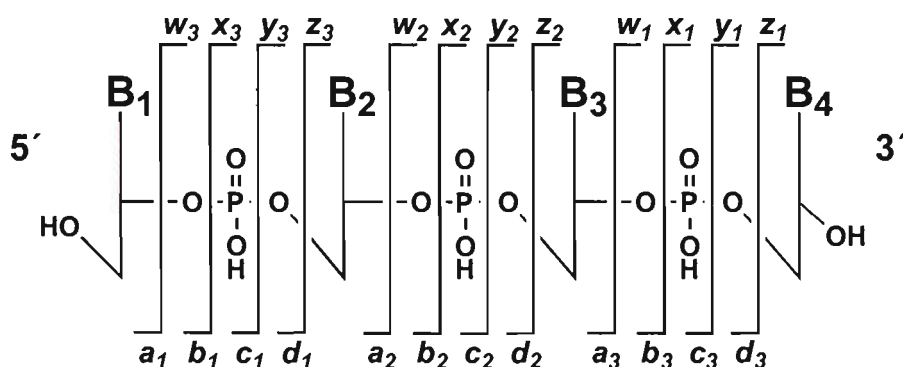
This chapter describes a preliminary investigation of the fragmentation observed in the negative and positive ion ESI-MS/MS spectra of unmodified oligonucleotides on the hybrid sector-TOF instrument in our laboratory. Prior to the commencement of this work, there had only been a few reports of the applications of ESI-MS/MS to oligonucleotides (see section 1.6 in chapter one). Furthermore, as outlined in chapter one, the majority of these studies with quadrupole ion trap instrumentation were aimed at deriving information regarding oligonucleotide fragmentation pathways^{1,2}. These data showed distinctly different oligonucleotide fragmentation pathways compared with subsequent reports utilising triple quadrupole³⁻⁶ and FT mass spectrometers⁷⁻¹¹. The greater degree of fragmentation observed in the MS/MS spectra obtained in these later studies was attributed to the more energetic collision conditions associated with analysis on these instruments. Since no previous studies had been undertaken of the ESI-MS/MS of oligonucleotides on a hybrid magnetic-sector time-of-flight instrument (in which case an even higher range of collision energies is available and in which, the time over which the dissociation occurs will be very different to the ion trap or FTMS), part of the work of this thesis involved a detailed study of the ESI-MS/MS fragmentation of oligonucleotides obtained on this instrument. For this study, identification of the product ions observed in the MS/MS spectra of positively and negatively charged precursor ions of oligonucleotides was facilitated by experiments whereby product ions of interest were also generated in the source and their MS/MS spectra examined. The

characteristic product ion types observed in the MS/MS fragmentation of deprotonated and protonated oligonucleotides were then summarised in order to establish the types of sequence and structural information derived from each ionisation mode. The product ions of singly charged protonated and deprotonated oligonucleotide anions was then examined and compared. The work described in this chapter constituted an important preliminary step for the later work concerning the sequence analysis of alkylated oligonucleotides by MS/MS (which is described in chapter six), in which it was necessary to utilise both ionisation modes for characterisation of ligand/DNA adducts.

4.2 OLIGONUCLEOTIDE FRAGMENTATION NOMENCLATURE

Throughout this work, the symbol M is used to represent the neutral oligonucleotide i.e. with an uncharged phosphate. The symbol s indicates the deoxyribose sugar group less the H₂O from either the 5'- or 3'-terminus (molecular formula, C₅H₆O₂, $M_r=98.0368$), and p is used to indicate the phosphate group, PO₃H ($M_r=79.9663$). Uncharged mononucleotides of either structure 5'-psB-3' or 5'-spB-3' are denoted as B_{nt} where B is one of the four bases, and uncharged polynucleotides are indicated by their constituent bases in parenthesis, for example (ACT) for the trinucleotide 5'-psApsCpsT-3' or 5'-sApsCpsTp-3'. The nomenclature employed for the 3'- and 5'-terminal sequence ions arising from cleavages along the phosphodiester backbone is derived from that originally proposed by McLuckey *et al.* and is shown in scheme 4.1 with the variation that the oligonucleotide is represented as a neutral species. Cleavages giving rise to sequence ions containing the 5'-terminus are denoted a, b, c, and d, and cleavages yielding ions with the 3'-terminus are denoted w, x, y, and z. The formation of singly charged b, d, w, and y ions, results in one hydrogen being transferred from the neutral fragment to the charged fragment upon cleavage. Therefore for singly charged anions from these cleavages there is no net hydrogen gain since the hydrogen lost upon ionisation is balanced by the hydrogen gained on transfer during dissociation. On the other hand, formation of singly charged cations of these fragments involves a net gain of two hydrogens i.e. a hydrogen gained from ionisation

in addition to the hydrogen transferred to the charged fragment upon cleavage. The formation of singly charged ions from a, c, x, and z type backbone cleavages results in the transfer of one hydrogen from the charged to the neutral fragment. The formation of singly charged anions as a result of these cleavages is therefore accompanied by a net loss of two hydrogens since one hydrogen is transferred from the charged to neutral fragment upon cleavage of the backbone in addition to the hydrogen initially lost upon ionisation. Singly charged cations formed from these fragments will experience no net hydrogen gain since the hydrogen gained from ionisation will be balanced by the hydrogen transferred upon cleavage from the charged to the uncharged fragment. A summary of the net hydrogen transfer associated with the generation of singly and multiply charged sequence ions resulting from each cleavage type is given in table 4.1.



Scheme 4.1: Nomenclature employed for the 3'- and 5'-terminal sequence ions arising from cleavages along the phosphodiester backbone analogous to that proposed by McLuckey et al.¹ with the variation that the oligonucleotide is represented as a neutral species.

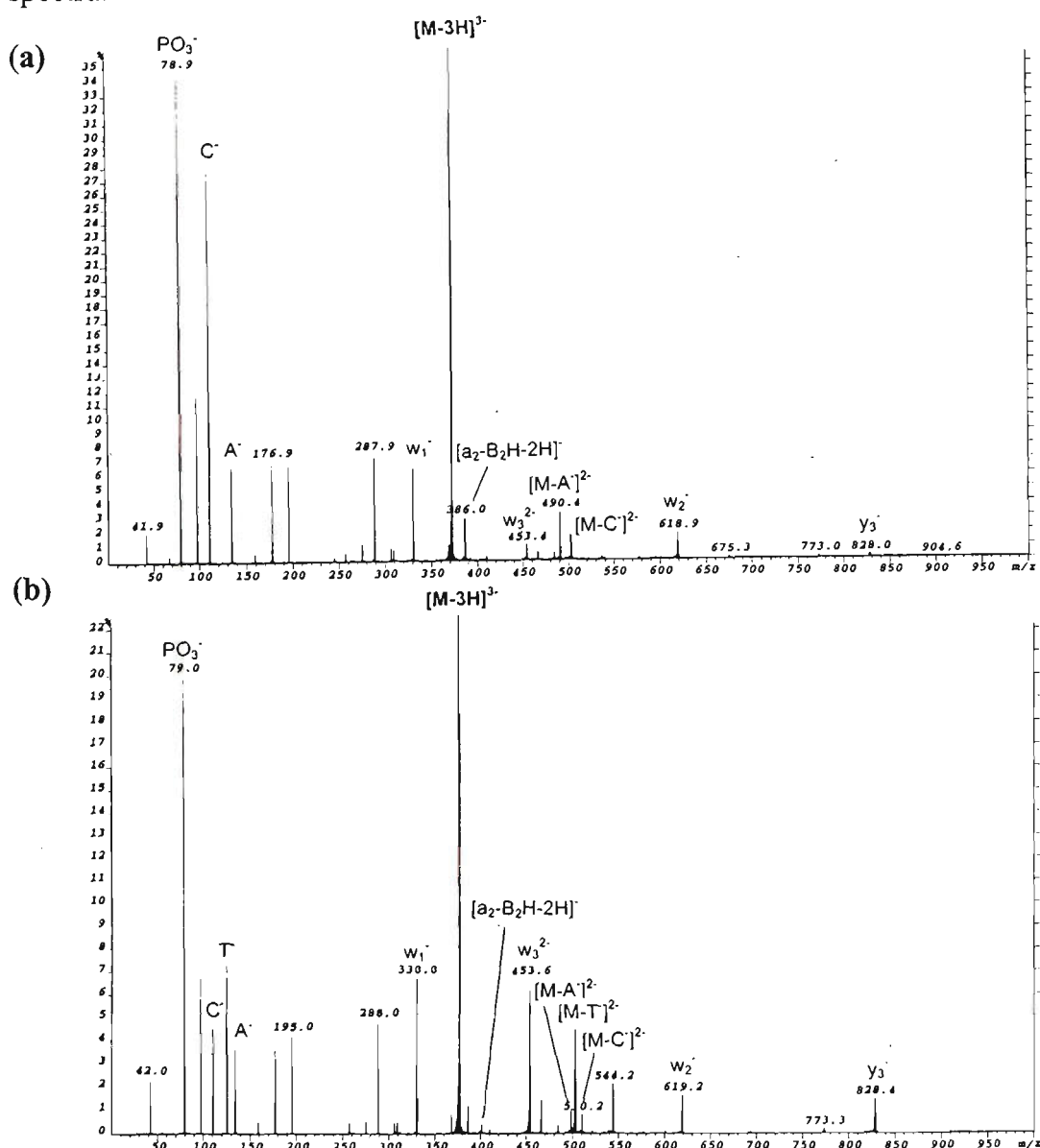
Table 4.1: Net hydrogen transfer associated with the formation of singly and multiply charged sequence ions.

Cleavage type	Singly charged anions	Multiply charged anions*	Singly charged cations	Multiply charged cations*
a	$[a-2H]^-$	$[\{a-(n+1)H\}/n]^{n-}$	a^+	$[\{a+(n-1)H\}/n]^{n+}$
b	b^-	$[\{b-(n-1)H\}/n]^{n-}$	$[b+2H]^+$	$[\{b+(n+1)H\}/n]^{n+}$
c	$[c-2H]^-$	$[\{c-(n+1)H\}/n]^{n-}$	c^+	$[\{c+(n-1)H\}/n]^{n+}$
d	d^-	$[\{d-(n-1)H\}/n]^{n-}$	$[d+2H]^+$	$[\{d+(n+1)H\}/n]^{n+}$
w	w^-	$[\{w-(n-1)H\}/n]^{n-}$	$[w+2H]^+$	$[\{w+(n+1)H\}/n]^{n+}$
x	$[x-2H]^-$	$[\{x-(n+1)H\}/n]^{n-}$	x^+	$[\{x+(n-1)H\}/n]^{n+}$
y	y^-	$[\{y-(n-1)H\}/n]^{n-}$	$[y+2H]^+$	$[\{y+(n+1)H\}/n]^{n+}$
z	$[z-2H]^-$	$[\{z-(n+1)H\}/n]^{n-}$	z^+	$[\{z+(n-1)H\}/n]^{n+}$

*where n is the number of charges

4.3 COMPARISON OF HYBRID-TOF AND ION TRAP TANDEM MASS SPECTRA OF OLIGONUCLEOTIDE ANIONS

In MS/MS studies conducted on the gentle collision conditions of ion trap instrumentation, the fragmentation of oligonucleotides is found to proceed almost entirely via nucleobase loss followed by cleavage of the 3' C-O bond of the adjacent ribose residues^{1,2,12-15}. A preliminary investigation was conducted to compare MS/MS spectra of oligonucleotide anions obtained on a hybrid mag-tof with those obtained on a quadrupole ion trap. For this study, the MS/MS spectra of the $[M-3H]^{3-}$ ions of the tetramers 5'-d(CCCA)-3', 5'-d(TCCA)-3', 5'-d(ACCT)-3' and 5'-d(CCCC)-3' were acquired and compared to the ion trap MS/MS spectra of these sequences reported by McLuckey *et al.*². The MS/MS spectra of these oligonucleotides, shown in figure 4.1, yield a number of significant differences compared with published ion trap MS/MS spectra.



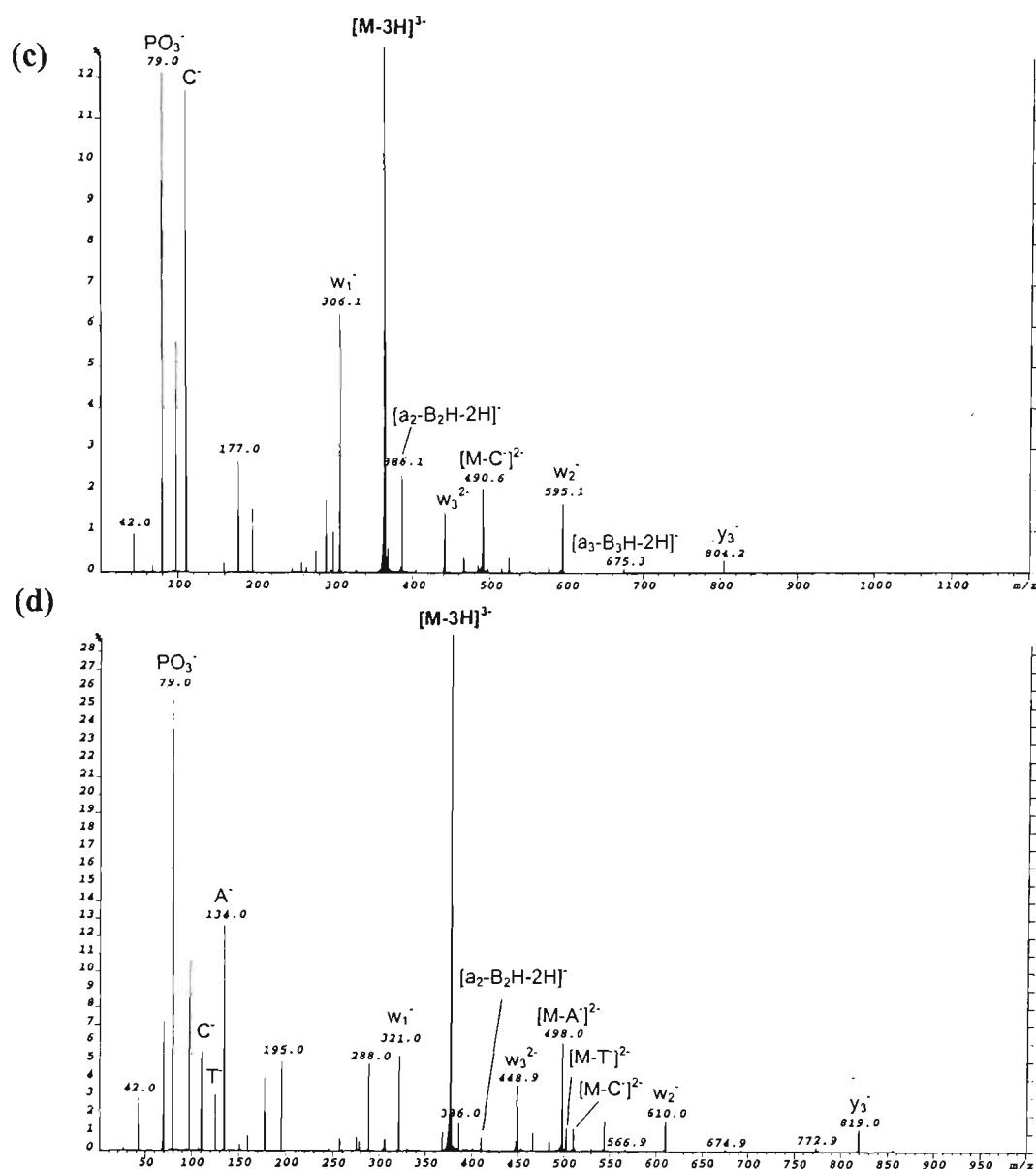


Figure 4.1: ESI-MS/MS spectra of the $[M-3H]^{3-}$ ions of the oligonucleotides (a) 5'-d(CCCA)-3'; (b) 5'-d(TCCA)-3'; (c) 5'-d(CCCC)-3'; (d) 5'-d(ACCT)-3'.

The MS/MS spectra of 5'-d(CCCA)-3', 5'-d(ACCT)-3', 5'-d(TCCA)-3' obtained on ion trap instrumentation were similar with the dominant fragmentation pathway involving the loss of a charged nucleobase. In the MS/MS spectra of 5'-d(CCCA)-3' and 5'-d(ACCT)-3', the ion arising from loss of A⁻ was the most intense (100% relative intensity). The ion arising from loss of T⁻ was the most abundant ion in the MS/MS spectrum of 5'-d(TCCA)-3'. In each case, predominant ions were also observed corresponding to the charged nucleobase anions (A⁻ and T⁻), along with w sequence ions formed by cleavage of the ribose C-O bond to the 3' side of base loss. In contrast, the corresponding MS/MS spectra of these oligonucleotides obtained with a mag-tof instrument show a more extensive array of product ions from backbone fragmentation. Moreover, there is a marked difference in the relative

intensities of the major product ions generated and in various trends for propensity of ion formation. The MS/MS spectra on ion trap instrumentation shows no ions at m/z 79 for PO_3^- , whereas this ion was the major product ion in the MS/MS spectrum of each oligonucleotide on mag-tof instrumentation, with the next more abundant product ions being the charged nucleobases (A^- and T^- etc.). In contrast to the trends observed with ion trap instrumentation which showed facile loss of charged adenine for both 5'-d(CCCA)-3' and 5'-d(ACCT)-3', the most abundant ions resulting from loss of nucleobase anions in the mag-tof MS/MS spectra were those arising from loss of the nucleobase at the 5'-terminus. The oligonucleotide 5'-d(CCCC)-3' is an example of a case where the fragmentation would not be driven by the effects of base identity, and the mag-tof MS/MS spectrum obtained for this oligonucleotide bore the most similarity to that observed on an ion trap. Incidentally, even though the ion trap MS/MS spectrum shows no marked distinctions arising from base loss propensity, differences in abundances of the various sequence ions are observed. The mag-tof MS/MS spectrum of 5'-d(CCCC)-3' shows the w_1^- ion as the most abundance sequence ion present at m/z 306.1, followed by $[\text{a}_2\text{-B}_2\text{H-2H}]^-$ (m/z 386.1), w_2^- (m/z 595.1), $[\text{w}_3\text{-H}]^{2-}$ (m/z 441.6), $[\text{w}_2\text{-H}]^{2-}$ (m/z 297.1), $[\text{a}_3\text{-B}_3\text{H-2H}]^-$ (m/z 675.3) in decreasing order. In contrast, the most intense ion generated by the MS/MS spectrum of the $[\text{M-3H}]^{3-}$ ion of 5'-d(CCCC)-3' acquired on an ion trap is the w_2^- ion followed in order by $[\text{a}_2\text{-B}_2\text{H-2H}]^-$, $[\text{w}_3\text{-H}]^{2-}$, w_1^- , $[\text{a}_3\text{-B}_3\text{H-2H}]^-$ and $[\text{w}_2\text{-H}]^{2-}$. The ion trap MS/MS spectrum also shows the ion for C^- to be of relatively low abundance, whereas the C^- ion is the second most intense ion observed in the mag-tof MS/MS spectrum.

The differences observed between ion trap and mag-tof MS/MS spectra in each case are probably a consequence of both the higher collision energy and/or differences in the time scale for dissociation on the mag-tof instrument. MS/MS studies of oligonucleotides using triple quadrupole instrumentation have likewise produced different results to that of ion traps owing to the comparatively harsher collision conditions associated with this instrument type. In the mag-tof MS/MS spectra of oligonucleotide anions, the most favourable location for a base to be lost as a

charged species is at the 5'-terminus. Under harsh collision conditions, the cleavage of all bases would be expected to occur so that the fact that the most intense base anion in each case is that of the 5'-terminal base can be explained on the basis that charged base loss is preferable from this position. In contrast, the gentle collision conditions of ion trap instrumentation yield only the lowest energy fragmentation pathways, so that the ions for charged base loss observed are representative of the most facile bond cleavages that can occur within the oligonucleotide anion. Moreover, the most intense sequence ions observed in the ion trap MS/MS of these oligonucleotides are those generated from the subsequent ribose C-O bond cleavage to the 3' side of the sites of preferential charged nucleobase loss. In the MS/MS spectra on a mag-tof, the w_1^- ion is the most intense w cleavage observed in each case, rather than the $[w_3-H]^{2-}$ ion which would arise from subsequent ribose 3' C-O bond cleavage adjacent to the 5'-terminal base. The higher abundance of the w_1^- ion may arise from multiple collisions of larger w ions.

In order to examine whether the collision conditions on mag-tof instrumentation could generate fragmentation similar to that observed with an ion trap, the MS/MS spectrum of the $[M-3H]^{3-}$ ion of the oligonucleotide 5'-d(TCCA)-3' was acquired with a collision gas cell pressure as low as possible so as to generate a minimal degree of fragmentation. This MS/MS spectrum, which is shown in figure 4.2, produces a substantially different fragmentation pattern showing only a small number of product ions compared with that observed under typical MS/MS conditions upon which the precursor ion was reduced by 30% of its incident intensity with the introduction of collision gas (figure 4.1b). The MS/MS spectrum obtained under low pressure conditions (figure 4.2) more closely resembles that reported for the ion trap² with the loss of charged T⁻ generating the most intense ion at m/z 502.6. As observed in the ion trap MS/MS spectrum, ions for the loss of A⁻ and C⁻ are also present in very similar abundances relative to the ion for loss of T⁻. In addition, the ion for T⁻ is the most intense base anion observed in the MS/MS spectra acquired on both instruments. Further consistent with the observations in the ion trap MS/MS spectrum, is the

observation of a predominant $[w_3-H]^{2-}$ ion at m/z 453.5. Although the MS/MS spectrum of 5'-d(TCCA)-3' acquired at low collision gas pressure shows a very similar type of fragmentation to that observed with ion trap instrumentation, using the collision conditions gentle enough to observe this fragmentation pattern, however, yields a significantly lower relative abundance of product ions which are less than 1% of the precursor ion. In contrast, the MS/MS spectrum reported on ion trap instrumentation yields apparently much more highly efficient fragmentation, which results in the product ion for charged base loss at a relative intensity of 100%.

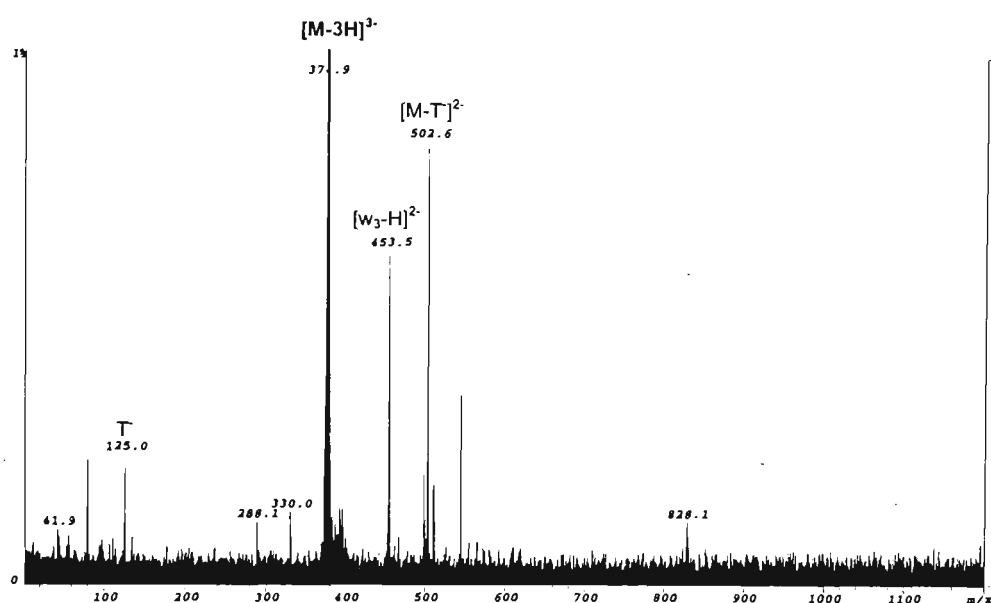


Figure 4.2: ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the oligonucleotide 5'-d(TCCA)-3' acquired with very low collision cell gas pressure.

4.4 CHARACTERISATION OF NEGATIVE ION ELECTROSPRAY TANDEM MASS SPECTRA OF OLIGONUCLEOTIDES

For instrumentation employing harsher collision conditions such as that used for the present work, the fragmentation behaviour of oligonucleotides is significantly more complex owing to the occurrence of multiple and higher energy collisions with no effects of collisional cooling by the bath gas. In order to gain an understanding of the MS/MS fragmentation generated on hybrid mag-tof instruments, a detailed characterisation of the types of product ions generated in MS/MS spectra was undertaken at the outset of this study. The oligonucleotides selected for this

examination were the two hexanucleotides, 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. There were three reasons for the choice of these hexamers. First and foremost, these oligonucleotides constituted the two major sequences used in the work conducted on the sequence analysis of ligand binding to DNA, and therefore a detailed characterisation of the MS/MS fragmentation of the unmodified oligonucleotides was an important preliminary stage for this aspect of the work. Secondly, each oligonucleotide incorporates all four of the major bases thus providing a suitable system for the examination of the effect of base identity on the product ion types observed. Thirdly, owing to the small size of these oligonucleotides, the extent of fragmentation generated in their MS/MS spectra makes the task of detailed product ion analysis a reasonable one. In this study, the MS/MS spectra of singly charged oligonucleotide precursor ions were examined since this provides the most facile system for product ion analysis with only singly charged species generated in resulting MS/MS spectra. This avoids the problems associated with determining product ion charge states frequently encountered with analysis of MS/MS spectra of multiply charged precursor ions.

4.4.1 Product Ion Types Observed

Figures 4.3 and 4.4 show the MS/MS spectra of the $[M-H]^-$ ions of the oligonucleotides 5'-d(CACGTG)-3' (m/z 1790.3) and 5'-d(CGTACG)-3' (m/z 1790.3) respectively. The MS/MS spectra of these oligonucleotides show extensive backbone fragmentation, which yields over seventy identifiable product ions in each case. For the task of characterising the product ion types generated, the fragmentation of individual product ions were analysed whereby each product ion of interest was first generated in MS-1 by increasing the potential on the sampling cone and skimmer lens in the ESI interface (as described in chapter 3, section 3.3.1), and then selected as the precursor ion in MS-2 where they were collisionally activated and their MS/MS spectrum recorded. The assignments of the product ions appearing in the MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' determined from multiple stage mass spectrometry analysis are given in tables 4.2 and 4.3.

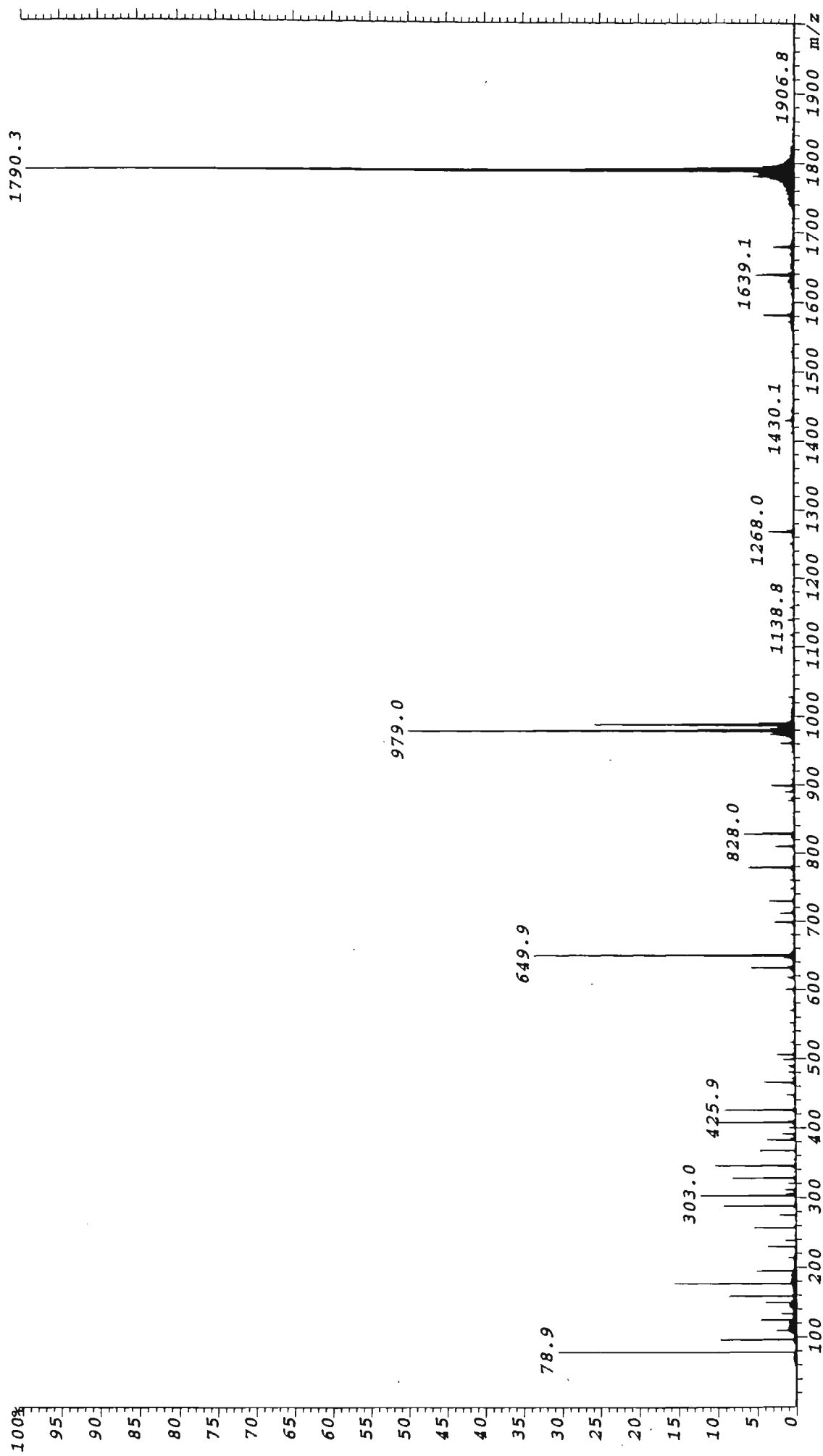


Figure 4.3: Negative Ion ESI-MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CACGTG)-3' (m/z 1790.3).

Table 4.2 : Assignment of product ions observed in the negative ion ESI-MS/MS spectrum of the oligonucleotide 5'-d(CACGTG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1790.3	100	1790.3	[M-H]⁻
1679.3	2.8	1679.3	loss of Cytosine
1639.1	5.2	1639.3	loss of Guanine
1581.1	4	1581.3	w ₅ ⁻
1430.1	1.2	1430.2	[w ₅ -GH] ⁻ and/or [(ACGT)+s+p+H ₂ O-H] ⁻
1268.0	3.5	1268.2	w ₄ ⁻
1250.1	0.5	1250.2	[x ₄ -2H] ⁻
1138.8	0.3	1139.1	[a ₄ -2H] ⁻
1117.2	0.2	1117.2	[w ₄ -GH] ⁻ and/or [(CGT)+s+p+H ₂ O-H] ⁻
1099.1	0.1	1099.1	[x ₄ -GH-2H] ⁻ and/or [(CGT)+s+p-H] ⁻
1027.8	0.3	1028.2	[a ₄ -CH-2H] ⁻ and/or [(ACG)+s-H] ⁻
		or 1028.1	[(ACG)+p+H ₂ O-H] ⁻
1019.1	0.1	1019.1	[z ₄ -GH-2H] ⁻ and/or [(CGT)+s-H] ⁻
		or 1019.2	[(CGT)+p+H ₂ O-H] ⁻
988.0	26.0	988.2	[a ₄ -B ₄ H-2H] ⁻
979.0	50.5	979.2	w ₃ ⁻
961.1	2.0	961.1	[x ₃ -2H] ⁻
930.1	0.1	930.1	[(ACG)-H] ⁻
921.1	0.1	921.1	[(CGT)-H] ⁻
908.2	0.1	908.2	[b ₄ -B ₄ H] ⁻
899.2	3.0	899.2	y ₃ ⁻
890.2	0.5	890.1	[c ₃ -2H] ⁻
881.2	0.2	881.2	[z ₃ -2H] ⁻
877.1	0.3	877.1	[a ₄ -GH-CH-2H] ⁻ and/or [(AC)+s+p+s-H] ⁻
		or 877.2	[(AC)+p+s+p+H ₂ O-H] ⁻
828.1	6.4	828.1	[w ₃ -GH] ⁻ and/or [(GT)+s+p+H ₂ O-H] ⁻
810.0	2.3	810.2	[a ₃ -2H] ⁻
		or 810.1	[x ₃ -GH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
795.0	0.1	795.1	[(CG)+s+p-H] ⁻
778.8	6.0	779.1	[c ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s+p-H] ⁻
761.0	0.3	761.1	[(CA and/or AC)+s+p-H ₂ O-H] ⁻
748.0	0.3	748.1	[y ₃ -GH] ⁻ and/or [(GT)+s+H ₂ O-H] ⁻
729.9	3.3	730.1	[z ₃ -GH-2H] ⁻ and/or [(GT)+s-H] ⁻
		or 730.2	[(GT)+p+H ₂ O-H] ⁻
712.0	2.0	712.1	[(GT)+s-H ₂ O-H] ⁻
		or 712.1	[(GT)+p-H] ⁻
699.1	2.3	699.1	[a ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s-H] ⁻
		or 699.1	[(AC)+p+H ₂ O-H] ⁻
681.1	0.4	681.1	[(AC)+p-H] ⁻
		or 681.1	[(AC)+s-H ₂ O-H] ⁻
649.9	33.9	650.1	w ₂ ⁻ and/or [(GT)+H ₂ O-H] ⁻
631.9	5.8	632.1	[x ₂ -2H] ⁻ and/or [(GT)-H] ⁻
619.0	1.0	619.1	d ₂ ⁻ and/or [(AC)+H ₂ O-H] ⁻
617.0	0.7	617.1	[(CG)-H] ⁻
601.0	1.3	601.1	[c ₂ -2H] ⁻ and/or [(AC)-H] ⁻
570.0	0.8	570.1	y ₂ ⁻ and/or [(GT)-H ₂ O-H] ⁻
552.0	0.8	552.1	[z ₂ -2H] ⁻ and/or [(GT)-p-H] ⁻
539.0	0.4	539.1	b ₂ ⁻ and/or [(AC)-p+H ₂ O-H] ⁻
524.1	0.5	524.1	[w ₂ -TH-2H] ⁻ and/or [G _{nt} +s+p+H ₂ O-H] ⁻
506.0	2.5	506.0	[x ₂ -TH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
499.0	1.5	499.0	[w ₂ -GH-2H] ⁻ and/or [T _{nt} +s+p+H ₂ O-H] ⁻
490.0	1.1	490.1	[c ₂ -CH-2H] ⁻ and/or [A _{nt} +s+p-H] ⁻
481.0	1.0	481.0	[x ₂ -GH-2H] ⁻ and/or [T _{nt} +s+p-H] ⁻
465.9	4.3	466.0	[c ₂ -AH-2H] ⁻ and/or [C _{nt} +s+p-H] ⁻
448.1	1.5	448.0	[C _{nt} +s+p-H ₂ O-H] ⁻
426.0	9.4	426.1	[z ₂ -TH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
		or 426.1	[G _{nt} +p+H ₂ O-H] ⁻
410.0	1.3	410.1	[a ₂ -CH-2H] ⁻ and/or [A _{nt} +s-H] ⁻

408.0	10.2	410.0	[A _{nt} +p+H ₂ O-H] ⁻
		408.1	[G _{nt} +s-H ₂ O-H] ⁻
		or 408.0	[G _{nt} +p-H] ⁻
401.0	1.0	401.1	[z ₂ -GH-2H] ⁻ and/or [T _{nt} +s-H] ⁻
		or 401.0	[T _{nt} +p+H ₂ O-H] ⁻
392.1	1.9	392.1	[A _{nt} +s-H ₂ O-H] ⁻
		or 392.0	[A _{nt} +p-H] ⁻
386.0	0.7	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
		or 386.0	[C _{nt} +p+H ₂ O-H] ⁻
382.9	3.9	383.1	[T _{nt} +s-H ₂ O-H] ⁻
		or 383.0	[T _{nt} +p-H] ⁻
368.0	4.6	368.1	[C _{nt} +s-H ₂ O-H] ⁻
		or 368.0	[C _{nt} +p-H] ⁻
346.0	10.5	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.0	8.4	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
321.0	1.3	321.0	[T _{nt} +H ₂ O-H] ⁻
312.0	1.5	312.0	[A _{nt} -H] ⁻
306.0	1.4	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.0	12.5	303.0	[T _{nt} -H] ⁻
288.0	9.3	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.0	2.2	275.1	[sps-H] ⁻
		or 275.0	[psp+H ₂ O-H] ⁻
256.9	5.5	257.1	[sps-H ₂ O-H] ⁻
		or 257.0	[psp-H] ⁻
239.0	1.6	239.0	[psp-H ₂ O-H] ⁻
229.9	3.9	230.1	[p+G-H] ⁻
213.9	1.2	214.0	[p+A-H] ⁻
195.0	5.3	195.0	[ps+H ₂ O-H] ⁻
176.9	15.8	177.0	[ps-H] ⁻
158.9	8.9	159.0	[ps-H ₂ O-H] ⁻
150.0	4.2	150.0	G ⁻
134.0	2	134.0	A ⁻
124.9	4.5	125.0	T ⁻
110.0	2.8	110.0	C ⁻
96.9	9.9	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
		or 97.0	[p+H ₂ O-H] ⁻
78.9	30.8	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

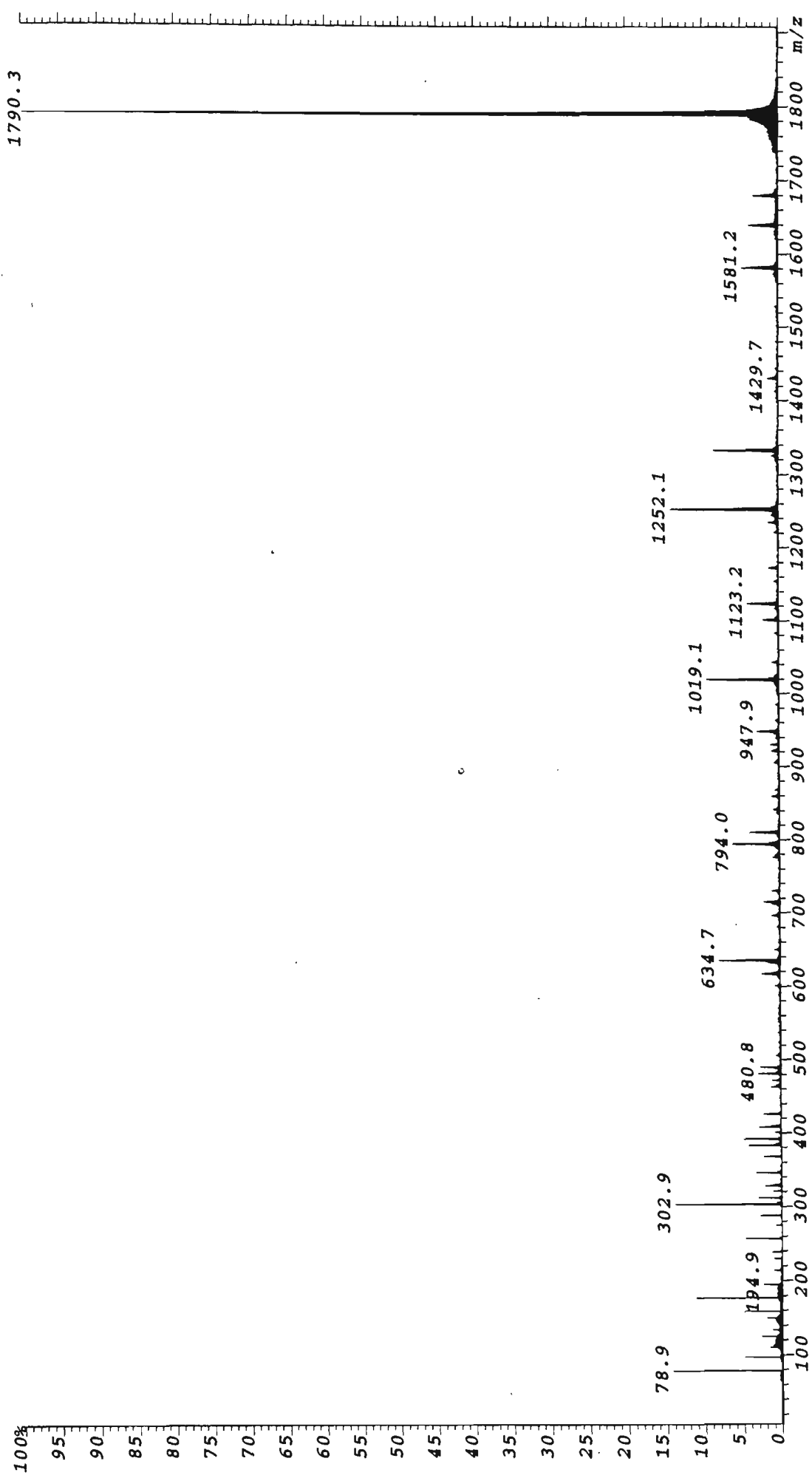


Figure 4.4: Negative ion ESI-MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CGTACG)-3' (m/z 1790.3).

Table 4.3 : Assignment of product ions observed in the negative ion ESI-MS/MS spectrum of the oligonucleotide 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1790.3	100	1790.3	[M-H]⁻
1679.4	3.3	1679.3	loss of Cytosine
1639.3	3.9	1639.3	loss of Guanine
1581.2	4.9	1581.3	w ₅ ⁻ and/or d ₅ ⁻
1429.7	1.5	1430.2	[w ₅ -GH] ⁻ and/or [d ₅ -GH] ⁻
1332.3	8.6	1332.3	[a ₅ -B ₅ H-2H] ⁻ and/or [z ₅ -GH-2H] ⁻
		1332.2	[(GTAC)+p+H ₂ O-H] ⁻
1252.1	14.4	1252.2	w ₄ ⁻ and/or d ₄ ⁻ and/or [(GTAC)+H ₂ O-H] ⁻
1234.1	1.5	1234.2	[x ₄ -2H] ⁻ and/or [c ₄ -2H] ⁻ and/or [(GTAC)-H] ⁻
1172.1	1.4	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GTAC)-p+H ₂ O-H] ⁻
1154.4	0.7	1154.2	[a ₄ -2H] ⁻ and/or [z ₄ -2H] ⁻ and/or [(GTAC)-p-H] ⁻
1123.2	4.2	1123.1	[x ₄ -CH-2H] ⁻ and/or [c ₄ -CH-2H] ⁻ and/or [(GTA)+s+p-H] ⁻
1101.0	2.1	1101.1	[w ₄ -GH] ⁻ and/or [d ₄ -GH] ⁻ and/or [(TAC)+s+p+H ₂ O-H] ⁻
1083.3	0.7	1083.1	[x ₄ -GH-2H] ⁻ and/or [c ₄ -GH-2H] ⁻ and/or [(TAC)+s+p-H] ⁻
1043.1	0.9	1043.2	[a ₄ -CH-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(GTA)+s-H] ⁻
		1043.1	[(GTA)+p+H ₂ O-H] ⁻
1019.1	9.6	1019.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -AH-2H] ⁻ and/or [(GTAC)-AH-p-H] ⁻
985.2	0.5	985.2	[(TAC)+s-H ₂ O-H] ⁻
		985.1	[(TAC)+p-H] ⁻
947.9	2.9	948.2	w ₃ ⁻
930.0	1.2	930.1	[x ₃ -2H] ⁻
921.0	1.1	921.0	[c ₃ -2H] ⁻
905.1	0.8	905.1	[(TAC)-H] ⁻
868.1	0.7	868.2	y ₃ ⁻
859.1	1.1	859.2	b ₃ ⁻
841.3	0.8	841.2	[a ₃ -2H] ⁻
810.2	3.9	810.1	[c ₃ -CH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
794.1	6.2	794.1	[(TA)+s+p-H] ⁻
776.1	1.0	776.1	[(TA)+s+p-H ₂ O-H] ⁻
730.2	1.1	730.1	[(GT)+s-H] ⁻
		730.0	[(GT)+p+H ₂ O-H] ⁻
715.0	2.2	715.0	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -AH-2H] ⁻
714.2	1.2	714.1	[(TA)+s-H] ⁻
		714.1	[(TA)+p+H ₂ O-H] ⁻
696.1	1.2	696.1	[(TA)+s-H ₂ O-H] ⁻
		696.1	[(TA)+p-H] ⁻
680.9	0.4	681.1	[(AC)+s-H ₂ O-H] ⁻
		681.1	[(AC)+p-H] ⁻
650.2	0.9	650.1	[(GT)+H ₂ O-H] ⁻
634.7	8.2	635.1	w ₂ ⁻ and/or d ₂ ⁻
617.3	2.7	617.1	[x ₂ -2H] ⁻ and/or [c ₂ -2H] ⁻
616.3	1.5	616.1	[(TA)-H] ⁻
601.1	0.8	601.1	[(AC)-H] ⁻
506.0	0.7	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
490.0	2.8	490.1	[A _{nt} +s+p-H] ⁻
480.8	3.0	481.0	[T _{nt} +s+p-H] ⁻
472.0	1.2	472.0	[A _{nt} +s+p-H ₂ O-H] ⁻
465.9	0.8	466.0	[x ₂ -GH-2H] ⁻ and/or [c ₂ -GH-2H] ⁻ and/or [C _{nt} +s+p-H] ⁻
462.9	1.3	463.0	[T _{nt} +s+p-H ₂ O-H] ⁻
426.0	2.5	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻

		or 426.1	[G _{nt} +p+H ₂ O-H] ⁻
410.0	1.5	410.1	[A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	2.9	408.1	[G _{nt} +s-H ₂ O-H] ⁻
		or 408.0	[G _{nt} +p-H] ⁻
401.1	0.9	401.1	[T _{nt} +s-H] ⁻
		or 401.0	[T _{nt} +p+H ₂ O-H] ⁻
392.1	4.9	392.1	[A _{nt} +s-H ₂ O-H] ⁻
		or 392.0	[A _{nt} +p-H] ⁻
386.0	0.8	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
		or 386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.1	4.4	383.1	[T _{nt} +s-H ₂ O-H] ⁻
		or 383.0	[T _{nt} +p-H] ⁻
368.1	2.5	368.1	[C _{nt} +s-H ₂ O-H] ⁻
		or 368.0	[C _{nt} +p-H] ⁻
346.1	3.5	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
330.1	1.2	330.2	[A _{nt} +H ₂ O-H] ⁻
328.0	2.2	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
321.0	1.2	321.0	[T _{nt} +H ₂ O-H] ⁻
312.0	3.1	312.0	[A _{nt} -H] ⁻
302.9	14.1	303.0	[T _{nt} -H] ⁻
288.0	2.8	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.0	0.8	275.1	[sps-H] ⁻
		or 275.0	[psp+H ₂ O-H] ⁻
256.9	4.8	257.1	[sps-H ₂ O-H] ⁻
		or 257.0	[psp-H] ⁻
239.0	1.3	239.0	[psp-H ₂ O-H] ⁻
229.9	1.1	230.1	[p+G-H] ⁻
214.0	1.1	214.0	[p+A-H] ⁻
194.9	2.6	195.0	[ps+H ₂ O-H] ⁻
176.9	11.3	177.0	[ps-H] ⁻
158.9	5.1	159.0	[ps-H ₂ O-H] ⁻
150.0	2.2	150.0	G ⁻
134.0	1.3	134.0	A ⁻
124.9	2.8	125.0	T ⁻
110.0	1.7	110.0	C ⁻
97.0	5.1	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
		or 97.0	[p+H ₂ O-H] ⁻
78.9	14.5	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denoted a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

(a) **Backbone and Nucleotide Fragments**

The MS/MS spectra of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are characterised by a variety of product ions in the low *m/z* region of their spectra (i.e. below *m/z* 530). These ions, which comprise backbone fragments, monomeric oligonucleotide constituents and mononucleotide ion types are summarised in table 4.4. In the MS/MS spectrum of each oligonucleotide, the most intense ion in the low *m/z* region arises from phosphate (PO_3^-) at *m/z* 78.9. The MS/MS spectrum of this ion is shown in figure 4.5 and gives rise to two main ions arising from cleavage of an oxygen atom to yield PO_2^- at *m/z* 63 and an ion for O^- at *m/z* 16.

Table 4.4: Nucleotide product ion types generated in the negative ion MS/MS spectra of oligonucleotides.

Product Ion Types of Nucleotide Fragments			
Backbone fragments*	<i>m/z</i>	Nucleoside fragments	<i>m/z</i>
PO_3^- { [p-H] $^-$ }	79	C $^-$	110
$\text{C}_5\text{H}_5\text{O}_2^-$ { [s-H] $^-$ } or [p+H ₂ O-H] $^-$	97	T $^-$	125
[p+s-H ₂ O-H] $^-$	159	A $^-$	134
[p+s-H] $^-$	177	G $^-$	150
[p+s+H ₂ O-H] $^-$	195	[p+A-H] $^-$	214
[p+s+p-H ₂ O-H] $^-$	239	[p+G-H] $^-$	230
[p+s+p-H] $^-$ or [s+p+s-H ₂ O-H] $^-$	257		
[p+s+p+H ₂ O-H] $^-$ or [s+p+s-H] $^-$	276		

*where p= PO_3H , s=deoxyribose-H₂O ($\text{C}_5\text{H}_6\text{O}_2$)

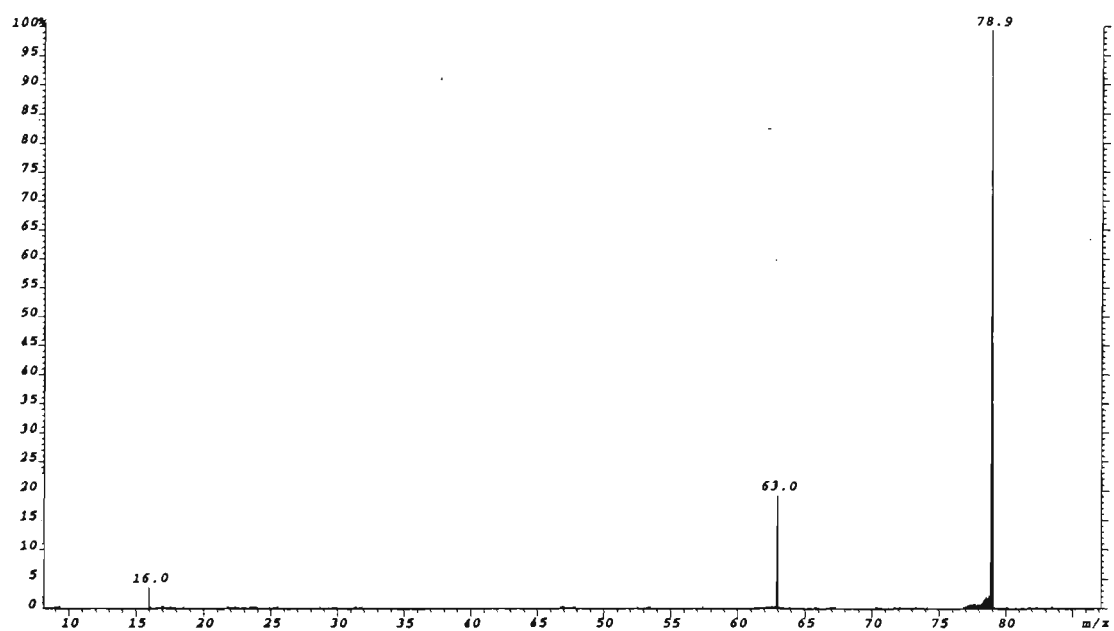
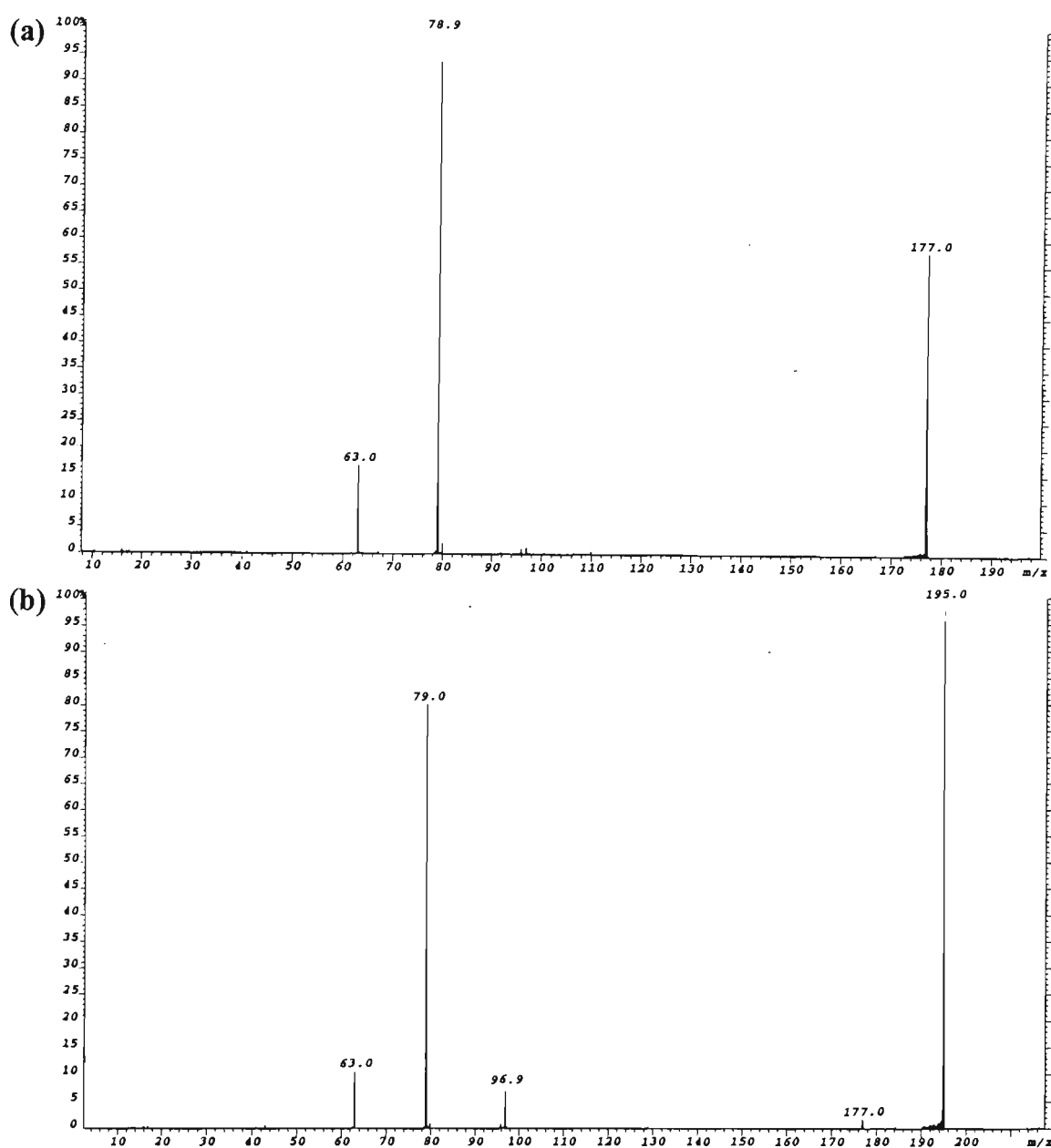


Figure 4.5: ESI-MS/MS spectrum of the ion at *m/z* 78.9 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Figure 4.6 shows the MS/MS spectra of the ions at m/z 177.0, 195.0, 239.0, 256.9 and 275.0 in the MS/MS spectrum of 5'-d(CACGTG)-3'[#]. Fragmentation of these ions yields product ions consistent with the dissociation of backbone fragments comprising sugar and phosphate groups. The MS/MS spectrum of each of these ions gives rise to an intense ion for PO_3^- at m/z 79 and a weak ion for PO_2^- at m/z 63. In most cases, an ion at m/z 97 is also detected, which is either $[\text{s-H}]^-$ and/or $[\text{p+H}_2\text{O-H}]^-$ (attempts to analyse this ion by multiple stage mass spectrometry proved unsuccessful). The MS/MS spectra of the ions at m/z 256.9 and 275.0 also show the backbone fragment ions at m/z 177 for $[\text{ps-H}]^-$, and at m/z 159 for $[\text{p+s-H}_2\text{O-H}]^-$.



[#]These ions are generated in the source and owing to small differences between the calibration of the magnet and the TOF and/or between TOF spectra obtained at different times, the m/z of product ions observed in MS/MS spectra may vary between 0.1-0.5Da.

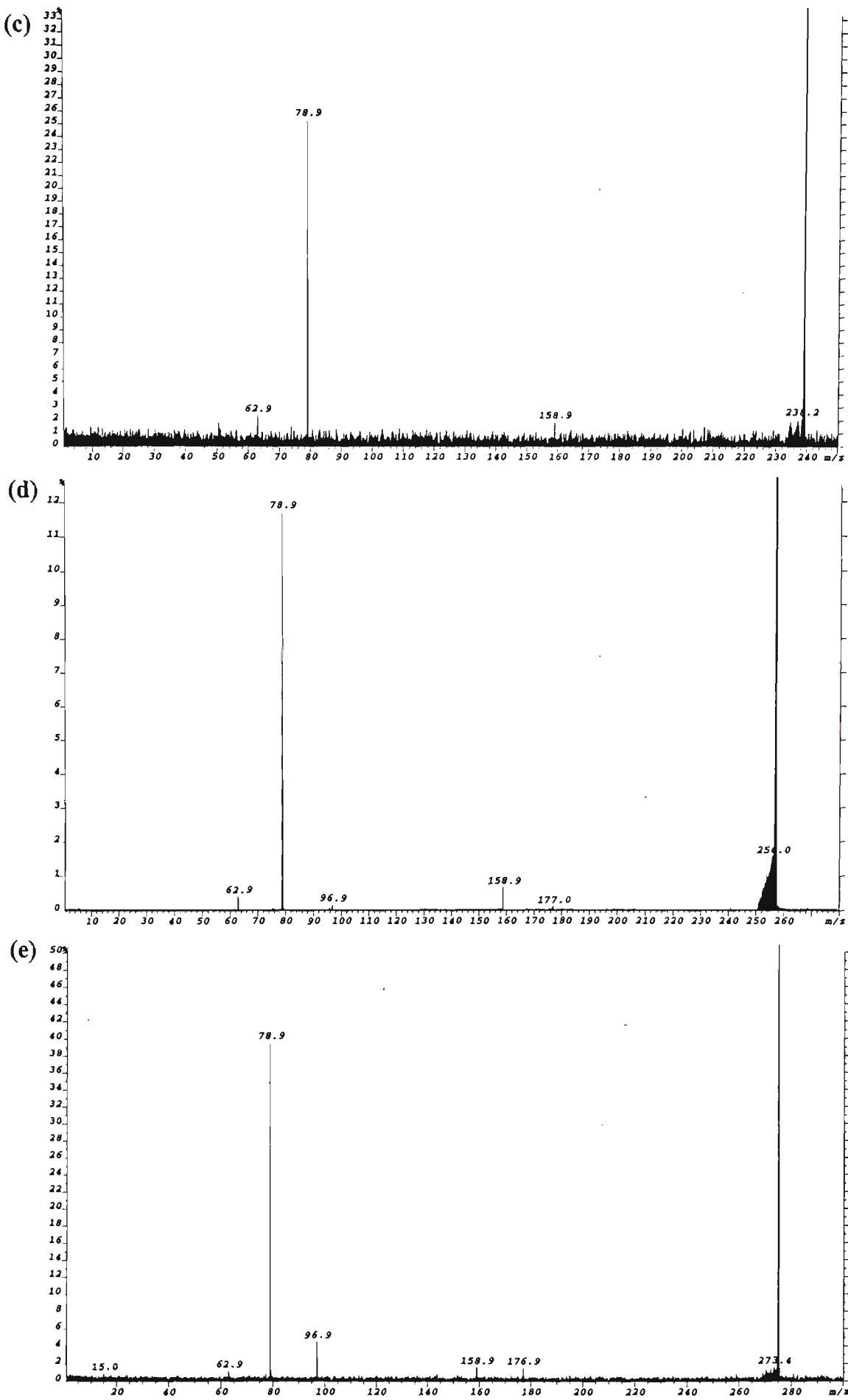


Figure 4.6: ESI-MS/MS spectra of the ions at (a) m/z 177.0; (b) m/z 195.0; (c) m/z 239.0; (d) m/z 256.9; and (e) m/z 275.0 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The ion appearing at m/z 256.9 corresponds in mass to either of the two backbone fragments $[p+s+p-H]^-$ or $[s+p+s-H_2O-H]^-$, and likewise the ion at m/z 275.0 can be attributed to either $[p+s+p+H_2O-H]^-$ or $[s+p+s-H]^-$. In each case, the two possible fragments differ by the addition of either a phosphate PO_3H group (p), or a sugar C_5H_4O group ($s-H_2O$) which have a mass difference of 0.0599Da. It was not possible to resolve for either or both of these species in the MS/MS spectra acquired of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', owing to insufficient resolution in the TOF mass analyser used for this work (maximum resolution of 500 peak width at half height). Figure 4.7 shows the MS spectrum acquired at a resolution of approximately 10000 at 5% valley peak-to-peak of the in-source generated product ion at m/z 257.

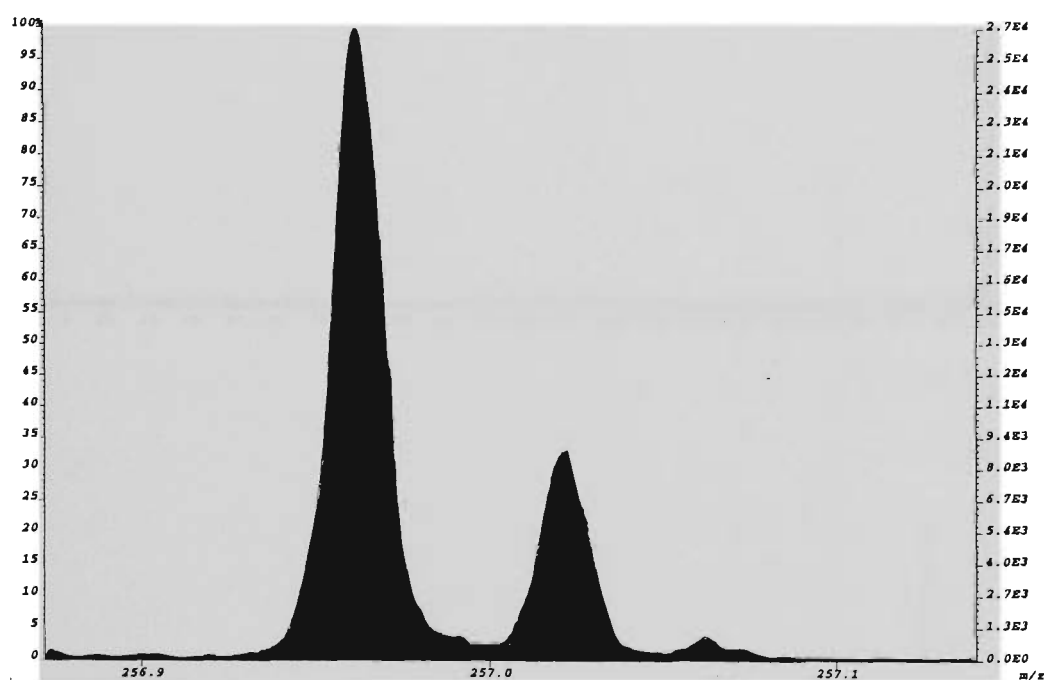
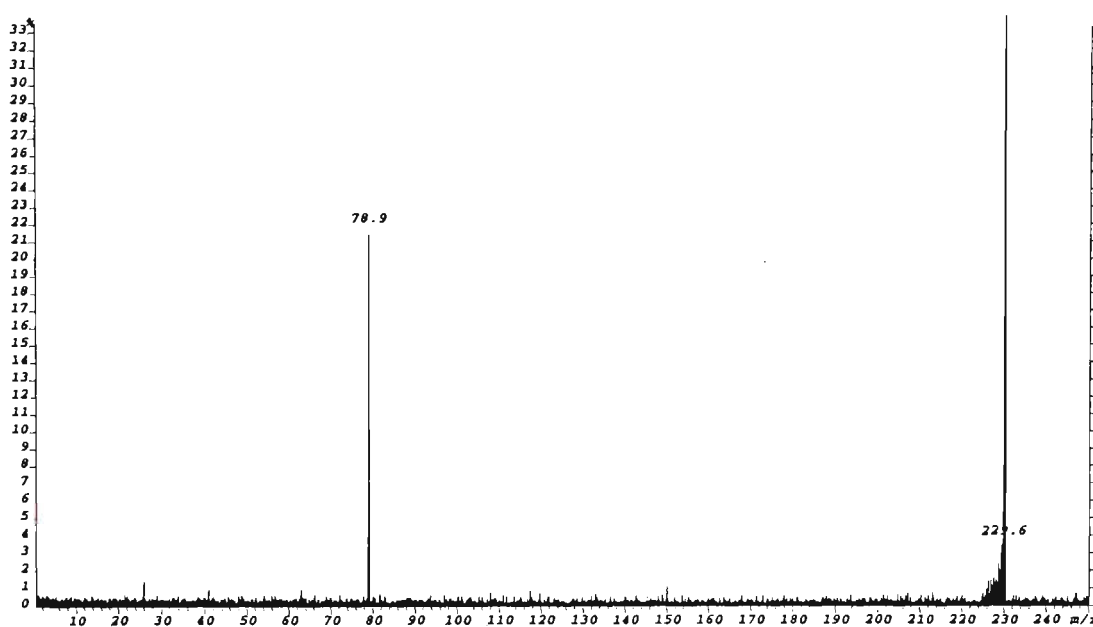


Figure 4.7: ESI-MS spectrum of the ion at m/z 257 generated by 'in-source' fragmentation of 5'-d(CACGTG)-3' acquired at a resolution of 10000 at 5% valley peak-to-peak.

In addition to the observation of these backbone fragments, the other types of nucleotide fragments observed in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' include ions for all four nucleobases (C^- at m/z 110, T^- at m/z 125, A^- at m/z 134, and G^- at m/z 150), in addition to ions at m/z 230 and 214. The MS/MS spectra of the ions at m/z 229.9 and 213.9 generated from the in-source fragmentation of 5'-d(CACGTG)-3' are shown in figure 4.8(a) and (b) respectively. In each case, fragmentation of these ions results in the predominant observation of PO_3^- at m/z 78.9,

with the MS/MS spectrum of the m/z 229.9 ion yielding an ion of very weak abundance at m/z 150.0 indicative of the guanine anion. These spectra seem to suggest that the ions at m/z 230 and 214 correspond to phosphate-nucleobase adducts of structure $[p+G-H]^-$ and $[p+A-H]^-$ respectively.

(a)



(b)

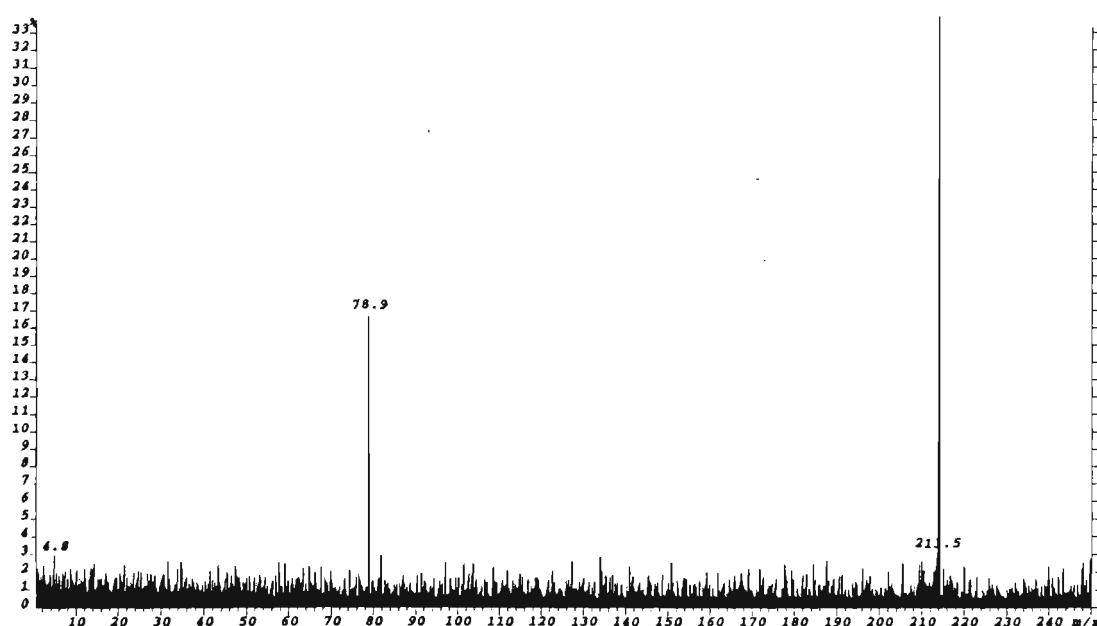


Figure 4.8: ESI-MS/MS spectra of the product ions appearing at (a) m/z 229.9; (b) m/z 213.9 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

(b) **Mononucleotide Ion Types**

The m/z 280-530 region of the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' yields a characteristic series of nucleotide ions for each of the four nucleobases incorporated in the sequence. A summary of the types of product ions observed is given in table 4.5.

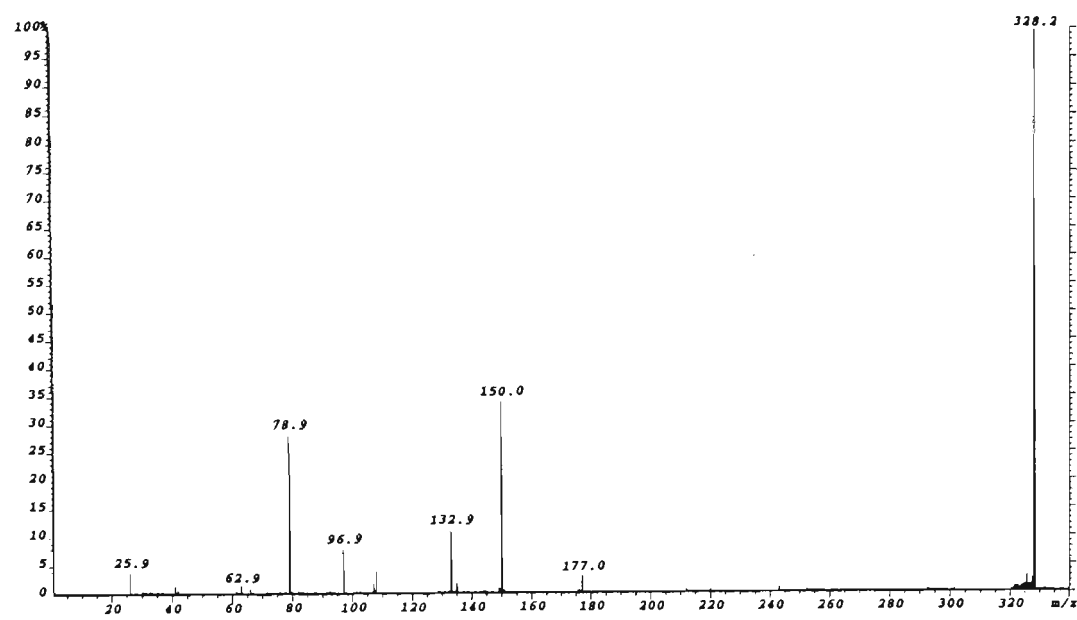
Table 4.5: Mononucleotide ion types generated in the negative ion MS/MS spectra of oligonucleotides.

Product Ion Types of Nucleotides*					
Mononucleotide ions:	X=	C	T	A	G
X_{nt}		288	303	312	328
$X_{nt}+H_2O$		306	321	-	346
$X_{nt}+p$ or $+(s-H_2O)$		368	383	392	408
$X_{nt}+s$ or $+(p+H_2O)$		386	401	410	426
$X_{nt}+p+s-H_2O$		448	-	-	-
$X_{nt}+p+s$		466	481	490	506
$X_{nt}+p+s+H_2O$		-	499	-	524

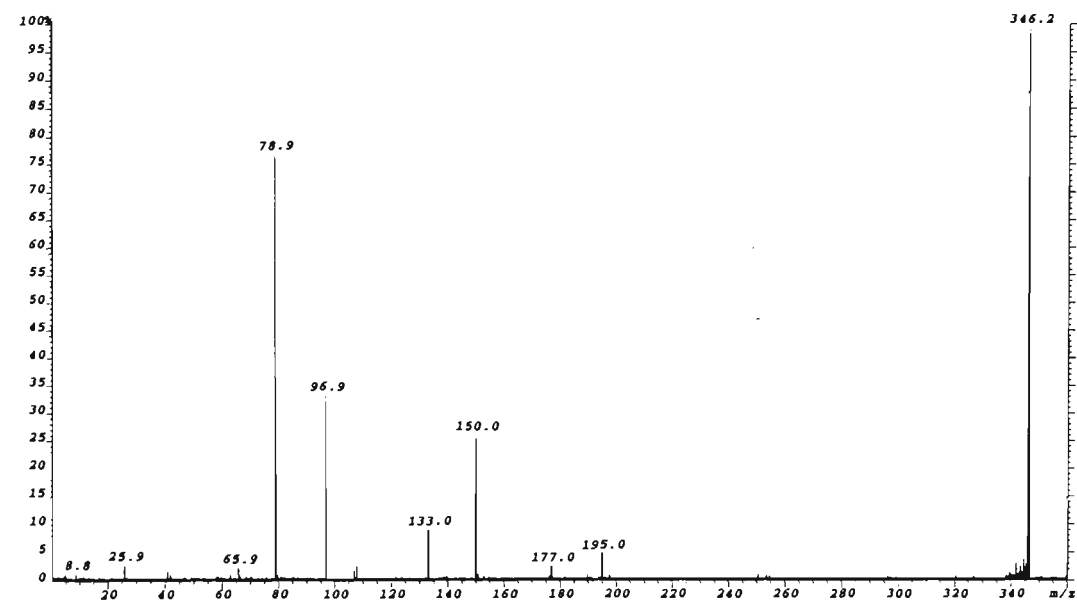
*where $p=PO_3H$, $s=deoxyribose-H_2O$ ($C_5H_6O_2$)

Figure 4.9 (a)-(f) shows the MS/MS spectra of the series of nucleotide product ion types appearing at m/z 328, 346, 408, 426, 506, and 524 generated by in-source fragmentation of 5'-d(CACGTG)-3'. Fragmentation of these mononucleotide ions yields an ion at m/z 150 indicative of deprotonated guanine in addition to a variety of the typically observed backbone fragment ions discussed above and shown in table 4.4. The MS/MS spectra of the series of nucleotide ion types observed for cytosine, adenine and thymine are given in appendix 4.1 and showed very similar fragmentation characteristics as that observed for the analogous nucleotide ion types of guanine. As observed for the backbone product ions at m/z 257, and 276 two different structures are also possible for the mononucleotide product ion types comprising the addition of 80Da ($[B_{nt}+p-H]^-$ or $[B_{nt}+s-H_2O]^-$) and 98Da ($[B_{nt}+s-H]^-$ or $[B_{nt}+p+H_2O-H]^-$). These ions each differ in mass by 0.0599Da and are therefore indistinguishable in MS/MS spectra.

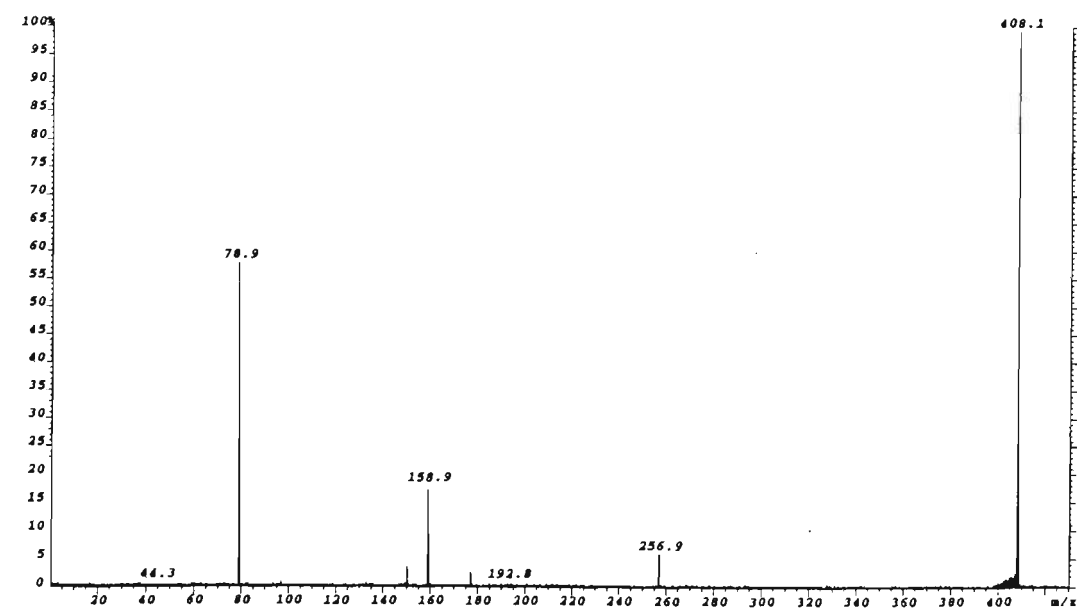
(a)



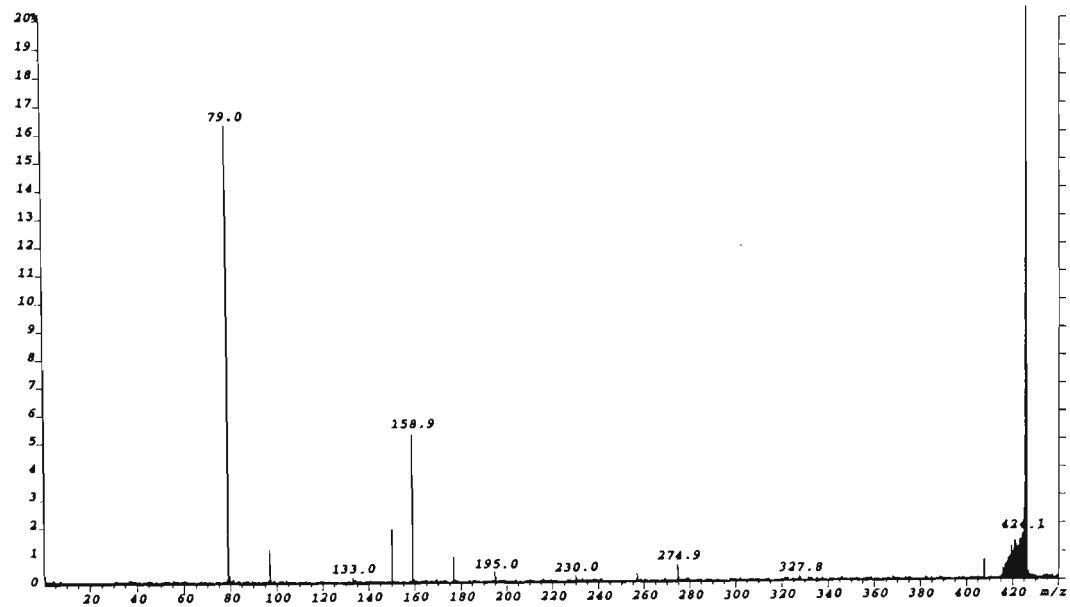
(b)



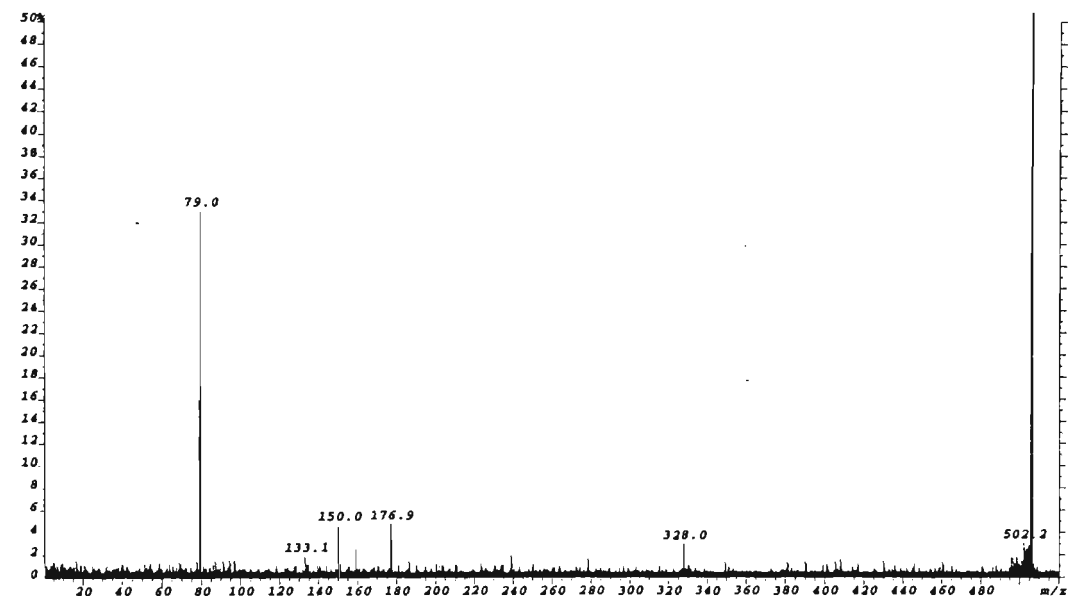
(c)



(d)



(e)



(f)

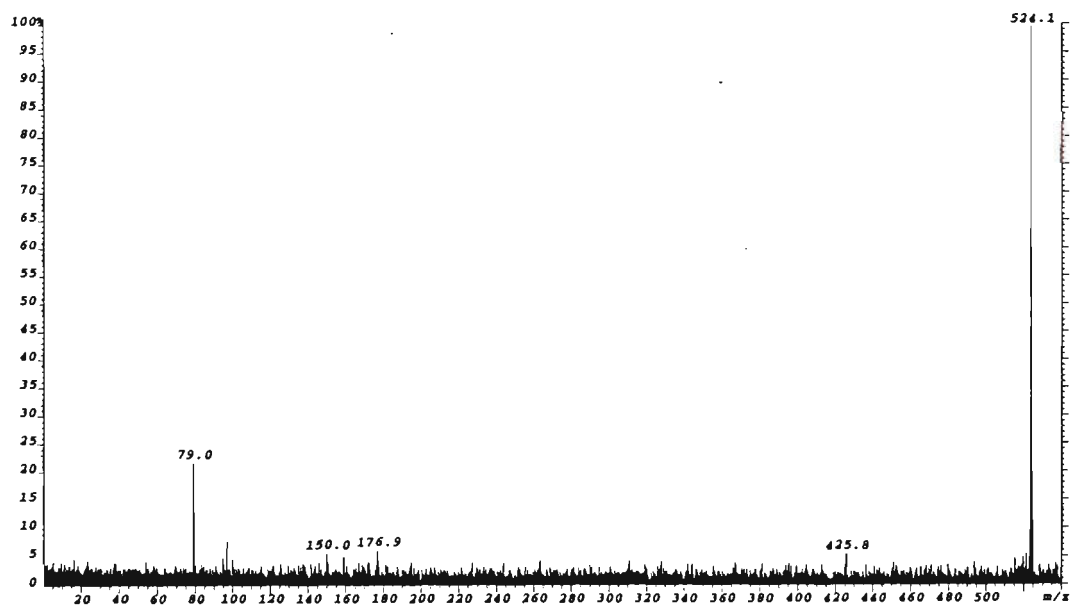


Figure 4.9: ESI-MS/MS spectra of the ions at (a) m/z 328.2; (b) m/z 346.2; (c) m/z 408.1; (d) m/z 426.0; (e) m/z 506.0 and (f) m/z 524.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Figure 4.10 shows the MS spectrum acquired at a resolution of approximately 10000 (5% valley peak-to-peak) of the product ion at m/z 426. The spectrum shows the presence of two peaks in the m/z region between 426.0 and 426.1, the largest of which at m/z 426.019 arises from the ion for the fragment $[G_{nt}+p+H_2O-H]^-$ ($m/z(\text{calc.})=426.021$). The less intense peak at m/z 426.080 (23% the intensity of the most intense ion), arises from the ion $[G_{nt}+s-H]^-$ ($m/z(\text{calc.})=426.081$). These observations are consistent with those obtained from the high resolution analysis of the ions at m/z 257 and 276 and suggests the existence of both fragments for analogous nucleotide ion types.

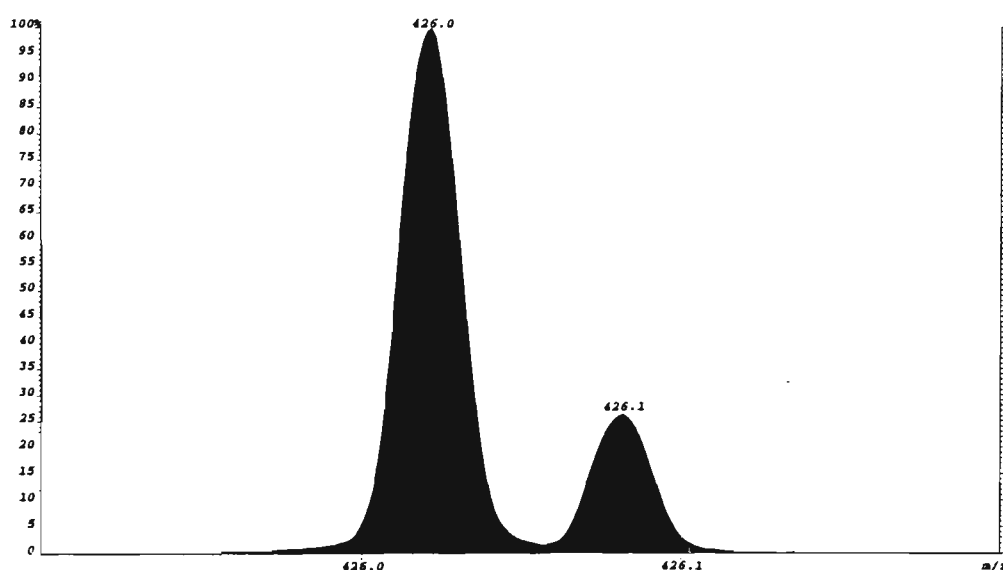


Figure 4.10: ESI-MS spectrum of the ion at m/z 426 generated by in-source fragmentation of 5'-d(CACGTG)-3' acquired at a resolution of 10000 at 5% valley peak-to-peak.

(c) Polynucleotide Ion Types - Sequence Ions and Internal Ions

The m/z region above 500 in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', shows an extensive number of polynucleotide product ions which includes both terminal ions (sequence ions) and internal ions. The predominant ion types observed in each case arise from nucleobase loss via N-glycosidic bond cleavage followed by cleavage of the 3' C-O bond of the sugar from which the base has been lost. This fragmentation pathway yields intense series of w and (a-B) product ions. In the MS/MS spectrum of 5'-d(CACGTG)-3', the entire w ion series is present with the

most intense ion in the spectrum corresponding to the w_3^- ion at m/z 979.0, followed by the w_2^- ion at m/z 649.9. Less intense ions are observed for w_4^- and w_5^- at m/z 1268.0 and 1581.1 respectively, and for w_1^- at m/z 346.0. The (a-B) series of ions, in comparison, are of lower relative abundance with the most intense ion of the series being the $[a_4-B_4H-2H]^-$ ion at m/z 988.0. Other (a-B) sequence ions are relatively weak in intensity with the $[a_5-B_5H-2H]^-$ ion absent from the spectrum. The spectrum also shows an extensive number of sequence ions arising from other backbone cleavages although these ions are significantly less pronounced (less than 10% relative abundance). Series of sequence ions from which bases have been lost are also present in the spectrum. In particular, loss of guanine from 3'-terminal sequence ions (w_2 to w_5 , x_2 to x_4 , y_2 and y_3 , and z_2 to z_4) is prevalent, as is loss of cytosine from 5'-terminal ions (a_2 to a_4 , and c_2 to c_3). Facile loss of neutral guanine and cytosine from the molecular ion yield ions at m/z 1639.1 and 1679.3 respectively, although no ions are detected for loss of neutral adenine and thymine. The MS/MS spectrum of 5'-d(CGTACG)-3' also shows the complete w-ion series with the most intense peak in the spectrum arising from the w_4^- ion at m/z 1252.1 with relatively abundant peaks also present for the w_2^- at m/z 634.7 and w_5^- at m/z 1581.2. In contrast, peaks of much lower intensity are generated for the w_3^- ion at m/z 947.9 and the w_1^- ion at m/z 346.1. The spectrum also shows a relatively intense series of (a-B) ions, in particular the $[a_4-B_4H-2H]^-$ and $[a_5-B_5H-2H]^-$ ions at m/z 1019.1 and m/z 1332.3 respectively. As with the oligonucleotide 5'-d(CACGTG)-3', the MS/MS spectrum of 5'-d(CGTACG)-3' contains a series of sequence ions from which neutral bases, in particular cytosine and guanine, have been lost. Moderately intense ions owing to loss of neutral guanine and cytosine from the molecular ion are also present in the spectrum.

In addition to these major ion types, further decompositions involving numerous other backbone cleavages are also observed resulting in a variety of internal ions. The general structures of these types of ions are summarised in table 4.6, and arise from various strings of phosphate and sugar groups attached similar to the mononucleotide ion types. As observed with mononucleotide ions, polynucleotide ions

with additions of 80Da, 98Da, 258Da and 276Da can have two possible structural assignments which differ in the addition of a phosphate or a sugar group. This is provided that both structures are possible from the sequence of the oligonucleotide.

Table 4.6: Polynucleotide product ion types generated in the negative ion MS/MS spectra of oligonucleotides.

Product ion types of polynucleotides* Observed Mass shift (Da)	
$[(B_1...B_n)-H]^-$	$[M(psB_1)+...+M(psB_n)-H]^-$
-p	-80
-p+H ₂ O	-62
+H ₂ O	+18
+p-H ₂ O	+62
+p	+80 (79.9663)
+s-H ₂ O	+80 (80.0262)
+p+H ₂ O	+98 (97.9769)
+s	+98 (98.0368)
+s+H ₂ O	+116 (116.0473)
+s+2H ₂ O	+134 (134.0579)
+p+s-H ₂ O	+160 (159.9925)
+p+s	+178 (178.0031)
+p+s+H ₂ O	+196 (196.0137)
+p+s+p	+258 (257.9694)
+s+p+s-H ₂ O	+258 (258.0293)
+p+s+p+H ₂ O	+276 (275.9800)
+s+p+s	+276 (276.0399)

*where p=PO₃H, s=deoxyribose-H₂O (C₅H₆O₂)

The MS/MS spectra of the terminal and internal polynucleotide product ions generated in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are given in appendices 4.2 and 4.3 respectively. Fragmentation of these ions in almost all cases gives rise to backbone, nucleobase and mononucleotide ion types in the low *m/z* region of the spectrum which are indicative of the constituent nucleotides in the fragment. The structural information provided by the observation of these previously characterised ion types, therefore enables the identity of the fragment to be 'pieced' together.

From the product ion assignments listed in tables 4.2 and 4.3 it can be seen that many of the terminal sequence ions can alternatively be assigned as internal ions where the internal sequence of the oligonucleotide comprises fragments of the same molecular compositions as certain terminal fragments. Furthermore, sequence ions arising from a, c, x, and z cleavage cannot be unambiguously assigned as terminal fragment ions in the present study since these ions can be alternatively assigned as b, d,

w, and y ions respectively with cleavage of the terminal hydroxyl group. It should be pointed out that these possibilities are not listed in the product ion assignments given in tables 4.2 and 4.3 to avoid added complexity. Terminal and internal ions of the same molecular composition are difficult to distinguish from in-source collisional activation experiments undertaken in this work, since further decomposition of these ions yields product ions which can arise from either structure. An example of this occurs the ion at m/z 649.9 in the MS/MS spectrum of 5'-d(CACGTG)-3' (figure 4.3) which can be attributed to the w_2^- sequence or the internal fragment $[(GT)+H_2O-H]^-$. In the MS/MS spectrum of the ion at m/z 650.1 shown in figure 4.11, the product ions corresponding to terminal fragments at m/z 570.0, 524.2, 425.9, 346.0 (see assignments of these product ions given in table 4.2) can also be assigned as internal ions. Likewise, the $[a_3-B_3H-2H]^-$ ion at m/z 699.1 can also occur internally as the fragment $[(AC)+p+H_2O-H]^-$.

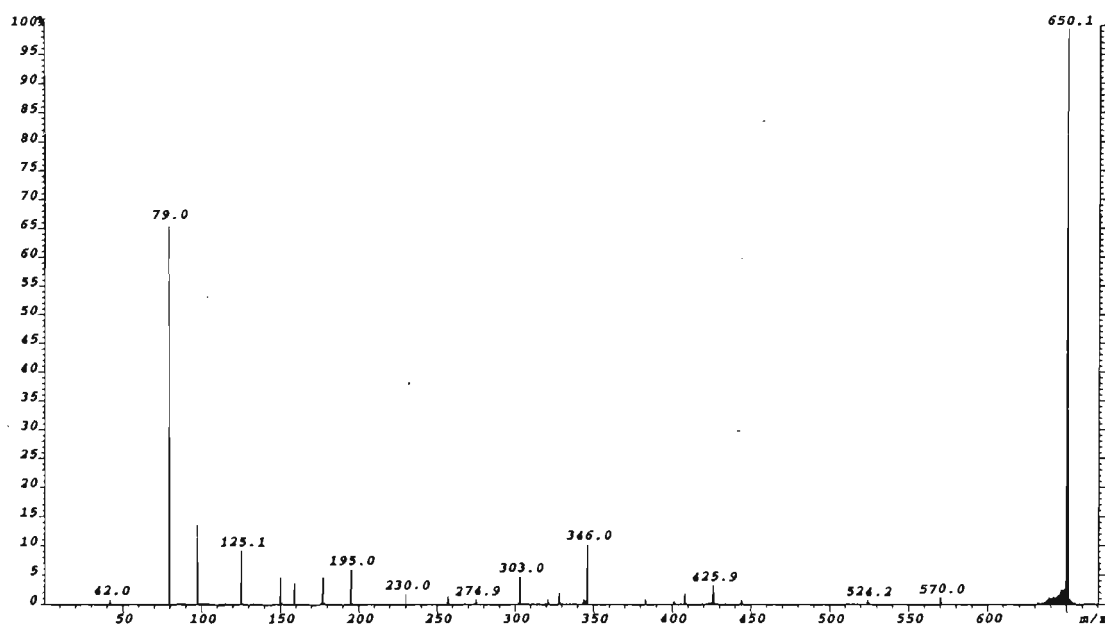


Figure 4.11: ESI-MS/MS spectrum of the ion at m/z 650.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Table 4.7 shows the types of sequence ions observed and the analogous internal fragment ions corresponding to the mass of each respective sequence ion. For the purposes of simplifying spectral interpretation, ions are assigned as w and (a-B) sequence ions in the discussion of MS/MS spectra of oligonucleotides where internal

ion assignments are also possible. These assignments are made on the basis of ion intensity since numerous previous studies have identified the backbone cleavages giving rise to these two ion types as dominant fragmentation pathways in the ESI-MS/MS of oligonucleotides.

Table 4.7: Alternative internal ion assignments for sequence ions based on molecular composition

Sequence ion	Analogous Internal Ion*
[a-2H] ⁻	[(B _x ..B _n)-p-H] ⁻
b ⁻	[(B _x ..B _n)-p+H ₂ O-H] ⁻
[c-2H] ⁻	[(B _x ..B _n)-H] ⁻
d ⁻	[(B _x ..B _n)+H ₂ O-H] ⁻
w ⁻	[(B _x ..B _n)+H ₂ O-H] ⁻
[x-2H] ⁻	[(B _x ..B _n)-H] ⁻
y ⁻	[(B _x ..B _n)-p+H ₂ O-H] ⁻
[z-2H] ⁻	[(B _x ..B _n)-p-H] ⁻

*where (B_x..B_n) is an internal polynucleotide ion

Further difficulties are encountered in the assignment of sequence ions in the MS/MS spectrum of oligonucleotides where it is not possible to determine from which terminus the sequence ion is generated. For the oligonucleotide 5'-d(CGTACG)-3', phosphodiester bond cleavage between the nucleosides containing G-2 and T-3 yields 5'-terminus sequence ions which can be alternatively assigned as deriving from the 3'-terminus sequence ions arising from phosphodiester bond cleavage between the nucleosides containing A-4 and C-5. More specifically, [a₂-2H]⁻, b₂⁻, [c₂-2H]⁻, and d₂⁻ ions can also be assigned as [z₂-2H]⁻, y₂⁻, [x₂-2H]⁻, w₂⁻ ions respectively, and [a₄-2H]⁻, b₄⁻, [c₄-2H]⁻, and d₄⁻ fragments could alternatively be assigned as [z₄-2H]⁻, y₄⁻, [x₄-2H]⁻, w₄⁻ ions respectively. Similar problems are also encountered with (a-B) product ions since [a₅-B₅H-2H]⁻, [a₄-B₄H-2H]⁻, and [a₃-B₃H-2H]⁻ have the same composition as the 3'-terminal ions [z₅-GH-2H]⁻, [z₄-AH-2H]⁻, and [z₃-AH-2H]⁻.

Regarding the assignment of the internal ions generated in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' (which can be unambiguously assigned as such), it should be noted that there are many cases where the internal ion can be attributed to more than one possible structure. This problem arises because the exact

points of cleavage across the oligonucleotide backbone can not be ascertained with certainty. For example, the ion at m/z 795.0 in the MS/MS spectrum of 5'-d(CACGTG)-3' assigned as $[(CG)+s+p-H]^+$ can be assigned as any of the following structures: pspCpsG; spsCpsGp; psCpsGps; sCpsGpsp. Furthermore, in each case there are two added possibilities depending on whether the sugar group has its hydroxyl group attached to the 3'- or 5'-terminus. Although these structures are indistinguishable from the experiments conducted in the present study, these problems could be overcome by conducting further studies utilising O^{18} labeled oligonucleotides and/or oligonucleotides with labeled termini.

A significant difficulty in assigning product ions in MS/MS spectra acquired at low resolution is the ability to distinguish between two or more fragments of different base compositions having m/z values which cannot be resolved. An example of this appears with the product ion at m/z 810.0 in the MS/MS spectrum of 5'-d(CACGTG)-3', which can be assigned as either: (a) the addition of 178Da (s+p) to a dinucleotide ion comprising the bases G and T (m/z 810.09) or; (b) the (CAC) trinucleotide less PO_3H (m/z 810.17). Similarly, two possible product ion assignments also exist for the ions at m/z 890.2 and m/z 988.2 which constitute additions of 80 Da (p or s- H_2O) and 98Da (p+ H_2O or s) respectively to the fragment at m/z 810.0. In each case, the mass difference between the two possibilities is 0.08238Da which thus requires a resolution of approximately 10000 for their correct identities to be discerned. Further analysis of product ions of these masses generated in the source enabled their correct identification owing to the differing base compositions of the two species. The MS/MS spectrum obtained of the species at m/z 810.0 in the MS/MS spectrum of 5'-d(CACGTG)-3' (figure 4.3) is shown in figure 4.12. It can be seen that the spectrum shows the presence of C^- (m/z 110.0), T^- (m/z 125.0), A^- (m/z 134.0) and G^- (m/z 150.0) nucleobase ions, and mononucleotide product ions comprising C (m/z 287.9), T (at m/z 302.9, 383.0, 463.1 and 480.8) and G nucleobases (m/z 327.8, 425.8, and 505.8) (see the assignments of these various ions listed in table 4.5). Therefore, it is apparent that the two different unresolved species are present since a nucleotide product ion

comprising all four residues would have a mass greater than that of the product ion at m/z 810.0.

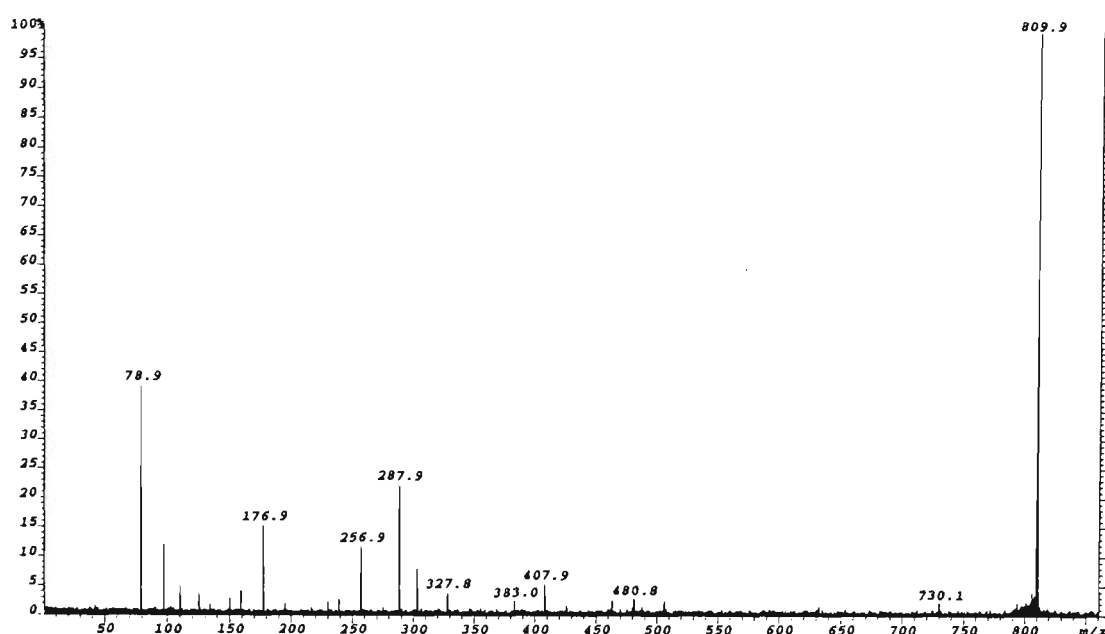


Figure 4.12: ESI-MS/MS spectrum of the ion at m/z 809.9 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The product ion at m/z 890.2 from the MS/MS spectrum of 5'-d(CACGTG)-3' (figure 4.3) corresponds to a species comprising the ion at m/z 810.0 with either an added p or s- H_2O group. The MS/MS spectrum of this product ion is shown in figure 4.13. The spectrum shows only C⁻ and A⁻ nucleobase ions (at m/z 110.0 and 134.0 respectively), and mononucleotide fragments comprising only the C and A nucleobases. The most intense ion in the spectrum at m/z 368.0 arises from the C nucleotide to which either a p or s- H_2O group is attached. The analogous fragment for the A nucleotide is also present at m/z 392.0. The ions at m/z 601.1 correspond to the (AC) dinucleotide with the ion at m/z 681.1 arising from an added p or s- H_2O group (assignments are given in table 4.2). The spectrum of this product ion thus provides clear evidence that the ion observed at m/z 890.2 arises only from the species $[c_3-2H]^-$. Similarly the MS/MS spectrum of the product ion at m/z 988.0 (appendix 4.2) only shows C⁻ and A⁻ nucleobases and mononucleotide ions. Therefore, although this ion could correspond to either $[a_4-B_4H-2H]^-$ (m/z 988.179) or $[(GT)+s+p+s+p-H]^-$ (m/z 988.097), the MS/MS spectrum is consistent only with a species comprising C

and A nucleobases which is the $[a_4-B_4H-2H]^-$ ion. Hence, in the absence of accurate mass measurement capabilities in MS-2, product ions having an m/z which corresponds to that of two or more fragments of different base compositions can be effectively distinguished by MS/MS spectra of source-generated product ions.

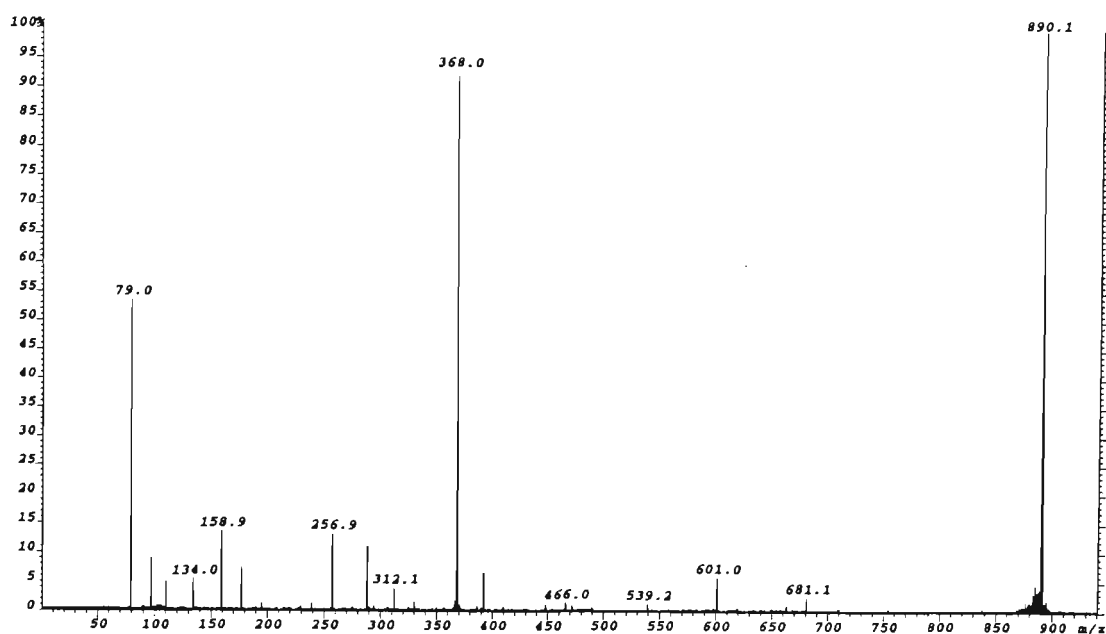


Figure 4.13: ESI-MS/MS spectrum of the ion at m/z 890.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

4.4.2 Analysis of Product Ion Fragmentation Pathways

Unlike the gentle fragmentation produced by ion trap instrumentation which enables monitoring of consecutive fragmentation pathways, the higher energy collisions associated with MS/MS on magnetic sector instrumentation makes it difficult to deduce information about the various fragmentation pathways of oligonucleotides owing to the occurrence of multiple collisions without the aid of collisional cooling. Thus, in order to obtain an insight into the fragmentation of oligonucleotides observed in this study, an analysis was undertaken of the product ions detected in MS/MS spectra by examining MS/MS spectra of ions of the same m/z generated in the source. By investigating the initial decomposition products of the major product ion types observed in MS/MS spectra of oligonucleotides, information can then be obtained of the pathways resulting from successive collisions of these ions.

Furthermore, the origin of the extensive number of internal product ions generated by multiple collisions of sequence ions in MS/MS spectra can be discerned.

(a) *5'-d(CACGTG)-3'*

In the high m/z region of the MS/MS spectrum of 5'-d(CACGTG)-3' favourable loss of both neutral guanine and cytosine is observed. The MS/MS spectrum of the ion corresponding to loss of neutral guanine from the molecular ion (at m/z 1639.1[#]) is shown in figure 4.14. Fragmentation of this ion yields the $[a_4-B_4H-2H]^-$ ion (m/z 989.0) and an ion at m/z 650.3 which can be attributed either to w_2^- if the loss of neutral guanine occurs from the G-4 position or to $[(GT)_n+H_2O-H]^-$ if it is the 3'-terminal guanine that has been lost. The ion at m/z 828.4 corresponds to the species $[w_3-GH]^-$ which arises from loss of neutral cytosine C-3 followed by ribose 3' C-O bond cleavage of the $[M-GH-H]^-$ ion. Hence it appears that loss of neutral guanine and cytosine are the most facile fragmentation pathways occurring, followed by cleavage of the adjacent ribose 3' C-O bonds of the residues from which the bases are lost.

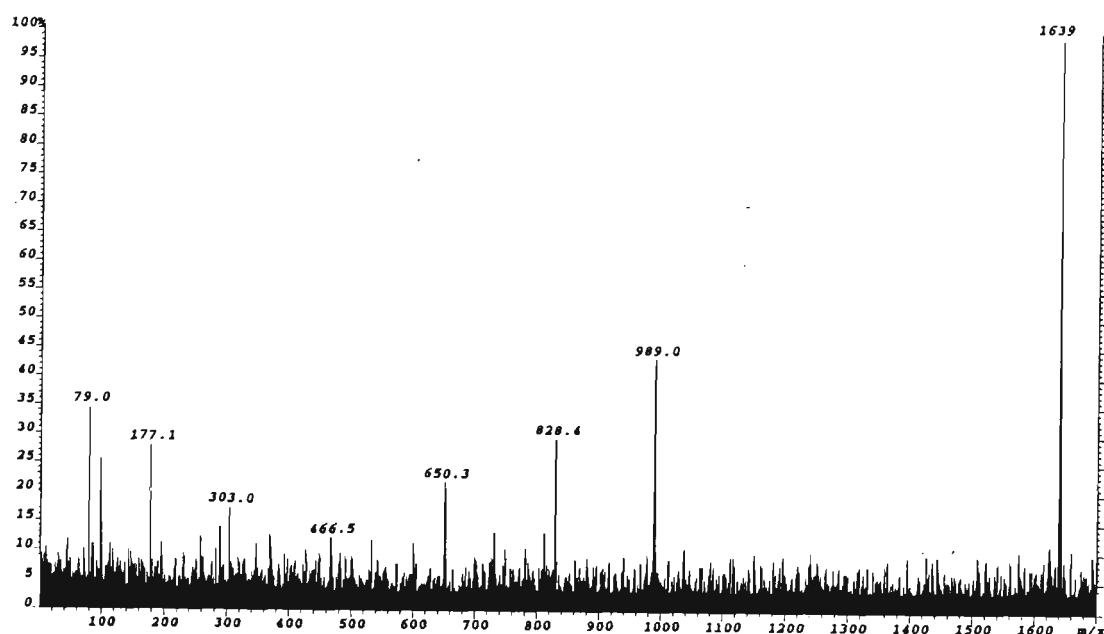


Figure 4.14: ESI-MS/MS spectrum of the $[M-GH-H]^-$ ion (m/z 1639.1) generated by in-source fragmentation of 5'-d(CACGTG)-3'.

[#]These ions are generated in the source and owing to small differences between the calibration of the magnet and the TOF and/or between TOF spectra obtained at different times, the m/z of product ions observed in MS/MS spectra may vary between 0.1-0.5 Da.

The MS/MS spectrum of the w_5^- ion (m/z 1581.3) is shown in figure 4.15 and the main dissociation pathways observed are illustrated in scheme 4.2.[#] Fragmentation of this ion favourably produces w_2^- at m/z 649.9, and w_3^- ion at m/z 978.8. An intense ion is observed at m/z 779.0 which corresponds to the complementary 5' fragment, [(AC)+p+s-H]⁻, arising from loss of neutral guanine G-4 followed by cleavage of the adjacent ribose 3' C-O bond which generates w_2^- . Hence the moderate abundance of this ion in the MS/MS spectrum of 5'-d(CACGTG)-3' can be explained by its favourable production from the fragmentation of w_5^- .

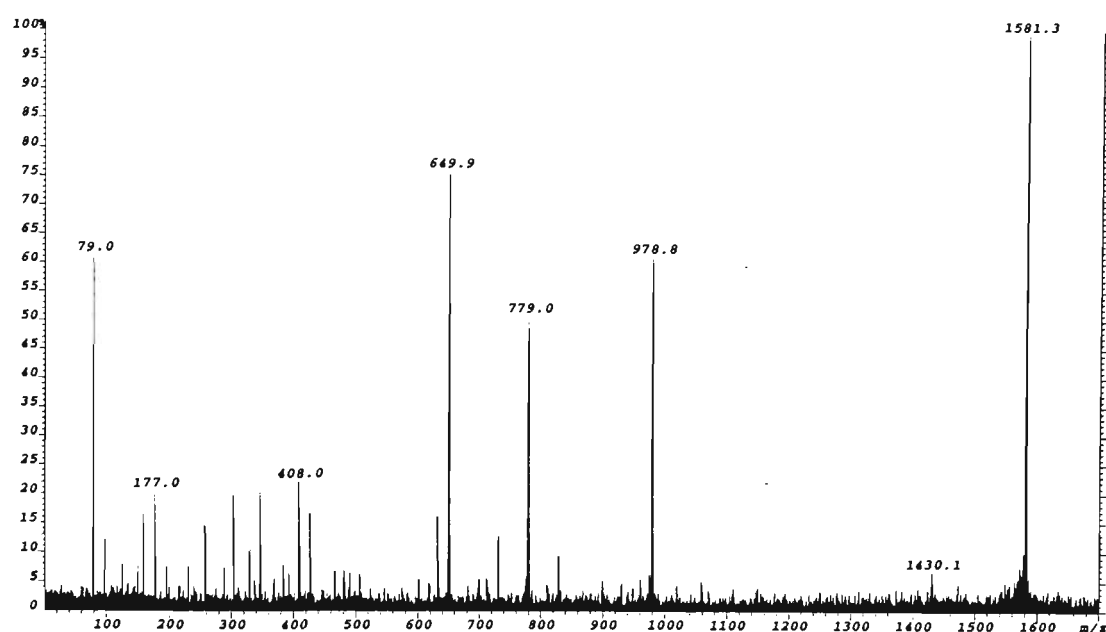
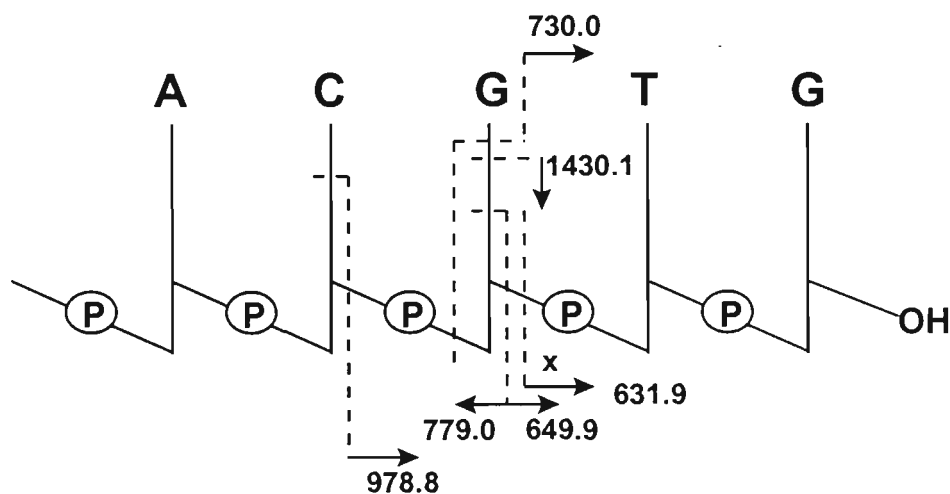


Figure 4.15: ESI-MS/MS spectrum of the w_5^- ion (m/z 1581.3) generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.2: Fragmentation pathways of the w_5^- ion (m/z 1581.2) generated by in-source fragmentation of 5'-d(CACGTG)-3'.

[#] To reduce complexity, the hydrogen transfer associated with these cleavages is not shown on this and all the following schemes. The net hydrogen transfer associated with the formation of difference sequence ions is given in table 4.1.

The MS/MS spectrum arising from further decomposition of the ion at m/z 778.8 is shown in figure 4.16. The spectrum shows the m/z 698.9 ion which may be attributed either to the ion, $[(AC)+p+H_2O-H]^-$, from d cleavage of the ribose at the 3'-terminus, or alternatively to the ion, $[(AC)+s-H]^-$, from y cleavage to the 5' side of the adenine-containing ribose. These two pathways are shown in scheme 4.3.

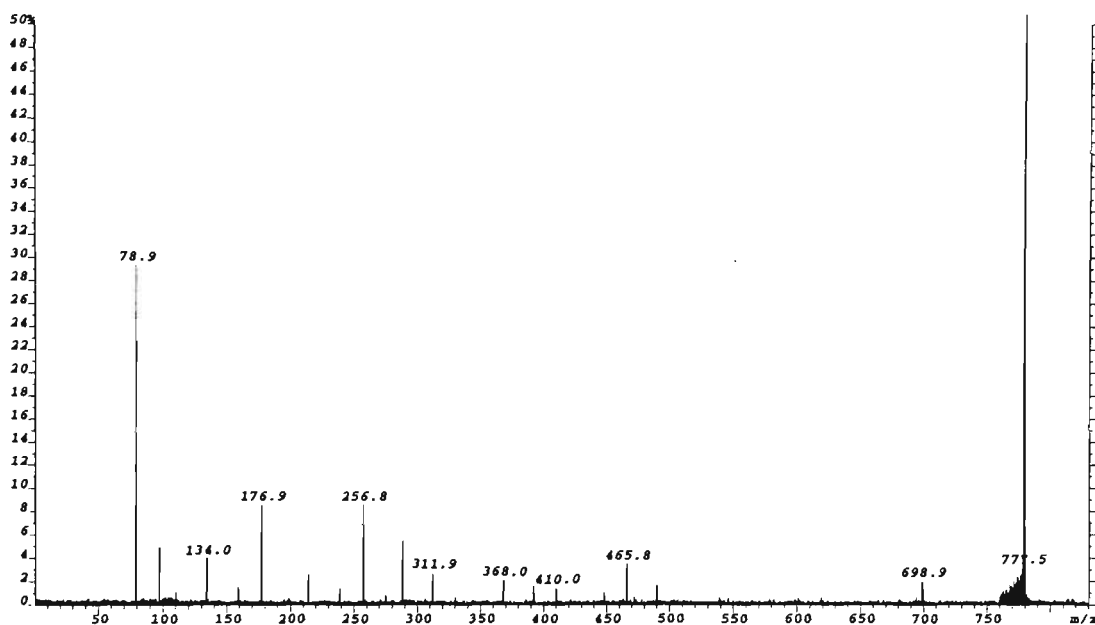
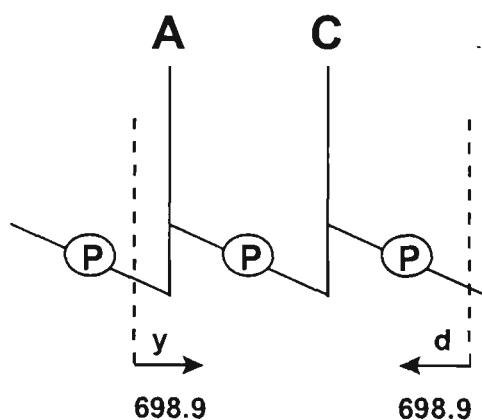


Figure 4.16: ESI-MS/MS spectrum of the ion at m/z 778.8 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.3: Alternate fragmentation pathways of the ion at m/z 778.8 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Fragmentation of the w_5^- ion (figure 4.15) also yielded the ions m/z 631.9, 730.0 and 828.0 which are moderately intense in the MS/MS spectrum of 5'-d(CACGTG)-3'. The ion at m/z 631.9 could arise from $[x_2-2H]^-$ or alternatively could result from phosphodiester bond cleavage to the 3' side of C-3 and T-5 generating $[(GT)-H]^-$. The ion at m/z 730.0 can correspond to $[z_3-GH-2H]^-$ or may be attributed to either of the internal ions $[(GT)+s-H]^-$ or $[(GT)+p-H_2O-H]^-$. The ion at m/z 828.0 is also observed to be generated from further fragmentation of the product ion $[M-GH-H]^-$ at m/z 1639.1 and would appear to arise from the loss of neutral guanine from w_3^- . The MS/MS spectrum of the w_5^- ion also yields a small ion for the loss of neutral guanine at m/z 1430.1 and the MS/MS spectrum obtained for this ion is shown in figure 4.17. The fragmentation illustrated in scheme 4.4 shows this ion predominantly yields the $[(AC)+p+s-H]^-$ ion (m/z 779.1) and the w_2^- ion (m/z 649.9) which would be the predominant fragmentation pathway expected to follow the loss of neutral guanine at the G-4 position. The fragmentation of $[w_5-GH]^-$ also yields the product ions at m/z 828.1, m/z 810.0 and m/z 730.1 for the 3'-terminal ions arising from w, x and z cleavages respectively to the 3' side of the ribose to which cytosine is attached.

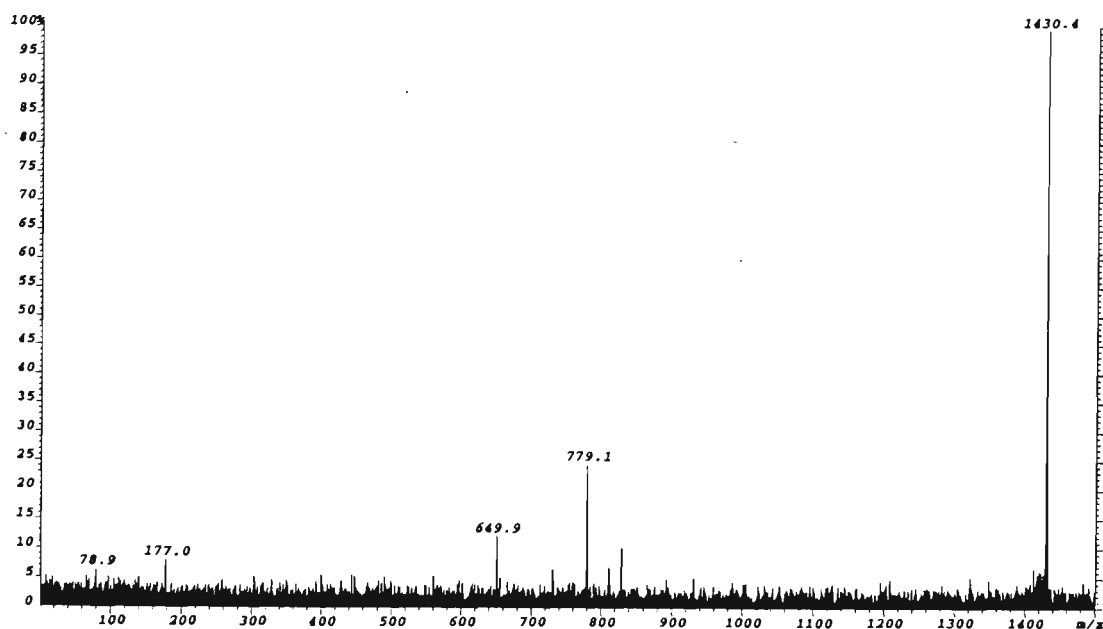
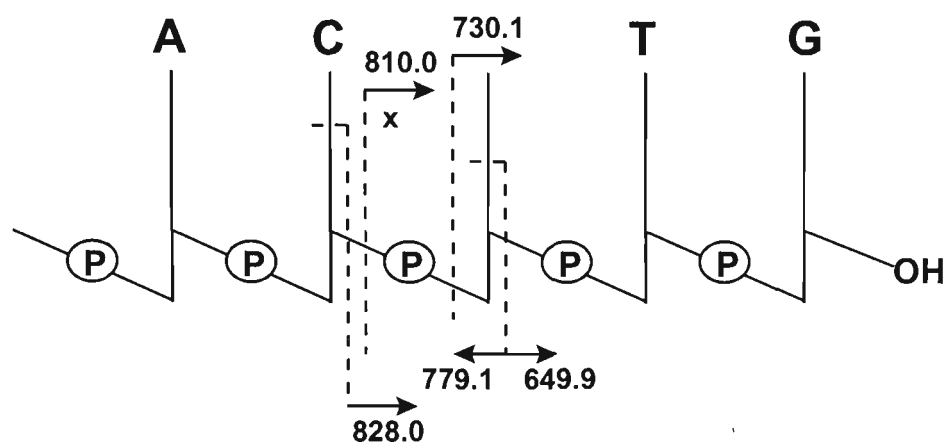


Figure 4.17: ESI-MS/MS spectrum of the ion at m/z 1430.4 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.4: Fragmentation pathways of the ion at m/z 1430.4 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Figure 4.18 shows the MS/MS spectrum of the w_4^- ion (m/z 1268.2) and the main fragmentation pathways observed in the spectrum are illustrated in scheme 4.5. Fragmentation of this ion is similar to that observed for w_5^- , which generates a predominant ion for w_2^- at m/z 650.0 as well as w_3^- at m/z 978.8, and less intense ions are also observed at m/z 632.0, m/z 730.1. The ion at m/z 1117.2 corresponds to the loss of neutral guanine from w_4^- . The MS/MS spectrum obtained for this ion is shown in figure 4.18 and is indicative of highly specific fragmentation, illustrated in scheme 4.6, which yields intense ions at m/z 650.0 owing to w_2^- , and at m/z 498.8 which would appear to arise from the loss of the 3'-terminal guanine from the w_2^- ion. An ion at m/z 465.9 is also present for the complementary 5'-terminal fragment arising from loss of guanine G-4 followed by cleavage of the ribose 3' C-O bond. The fact that formation of the w_2^- ion is observed to be a major fragmentation pathway is consistent with the loss of neutral guanine occurring from the G-4 position of the precursor ion, since the subsequent cleavage of the adjacent ribose 3' C-O bond yielding w_2^- would be expected to be a highly favourable fragmentation pathway following the loss of neutral G-4. Weak intensity ions are observed at m/z 729.7 and m/z 828.2 for $[z_3\text{-GH-2H}]^-$ and $[w_3\text{-GH}]^-$ respectively.

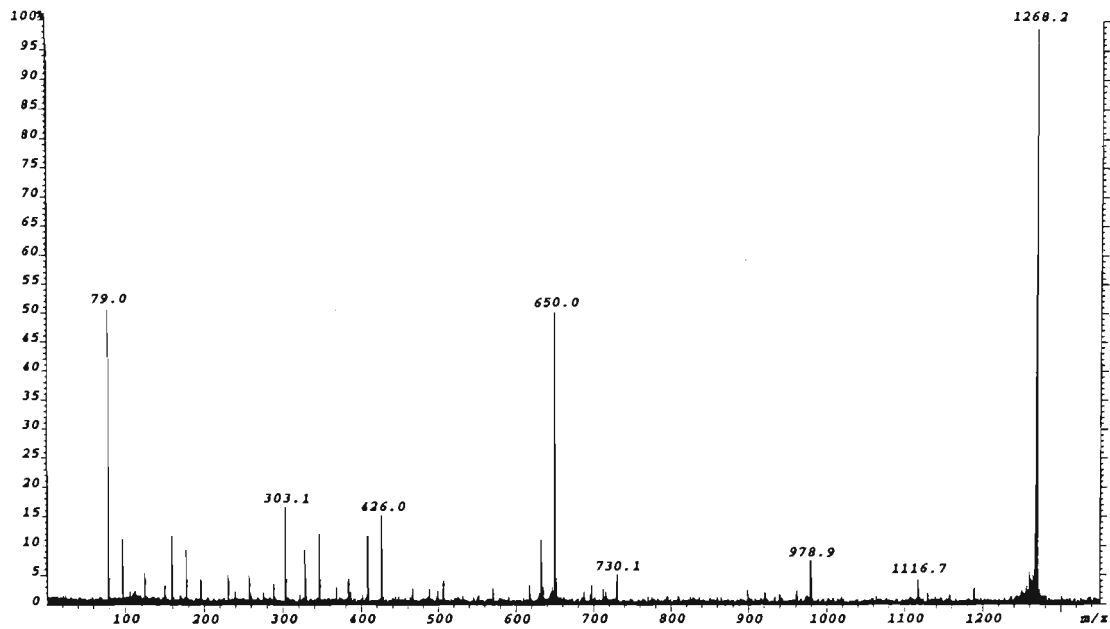
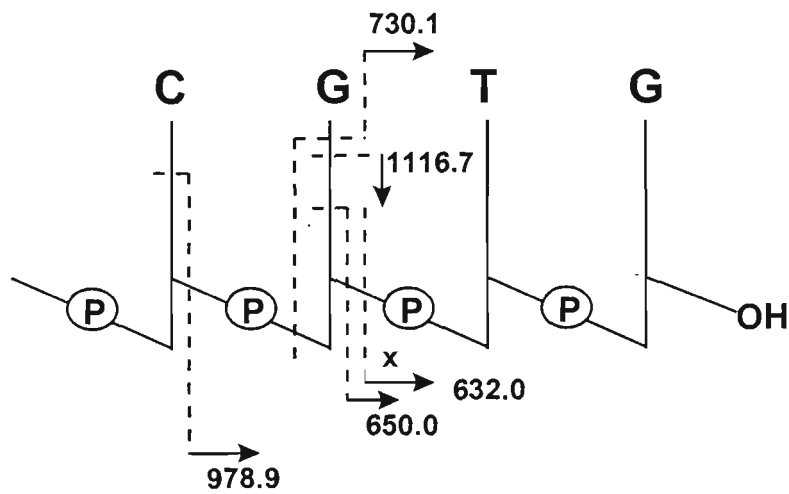


Figure 4.18: ESI-MS/MS spectrum of the w_4^- ion (m/z 1268.2) generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.5: Fragmentation pathways of the w_4^- ion (m/z 1268.2) generated by in-source fragmentation of 5'-d(CACGTG)-3'.

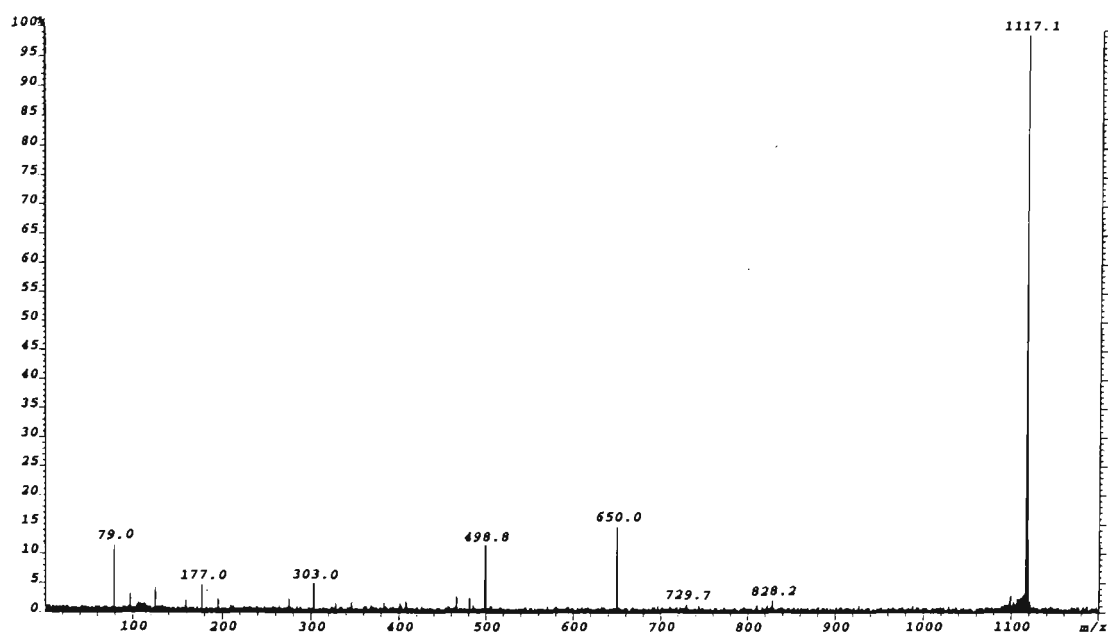
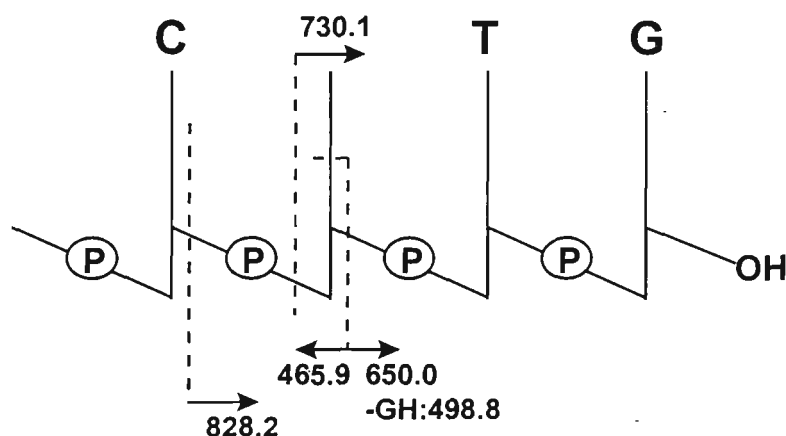


Figure 4.19: ESI-MS/MS spectrum of the ion at m/z 1117.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.6: Fragmentation pathways of the ion at m/z 1117.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Decomposition of the w_3^- ion at m/z 979.3 (figure 4.20) was found to produce no further facile fragmentation pathway with the MS/MS spectrum of this ion showing predominantly PO_3^- in addition to weakly abundant ions owing to G $^-$ and T $^-$ and various nucleotide ion types containing these bases. In contrast to the fragmentation of w_4^- and w_5^- , the loss of neutral guanine (G-4) and the formation of w_2^- is not seen to occur in the MS/MS spectrum of w_3^- . The observed lack of preference for this fragmentation

pathway is perhaps owing to competition from a more facile pathway involving loss of charged guanine, which yields G^- as the second most intense product ion observed. The weak abundance of the w_1^- ion at m/z 346.1 may arise from the lack of preference for loss of neutral thymine which would make this w fragmentation channel unfavourable (discussed in more detail in chapter five). The apparent stability of the w_3^- ion to further fragmentation is a likely factor contributing to its strong relative abundance in the MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CACGTG)-3'. In the MS/MS spectrum of w_2^- ion (not shown), the PO_3^- ion was also the major product. The w_1^- ion was also present in that spectrum with a small ion observed for loss of neutral thymine at m/z 524.1.

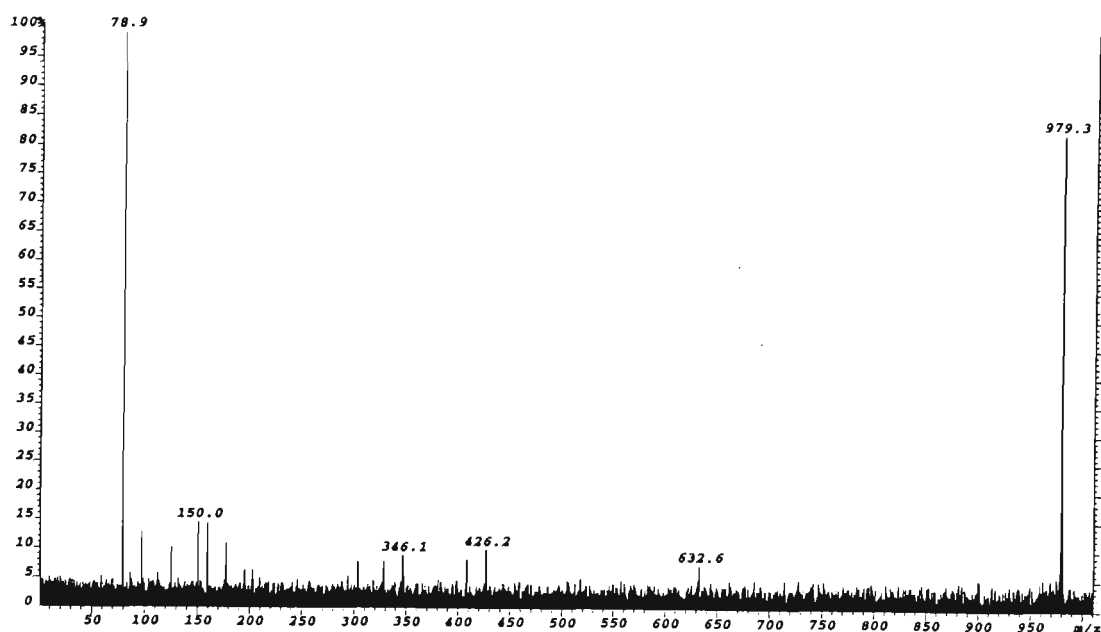


Figure 4.20: ESI-MS/MS spectrum of the w_3^- ion (m/z 979.3) generated by in-source fragmentation of 5'-d(CACGTG)-3'.

In addition to the predominant w ions observed in the MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CACGTG)-3', a favourable fragmentation pathway involving (a-B)-type cleavage at the G-4 position resulted in an intense $[a_4-B_4H-2H]^-$ ion at m/z 988.0. Collisional activation of the $[a_4-B_4H-2H]^-$ ion generated in the source yielded the MS/MS spectrum shown in figure 4.21. The major fragmentation pathways of this ion are shown in scheme 4.7 and involves base loss followed by w -type cleavage.

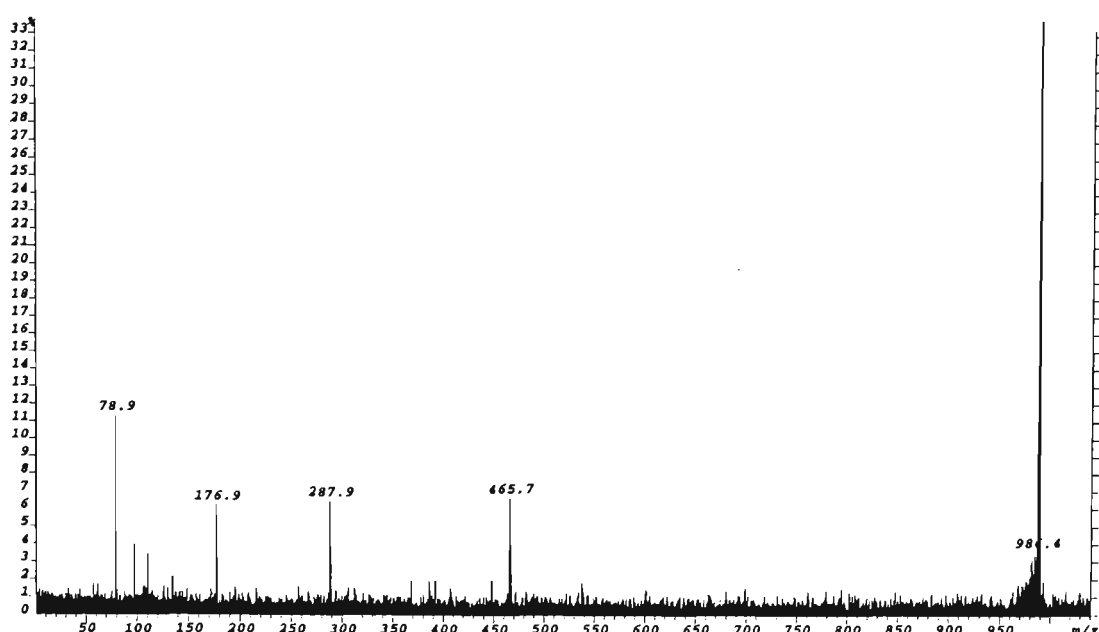
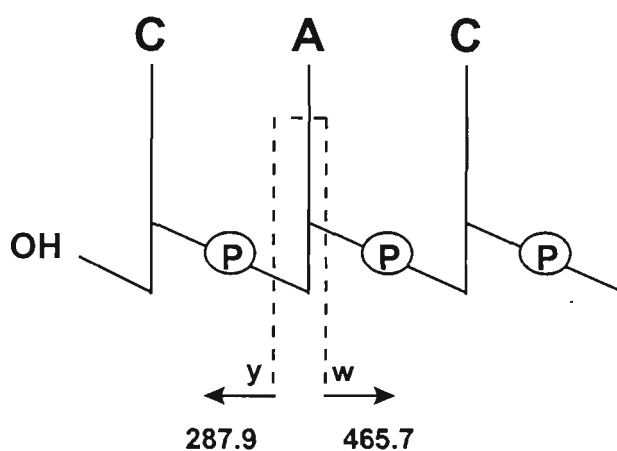


Figure 4.21: ESI-MS/MS spectrum of the $[a_4-B_4H-2H]^-$ ion (m/z 988.0) generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.7: Fragmentation pathways of the ion at m/z 988.0 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The MS/MS spectrum of the $[a_4-B_4H-2H]^-$ ion yields a major ion at m/z 465.7 owing to $[C_{nt}+p+s-H]^-$ which corresponds to the w ion generated from initial loss of neutral adenine followed by cleavage of the ribose 3' C-O bond. The $[C_{nt}-H]^-$ ion at m/z 287.9 could arise either from further decomposition of $[C_{nt}+p+s-H]^-$ via cleavages of the ribose 3' C-O bond without initial base loss, or y_1 -type cleavage of the 5'-terminal cytosine. The former fragmentation pathway could also yield the complementary (s+p) ion at m/z 176.9 although this species could be generated from numerous other backbone fragmentations. It can be seen that the $[a_4-B_4H-2H]^-$ ion does not fragment to yield the $[a_3-B_3H-2H]^-$ ion (m/z 699.1) which would be possible following a cleavage

arising from initial loss of neutral cytosine C-3. In addition, no product ions are present arising from w-type cleavage generated by loss of the 5'-terminal cytosine. These observations could be explained on the basis that facile loss of adenine is occurring owing to the negative charge being preferentially located on the phosphate to the 3' side of the ribose to which adenine is attached catalysing base loss at this position.

The MS/MS spectrum of 5'-d(CACGTG)-3' also showed a small peak for the $[a_3-B_3H-2H]^-$ ion although this product ion was much less pronounced than that of the $[a_4-B_4H-2H]^-$ ion. The MS/MS spectrum of the $[a_3-B_3H-2H]^-$ ion (m/z 699.0), shown in figure 4.22, shows an ion at m/z 489.9 owing to $[A_n+p+s-H]^-$, and corresponds to the 3' fragment arising from cleavages of the ribose 3' C-O bond of the ribose containing C-1. This fragmentation pathway is analogous to that observed in the MS/MS spectrum of the $[a_4-B_4H-2H]^-$ product ion involving loss of neutral adenine, and as previously discussed this fragmentation pathway could also occur as a result of preferential location of the negative charge on the phosphate group to the 3'-terminal side of the residue catalysing the loss of the 5'-terminal cytosine.

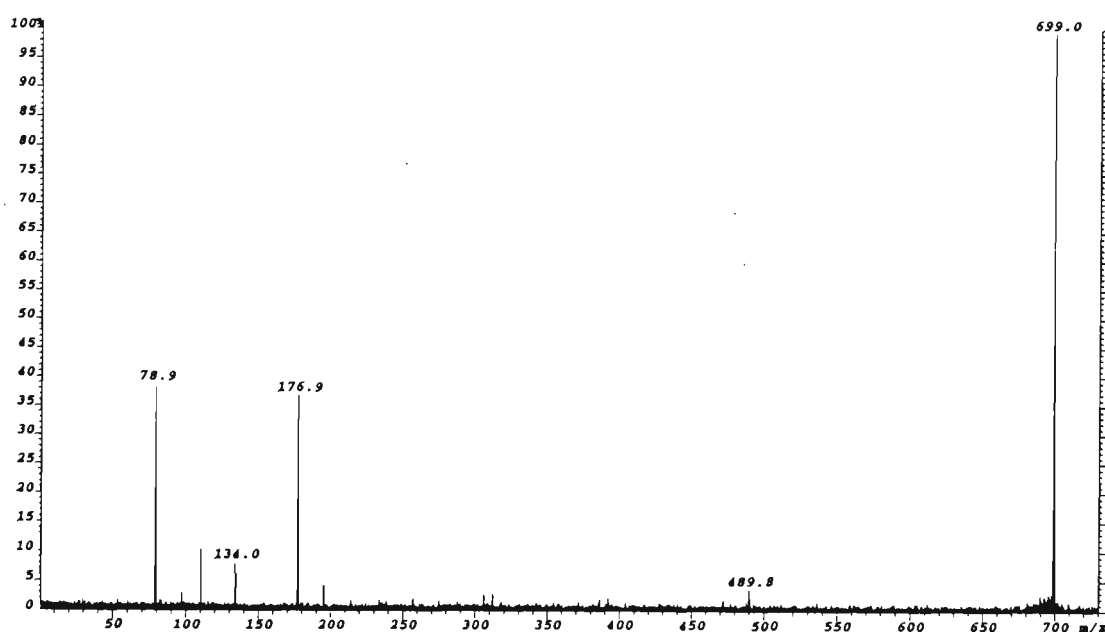


Figure 4.22: ESI-MS/MS spectrum of the $[a_3-B_3H-2H]^-$ ion (m/z 699.0) generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The MS/MS spectrum of $[a_3-B_3H-2H]^-$ also showed predominant ions in the low m/z region of the spectrum for PO_3^- (m/z 78.9) and $[ps-H]^-$ (m/z 176.9), and loss of C^- and A^- . The absence of the $[a_2-B_2H-2H]^-$ ion as a decomposition product in the MS/MS spectra of either the $[a_3-B_3H-2H]^-$ or the $[a_4-B_4H-2H]^-$ product ions is not surprising given that this species was of low abundance in the MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CACGTG)-3'.

(b) 5'-d(CGTACG)-3'

Similar to 5'-d(CACGTG)-3', the MS/MS spectrum of 5'-d(CGTACG)-3' (figure 4.4) also showed the loss of neutral guanine and cytosine to be the initial fragmentation pathways generated. The MS/MS spectrum of the ion at m/z 1639.3 corresponding to loss of neutral guanine is shown in figure 4.23.

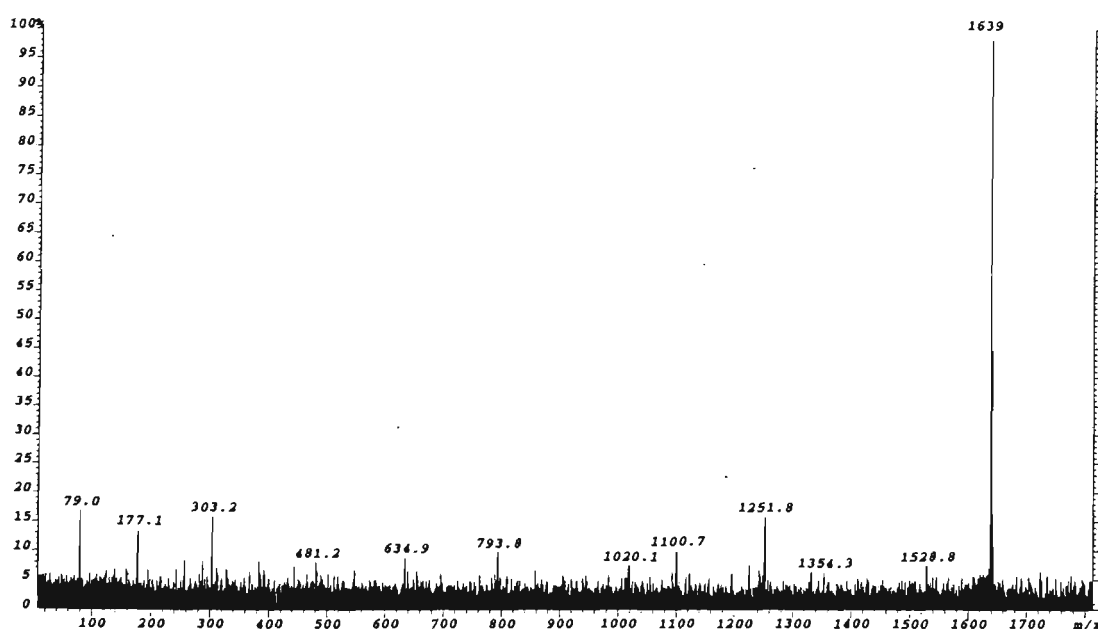


Figure 4.23: ESI-MS/MS spectrum of the $[M-GH-H]^-$ ion (m/z 1639.3) generated by in-source fragmentation of 5'-d(CGTACG)-3'.

This ion yields product ions which could be attributed to a precursor ion having lost either the guanine at the G-2 position or the 3'-terminal guanine. The ion at m/z 1251.8 can be attributed either to the w_4^- ion arising from the loss of neutral guanine at the G-2 position followed by ribose C-O bond cleavage, or to the d_4^- ion if the precursor m/z 1639.3 ion corresponds to the loss of the 3'-terminal guanine. If this were in fact

the case, a further possibility is that the m/z 1251.8 product ion can correspond to the internal fragment $[(GTAC)+H_2O-H]^-$, although this is an unlikely contribution given the observation of the m/z 635.0 ion which only corresponds to either the terminal w_2^- or d_2^- ion. The ion at m/z 1100.7 corresponds to the further loss of guanine from the species at m/z 1251.8. The product ion for loss of neutral cytosine (appendix 4.3) was found to be quite stable since the MS/MS spectrum of this ion showed only a small product ion generated at m/z 1581.2 which can be attributed either to w_5^- or to d_5^- depending on whether the loss of cytosine is the 5'-terminal C, or C-5 respectively.

The relatively intense ion at m/z 1581.2 in the MS/MS spectrum of 5'-d(CGTACG)-3' has been attributed to the w_5^- sequence ion and the MS/MS spectrum of this ion is shown in figure 4.24. The spectrum shows abundant ions at m/z 1123.0 and m/z 810.0 which correspond to the 5'-terminal fragments arising from ribose 3' C-O bond cleavage to the 3' side of cytosine and adenine respectively following prior base loss. Alternatively, the ion at m/z 1123.0 may also be attributed to the 3' fragment ion $[x_4-CH-2H]^-$, however, this is considered less likely since x_n ions are not dominant ions in the MS/MS spectra of intact oligonucleotides. These fragmentation pathways are illustrated in scheme 4.8.

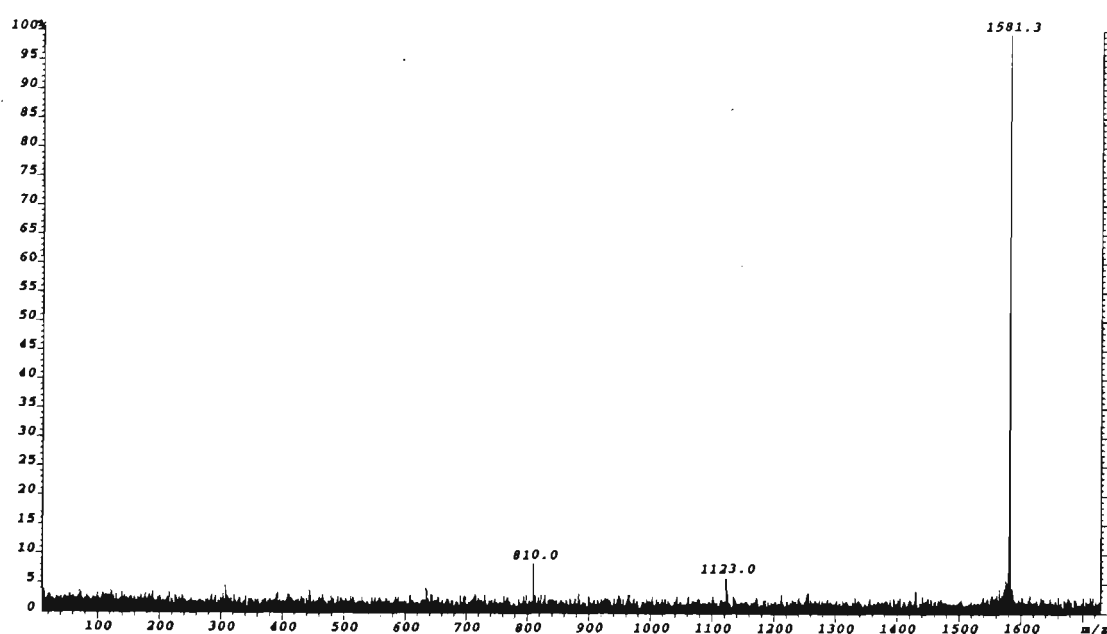
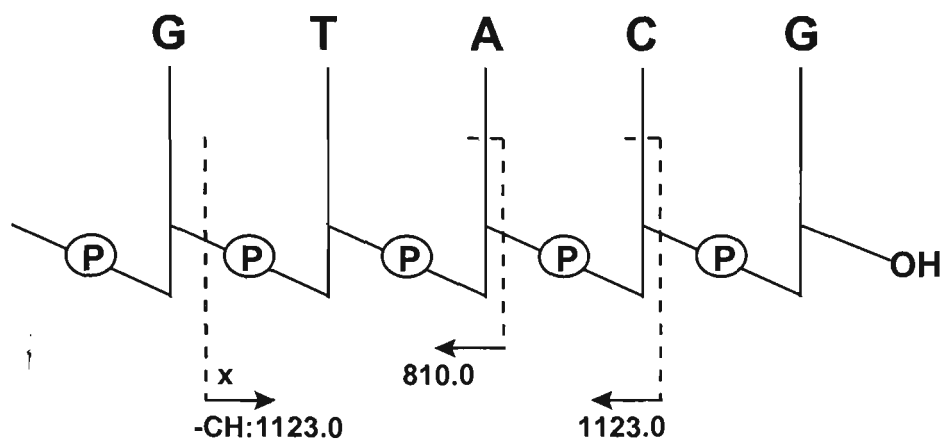


Figure 4.24: ESI-MS/MS spectrum of the w_5^- ion (m/z 1581.3) generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.8: Fragmentation pathways of the ion at m/z 1581.3 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of the ion at m/z 1123.1 is shown in figure 4.25 and the main dissociation pathways generated are represented in scheme 4.9. The MS/MS spectrum of this ion shows an ion at m/z 1043.0 which can be assigned either as $[z_4\text{-CH-2H}]^-$ or as the fragment $[(\text{GTA})+\text{p}+\text{H}_2\text{O-H}]^-$, and the product ions at m/z 793.9 and 775.8 which correspond to $[(\text{TA})+\text{p}+\text{s-H}]^-$ and $[(\text{TA})+\text{p}+\text{s-H}_2\text{O-H}]^-$ respectively. All of these ions could be derived from precursor ions having either of the two structures $[(\text{GTA})+\text{p}+\text{s-H}]^-$ or $[x_4\text{-CH-2H}]^-$, both of which are possible assignments for the m/z 1123.2 as described previously. The presence of the $[(\text{GT})-\text{H}]^-$ ion at m/z 632.2, and the $[(\text{GT})-\text{p}+\text{H}_2\text{O-H}]^-$ ion at m/z 570.1 is, however, consistent with fragmentation of a precursor ion having neighbouring G and T bases, i.e. $[(\text{GTA})+\text{p}+\text{s-H}]^-$. Furthermore, the fact that the ion at m/z 1123.2 is also observed to be relatively intense in the MS/MS spectrum of the $[a_5\text{-B}_5\text{H-2H}]^-$ sequence ion (discussed in more detail below) is consistent with its assignment as $[(\text{GTA})+\text{p}+\text{s-H}]^-$ rather than $[x_4\text{-CH-2H}]^-$. The relatively intense ions at m/z 490.1 and m/z 392.0 correspond to the species $[A_{nt}+\text{p}+\text{s-H}]^-$ and $[A_{nt}+\text{s-H}_2\text{O-H}]^-$ and may arise from w-type and z-type cleavages respectively to the 3' side of thymine.

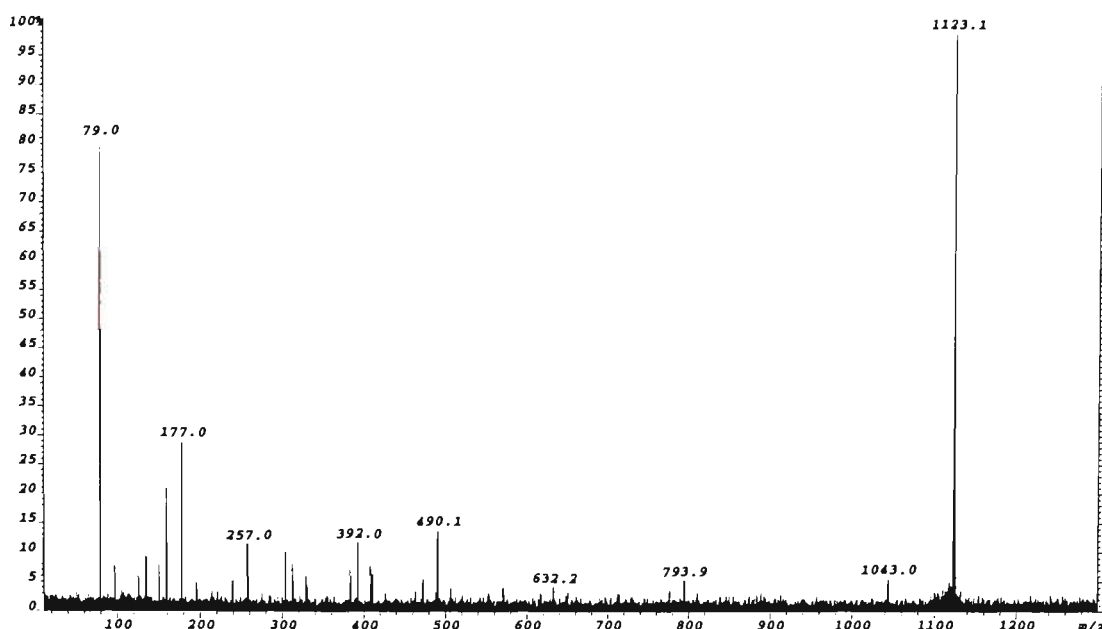
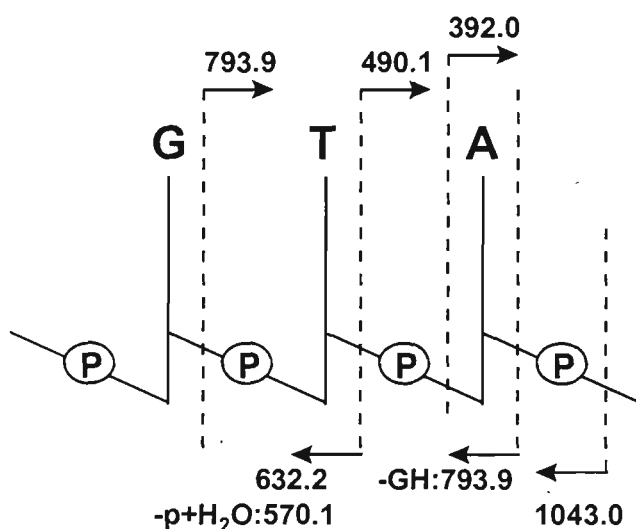


Figure 4.25: ESI-MS/MS spectrum of the ion at m/z 1123.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.9: Fragmentation pathways of the ion at m/z 1123.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of w_5^- (figure 4.24) also shows an extremely weak ion at m/z 1430.1 corresponding to the loss of neutral guanine (i.e. either G-2 or G-6). The MS/MS spectrum of this ion is shown in figure 4.26 and the major fragmentation pathways observed are illustrated in scheme 4.10. The spectrum shows an ion at m/z 810.0 which is also the most intense product ion generated in the MS/MS spectrum

of w_5^- . This ion corresponds to the species $[(GT)+p+s-H]^-$ i.e. the 5'-terminal fragment arising from neutral adenine loss followed by cleavages of the ribose 3' C-O bond. The observation of this ion suggests that it is the 3'-terminal guanine which is predominantly lost from w_5^- to generate this product ion. Consistent with this conclusion is the presence of the ion at m/z 730.1 in the MS/MS spectrum of the $[w_5-GH]^-$ ion which is the 5'-terminal fragment $[(GT)+p+H_2O-H]^-$, generated from z-type cleavage to the 5' side of adenine.

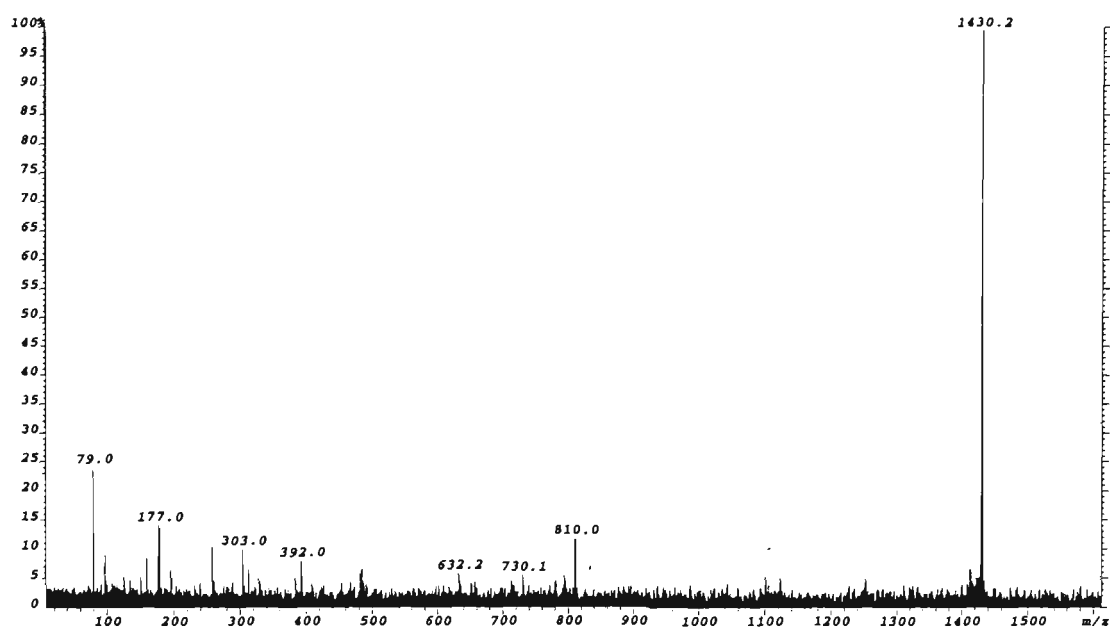
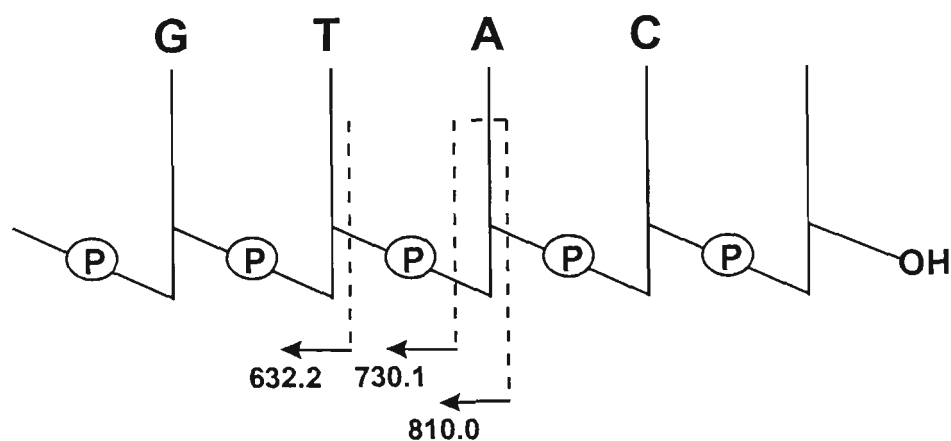


Figure 4.26: ESI-MS/MS spectrum of the ion at m/z 1430.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.10: Fragmentation pathways of the ion at m/z 1430.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

Figure 4.27 shows the MS/MS spectrum of w_4^- at m/z 1252.1 and scheme 4.11 illustrates of the main fragmentation pathways observed in the spectrum.

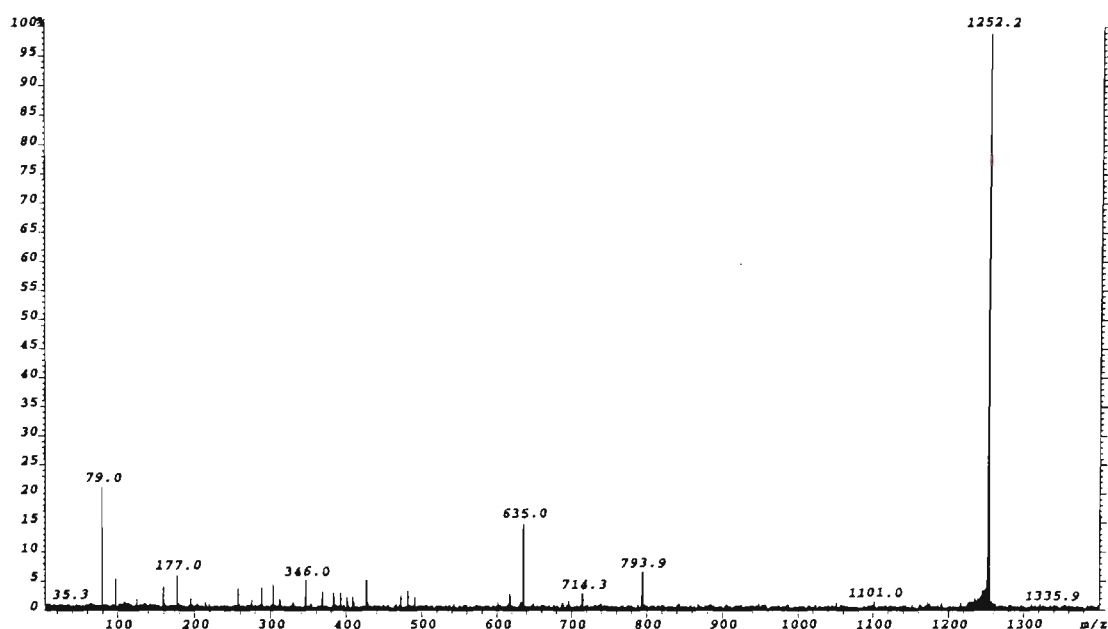
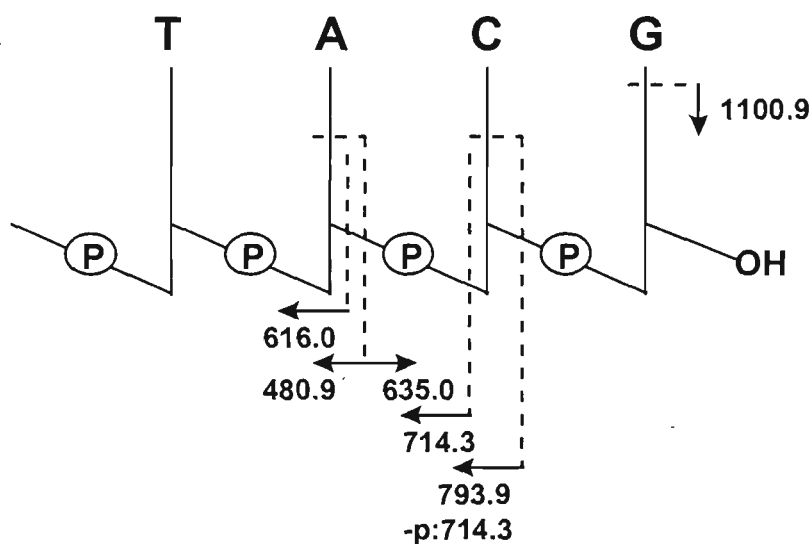


Figure 4.27: ESI-MS/MS spectrum of the w_4^- ion at m/z 1252.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.11: Fragmentation pathways of the ion at m/z 1252.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The spectrum shows a major fragmentation pathway owing to the facile loss of neutral adenine followed by ribose 3' C-O bond cleavage generating w_2^- at m/z 635.0. The complementary 5' fragment from this fragmentation, $[psT+p+s-H]^-$, is also present at m/z 480.9. As was observed with the MS/MS fragmentation of w_5^- , the MS/MS spectrum of w_4^- shows the formation of the w_3^- ion (m/z 947.9) to be strongly

disfavoured. This is again consistent with previous observations suggesting a strong lack of preference for ribose 3' C-O bond cleavage to the 3' side of a thymine containing residue. The ion at m/z 616.0 in the MS/MS spectrum of w_4^- is attributed to $[(TA)-H]^-$ and would appear to be formed from cleavage of the ribose 3' C-O bond to the 3' side of adenine without prior base loss. The relatively intense ion at m/z 793.9 corresponds to the species $[(TA)+p+s-H]^-$ which is the 5' fragment formed from loss of neutral cytosine C-5 from the w_4^- ion followed by ribose 3' C-O bond cleavage. The MS/MS spectrum of this ion, given in figure 4.28, shows formation of the ion at m/z 714. This ion is also observed in the MS/MS spectrum of the w_4^- ion and can arise from either the loss of the 5'-PO₃H group or the 3'-C₅H₄O ribose group from the ion $[(TA)+p+s-H]^-$.

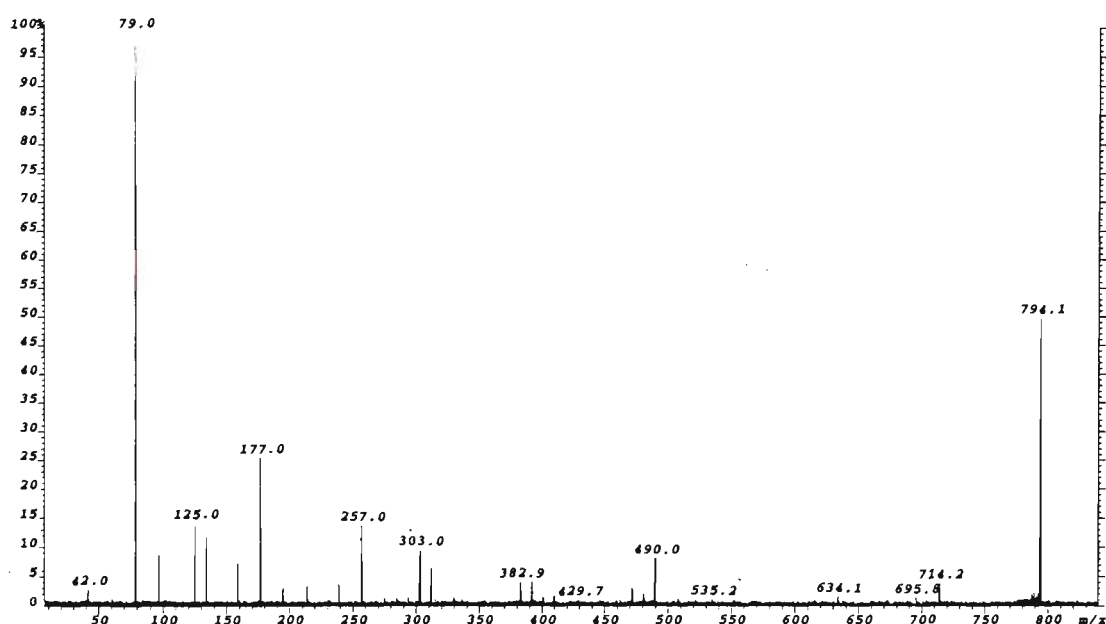


Figure 4.28: ESI-MS/MS spectrum of the ion at m/z 794.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of w_4^- (figure 4.27) also shows a small ion arising from the loss of the 3'-terminal guanine from w_4^- at m/z 1101.0, and the MS/MS spectrum of this ion is shown in figure 4.29. Fragmentation of this ion results in the formation a weak series of ions, illustrated in scheme 4.12, i.e. $[x_4-GH-2H]^-$ at m/z 1083.4, $[y_4-GH]^-$ at m/z 1021.0, and $[z_4-GH]^-$ at m/z 1003.0. The ion observed at m/z 793.7 is also generated from fragmentation of $[a_5-B_5H-2H]^-$ and as described previously, results

from loss of cytosine C-5 followed by cleavage of the ribose 3' C-O bond. The ions at m/z 713.8, 633.9 and 616.0 arise from $[(TA)+p+H_2O-H]^-$, $[(TA)+H_2O-H]^-$, and $[(TA)-H]^-$ respectively and are indicative of fragmentation occurring across the phosphodiester group to the 3' side of adenine. The relatively intense ion at m/z 484.0 arises from the loss of adenine and subsequent cleavage of the ribose 3' C-O bond yielding $[w_2-GH]^-$. The absence of the $[w_3-GH]^-$ ion is once again consistent with the lack of preference for loss of neutral thymine and ribose 3' C-O bond cleavage to the 3' side of thymine.

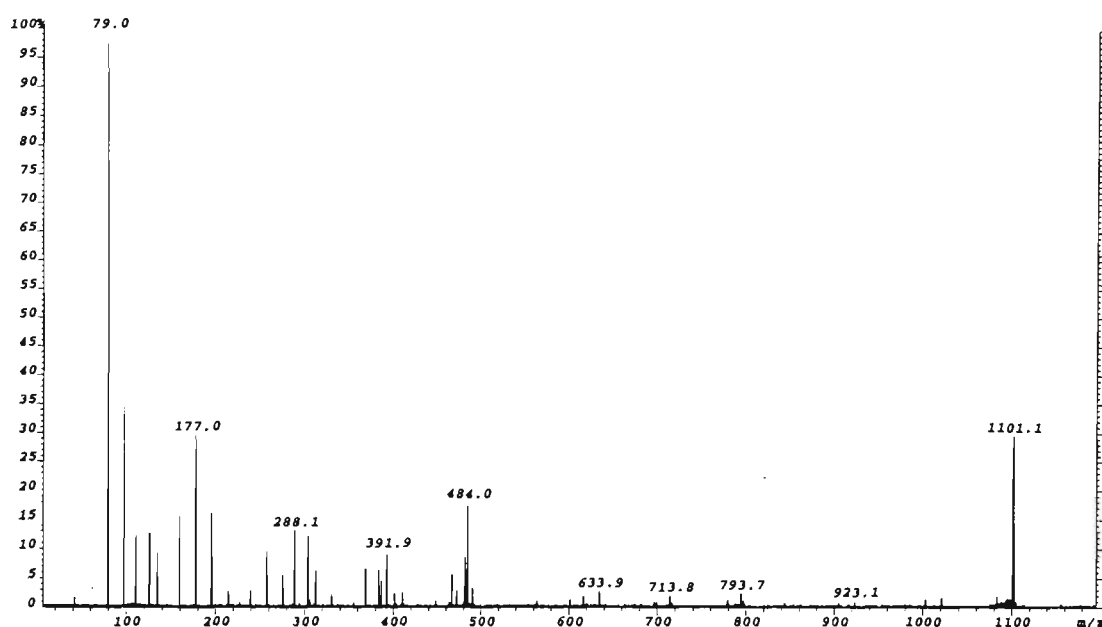
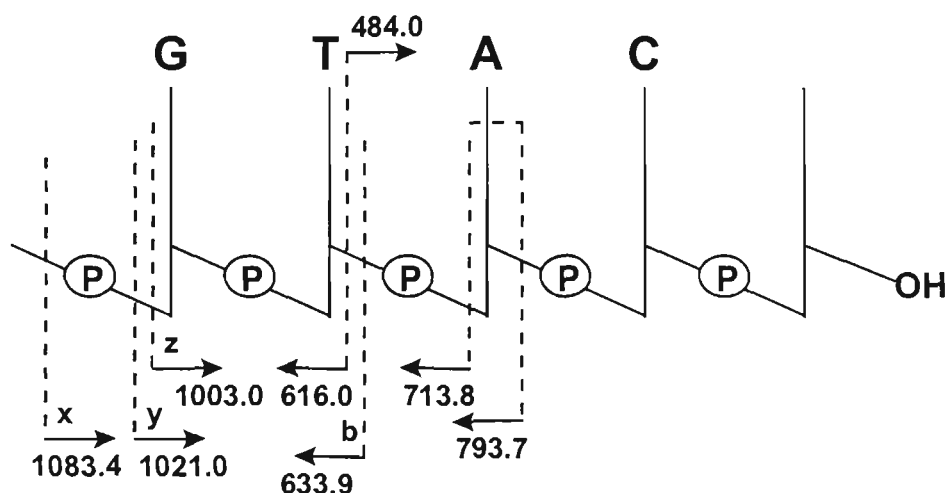


Figure 4.29: ESI-MS/MS spectrum of the ion at m/z 1101.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.12: Fragmentation pathways of the ion at m/z 1101.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of the w_3^- ion (m/z 948.0) is shown in figure 4.30 and the main fragmentation pathways observed are illustrated in scheme 4.13.

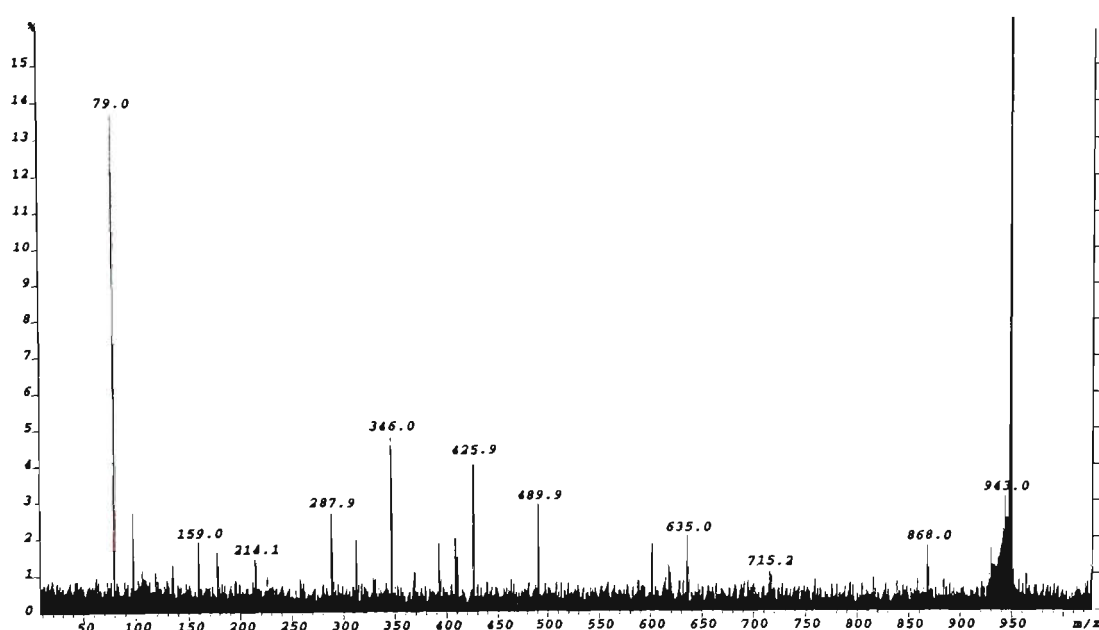
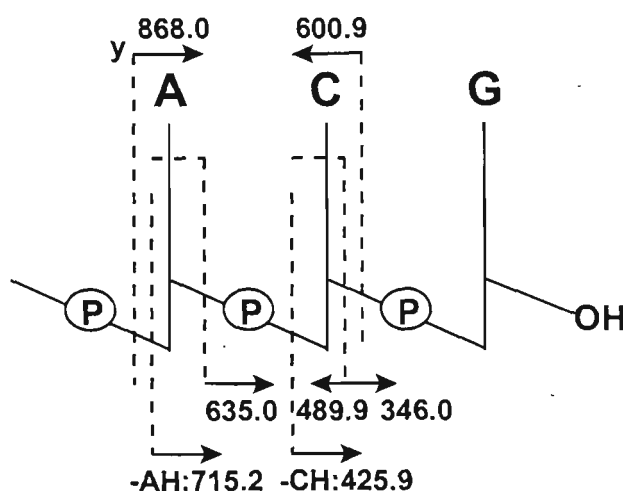


Figure 4.30: ESI-MS/MS spectrum of the w_3^- ion at m/z 948.0 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.13: Fragmentation pathways of the ion at m/z 948.0 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The spectrum is dominated by ions from the loss of neutral cytosine or adenine followed by cleavage of the ribose 3' C-O bond. A relatively intense ion is observed for w_1^- (m/z 346.0) in addition to the complementary 5'-terminal ion $[A_n + s + p - H]^-$ at m/z 489.9 arising from neutral cytosine loss and cleavage of the ribose 3' C-O bond. The w_2^- ion is also present at m/z 635.0. The ion at m/z 600.9 owing to $[(AC) - H]^-$ results from ribose 3' C-O bond cleavage to the 3' side of cytosine without initial loss

of the neutral base. Ions owing to neutral base loss and z-type backbone cleavage are also observed at m/z 425.9 for $[z_2\text{-CH-2H}]^-$ and at m/z 715.2 for $[z_3\text{-AH-2H}]^-$. A small contribution from the y_3^- ion which arises from the loss of the 5'-terminus PO_3H is present at m/z 868.0. Further fragmentation of this ion (appendix 4.3) predominantly results in the w_1^- at m/z 346.0. MS/MS of the w_2^- ion produced a low intensity ion at m/z 554.9 (i.e. y_2^-) although this was not observed in the MS/MS spectrum of the $[\text{M-H}]^-$ ion of 5'-d(CGTACG)-3'. A moderately intense w_1^- ion is also present at m/z 345.9. In the MS/MS spectra of the w_4^- , w_3^- , and w_2^- ions, ions for the loss of neutral guanine are not observed although the loss of neutral guanine is observed to be quite facile in the MS/MS spectrum of 5'-d(CGTACG)-3' and is also observed in low abundance in the MS/MS spectrum of the w_5^- ion. In each case the results from product ion analysis suggest it is the 3'-terminal base that is lost. The absence of neutral guanine loss in the case of these smaller w ions may indicate unfavourable loss of the 3'-terminal base upon decreasing size of the product ion. The ion owing to loss of neutral cytosine, although quite intense in the MS/MS spectrum of the $[\text{M-H}]^-$ ion of 5'-d(CGTACG)-3', was not observed in any of the MS/MS spectra of the w ions and would appear to suggest that the loss of the 5'-terminal cytosine predominantly contributes to the $[\text{M-CH-H}]^-$ ion at m/z 1679.4 observed in the MS/MS spectrum of 5'-d(CGTACG)-3'.

Figure 4.31 is the MS/MS spectrum of the $[a_5\text{-B}_5\text{H-2H}]^-$ ion at m/z 1332.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'. The fragmentation pathways generated are represented in scheme 4.14. The spectrum shows predominantly the internal ions $[A_n+p+s\text{-H}]^-$ at m/z 490.0, $[(\text{TA})+p+s\text{-H}]^-$ at m/z 794.0, and $[(\text{GTA})+p+s\text{-H}]^-$ at m/z 1123.0. These ions correspond to the series of 3' fragments resulting from w-type cleavage to the 3' side of thymine, guanine and cytosine respectively. The facile production of these ions from decomposition of $[a_5\text{-B}_5\text{H-2H}]^-$ would explain their moderate abundance in the MS/MS spectrum of 5'-d(CGTACG)-3'. Furthermore, the fact that the ion at m/z 794.0 is favourably generated from decomposition of w_4^- would additionally contribute to its ion intensity.

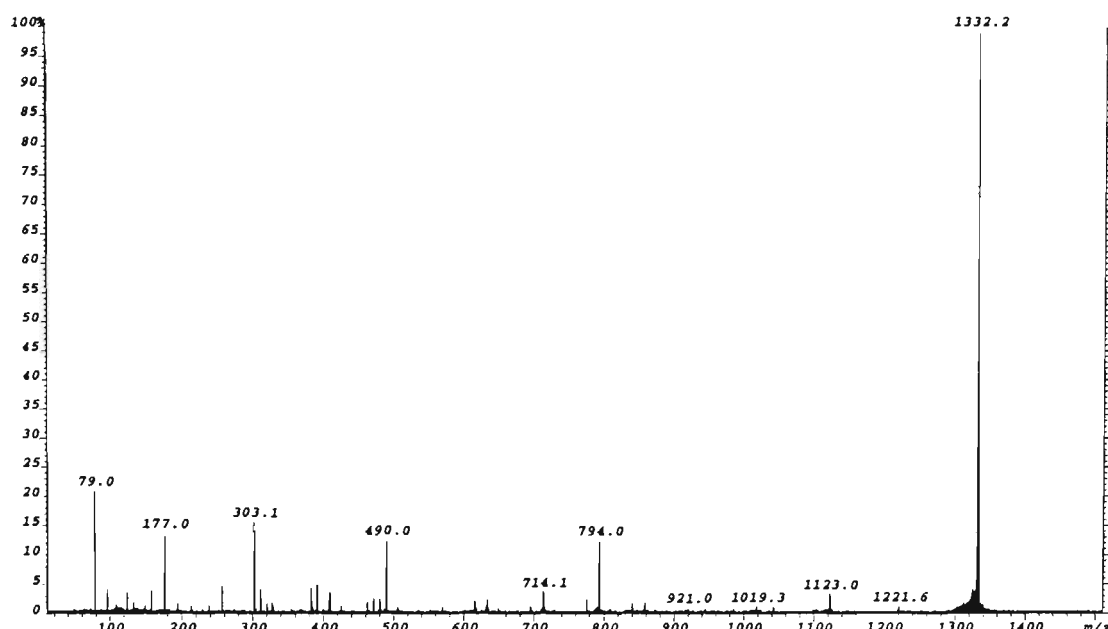
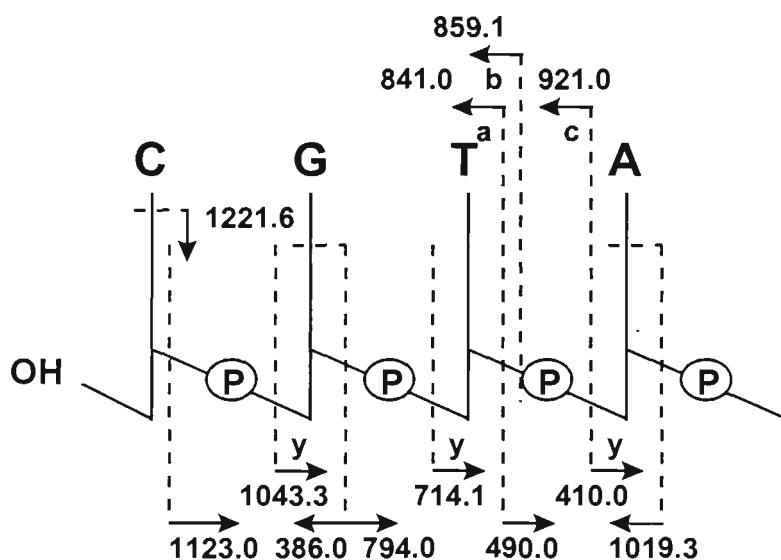


Figure 4.31: ESI-MS/MS spectrum of the $[a_5-B_5H-2H]^-$ ion at m/z 1332.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.14: Fragmentation pathways of the ion at m/z 1332.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

An analogous series of ions are also present owing to y-type cleavage which yields ion at m/z 410.0 for $[A_n+p+H_2O-H]^-$, m/z 714.1 for $[(TA)+p+H_2O-H]^-$, and m/z 1043.3 for $[(GTA)+p+H_2O-H]^-$. An ion for loss of neutral cytosine (C-1) from the precursor ion is also produced at m/z 1221.6. The $[a_4-B_4H-2H]^-$ (m/z 1019.3) and $[a_2-B_2H-2H]^-$ (m/z 386.0) ions are both present, although the $[a_3-B_3H-2H]^-$ ion is not observed which can be explained by the lack of preference for neutral thymine loss. A relatively intense ion for the intact thymine nucleotide at m/z 303.1 is observed in contrast to nucleotide ions

of other bases, which is likely to result from this strong lack of tendency for loss of neutral thymine in comparison to the other bases. The abundant ion for loss of charged thymine (m/z 125.0) in comparison to charged base loss for other nucleobases, can also be explained by the poor competition from the fragmentation channel producing neutral base loss in the case of thymine. Although the MS/MS spectrum of the $[a_3-B_5H-2H]^-$ ion shows fragmentation pathways involving neutral thymine loss to be disfavoured, backbone cleavage of the phosphodiester group to the 3' side of thymine without prior thymine loss yields the ions $[a_3-2H]^-$ at m/z 841.0, b_3^- at m/z 859.1, and $[c_3-2H]^-$ at m/z 921.0. Subsequent fragmentation of these product ions (appendix 4.3) also showed that the loss of neutral thymine was unfavourable.

Figure 4.32 is the MS/MS spectrum of the $[a_3-2H]^-$ ion at m/z 841.2. The spectrum shows predominantly ions from w-type cleavages to the 3' side of G-2 yielding a substantial ion at m/z 303.1 for $[T_{nt}+H_2O-H]^-$. An intense ion is also present at m/z 125.0 for T^- . The same fragmentation pathway as observed for the $[a_3-2H]^-$ ion are also observed in the MS/MS spectrum of b_3^- at m/z 859.4 (appendix 4.3), which, in this case, results in the $[T_{nt}+H_2O-H]^-$ ion at m/z 321.0. Similarly, the most intense ion observed in the MS/MS spectrum of $[c_3-2H]^-$ at m/z 921.0 is the $[T_{nt}+p-H]^-$ ion at m/z 382.9 which arises from w-type cleavage to the 3' side of G-2.

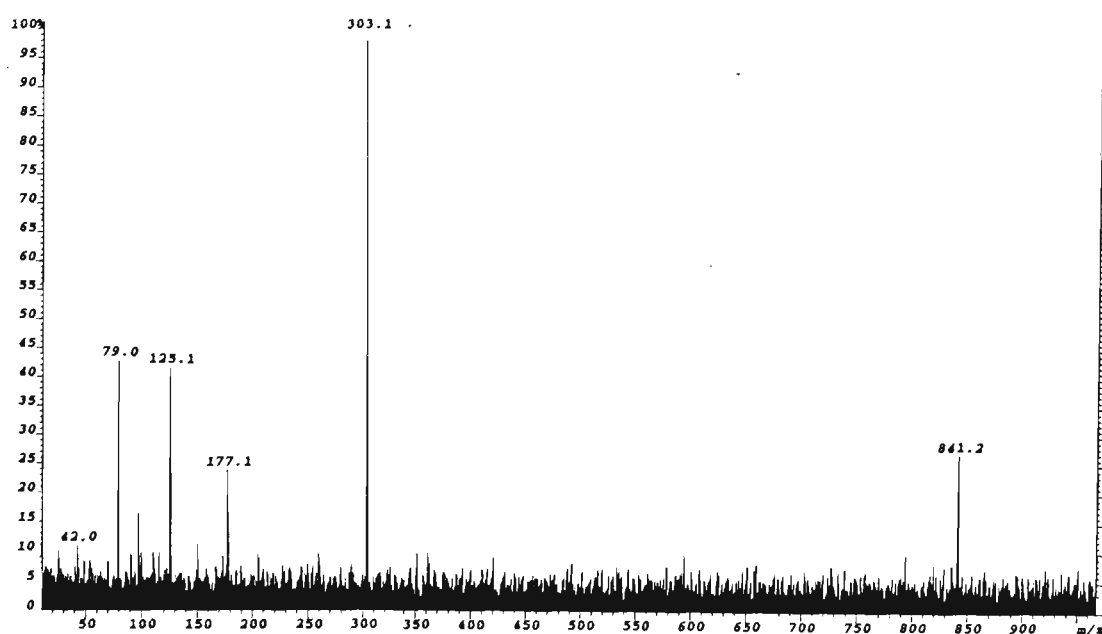


Figure 4.32: ESI-MS/MS spectrum of the $[a_3-2H]^-$ ion at m/z 841.2, generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of the $[a_4-B_4H-2H]^-$ ion at m/z 1019.1 is shown in figure 4.33 and the main fragmentation pathways of this ion are shown in scheme 4.15. The spectrum yields ions from cleavage of the ribose 3' C-O bond to the 3' side of the ribose-containing guanine which generates the $[T_{nt}+p+s-H]^-$ ion at m/z 480.8. An additional w-type cleavage to the 3' side of the T-containing residue of this product ion results in the $[T_{nt}-H]^-$ ion at m/z 303.3, and the complementary fragment ion, (p+s), at m/z 176.9. The ion at m/z 810.1 corresponds to the 3'-terminal fragment arising from cleavage of the ribose C-O bond to the 3' side of cytosine C-1.

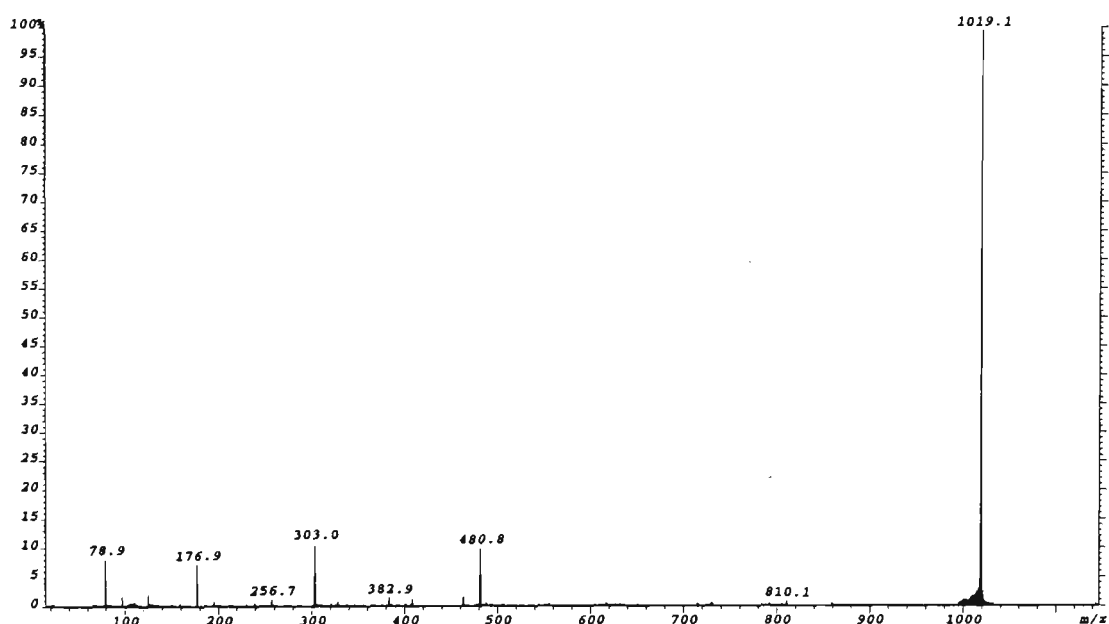
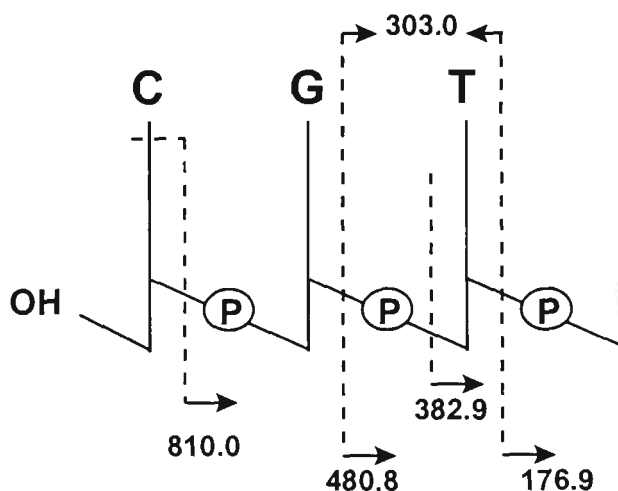


Figure 4.33: ESI-MS/MS spectrum of the $[a_4-B_4H-2H]^-$ ion at m/z 1019.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.15: Fragmentation pathways of the ion at m/z 1019.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of the $[a_3-B_3H-2H]^-$ (m/z 715.1) (shown in figure 4.34) yields fragmentation similar to that generated from the $[a_4-B_4H-2H]^-$ ion in the MS/MS spectrum of 5'-d(CACGTG)-3' resulting in intense ions detected for the backbone fragment ions PO_3^- (m/z 79.0) and $[ps-H]^-$ (m/z 176.9), in addition to ions for C⁻ and G⁻.

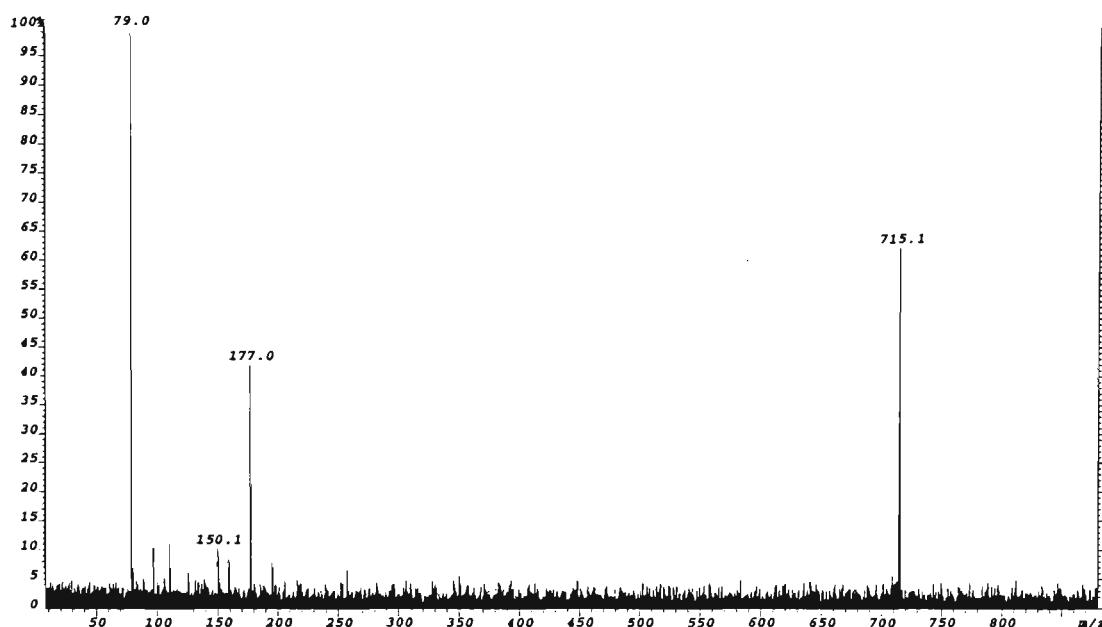


Figure 4.34: ESI-MS/MS spectrum of the $[a_3-B_3H-2H]^-$ ion at m/z 715.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

(c) Summary

From analysis of the MS/MS spectra of the major product ions observed in the MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', it can be seen that the major internal ions present in MS/MS spectra originate from the facile fragmentation pathways occurring upon multiple collisions of the major sequence ions. The MS/MS spectra of the major w ions for both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' showed fragmentation consistent with additional nucleobase loss followed by cleavage of the ribose 3' C-O bond to the 3' side of the ribose from which the base is lost. This dominant fragmentation pathway yielded smaller w ions and ions for the complementary 5' fragments of the structure $[(B..B_n)+p+s-H]^-$, which correspond to internal ions observed in the MS/MS spectra of oligonucleotides. The same predominant fragmentation pathway was also observed to occur from

decomposition of the major (a-B) product ions although ions were only observed for the 3' fragments arising from this pathway generating ions of the structure $[(B..B_n)+p+s-H]^-$ analogous to the 5' fragments generated from the further decomposition of w-types ions. The fact that smaller (a-B) ions were not found to be produced upon further dissociation of (a-B) product ions may be a consequence of location of the negative charge on the phosphate to the 3' side of the cleavage site being favoured. As was observed with sequence ions in MS/MS spectra of oligonucleotides, the formation of product ions from the further decomposition of sequence ions was disfavoured for fragmentation pathways involving the loss of neutral thymine and subsequent ribose 3' C-O bond cleavage to the 3' side of a ribose containing thymine.

4.5 CHARACTERISATION OF POSITIVE ION ELECTROSPRAY TANDEM MASS SPECTRA OF OLIGONUCLEOTIDES

Since the ability to efficiently generate positive oligonucleotide ions from ESI was demonstrated by Loo *et al.*¹⁶, the utility of positive ion MS/MS in the analysis of oligonucleotides is attracting increasing attention. Much of this interest has been driven by the applications of positive ion MS/MS in expanding oligonucleotide sequencing capabilities by enabling complementary structural information to be obtained. Moreover, the ability to effectively analyse oligonucleotides in the positive ion mode may provide a means for probing the interactions of DNA with proteins and ligands. The work described in this section is aimed at investigating the type of fragmentation apparent in the positive ion ESI-MS/MS spectra of oligonucleotides, with particular emphasis being placed on the comparisons that can be drawn between both positive and negative ion MS/MS spectra.

4.5.1 Product Ion Types Observed

The positive ion MS/MS spectra of the singly charged ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are shown in figures 4.35 and 4.36 respectively. Although these spectra pose some similarities to the corresponding ESI-MS/MS spectra of these oligonucleotides obtained in the negative ion mode, there are some marked differences in the types of product ions observed as well as in the relative abundances of certain product ions. In order to confirm the assignments of these product ions, and to gain an insight into the positive ion MS/MS fragmentation observed in the MS/MS spectra of these oligonucleotides, product ions appearing in each spectrum were generated by in-source decomposition and analysed by MS/MS in the same way as the product ions in the negative ion mode. Tables 4.8 and 4.9 list the assignments of the product ions appearing in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' respectively based on these experiments.

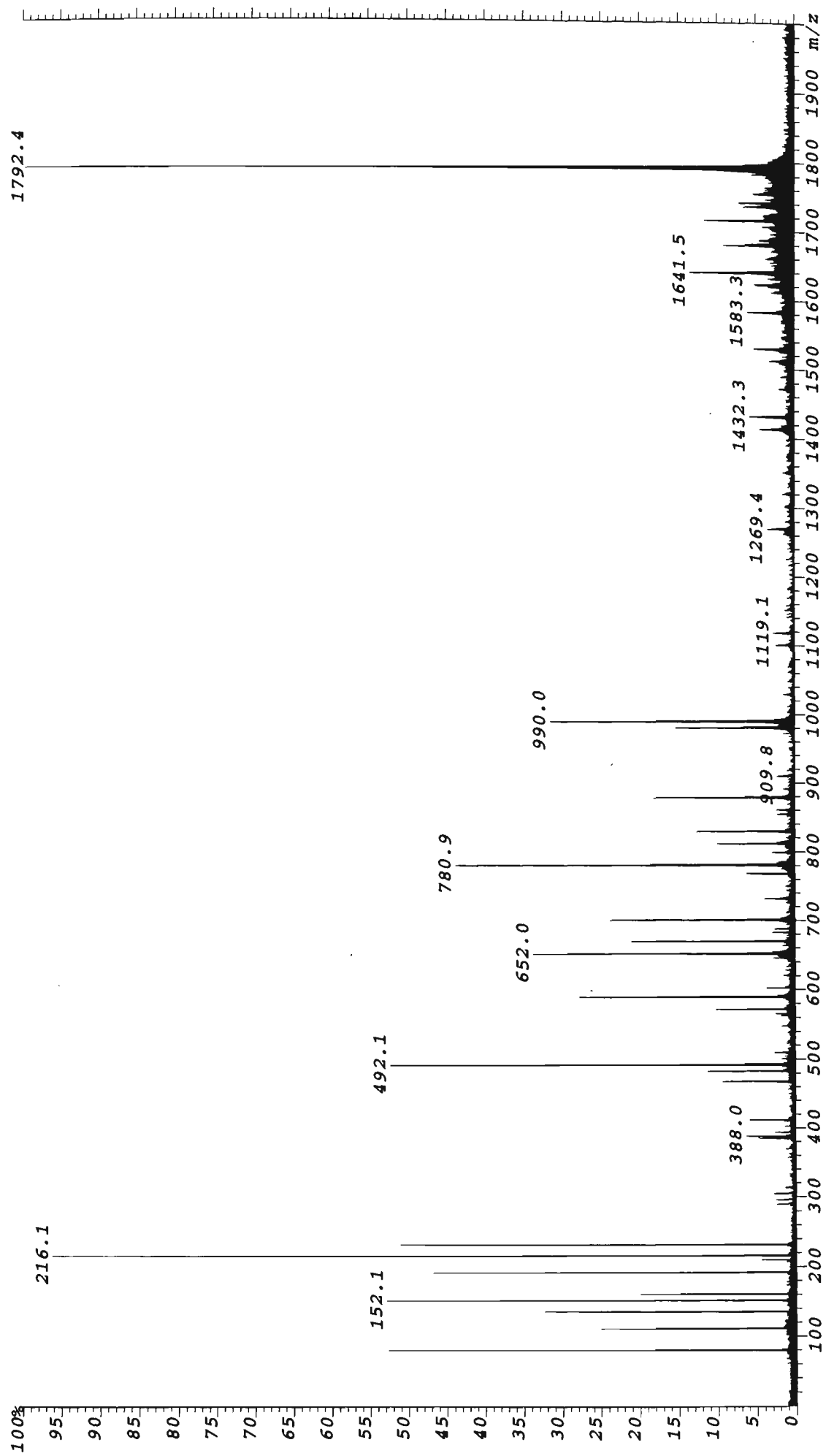


Figure 4.35: Positive ion ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CACGTG)-3' (m/z 1792.4).

Table 4.8: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CACGTG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>		<i>m/z(calc.)</i>	<i>Assignment</i>
1792.4	100		1792.4	$[M+H]^+$
1717.0	11.7		1716.8	$[2M-GH+2H]^{2+}$
1681.3	9.0		1681.3	loss of CH
1641.5	13.5		1641.3	loss of GH
1623.3	5.3		1623.3	loss of GH and H ₂ O
1583.3	6.2		1583.3	$[w_5+2H]^+$
1530.3	5.1		1530.3	loss of CH and GH
1512.3	3.1		1512.3	loss of GH, CH and H ₂ O
1432.3	5.8		1432.2	$[w_5-GH+2H]^+$ and/or $[d_5-CH+2H]^+$
1414.2	4.5		1414.2	$[c_5-CH]^+$ and/or $[x_5-GH]^+$
1269.4	3.8		1270.2	$[w_4+2H]^+$
1119.1	2.9		1119.2	$[w_4-GH+2H]^+$ and/or $[(CGT)+p+s+H_2O+H]^+$
1011.1	2.5		1011.2	$[x_4-GH]^+$ and/or $[(CGT)+p+s+H]^+$
990.0	31.5		990.2	$[a_4-B_4H]^+$
981.1	15.5		981.2	$[w_3+2H]^+$
909.8	2.4		910.2	$[z_4-GH-CH]^+$ and/or $[(GT \text{ or } TG)+s+p+s+H]^+$
			910.1	$[(GT \text{ or } TG)+p+s+p+H_2O+H]^+$
892.1	1.5		892.1	$[(GT \text{ or } TG)+s+p+s-H_2O+H]^+$
		or	892.1	$[(GT)+p+s+p+H]^+$
879.0	18.4		879.1	$[a_4-B_4H-CH]^+$ and/or $[(AC)+s+p+s+H]^+$
			879.1	$[(AC)+p+s+p+H_2O+H]^+$
861.1	2.4		861.1	$[(CA \text{ or } CA)+s+p+s-H_2O+H]^+$
			861.1	$[(AC)+p+s+p+H]^+$
855.1	2.3		855.1	$[a_4-B_4H-AH]^+$
830.1	12.9		830.1	$[w_3-GH+H]^+$ and/or $[(GT)+p+s+H_2O+H]^+$
812.1	10.2		812.1	$[x_3-GH+2H]^+$ and/or $[(GT)+p+s+H]^+$
799.2	2.9		799.1	$[d_3-CH+2H]^+$ and/or $[(AC)+p+s+H_2O+H]^+$
780.9	44.0		781.1	$[c_3-CH]^+$ and/or $[(AC)+p+s+H]^+$
768.1	6.3		768.1	$[a_4-B_4H-CH-CH]^+$ and/or
				$[A_{nt}+s+p+s+p+s+H]^+$
		or	768.1	$[A_{nt}+p+s+p+s+p+H_2O+H]^+$
732.1	3.9		732.1	$[z_3-GH]^+$ of $[(GT)+s+H]^+$
		or	732.1	$[(GT)+p+H_2O+H]^+$
701.1	23.9		701.1	$[a_3-B_3H]^+$
		or	701.1	$[(AC)+p+H_2O+H]^+$
688.1	2.6		688.1	$[d_3-CH-CH+2H]^+$ and/or
				$[A_{nt}+s+p+s+p+H_2O+H]^+$
683.1	3.0		683.1	$[(AC \text{ or } CA)+s-H_2O+H]^+$
		or	683.1	$[(AC)+p+H]^+$
670.1	21.2		670.1	$[c_3-CH-CH]^+$ and/or $[A_{nt}+p+s+p+s+H]^+$
652.0	33.9		652.1	$[w_2+2H]^+$ and/or $[(GT)+H_2O+H]^+$
		or	652.1	$[A_{nt}+p+s+p+s-H_2O+H]^+$
603.1	3.8		603.1	c_2^+ and/or $[(AC)+H]^+$
590.1	27.8		590.1	$[a_3-CH-CH]^+$ and/or $[A_{nt}+s+p+s+H]^+$
		or	590.1	$[A_{nt}+p+s+p+H_2O+H]^+$
572.1	10.2		572.1	$[A_{nt}+s+p+s-H_2O+H]^+$
		or	572.0	$[A_{nt}+p+s+p+H]^+$
566.1	2.5		566.1	$[a_3-AH-CH]^+$ and/or $[C_{nt}+s+p+s+H]^+$
		or	566.0	$[C_{nt}+p+s+p+H_2O+H]^+$
510.1	2.7		510.1	$[d_2-CH+2H]^+$ and/or $[A_{nt}+p+s+H_2O+H]^+$
492.1	52.5		492.1	$[c_2-CH]^+$ and/or $[A_{nt}+p+s+H]^+$
483.1	11.4		483.1	$[x_2-GH]^+$ and/or $[T_{nt}+p+s+H]^+$
468.1	9.4		468.1	$[c_2-AH]^+$ and/or $[C_{nt}+p+s+H]^+$
412.1	6.0		412.1	$[a_2-CH]^+$ and/or $[A_{nt}+s+H]^+$
		or	412.0	$[A_{nt}+p+H_2O+H]^+$
394.1	2.9		394.1	$[A_{nt}+s-H_2O+H]^+$
		or	394.0	$[A_{nt}+p+H]^+$
388.0	6.3		388.1	$[a_2-B_2H]^+$

		or	388.0	$[C_{nt}+p+H_2O+H]^+$
385.1	4.8		385.1	$[T_{nt}+s-H_2O+H]^+$
		or	385.0	$[T_{nt}+p+H]^+$
305.2	3.0		305.1	$[T_{nt}+H]^+$
296.2	2.5		296.1	$[A_{nt}-H_2O+H]^+$
290.1	2.5		290.1	$[C_{nt}+H]^+$
232.1	50.9		232.1	$[sG-H_2O+H]^+$
216.1	96.1		216.1	$[sA-H_2O+H]^+$
210.1	4.71		210.1	$[sC+H]^+$
192.1	46.7		192.1	$[sC-H_2O+H]^+$
161.1	20.4		161.0	$[s+p-H_2O+H]^+$
152.1	52.6		152.1	$[GH+H]^+$
136.0	32.4		136.1	$[AH+H]^+$
112.0	25.1		112.1	$[CH+H]^+$
81.0	52.5		81.0	$[s-H_2O+H]^+$

*calculated monoisotopic mass
s = deoxyribose- H_2O ($C_5H_8O_2$), M_r = 98.0368 Da
p = PO_3H , M_r = 79.9663 Da
 B_{nt} denotes a mononucleotide which may be either psB or sBp
[($B_1...B_n$)] denotes a polynucleotide which may be either (ps $B_1...+psB_n$) or (s $B_1p...+sB_np$)

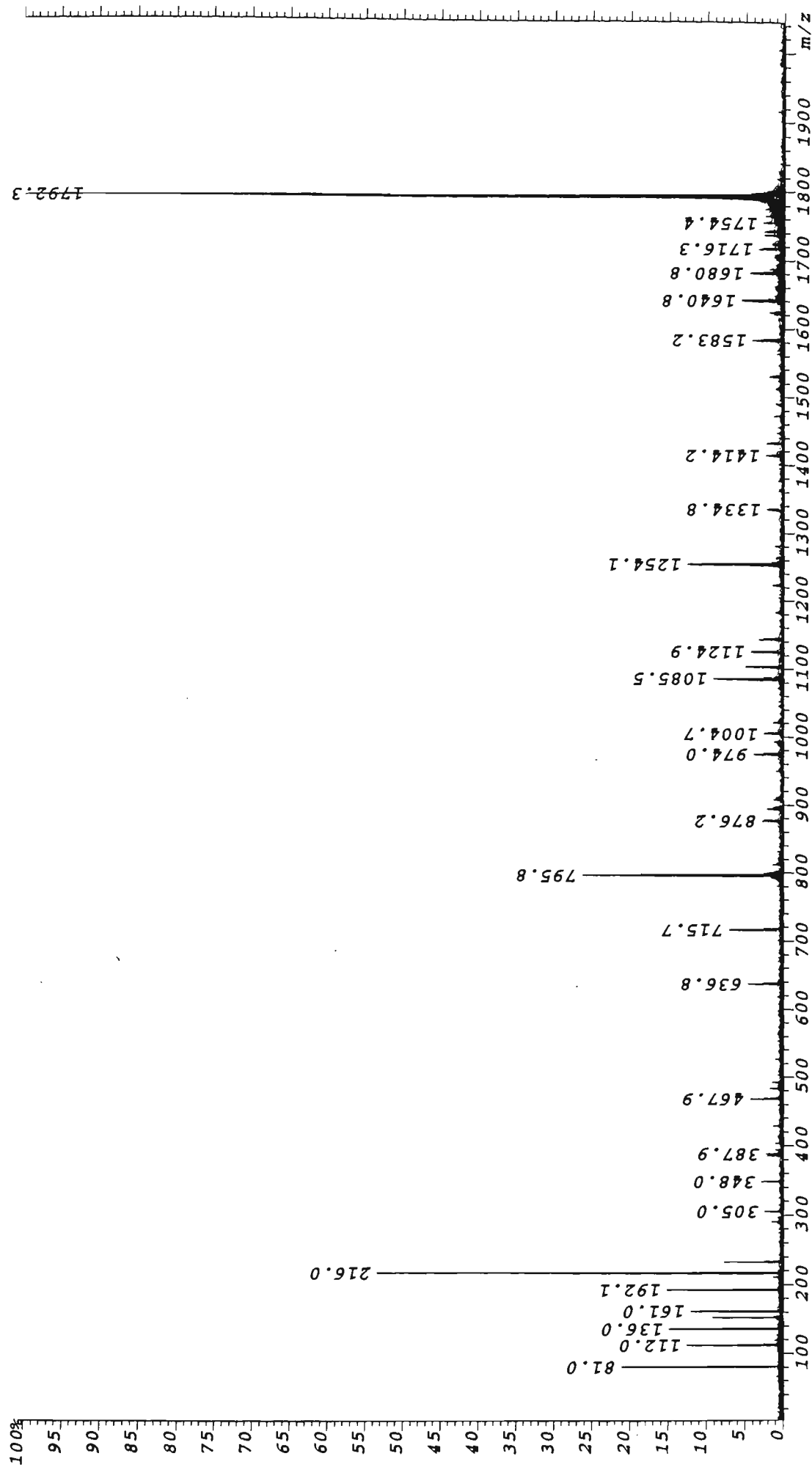


Figure 4.36: Positive Ion ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGTCACG)-3' (m/z 1792.3).

Table 4.9: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc.)</i>	<i>Assignment</i>
1792.3	100	1792.4	$[M+H]^+$
1716.3	3.6	1716.8	$[2M-GH+2H]^2+$
1680.8	4.8	1681.3	loss of CH
1640.8	5.6	1641.3	loss of GH
1623.3	2.0	1623.3	loss of GH and H ₂ O
1583.2	4.2	1583.3	$[w_5+H]^+$
1530.3	1.9	1530.3	loss of CH and GH
1472.1	1.4	1472.2	loss of GH, GH, H ₂ O and/or $[w_5-CH+2H]^+$
1432.3	2.4	1432.2	$[w_5-GH+2H]^+$ and/or $[d_5-CH+2H]^+$
1414.2	2.4	1414.2	$[c_5-CH]^+$ and/or $[x_5-GH]^+$
1334.8	2.4	1334.2	$[a_5-B_5H]^+$ and/or $[z_5-GH]^+$ and/or $[(GTAC)+p+H_2O+H]^+$
1281.0	1.1	1281.1	$[w_5-GH-GH+2H]^+$ and/or $[d_5-CH-GH+2H]^+$
1254.1	12.8	1254.2	$[w_4+2H]^+$ and/or $[d_4+2H]^+$ and/or $[(GTAC)+H_2O+H]^+$
1223.1	1.4	1223.2	$[z_5-CH-GH]^+$ and/or $[a_5-CH-CH]^+$ and/or $[(GTA)+p+s+p+H_2O+H]^+$
1183.0	1.0	1183.2	$[z_5-GH-GH]^+$ and/or $[a_5-CH-GH]^+$ and/or $[(TAC)+p+s+p+H_2O+H]^+$
1143.0	3.5	1143.2	$[w_4-CH+2H]^+$ and/or $[d_4-CH+2H]^+$ and/or $[(GTA)+p+s+H_2O+H]^+$
1124.9	4.4	1125.2	$[x_4-CH]^+$ and/or $[c_4-CH]^+$ and/or $[(GTA)+p+s+H]^+$
1103.1	4.9	1103.2	$[w_4-GH+2H]^+$ and/or $[d_4-GH+2H]^+$ and/or $[(TAC)+p+s+H_2O+H]^+$
1085.5	5.1	1085.2	$[x_4-GH]^+$ and/or $[c_4-GH]^+$ and/or $[(TAC)+s+p+H]^+$
1020.9	1.5	1021.1	$[a_4-B_4H]^+$ and/or $[z_4-AH]^+$ and/or $[(GT)+C_{nt}+s+H]^+$
1004.7	2.5	1005.2	$[a_4-GH]^+$ and/or $[z_4-GH]^+$ and/or $[(TAC)+s+H]^+$
		or 1005.1	$[(TAC)+p+H_2O+H]^+$
992.1	1.2	992.1	$[w_4-GH-CH+2H]^+$ and/or $[d_4-GH-CH+2H]^+$ and/or $[(TA)+p+s+p+s+H_2O+H]^+$
974.0	4.0	974.1	$[x_4-GH-CH]^+$ and/or $[c_4-GH-CH]^+$ and/or $[(TA)+p+s+p+s+H]^+$
893.8	2.3	894.1	$[z_4-GH-CH]^+$ and/or $[a_4-GH-CH]^+$ and/or $[(TA)+s+p+s+H]^+$
		or 894.1	$[(TA)+p+s+p+H_2O+H]^+$
		or 894.1	$[(GT)+p+s+p+H]^+$
876.2	2.7	876.1	$[(TA)+s+p+s-H_2O+H]^+$
		or 876.1	$[(TA)+p+s+p+H]^+$
811.9	1.5	812.1	$[c_3-CH+2H]^+$ and/or $[(GT)+p+s+H]^+$
795.8	26.5	796.1	$[(TA)+p+s+H]^+$
715.7	7.2	716.1	$[(TA)+s+H]^+$
		or 716.1	$[(TA)+p+H_2O+H]^+$
636.8	4.8	637.1	$[w_2+2H]^+$ and/or $[d_2+2H]^+$
618.1	0.89	618.1	$[(TA)+H]^+$
526.1	0.88	526.1	$[w_2-CH+2H]^+$ and/or $[d_2-CH+2H]^+$
		or 526.1	$[G_{nt}+p+s+H_2O+H]^+$
492.1	1.5	492.1	$[A_{nt}+p+s+H]^+$
483.1	1.7	483.1	$[T_{nt}+p+s+H]^+$
467.9	4.4	468.1	$[c_2-GH]^+$ and/or $[C_{nt}+p+s+H]^+$
428.1	1.2	428.1	$[z_2-CH]^+$ and/or $[a_2-CH]^+$ and/or $[G_{nt}+s+H]^+$
		428.0	$[G_{nt}+p+H_2O+H]^+$
403.1	0.97	403.1	$[T_{nt}+s+H]^+$
		403.0	$[T_{nt}+p+H_2O+H]^+$
394.1	0.97	394.1	$[A_{nt}+s-H_2O+H]^+$

		or	394.0	$[A_{nt}+p+H]^+$
387.9	2.4		388.1	$[a_2-B_2H]^+$
		or	388.0	$[C_{nt}+p+H_2O+H]^+$
385.1	2.3		385.1	$[T_{nt}+s-H_2O+H]^+$
		or	385.0	$[T_{nt}+p+H]^+$
348.0	2.9		348.1	$[w_1+2H]^+$ and/or $[G_{nt}+H_2O+H]^+$
305.2	2.5		305.1	$[T_{nt}+H]^+$
290.1	1.5		290.0	$[C_{nt}+H]^+$
232.0	7.3		232.1	$[sG-H_2O+H]^+$
216.0	53.3		216.1	$[sA-H_2O+H]^+$
210.0	1.5		210.1	$[sC+H]^+$
192.1	16.0		192.1	$[sC-H_2O+H]^+$
161.0	12.1		161.0	$[s+p-H_2O+H]^+$
152.0	9.1		152.1	$[GH+H]^+$
136.0	15.2		136.1	$[AH+H]^+$
112.0	12.8		112.1	$[CH+H]^+$
81.0	21.5		81.0	$[s-H_2O+H]^+$

*calculated monoisotopic mass
s = deoxyribose- H_2O ($C_5H_6O_2$), M_r = 98.0368 Da
p = PO_3H , M_r = 79.9663 Da
 B_{nt} denotes a mononucleotide which may be either psB or sBp
[[$B_1...B_n$]] denotes a polynucleotide which may be either (ps $B_1...+psB_n$) or (s $B_1p...+sB_np$)

(a) Backbone and Nucleotide Fragments

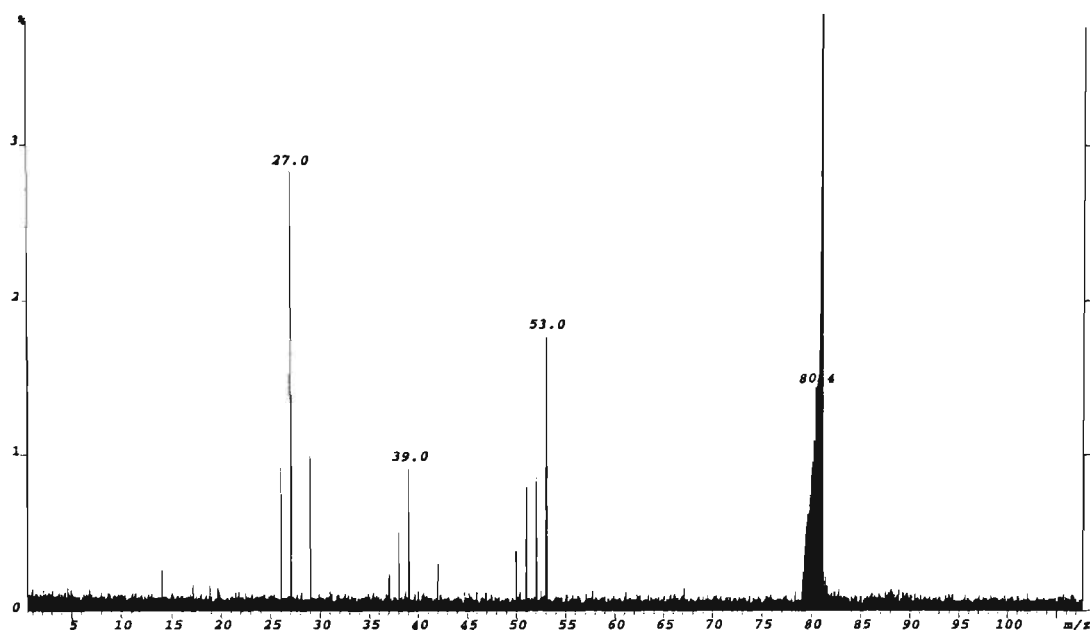
Table 4.10 summarises the types of ions arising from backbone and nucleotide fragments observed in the low m/z region of the positive ion MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. Whereas the MS/MS spectra acquired in the negative ion mode showed a series of backbone fragments comprising phosphate and sugar groups, the MS/MS spectra acquired in the positive ion mode only yields two types of backbone fragment ions at m/z 81 and at m/z 161. The MS/MS spectra of these ions generated from the in-source collisional activation of 5'-d(CGTACG)-3' are shown in figures 4.37(a) and (b) respectively.

Table 4.10: Nucleotide product ion types generated in the positive ion MS/MS spectra of oligonucleotides.

Product Ion Types of Nucleotide Fragments			
Backbone fragments*	m/z	Nucleoside fragments	m/z
$[s-H_2O+H]^+$	81	$[CH+H]^+$	112
$[p+s-H_2O+H]^+$	161	$[AH+H]^+$	136
		$[GH+H]^+$	152
		$[sC-H_2O+H]^+$	192
		$[sC+H]^+$	192
		$[sA-H_2O+H]^+$	216
		$[sG-H_2O+H]^+$	232

*where p= PO_3H , s=deoxyribose- H_2O ($C_5H_6O_2$)

(a)



(b)

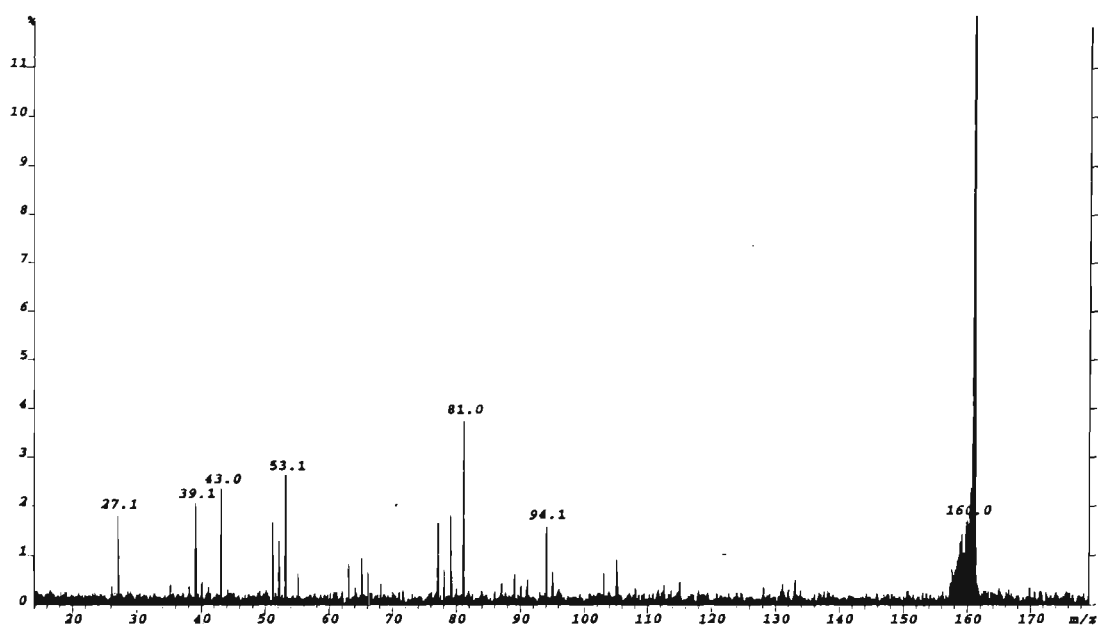


Figure 4.37: ESI-MS/MS spectra of the product ions at (a) m/z 81.0 and (b) m/z 161.0 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The MS/MS spectrum of the m/z 81.0 ion shows product ions representative of the cleavage of a furan ring structure. The most intense ion in the spectrum at m/z 27.0 may correspond to the ion, $C_2H_3^+$. The ion at m/z 26.0 may be attributed to $C_2H_2^+$, and the ion at m/z 29.0 would seem to arise from HCO^+ . The ion at m/z 53.0 may arise from loss of CO from the precursor ion, with the series of ions at m/z 52.0, m/z 51.0 and

m/z 50.0 owing to the respective neutral losses of HCO , CH_2O , and CH_3O . The ion at m/z 39.0 corresponds to the species C_3H_3^+ which would result from elimination of CH_2CO . The less intense ions at m/z 38.0 and m/z 37.0 may be owing to losses of the neutral species $\text{C}_2\text{H}_3\text{CO}$ and $\text{C}_2\text{H}_4\text{CO}$ respectively. Therefore from the fragmentation observed in the MS/MS spectrum of the product ion at m/z 81.0 (figure 4.37(a)), the structure of this ion would appear to be consistent with that of the sugar ion, $\text{C}_5\text{H}_5\text{O}^+$ ($[\text{s-H}_2\text{O-H}]^+$) ($m/z(\text{calc.})=81.034$), and not from a protonated phosphate, H_2PO_3^+ as reported by Ni *et al.*¹⁷ which would also give rise to a m/z value of 81 ($m/z(\text{calc.})=80.974$). Hence, these results support the findings of Phillips and McCloskey¹⁸ in their work on the collision-induced dissociation of dinucleotide monophosphates. The MS/MS spectrum of the ion at m/z 161.0 (figure 4.37(b)) gives rise to the ion at m/z 81.0 in addition to the product ions arising from its fragmentation. The ion at m/z 94.1 corresponds to the mass of the ion $\text{HPO}_3\text{CH}_2^+$ which would be an observation consistent with the dissociation of the backbone fragment, $[\text{p+s-H}_2\text{O+H}]^+$ ($m/z(\text{calc.})=161.0$). The absence of any longer backbone fragments in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' analogous to the ion types observed in the negative ion mode would appear to support the expected preference for charge location in protonated oligonucleotides as residing on the nucleobases rather than on the phosphate groups.

From the list of the types of nucleobase ions generated in positive ion MS/MS spectra (table 4.10), it can be seen that protonated base ions are present for C, A, and G, and no ion is detected for protonated thymine in either spectrum. Similarly, nucleoside ions of the type $[\text{sB-H}_2\text{O+H}]^+$ are detected for C, A, and G with the absence of an ion of this type for T. MS/MS fragmentation of these ions (appendix 4.4) yields an intense ion for $[\text{s-H}_2\text{O+H}]^+$ at m/z 81 with a very weak signal present for the respective base ion. The lack of observation of nucleobase and nucleoside ion types for thymine is in agreement with the values estimated for the proton affinities of the nucleobases by Rogers *et al.*¹⁹ using semiempirical calculations: C (225.7 kcal/mol) > G (222.1 kcal/mol) > A (220.7 kcal/mol) > T (203.5 kcal/mol). The large disparity

between the proton affinity of thymine compared with that of the other three nucleobases C, A and G predicts that the location of the single charge on the protonated oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' is at either G, A, or C. The MS/MS spectra obtained here do not show clear trends for the relative intensities of the other three protonated bases with the MS/MS spectrum of 5'-d(CACGTG)-3' showing an order of $G > A > C$, compared with $A > C > G$ in the MS/MS spectrum of 5'-d(CGTACG)-3'. These differences probably reflect the marginal differences in the values for the proton affinities of C, A and G. The two oligonucleotides examined have the same composition but different sequence which suggests that intramolecular interactions could alter these proton affinity values sufficiently so as to cause a change in the overall preference for observed base loss.

The types of nucleotide product ions common to the MS/MS spectra of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are summarised in table 4.11 and the MS/MS spectra confirming these assignments are given in appendix 4.5. In contrast to the MS/MS spectra obtained in negative ion mode which showed a series of nucleotide ion types for each constituent base, the formation of these ion types in the positive ion MS/MS spectra is largely influenced by the identity of the attached base with certain ion types completely absent from the spectra. It is interesting that the MS/MS spectra of both oligonucleotides show nucleotide ion types predominantly for thymine followed by cytosine, and then adenine, with no nucleotide ions for guanine observed to be common to both spectra. Some nucleotide ion types were found to be present in the MS/MS spectrum of one oligonucleotide and not the other. For instance, although no adenine nucleotide ion types are observed in the MS/MS spectrum of 5'-d(CGTACG)-3', the MS/MS spectrum of 5'-d(CACGTG)-3' shows a series of additional nucleotide ion types at m/z 296.2 for $[A_{nt}-H_2O+H]^+$, 412.1 for $[A_{nt}+p+H_2O+H]^+$ or $[A_{nt}+s+H]^+$, m/z 510.3 for $[A_{nt}+p+s+H_2O+H]^+$, 572.1 for $[A_{nt}+p+s+p+H]^+$ or $[A_{nt}+s+p+s-H_2O+H]^+$, m/z 590.1 for $[A_{nt}+p+s+p+H_2O+H]^+$ or $[A_{nt}+s+p+s+H]^+$, m/z 652 for $[A_{nt}+p+s+p+s-H_2O+H]^+$, m/z 670 for $[A_{nt}+s+p+s+p+H]^+$, and at m/z 768 for $[A_{nt}+p+s+p+s+p+H_2O+H]^+$ or $[A_{nt}+s+p+s+p+s+p+H]^+$ (the MS/MS

spectra of these ions are given in appendix 4.6). Furthermore, whereas nucleotide ions for guanine are absent in the MS/MS spectrum of 5'-d(CACGTG)-3', a few ions are detected in the MS/MS spectrum of 5'-d(CGTACG)-3' at m/z 348.1 for $[G_{nt}+H_2O+H]^+$, at m/z 427.9 for $[G_{nt}+p+H_2O+H]^+$ or $[G_{nt}+s+H]^+$, and at m/z 526.3 for $[G_{nt}+p+s+H_2O+H]^+$ (the MS/MS spectra of these ions are given in appendix 4.7). These results would seem to indicate that the types of product ions formed from backbone cleavages in the MS/MS fragmentation of oligonucleotides in the positive ion mode is primarily influenced by base identity.

Table 4.11: Mononucleotide ion types generated in the positive ion MS/MS spectra of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

Mononucleotide ions:	Product Ion Types of Nucleotides*			
	C _{nt}	T _{nt}	A _{nt}	G _{nt}
	290	305	-	-
+p or +(s-H ₂ O)	-	385	394	-
+s or +(p+H ₂ O)	388	403	-	-
+p+s	468	483	492	-

*where p=PO₃H, s=deoxyribose-H₂O (C₅H₆O₂)

(b) Polynucleotide Ion Types

In negative ion MS/MS spectra, a predominant series of w and (a-B) sequence ions were observed for the fragmentation pathway involving nucleobase loss followed by cleavage of the ribose C-O bond to the 3' side of the residue from which the base is lost. In addition, numerous other polynucleotide ion types were also present at generally much lower relative abundance with various attachments of backbone fragments. From multiple stage mass spectrometry analysis of these ions, it would appear that these ions arise from the further dissociation of the major ion types owing to multiple collisions. The positive ion MS/MS spectra of 5'-d(CACGTG)-3' (figure 4.35) and 5'-d(CGTACG)-3' (figure 4.36) yield very different fragmentation patterns, with some intense ions present corresponding to w and (a-B) sequence ions, in addition

to numerous other types of intense backbone fragments arising from other fragmentation pathways. The MS/MS spectra of the polynucleotide product ions observed in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are given in appendices 4.8 and 4.9. In the MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CACGTG)-3', ions are observed for all w-type cleavages except for ribose C-O bond cleavage to the 3' side of thymine which would generate $[w_1+2H]^+$. The formation of $[w_2+2H]^+$ (m/z 652.0) yields the most intense w ion followed by $[w_3+2H]^+$ (m/z 981.1), $[w_5+2H]^+$ (m/z 1583.3), and $[w_4+2H]^+$ (m/z 1269.4). Similarly, in the case of (a-B) ion formation, the complete series is observed except for $[a_5-B_5H]^+$ which requires the loss of thymine followed by ribose 3' C-O bond cleavage. The most abundant of the series is the $[a_4-B_4H]^+$ ion (m/z 990.0) followed by $[a_3-B_3H]^+$ (m/z 701.1) and $[a_2-B_2H]^+$ (m/z 388.1). The MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGTACG)-3' (figure 4.36) shows dominant series of w and (a-B) ions with ions of these series observed for all cleavages except for $[w_3+2H]^+$ and $[a_3-B_3H]^+$ which are the complementary pair of ions formed from loss of thymine T-3 followed by cleavage of the adjacent ribose 3' C-O bond. The most intense w ion generated is $[w_4+2H]^+$ (m/z 1254.1) followed by $[w_2+2H]^+$ (m/z 636.8), $[w_5+2H]^+$ (m/z 1583.2), and $[w_1+2H]^+$ (m/z 348.0). The formation of $[a_2-B_2H]^+$ (m/z 387.9) and $[a_5-B_5H]^+$ (m/z 1334.8) yields the most abundant (a-B) ions followed by $[a_4-B_4H]^+$ (m/z 1020.9). Once again, the lack of observation of cleavage products arising from the initial loss of neutral thymine in the MS/MS spectra of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' can be rationalised on the basis of both the relatively low proton affinity of thymine in addition to its poor ability to act as a good leaving group in comparison to the other nucleobases¹⁹. Consequently, these factors which make the loss of thymine an energetically unfavourable event would also affect the formation of the w-type and (a-B)-type cleavage products.

The MS/MS spectra of both oligonucleotides also show moderately abundant ions for the loss of neutral cytosine and guanine from the precursor ion. These losses are quite a distinct observation throughout the entire region of the MS/MS spectra,

with the majority of sequence and internal ions formed showing a series of ions arising from the loss of up to two cytosine and guanine nucleobases from the fragment. The facile loss of cytosine and guanine is also a general observation in the MS/MS spectra acquired in the negative ion mode and may be attributed to the stability of neutral cytosine and guanine as leaving groups in comparison to adenine and thymine¹⁹ thus making their loss more facile (discussed in more detail below). A common theme observed in the types of ions generated in the positive ion mode was the predominant observation of distinct series of individual nucleotide or polynucleotide ions arising from consecutive additions of sugar and phosphate backbone moieties. In the MS/MS spectrum of 5'-d(CACGTG)-3', a host of product ions are generated owing to the A mononucleotide with various portions of backbone fragments attached. A predominant series is also present for the CA dinucleotide. Less significant series are generated for the (GT) dinucleotide. The MS/MS spectrum of 5'-d(CGTACG)-3' shows a significant ion series incorporating the (TA) dinucleotide, and less predominant series are also observed for the (TAC) and (GTA) trinucleotides. A summary of these ion series is given in table 4.12. The origin of formation of the predominant series of nucleotide and polynucleotide ions in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' can be derived from analysis of the dissociation pathways of product ion types observed in MS/MS spectra of fragment ions generated in the source.

Table 4.12: Series of mono- and polynucleotide product ion types generated in the positive ion MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

Product Ion Types	5'-d(CACGTG)-3'			5'-d(CGTACG)-3'		
	A _{nt}	(CA)	(GT)	(TA)	(TAC)	(GTA)
	-	603	-	618	-	-
+p or +(s-H ₂ O)	412	683	-	-	-	-
+s or +(p+H ₂ O)	-	701	732	716	1005	-
+p+s	492	781	812	796	1086	1124
+p+s+H ₂ O	510	799	-	-	1103	1143
+p+s+p or +s+p+s-H ₂ O	572	861	892	876	-	1223
+p+s+p+H ₂ O or +s+p+s	590	879	910	894	1183	-
+p+s+p+s-H ₂ O	652	-	-	-	1281	-
+p+s+p+s	670	-	-	974	-	-
+p+s+p+s+p+H ₂ O or +s+p+s+p+s	768	-	-	993	-	-

*where p=PO₃H, s=deoxyribose-H₂O (C₅H₆O₂)

4.5.2 Analysis of Product Ion Fragmentation Pathways

(a) 5'-d(CACGTG)-3'

The MS/MS spectrum of 5'-d(CACGTG)-3' shows an intense ion owing to loss of neutral guanine at m/z 1641.5. The MS/MS spectrum of this ion is shown in figure 4.38, with the main fragmentation pathways observed illustrated in scheme 4.16.

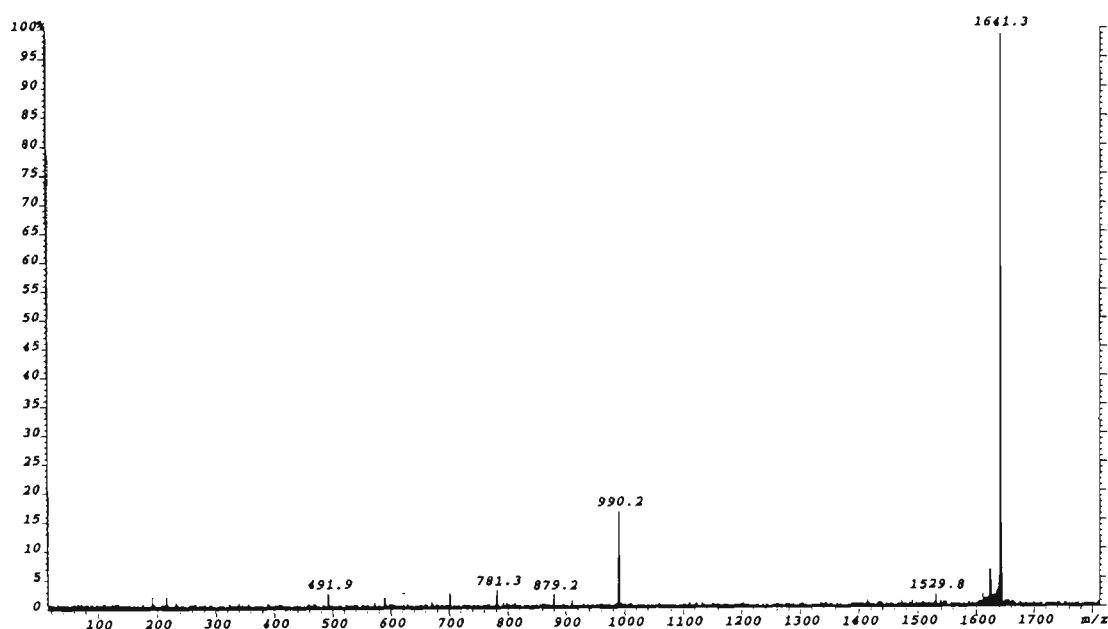
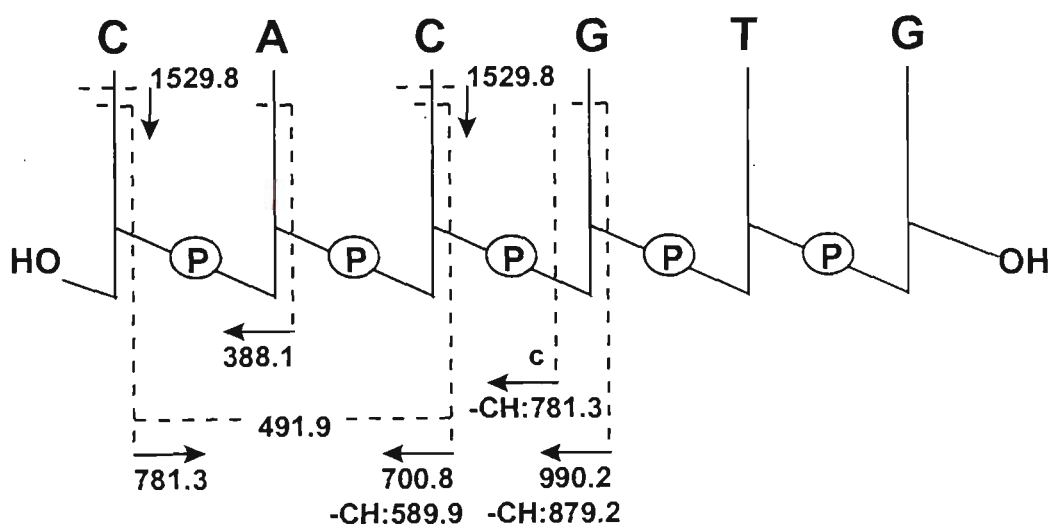


Figure 4.38: ESI-MS/MS spectrum of the ion at m/z 1641.3 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.16: Proposed fragmentation pathways of the ion at m/z 1641.3 arising from loss of guanine generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Fragmentation of this ion gives rise to an intense ion at m/z 990.2 corresponding to $[a_4-B_4H]^+$ with a less intense ion present at m/z 700.8 for $[a_3-B_3H]^+$. Ions are also

observed owing to the loss of neutral cytosine from these ions at m/z 879.2 and m/z 589.9 respectively. The ion at m/z 781.3 may arise from further cleavage of the $[a_4-B_4H]^+$ ion at the ribose C-O bond to the 3' side of cytosine C-1. Similar fragmentation of the $[a_3-B_3H]^+$ ion may also be responsible for the ion at m/z 491.9. Alternatively the ion at m/z 781.3 may also be attributed to $[c_3-CH]^+$. A small ion is also detected owing to the loss of neutral cytosine from the precursor ion at m/z 1529.8. It was not possible to distinguish between the loss of either G-4 or G-6 based on the MS/MS spectrum, hence loss of guanine is not shown in the scheme.

The MS/MS spectrum of the $[a_4-B_4H]^+$ ion at m/z 990.2, shown in figure 4.39, gives rise to the ions at m/z 879.2, m/z 780.9, m/z 590.0 and m/z 492.1 similar to those observed in the MS/MS spectrum of the product ion for neutral guanine loss (figure 4.38). The ion detected at m/z 670.0 corresponds to the species $[A_{nt}+p+s+p+s+H]^+$ and may arise from fragmentation of the $[a_4-B_4H]^+$ ion via ribose 3' C-O bond cleavage adjacent to cytosine C-1. The spectrum also shows an ion at m/z 388.1 for $[a_2-B_2H]^+$. The intense ion at m/z 216.1 for $[A+s-H_2O+H]^+$ can be generated from the further decomposition of any of the product ions observed in the spectrum, since this ion is generated in the MS/MS spectra of each of these product ions (see appendix 4.5). These major fragmentation pathways are shown in scheme 4.17.

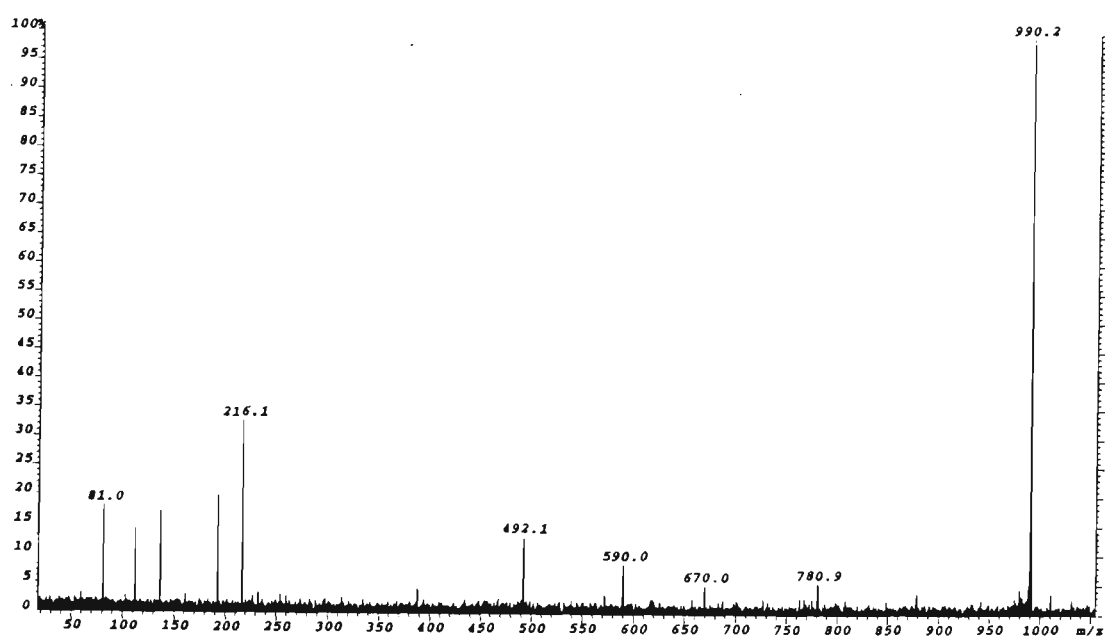
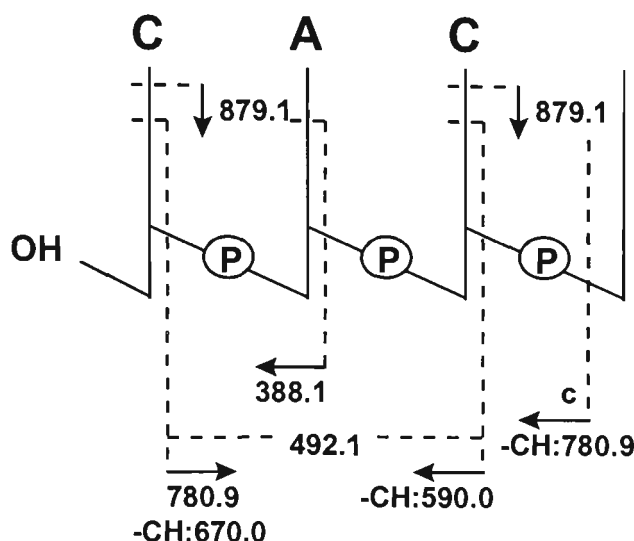


Figure 4.39: ESI-MS/MS spectrum of the ion at m/z 990.2 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.17: Proposed fragmentation pathways of the ion at m/z 990.2 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The ion at m/z 780.9 is observed to be relatively intense in the MS/MS spectrum of 5'-d(CACGTG)-3' and may arise from either of the structures $[(AC)+p+s+H]^+$ or $[c_3-CH]^+$. The MS/MS spectrum of this ion, shown in figure 4.40, gives rise predominantly to ions at m/z 81.0 for $[s-H_2O+H]^+$ and at m/z 216.1 for $[A+s-H_2O+H]^+$. Loss of neutral cytosine from this ion yields an ion at m/z 669.8. If the ion at m/z 780.9 is of the structure $[(AC)+p+s+H]^+$, the formation of the ion at m/z 492.0 may arise from subsequent cleavage of the adjacent ribose 3' C-O bond following loss of C-3. The ion at m/z 572.1 may be produced from either z-type cleavage at the 5'-terminus of this fragment, or c-type cleavage at the 3'-terminus. Corresponding to this, the small ion at m/z 590.1 may be attributed either to y-type cleavage at the 5'-terminus, or to d-type cleavage at the 3'-terminus. The fragmentation pathways are shown in scheme 4.18(a). Alternatively, if the ion at m/z 780.9 arises from $[c_3-CH]^+$, the ion at m/z 572.1 can result from loss of cytosine C-1 followed by ribose 3' C-O bond cleavage as illustrated in scheme 4.18(b). Additional fragmentation of this ion via cleavage of the ribose C-O bond at the 3'-terminus yields the ion at m/z 492.0. The ion at m/z 590.1 may arise from ribose C-O bond cleavage to the 3' side of

the residue from which the base is lost combined with the loss of the 5'-terminal cytosine. Hence, in this example, the in-source dissociation of the fragment ion does not distinguish between these two possibilities.

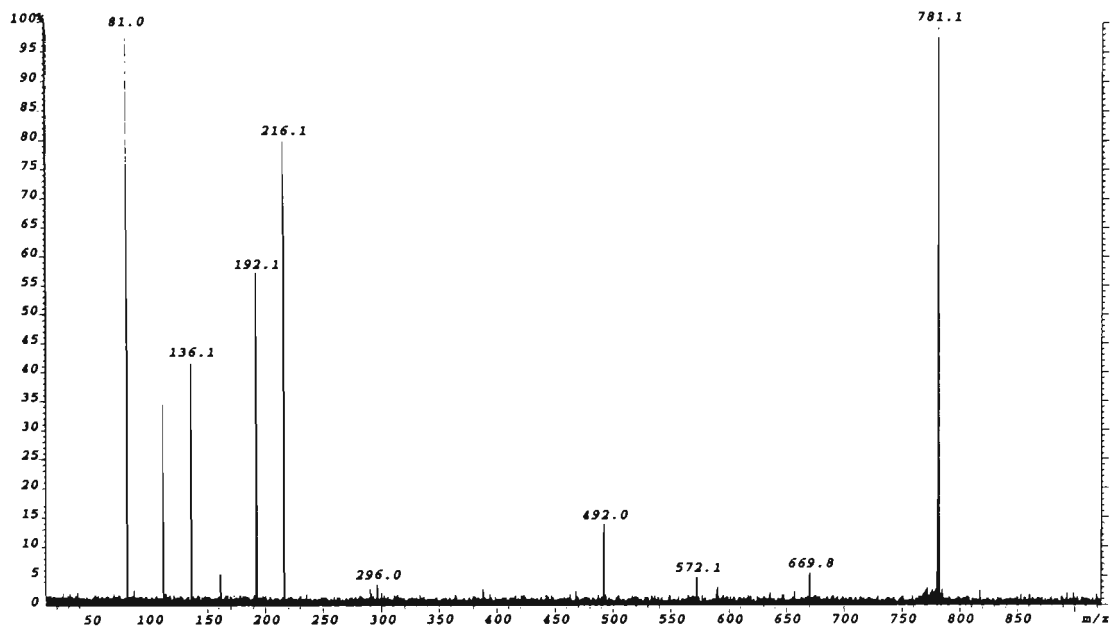
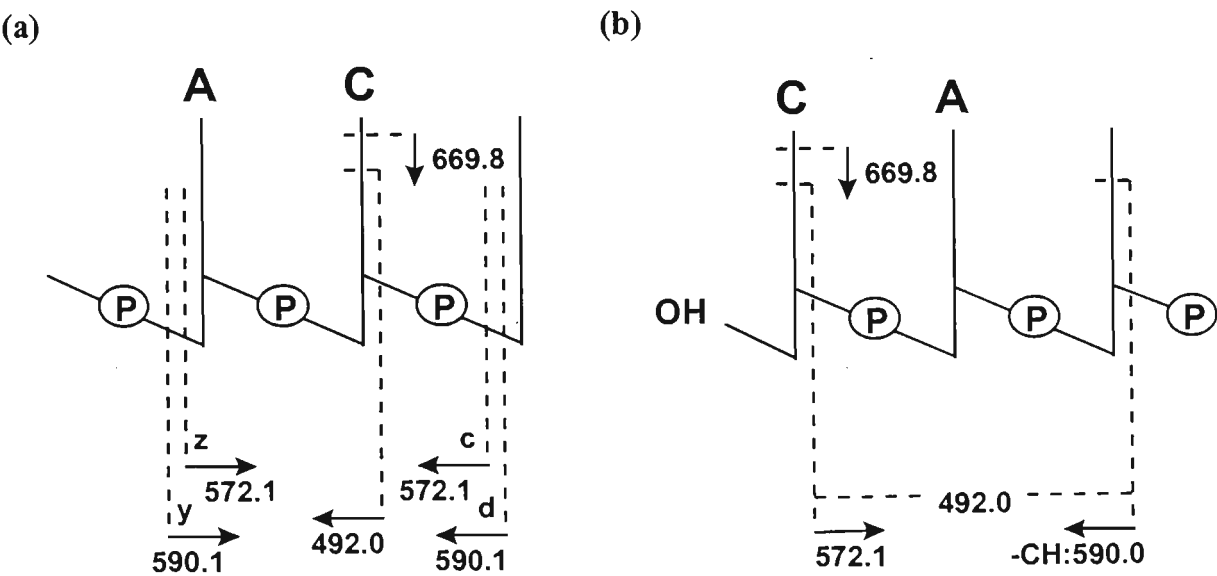


Figure 4.40: ESI-MS/MS spectrum of the ion at m/z 781.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.18: Proposed fragmentation pathways of the two structures (a) $[(AC)+p+s+H]^+$ and (b) $[c_3-CH]^+$ which can be attributed to the ion at m/z 781.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

The MS/MS spectrum of the $[a_3-B_3H]^+$ ion at m/z 701.1 is shown in figure 4.41, and yields fragmentation very similar to that observed for the $[a_4-B_4H]^+$ ion (figure 4.38). The main fragmentation pathways observed upon dissociation of this ion

are represented in scheme 4.19. The spectrum shows an ion at m/z 590.2 corresponding to the loss of neutral cytosine C-1. Subsequent cleavage of the adjacent ribose 3' C-O bond gives rise to the ion at m/z 492.0. The ion at m/z 412.0 may arise from loss of cytosine in addition to ribose C-O bond cleavage to the 3' side of adenine without prior adenine loss. The low m/z region of the spectrum shows intense ions at m/z 81.0 for $[s-H_2O+H]^+$, m/z 136.1 for AH_2^+ , and at m/z 216.1 for $[A+s-H_2O+H]^+$. These ions are observed to be the predominant dissociation products generated in the MS/MS spectra of the ions at m/z 492.0 ($[A_{nt}+p+s+H]^+$), and m/z 590.1 ($[A_{nt}+s+p+s+H]^+$) which are given in figures 4.42(a) and (b) respectively.

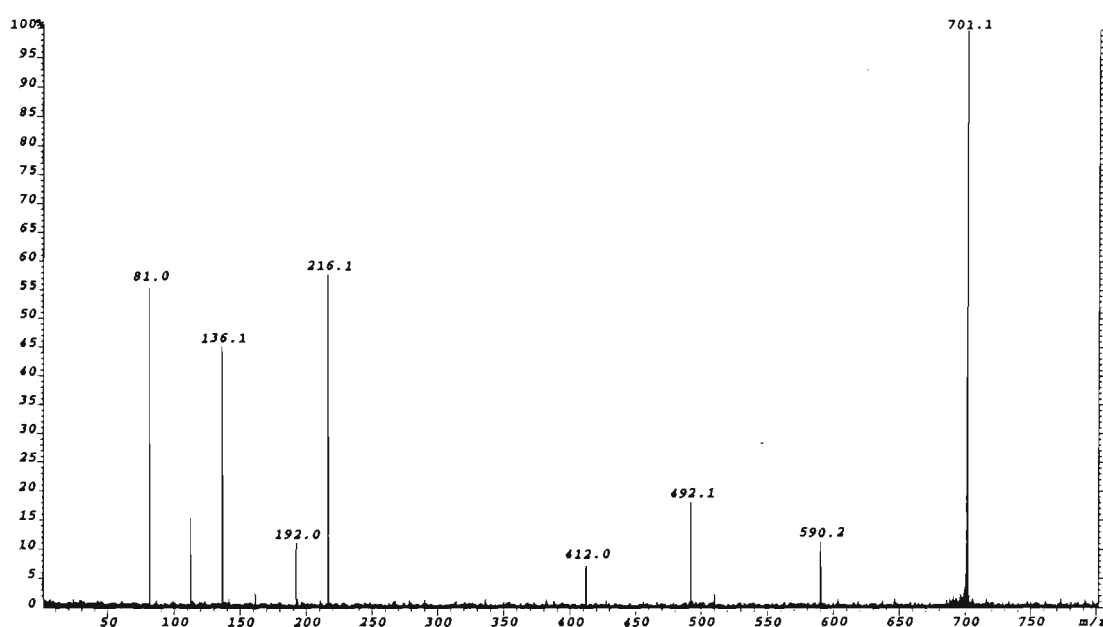
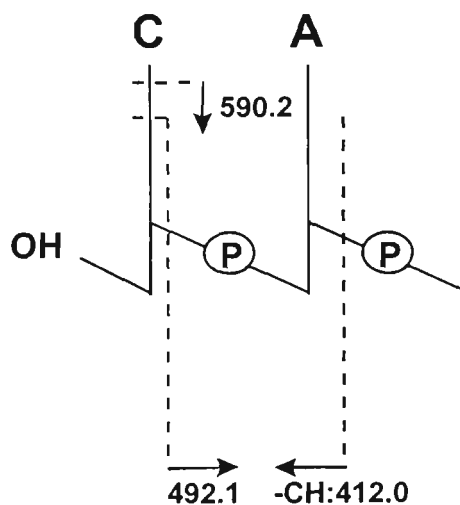
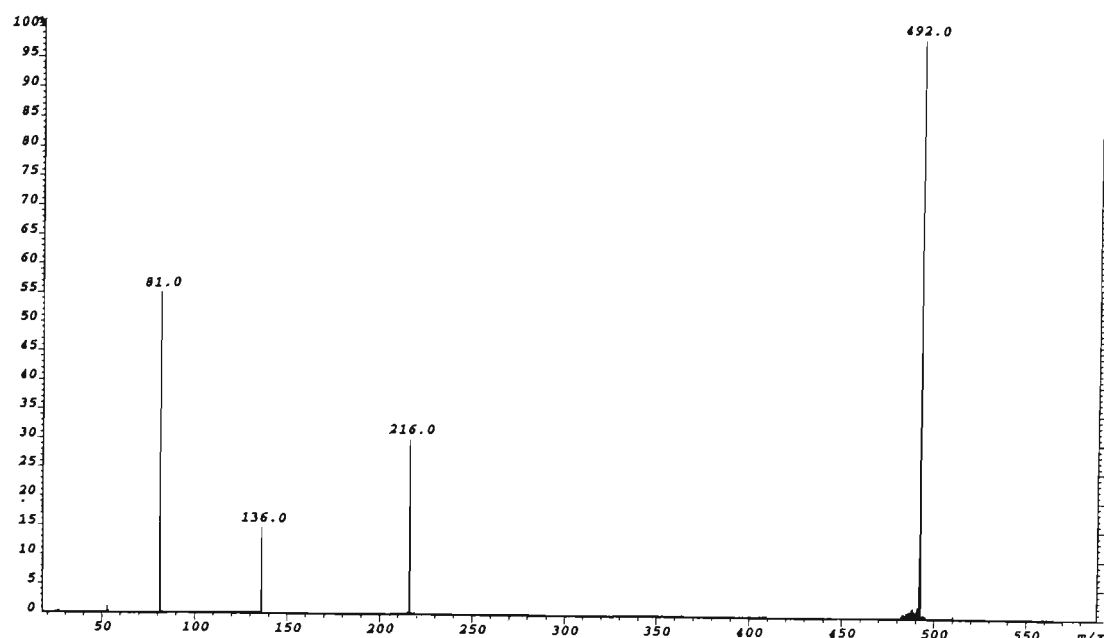


Figure 4.41: ESI-MS/MS spectrum of the ion at m/z 701.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.



Scheme 4.19: Proposed fragmentation pathways of the ion at m/z 701.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

(a)



(b)

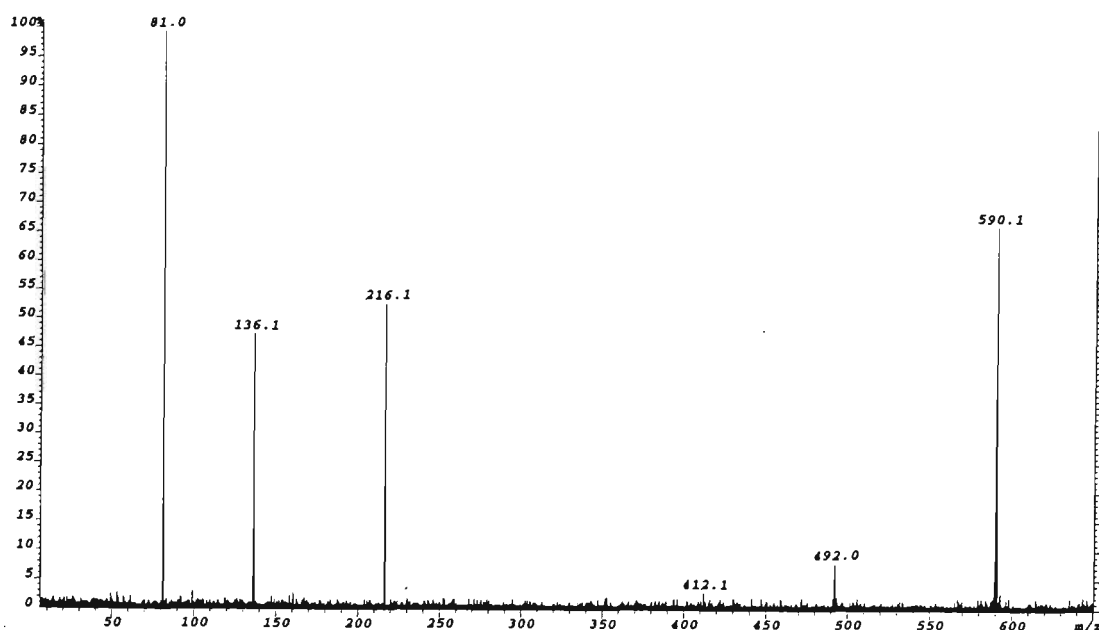


Figure 4.42: ESI-MS/MS spectra of the ions at (a) m/z 492.0 and (b) m/z 590.1 generated by in-source fragmentation of 5'-d(CACGTG)-3'.

Analysis of the MS/MS spectra of product ions in the positive ion MS/MS spectrum of 5'-d(CACGTG)-3', shows dissociation to proceed via base loss followed by cleavage of the adjacent 3' C-O bond analogous to the major fragmentation pathway generated in the negative ion mode. A distinctly different feature of the dissociation of positively charged 5'-d(CACGTG)-3', however, is the prevalent loss of cytosine and guanine nucleobases which strongly affects the types of backbone fragments observed. The MS/MS spectrum of 5'-d(CACGTG)-3' (figure 4.35) generates quite an intense

ion for the loss of guanine at m/z 1641.5 and formation of the $[a_4-B_4H]^+$ ion at m/z 990.0 from ribose 3' C-O bond cleavage following the loss of G-4. Further dissociation of these ions predominantly results in the facile loss of cytosine C-1 and C-3 in addition to cleavage of the respective ribose 3' C-O bonds adjacent to the residue from which the base has been lost. These fragmentation pathways give rise to the variety of relatively intense ions in the MS/MS spectrum of 5'-d(CACGTG)-3' owing to backbone fragments incorporating the adenine nucleobase.

(b) 5'-d(CGTACG)-3'

The MS/MS spectrum of 5'-d(CGTACG)-3' yields major series of ions incorporating the (TA) dinucleotide and the (TAC) trinucleotide, with the most intense ions of each series owing to the ions at m/z 795.8 for $[(TA)+p+s+H]^+$ and at m/z 1085.5 for $[(TAC)+p+s+H]^+$. In-source dissociation followed by MS/MS of the ions observed in the high m/z region of the MS/MS spectrum of 5'-d(CGTACG)-3' showed both of these ions in relatively high abundance, as for example, in the MS/MS spectrum of the ion at m/z 1472.4 which is shown in figure 4.43. This product ion can be attributed to two different species, namely, the ion $[M-GH-GH-H_2O+H]^+$ ($m/z(\text{calc.})=1472.24$) or the

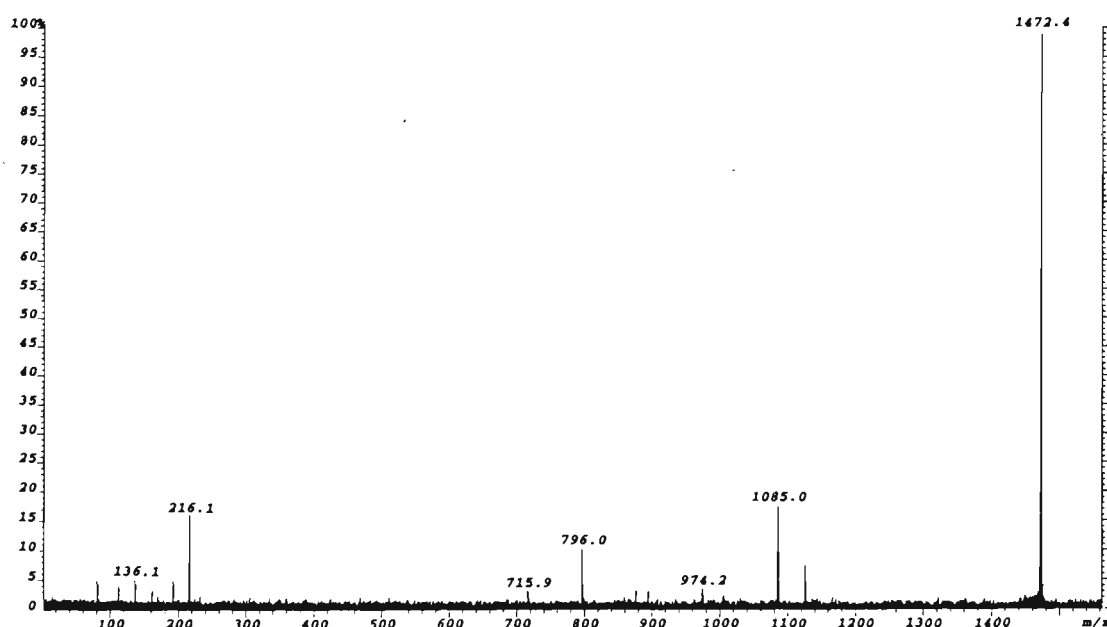
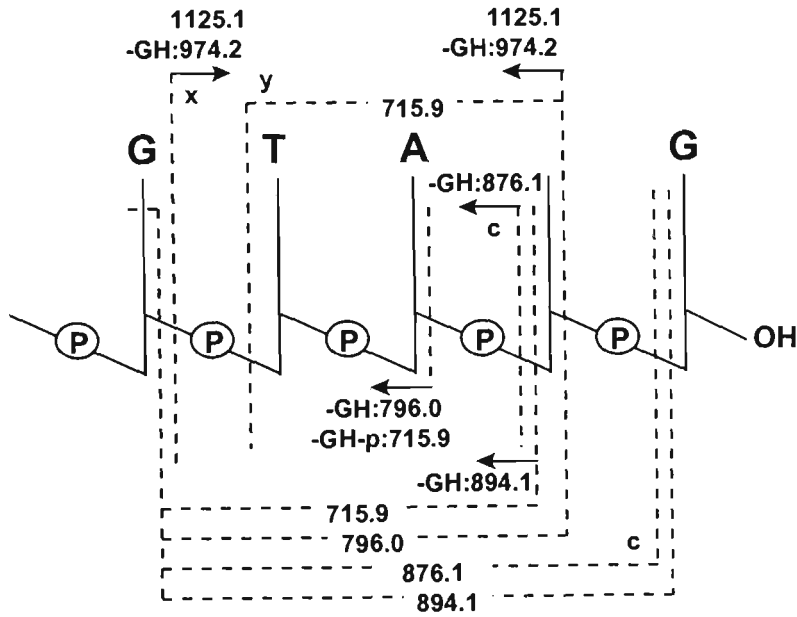


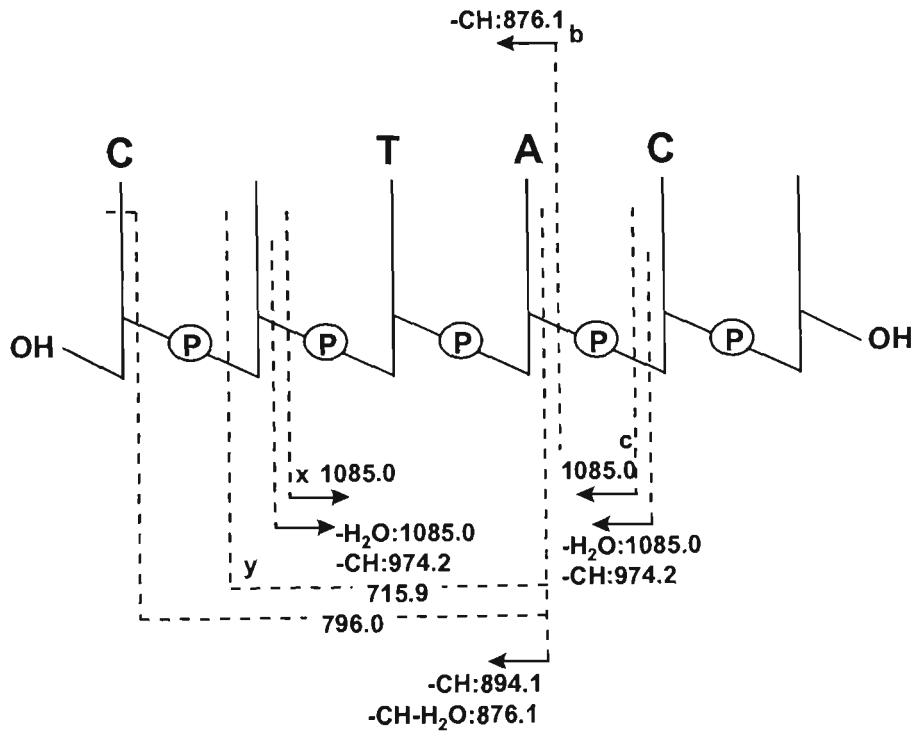
Figure 4.43: ESI-MS/MS spectrum of the ion at m/z 1472.4 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

ion $[w_5\text{-CH}+2\text{H}]^+$ ($m/z(\text{calc.})=1472.23$), since these ions would appear unresolved in the MS/MS spectrum of 5'-d(CGTACG)-3'. The MS/MS spectrum of this ion confirms the presence of both fragments from the observation of the ions at m/z 1125.2 for $[(\text{GTA})+\text{p}+\text{s}+\text{H}]^+$ and at m/z 1085.0 for $[(\text{TAC})+\text{p}+\text{s}+\text{H}]^+$. The ion at m/z 1125.2 can only arise from fragmentation of the species, $[w_5\text{-CH}+2\text{H}]^+$ (scheme 4.20(a)), whereas the ion at m/z 1085.0 can only be a product ion of $[M\text{-GH-GH-H}_2\text{O}+\text{H}]^+$ (scheme 4.20(b)).

(a)



(b)



Scheme 4.20: Proposed fragmentation pathways of the two structures (a) $[w_5\text{-CH}+2\text{H}]^+$ and (b) $[M\text{-GH-GH-H}_2\text{O}+\text{H}]^+$ which can be attributed to the ion at m/z 1472.4 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The formation of the ion at m/z 1125.2 from fragmentation of $[w_5\text{-CH}+2\text{H}]^+$ (shown in scheme 4.20(b)) may arise from cleavage of the ribose C-O bond from which cytosine C-4 is lost, or alternatively from x-type cleavage to the 3' side of ribose containing guanine G-2. Loss of guanine from either fragment yields the ion at m/z 974.2. A variety of possible structures exist for the remaining product ions at m/z 894.1, m/z 876.1, m/z 796.0 and m/z 715.9 from backbone cleavage of this fragment, and some proposed structures of these ions are illustrated in the scheme. The MS/MS spectrum of the ion at m/z 1125.2 shown in figure 4.44 gives rise predominantly to the two (TA)-containing dinucleotide product ion-types at m/z 716.0 and 796.1.

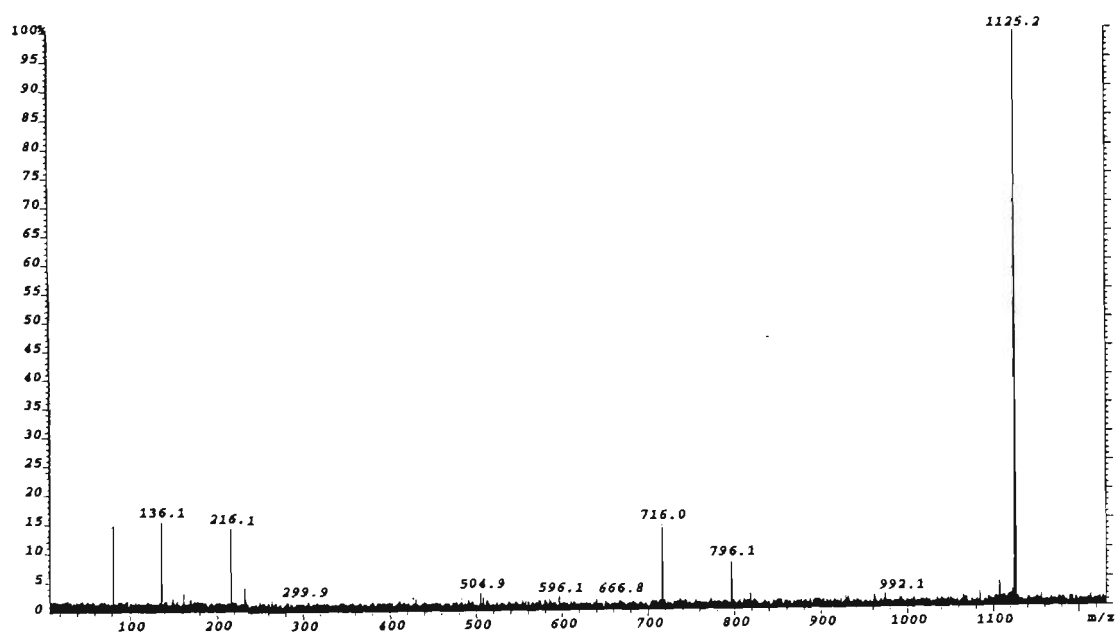


Figure 4.44: ESI-MS/MS spectrum of the ion at m/z 1125.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

Dissociation of the fragment ion $[M\text{-GH-GH-H}_2\text{O}+\text{H}]^+$, can yield the ion at m/z 1085.0 from a number of backbone cleavages as illustrated in scheme 4.20(b). It should be noted that the loss of H_2O is not indicated on the structure but rather with each individual cleavage. Firstly, the ion at m/z 1085.0 can be generated from loss of H_2O in addition to cleavage of the ribose C-O bond to the 3' side of the residue from which guanine G-2 is lost or from cleavage at the d_4 position. Alternatively, this ion can be attributed to c_4 cleavage or x_4 cleavage. Loss of CH from the ion at m/z 1085.0

gives rise to the ion at m/z 974.2. Although all of the possible structures of the remaining product ions comprising the (TA) dinucleotide are not indicated in scheme 4.20(b), these can be formed from the same backbone fragmentation of the precursor ion $[w_5\text{-CH}+2\text{H}]^+$ as indicated in scheme 4.20(a). The MS/MS spectrum of the ion at m/z 1085.0 (shown in figure 4.45) gives rise to the ions at m/z 974.3, m/z 876.2 and m/z 796.3.

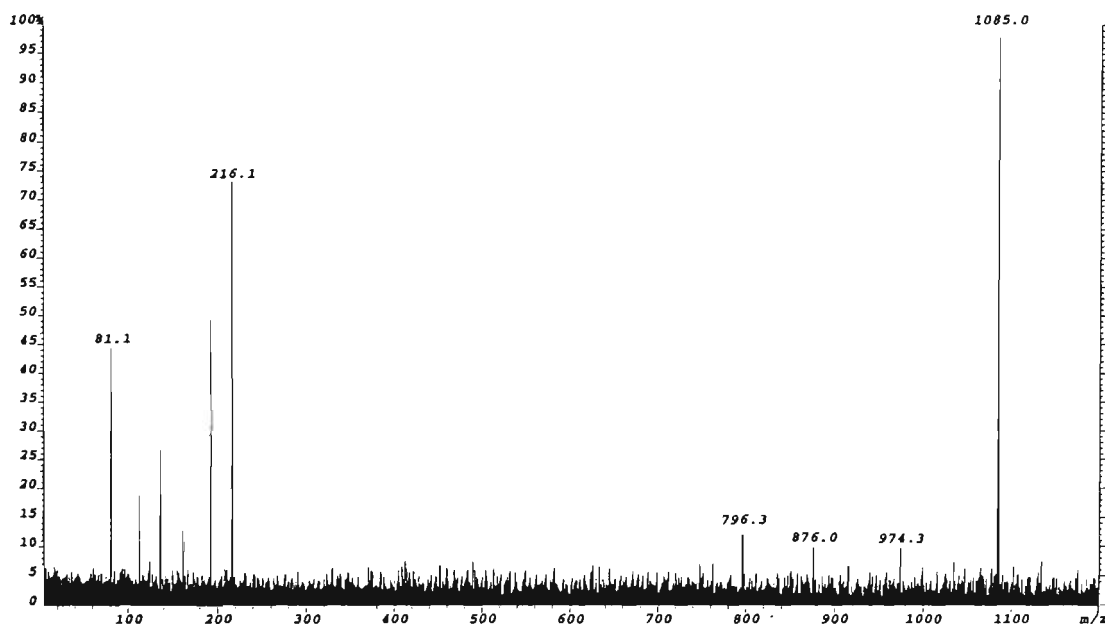


Figure 4.45: ESI-MS/MS spectrum of the ion at m/z 1085.0 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The product ion at m/z 795.8 owing to $[(\text{TA})+\text{p}+\text{s}+\text{H}]^+$, is found to be the major ion following in-source dissociation of a number of product ions (see appendix 4.9). These include the ions at m/z 1432.2, m/z 1414.2, m/z 1334.8, m/z 1281.8, m/z 1223.3, m/z 1183.1, m/z 1143.2, m/z 1125.2, m/z 1103.2, m/z 1085.0, m/z 1005.1, m/z 974.1, and m/z 894.1. The favourable production of this ion from the decomposition of other product ions would appear to be a consequence of the facile loss of guanine G-2 and cytosine C-1 followed by backbone cleavage. This is evident in the MS/MS spectrum of the $[a_5\text{-B}_5\text{H}]^+$ ion at m/z 1334.2 which is shown in figure 4.46. Scheme 4.21 summarises the main fragmentation pathways giving rise to the ions observed in the spectrum. Fragmentation of this ion gives rise to ions for the loss of neutral cytosine at m/z 1223.2 and neutral guanine at m/z 1183.1. Relatively intense

ions are observed at m/z 715.7 and m/z 796.0 and, as shown in scheme 4.21, these can be produced from more than one type of cleavage.

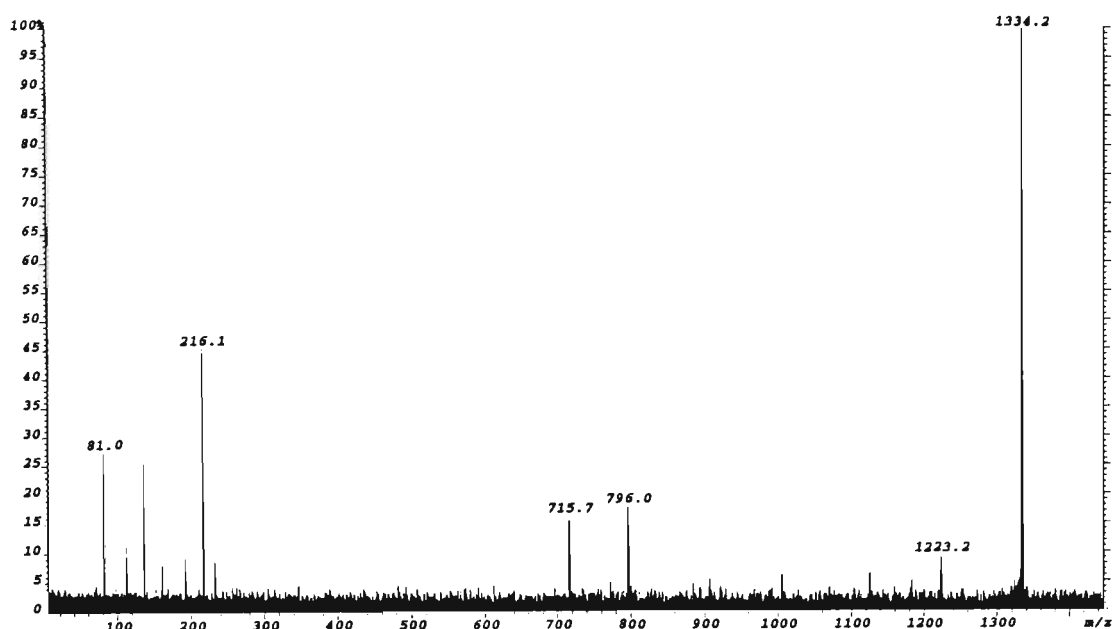
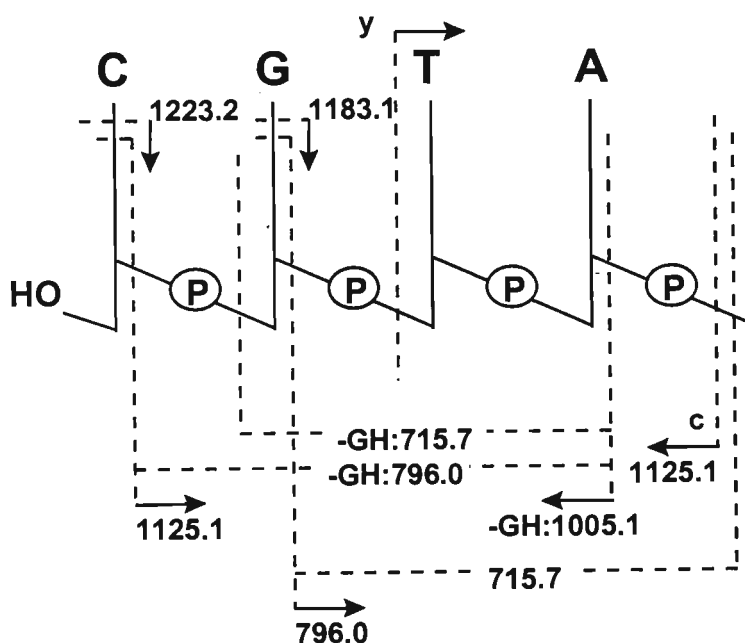


Figure 4.46: ESI-MS/MS spectrum of the ion at m/z 1334.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.21: Proposed fragmentation pathways of the product ion at m/z 1334.2 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

The ion at m/z 1125.1 can arise from subsequent cleavage of the ribose C-O bond to the 3' side of the residue from which cytosine C-1 was lost. Alternatively, this ion

could also result from a c-type cleavage to the 3' side of adenine. The ion at m/z 1005.1 is the 5'-terminal fragment arising from cleavage of the ribose 3' C-O bond of adenine without loss of adenine, in addition to loss of neutral guanine. The MS/MS spectrum of this ion is shown in figure 4.47 and the observed fragmentation produced is illustrated in scheme 4.22. Dissociation of this ion shows facile loss of cytosine indicated by the presence of the ion at m/z 894.0. Subsequent cleavage of the adjacent ribose 3' C-O bond yields the ion at m/z 796.1 for the 3'-terminal fragment. The ion at m/z 387.9 is the $[a_2-B_2H]^+$ ion.

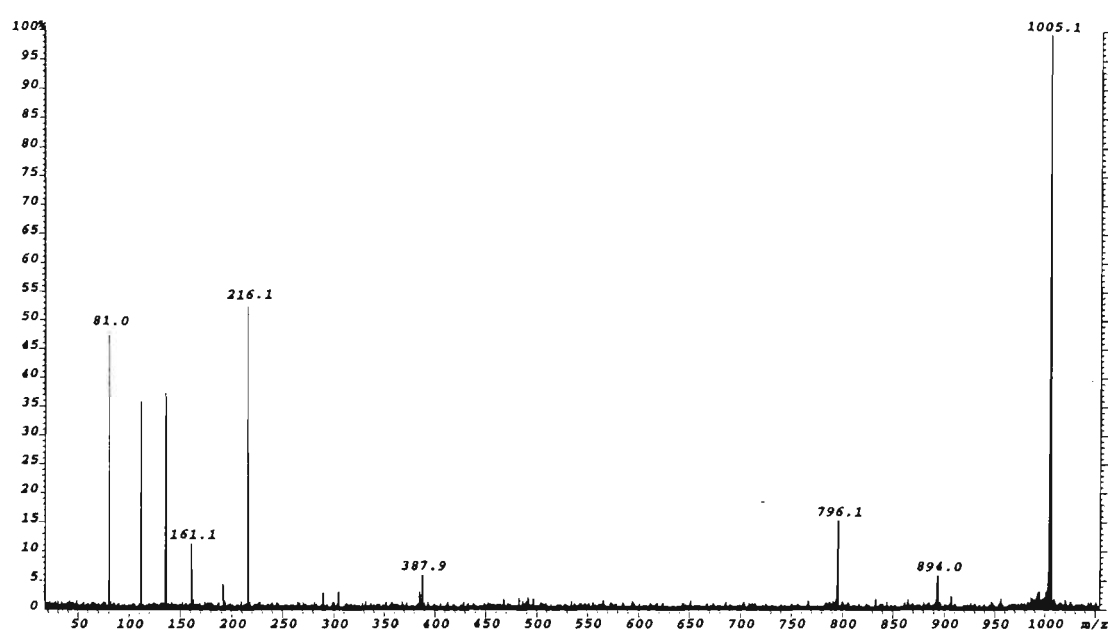
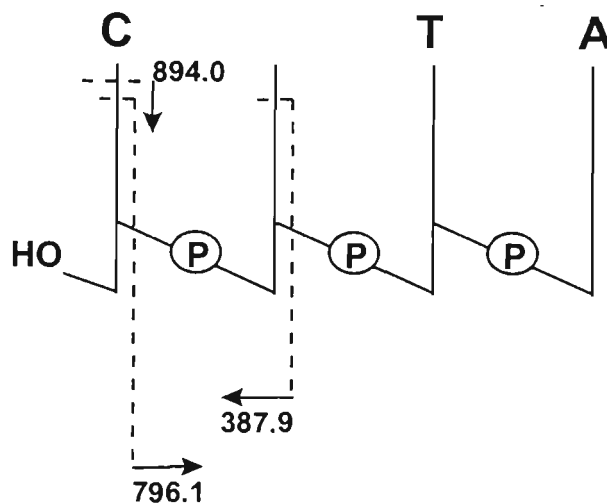


Figure 4.47: ESI-MS/MS spectrum of the ion at m/z 1005.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.



Scheme 4.22: Proposed fragmentation pathways of the ion at m/z 1005.1 generated by in-source fragmentation of 5'-d(CGTACG)-3'.

Therefore, as observed in the case of 5'-d(CACGTG)-3', the analysis of MS/MS spectra of product ions for 5'-d(CGTACG)-3' also shows dissociation to proceed predominantly via nucleobase loss followed by cleavage of the adjacent ribose 3' C-O bond, with the loss of neutral cytosine and guanine nucleobases observed to be the most facile of the nucleobases. Furthermore, dissociation of the product ions formed from this fragmentation pathway yields the facile loss of neutral cytosine and guanine in addition to cleavage of the ribose 3' C-O bond of the residue from which the base is lost. Similar fragmentation pathways are observed in the case of 5'-d(CGTACG)-3', which dissociates to yield intense ions corresponding to various backbone fragments with T and A attached, and a less intense series of backbone fragments incorporating the nucleobases, T, A and C. These observations of facile loss of neutral cytosine and guanine can be understood in terms of the relative stabilities of the tautomeric forms of these nucleobases, in addition to their higher proton affinities relative to that of thymine, resulting in protonation of these bases being more favourable. In a study by Rodgers *et al.*¹⁹ on the dissociation pathways of protonated dinucleotides, these workers calculated the relative energies of the most stable tautomeric structures of each of the nucleobases upon elimination as a tautomer BH, following protonation and glycosidic bond cleavage. From their calculations, it was found that the energies of the most stable tautomeric forms of cytosine and guanine in comparison to their normal, commonly occurring structures are 0.6 kcal mol⁻¹ and 1.5 kcal mol⁻¹ respectively, followed by that of adenine which has an energy of 6.7 kcal mol⁻¹ greater than its common structure, with the most stable tautomer of thymine being 9.0 kcal mol⁻¹ higher in energy than its normal form. These calculations indicate that cytosine and guanine are significantly better leaving groups than adenine, with thymine being the least stable leaving group of all the bases. The results obtained for the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' are consistent with these calculations.

4.6 COMPARISON OF POSITIVE ION AND NEGATIVE ION FRAGMENTATION OF OLIGONUCLEOTIDES

4.6.1 Fragmentation Pathways of Singly Charged Oligonucleotide Anions

To investigate further some of the characterisitics of the fragmentation of singly charged oligonucleotide anions, MS/MS spectra of the [M-H]⁻ ions of the oligonucleotides 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3' were acquired and are shown in figures 4.48, 4.49, 4.50, and 4.51 respectively. The corresponding product ion assignments for these spectra are given in tables 4.13, 4.14, 4.15, and 4.16.

Table 4.13: Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-H]⁻ ion oligonucleotide 5'-d(ATGCAT)-3'.

m/z	Rel. Int.(%)	m/z(calc).	Assignment
1789.5	100	1789.4	[M-H] ⁻
1653.9	2.5	1654.2	loss of Adenine
1556.7	3.8	1556.3	w ₅ ⁻
1332.1	4.4	1332.2	[a ₅ -B ₅ H-2H] ⁻ and/or [z ₅ -AH-2H] ⁻ and/or [(TGCA)+s-H] ⁻
		or 1332.2	[(TGCA)+p+ H ₂ O -H] ⁻
1234.0	2.9	1234.2	[x ₄ -2H] ⁻ and/or [c ₄ -2H] ⁻ and/or [(TCGA)-H] ⁻
1172.5	2.1	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(TCGA)-p+H ₂ O -H] ⁻
1154.1	3.0	1154.2	[a ₄ -2H] ⁻ and/or [z ₄ -2H] ⁻ and/or [(TCGA)-p-H] ⁻
1043.1	40.9	1043.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(TGCA)-CH-H] ⁻
945.0	3.1	945.1	[c ₃ -2H] ⁻
923.1	70.3	923.2	w ₃ ⁻
905.2	3.5	905.1	[x ₃ -2H] ⁻
843.2	4.0	843.2	y ₃ ⁻
810.0	7.5	810.1	[c ₃ -AH-2H] ⁻ and/or [(TG)+s+p-H] ⁻
714.1	7.6	714.1	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -CH-2H] ⁻
681.1	2.2	681.1	[(CA)+p-2H] ⁻
		or 681.1	[(CA)+s-H ₂ O-H] ⁻
634.0	35.7	634.1	w ₂ ⁻ and/or d ₂ ⁻
617.1	4.6	617.1	[(CG)-H] ⁻
616.1	7.0	616.1	[c ₂ -2H] ⁻ and/or [x ₂ -2H] ⁻
554.0	2.3	554.1	y ₂ ⁻ and/or b ₂ ⁻
506.0	9.2	506.0	[G _{nt} +s+p-H] ⁻
490.0	2.5	490.1	[A _{nt} +s+p-H] ⁻ and/or [c ₂ -TH-2H] ⁻ and/or [x ₂ -TH-2H] ⁻
488.3	2.3	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
480.9	2.1	481.0	[T _{nt} +s+p-H] ⁻ and/or [c ₂ -AH-2H] ⁻ and/or [x ₂ -AH-2H] ⁻
466.1	7.1	466.0	[C _{nt} +s+p-H] ⁻
448.1	1.8	448.0	[C _{nt} +s+p- H ₂ O-H] ⁻
426.0	1.9	426.1	[G _{nt} +s-H] ⁻
		or 426.1	[G _{nt} +p+H ₂ O-H] ⁻
410.0	1.3	410.1	[a ₂ -B ₂ H-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	8.0	408.1	[G _{nt} +s- H ₂ O-H] ⁻

401.0	6.2	or 408.0	[G _{nt} +p-H] ⁻
		401.1	[T _{nt} +s-H] ⁻ and/or [a ₂ -AH-2H] ⁻ and/or [z ₂ -AH-2H] ⁻
392.1	3.6	or 401.0	[T _{nt} +p+H ₂ O-H] ⁻
		392.1	[A _{nt} +s-H ₂ O-H] ⁻
386.0	2.8	or 392.0	[A _{nt} +p-H] ⁻
		386.1	[C _{nt} +s-H] ⁻
383.2	4.7	or 386.0	[C _{nt} +p+H ₂ O-H] ⁻
		383.1	[T _{nt} +s-H ₂ O-H] ⁻
368.1	7.0	or 383.0	[T _{nt} +p-H] ⁻
		368.1	[C _{nt} +s-H ₂ O-H] ⁻
346.1	1.7	or 368.0	[C _{nt} +p-H] ⁻
328.0	12.8	346.1	[G _{nt} +H ₂ O-H] ⁻
321.1	11.4	328.0	[G _{nt} -H] ⁻
312.0	4.0	321.0	w ₁ ⁻ and/or [T _{nt} + H ₂ O-H] ⁻
303.1	5.8	312.0	[c ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
288.0	5.9	303.0	[x ₁ -2H] ⁻ and/or [T _{nt} -H] ⁻
257.0	4.9	288.0	[C _{nt} -H] ⁻
		257.1	[sps- H ₂ O-H] ⁻
239.0	1.8	or 257.0	[psp-H] ⁻
230.0	2.8	239.0	[psp- H ₂ O-H] ⁻
214.0	4.7	230.1	[sG-H ₂ O-H] ⁻
195.0	4.2	214.0	[sA-H ₂ O-H] ⁻
177.0	15.7	195.0	[ps+ H ₂ O-H] ⁻
158.9	4.1	177.0	[ps-H] ⁻
134.1	6.8	159.0	[ps-H ₂ O-H] ⁻
125.0	2.9	134.0	A ⁻
96.9	7.2	125.0	T ⁻
		97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
79.0	15.9	or 97.0	[p+H ₂ O-H] ⁻
		79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

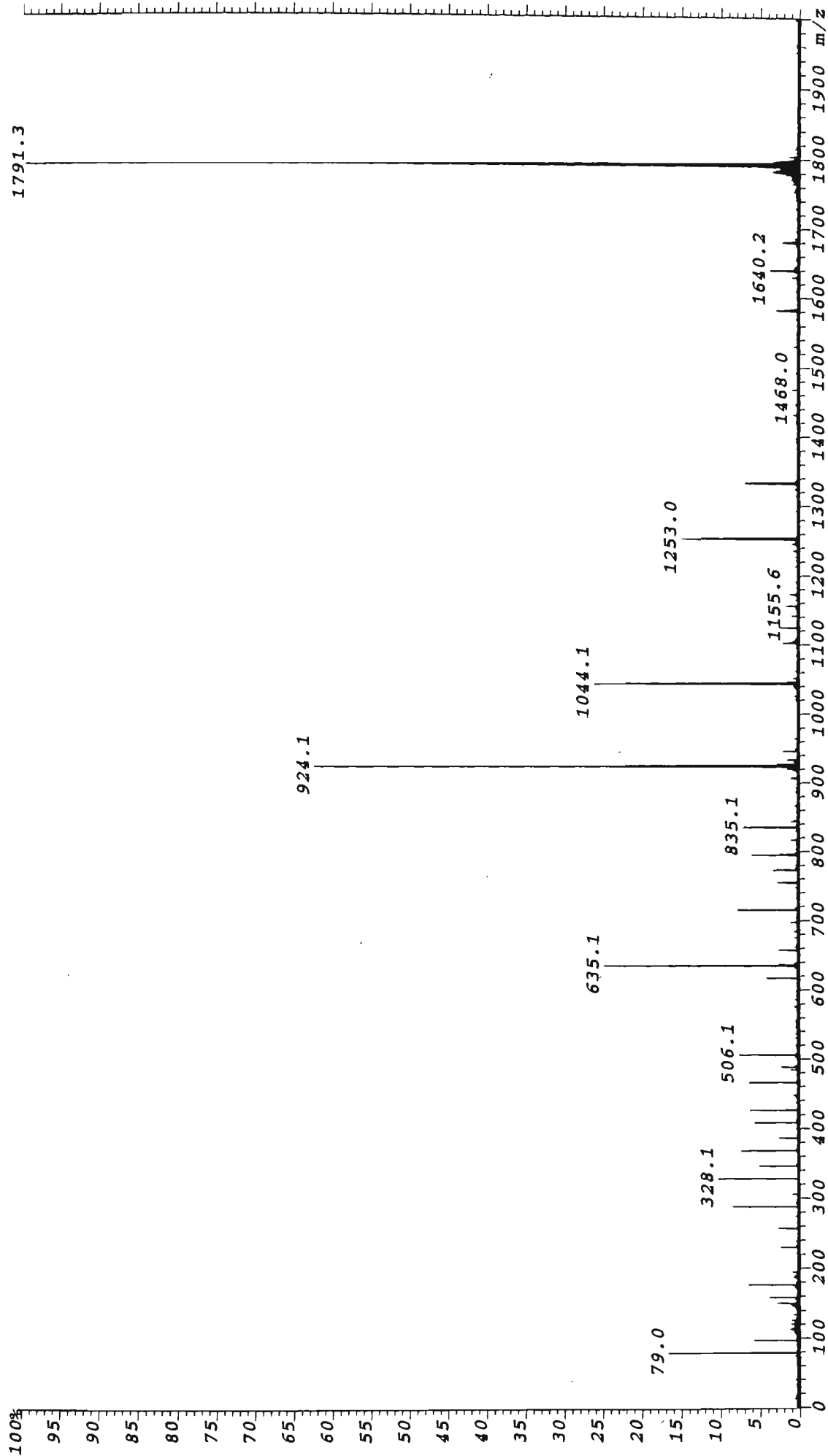


Figure 4.48: Negative Ion ESI-MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(ATGCAT)-3' (m/z 1789.5).

Table 4.14: Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-H]^-$ ion of the oligonucleotide 5'-d(TCACGA)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc.)</i>	<i>Assignment</i>
1774.3	100	1774.3	$[M-H]^-$
1638.3	1.8	1638.3	loss of Adenine
1292.0	10.2	1292.2	$[a_5-B_5H-2H]^-$
1261.3	11.8	1261.2	w_4^-
1003.1	42.1	1003.2	$[a_4-B_4H-2H]^-$
948.1	18.0	948.2	w_3^- and/or $[(ACG)+H_2O-H]^-$
930.3	2.1	930.1	$[x_3-2H]^-$ and/or $[(ACG)-H]^-$
904.8	2.1	905.1	$[c_3-2H]^-$
779.0	10.2	779.1	$[c_3-TH-2H]^-$ and/or $[x_3-GH-2H]^-$ and/or $[(CA)+s+p-H]^-$ and/or $[(AC)+s+p-H]^-$
699.1	4.1	699.1	$[a_3-TH-2H]^-$ and/or $[z_3-GH-2H]^-$ and/or $[(CA)+s-H]^-$ and/or $[(AC)+s-H]^-$
		or 699.1	$[(CA)+p+H_2O-H]^-$ and/or $[(AC)+p+H_2O-H]^-$
690.1	4.3	690.1	$[a_3-B_3H-2H]^-$
681.5	2.1	681.1	$[(CA)+s-H_2O-H]^-$
		or 681.1	$[(CA)+p-H]^-$
659.0	32.4	659.1	w_2^-
641.0	3.8	641.1	$[x_2-2H]^-$
619.1	2.8	619.1	$[(CA)+H_2O-H]^-$
592.2	1.9	592.1	$[c_2-2H]^-$
530.0	1.3	530.1	b_2^-
490.0	6.1	490.1	$[A_{nt}+s+p-H]^-$ and/or $[x_2-GH-2H]^-$
472.5	2.7	472.0	$[A_{nt}+s+p-H_2O-H]^-$
465.9	4.6	466.0	$[C_{nt}+s+p-H]^-$ and/or $[c_2-TH-2H]^-$
447.9	2.0	448.0	$[C_{nt}+s+p-H_2O-H]^-$
426.0	1.1	426.1	$[G_{nt}+s-H]^-$
		or 426.1	$[G_{nt}+p+H_2O-H]^-$
410.1	4.0	410.1	$[z_2-GH-2H]^-$ and/or $[A_{nt}+s-H]^-$
		410.0	$[A_{nt}+p+H_2O-H]^-$
408.0	3.2	408.1	$[G_{nt}+s-H_2O-H]^-$
		or 408.0	$[G_{nt}+p-H]^-$
391.9	6.1	392.1	$[A_{nt}+s-H_2O-H]^-$
		or 392.0	$[A_{nt}+p-H]^-$
386.0	1.6	386.1	$[a_2-TH-2H]^-$ and/or $[C_{nt}+s-H]^-$
		or 386.0	$[C_{nt}+p+H_2O-H]^-$
368.1	6.1	368.1	$[C_{nt}+s-H_2O-H]^-$
		or 368.0	$[C_{nt}+p-H]^-$
330.0	4.2	330.1	w_1^- and/or $[A_{nt}+H_2O-H]^-$
328.0	2.8	328.0	$[G_{nt}-H]^-$
321.1	11.4	321.0	$[T_{nt}+H_2O-H]^-$
312.0	4.0	312.0	$[x_1-2H]^-$ and/or $[A_{nt}-H]^-$
306.0	1.8	306.0	$[C_{nt}+H_2O-H]^-$
288.0	5.9	288.0	$[C_{nt}-H]^-$
256.9	4.3	257.1	$[sps-H_2O-H]^-$
		or 257.0	$[psp-H]^-$
230.0	2.1	230.1	$[sG-H_2O-H]^-$
214.0	3.9	214.0	$[sA-H_2O-H]^-$
195.1	2.7	195.0	$[ps+H_2O-H]^-$
177.1	9.5	177.0	$[ps-H]^-$
159.1	2.8	159.0	$[ps-H_2O-H]^-$
150.1	2.5	150.0	G^-
134.0	2.8	134.0	A^-
125.1	2.9	125.0	T^-
97.0	5.5	97.0	$[a_1-B_1H-2H]^-$ and/or $[s-H]^-$
		or 97.0	$[p+H_2O-H]^-$
79.0	14.8	79.0	$[p-H]^-$ (PO_3^-)

*calculated monoisotopic mass

 $s = \text{deoxyribose-H}_2\text{O}$ ($C_5H_6O_2$), $M_r = 98.0368$ Da; $p = PO_3H$, $M_r = 79.9663$ Da B_{nt} denotes a mononucleotide which may be either psB or sBp ; $[(B_1..B_n)]$ denotes a polynucleotide which may be either $(psB_1...+psB_n)$ or $(sB_1p...+sB_np)$

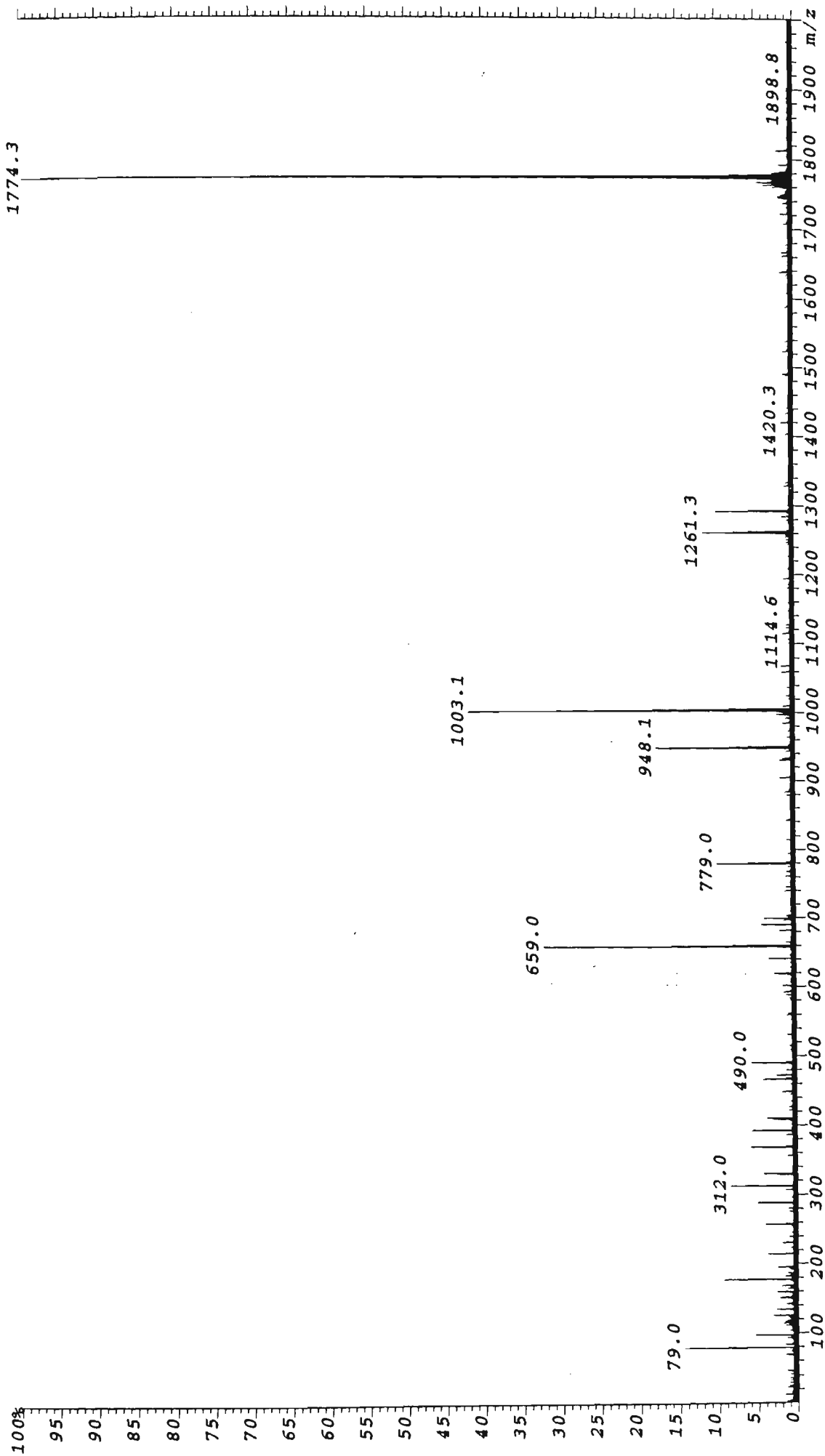


Figure 4.49: Negative Ion ESI-MS/MS spectrum of the [M-H]⁻ ion of 5'-d(TCACGA)-3' (m/z 1774.3).

Table 4.15 : Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-H]^-$ ion of the oligonucleotide 5'-d(CG GCCG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1791.3	100	1791.3	$[M-H]^-$
1680.2	2.2	1680.3	loss of Cytosine
1640.2	4.0	1640.3	loss of Guanine
1582.2	3.1	1582.3	w_5^-
1333.2	7.1	1333.2	$[a_5-B_5H-2H]^-$ and/or $[z_5-GH-2H]^-$
1253.0	15.2	1253.2	w_4^- and/or d_4^- and/or $[(GGCC)+H_2O-H]^-$
1173.1	1.3	1173.2	y_4^- and/or b_4^- and/or $[(GGCC)-p+H_2O-H]^-$
1155.1	2.0	1155.2	$[z_4-2H]^-$ and/or $[a_4-2H]^-$ and/or $[(GGCC)-p-H]^-$
1141.8	1.0	1142.1	$[w_4-CH]^-$ and/or $[d_4-CH]^-$ and/or $[(GGC)+s+p+H_2O-H]^-$
1124.0	2.5	1124.1	$[x_4-CH-2H]^-$ and/or $[c_4-CH-2H]^-$ and/or $[(GGC)+s+p-H]^-$
1102.1	2.1	1102.2	$[w_4-GH]^-$ and/or $[d_4-GH]^-$ and/or $[(GCC)+s+p+H_2O-H]^-$
1044.1	25.1	1044.2	$[a_4-B_4H-2H]^-$ and/or $[z_4-CH-2H]^-$ and/or $[(GGC)+s-H]^-$
964.2	0.8	964.2	d_3^- and/or $[(GGC)+H_2O-H]^-$
945.9	2.1	946.1	$[c_3-2H]^-$ and/or $[(GGC)-H]^-$
933.3	1.8	933.1	$[(CC)+s+p+s+p-H]^-$
		or 933.1	$[(GG)+s+p+s-H]^-$
		or 933.1	$[(GG)+p+s+p-H_2O-H]^-$
924.1	63.0	924.1	w_3^- and/or $[(GCC)+H_2O-H]^-$
906.1	1.2	906.1	$[x_3-2H]^-$ and/or $[(GCC)-H]^-$
866.6	0.7	866.2	$[a_3-2H]^-$ and/or $[(GGC)-p-H]^-$
843.8	1.2	844.2	y_3^- and/or $[(GCC)-p+H_2O-H]^-$
835.1	7.5	835.1	$[c_3-CH-2H]^-$ and/or $[(GG)+s+p-H]^-$
		or 835.0	$[(CC)+p+s+p-H]^-$
		or 835.1	$[(CC)+s+p+s-H_2O-H]^-$
816.9	1.5	817.1	$[(GG)+s+p-H_2O-H]^-$
795.0	6.2	795.1	$[c_3-GH-2H]^-$ and/or $[x_3-CH-2H]^-$ and/or $[(CG)+s+p-H]^-$
773.3	3.6	773.1	$[w_3-GH]^-$ and/or $[(CC)+s+p+H_2O-H]^-$
		or 773.1	$[b_3-CH]^-$ and/or $[(GG)+s+H_2O-H]^-$
754.9	3.0	755.1	$[x_3-GH]^-$ and/or $[(CC)+s+p-H]^-$
		or 755.1	$[a_3-CH]^-$ and/or $[(GG)+s-H]^-$
		or 755.1	$[(GG)+p+H_2O-H]^-$
715.0	8.1	715.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-CH-2H]^-$ and/or $[(GC)+s-H]^-$
		or 715.1	$[(GC)+p+H_2O-H]^-$
696.9	1.2	697.1	$[(GC)+p-H]^-$
		or 697.1	$[(GC)+s-H_2O-H]^-$
656.8	2.8	657.1	$[(CC)+s-H_2O-H]^-$
		or 657.1	$[(CC)+p-H]^-$
		or 657.1	$[(GG)-H]^-$
635.1	25.1	635.1	w_2^- and/or d_2^- and/or $[(GC)+H_2O-H]^-$
617.1	4.4	617.1	$[c_2-2H]^-$ and/or $[x_2-2H]^-$ and/or $[(GC)-H]^-$
576.9	0.9	577.1	$[(CC)-H]^-$
555.4	1.9	555.1	y_2^- and/or b_2^- and/or $[(CG)-p+H_2O-H]^-$
506.1	7.9	506.0	$[x_2-CH-2H]^-$ and/or $[c_2-CH-2H]^-$ and/or $[G_{nt}+s+p-H]^-$
488.2	2.3	488.0	$[G_{nt}+s+p-H_2O-H]^-$
447.9	1.0	448.0	$[C_{nt}+s+p-H_2O-H]^-$
426.1	6.6	426.1	$[a_2-CH-2H]^-$ and/or $[z_2-GH-2H]^-$ and/or $[G_{nt}+s-H]^-$
		or 426.1	$[G_{nt}+p+H_2O-H]^-$
408.1	5.9	408.1	$[G_{nt}+s-H_2O-H]^-$
		or 408.0	$[G_{nt}+p-H]^-$

386.1	2.8	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
		or 386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.0	7.5	368.1	[C _{nt} +s-H ₂ O-H] ⁻
		or 368.0	[C _{nt} +p-H] ⁻
346.1	5.3	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	10.5	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
306.1	1.1	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
288.1	8.7	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
274.9	0.7	275.1	[sps-H] ⁻
		or 275.1	[psp+ H ₂ O-H] ⁻
257.0	2.8	257.1	[sps- H ₂ O-H] ⁻
		or 257.0	[psp-H] ⁻
239.0	0.6	239.1	[psp-H ₂ O-H] ⁻
230.0	2.6	230.1	[sG-H ₂ O-H] ⁻
195.1	1.1	195.0	[ps+H ₂ O-H] ⁻
190.1	0.8	190.1	[sC-H ₂ O-H] ⁻
177.0	6.7	177.0	[ps-H] ⁻
159.0	4.0	159.0	[ps-H ₂ O-H] ⁻
150.1	3.0	150.0	G ⁻
110.0	1.2	110.0	C ⁻
97.0	6.1	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
		or 97.0	[p+H ₂ O-H] ⁻
79.0	16.8	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_n denotes a mononucleotide which may be either psB or sBp
[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

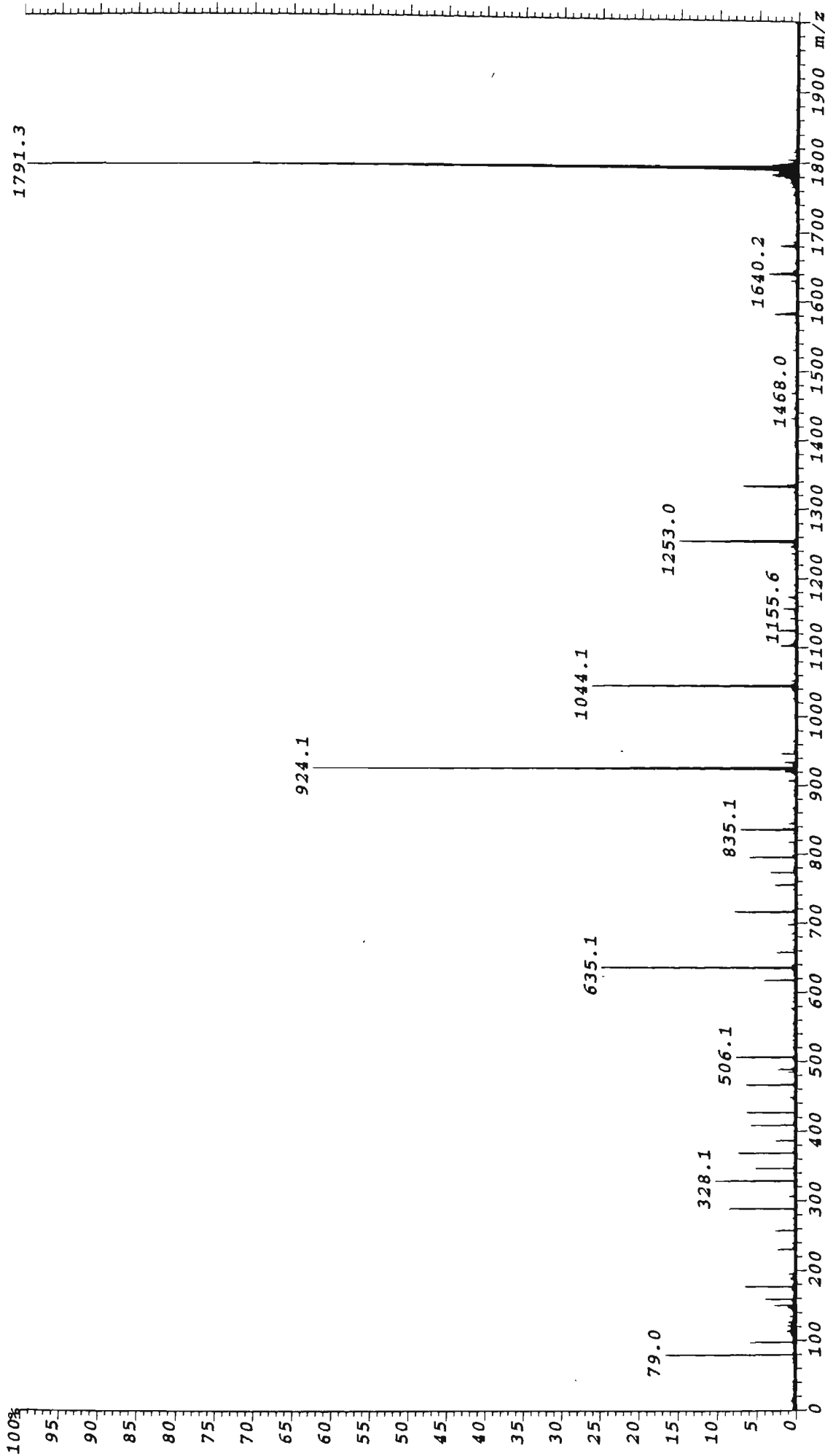


Figure 4.50: Negative Ion ESI-MS/MS spectrum of the $[M-H]^-$ ion of 5'-d(CGCGCCG)-3' (m/z 1791.3).

Table 4.16: Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-H]^-$ ion of the oligonucleotide 5'-d(GCCGGC)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1791.3	100	1791.3	$[M-H]^-$
1680.6	3.7	1680.3	loss of Cytosine
1640.2	3.5	1640.3	loss of Guanine
1542.1	3.6	1542.2	w_5^-
1333.0	7.7	1333.2	$[a_5-B_5H-2H]^-$ and/or $[z_5-CH-2H]^-$
1253.3	10.1	1253.2	w_4^- and/or d_4^- and/or $[(CCGG)+H_2O-H]^-$
1235.2	1.8	1235.2	$[x_4-2H]^-$ and/or $[c_4-2H]^-$ and/or $[(CCGG)-H]^-$
1173.1	1.8	1173.2	y_4^- and/or b_4^- and/or $[(CCGG)-p+H_2O-H]^-$
1155.1	1.7	1155.2	$[z_4-2H]^-$ and/or $[a_4-2H]^-$ and/or $[(CCGG)-p-H]^-$
1141.8	2.0	1142.1	$[w_4-CH]^-$ and/or $[d_4-CH]^-$ and/or $[(CCGG)+s+p+H_2O-H]^-$
1084.3	2.6	1084.1	$[x_4-GH-2H]^-$ and/or $[c_4-GH-2H]^-$ and/or $[(CCG)+s+p-H]^-$
1004.2	53.1	1004.2	$[a_4-B_4H-2H]^-$ and/or $[z_4-GH-2H]^-$ and/or $[(CCG)+s-H]^-$
		or 1004.1	$[(CCG)+p+H_2O-H]^-$
964.2	66.2	964.2	w_3^- and/or $[(CCG)+H_2O-H]^-$
946.3	4.1	946.1	$[x_3-2H]^-$ and/or $[(CCG)-H]^-$
924.2	1.6	924.1	d_3^- and/or $[(CCG)+H_2O-H]^-$
906.2	2.1	906.1	$[c_3-2H]^-$ and/or $[(CCG)-H]^-$
884.2	6.0	884.2	y_3^- and/or $[(CCG)-p+H_2O-H]^-$
853.2	7.2	853.1	$[z_3-CH-2H]^-$ and/or $[a_3-CH-2H]^-$ and/or $[(GG)+s+p+H_2O-H]^-$
844.2	1.1	844.2	b_3^- and/or $[(CCG)-p+H_2O-H]^-$
835.3	1.9	835.1	$[x_3-CH-2H]^-$ and/or $[(GG)+s+p-H]^-$
795.1	8.9	795.1	$[c_3-CH-2H]^-$ and/or $[x_3-GH-2H]^-$ and/or $[(CG)+s+p-H]^-$
777.0	0.9	777.1	$[(CG)+s+p-H_2O-H]^-$
755.2	11.0	755.1	$[c_3-GH-2H]^-$ and/or $[x_3-CH-2H]^-$ and/or $[(CC)+s+p-H]^-$
737.1	3.8	737.1	$[(CC)+s+p-H_2O-H]^-$
		or 737.1	$[(GG)+s-H_2O-H]^-$
		or 737.1	$[(GG)+p-H]^-$
715.1	15.2	715.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-GH-2H]^-$ and/or $[(CG)+s-H]^-$
		or 715.1	$[(CG)+p+H_2O-H]^-$
696.8	3.0	697.1	$[(CG)+p-H]^-$
		or 697.1	$[(CG)+s-H_2O-H]^-$
675.1	3.5	675.1	$[(CC)+p+H_2O-H]^-$
		or 675.1	$[a_3-GH-2H]^-$ or $[(CC)+s-H]^-$
657.0	5.9	657.1	$[(CC)+s-H_2O-H]^-$
		or 657.1	$[(CC)+p-H]^-$
		or 657.1	$[(GG)-H]^-$
635.1	44.2	635.1	w_2^- and/or d_2^- and/or $[(CG)+H_2O-H]^-$
617.0	12.5	617.1	$[c_2-2H]^-$ and/or $[x_2-2H]^-$ and/or $[(CG)-H]^-$
577.4	1.8	577.1	$[(CC)-H]^-$
555.4	1.9	555.1	y_2^- and/or b_2^- and/or $[(CG)-p+H_2O-H]^-$
537.1	1.7	537.1	$[z_2-2H]^-$ and/or $[a_2-2H]^-$ and/or $[(CG)-p-H]^-$
523.9	2.7	524.1	$[w_2-CH]^-$ and/or $[d_2-CH]^-$ and/or $[G_{nt}+s+p+H_2O-H]^-$
506.1	18.0	506.0	$[x_2-CH-2H]^-$ and/or $[c_2-CH-2H]^-$ and/or $[G_{nt}+s+p-H]^-$
488.2	3.9	488.0	$[G_{nt}+s+p-H_2O-H]^-$
426.2	8.7	426.1	$[a_2-B_2H-2H]^-$ and/or $[z_2-CH-2H]^-$ and/or $[G_{nt}+s-H]^-$
		or 426.1	$[G_{nt}+p+H_2O-H]^-$
408.1	19.8	408.1	$[G_{nt}+s-H_2O-H]^-$

386.0	1.6	or 408.0 386.1	[G _{nt} +p-H] ⁻ [a ₂ -GH-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
368.1	18.7	or 386.0 368.1	[C _{nt} +p+H ₂ O-H] ⁻ [C _{nt} +s-H ₂ O-H] ⁻
346.1	5.2	or 368.0 346.1	[C _{nt} +p-H] ⁻ d ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	23.5	328.0	[c ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
310.1	2.1	310.0	[G _{nt} -H ₂ O-H] ⁻
306.1	18.5	306.0	w ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
288.0	27.1	288.0	[x ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.1	3.8	275.1	[sps-H] ⁻
257.1	8.2	or 275.1 257.1	[psp+H ₂ O-H] ⁻ [sps-H ₂ O-H] ⁻
239.0	2.4	or 257.0 239.1	[psp-H] ⁻ [psp-H ₂ O-H] ⁻
230.0	14.2	230.1	[sG-H ₂ O-H] ⁻
195.1	8.9	195.0	[ps+H ₂ O-H] ⁻
190.1	1.7	190.1	[sC-H ₂ O-H] ⁻
177.1	30.0	177.0	[ps-H] ⁻
158.9	10.5	159.0	[ps-H ₂ O-H] ⁻
150.0	19.2	150.0	G ⁻
110.1	3.8	110.0	C ⁻
97.0	19.4	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
79.0	43.2	or 97.0 79.0	[p+H ₂ O-H] ⁻ [p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

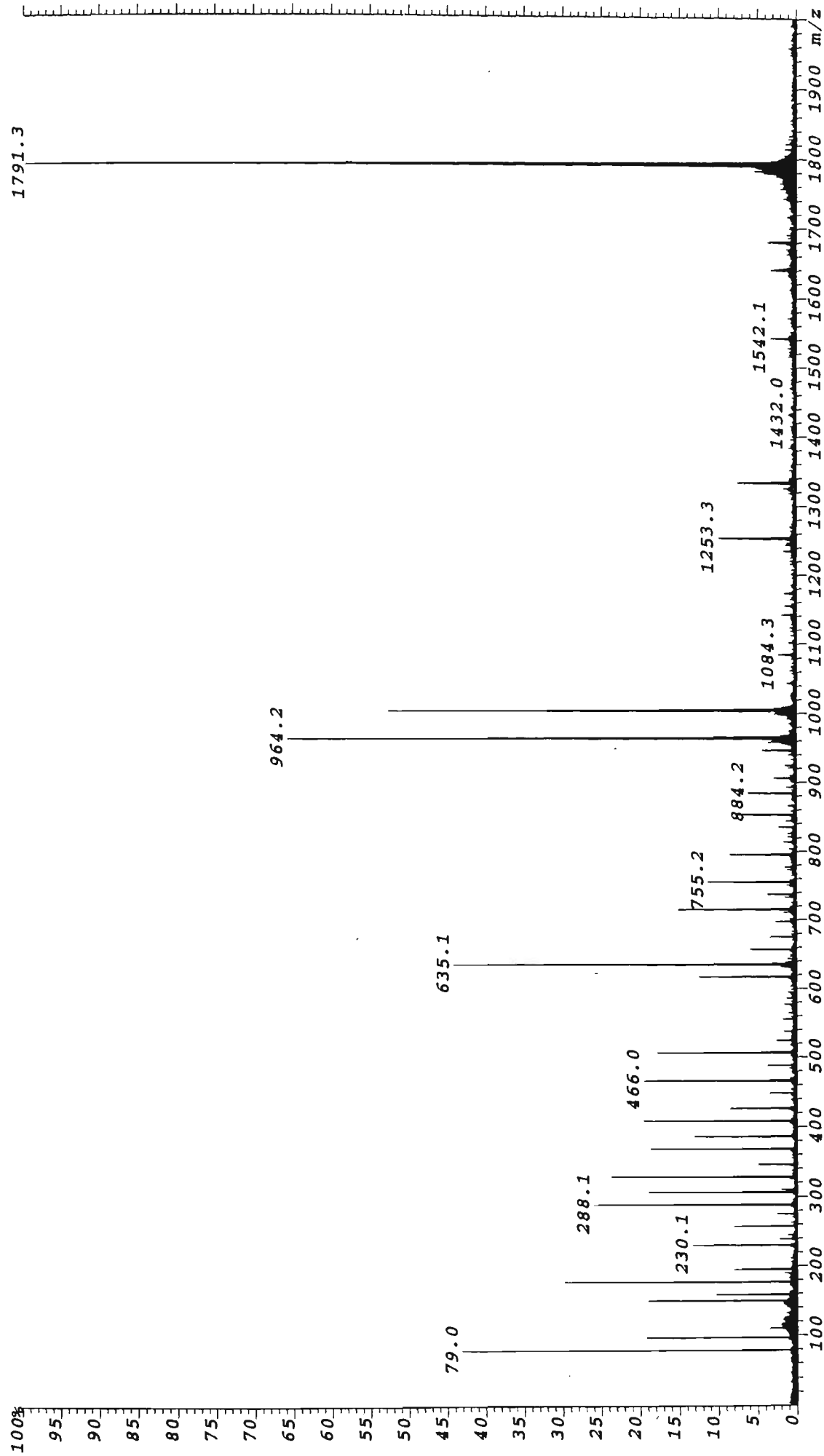


Figure 4.51: Negative Ion ESI-MS/MS spectrum of the [M-H]⁻ ion of 5'-d(GCCGGC)-3' (m/z 1791.3).

The propensity for base anion formation in the MS/MS of singly charged oligonucleotide anions is summarised in figure 4.52. The MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' both show thymine to be the most intense base anion with the overall propensity for charged base loss following the order, $T^- > G^- > C^- > A^-$. Charged thymine loss also generates the most intense base anion in the MS/MS spectrum of the oligonucleotide 5'-d(TCACGA)-3' followed by A^- and G^- with no ion present for C^- . In contrast, charged adenine loss yields the most intense base anion for 5'-d(ATGCAT)-3' followed by T^- and then G^- . These results may indicate a preference for the loss of charged adenine over thymine where adenine is the 5'-terminal base. The absence of ions for the loss of charged cytosine in both the MS/MS spectra of 5'-d(ATGCAT)-3' and 5'-d(TCACGA)-3' may indicate the lack of preference for the loss of C^- where it is not the 5'-terminal base. A small ion for C^- is present in the MS/MS spectrum of 5'-d(GCCGGC)-3' which may be attributed to the fact that this oligonucleotide comprises three C nucleobases with the only competition from nucleobase loss coming from G.

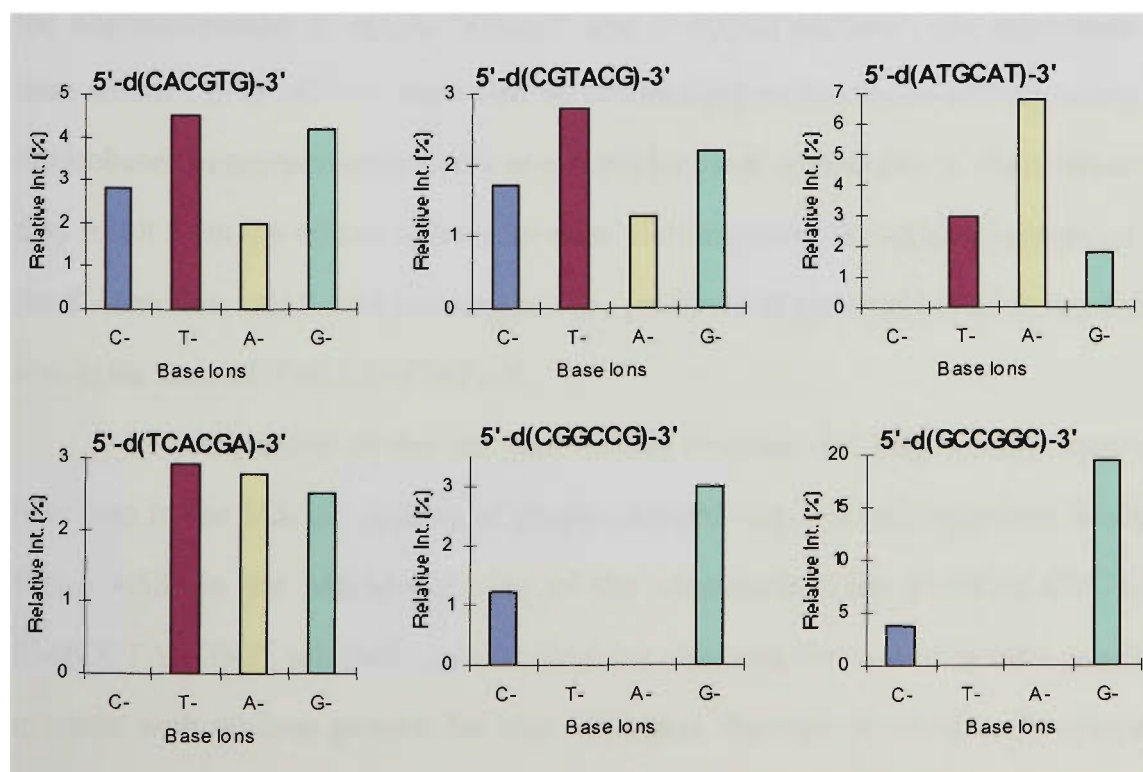


Figure 4.52: Comparisons of the relative intensities of base anions generated in the ESI-MS/MS spectra of singly charged oligonucleotide anions.

The observation that loss of charged thymine is favoured in the MS/MS spectra of singly charged oligonucleotides might be explained by the greater relative acidity of thymine compared with that of the other nucleobases. If base loss proceeds by the mechanism proposed by Rodgers *et al.*²⁰, following glycosidic bond cleavage, a proton-bound intermediate complex is formed between the cleaved base and a deprotonated phosphate. The ability of the nucleobase to be lost as an anion following dissociation of this proton-bound intermediate will thus depend on the stability of the base anion and thus the acidity of the base. From semiempirical calculations, Rodgers *et al.*²⁰ determined the relative acidities of the nitrogen bound to the ribose for each of the nucleobases to be T (328.0 kcalmol⁻¹) > G (330.5 kcalmol⁻¹) > A (332.9 kcalmol⁻¹) > C (338.9 kcalmol⁻¹). The results obtained for the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(TCACGA)-3' with regard to the loss of charged thymine are in agreement with these calculations. The observation that A⁻ was slightly more intense than G⁻ in the MS/MS spectrum for the oligonucleotide 5'-d(TCACGA)-3' does not follow this order and may be a consequence of the presence of two adenines and only one guanine in this oligonucleotide. In the MS/MS spectra of the [M-H]⁻ ions of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', the intensities of the base anions i.e. G⁻ > C⁻ > A⁻ may also be rationalised on the basis of their being two C nucleobases in comparison to only one A nucleobase. Alternatively, these observations may result from the effects of base location yielding more favourable base anion loss at the 5'-terminus, and would also explain the preferential observation of charged adenine loss in the case of 5'-d(ATGCAT)-3'.

A comparison of the ion abundances obtained for ions arising from neutral base loss in the MS/MS spectra of singly charged oligonucleotide anions is shown in figure 4.53. In the MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', relatively intense ions are observed for loss of neutral guanine and cytosine with no ions present for loss of neutral thymine or adenine. In contrast, the MS/MS spectra of 5'-d(ATGCAT)-3' and 5'-d(TCACGA)-3' each show an ion for the loss of neutral adenine with no ions present for the loss of neutral guanine and

cytosine. In all cases, no ions are detected for the loss of neutral thymine. This can be rationalised on the basis of the lower stability of neutral thymine as a leaving group compared with the other bases as discussed for the positive ion spectra in section 4.5.2.

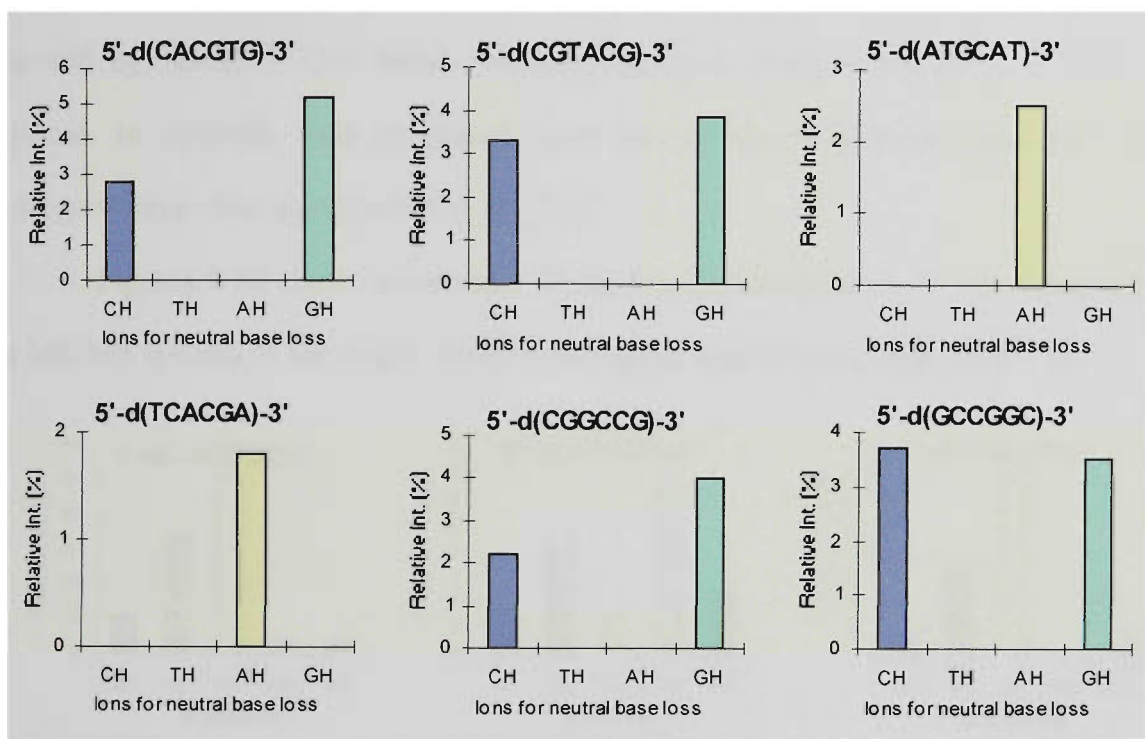


Figure 4.53: Comparisons of the relative intensities of the ions for neutral base loss generated in the ESI-MS/MS spectra of singly charged oligonucleotide anions.

The results obtained appear to suggest a trend whereby a distinct preference exists for the formation of ions owing to the loss of the 3'-terminal base as a neutral followed by the 5'-terminal base except in the case of thymine which is not observed to be lost as a neutral when located at either terminus. Furthermore, the results also suggest a lack of preference for the neutral loss of bases located internally within the oligonucleotide. In the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' loss of neutral guanine produces a marginally more intense ion than that of cytosine. Consistent with these observations are the MS/MS spectra of 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3' where in each case intense ions are observed for the loss of both neutral guanine and cytosine with the loss of neutral G giving rise to a more intense ion for 5'-d(CGGCCG)-3', and the loss of neutral C yielding a more intense ion for 5'-d(GCCGGC)-3'. The preference for loss of the 3'-terminal base as a neutral may be a

consequence of the greater stability of this ion to further decomposition in comparison to those formed via loss of a neutral base at other positions. The absence of a charged phosphate group to the 3' side of the 3'-terminal ribose would mean that for this ion to undergo further decomposition by the major decomposition route of nucleobase loss followed by ribose 3' C-O bond cleavage, the loss of an additional neutral base is required. In contrast, ions generated from loss of the 5'-terminal base can readily undergo w-type cleavage generating w_5^- ions.

Figures 4.54 shows a summary of the relative intensities of w ions generated in the MS/MS spectra of the singly charged precursor ions of each oligonucleotide.

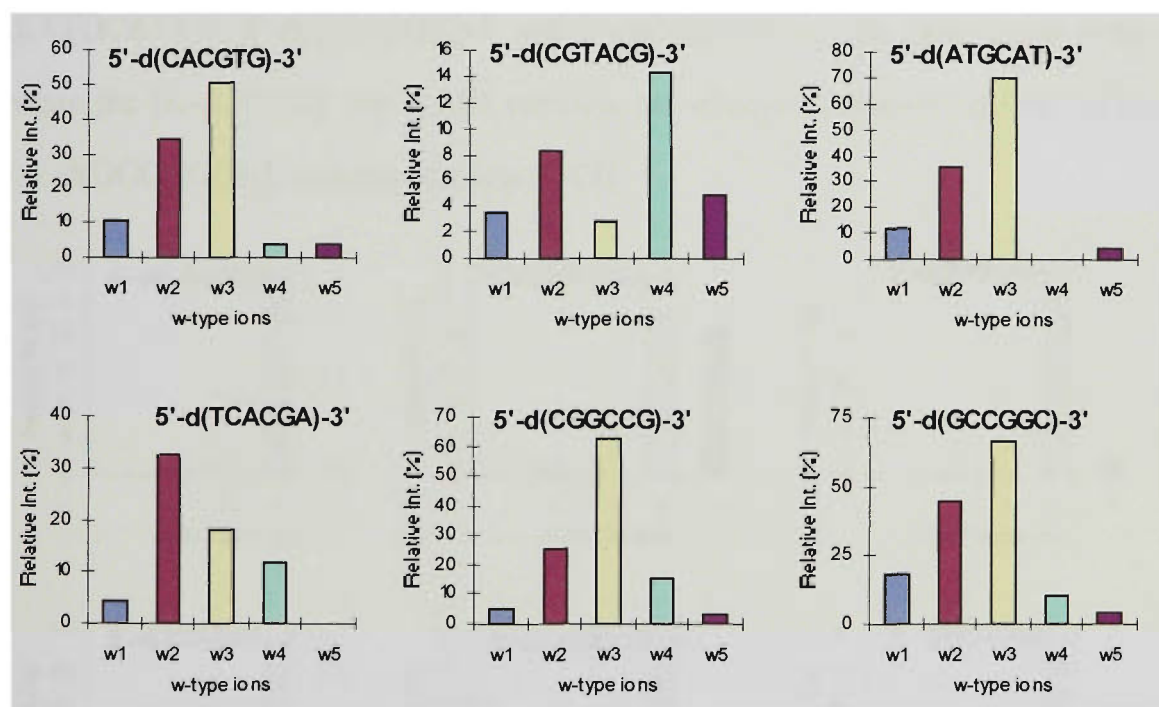


Figure 4.54: Comparison of the relative intensities of the various w ions generated in the ESI-MS/MS spectra of singly charged oligonucleotide anions.

The MS/MS spectra of 5'-d(CACGTG)-3', 5'-d(ATGCAT)-3', 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3', show predominant w_3^- ions followed by w_2^- ions. In the MS/MS spectrum of 5'-d(TCACGA)-3' a more intense ion for w_2^- formation is observed in preference to w_3^- . In contrast, the MS/MS spectrum of 5'-d(CGTACG)-3' shows the production of w_3^- to be the least favourable with the most intense w ions being w_4^- followed by w_2^- . This observation can be rationalised on the basis of the lack of preference for the loss of neutral thymine in the MS/MS of singly charged

oligonucleotide anions since the loss of TH would be the initial step preceding the formation of w_3^- for the oligonucleotide 5'-d(CGTACG)-3'. Similarly, the strong lack of preference for the formation of w_4^- in the MS/MS spectrum of 5'-d(ATGCAT)-3', and w_5^- in the MS/MS spectrum of 5'-d(TCACGA)-3' can also be explained by the fact that loss of TH would be the preliminary step preceding the formation of these w ions.

A comparison of the relative intensities of (a-B) ions generated in the MS/MS of singly charged precursor ions is shown in figure 4.55. The results show a distinct preference for the formation of $[a_4-B_4H-2H]^-$ ions for all oligonucleotides with the $[a_2-B_2H-2H]^-$ generally being the least favoured. For the oligonucleotides 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3' and 5'-d(CGGCCG)-3' the base lost in order to generate the $[a_4-B_4H-2H]^-$ ion is GH whereas the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(GCCGGC)-3' undergo the loss of CH.

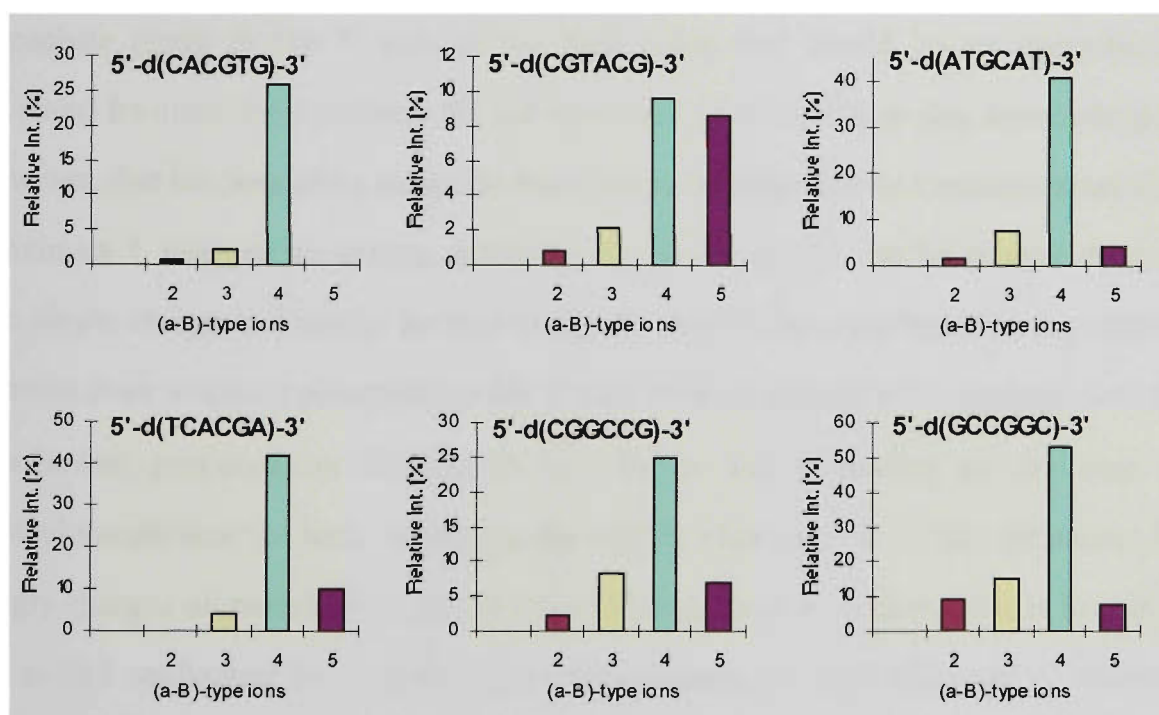


Figure 4.55: Comparison of the ion intensities of the various (a-B) ions generated in the ESI-MS/MS spectra of singly charged oligonucleotide anions.

In the case of 5'-d(CGTACG)-3' the loss of AH precedes formation of the $[a_4-B_4H-2H]^-$ ion, and compared with the other oligonucleotides, the intensity of this ion was found to be lower in comparison to the second and third most abundant (a-B) ions. This suggests that although base loss to generate the $[a_4-B_4H-2H]^-$ ion is favoured in

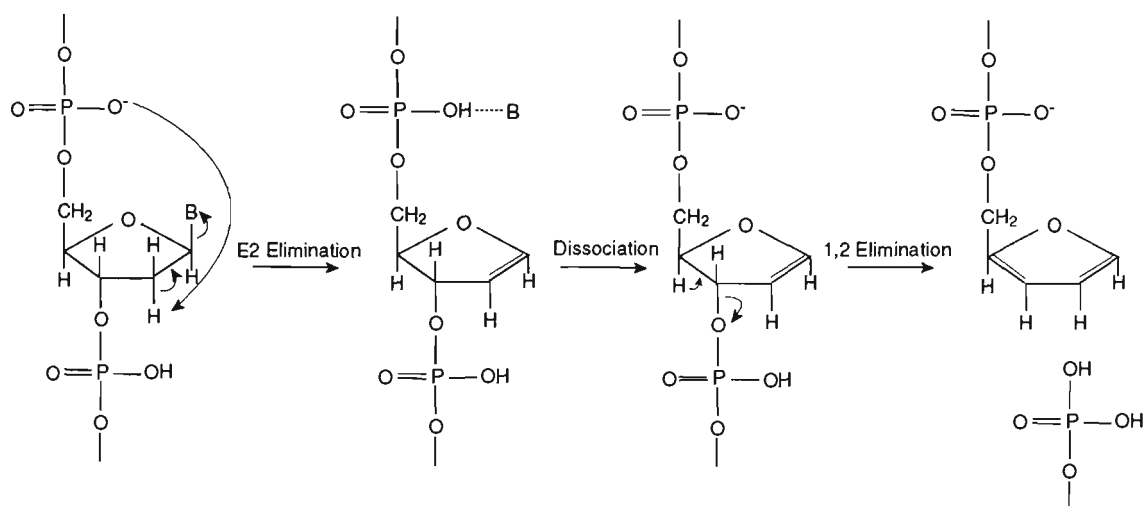
comparison to the formation of other (a-B) ions, this fragmentation pathway is more facile when it involves the loss of neutral CH or GH. As observed for w ions, the formation of (a-B) ions involving the loss of neutral TH is strongly disfavoured. This is particularly evident in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(ATGCAT)-3' in which the $[a_5-B_5H-2H]^-$ and $[a_2-B_2H-2H]^-$ ions respectively are the weakest (a-B) ions observed.

In the case of singly charged oligonucleotides, if initial base loss is assumed to catalyse ribose 3' C-O bond cleavage, the dominant w and (a-B) ions arising from this pathway will not only depend on the propensity for the base to be lost (via cleavage of the glycosidic bond), but will also be determined by the ability of the base to be lost as a neutral species (base acidity). Furthermore, the formation of w ions from initial base loss would be thought to be influenced by the location of the charge on the oligonucleotide backbone since base elimination catalysed by a deprotonated phosphate group to the 3' side of the base being lost would be an energetically favoured fragmentation pathway for the formation of w ions. It is also acknowledged, however, that the possibility exists for base loss to be catalysed by a remote rather than proximate 3' charged phosphate. Alternatively, w ions could also be generated where the single charge is initially located to the 5' side of the cleavage site e.g. charge transfer from a remote phosphate to the 5' side of the cleavage point analogous to the mechanism proposed by Bartlett *et al.*⁵ for w ion formation in the case of methylphosphonate w ions. Although the results obtained for w ion formation for singly charged oligonucleotide anions shows formation of w_3^- and w_2^- ions is favoured, no overall preference for specific base loss generating w-type cleavage is observed since the most intense w ions generated in the MS/MS spectrum of each oligonucleotide examined arises from preliminary loss of either CH or GH. Furthermore, the w_3^- ion was observed to be the most intense w ion unless ribose 3' C-O bond cleavage at this position involved either the loss of TH in the case of 5'-d(CGTACG)-3' or AH in the case of 5'-d(TCACGA)-3'. The most intense w ion generated in these cases was found to be w_4^- for 5'-d(CGTACG)-3' which arises from

initial loss of GH and w_2^- for 5'-d(TCACGA)-3' which arises from initial loss of CH. Thus, although base identity plays a role in influencing the extent of formation of w ions, other factors appear to be contributing to the strong preference for w_3^- and w_2^- ion formation. These trends may suggest that the charge may be preferentially located on the third and fourth phosphate groups which catalyses base elimination at these positions. Alternatively, the favourable production of w ions of this size may also be a result of energetically favoured fragmentation pathways occurring under the collision conditions employed. The effects of collision conditions on w ion formation are discussed in more detail in chapter five.

The formation of complementary (a-B) ions in the case of singly charged oligonucleotide precursor ions would appear to involve a different charge location on the oligonucleotide backbone than that involved in generating the corresponding w ions. This is since if the charge is located at the phosphate group to the 3' side of the base being lost as has been implicated in w ion formation, then the complementary (a-B)-type fragment formed will be a neutral species. Although the location of the charge to the 3' side of the cleavage site has been implicated in driving cleavage of the ribose 3' C-O bond, the formation of (a-B) ions from singly charged species would appear to require the location of the charge to the 5' side of the cleavage point. The formation of (a-B) ions from initial base loss upon MS/MS fragmentation of singly charged oligonucleotide anions can be rationalised via the mechanism shown in scheme 4.23. The mechanism by which base elimination is initiated is analogous to that proposed by Rodgers *et al.*²⁰ which involves proton transfer from the 2' ribose hydrogen to the phosphate anion to the 5' side of the base (see chapter one, scheme 1.2). It is acknowledged that base elimination could also be catalysed by other remote phosphate anions to the 5' side of the ribose from which the base is lost in an analogous mechanism to that proposed by Bartlett *et al.*⁵. For the charge to remain on the phosphate moiety, loss of a neutral base must occur upon dissociation of the proton-bound intermediate formed between the base and phosphate moiety. Subsequent 1,2 elimination of the phosphate group to the 3' side of the ribose from

which the base was lost with the charge residing on the 5' fragment results in the formation of an (a-B) ion. As with the formation of w ions, the driving force for the cleavage of the ribose 3' C-O bond is the formation of a stable furan product.



Scheme 4.23: Proposed mechanism for the formation of (a-B) ions in the ESI-MS/MS spectra of singly charged oligonucleotide anions

The more pronounced observation of w ions in the MS/MS spectra of singly charged oligonucleotide anions could be rationalised on the basis that the negative charge on the phosphate preferably initiates loss of the base to the 5' side rather than that to the 3' side owing to the former fragmentation pathway being the more favourable one in terms of activation energy. A possible rationalisation for the observations of abundant (a-B) ions from cleavage of the fourth base could be preferable charge location on the phosphate to the 5' side of the lost base which catalyses base loss as described in the mechanism above. The preferable location of the charge at this third phosphate group is also consistent with the observations for w ion formation which in almost all circumstances showed w_3^- ion formation to be the most

abundant. This is since the most abundant w ions formed where the charge is located on the third phosphate group would be those involving base loss and ribose 3' C-O bond cleavage to the 5' side of the deprotonated phosphate, i.e. w_3^- ions, since this phosphate can also catalyse the loss of the base to the 5' side.

4.6.2 *Fragmentation Pathways of Singly Charged Oligonucleotide Cations*

In order to investigate further the features observed in the positive ion MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', positive ion MS/MS spectra were acquired of the singly charged ions of two additional oligonucleotides, one of heterogenous base composition, 5'-d(ATGCAT)-3', and one consisting solely of C and G bases, 5'-d(CGGCCG)-3'. The oligonucleotide 5'-d(ATGCAT)-3' was examined to establish further the effects of base identity on the trends exhibited for the types of product ions observed, whereas the oligonucleotide 5'-d(CGGCCG)-3' was of interest because our previous studies had shown that G and C are the nucleobases which are most easily lost from positively charged precursor ions. The positive ion MS/MS spectra of the $[M+H]^+$ ions of 5'-d(ATGCAT)-3' and 5'-d(CGGCCG)-3' are shown in figures 4.56 and 4.57 respectively. The assignments of product ions observed in the MS/MS spectrum of 5'-d(ATGCAT)-3' are given in table 4.17, and the product ion assignments for the MS/MS spectrum of 5'-d(CGGCCG)-3' are given in table 4.18.

One obvious feature of the positive ion MS/MS spectrum of each oligonucleotide is the marked different appearance of the MS/MS spectrum of 5'-d(CGGCCG)-3' which shows fewer product ions generated with an overall poor signal-to-noise ratio. This may be explained by the facile loss of cytosine and guanine nucleobases as both charged and neutral species which can be understood in terms of their favourable leaving group ability and proton affinities (as described previously). Consequently, MS/MS fragmentation of 5'-d(CGGCCG)-3' predominantly gives rise to the observation of nucleobase ions, backbone fragments and smaller sequence and polynucleotide ions since the observation of larger product ion types incorporating bound cytosine and guanine nucleobases is unfavourable owing to the ease of loss of these nucleobases.

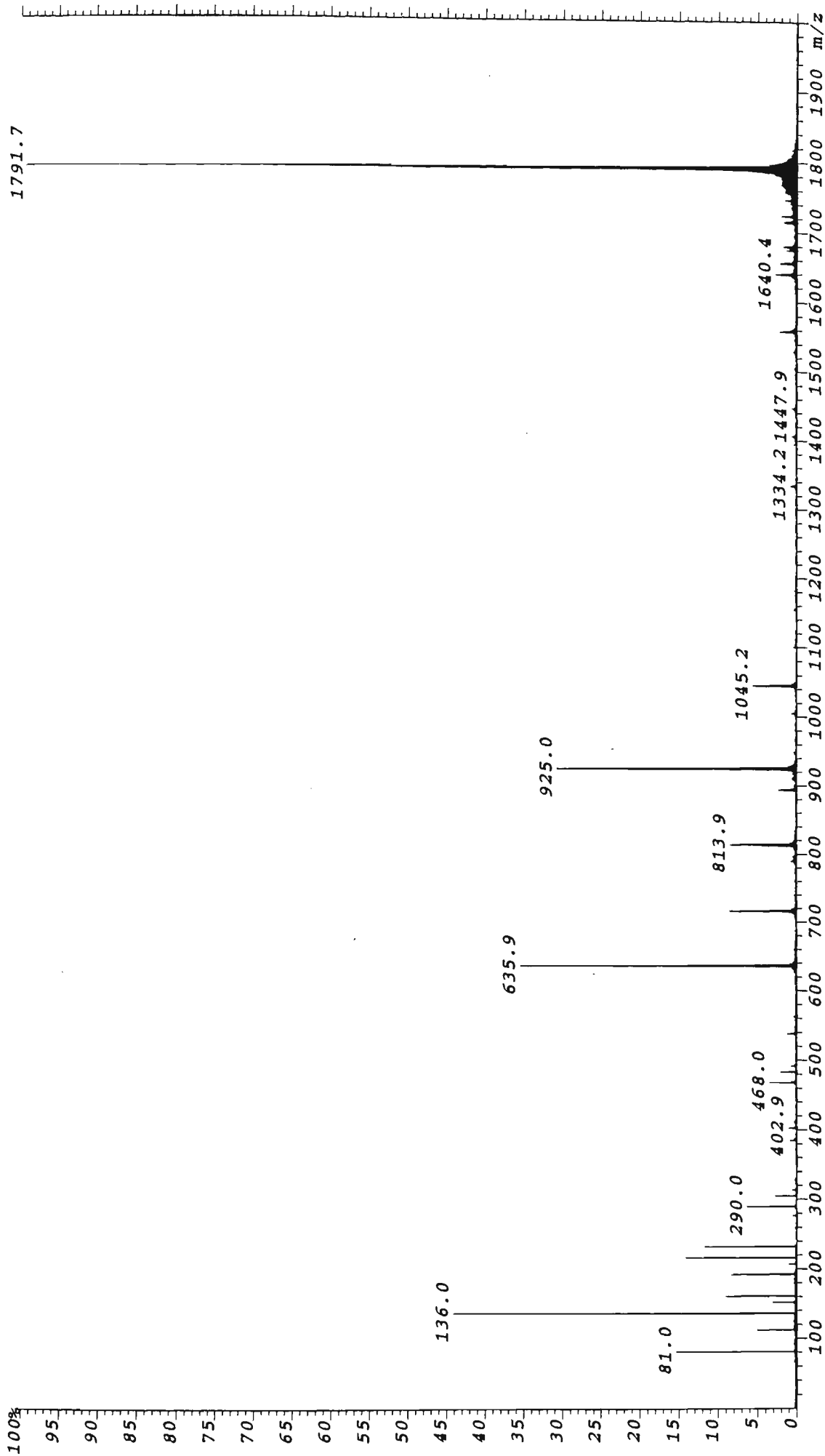


Figure 4.56: Positive Ion ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(ATGCAT)-3' (m/z 1791.7).

Table 4.17: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(ATGCAT)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1791.7	100	1791.4	$[M+H]^+$
1715.8	1.6	1715.8	$[2M-GH+2H]^2+$
1680.5	1.6	1680.3	loss of CH
1655.2	2.0	1655.3	loss of AH
1640.4	2.7	1640.3	loss of GH
1558.2	2.3	1558.3	$[w_5+2H]^+$
1334.2	0.8	1334.2	$[a_5-B_5H]^+$ and/or $[z_5-AH]^+$ and/or $[(TCGA)+s+H]^+$
		or 1334.2	$[(TCGA)+p+H_2O+H]^+$
1045.2	5.8	1045.2	$[a_4-B_4H]^+$ and/or $[z_4-CH]^+$ and/or $[(TCGA)-CH+H]^+$
925.0	31.0	925.1	$[w_3+2H]^+$
894.0	2.4	894.1	$[a_4-B_4H-GH]^+$ and/or $[z_4-CH-GH]^+$
813.9	8.6	814.1	$[d_3-GH+2H]^+$ and/or $[w_3-CH+2H]^+$
716.1	8.6	716.1	$[a_3-B_3H]^+$ and/or $[z_3-CH]^+$
635.9	35.6	636.1	$[w_2+2H]^+$ and/or $[d_2+2H]^+$
538.1	1.4	538.1	a_2^+ and/or z_2^+
492.1	0.8	492.1	$[A_{nt}+p+s+H]^+$
483.0	2.2	483.1	$[c_2-AH]^+$ and/or $[x_2-AH]^+$ and/or $[T_{nt}+p+s+H]^+$
468.0	4.6	468.1	$[C_{nt}+p+s+H]^+$
402.9	1.1	403.1	$[a_2-AH]^+$ and/or $[z_2-AH]^+$ and/or $[T_{nt}+s+H]^+$
		403.0	$[T_{nt}+p+H_2O+H]^+$
385.1	0.9	385.1	$[T_{nt}+s-H_2O+H]^+$
		or 385.0	$[T_{nt}+p+H]^+$
314.1	0.6	314.1	c_1^+ and/or $[A_{nt}+H]^+$
305.1	2.9	305.1	$[T_{nt}+H]^+$
290.0	6.3	290.1	$[C_{nt}+H]^+$
277.0	0.6	277.0	$[s+p+s+H]^+$ and/or $[p+s+p+H_2O+H]^+$
232.0	11.9	232.1	$[sG-H_2O+H]^+$
216.0	14.4	216.1	$[sA-H_2O+H]^+$
207.0	1.0	207.1	$[sT+H]^+$
192.0	8.5	192.1	$[sC-H_2O+H]^+$
161.0	9.3	161.0	$[s+p-H_2O+H]^+$
152.0	3.2	152.1	$[GH+H]^+$
136.0	44.1	136.1	$[AH+H]^+$
112.0	5.3	112.1	$[CH+H]^+$
81.0	15.6	81.0	$[s-H_2O+H]^+$

*calculated monoisotopic mass

s = deoxyribose- H_2O ($C_5H_8O_2$), M_r = 98.0368 Da

p = PO_3H , M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

$[(B_1...B_n)]$ denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

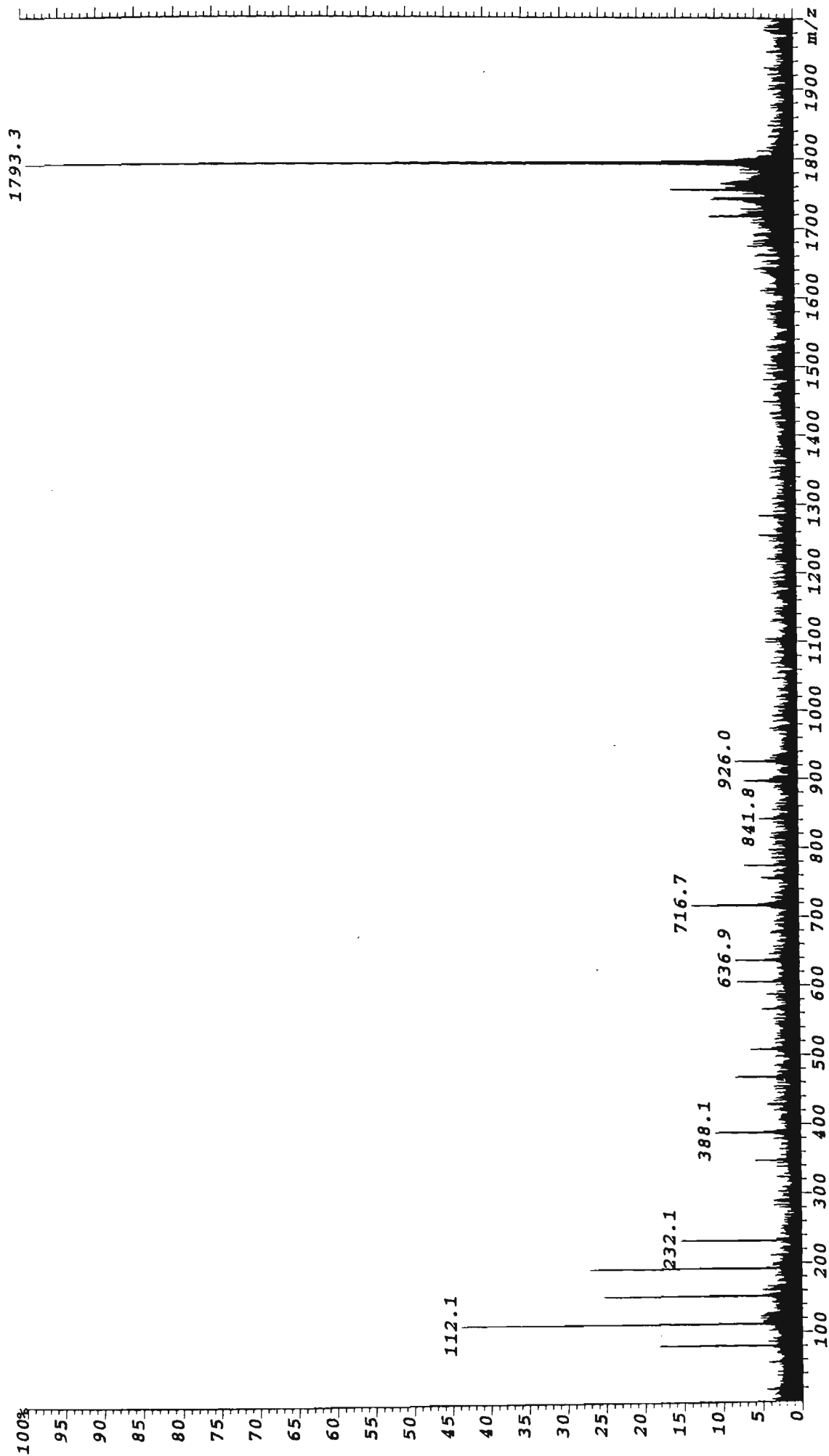


Figure 4.57: Positive Ion ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGGCCG)-3' (m/z 1793.3).

Table 4.18: Assignment of product ions observed in the ESI-MS/MS spectrum of the oligonucleotide 5'-d(CGGCCG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z (calc.)</i>	<i>Assignment</i>
1793.3	100	1793.3	[M+H] ⁺
1737.8	10.9	1737.8	[2M-CH+2H] ²⁺
1717.8	11.0	1717.8	[2M-GH+2H] ²⁺
1255.2	5.1	1255.2	[w ₄ +2H] ⁺ and/or [d ₄ +2H] ⁺ and/or [(GGCC)+H ₂ O+H] ⁺
926.0	8.1	926.1	[w ₃ +2H] ⁺ and/or [(GCC)+H ₂ O+H] ⁺
899.0	7.0	899.1	[(GG)+p+s+p-H ₂ O+H] ⁺ and/or [(CC)+p+s+p+s-2H ₂ O+H] ⁺
775.1	7.0	775.1	[w ₃ -GH+2H] ⁺ and/or [(CC)+s+p+H ₂ O+H] ⁺
		or 775.1	[b ₃ -CH+2H] ⁺ and/or [(GG)+s+H ₂ O+H] ⁺
756.9	4.9	757.1	[x ₃ -GH] ⁺ and/or [(CC)+s+p+H] ⁺
		or 757.1	[a ₃ -CH] ⁺ and/or [(GG)+s+H] ⁺
		or 757.1	[(GG)+p+H ₂ O+H] ⁺
716.7	13.7	717.1	[a ₃ -B ₃ H] ⁺ and/or [z ₃ -CH] ⁺ and/or [(GC)+s+H] ⁺
		or 717.1	[(GC)+p+H ₂ O+H] ⁺
636.9	8.3	637.1	[w ₂ +2H] ⁺ and/or [d ₂ +2H] ⁺ and/or [(GC)+H ₂ O+H] ⁺
606.1	8.1	606.1	[a ₃ -CH-GH] ⁺ and/or [z ₃ -CH-CH] ⁺ and/or [G _{nt} +s+p+s+H] ⁺
566.1	4.9	566.1	[a ₃ -GH-GH] ⁺ or [z ₄ -GH-CH] ⁺ or [C _{nt} +s+p+s+H] ⁺
508.1	5.7	508.0	[x ₂ -CH] ⁺ and/or [c ₂ -CH] ⁺ and/or [G _{nt} +s+p+H] ⁺
468.1	8.5	468.0	[C _{nt} +s+p-H ₂ O+H] ⁺
388.1	11.0	388.1	[a ₂ -B ₂ H] ⁺ and/or [z ₂ -GH] ⁺ and/or [C _{nt} +s+H] ⁺
		or 388.0	[C _{nt} +p+H ₂ O+H] ⁺
348.1	5.8	348.1	[w ₁ +2H] ⁺ and/or [G _{nt} +H ₂ O+H] ⁺
232.1	15.6	232.1	[sG-H ₂ O+H] ⁺
192.1	27.3	192.1	[sC-H ₂ O+H] ⁺
161.1	4.9	161.0	[s+p-H ₂ O+H] ⁺
152.1	25.5	152.1	[GH+H] ⁺
112.1	44.0	112.1	[CH+H] ⁺
81.0	18.5	81.0	[s-H ₂ O+H] ⁺

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₈O₂), M_r = 98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Figure 4.58 shows a summary of the relative intensities of nucleobase ions generated in the MS/MS spectra of protonated 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3' and 5'-d(CGGCCG)-3'. The first three of these oligonucleotides, which comprise all four of the major nucleobases, show ions only for protonated cytosine, guanine and adenine. The absence of ions for protonated thymine is consistent with the significantly lower proton affinity of thymine in comparison to the other nucleobases. A comparison of the relative intensities of the protonated bases observed in the MS/MS spectrum of each oligonucleotide gives rise to a different

ordering for the propensity of formation of protonated C, A and G in each case. As explained previously, these observations can be attributed to the fact that the differences in proton affinities of these nucleobases is so small that intramolecular interactions of these nucleobases within the oligonucleotide could alter their proton affinity sufficiently so as to cause a change in the overall preference for base cation formation.

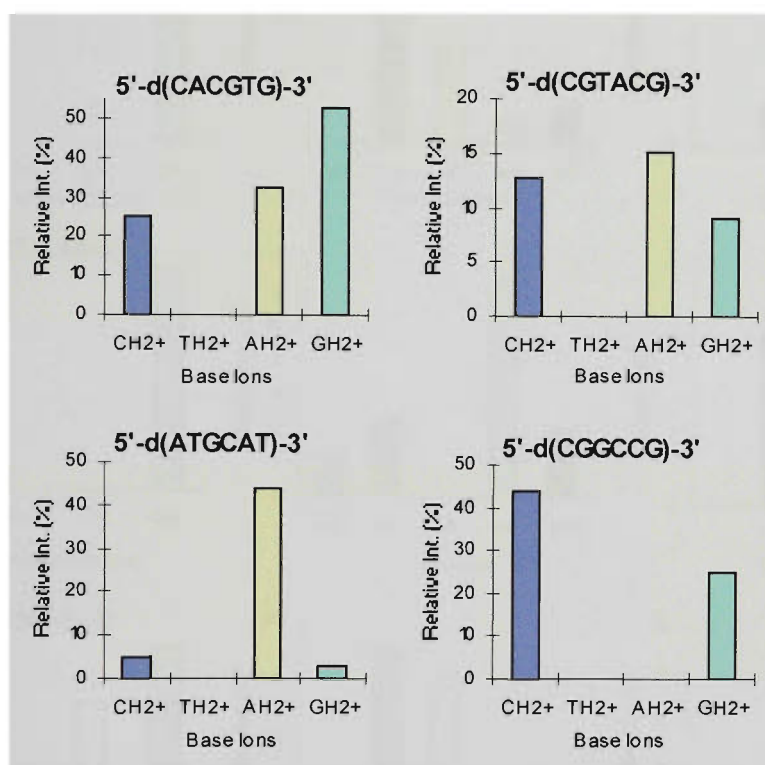


Figure 4.58: Comparisons of the relative intensities of base ions generated in the ESI-MS/MS spectra of singly charged oligonucleotide cations.

Figure 4.59 summarises the relative intensities of the product ion types arising from the fragmentations pathways of neutral nucleobase loss and the formation of the w and (a-B) sequence ions observed in the MS/MS spectrum of each of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3' and 5'-d(CGGCCG)-3'. In the case of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', ions for neutral base loss are only observed for the loss of cytosine and guanine with an order of GH > CH. The absence of ions for loss of adenine and thymine may be rationalised by the comparatively lower stability of the tautomeric forms of these nucleobases as leaving groups as described previously. It is interesting that the MS/MS spectrum of

5'-d(ATGCAT)-3' shows an ion for the loss of neutral adenine in addition to ions for the loss of neutral guanine and cytosine which may suggest that the location of adenine at the 5'-terminal position promotes favourable loss of neutral adenine.

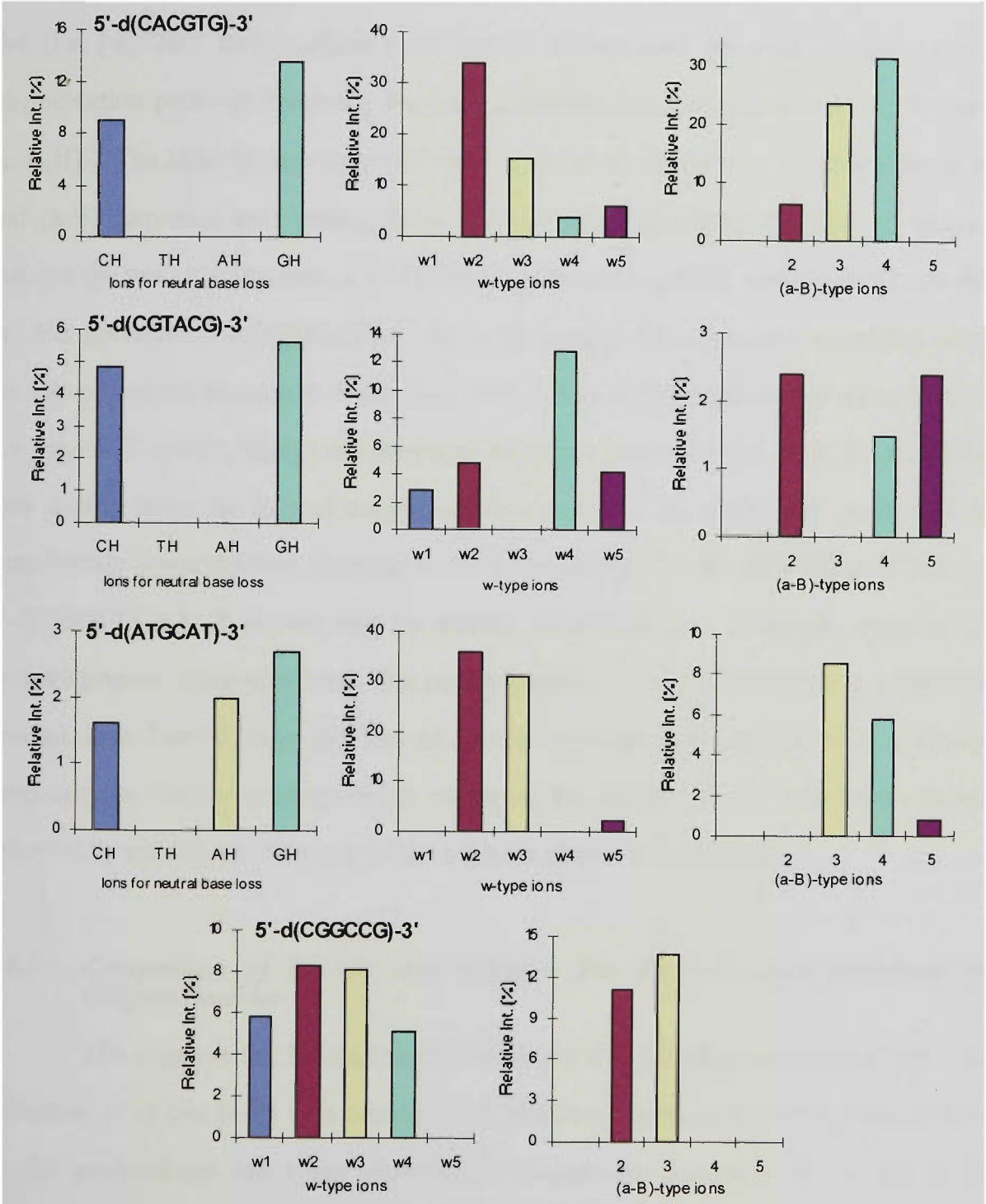


Figure 4.59: Relative intensities of product ions arising from neutral base loss and the formation of w and (a-B) ions in the ESI-MS/MS spectra of singly charged oligonucleotide cations.

From the trends for formation of w and (a-B) ions, it can be seen that for all three oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(ATGCAT)-3', w and

(a-B) ions are not observed for cleavages to the 3' side of a thymine-containing ribose. These observations are further consistent with the lack of preference for loss of thymine as discussed above. The MS/MS spectrum of 5'-d(CACGTG)-3' yields the most intense w and (a-B) ions from fragmentation involving the loss of neutral guanine G-4 (i.e. $[w_2+2H]^+$ and $[a_4-B_4H]^+$). Relatively intense ions are also present for the fragmentation pathway involving the loss of neutral cytosine C-3 (i.e. $[w_3+2H]^+$ and $[a_3-B_3H]^+$). The MS/MS spectrum of 5'-d(CGTACG)-3', also shows the predominant w and (a-B) sequence ions arising from fragmentation involving the loss of neutral guanine (in this case the loss of G-2) which generates $[w_4+2H]^+$ and $[a_2-B_2H]^+$. In the MS/MS spectrum 5'-d(ATGCAT)-3', the w and (a-B) sequence ions arising from the loss of neutral guanine G-3 (i.e. $[w_3+2H]^+$ and $[a_2-B_2H]^+$) and neutral cytosine C-4 (i.e. $[w_2+2H]^+$ and $[a_4-B_4H]^+$) are the major sequence ions observed, with the sequence ions arising from the loss of neutral adenine A-5 (i.e. $[w_1+2H]^+$ and $[a_5-B_5H]^+$) of significantly lower relative abundance. It is interesting that the MS/MS spectrum of 5'-d(CGGCCG)-3' shows no ions for neutral nucleobase loss, although w and (a-B) ions are present in the spectrum. One possible reason for this observation may be that the product ions formed from the loss of neutral cytosine and guanine have a strong propensity to further decomposition owing to the facile loss of these nucleobases followed by subsequent cleavage of the adjacent ribose 3' C-O bond.

4.6.3 *Comparison of Positive and Negative Ion Fragmentation Pathways of Oligonucleotides*

The negative ion MS/MS spectra of singly charged oligonucleotides show the formation of w and (a-B) ions arising from backbone cleavage following base loss to be the predominant ion types observed. Although the formation of w and (a-B) sequence ions is still observed to be a significant fragmentation pathway in the positive ion mode, the observation of these sequence ions as intact species is not as prevalent owing to facile nucleobase loss from these fragments. Consequently, the product ions that tend to dominate positive ion MS/MS spectra of singly charged oligonucleotides are ions arising from multiple losses of nucleobases (predominantly cytosine and

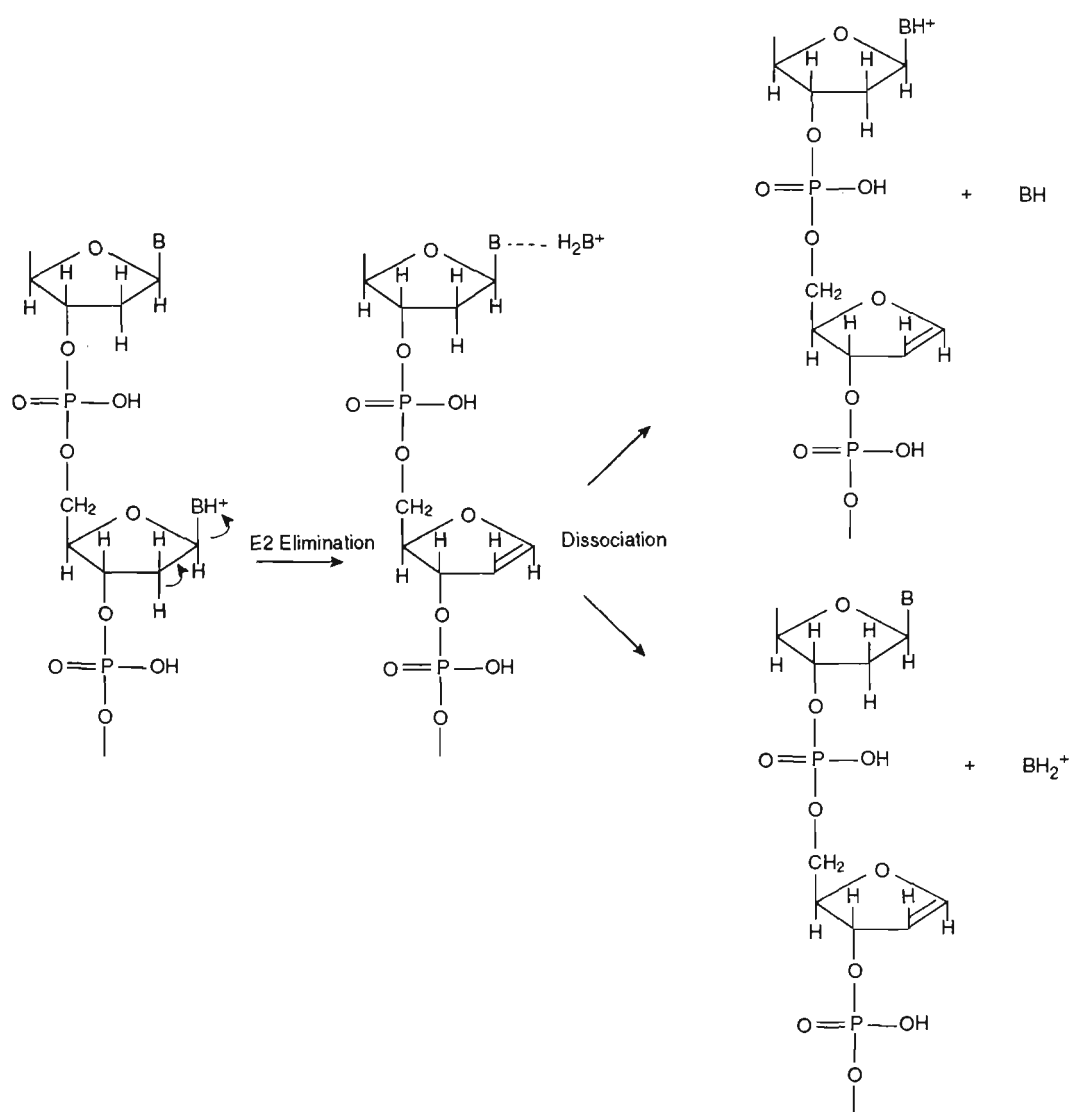
guanine). The observed differences in the positive and negative ion MS/MS of oligonucleotides are consistent with the effect of charge location in directing fragmentation pathways.

In the case of oligonucleotide anions, the location of the negative charge on the phosphodiester group has been implicated in catalysing base loss²⁰. Subsequent cleavage of the adjacent ribose 3' C-O bond to yield w and (a-B) sequence ions is thought to be driven by the formation of a stable furan product². The mechanism proposed by Rodgers *et al.*²⁰ for nucleobase loss involves a base-induced elimination process (see scheme 1.2, chapter one) and the formation of an intermediate proton-bound collision complex between the cleaved base and the 3' protonated phosphate group. Upon dissociation of the complex, retention of the proton on the phosphate group gives rise to loss of a base anion, whereas transfer of the proton to the base gives rise to neutral base loss and charge retention on the phosphate group of the oligonucleotide. The dominant observation of w and (a-B) sequence ions in the MS/MS spectra of singly charged oligonucleotide anions indicate preferable retention of the negative charge on the phosphate group following nucleobase loss. These observations for singly charged anions of hexamers would likely arise from the gas-phase basicity of the cleaved base, B⁻, being greater than that of the oligonucleotide fragment, [M-BH-H]⁻. Consequently, the cleaved base would compete more favourably for the proton than the phosphate group upon dissociation of the proton-bound intermediate complex. The relative abundances of base ions observed in the MS/MS spectra of singly charged deprotonated oligonucleotides show good correlation with calculated values of the relative acidities of the bases determined by Rodgers *et al.*²⁰ (T > G > A > C) although base location is also an important factor. The loss of charged thymine is observed to be highly favoured, and a lack of preference is exhibited for the loss of C⁻ except when located at the 5'-terminal position. Base location is also seen to have an effect in the case of charged adenine loss which is observed to be highly favourable when it is located at the 5'-terminal position. The propensities for neutral nucleobase base loss on the other hand are observed to be strongly influenced by base

location. A distinct preference is exhibited for loss of the 3'-terminal base as a neutral followed by the 5'-terminal base, with the loss of bases located internally as neutrals not favourable presumably because these ions if formed fragment further to yield w and (a-B) ions. These trends for neutral base loss are, however, overridden by the strong lack of preference for the loss of neutral thymine, which is not observed to be lost as a neutral regardless of its location. The results obtained for w and (a-B) ion formation in the negative ion MS/MS spectra of singly charged oligonucleotides, suggest charge location to be an important, if not the major, factor influencing the major ion types observed. The MS/MS spectra of almost all oligonucleotides yield relatively intense w_3^- ions suggesting favourable ribose C-O bond cleavage to the 3' side of the third residue. An exception is when thymine is located to the 5' side of the cleavage site which results in a significant reduction in the intensity of the w_3^- ion. These observations are consistent with the lack of preference for loss of neutral thymine observed in the corresponding MS/MS spectra of these oligonucleotides. In the case of (a-B) ions, a preference is exhibited for base loss and ribose 3' C-O bond cleavage of the fourth residue resulting in relatively intense $[a_4-B_4H-2H]^-$ ions. These observations may be rationalised by a preference for location of the negative charge in the case of singly charged hexamers on the phosphate group of the third residue which could catalyse loss of the adjacent bases to the 5' and 3' side. Therefore, where the phosphate group catalyses loss of the adjacent 5' base (the third base), subsequent cleavage of the ribose 3' C-O bond of this residue would result in the formation of w_3^- ions (assuming neutral base loss and retention of the negative charge on the phosphate group). On the other hand, if the deprotonated phosphate group catalyses loss of the adjacent 3' base (the fourth base), subsequent ribose 3' C-O bond cleavage following neutral base loss would give rise to $[a_4-B_4H-2H]^-$ ions.

The MS/MS spectra of singly charged protonated oligonucleotides, on the other hand, show nucleobase loss from the precursor ion, and from fragments formed from backbone cleavage to dominate the types of ions observed. These observations are consistent with the positive charge being located on the nucleobase and facilitating

N-glycosidic bond cleavage and base elimination, analogous to the mechanism for nucleobase loss in singly charged protonated oligonucleotides proposed by Rodgers *et al.*¹⁹, which is shown in scheme 4.24. This mechanism involves a syn E2 elimination mechanism whereby base loss is driven by protonation and the formation of a stable tautomer of the base. Following elimination of the base, a collision complex is formed between the protonated base and the remaining oligonucleotide which will dissociate to yield either the protonated nucleobase and a neutral molecule for the oligonucleotide, or the neutral base and an ion for the protonated oligonucleotide. Dissociation of this complex will thus yield either the charged or neutral nucleobase depending on the propensity for the charge to be retained on the nucleobase or oligonucleotide respectively.



Scheme 4.24: Proposed mechanism for base loss in the ESI-MS/MS spectra of singly charged protonated oligonucleotides¹⁹.

The characteristic series of product ion types generated in the MS/MS spectra of protonated oligonucleotides can therefore be rationalised from both the relative proton affinities of the bases, in addition to their ability to act as leaving groups. In the MS/MS spectrum of 5'-d(CACGTG)-3', a predominant series of product ions is generated comprising adenine attached to varying portions of sugar-phosphate backbone owing to the loss of CH and/or GH on neighbouring residues. Similarly, in the MS/MS spectrum of 5'-d(CGTACG)-3', the loss of CH and GH gives rise to a major product ion series owing to the (TA) dinucleotide attached to varying portions of sugar-phosphate backbone. In the MS/MS spectrum of 5'-d(ATGCAT)-3', series of sequence ions are observed owing to the loss of CH and GH. Therefore, in each case the types of product ions formed are found to be strongly affected by the facile loss of guanine and cytosine which is consistent with both the calculated proton affinities ($C > G > A \gg T$)¹⁹ and leaving group abilities of the nucleobases ($C > G \gg A \gg T$)¹⁹. The results for the relative intensities of ions owing to charged and neutral nucleobase loss in the MS/MS spectra of protonated oligonucleotides are also consistent with both the calculated proton affinities and leaving group abilities of the nucleobases. In all MS/MS spectra of all thymine-containing oligonucleotides a strong lack of preference is exhibited for the loss of thymine with no ions detected for protonated thymine or for loss of neutral thymine from the precursor ion. Regarding base ion formation, ions for protonated cytosine, guanine and adenine are observed in each case, although different orders for the relative abundances of these ions is observed for each oligonucleotide which can be rationalised by the small differences in proton affinities of these three bases. In the case of neutral base loss with the chain remaining intact, a preference is exhibited for loss of neutral guanine with the loss of neutral cytosine also highly favoured. Where the oligonucleotide only comprises cytosine and guanine, 5'-d(CGGCCG)-3', the facile loss of these nucleobases gives rise to extensive decomposition with no ions detected for neutral nucleobase loss. In addition to the lack of preference for the loss of neutral thymine, no ions were observed for the loss of neutral adenine except when it was the 5'-terminal base, suggesting that base location

may also be an important factor influencing the propensities for neutral base loss in the MS/MS spectra of protonated oligonucleotides. In contrast to deprotonated oligonucleotides, the relative intensities of the various w and (a-B) sequence ions in the MS/MS spectra of protonated oligonucleotides are observed to be strongly influenced by the propensities for neutral nucleobase loss. The MS/MS spectra for all protonated oligonucleotides show abundant w and (a-B) ions for cleavages to the 3' side of residues containing guanine and cytosine, which is consistent with the observed facile neutral loss of these bases. In all MS/MS spectra of thymine-containing oligonucleotides, w and (a-B) ions which involve ribose 3' C-O bond cleavage to the 3' side of a thymine-containing residue are absent.

The preferences for charge location for protonated and deprotonated oligonucleotides give rise to some major differences in the predominant fragmentation pathways and ion types observed. In the case of singly charged oligonucleotide anions, charge location on the phosphate backbone directs the major fragmentation pathways observed. The results obtained show a general preference for the same w and (a-B) sequence cleavages which may arise from a preference for charge location in the case of singly charged hexamers being on the middle phosphodiester residue. In the case of protonated oligonucleotides, the location of the positive charge on the nucleobases which facilitates base loss, influences the types of sequence cleavages observed. The major fragmentation pathways generated are therefore indicative of the preferential location of the positive charge on the oligonucleotide, and are consistent with the relative proton affinities and leaving group abilities of the nucleobases.

REFERENCES

- (1) McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.
- (2) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.
- (3) Barry, J. P.; Vouros, P.; Van Schepdael, A.; Law, S.-J. *J. Mass Spectrom* **1995**, *30*, 993-1006.
- (4) Gentil, E.; Banoub, J. *J. Mass Spectrom* **1996**, *31*, 83-94.
- (5) Bartlett, M. G.; McCloskey, J. A.; Manalili, S.; Griffey, R. H. *J. Mass Spectrom* **1996**, *31*, 1277-1283.
- (6) Ni, J.; Pomerantz, S. C.; McCloskey, J. A. *Nucleic Acids Symposium Series* **1996**, *35*, 113-114.
- (7) Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893-4897.
- (8) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
- (9) Little, D. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **1995**, *117*, 6783-6784.
- (10) Little, D. P.; Thannhauser, T. W.; McLafferty, F. W. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2318-2322.
- (11) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A.; McLafferty, F. W. *J. Am. Chem. Soc.* **1996**, *118*, 9352-9359.
- (12) McLuckey, S. A.; Habibi-Gourdarzi, S. *Journal of the American Society for Mass Spectrometry* **1994**, *5*, 740-747.
- (13) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom* **1995**, *30*, 1222-1229.
- (14) McLuckey, S. A.; Goeringer, D. E. *Anal. Chem.* **1995**, *67*, 2493-2497.
- (15) McLuckey, S. A.; Vaidyanathan, G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 1-16.
- (16) Loo, J. A.; Sannes, K. A.; Hu, P.; Mei, H. Y.; Mack, D. *Studying Non Covalent Protein-RNA Interactions by Electrospray Ionization Mass Spectrometry*; 44th ASMS Conference on Mass Spectrometry and Allied Topics, Santa Fe: Portland, Oregon, USA, 1996, pp 1.
- (17) Ni, J.; Mathews, M. A. A.; McCloskey, J. A. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 535-540.
- (18) Phillips, D. R.; McCloskey, J. A. *Int. J. Mass Spectrom. Ion Proc.* **1993**, *2*, 402-412.
- (19) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1995**, *148*, 1-23.
- (20) Rodgers, M. T.; Campbell, S.; Marzluff, M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.

Chapter 5

FACTORS INFLUENCING THE FRAGMENTATION OF OLIGONUCLEOTIDE ANIONS IN ELECTROSPRAY TANDEM MASS SPECTRA

5.1 INTRODUCTION

This chapter describes a detailed examination of the influence of oligonucleotide sequence and charge state on the fragmentation observed in tandem mass spectra of oligonucleotide anions. Since the observation that nucleobase loss is the first major step in the decomposition of oligonucleotides followed by the subsequent cleavage of the ribose C-O bond to the 3' side of the ribose from which the base has been lost, the exact mechanism by which this occurs has been the subject of a number of studies¹⁻³. The charge state of the precursor ion, and the location of charges along the oligonucleotide backbone have been shown to be important factors implicated in directing the major fragmentation pathways observed in the MS/MS of oligonucleotides. This has derived from previous observations, such as the preference for 5' terminal base loss compared with 3'-terminal base loss, which show that the location of a charged phosphate group to the 3' side of the residue from which the base has been lost plays an important role in driving this fragmentation pathway^{1,4-6}. Consequently, mechanisms have been proposed involving the catalysis of base elimination by initial abstraction of the 2'-ribose hydrogen to the phosphate anion^{1,7}. Other workers, however, have shown that the presence of a 3' ribose phosphate is not essential for the occurrence of this pathway and have suggested that it is possible for base loss to be catalysed by other phosphate groups or a 3' hydroxyl group^{3,8}. The facile subsequent cleavage of the ribose

3' C-O bond has been rationalised on the basis of the formation of a stable furan product following 1,2 elimination of the 3' phosphate group⁹. The precursor ion charge state has also been shown to affect the propensity for a base to be lost as a charged or neutral species. This has led to the suggestion that following glycosidic bond cleavage, an intramolecular proton-bound complex is formed between the base and the phosphodiester linkage. Dissociation of this complex then yields either the neutral or charged base depending on the relative proton affinities of the products formed. In addition to charge state effects, several studies have also shown base identity to play an important role^{1,7} in influencing base loss and also the propensity for a base to be lost as a charged or neutral species. This is because the propensity for loss of a base anion is also dependent on the relative acidities of the nucleic acid bases which is determined by the energy of dissociation of the N-H bond. The role played by each of these factors in the fragmentation of oligonucleotides observed in MS/MS spectra generated on hybrid-sector instrumentation has been explored here and compared to the observations made in studies conducted on other types of instruments in which the collision energy and/or the dissociation times may be very different.

5.2 THE EFFECTS OF OLIGONUCLEOTIDE SEQUENCE

The effects of oligonucleotide sequence such as base identity and base location, on the fragmentation of oligonucleotide anions were investigated from analysis of the MS/MS spectra of the doubly charged anions ($[M-2H]^{2-}$) of fourteen hexamers. These included three oligonucleotides of unibase sequence (5'-d(AAAAAA)-3', 5'-d(CCCCCC)-3', 5'-d(TTTTTT)-3'), nine oligonucleotides comprising each of the four major nucleobases (5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(AGGCCT)-3',

5'-d(ATGCAT)-3', 5'-d(ATCGAT)-3', 5'-d(GCCATG)-3', 5'-d(TACGTA)-3', 5'-d(TCACGA)-3', 5'-d(TCGTGA)-3'), and two oligonucleotides comprising only cytosine and guanine nucleobases (5'-d(CGGCCG)-3', 5'-d(GCCGGC)-3'). The results from the MS/MS spectra were analysed in the context of the major dissociation pathways of nucleobase loss and cleavage of the ribose 3' C-O bond resulting in the formation of w and (a-B) sequence ions. The fragmentation of unibase oligonucleotides was explored in order to ascertain the effects of charge location in directing the major fragmentation pathways observed, since, with oligonucleotides of heterogeneous sequence, the competition for different fragmentation pathways is also influenced by the differing proton affinities and leaving group abilities of the various nucleobases. The fragmentation of the set of oligonucleotides of heterogeneous nucleobase composition was then investigated in order to examine the contributions of the effects of base identity in addition to those of charge location.

5.2.1 Fragmentation of Unibase Oligonucleotides

Figure 5.1(a)-(c) shows the MS/MS spectra of the $[M-2H]^{2-}$ ions of the unibase oligonucleotides 5'-d(AAAAAA)-3', 5'-d(CCCCCC)-3', and 5'-d(TTTTTT)-3'. The assignments and relative abundance of the product ions arising from nucleobase loss, and w and (a-B) sequence ions are given table 5.1 (the complete product ion assignments for each spectrum is given in appendix 5.1). The MS/MS spectrum of each oligonucleotide shows dissociation to occur predominantly via nucleobase loss and cleavage of the adjacent ribose 3' C-O bond giving rise to w and (a-B) ions. Sequence ions owing to other backbone cleavages are also present although these are generally much less intense. An extensive number of product ions are generated in the low m/z

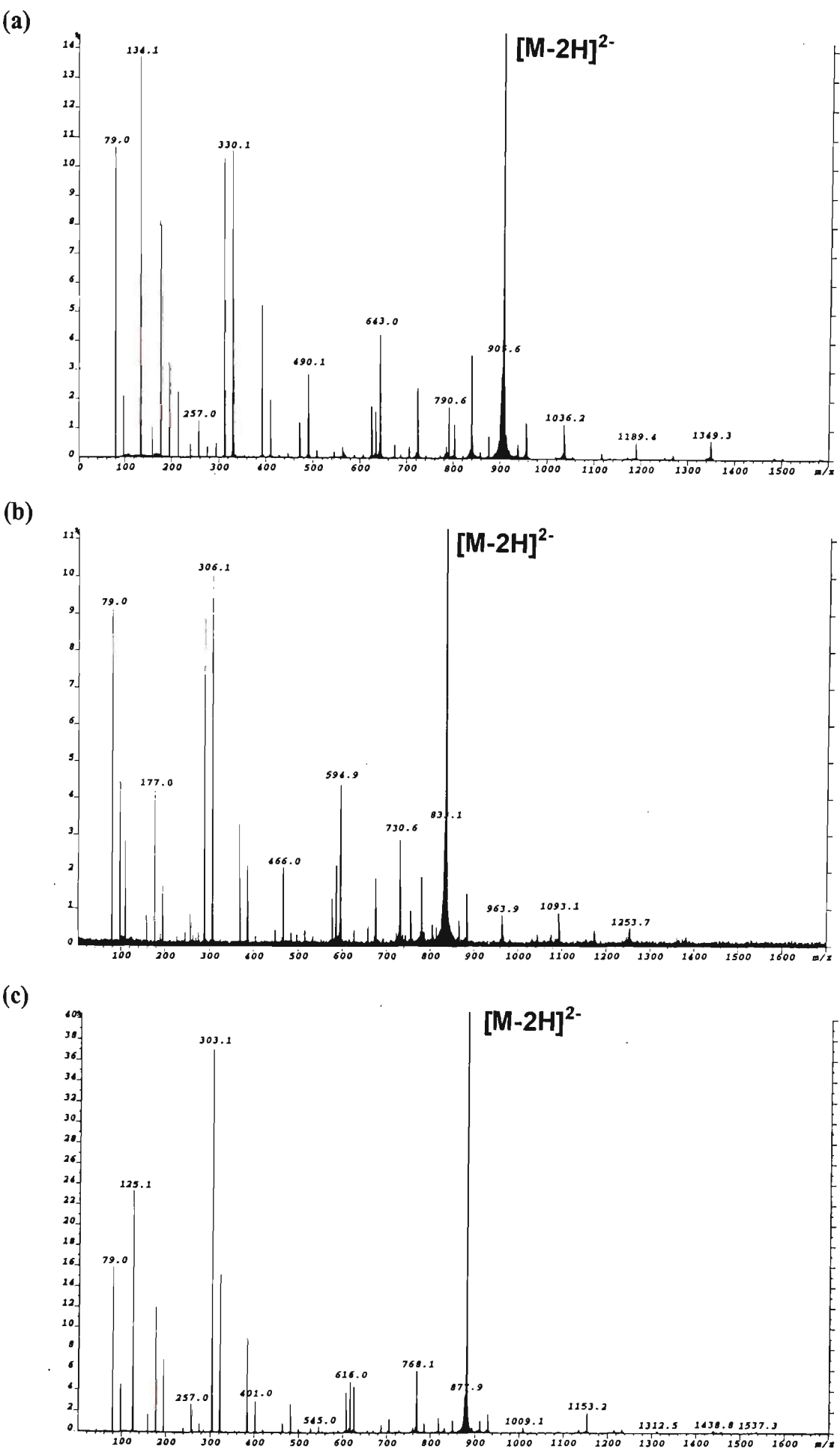


Figure 5.1: ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of (a) 5'-d(AAAAAA)-3'; (b) 5'-d(CCCCCC)-3'; and (c) 5'-d(TTTTTT)-3'.

Table 5.1: Assignments of product ions observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of the 5'-d(AAAAAA)-3', 5'-d(CCCCCC)-3', 5'-d(TTTTTT)-3'.

Product Ions	5'-d(AAAAAA)-3'		5'-d(CCCCCC)-3'		5'-d(TTTTTT)-3'	
	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)
[M-2H]²⁻	907.2	100	835.1	100	880.2	100
[M-BH-2H] ²⁻	839.7	3.60	779.6	1.94	818.1	1.54
w ₅ ⁻	1582.3	0.06	-	-	1537.3	0.28
[w ₅ -H] ²⁻	790.6	1.82	730.6	3.06	768.1	6.03
w ₄ ⁻	1269.3	0.17	1173.3	0.44	1233.3	0.38
[w ₄ -H] ²⁻	634.1	1.69	586.1	2.23	616.0	5.07
w ₃ ⁻	956.2	1.24	884.1	1.45	929.1	1.92
w ₂ ⁻	643.0	4.51	594.9	4.46	625.0	4.72
w ₁ ⁻	330.1	10.5	306.1	10.1	321.1	15.4
[a ₅ -B ₅ H-2H] ⁻	1349.3	0.68	1253.7	0.52	1312.5	0.31
[a ₅ -B ₅ H-3H] ²⁻	674.1	0.47	626.1	0.45	656.1	0.35
[a ₄ -B ₄ H-2H] ⁻	1036.2	1.24	963.9	0.86	1009.1	0.60
[a ₃ -B ₃ H-2H] ⁻	723.1	2.47	675.1	1.86	705.1	1.45
[a ₂ -B ₂ H-2H] ⁻	410.0	2.09	386.1	2.43	401.0	3.25
[a ₁ -B ₁ H-2H] ⁻	97.0	2.13	97.0	4.46	97.0	4.7
[B _{nt} -H] ⁻	312.0	10.3	288.1	9.02	303.1	37.3
B ⁻	134.1	13.7	110.0	2.87	125.1	23.3
PO ₃ ⁻	79.0	10.7	79.0	9.10	79.0	15.9

*calculated monoisotopic mass
B_{nt} denotes a mononucleotide which may be either psB or sBp.

region of the spectrum owing to nucleotide fragments and backbone fragments (the detailed characterisation of these product ions types has been described previously in chapter four), with relatively intense ions observed for the constituent nucleobase and mononucleotide anions, the PO₃⁻ ion (*m/z* 79.0) and the w₁⁻ ion although the relative intensities of these ions is very different in each spectrum. The MS/MS spectrum of 5'-d(AAAAAA)-3' shows A⁻ (*m/z* 134.1) as the most abundant ion, and relatively intense ions are also observed for PO₃⁻ (*m/z* 79.0), w₁⁻ (*m/z* 330.1), and [A_{nt}-H]⁻ (*m/z* 312.1). In the MS/MS spectrum of 5'-d(CCCCCC)-3', the w₁⁻ ion (*m/z* 306.1) produced the most abundant signal followed by PO₃⁻ (*m/z* 79.0), [C_{nt}-H]⁻ (*m/z* 288.1), with the ion for C⁻ (*m/z* 110.0) of comparatively lower relative abundance. It is

interesting that in the MS/MS spectrum of 5'-d(TTTTTT)-3', the mononucleotide ion, $[T_m-H]^-$ (m/z 303.1) is the most abundant product ion, followed by T^- (m/z 125.1), PO_3^- (m/z 79.0), and w_1^- (m/z 321.1). The observed differences in the propensities for base anion formation (ie. $A^- > T^- > C^-$) may be rationalised on the basis of the relative acidities of these three nucleobases, and are in agreement with results obtained by Rodgers *et al.*¹ and McLuckey *et al.*² which show the general order of the acidities to be adenine > guanine > thymine > cytosine.

A comparison of the relative intensities of the various w and (a-B) ions observed in the MS/MS spectrum of each oligonucleotide is given in figure 5.2. It should be noted that the ions for $[a_1-B_1H-2H]^-$ were not included for the purposes of comparison of the (a-B) ions since this ion is also a common fragment generated from backbone cleavage, $[s-H]^-$ (m/z 97.0), and thus is not representative of fragmentation occurring predominantly via loss of the 5' terminal base followed by cleavage of the adjacent ribose 3' C-O bond.

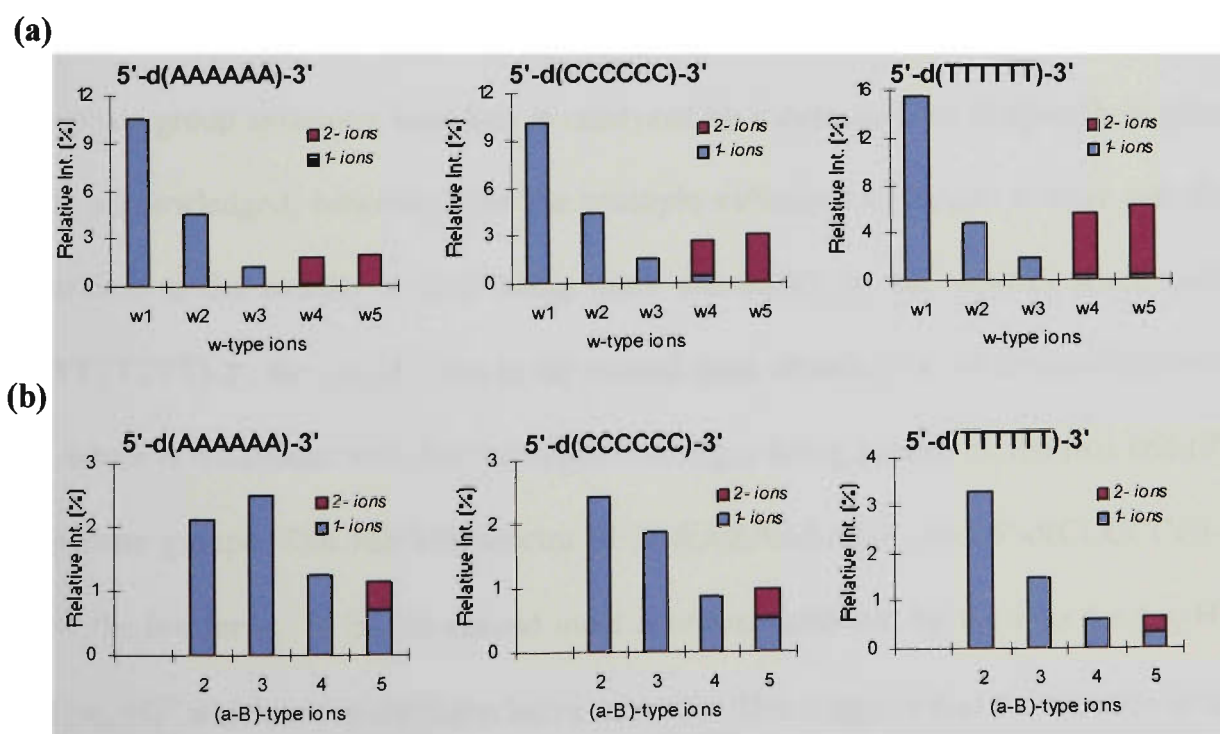


Figure 5.2: Comparison of the relative intensities of (a) w and (b) (a-B) ions observed in the MS/MS spectra of the $[M-2H]^{2-}$ ions of 5'-d(AAAAAA)-3', 5'-d(CCCCCC)-3', and 5'-d(TTTTTT)-3'.

The results for w ion formation (figure 5.2(a)) show similar trends in the MS/MS spectra of all three hexamers. In each case, the w_1 ion is the most abundant, with the w_3 being the weakest of all w ions. The trends observed in each case appear consistent with favourable location of the two negative charges on phosphate groups furthest apart on the backbone which would be expected to yield the most stable structure for a doubly charged precursor ion in terms of Coulombic repulsion forces. The formation of the w_3^- ion from base loss catalysed by a deprotonated 3' phosphate group would require one of the charges to be located on the middle phosphate group which means that the maximum spacing of the two charges would then be only one phosphate group apart (ie. the two charges would be located on the first and third phosphates, or the third and fifth phosphates). Compared with the production of all other w ions, formation of w_3^- involves a precursor ion with the two charges placed in closest proximity to each other and therefore can be expected to be unfavoured on the basis of Coulombic repulsion interactions. The high abundance of the w_1^- ion in the MS/MS spectrum of each hexamer suggests a preference for the charges to be located on the first phosphate group assuming base loss is catalysed by a deprotonated 3' phosphate group (it is acknowledged, however, that the multiple collisions of larger w ions can also contribute to the smaller w ions being more abundant). In the MS/MS spectrum of 5'-d(TTTTTT)-3', the $[w_5-H]^{2-}$ ion is the second most abundant w ion formed following w_1^- , which is consistent with the two negative charges being located on the first and fifth phosphate groups. The MS/MS spectra of 5'-d(AAAAAA)-3', and 5'-d(CCCCCC)-3' show the ion for w_2^- to be the second most abundant followed by the ions for $[w_5-H]^{2-}$ and $[w_4-H]^{2-}$ which are of similar relative intensity. This suggests that the location of the

two negative charges either on the first and fourth phosphate groups, or the second and fifth phosphate groups may also yield stable structures of these precursor ions.

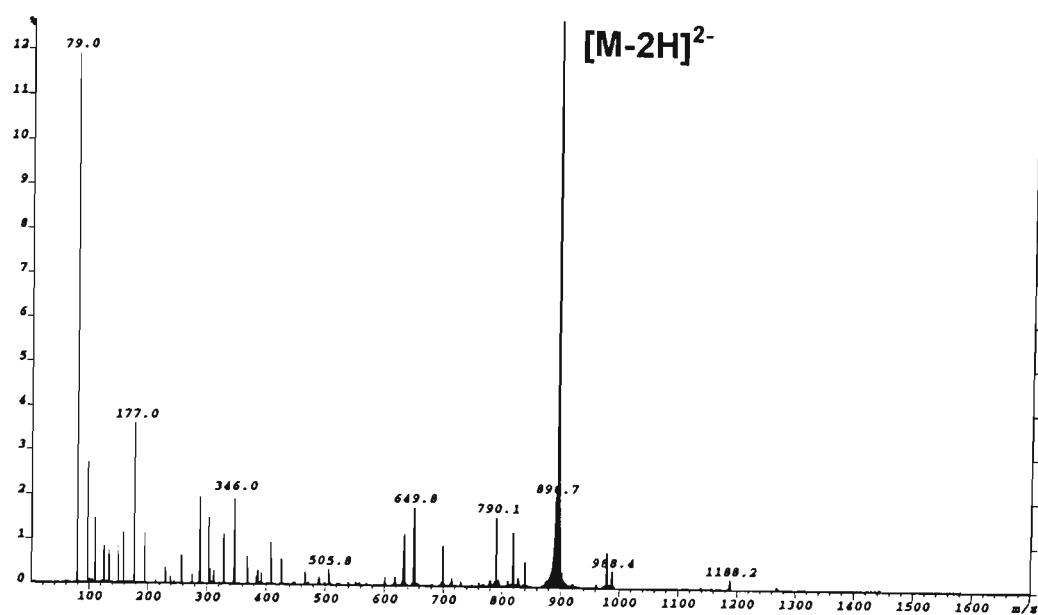
The results for (a-B) ion formation (figure 5.2(b)) in the MS/MS spectrum of each hexamer show distinctly different trends compared with those observed for formation of complementary w ions. The relative intensities of (a-B) ions in the MS/MS spectra of the $[M-2H]^{2-}$ ions of 5'-d(CCCCCC)-3', and 5'-d(TTTTTT)-3' are very similar with the ion for $[a_2-B_2H-2H]^-$ being the most intense followed by $[a_3-B_3H-2H]^-$, the combined intensities of $[a_5-B_5H-2H]^-$ and $[a_5-B_5H-3H]^{2-}$, and $[a_4-B_4H-2H]^-$. In contrast, the MS/MS spectrum of the $[M-2H]^{2-}$ ion of 5'-d(AAAAAA)-3' shows predominant formation of $[a_3-B_3H-2H]^-$, followed by $[a_2-B_2H-2H]^-$, $[a_4-B_4H-2H]^-$ and the combined intensities of $[a_5-B_5H-2H]^-$ and $[a_5-B_5H-3H]^{2-}$. The fact that doubly charged $[a_4-B_4H-3H]^{2-}$ ions are not observed in any case is consistent with the charges preferentially being located some distance apart, since observation of this ion would require the location of both negative charges amongst the first three phosphate groups. The favourable formation of $[a_2-B_2H-2H]^-$ ions for all hexamers, and $[a_3-B_3H-2H]^-$ ions in the case of 5'-d(AAAAAA)-3' and 5'-d(CCCCCC)-3' do not correspond with the propensities for formation of their complementary w ions, w_4^- and w_3^- , which are both observed in very weak relative abundance. Furthermore the comparatively lower relative intensity of the ions for $[a_5-B_5H-2H]^-$ and $[a_4-B_4H-2H]^-$ is also inconsistent with the favourable formation of complementary w_1^- and w_2^- ions. One possible explanation for the patterns observed in the intensities of the (a-B) ions in these MS/MS spectra is that they are a consequence of multiple collisions and the decomposition of the larger (a-B) ions from further nucleobase loss catalysed by the deprotonated phosphate on the ion

(discussed in more detail below in section 5.3). Alternatively, this data may indicate that the (a-B) and w ions are formed via different mechanisms.

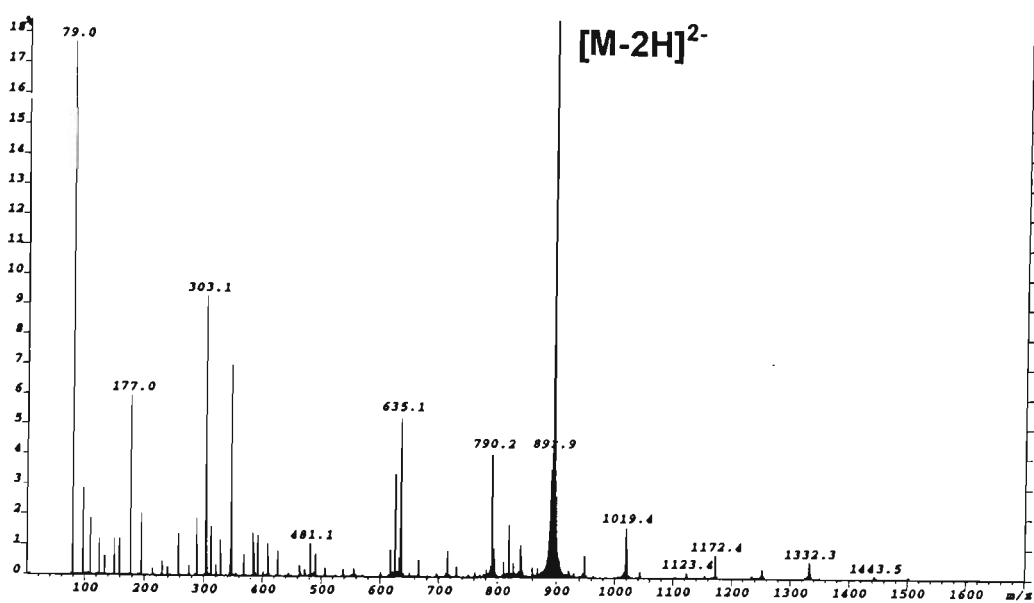
5.2.2 Fragmentation of Oligonucleotides of Heterogeneous Base Composition

Figures 5.3(a)-(k) show the MS/MS spectra of $[M-2H]^{2-}$ ions of the heterogeneous oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TACGTA)-3', 5'-d(TCGTGA)-3', 5'-d(TCACGA)-3', 5'-d(GCCATG)-3', 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3' respectively. The set of oligonucleotides chosen for this study includes all possibilities for the location of the four different nucleobases at each of the six base locations within the hexamer sequence so that the effect of base location versus base identity could be explored to some extent. The MS/MS spectra of these heterogeneous oligonucleotides were studied with the aim of first investigating the relative propensities of base anion formation and neutral nucleobase loss of the different nucleobases, and to assess and compare the influence of both base identity and base location on the relative abundance of these ions. Second, it was of interest to examine the extent to which the propensities for formation of the various w and (a-B) sequence ions are influenced by base identity, and to compare the results obtained with those of the unibase oligonucleotides in which base location is the predominant factor influencing the formation of these sequence ions. The assignments and relative abundances of the product ions arising from nucleobase loss, and w and (a-B) sequence ions are given in table 5.2.

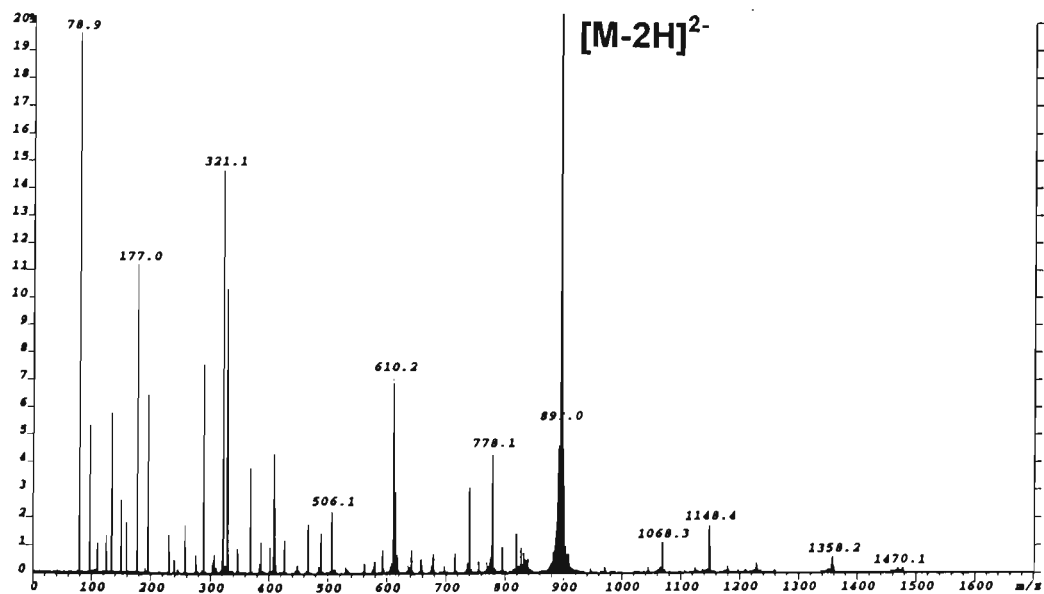
(a)

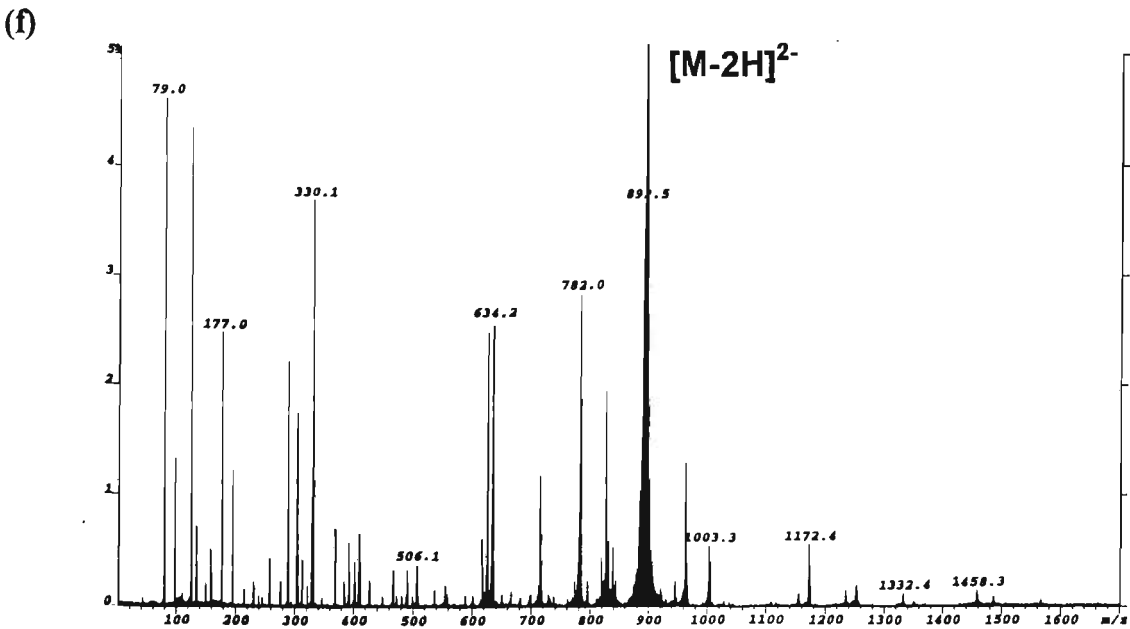
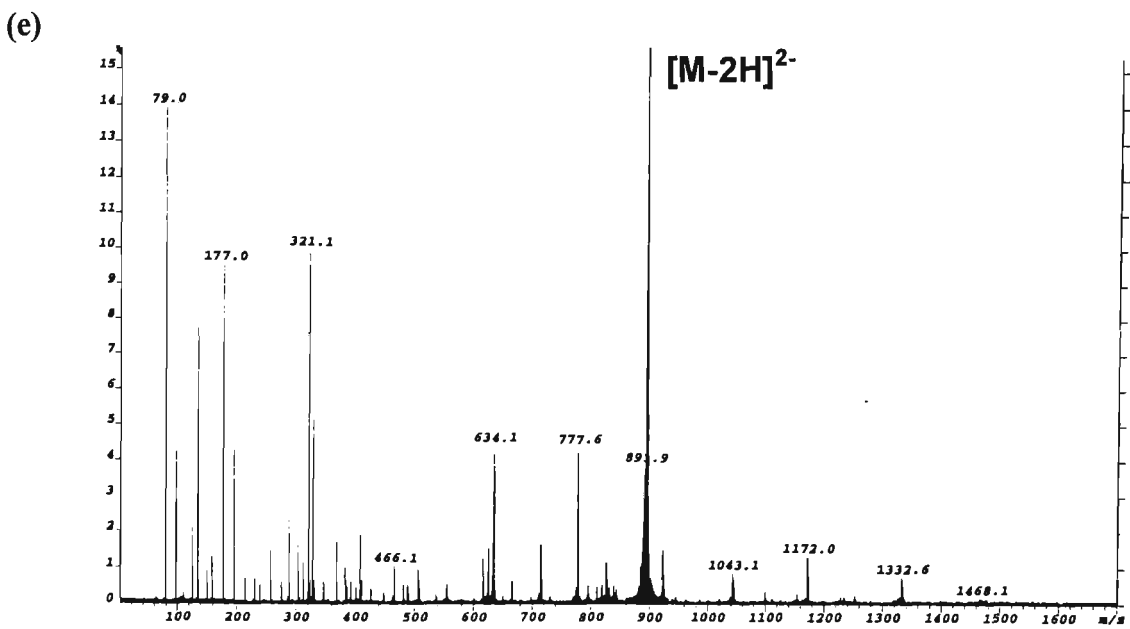
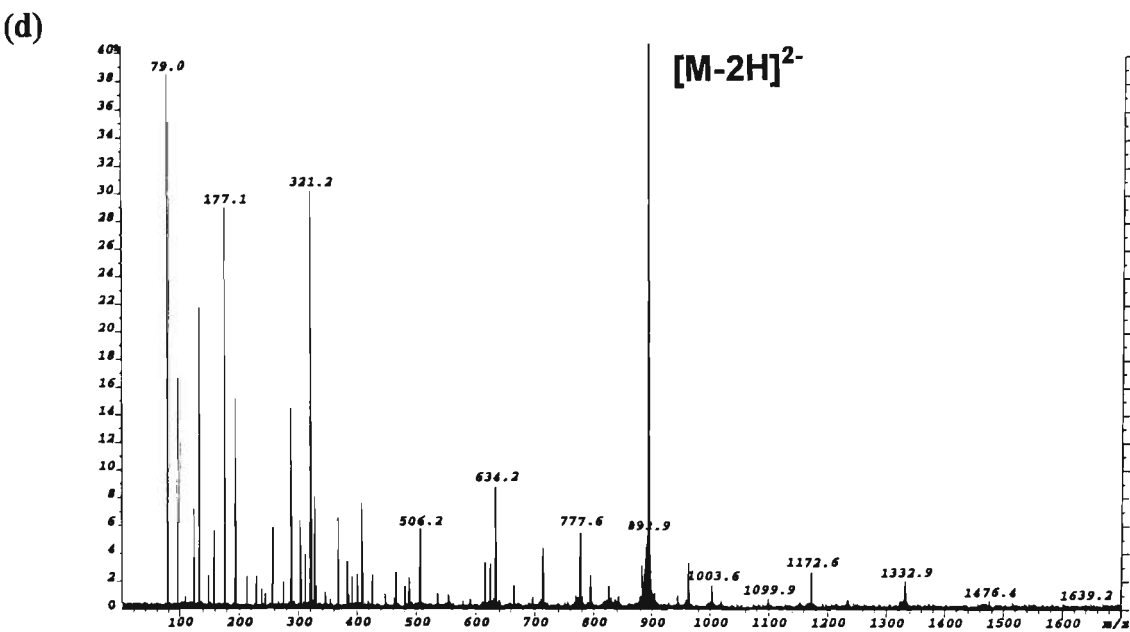


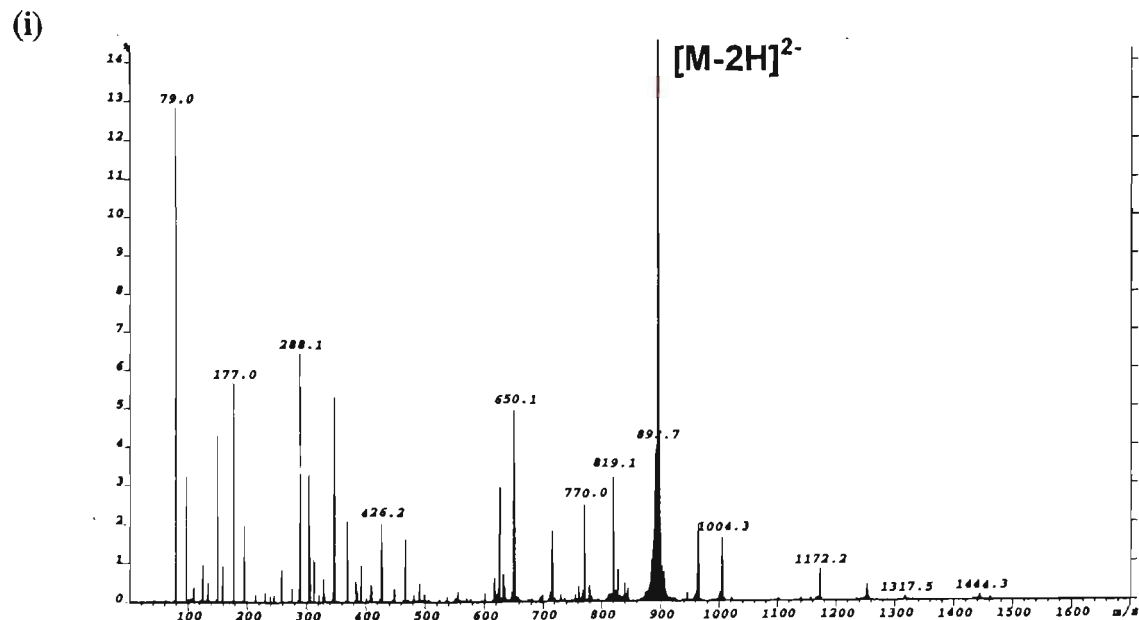
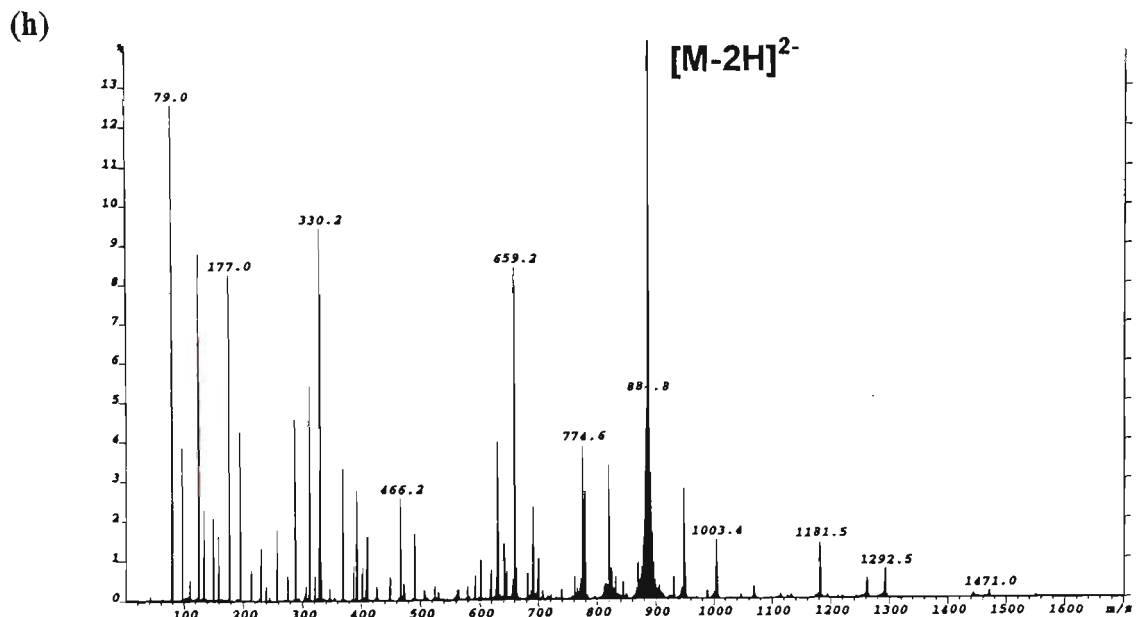
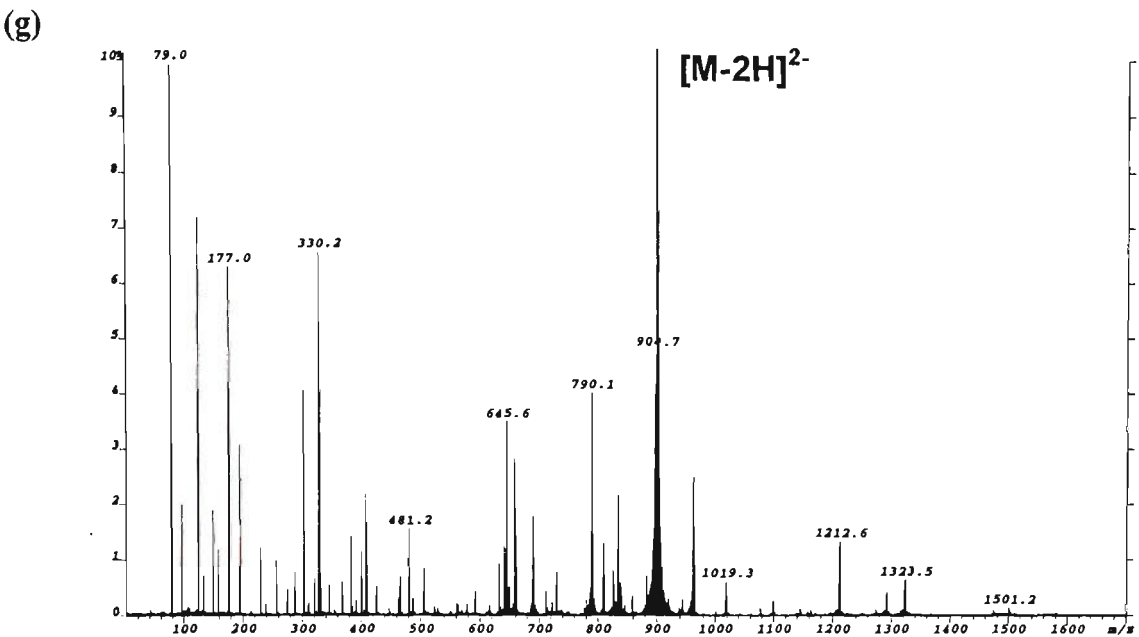
(b)



(c)







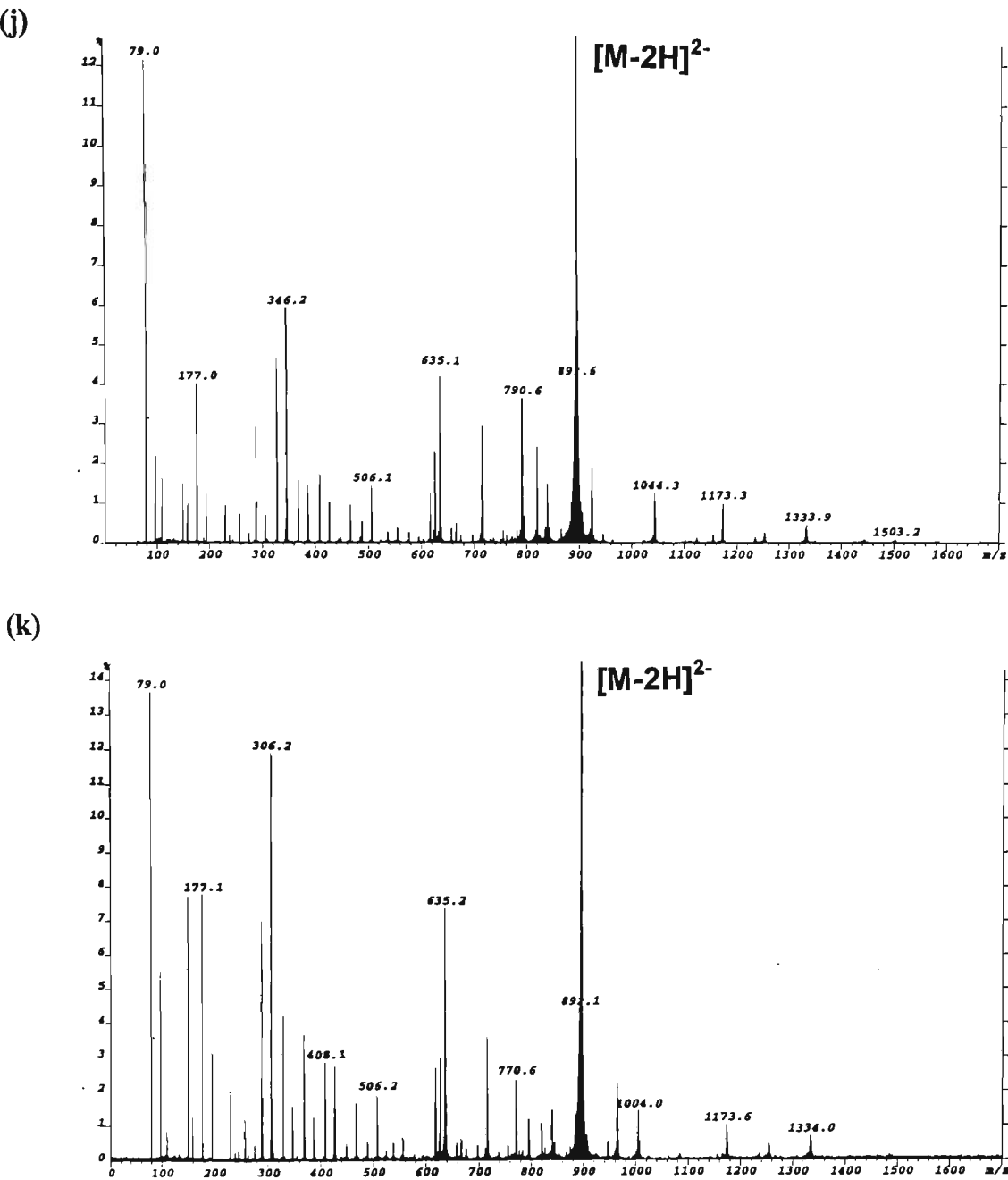


Figure 5.3: ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of (a) 5'-d(CACGTG)-3'; (b) 5'-d(CGTACG)-3'; (c) 5'-d(AGGCCT)-3'; (d) 5'-d(ATCGAT)-3'; (e) 5'-d(ATGCAT)-3'; (f) 5'-d(TACGTA)-3'; (g) 5'-d(TCGTGA)-3'; (h) 5'-d(TCACGA)-3'; (i) 5'-d(GCCATG)-3'; (j) 5'-d(CGGCCG)-3'; and (k) 5'-d(GCCGGC)-3'

Table 5.2 cont/...

Product Ions	5'-d(TCGTGA)-3'		5'-d(TCACGA)-3'		5'-d(GCCATG)-3'		5'-d(CGGCCG)-3'		5'-d(GCCGGC)-3'	
	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)	m/z	Rel. Int.(%)
[M-2H] ²⁻	902.2	100	886.7	100	894.7	100	895.2	100	895.2	100
[M-CH-2H] ²⁻	846.6	0.31	831.2	0.61	839.1	0.53	839.6	1.49	839.7	1.48
[M-TH-2H] ²⁻	839.2	0.63	823.5	0.82	-	-	-	-	-	-
[M-AH-2H] ²⁻	834.6	2.21	819.1	3.39	827.1	0.86	-	-	-	-
[M-GH-2H] ²⁻	826.6	0.86	811.1	0.39	819.1	3.41	819.7	2.45	819.7	1.13
w ₅ ⁻	-	-	1550.3	0.08	-	-	-	-	-	-
[w ₅ -H] ²⁻	790.0	4.03	774.7	3.78	770.0	2.60	790.6	3.67	770.6	2.40
w ₄ ⁻	1292.4	0.43	1261.4	0.50	1252.3	0.44	1253.4	0.27	1253.3	0.46
[w ₄ -H] ²⁻	645.6	3.51	630.1	4.05	625.6	3.03	626.0	2.33	626.0	3.23
w ₃ ⁻	963.3	2.51	948.3	2.76	963.2	2.02	924.5	1.90	964.4	2.22
w ₂ ⁻	659.2	2.89	659.2	8.24	650.1	5.03	635.1	4.41	635.2	7.39
w ₁ ⁻	330.2	6.55	330.2	9.44	346.1	5.32	346.2	5.92	306.2	11.90
[a ₅ -B ₃ H-2H] ⁻	1323.5	0.65	1292.1	0.71	1317.5	0.14	1333.9	0.48	1333.0	0.72
[a ₃ -B ₃ H-3H] ²⁻	661.6	1.38	645.6	0.71	658.1	0.16	666.0	0.54	666.1	0.62
[a ₄ -B ₄ H-2H] ⁻	1019.3	0.62	1003.4	1.45	1004.3	1.73	1044.3	1.29	1004.0	1.43
[a ₃ -B ₃ H-2H] ⁻	690.1	1.82	690.1	2.28	715.1	1.88	715.1	2.99	715.2	3.60
[a ₂ -B ₂ H-2H] ⁻	401.1	1.25	401.2	0.74	426.2	2.16	386.2	1.48	426.2	2.87
G ⁻	150.1	1.91	150.1	2.18	150.0	4.37	150.1	1.50	150.1	7.75
A ⁻	134.0	0.79	134.1	2.39	134.0	0.56	-	-	-	-
T ⁻	125.0	7.21	125.1	9.18	125.0	1.01	-	-	-	-
C ⁻	110.0	0.17	110.1	0.55	110.0	0.44	110.1	1.63	110.0	0.90
PO ₃ ⁻	79.0	9.95	79.0	13.38	79.0	12.8	79.0	12.15	79.0	13.69

The propensities for base anion formation in the MS/MS spectrum of each oligonucleotide is summarised in figure 5.4. In each MS/MS spectrum the ion for the deprotonated 5' terminal base yields the most intense signal of all base anions regardless of the identity of the 5' terminal base. These results are indicative of base location being the more influential factor in the propensities for base anion formation compared with base identity. The observed preference for charged base loss at the 5'-terminal position would appear consistent with the charges being favourably located on the two extreme phosphate groups (the first and fifth phosphate groups) whereby Coulombic repulsion between the two charges is minimised. Further, the preference for location of a charge at the first phosphate group (and to the 3' position of the first base) which can catalyse base loss at the 5' terminus would result in this process being more favourable than loss of the 3' base for which there is no 3' phosphate group.

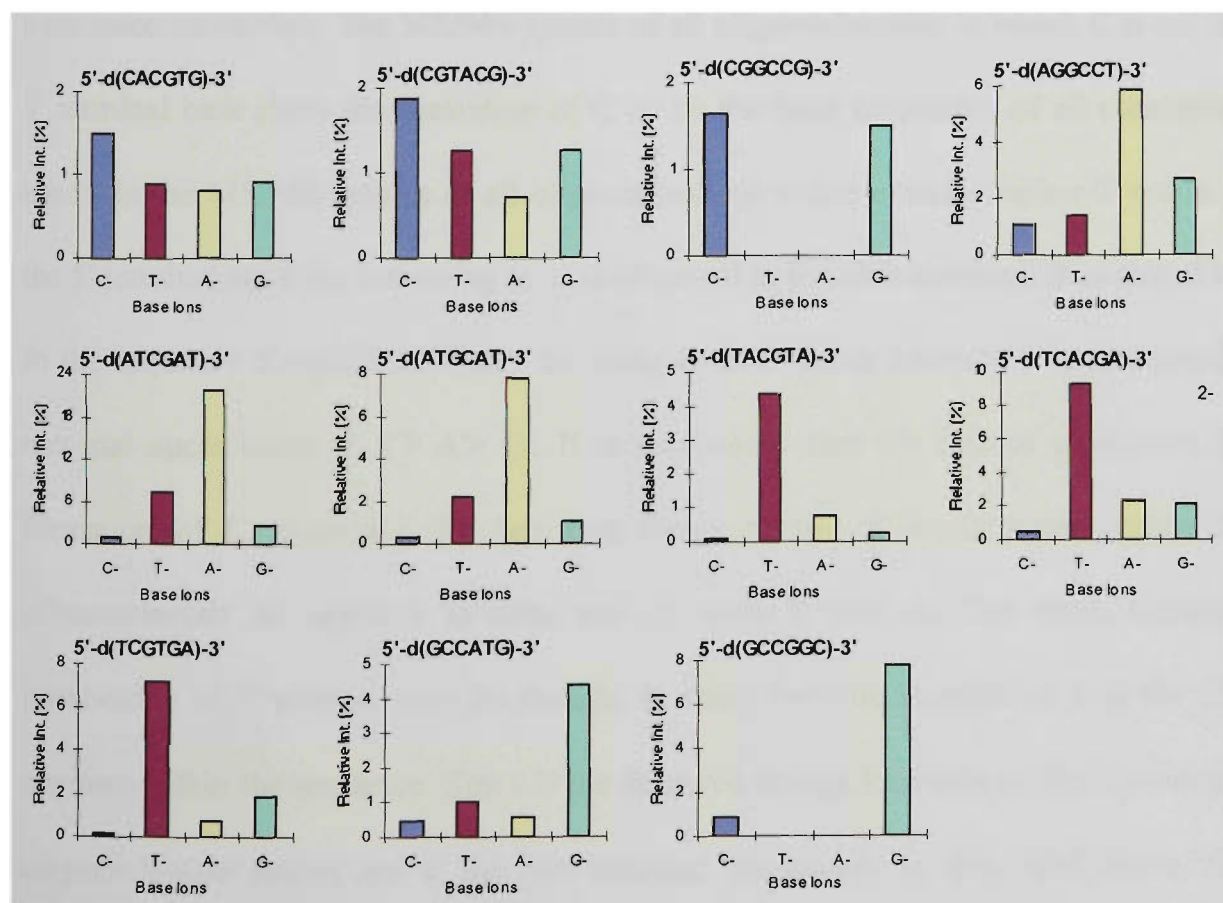


Figure 5.4: Comparison of the relative intensities of base anions observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(CGGCCG)-3', 5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TACGTA)-3', 5'-d(TCACGA)-3', 5'-d(TCGTGA)-3', 5'-d(GCCATG)-3', and 5'-d(GCCGGC)-3'.

Apart from the clear preference displayed for charged base loss at the 5' terminal position, base location does not appear to be an important factor influencing the propensities for base anion formation of the remaining bases. Although, the results obtained for the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TACGTA)-3', 5'-d(TCACGA)-3' may be thought to suggest a secondary preference for charged base loss of the 3' terminal base, the order for charged base loss obtained for the oligonucleotides 5'-d(AGGCCT)-3', and 5'-d(TCGTGA)-3' (which in each case show the ion for G⁻ to be the second most abundant base anion), are clearly not consistent with this. Rather, the results are more consistent with base identity being a predominant influential factor in the propensity for base anion formation of nucleobases not at the 5' terminal position. Furthermore, the number of a given nucleobase present in the oligonucleotide may also affect the order of base anion intensities. The MS/MS spectra of all oligonucleotides in which C is not the 5' terminal base show the formation of C⁻ to be the least favourable of all constituent bases. In the MS/MS spectra of all oligonucleotides which contain neither T nor A as the 5' terminal base, the ion owing to T⁻ is observed to be more abundant than that of A⁻. In the sequence 5'-d(GCCATG)-3', the order of base anion intensities for the non 5'-terminal nucleobases is T⁻ > A⁻ > C⁻. It is noteworthy that the lack of preference for formation of C⁻ overrides the fact that there are two C nucleobases within the oligonucleotide as opposed to only one of each T and A. The more abundant observation of T⁻ over A⁻ may be thought to arise from the location of T at the fifth position within the sequence. Since if the favoured charge locations in doubly charged oligonucleotide anions are at the two terminal phosphates as described above, this would result in the fifth base having a 3' deprotonated phosphate group which can

catalyse base loss at this position. The trend of $T^- > A^-$ is also observed for the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. In the case of 5'-d(CACGTG)-3', T is also located at the fifth residue so the preference for its loss over A may also be a result of base location. This is however not the case with 5'-d(CGTACG)-3' in which T and A are located at the third and fourth residues respectively. If the preference for charge location is that associated with minimal internal repulsion forces, then the location of a charged phosphate group on the third residue would be thought to be the least likely scenario since it would essentially result in the two charges being placed within one residue of each other. It would thus be thought that the location of a charged phosphate at the fourth residue that can catalyse loss of A, would be more favourable than at the third residue. The observation of T^- being more intense than A^- would, therefore, appear to arise from the effects of base identity rather than base location. The propensity for formation of G^- relative to that of T^- and A^- yields interesting results which appear to be influenced by base sequence. In the MS/MS spectra of oligonucleotides which comprise neither G nor T at the 5' terminal position, an order of $G^- > T^-$ is observed in the MS/MS spectra of 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(AGGCCT)-3', whereas the MS/MS spectra of 5'-d(ATGCAT)-3' and 5'-d(ATCGAT)-3' show the opposite order of $T^- > G^-$. These results for the relative abundances of T^- and G^- ions may be a consequence of the number of G and T bases present in each sequence since the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(AGGCCT)-3' comprise two G bases and one T base whereas in 5'-d(ATGCAT)-3' and 5'-d(ATCGAT)-3' there are two T bases compared to one G base. Similar results are also obtained for the relative abundances of G^- and A^- base anions in oligonucleotides where these bases are not at the

5' terminal position. In the MS/MS of oligonucleotides which comprise two G bases and one A base, ie. 5'-d(TCGTGA)-3', 5'-d(CACGTG)-3', and 5'-d(CGTACG)-3', a preference of $G^- > A^-$ is observed. On the other hand, the MS/MS spectrum of the oligonucleotide, 5'-d(TACGTA)-3' which comprises two A bases and one G base shows A^- to be more abundant than G^- .

Figure 5.5 summarises the relative intensities of the ions owing to neutral base loss observed in the MS/MS spectrum of the $[M-2H]^{2-}$ ion of each heterogenous hexamer. The MS/MS spectra of all oligonucleotides except those comprising thymine at the 3' terminal position (5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3') show a preference for loss of the 3' terminal base as a neutral. The high abundance of the ion for neutral loss of the 3' terminal base may be rationalised by the relative

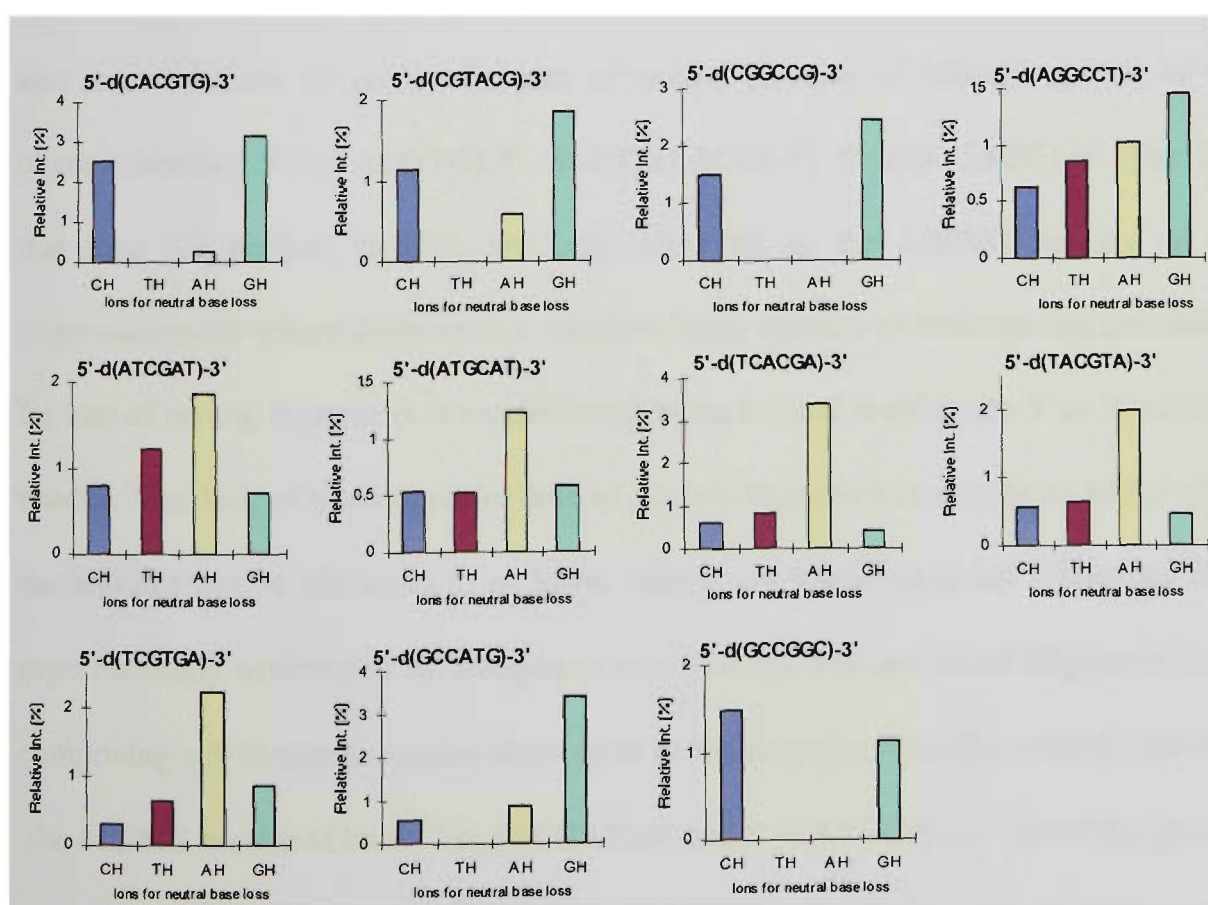


Figure 5.5: Comparison of the relative intensities of ions owing to neutral base loss observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(CGGCCG)-3', 5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(TACGTA)-3', 5'-d(TCGTGA)-3', 5'-d(GCCATG)-3', and 5'-d(GCCGGC)-3'.

stability of this ion to further fragmentation from multiple collisions, since fragmentation of this ion from subsequent cleavage of the adjacent ribose 3' C-O bond resulting in water loss is not observed to be a favourable fragmentation pathway in the results obtained in these studies nor from previous studies^{2,10}. Therefore the enhanced stability of the ion formed from loss of the 3' terminal base as a neutral would result in this ion being more abundant than those formed from the loss of the other bases which can further dissociate through phosphodiester bond cleavage of the ribose 3' C-O bond to generate w and (a-B) ions. The preference for neutral base loss of the 3' terminal base appears to be followed by neutral base loss of the 5'-terminal base, although the results from the MS/MS spectrum of 5'-d(TCGTGA)-3' do not follow this preference, and instead show the ion for loss of GH to produce the second most abundant of the ions arising from neutral base loss. A distinct observation from the results for neutral base loss is the absence of an ion for loss of neutral thymine in MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(GCCATG)-3'. The fact that ions for neutral thymine are only observed in the MS/MS spectra of the oligonucleotides where thymine is a terminal base, appears to indicate that the ability for loss of neutral thymine is dependent on it being located at either the 5' or 3' terminal residue. The lack of preference for loss of neutral thymine in comparison to the other nucleobases can be attributed to its lower stability as a leaving group which has been experimentally determined by Rodgers *et al.*¹. The MS/MS spectra of oligonucleotides comprising a 3' terminal thymine show quite different propensities for neutral base loss. The MS/MS spectra of both 5'-d(ATCGAT)-3' and 5'-d(ATGCAT)-3' show the greatest preference of neutral base loss for adenine, which may suggest a preference for neutral loss of the 5' terminal base where the 3' terminal base is thymine. A major difference is

observed, however, in the propensity for loss of TH in the MS/MS spectra of 5'-d(ATCGAT)-3' and 5'-d(ATGCAT)-3'. Whereas loss of TH is the second most abundant ion of those from neutral base loss in the case of 5'-d(ATCGAT)-3', it is the least favourable neutral base lost in the MS/MS spectrum of 5'-d(ATGCAT)-3'. These differences in neutral base loss propensity are surprising given the high similarity in base composition and sequence of both oligonucleotides. The MS/MS spectrum of 5'-d(AGGCCT)-3' shows the strongest propensity for loss of GH which is in contrast to the results obtained for 5'-d(ATGCAT)-3' and 5'-d(ATCGAT)-3' which yield preferential loss of AH. One possible explanation for these differences in propensity may arise from the fact the oligonucleotide 5'-d(AGGCCT)-3' comprises a greater number of G bases compared with A bases. This explanation would also account for the observed preferential loss of GH rather than AH in the MS/MS spectrum of 5'-d(TCGTGA)-3'.

The propensity for neutral base loss of bases located internally within the oligonucleotide are quite different depending on base identity and oligonucleotide sequence. As discussed above, in all cases where T appears as an internal base, no ions are present for loss of TH. In the case of oligonucleotides comprising C and G internally within their sequence, the MS/MS spectra of 5'-d(TCACGA)-3' and 5'-d(TACGTA)-3' and 5'-d(ATCGAT)-3' show the same preference for loss of CH > GH, whereas the reverse order of GH > CH is observed in the MS/MS spectra of 5'-d(AGGCCT)-3', 5'-d(TGCTGA)-3' and 5'-d(ATGCAT)-3'. Although the different relative intensities of ions for loss of CH and GH obtained for 5'-d(TCACGA)-3' and 5'-d(TCGTGA)-3' can be rationalised as arising from the differing proportions of C and G nucleobases in these oligonucleotides, differences in the propensities for loss of CH and GH are also

observed where the oligonucleotides contain the same number of C and G bases (ie. 5'-d(TACGTA)-3', 5'-d(ATCGAT)-3' and 5'-d(AGGCCT)-3', d(ATGCAT)-3'). The results would appear to suggest that base location is an important factor influencing the relative propensities for loss of CH and GH in these cases, with more preferable neutral base loss when the base is located at the third position compared with when the base is located at the fourth position. These effects may arise from the lack of preference for charge location at the third residue of a doubly charged anion since this would mean that the maximum spacing between the two charges could only be nucleotide unit apart. In this case the third residue would lack a 3' deprotonated phosphate group to catalyse base loss at this position, similar to the situation with the 3' terminal base. It may therefore be that base loss at these positions occurs predominantly via 1,2 base elimination giving rise to neutral base loss analogous to the mechanism for neutral base loss proposed by McLuckey *et al.*² (see chapter one, scheme 1.3(b)). Where there is a deprotonated phosphate group to the 3' side of a base, base loss can be thought to occur via the mechanism proposed by Rodgers *et al.*¹, and results in a competition for charged versus neutral base loss depending on whether the charge is transferred to the leaving base, or is retained with the phosphate group upon dissociation of the proton-bound intermediate (see chapter one, scheme 1.2). These observed effects of base location for neutral loss of internal bases are however outweighed by the effects of base identity in the case of loss of AH versus CH as observed in the MS/MS spectrum of 5'-d(GCCATG)-3'. Although this oligonucleotide comprises C at the third residue in addition to having two C bases compared and only one A base, the results show a clear preference for neutral base loss of AH > CH.

From the results obtained in these analyses it is difficult to rationalise the contributions of the factors of base identity, base location and sequence effects which give rise to the observed propensities in neutral base loss. Moreover, the observed propensities may also be a consequence of differences in the relative stability of the various ions formed to further decomposition from multiple collisions. Therefore, further studies are required in order to probe more fully the interplay of these effects.

The relative intensities of w ions observed in the MS/MS spectrum of each oligonucleotide are illustrated in figure 5.6. The trends for w ion formation for oligonucleotides of heterogeneous base composition are similar to those observed in the MS/MS spectra, of the unibase oligonucleotides. In all MS/MS spectra the w_1 ion is the most abundant w ion formed, and w_3 ions are the least abundant. These results indicate that charge location affects the formation of w ions, and a preference exists for location of the two negative charges as far apart as possible on the backbone to minimise charge repulsion as discussed above for the MS/MS results of the unibase oligonucleotides. Formation of either the w_2^- and $[w_5-H]^{2-}$ ions constitute the second most abundant cleavages in all cases. From comparison of the relative intensities of the w ions formed in the MS/MS spectra of all oligonucleotides, it can be seen that base identity also plays a role in influencing the formation of w ions when thymine is located to the 5' side of the cleavage site. The MS/MS spectrum of 5'-d(CGTACG)-3' shows the relative abundance of w_3^- is significantly lower than other w ions. In the MS/MS spectra of 5'-d(ATCGAT)-3', and 5'-d(ATGCAT)-3', the relative intensity of the w_4 ions are much lower than those of w_4 ions in the MS/MS spectra of the other oligonucleotides. These observations would appear to arise from the lack of preference for loss of neutral thymine as a consequence of its poor leaving group ability in comparison to the other

nucleobases (as discussed previously in chapter four, section 4.6). If the predominant mechanism for formation of w and (a-B) sequence ions involves initial loss of the base to the 5' side of the cleavage site, the relative propensities for base loss would thus be expected to be an important factor influencing sequence ion formation. The effect of the propensity for base loss however is not found to overcome the distinct preference for formation of the w_1^- ion as is evident from the relative intensities of w ions in the MS/MS spectra of 5'-d(CACGTG)-3', 5'-d(TACGTA)-3' and 5'-d(GCCATG)-3'. Although w_1^- ion formation in all of these oligonucleotides involves cleavage of the ribose 3' C-O bond to the 3' side of a T-containing residue, the w_1^- ion is still the most abundant of all w ions.

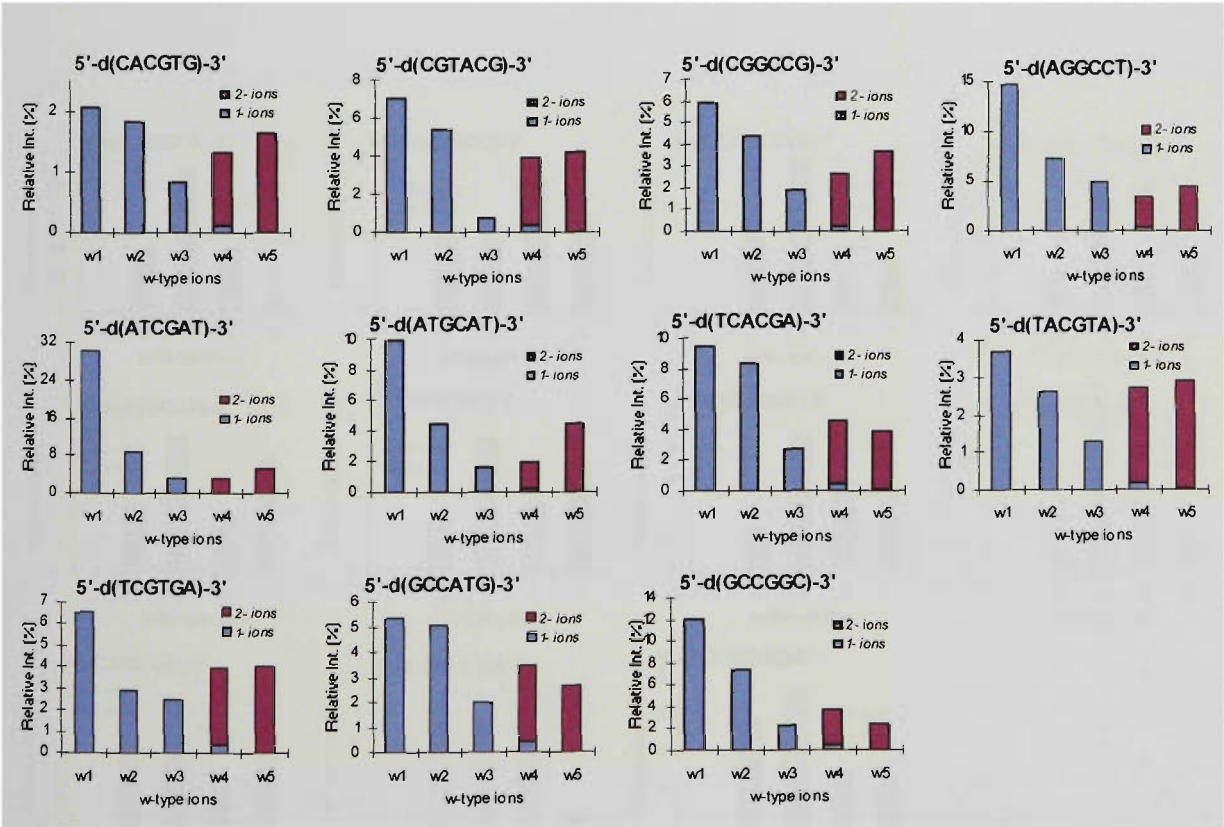


Figure 5.6: Comparison of the relative intensities of w ions observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(CGGCCG)-3', 5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(TACGTA)-3', 5'-d(TCGTGA)-3', 5'-d(GCCATG)-3', and 5'-d(GCCGGC)-3'.

The trends for the relative abundances of (a-B) ions generated in the MS/MS spectra of all heterogeneous oligonucleotides are summarised in figure 5.7. An overall strong preference is observed for formation of the $[a_3-B_3H-2H]^-$ ion which, in most cases, is the most intense (a-B) ion observed. Similar results for the $[a_3-B_3H-2H]^-$ ion were also observed in the MS/MS spectra of the unibase oligonucleotides 5'-d(AAAAAA)-3' and 5'-d(CCCCCC)-3'. This preference for base loss and cleavage of the ribose C-O bond to the 3' side of the third residue is in contrast to the trends for complementary w ion formation which show a distinct lack of preference for formation of w_3^- ions. As discussed previously in the case of the unibase oligonucleotides, this may be a consequence of multiple collisions of larger (a-B) ions which decompose into smaller (a-B) ions under the collision conditions used for these experiments.

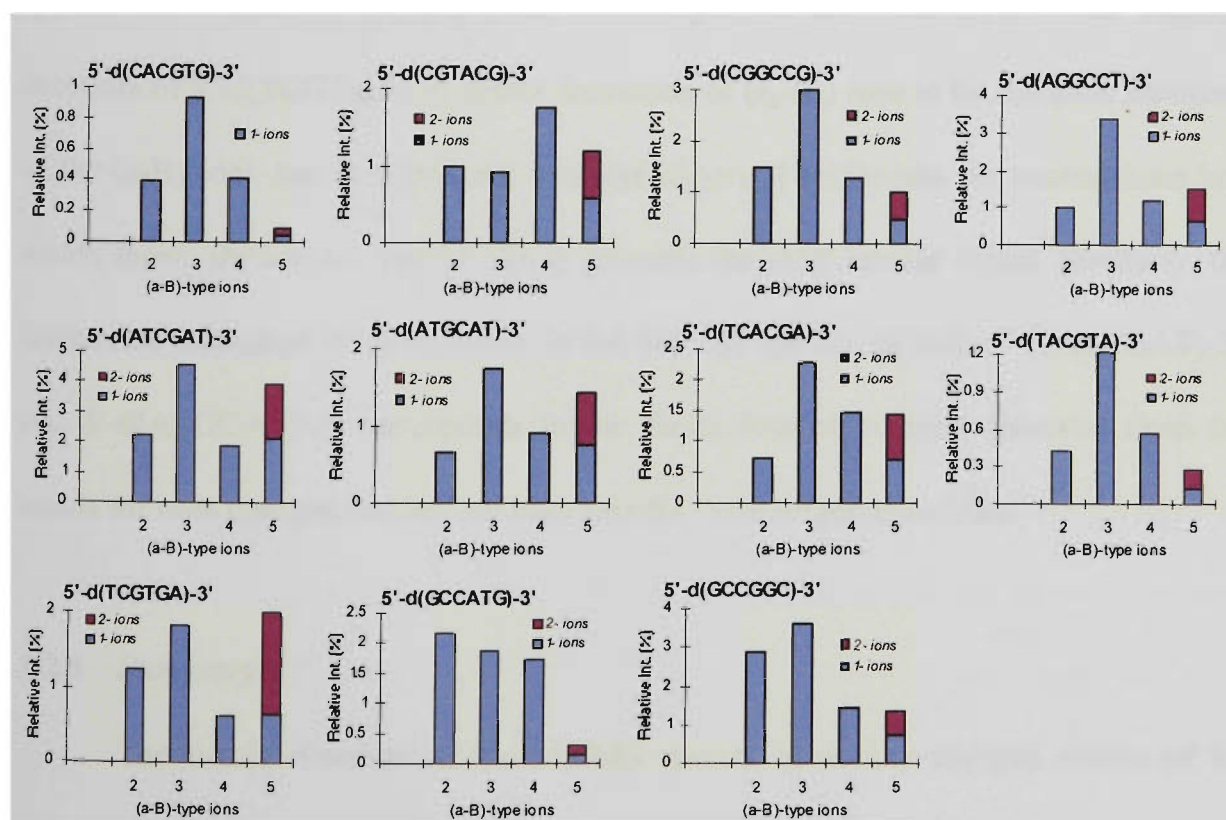


Figure 5.7: Comparison of the relative intensities of (a-B) ions observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(CGGCCG)-3', 5'-d(AGGCCT)-3', 5'-d(ATCGAT)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(TACGTA)-3', 5'-d(TCGTGA)-3', 5'-d(GCCATG)-3', and 5'-d(GCCGGC)-3'.

Alternatively, this may suggest a different mechanism (more below). An exception to the observed preference for $[a_3-B_3H-2H]^-$ ion formation occurs in the MS/MS spectrum of 5'-d(CGTACG)-3' which showed formation of the $[a_3-B_3H-2H]^-$ ions to be the least abundant of all (a-B) ions, with the most intense being $[a_4-B_4H-2H]^-$ ions. This can be rationalised by the lower propensity for loss of neutral thymine observed in the MS/MS spectrum of 5'-d(CGTACG)-3', and is further consistent with the strong lack of preference for w ion formation involving cleavage to the 3' side of thymine. The effects of the lack of preference for neutral thymine loss on (a-B) ion formation can also be seen in the results obtained in the MS/MS spectra of 5'-d(CACGTG)-3', 5'-d(TACGTA)-3' and 5'-d(GCCATG)-3' which all show formation of (a_5-B_5) ions to be highly unfavoured. In contrast, (a_5-B_5) ions are relatively abundant in the MS/MS spectra of 5'-d(TCGTGA)-3', 5'-d(ATGCAT)-3', 5'-d(ATCGAT)-3'. The MS/MS spectrum of 5'-d(TCGTGA)-3' shows formation of (a_5-B_5) ions to be the most abundant of the (a-B) ions and is consistent with the observed preference for neutral base loss which shows the ion for loss of GH to generate the most intense signal. Similarly, the favourable formation of (a_5-B_5) ions in the MS/MS spectra of both 5'-d(ATCGAT)-3', and 5'-d(ATGCAT)-3' corresponds to the facile loss of adenine observed from the trends for both charged and neutral base loss for these oligonucleotides.

5.2.3 Summary

The trends observed in the MS/MS spectra of doubly charged anions of the oligonucleotide hexamers with either unibase or heterogeneous sequences show the base location and the site(s) of deprotonation of the precursors to be important factors in directing the major fragmentation pathways, and therefore in determining the propensity

for formation of different sequence ions. This was particularly evident in the case of w ions. Cleavage of the ribose 3' C-O bond at the fifth residue resulting in w_1^- ions was clearly favoured whereas the w_3^- ions arising from cleavage of the ribose 3' C-O bond at the third residue were generally observed in low abundance. If the formation of these sequence ions occurs predominantly via a mechanism involving a deprotonated phosphate group located to the 3' side of the cleavage site, then the observed preferences for formation of different w ions would appear to arise from a greater probability of the charge being located on phosphate groups at these cleavage sites. Furthermore, these results are consistent with the expected distribution of the two charges whereby minimum Coulombic repulsion is experienced when the two charges are on the terminal phosphate groups (ie. furthest apart). In contrast, the structure that would result where one of the charges was residing on the third residue would be considerably less stable owing to the close proximity of the two charges. In the MS/MS spectra of heterogeneous oligonucleotides, the base identity was also observed to influence the relative intensities of w ions to some extent. Where thymine was located on the residue to the 5' side of the cleavage site, the resulting w ions were lower in abundance in comparison to the relative abundance of the same w ions generated in the MS/MS spectra of other oligonucleotides where thymine was not located at the cleavage site. The effects of base identity, however, were not found to overcome the trends arising from the effects of charge location in directing formation of the major ion types observed. The relative abundances of (a-B) ions showed little correlation with the relative abundances of the complementary w ions. For example, MS/MS spectra of the majority of oligonucleotides showed the formation of $[a_3-B_3H-2H]^-$ ions to be particularly favoured. The only exceptions to this were the MS/MS spectra of the

unibase oligonucleotides 5'-d(CCCCCC)-3' and 5'-d(TTTTTT)-3' in which the $[a_2-B_2H-2H]^-$ ions were the most abundant of the (a-B) series followed by the $[a_3-B_3H-2H]^-$ ions. This preference for loss of the base and cleavage of the ribose C-O bond to the 3' side of the third residue for the (a-B) series is in contrast to the trends for the complementary w ion series whereby w_1 and w_5 ions were the most abundant. This may be a consequence of the effects of multiple collisions of larger (a-B) ions into smaller (a-B) ions. However, it is difficult to explain why (a-B) ions would be influenced by this to a greater extent than the w ions. Hence an alternative explanation is that the w and (a-B) ions are formed via different mechanisms. The effects of base identities were also found to influence the relative intensities of (a-B) ions. In the majority of cases, a lack of preference is exhibited for the formation of (a-B) ions generated at thymine-containing residues. Furthermore, the propensities for (a-B) ion formation in some cases correlate with the observed preferences for nucleobase loss generated in corresponding MS/MS spectra.

The most favourable pathways involving loss of the nucleobase (either as an anion or a neutral) in the MS/MS spectra of doubly charged heterogeneous oligonucleotides were found to be predominantly directed either by charge location and/or base location. In all cases, the MS/MS spectra of the $[M-2H]^{2-}$ ions of heterogeneous hexamers showed the most intense base anion to arise from the base at the 5' terminal position. This is also consistent with a preference for the terminal phosphate groups being deprotonated which can catalyse loss of the 5' terminal base. Apart from the effect of base location influencing loss of the 5' terminal base, the order of base anion intensities of the non 5' terminal bases appeared to be more dependent on base identity rather than base location. In all cases, formation of C^- was observed to be

the least favourable of all base anions. Moreover, a preference is observed for $T^- > A^-$ in all cases where these bases are not at the 5' terminus. From the set of oligonucleotides examined in which G and T or G and A are not at the 5' terminus, the propensity for formation of G^- compared with that of T^- or A^- appears to be strongly biased by the relative number of each of these bases present in each oligonucleotide. Therefore, in oligonucleotides which comprise more T or A bases compared with the number of G bases, a preference of $T^- > G^-$ or $A^- > G^-$ was observed. Conversely, where the number of G bases exceeded that of either T or A in the oligonucleotide, the G^- ion was the more abundant.

The relative propensities for formation of ions owing to loss of a neutral nucleobase (with the oligonucleotide remaining intact), appeared to be strongly influenced by base location, with base identity also having a significant effect in the case of neutral thymine loss. In all MS/MS spectra, except where thymine was located at the 3' terminal position, the most intense ion arising from neutral base loss corresponded to loss of the 3' terminal base. This preference may arise from the stability of this ion to further decomposition via ribose 3' C-O bond cleavage resulting in the loss of H_2O , which has been generally observed to be a less facile process than ribose 3' C-O bond cleavage of ions formed from neutral base loss at other positions. The lack of preference for loss of neutral thymine was also a strong trend observed in the MS/MS spectra of all oligonucleotides in which thymine was not a terminal base. In these cases, no ions for loss of neutral thymine were observed, which appears to suggest that the location of thymine at the 5' or 3' terminus is integral to its ability to be lost as a neutral species. These observations in the case of neutral thymine loss are consistent with results obtained from other studies which showed that neutral thymine is a significantly less

stable leaving group in comparison to the other nucleobases¹. The preference for neutral base loss of the 3' terminal base appears to be followed in most case by loss of the 5' terminal base, although this was observed to be dependent on oligonucleotide sequence composition and the identity of the 5' terminal base. This is evident from the results obtained with 5'-d(TGCTGA)-3' which show loss of GH to be more favourable than loss of TH. The results obtained here also suggest that the propensities for neutral base loss for the internal bases are influenced by a interplay of the influences of sequence composition, base location, and base identity. The relative contribution of these effects and their overall influence on the observed propensities for neutral base loss is difficult to discern from the results obtained in this work and warrants further study.

5.3 THE EFFECTS OF PRECURSOR ION CHARGE STATE

In order to investigate the effects of charge state on oligonucleotide fragmentation in the negative ion mode, ESI-MS/MS spectra were acquired for all charge states (1⁻ to 5⁻) of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3'. The results obtained for the MS/MS spectra of singly charged and doubly charged precursor ions of these oligonucleotides have already been discussed in detail (see chapter four section 4.6 for the discussion of results for singly charged precursor ions, section 5.2 in this chapter for discussion of the results for doubly charged precursor ions). Figures 5.8-5.13 show the ESI-MS/MS spectra of the (a) [M-3H]³⁻, (b) [M-4H]⁴⁻ and (c) [M-5H]⁵⁻ ions of 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3' and 5'-d(GCCGGC)-3' respectively. The product ion assignments of all MS/MS spectra of each oligonucleotide are given in appendix 5.2.

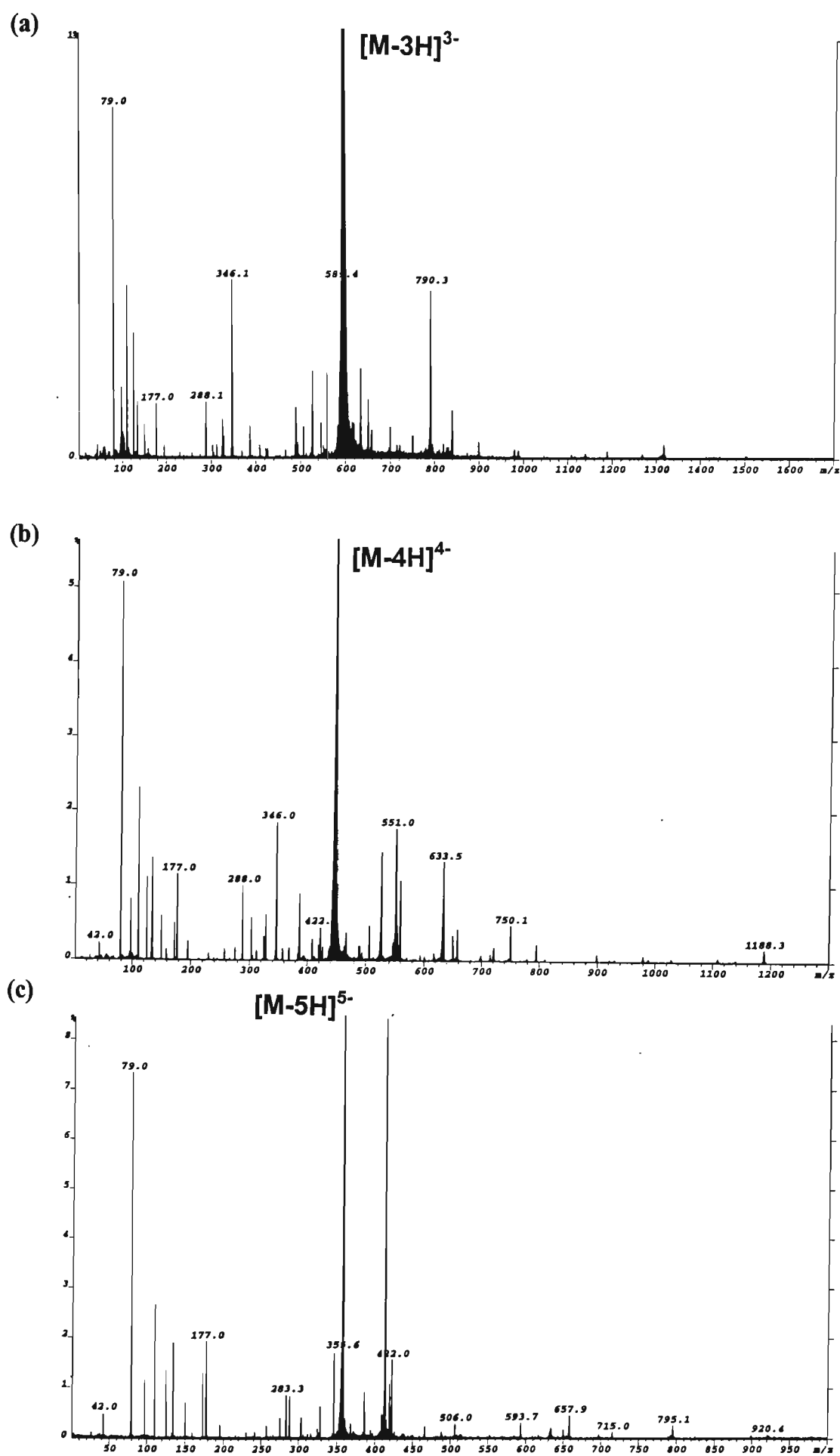


Figure 5.8: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(CACGTG)-3'.

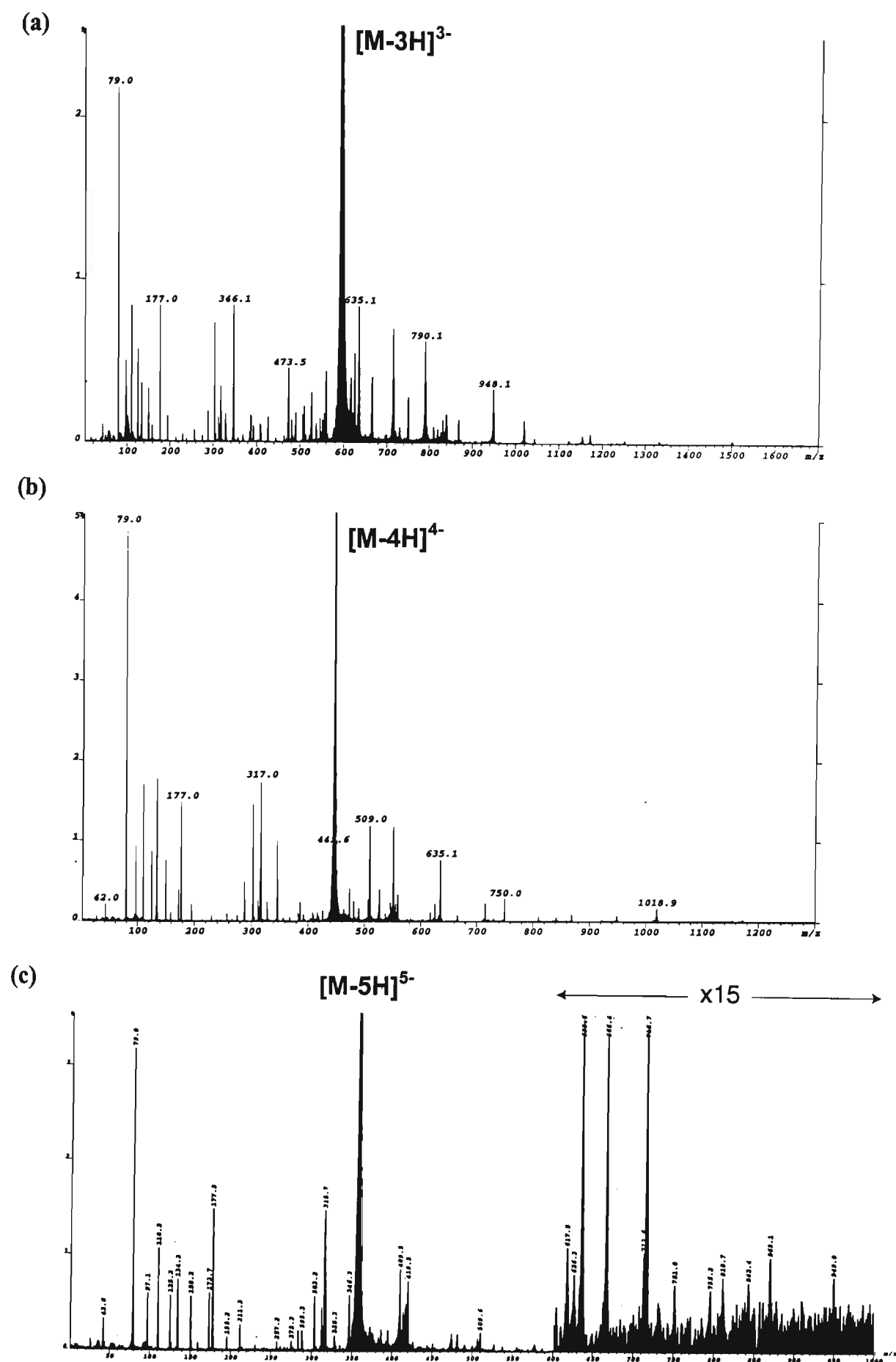


Figure 5.9: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(CGTACG)-3'.

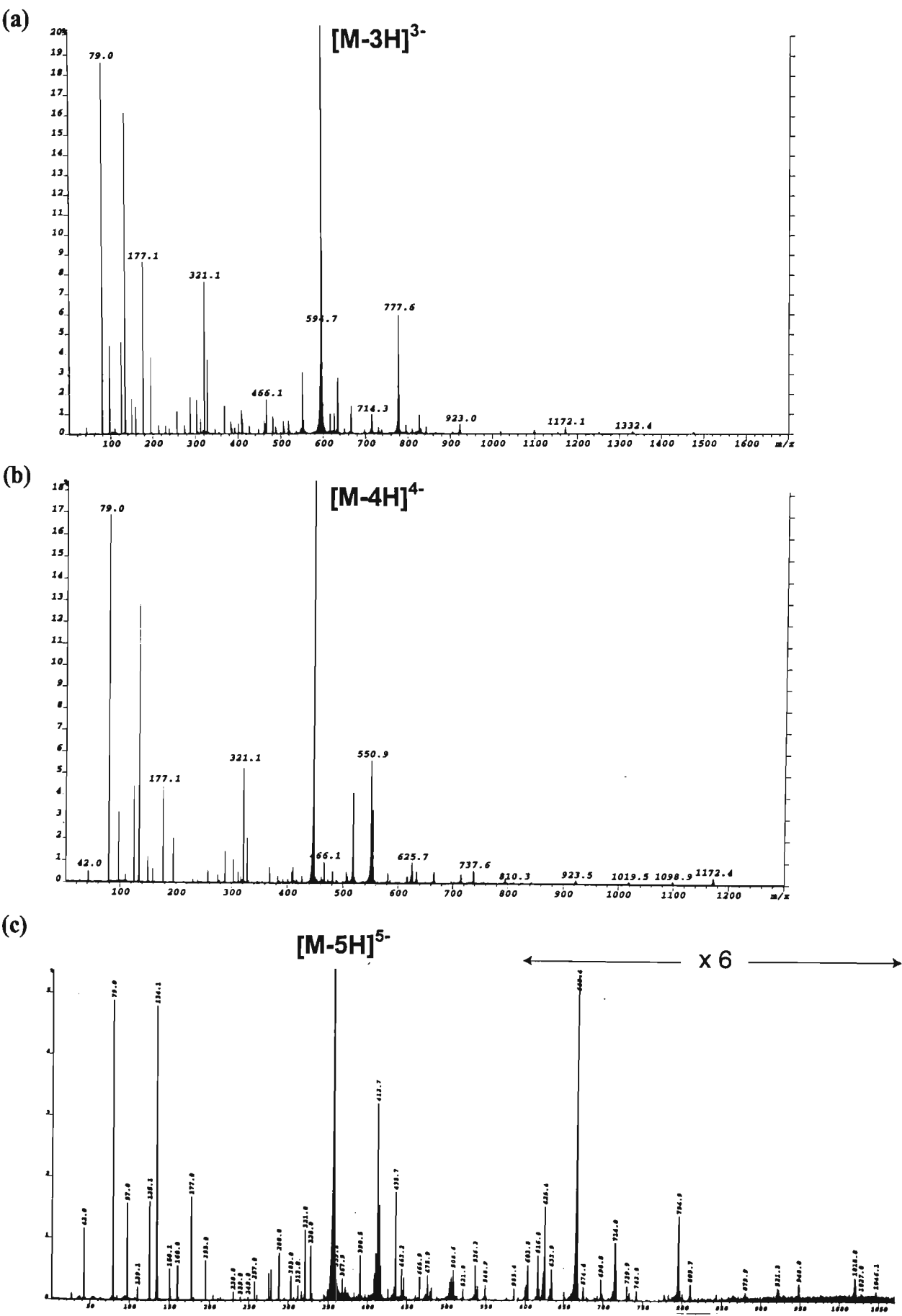


Figure 5.10: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(ATGCAT)-3'.

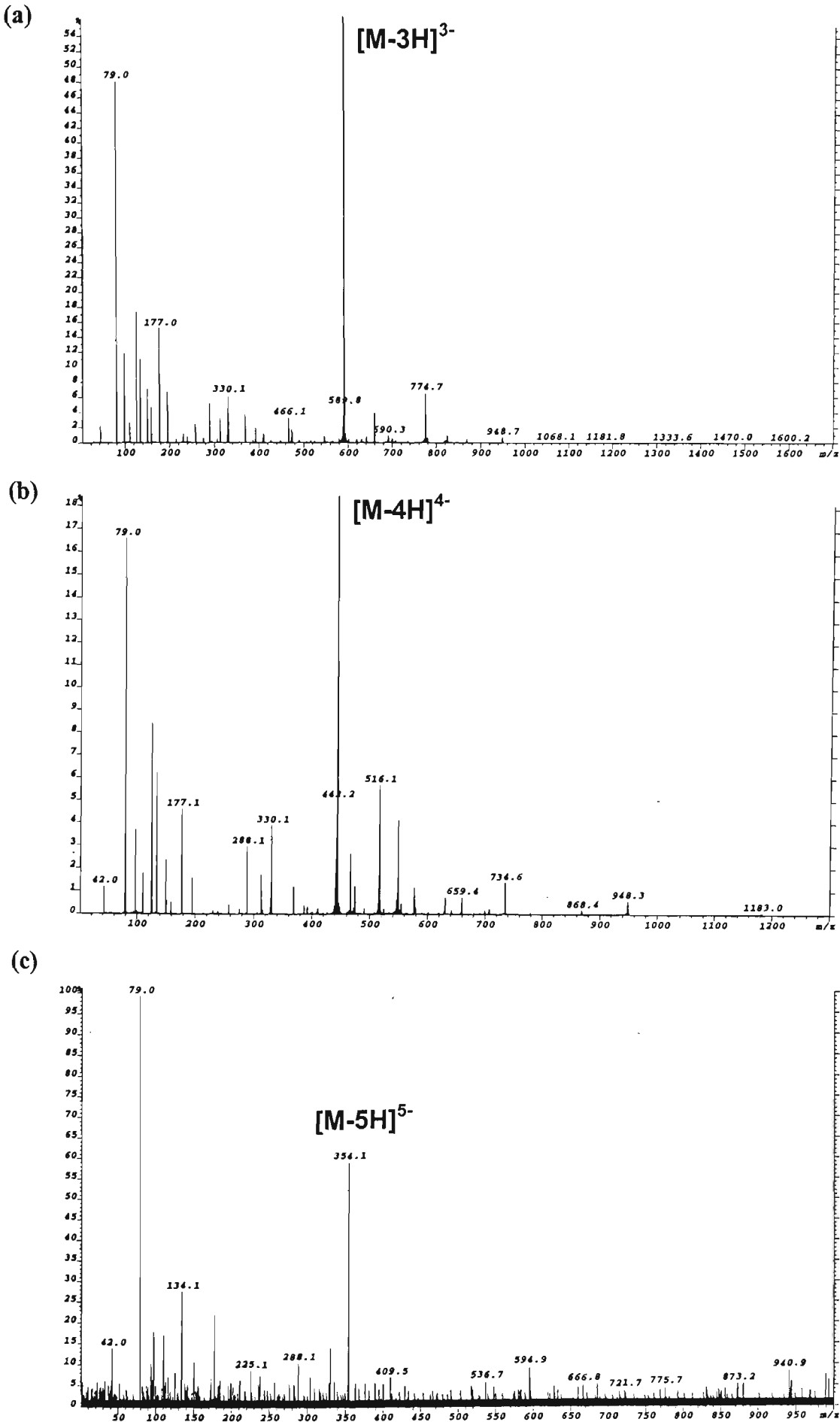


Figure 5.11: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(TCACGA)-3'.

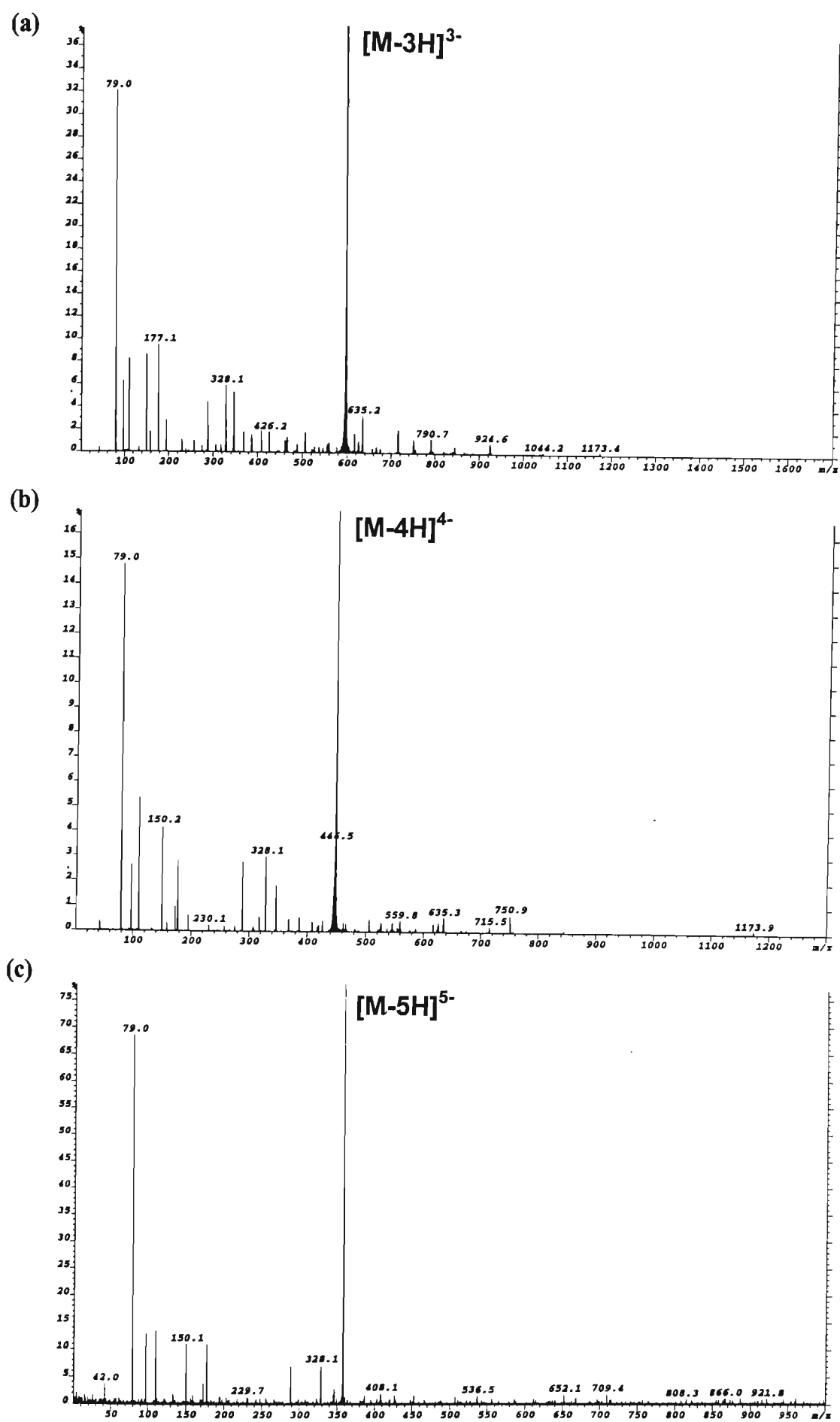


Figure 5.12: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(CGGCCG)-3'.

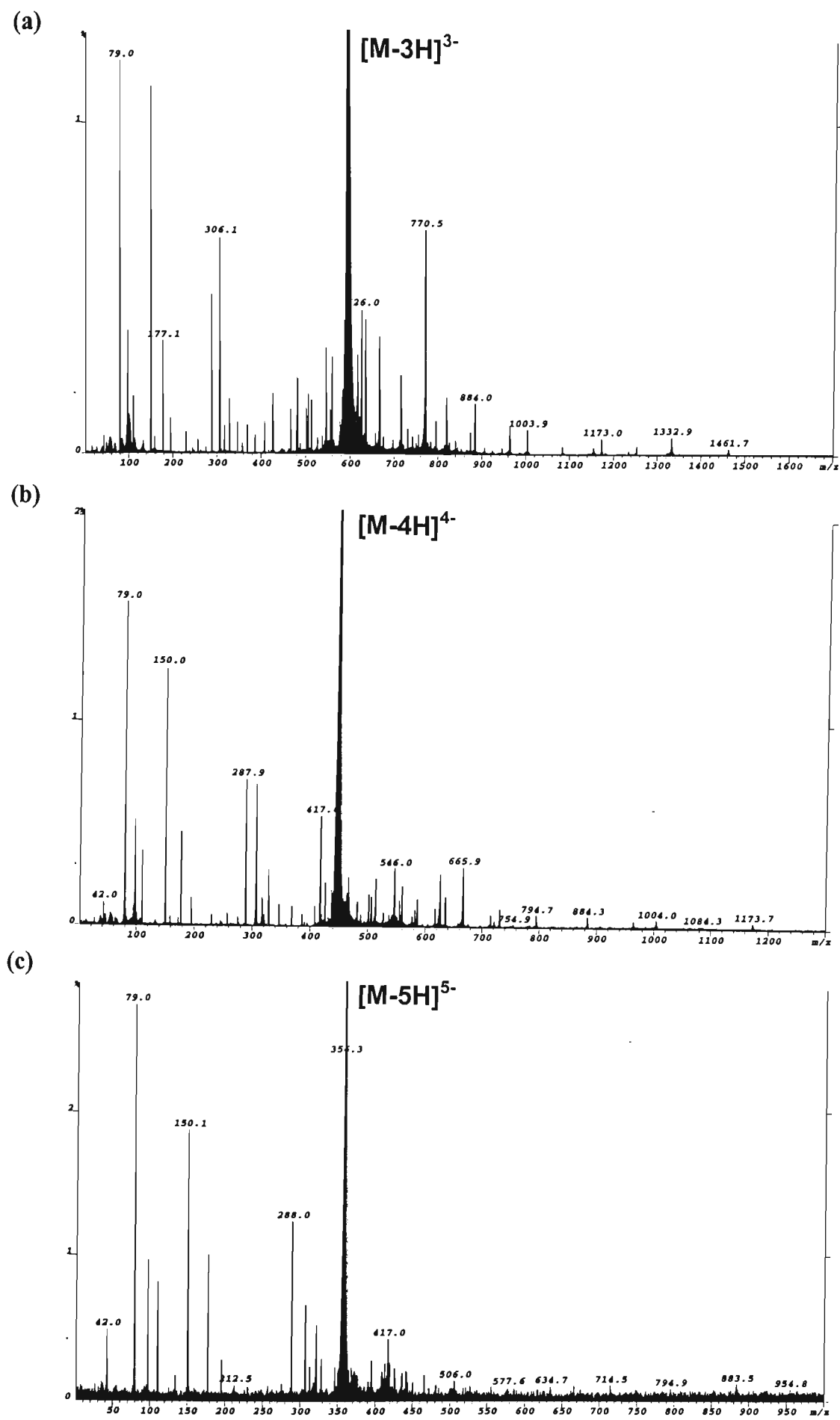


Figure 5.13: ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$, and (c) $[M-5H]^{5-}$ ions of the oligonucleotide 5'-d(GCCGGC)-3'.

Only hexamers were chosen for this study so as to remove the effects of oligonucleotide length on observed fragmentation behaviour. The set of oligonucleotides examined comprised all four possibilities for the 5' terminal base and 3' terminal base so that the effect of base identity and base location could also be examined to some degree. The MS/MS spectrum of each charge state anion for each oligonucleotide was analysed with regard to the relative intensities of ions arising from: (a) the loss of nucleobases, and (b) ribose 3' C-O bond cleavage of the residue from which a base is lost resulting in the formation of w and (a-B) sequence ions. In the case of the MS/MS spectra of multiply charged precursor oligonucleotide anions, the intensities of all the charge states product ions generated for any given w or (a-B) cleavage were also summed for comparison. For all charge states examined, it is assumed that the negative charges on the oligonucleotide are preferentially located on the phosphodiester backbone rather than on the nucleobases. This is based on semiempirical calculations made by Rodgers *et al.*¹ for the relative acidities of the most acidic "non-attachment" sites of nucleobases, G (333.4 kcalmol⁻¹), T (344.1 kcalmol⁻¹), C (345.6 kcalmol⁻¹), A (349.0 kcalmol⁻¹) compared with that determined for a phosphodiester bridge (329 kcalmol⁻¹). From these results, favourable deprotonation would be expected to occur at the phosphate group owing to its greater acidity.

5.3.1 Nucleobase Loss

The effect of increasing charge state upon the observation of product ions arising from charged base loss (ie. $[M-B^--nH]^{(n+1)-}$) and neutral base loss (ie. $[M-BH-nH]^{n-}$) in MS/MS spectra of oligonucleotide anions is illustrated in figure 5.14. For the MS/MS spectrum of different charge states of each oligonucleotide, the

intensities of product ions corresponding to the oligonucleotide minus a neutral nucleobase were summed and compared with the sum of the intensities of product ions owing to the oligonucleotide minus a charged base. Base anion intensities were not used for the purposes of this comparison since their relative abundances can include contributions from multiple decompositions of other product ions and so may not be representative of fragmentation predominantly involving base anion loss from the precursor ion.

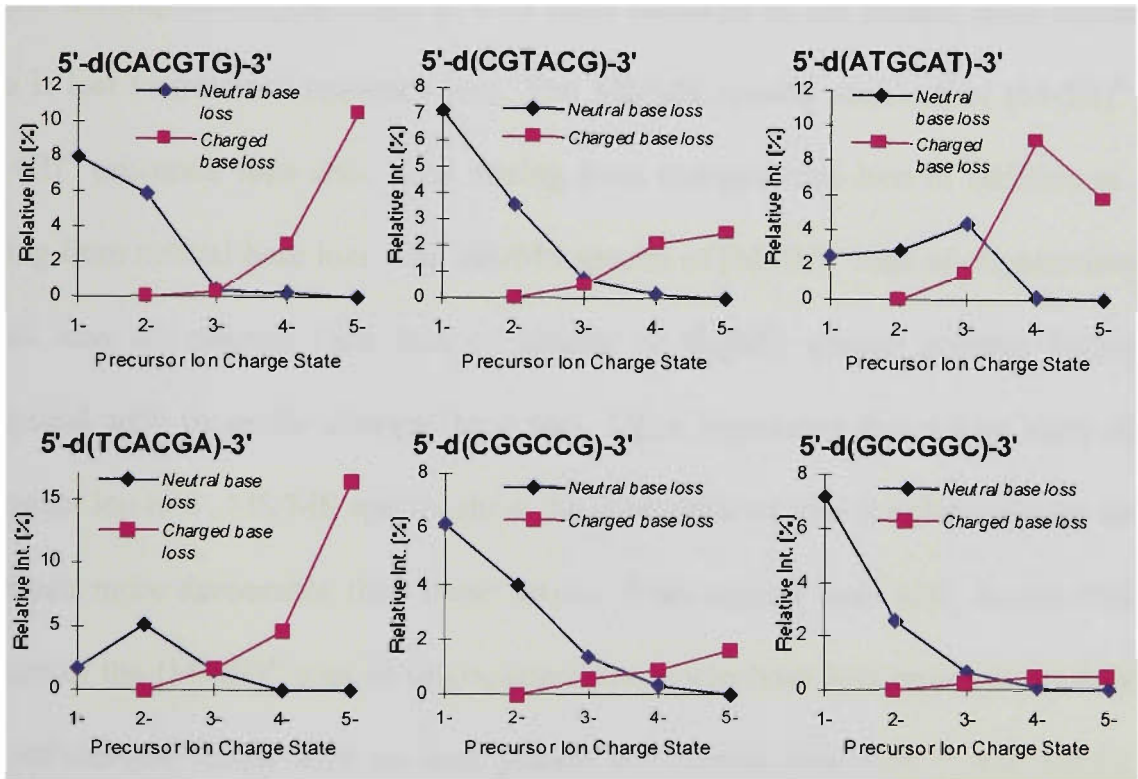
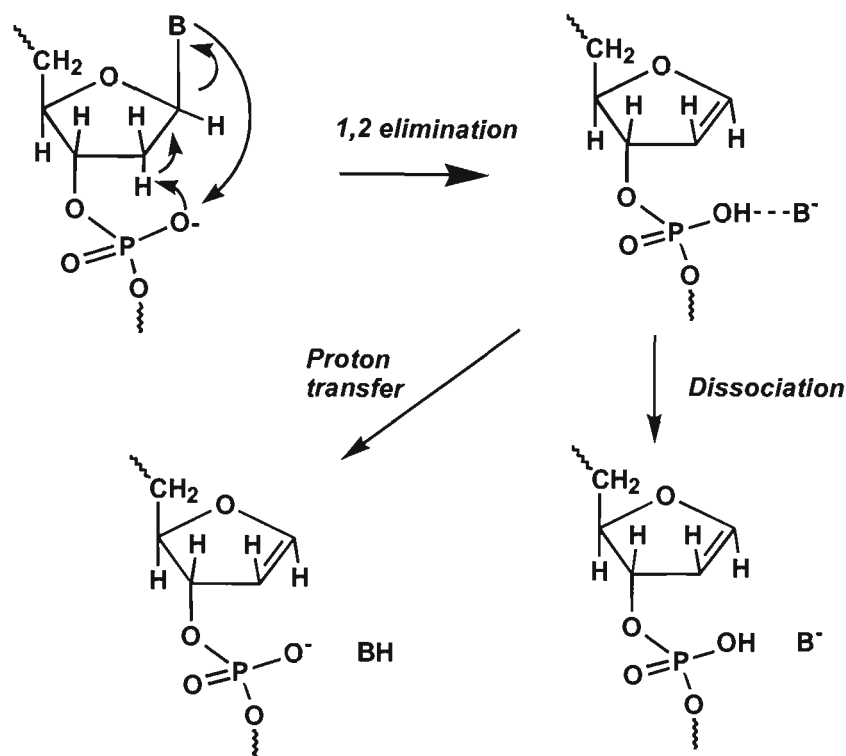


Figure 5.14: Comparison of the summed relative intensities of the product ions corresponding to neutral base loss versus charged base loss from the precursor ion in ESI-MS/MS spectra of oligonucleotides.

In the case of singly charged precursor ions, fragmentation involving base loss will yield an ion for either the charged base, or the oligonucleotide fragment minus a neutral base, depending on whether the charge remains with the base or the oligonucleotide respectively. For precursor ions having two or more charges, charged base loss from the precursor ion can yield ions for both the charged base and the complementary oligonucleotide fragment minus the charged base. In all MS/MS spectra

acquired of doubly charged oligonucleotide anions, no ions are observed for the oligonucleotide fragment resulting from loss of a base anion although nucleobase anions are observed in MS/MS spectra. The fact that these complementary ions for charged base loss are not present may indicate that charged base loss directly from the precursor ion is unfavourable for doubly charged ions, and that the base anions observed are owing to subsequent decomposition of product ions from further collisions. Alternatively, it may be that the ions formed from charged base loss are unstable to further decomposition via ribose 3' C-O bond cleavage of the residue from which the base is lost to generate sequence ions. The MS/MS spectra acquired of $[M-3H]^{3-}$ and $[M-4H]^{4-}$ precursor ions show ions arising from charged base loss in addition to ions arising from neutral base loss. The MS/MS spectra of $[M-3H]^{3-}$ ions of oligonucleotides yields ions for neutral base loss of similar or slightly greater relative intensities compared with those for charged base loss. Upon increasing the charge state of the precursor ion to 4⁻, MS/MS spectra show the production of ions for charged base loss to be much more favourable than those arising from neutral base loss. In the MS/MS spectra of the $[M-5H]^{5-}$ ions of oligonucleotides, nucleobase loss proceeds entirely via loss of charged bases, with no ions present for neutral base loss. These results are indicative of the significant effect of precursor ion charge state on the propensity for charged base loss versus neutral base loss with the loss of charged bases being more favourable upon increasing charge state in MS/MS spectra of all oligonucleotides examined. This is also consistent with previous studies conducted by McLuckey *et al.*⁷ in which precursor ion charge state was found to be a major factor influencing whether a base is lost as a charged or a neutral species. These observations were rationalised on the basis of fragmentation occurring via the formation and subsequent dissociation of an

intermediate proton-bound collision complex between the cleaved base and the oligonucleotide, derived from the mechanism for base loss proposed by Rodgers *et al.*¹ which is illustrated in scheme 5.1.



Scheme 5.1: Mechanism proposed by Rodgers *et al.*¹ for nucleobase loss in deprotonated dinucleotides.

The propensity for charged versus neutral base loss was suggested to involve competition between the leaving base and deprotonated phosphate group for the proton upon dissociation of the collision complex. For precursor ions with maximum deprotonation of the phosphodiester backbone, nucleobase loss was found to occur solely as a charged species consistent with the unfavourable dissociation of a proton-bound intermediate complex via proton transfer from the phosphate group to the base upon fragmentation of highly charged precursor ions. Conversely, the facile production of ions for neutral base loss in the MS/MS spectra of doubly charged oligonucleotide precursor ions, in combination with the absence of oligonucleotide product ions owing

to charged base loss, suggests preferential retention of the proton by the leaving base upon dissociation of a proton-bound complex in the MS/MS of doubly charged precursor ions.

(a) Neutral Base Loss

In addition to influencing the competition for charged and neutral base loss, precursor ion charge state has a significant effect on the relative propensities for both charged or neutral base loss of the four bases. The effect of increasing charge state on the propensity for neutral base loss is observed to vary depending on oligonucleotide sequence and base location. Figures 5.15 and 5.16 show the trends for neutral base loss observed in the MS/MS spectra of the 1^- to 4^- ions of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' respectively.

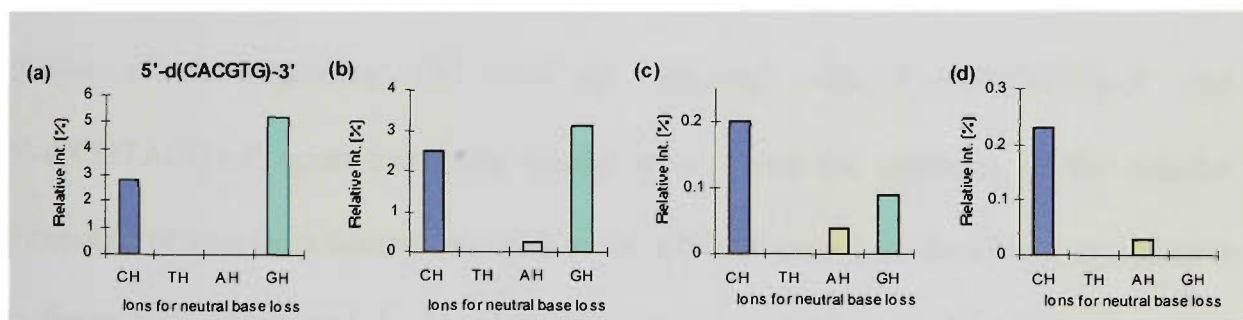


Figure 5.15: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a) $[M-H]^-$, (b) $[M-2H]^{2-}$, (c) $[M-3H]^{3-}$, and (d) $[M-4H]^{4-}$ ions of 5'-d(CACGTG)-3'.

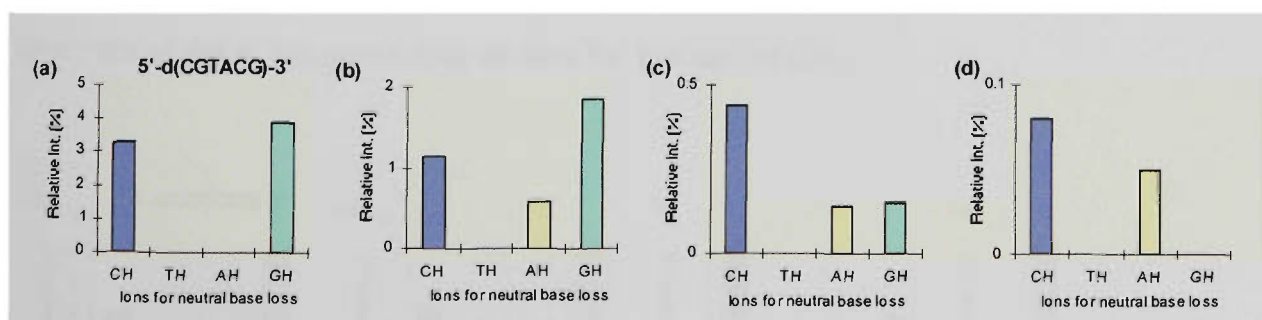


Figure 5.16: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a) $[M-H]^-$, (b) $[M-2H]^{2-}$, (c) $[M-3H]^{3-}$, and (d) $[M-4H]^{4-}$ ions of 5'-d(CGTACG)-3'.

The MS/MS spectra of both oligonucleotides show highly similar propensities for charged base loss for all four charge state anions. A striking feature of all MS/MS spectra obtained for these oligonucleotides is the absence of a signal for the loss of TH. The MS/MS spectra of the $[M-H]^-$ and $[M-H]^{2-}$ ions show the loss of GH to dominate followed by loss of CH. The MS/MS spectrum of the doubly charged anion also shows a small ion for loss of AH which is absent in the MS/MS spectrum of the singly charged anion. Upon increasing the precursor ion charge state to the 3^- , the strong preference for loss of GH observed in the MS/MS spectra of 1^- and 2^- precursor ions greatly diminishes, with instead the loss of CH producing ions of substantially greater relative intensity. This effect is even more significant in MS/MS spectra of 4^- precursor ions which show no ions for the loss of GH.

The MS/MS spectra obtained for the oligonucleotide 5'-d(CG GCCG)-3' shows similar effects regarding GH loss as observed with 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' upon increasing charge state. From the summary of the relative intensities of ions from neutral base loss in the MS/MS spectra of the 1^- to 4^- ions shown in figure 5.17, singly and doubly charged precursor ions show preferential loss of GH over loss of CH. In contrast, increasing the precursor ion charge state to 3^- results in more predominant ions owing to loss of CH relative to the loss of GH, with the MS/MS spectrum of the 4^- ion generating no ions for the loss of GH.

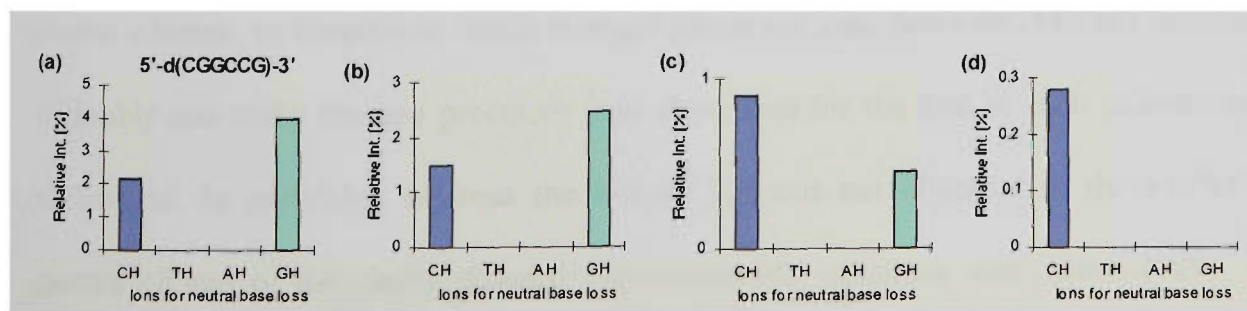


Figure 5.17: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a) $[M-H]^-$, (b) $[M-2H]^{2-}$, (c) $[M-3H]^{3-}$, and (d) $[M-4H]^{4-}$ ions of 5'-d(CG GCCG)-3'.

In the case of the oligonucleotide 5'-d(GCCGGC)-3', the trends for neutral base loss, which are shown in figure 5.18, are similar for the MS/MS spectra of the 1⁻ to 3⁻ precursor ions. Each spectrum shows ions of quite comparable intensities for the loss of CH and GH, with the loss of CH being slightly more favoured in the MS/MS spectra of the 1⁻ and 2⁻ charge states. The MS/MS spectrum of the 4⁻ precursor ion, however, shows a significant decrease in the relative signal intensity for ions owing to loss of GH. In this case it would appear that the preferential loss of the 5'-terminal base with increasing charge states is overcome by a lack of preference for loss of neutral guanine upon fragmentation of highly charged precursor ions.

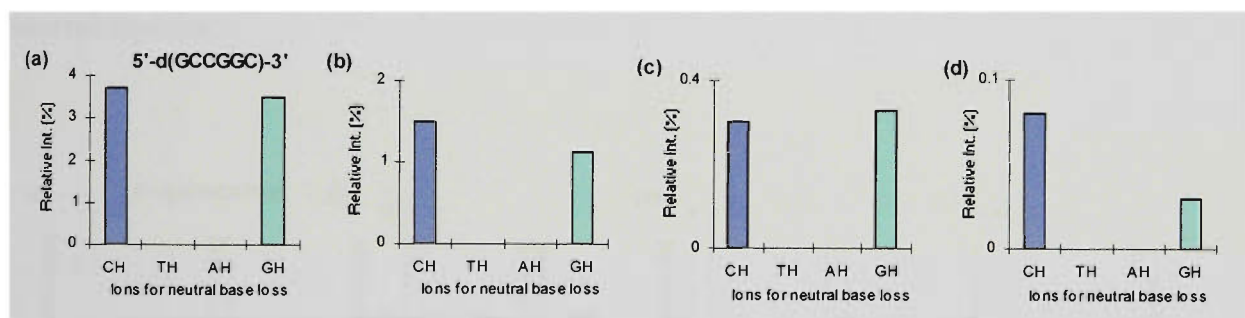


Figure 5.18: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, and (d)[M-4H]⁴⁻ ions of 5'-d(GCCGGC)-3'.

Figures 5.19 and 5.20 show the trends for neutral base loss observed in the MS/MS spectra of the 1⁻ to 4⁻ charge states of the oligonucleotides 5'-d(ATGCAT)-3' and 5'-d(TCACGA)-3'. The MS/MS spectra for the singly, doubly and triply charged precursor ions of 5'-d(ATGCAT)-3' and 5'-d(TCACGA)-3' show preferential loss of neutral adenine. In contrast to singly charged precursor ions, however, MS/MS spectra of doubly and triply charged precursor ions show ions for the loss of each constituent nucleobase. In particular, whereas the loss of TH was not observed in the MS/MS spectra of any of the singly charged oligonucleotide precursor ions, loss of TH is evident in the MS/MS spectra of doubly and triply charged ions of both

5'-d(ATGCAT)-3', and 5'-d(TCACGA)-3'. This may suggest that the location of thymine at the 3' or 5' terminus is important in overcoming the strong lack of preference for the loss of TH that was evident in the MS/MS spectra of all charge states of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. The fragmentation of the 4⁻ precursor ions shows highly specific neutral base loss similar to that generated in the MS/MS spectra of singly charged precursor ions. The MS/MS spectrum of 5'-d(ATGCAT)-3' yields ions for neutral base loss solely of adenine and in the MS/MS spectrum of 5'-d(TCACGA)-3' the observation of neutral base loss for all bases was found to be disfavoured with only a very weak intensity ion detected for the loss of neutral thymine.

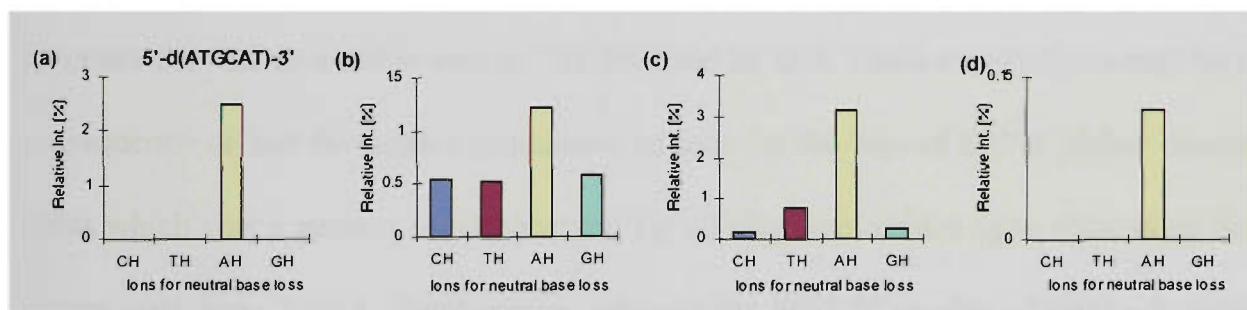


Figure 5.19: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, and (d)[M-4H]⁴⁻ ions of 5'-d(ATGCAT)-3'.

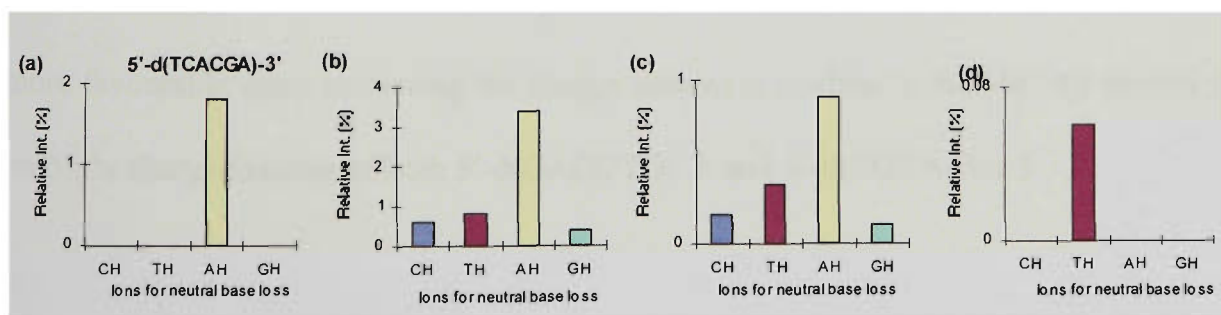


Figure 5.20: Relative intensities of ions for neutral base loss generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, and (d)[M-4H]⁴⁻ ions of 5'-d(TCACGA)-3'.

From the results for neutral base loss it can be seen that the fragmentation of the singly and doubly charged oligonucleotide anions generates similar trends for the

propensity for neutral base loss which generally show ions for neutral base loss of the 3' terminal base to be the most abundant followed by that of the 5' terminal base. These trends for neutral base loss are however overcome by the strong lack of preference for the loss of neutral thymine. Upon shifting to the 3⁻ charge state, more preferential observation of ions for neutral base loss of the 5' terminal base is observed except for the oligonucleotide 5'-d(TCACGA)-3' which is consistent with the unfavourable loss of neutral thymine. Fragmentation of 4⁻ precursor ions show a general lack of preference for the observation of ions from neutral base loss from the 3'-terminus with the neutral loss of the 5' terminal base generally observed to be highly favoured in MS/MS spectra. In all MS/MS spectra of 4⁻ oligonucleotide ions, no ions were observed for neutral base loss from the 3' terminus except in the case of 5'-d(GCCGGC)-3' for which the MS/MS spectrum showed favourable loss of CH followed by GH. These observations may be a consequence of less favourable production of ions for the loss of GH at higher charge states which was a general trend observed for all oligonucleotides upon increasing the charge state from 2⁻ to 4⁻. Furthermore, whereas the MS/MS spectra of singly charged precursor anions show the loss of AH only to occur for oligonucleotides containing adenine at the 5' or 3' terminus, the production of ions for AH loss was found to be more favourable upon increasing the charge state as is evident in the MS/MS spectra of multiply charged anions of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

(b) Base Anion Loss

The effect of increasing charge state on base anion loss in the MS/MS spectra of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' is shown in figures 5.21 and 5.22 respectively. The relative base anion intensities observed in the MS/MS

spectrum of each charge state anion yields similar results for both oligonucleotides, with some significant changes in the preference for base anion formation being observed upon increasing precursor ion charge states. In particular, a marked difference in the relative propensities of base anion formation is evident in the MS/MS spectra of singly charged versus multiply charged precursor ions. Whereas the MS/MS spectra of singly

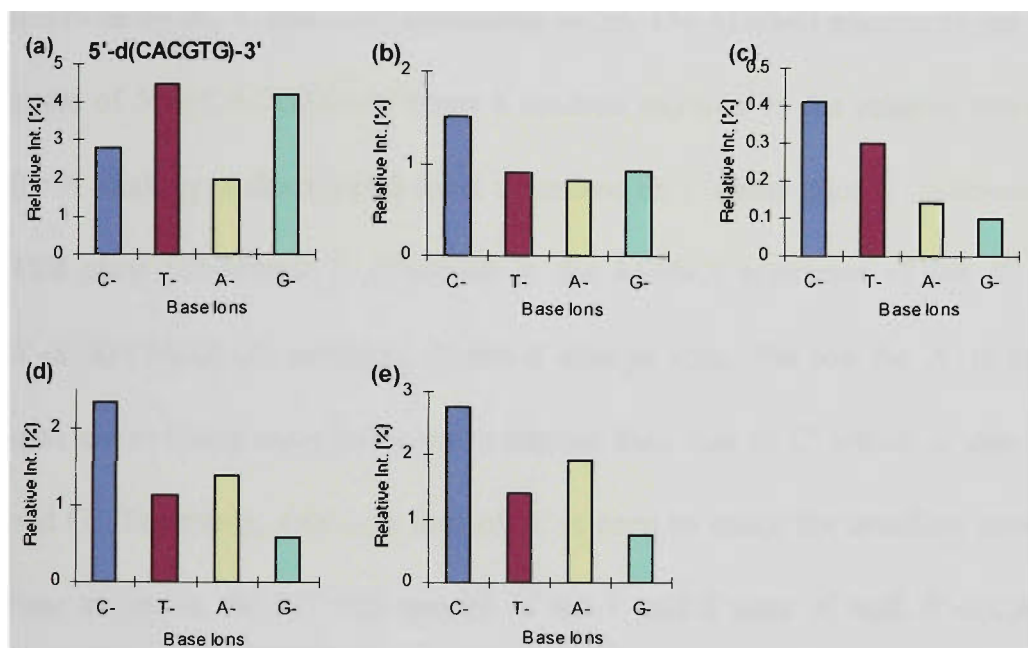


Figure 5.21: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(CACGTG)-3'.

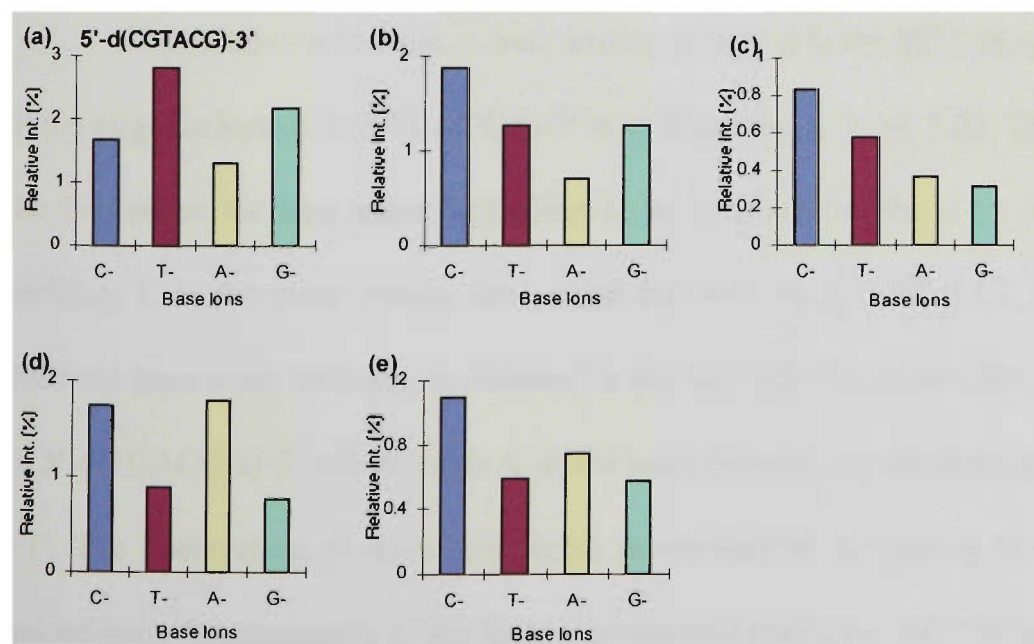


Figure 5.22: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(CGTACG)-3'.

charged anions of both oligonucleotides show T^- to be the most abundant base anion, loss of C^- is preferred for the 2^- to 5^- precursor ions. In the MS/MS spectra of their doubly charged anions, both oligonucleotides show the same order of base anion intensity being that of $C^- > G^- > T^- > A^-$. Upon increasing the charge state to 3^- , resulting MS/MS spectra show the G^- ion to produce the smallest signal of the base anions followed by A^- , T^- and C^- in increasing order. The MS/MS spectra of the 4^- and 5^- charge states of $5'$ -d(CACGTG)- $3'$ show a marked increase in the relative intensity of the ion for A^- making it the second most abundant base anion after C^- followed by T^- and G^- . This same preference is observed in the MS/MS spectrum of the 5^- charge state of $5'$ -d(CGTACG)- $3'$, however, in the 4^- charge state, the ion for A^- is the most intense base anion being marginally more intense than that of C^- which is then followed by T^- and G^- . Therefore, although loss of A^- is seen to make the smallest contribution of the base anions in the MS/MS spectra of the 1^- and 2^- ions of both $5'$ -d(CACGTG)- $3'$ and $5'$ -d(CGTACG)- $3'$, the formation of this ion is seen to become strongly favoured at high charge states.

The relative intensities of base anions generated in the MS/MS spectra obtained of the oligonucleotide $5'$ -d(TCACGA)- $3'$ is summarised in figure 5.23. The results show the preference for base anion intensities to be constant for the 1^- to 4^- charge states yielding T^- as the most intense base anion followed by $A^- > G^- > C^-$. A significantly different base anion ordering is obtained in the MS/MS spectrum of the 5^- charge state of $5'$ -d(TCACGA)- $3'$ which yields A^- as the most intense base anion followed by $C^- > G^- > T^-$. The observation of more favourable production of A^- relative to the other base anions upon fragmentation of the fully deprotonated precursor ion is consistent with the effects of increasing precursor ion charge state on relative base anion intensities

generated in the MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' described above.

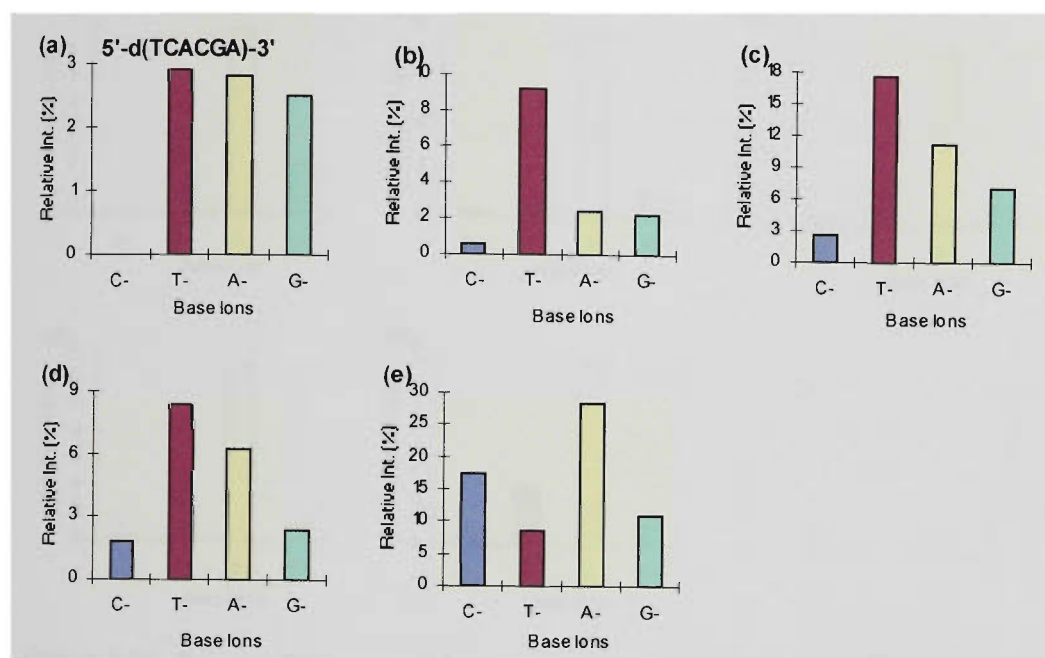


Figure 5.23: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(TCACGA)-3'.

Increasing precursor ion charge state for the oligonucleotide 5'-d(ATGCAT)-3' has minimal effect on the relative production of base anions in MS/MS spectra, as can be seen from the summarised data given in figure 5.24. The MS/MS spectra of the 1⁻ to 5⁻ charge states of 5'-d(ATGCAT)-3' show the same propensity for base anion formation of A⁻ > T⁻ > G⁻ > C⁻. The lack of preference for formation of C⁻ is observed to be strongest in the MS/MS spectrum of the singly charged precursor ion in which this ion is absent. The results for this oligonucleotide is the only case where the ordering for base anion intensity is observed to be in agreement with that obtained in low energy CID studies undertaken by Rodgers *et al.*¹ on deprotonated dinucleotides, and with low energy MS/MS work conducted by McLuckey *et al.*¹¹ on multiply charged oligonucleotides. These results however, are most likely to be a feature of

oligonucleotide sequence and base location (in particular, the location of the adenine at the 5'-terminus making loss of A⁻ the most facile process).

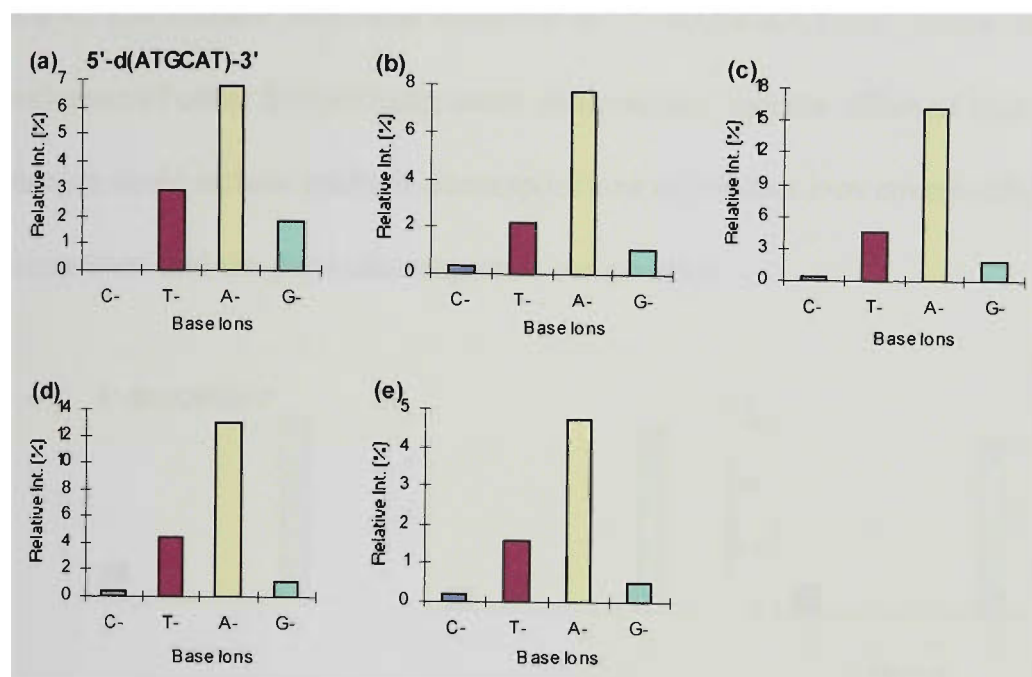


Figure 5.24: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(ATGCAT)-3'.

The results for base loss exhibited in the MS/MS spectrum of 5'-d(GCCGGC)-3' (figure 5.25) show a consistent preference for loss of G⁻ over C⁻ for all precursor ion charge states. The results obtained for 5'-d(CGGCCG)-3' (figure 5.26) show the most significant difference in propensities for base anion formation between singly and multiply charged precursor ions. Whereas favourable loss of G⁻ is observed for the singly charged precursor ion, the MS/MS spectra of the multiply charged ions show preferential loss of C⁻ in almost all cases. The MS/MS spectrum of the 3⁻ precursor ion yields very similar propensities for loss of both bases, with slightly favourable loss of G⁻ over C⁻. This ordering for base anion loss is not consistent with the propensities observed in the MS/MS spectra of 3⁻ ions of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' which comprise the same 5' and 3' terminal bases as 5'-d(CGGCCG)-3'. In those cases loss of C⁻ yields the most abundant base anion with

the ion for G⁻ generating the smallest base anion signal. Furthermore, the MS/MS spectra of all 3⁻ ions of oligonucleotides show the 5' base to yield the most intense base anion. The unusual behaviour observed for 5'-d(CGGCCG)-3' can be attributed to the influence of other factors being more predominant than the effect of base location. Such factors could include multiple decompositions of product ions contributing to base anion intensities, and the particular sequence composition.

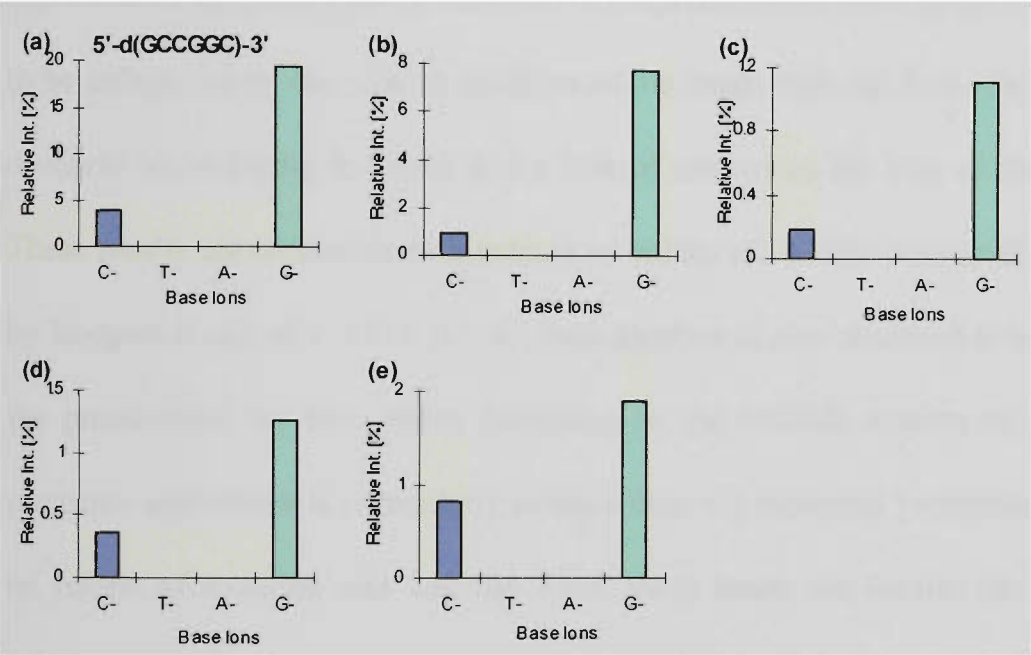


Figure 5.25: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(GCCGGC)-3'.

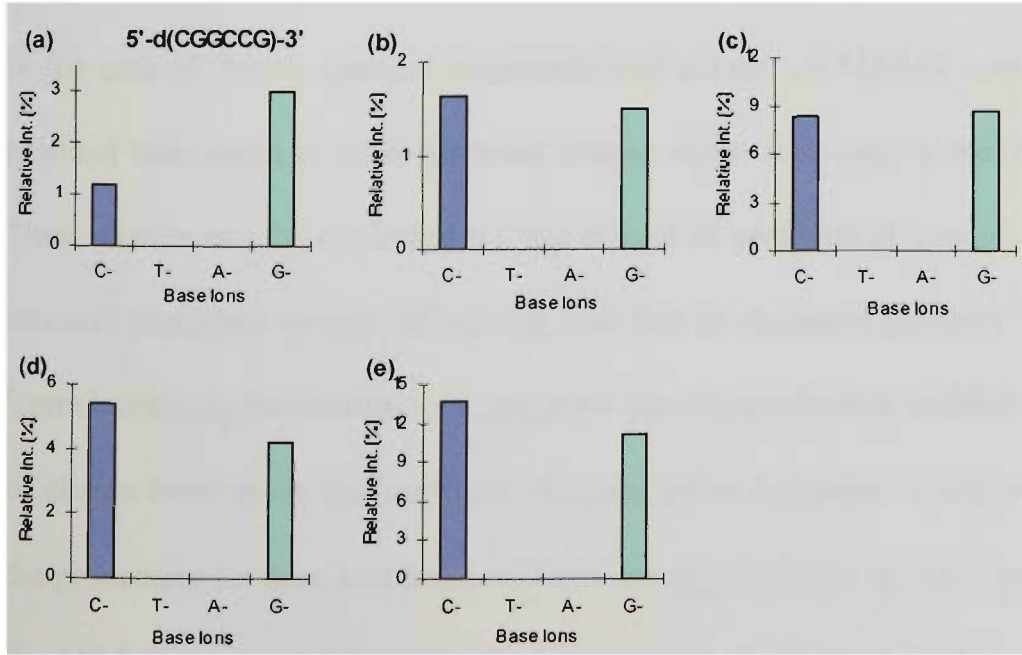


Figure 5.26: Relative intensities of base anions generated in the ESI-MS/MS spectra of the (a)[M-H]⁻, (b)[M-2H]²⁻, (c)[M-3H]³⁻, (d)[M-4H]⁴⁻, and (e)[M-5H]⁵⁻ ions of 5'-d(CGGCCG)-3'.

The results obtained for base anion loss as a function of oligonucleotide charge state show base loss to be directed by both charge location and base identity although the extent of the contribution of each of these factors is also a function of precursor ion charge state. The MS/MS spectra obtained for multiply charged oligonucleotide anions in almost all cases yielded markedly different preferences for base anion intensities from that observed for singly charged precursor ions. In the MS/MS spectra of singly charged deprotonated oligonucleotides the relative propensities for base anion formation appear to be influenced by the relative acidities of the bases with the loss of charged thymine observed to be highly favoured and a lack of preference for loss of charged cytosine. These results are consistent with calculated values of relative base acidities determined by Rodgers *et al.*¹ of $T > G > A > C$. Base location is also observed to have an effect in the propensities for base anion formation in the MS/MS spectra of singly charged precursor ions which is particularly evident from the increased preference for formation of anions of cytosine and adenine when these bases are located at the 5'-terminal position. Base location is also seen to have an effect in the case of charged adenine loss which is observed to be highly favourable when it is located at the 5'-terminal position. In the case of doubly charged oligonucleotide anions, all MS/MS spectra show the 5' terminal base anion to yield the most intense signal followed by the 3' terminal base. These results can be explained by the effects of preferential charge location on the terminal phosphate groups influencing base loss as discussed previously in section 5.2. Upon increasing the number of charges on the oligonucleotide backbone, the influence of charge location on the propensity for base anion formation is seen to diminish, and the propensity for base loss becomes more strongly dictated by base identity. Although the MS/MS spectra of the 2⁻ to 5⁻ charge states of oligonucleotides show an overall

preference for the loss of the 5'-terminal base anion (with only a few exceptions), the subsequent ordering for base anion intensities in MS/MS spectra is observed to vary significantly upon shifting to higher charge state precursor anions. The most striking observation is the substantial increase in propensity for loss of A⁻ in the MS/MS spectra of highly charged anions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(TCACGA)-3'. These findings would appear to be consistent with those obtained by McLuckey *et al.*² on ion trap instrumentation which show the preferential loss of A⁻ in the decomposition of highly charged species.

The effect of increasing charge state on the relative intensities of ions owing to charged nucleobase loss, ie. $[M-B-nH]^{(n+1)-}$ ions, in the MS/MS spectra of the 3⁻ and 4⁻ charge state ions corresponds reasonably well with the effects observed upon base anion intensities as discussed above. Figures 5.27 and 5.28 show the propensities for charged base loss generated in the MS/MS spectra of the 3⁻ to 5⁻ charge states of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. Fragmentation of the 3⁻ ions of the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' both show a predominant ion for the loss of C⁻ with the least abundant ion for charged base loss owing to the loss of A⁻. Increasing the charge state to 4⁻ was found to result in a marked increase in the propensity for the loss of A⁻ which is consistent with the effects observed on the production of base anions for these charge states. The MS/MS spectrum of the 5⁻ ion of 5'-d(CACGTG)-3' exhibited a substantial increase in the preference for loss of A⁻ over other nucleobases compared with that found in the MS/MS spectrum of the 4⁻ ion. In contrast, the MS/MS spectrum of the 5⁻ ion of 5'-d(CGTACG)-3', shows a distinct reversal in charged base loss to that generated from fragmentation of the 4⁻ ion with the loss of A⁻ becoming the least favourable base anion lost and the loss of G⁻ and C⁻

dominating. One possible explanation for the sharp decrease in relative intensity of the ion owing to loss of A⁻ upon shifting from the 4⁻ to 5⁻ charge state may be owing to facile further decomposition of these product ions.

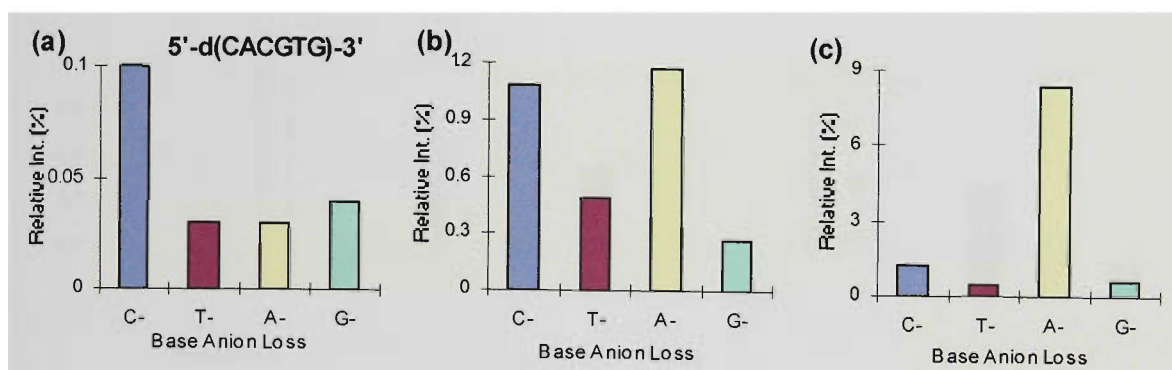


Figure 5.27: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$ and (c) $[M-5H]^{5-}$ ions of 5'-d(CACGTG)-3'.

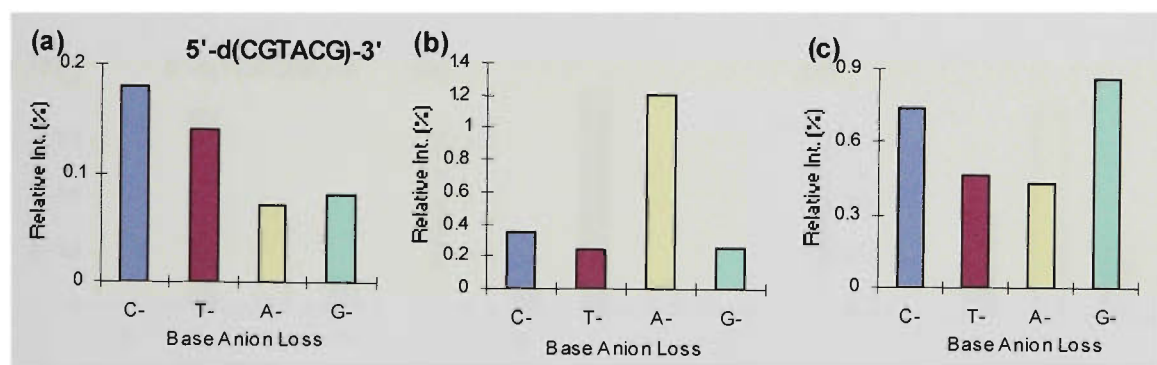


Figure 5.28: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a) $[M-3H]^{3-}$, (b) $[M-4H]^{4-}$ and (c) $[M-5H]^{5-}$ ions of 5'-d(CGTACG)-3'.

The production of ions for charged base loss in the MS/MS spectra of 5'-d(ATGCAT)-3' are summarised in figure 5.29. The same predominant trends are observed for the 3⁻ to 5⁻ charge states whereby adenine loss is the most favoured followed by that of charged thymine loss, with the formation of ions for charged cytosine loss strongly disfavoured. Slightly different preferences are observed in the case of the 4⁻ precursor ion in which no ions are detected for the loss of charged guanine although the corresponding ions are present in the MS/MS spectra of the 3⁻ and 5⁻ precursor ions. These observations regarding the loss of charged guanine in the MS/MS

spectrum of the 4⁻ precursor ion of 5'-d(ATGCAT)-3' are also evident in the trends for charged base loss observed in the MS/MS spectra of 5'-d(TCACGA)-3' which are summarised in figure 5.30.

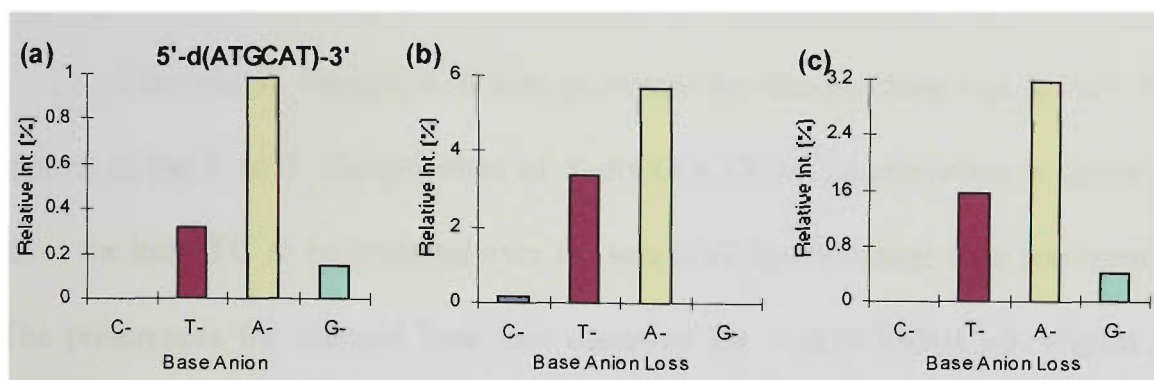


Figure 5.29: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a) [M-3H]³⁻, (b) [M-4H]⁴⁻ and (c) [M-5H]⁵⁻ ions of 5'-d(ATGCAT)-3'.

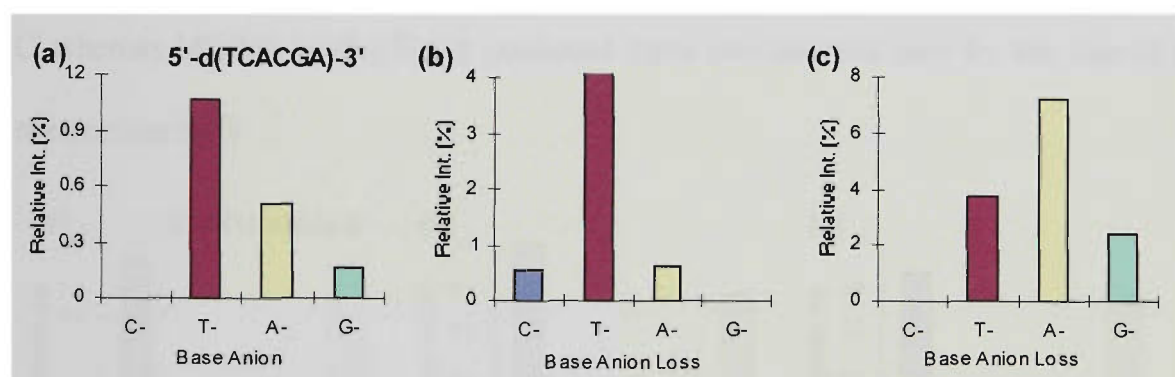


Figure 5.30: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a) [M-3H]³⁻, (b) [M-4H]⁴⁻ and (c) [M-5H]⁵⁻ ions of 5'-d(TCACGA)-3'.

The results for 5'-d(TCACGA)-3' show the same major preference in the MS/MS spectra of the 3⁻ and 4⁻ precursor ions with the loss of T⁻ being predominant followed by A⁻. The MS/MS spectrum of the 5⁻ charge state ion shows a significant change in the relative intensities of ions generated for charged base loss, with the loss of A⁻ yielding the most intense signal rather than T⁻. These observations of more favourable production of ions owing to the loss of charged adenine is consistent with the results obtained for charged base loss upon increasing precursor ion charge state in the MS/MS spectra for 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. Furthermore, these

results are consistent with the effect of increasing charge state on the propensities of formation of base anions and ions for neutral base loss for this oligonucleotide and suggests that the loss of adenine becomes more facile in the MS/MS of highly charged oligonucleotide anions.

The relative intensities of ions generated for charged base loss in the MS/MS spectra of the 3⁻ to 5⁻ charge states of 5'-d(CGGCCG)-3', summarised in figure 5.31, show the loss of C⁻ to be favoured over the loss of G⁻ for all charge state precursor ions. The preferences for charged base loss observed for 5'-d(GCCGGC)-3' (figure 5.32) vary with increasing charge state. The MS/MS spectra of the 3⁻ and 4⁻ charge states show loss of G⁻ to yield ions of greater relative intensity than those generated for loss of C⁻ whereas MS/MS of the 5⁻ ion produced more predominant ions for the loss of C⁻ in comparison to G⁻.

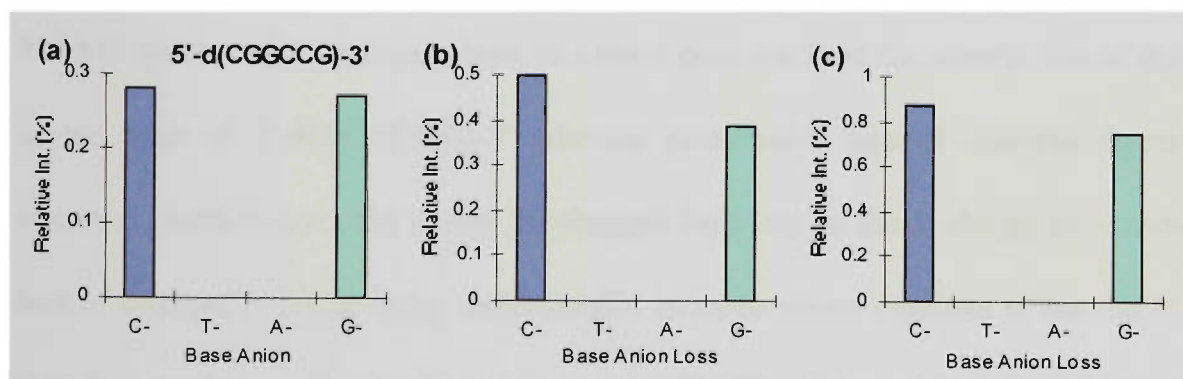


Figure 5.31: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a)[M-3H]³⁻, (b)[M-4H]⁴⁻ and (c)[M-5H]⁵⁻ ions of 5'-d(CGGCCG)-3'.

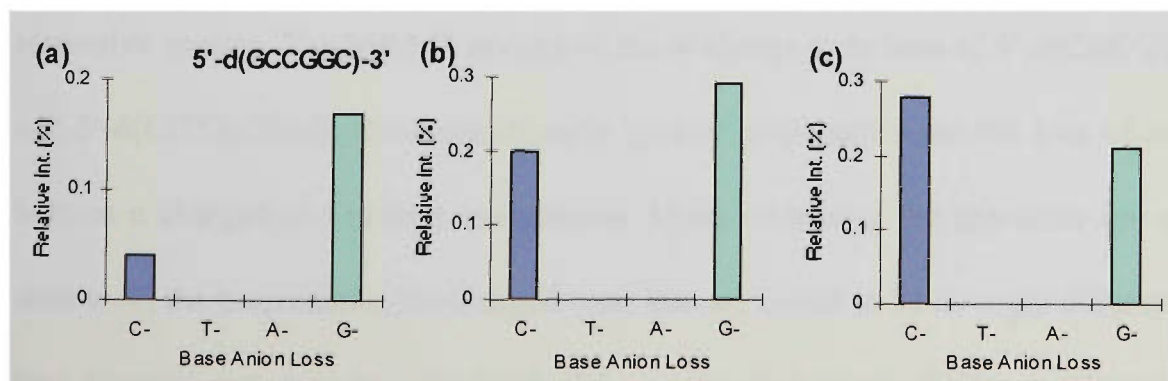


Figure 5.32: Comparison of the relative intensities of ions generated from charged base loss in the ESI-MS/MS spectra of the (a)[M-3H]³⁻, (b)[M-4H]⁴⁻ and (c)[M-5H]⁵⁻ ions of 5'-d(GCCGGC)-3'.

The propensities for charged base loss observed in the MS/MS spectra of each of the 3⁻ charge state oligonucleotide anions show preferential production of ions for the charged loss of the 5'-terminal base. The ordering for charged base loss observed in the MS/MS spectrum of the 3⁻ ion of each oligonucleotide is in most cases consistent with the relative intensities generated from the formation of base anions in each respective spectrum, and furthermore, are markedly similar to the propensities for ion formation from neutral base loss. The major difference in propensities in the case of neutral base loss compared with charged base loss for 3⁻ precursor ions arises from the unfavourable loss of neutral thymine. The corresponding ion for its loss is therefore absent in the MS/MS spectra of the 3⁻ ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' whereas loss of charged thymine is observed. Similarly, although the loss of the 5'-terminal base is seen to be favoured for pathways involving both charged and neutral base loss in the MS/MS spectra of 3⁻ precursor ions, this trend does not hold for neutral loss of thymine in the case of 5'-d(TCACGA)-3' whereas preferential loss of charged thymine is observed. Furthermore, the trends for charged base loss in the 3⁻ charge state show the loss of charged cytosine to be unfavourable in cases where cytosine is not the 5' base therefore rendering this ion absent from the MS/MS spectra of 5'-d(ATGCAT)-3' and 5'-d(TCACGA)-3' when ions for the neutral loss of cytosine are present in these respective spectra. The MS/MS spectra of the 4⁻ charge state ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' show significantly greater propensities for the loss of adenine both as a charged and neutral nucleobases. Upon increasing the precursor ion charge state to 4⁻, the propensities for charged base loss are found to be strongly influenced by base identity and sequence, in particular, enhanced loss of adenine is observed for pathways involving charged and neutral base loss for the oligonucleotides

5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. This trend is consistent with observations of the relative abundances of base anions, and is likewise reflected in the ion abundances from fragmentation owing to neutral base loss. The propensities for ions formed from charged base loss in the MS/MS of 5⁻ precursor ions showed predominant productions of ions for adenine loss for both 5'-d(CACGTG)-3' and 5'-d(TCACGA)-3'. The correlation between the propensities for charged base loss with that of base anion formation generated in MS/MS spectra of 5⁻ ions was found to be much less reliable than that generated for the 3⁻ and 4⁻ precursor ions. In each case, except for the MS/MS spectrum of the 5⁻ ion of 5'-d(TCACGA)-3', predominant formation of base anions for the 5'-terminal base were observed. The results for base anion intensities were consistent with results obtained for charged base loss in the MS/MS spectra of the oligonucleotides 5'-d(CGGCCG)-3' and 5'-d(ATGCAT)-3', and were also consistent for 5'-d(TCACGA)-3' which showed favourable loss of A⁻. Substantially different preferences for charged base loss and base anion formation were observed in the MS/MS spectra of the 5⁻ ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' which showed predominant ions for loss of A⁻ and G⁻ respectively, whereas the production of C⁻ was favoured in both MS/MS spectra. The MS/MS spectra of the 5⁻ ion of 5'-d(GCCGGC)-3' was found to generate reverse orderings for the propensities of charged base loss and base anion intensities. These inconsistencies may be a result of the effects of multiple collisions of product ions for charged base loss and the further decomposition of product ions contributing to relative intensities of base anions. This would also explain the unusual behaviour observed for charged base loss in the case of the MS/MS spectrum of the 5⁻ ion of 5'-d(CGTACG)-3'. Whereas the MS/MS spectra of the 3⁻ and 4⁻ charge state precursor ions showed increasingly favourable loss of A⁻

similar to that of 5'-d(CACGTG)-3', loss of A⁻ was found to generate the smallest contribution for charged base loss in the 5⁻ charge state. It is interesting to note, however, that the propensities for base anion formation in the MS/MS spectrum of the 5⁻ ion of 5'-d(CGTACG)-3' yielded A⁻ as the second most abundant consistent with the results of the MS/MS of the 5⁻ ion of 5'-d(CACGTG)-3' although this spectrum showed predominant ions for the loss for A⁻. The results for 5'-d(CGTACG)-3' might therefore arise from further fragmentation of ions for charged adenine loss formed from the MS/MS of fully deprotonated 5'-d(CGTACG)-3'.

5.3.2 Formation of *w* and (*a*-*B*) ions

The MS/MS spectra of all oligonucleotide anions show series of *w* and (*a*-*B*) sequence ions consistent with ribose 3' C-O bond cleavage following the loss of a base. Illustrated in figure 5.33 is a comparison of the total ion current arising from (*a*-*B*) ions expressed as a percentage of the ion current owing to *w* ions in the MS/MS spectra of the 1⁻ to 5⁻ charge states of each oligonucleotide. For the purposes of this comparison and all following comparisons of (*a*-*B*) ions, the ion for [*a*₁-*B*₁H-2H]⁻ (*m/z* 97.0), which corresponds to a sugar unit, was not included in the calculations since product ion analysis studies have shown that this ion is a commonly produced backbone fragment. Therefore it may not be representative of fragmentation generated predominantly from base loss and ribose 3' C-O bond cleavage at the 5'-terminal residue. As can be seen, the production of *w* ions in MS/MS spectra is clearly a more facile process than that of (*a*-*B*) ions for each precursor ion charge state of all oligonucleotides examined. These observations of the general lower overall abundance of (*a*-*B*) ions compared with *w* ions are consistent with the mechanism for base loss which is catalysed by hydrogen abstraction via a deprotonated phosphate group to the 3' side of the cleavage site. Based

on this mechanism, and in the absence of charge transfer to the 5' side of the cleavage site, subsequent cleavage of the ribose 3' C-O bond will generate a w ion in all cases of neutral base loss. The formation of (a-B) ions from this mechanism however, depends on the location of a charge to the 5' side of the cleavage site in addition to the charge on the phosphate to the 3' side of the cleavage site, thus making their production less facile than that of w ions.

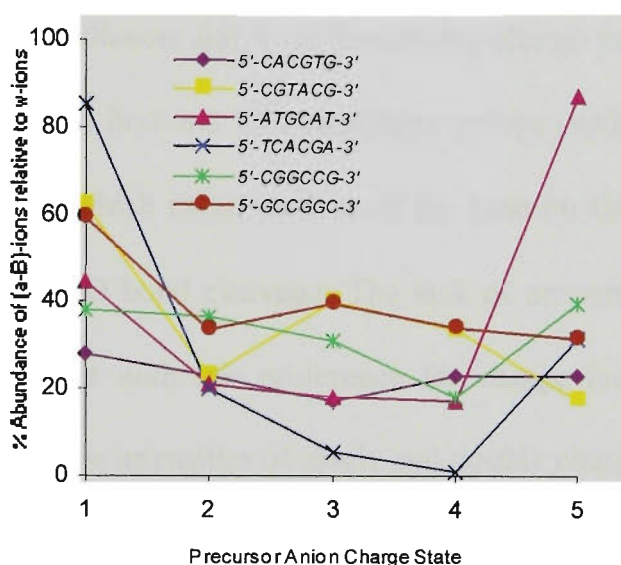


Figure 5.33: Comparison of the total abundance of (a-B) ions relative to the total abundance of w ions generated in MS/MS spectra of the 1⁺ to 5⁺ charge states of each oligonucleotide anion.

The trends observed for the formation of w and (a-B) ions in the MS/MS spectra of singly charged anions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3', 5'-d(GCCGGC)-3' has been discussed previously in chapter four, section 4.6, and those of the corresponding doubly charged precursor ions have been discussed previously in section 5.2 of this chapter. The MS/MS spectra of singly charged oligonucleotide precursor anions yield relatively intense w₃⁻ ions and [a₄-B₄H-2H]⁻ ions except when thymine is located to the 5' side of the cleavage site which results in a significant reduction in signal intensity of the cleavage product. As discussed in chapter four, these observations may arise from a

preference for location of the negative charge in the case of singly charged hexamers on the phosphate group of the third residue which could catalyse loss of the adjacent bases to the 5' and 3' side. The MS/MS spectra of doubly charged oligonucleotides anions, on the other hand, yield favourable formation of w_1^- ions, with a distinct lack of preference for formation of w ions from ribose 3' C-O bond cleavage at the third residue. As discussed previously, the favourable formation of w_1^- ions may arise from a combination of the effects of multiple collisions and a preference for charge location in the case of doubly charged anions on the first and fifth phosphate groups (which would give rise to the least charge repulsion) which catalyses loss of the base on the adjacent 5' residue and subsequent ribose 3' C-O bond cleavage. The lack of propensity for formation of w_3^- ions is further consistent with this preference for charge location. The effect of charge location on the relative intensities of singly and doubly charged w ions generated in the MS/MS spectra of doubly charged oligonucleotide anions is summarised in figure 5.34.

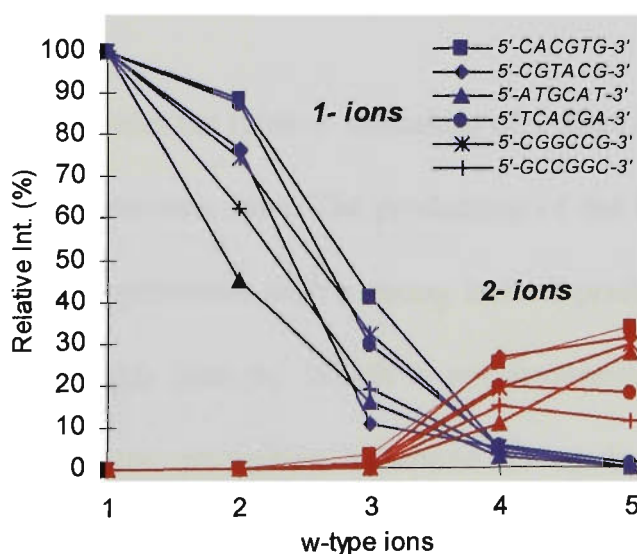


Figure 5.34: Relative intensities of 1⁻ and 2⁻ charge states of the various w ions produced in the MS/MS spectra of the $[M-2H]^{2-}$ ions of oligonucleotide anions.

As can be seen, the relative abundance of singly charged w ions decreases upon increasing size of the w fragment. The formation of w_5^- yields the smallest contribution

of all singly charged w ions which is understandable since, for a hexamer with 2 charges lying on the phosphodiester backbone, the formation of singly charged species would be expected to be less likely upon increasing size of w fragments which have a greater number of phosphates. On the other hand, doubly charged w ions are most predominant for the w_5 fragments and decrease in abundance with decreasing w ion size. These observations are consistent with a preference for location of the two charges on the phosphate groups far apart on the oligonucleotide backbone.

The formation of (a-B) ions in the MS/MS spectra of doubly charged oligonucleotide anions is not consistent with the propensities observed for the formation of complementary w ions. In the majority of cases, the $[a_3-B_3H-2H]^-$ ion is the most intense (a-B) ion observed. These results may be owing to the effects of multiple collisions of larger (a-B) ions into smaller (a-B) ions. The propensities for formation of both w and (a-B) ions in the MS/MS spectra of doubly charged oligonucleotide anions are observed to be adversely affected by the location of a thymine-containing residue at the cleavage site.

Figure 5.35 summarises the relative intensities of w ions observed in the MS/MS spectra of triply charged precursor ions. The production of the singly charged ions w_1^- and w_2^- is still found to be prevalent with a strong lack of preference for formation of singly charged w ions larger than w_3^- which is not surprising if the 3 charges are preferentially located on the phosphate backbone so as to minimise Coulombic repulsion. Doubly charged w ions are observed for w fragments from $[w_2-H]^{2-}$ to $[w_5-H]^{2-}$ with the production of $[w_5-H]^{2-}$ ions generally producing the most intense signals. These observations are consistent with the trends for charged base loss generated in the MS/MS spectra of triply charged precursor ions. In each case, MS/MS

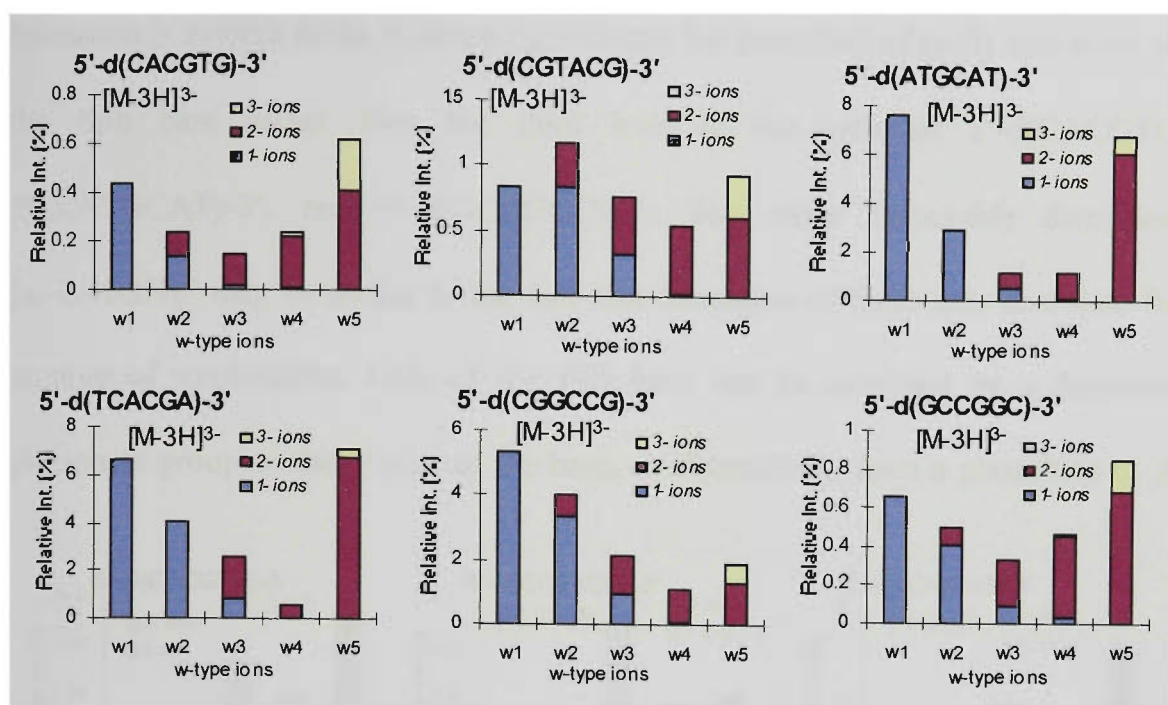


Figure 5.35: Comparison of the ion intensities of the various w ions generated in the ESI-MS/MS spectra of triply charged oligonucleotide anions.

spectra show preferable production of ions for charged base loss of the 5' terminal base which will give rise to formation of $[w_5-H]^2-$ ions following cleavage of the adjacent ribose 3' C-O bond. The relatively intense observation of ions for w_1 and w_5 cleavage is further consistent with charge location at the 5' and 3' terminal phosphate groups catalysing the loss of the first and forth bases respectively. MS/MS spectra of all triply charged oligonucleotide anions also show the presence of triply charged w ions for $[w_5-2H]^3-$. A weak contribution from $[w_4-2H]^3-$ is also detected in the MS/MS spectra of some oligonucleotides. The general lack of preference for formation of triply charged w ions smaller than w_5 is also consistent with a distribution of the charges along the backbone that yields minimal Coulombic repulsion ie. deprotonation of the first, third and fifth phosphate groups.

The relative intensities of (a-B) ions observed in the MS/MS spectra of triply charged oligonucleotide anions is shown in figure 5.36. Upon shifting charge state from doubly to triply charged precursor ions a noticeable change in propensity for (a-B) ion

formation is evident in the enhanced preference for formation of (a-B) ions from loss of the fifth base rather than the third base in the case of 5'-d(CACGTG)-3', 5'-d(ATGCAT)-3', and 5'-d(GCCGGC)-3'. The more favourable formation of $[a_5-B_5H-3H]^{2-}$ may be owing to the fact that formation of these ions can arise from a number of mechanisms. Loss of the fifth base can be catalysed by a deprotonated phosphate group to the 3' side of the base, or alternatively from a phosphate to the 5'

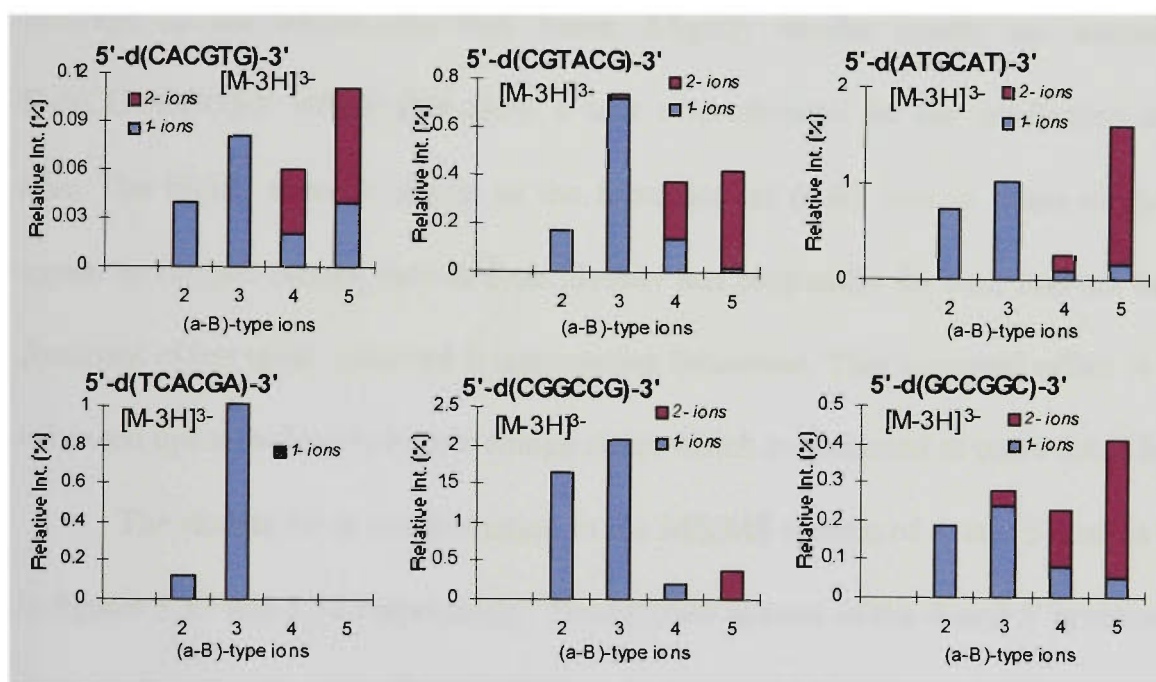


Figure 5.36: Comparison of the ion intensities of the various (a-B) ions generated in the ESI-MS/MS spectra of triply charged oligonucleotide anions.

side. Where all three charges are located to the 5' side of the cleavage site, the $[a_5-B_5H-3H]^{2-}$ ion can also be formed from initial charged base loss. The absence of ions owing to triply charged $[a_5-B_5H-4H]^{3-}$ may suggest that either the location of 3 charges to the 5' side of the fifth base is unfavourable, or indicates a lack of preference for neutral base loss upon cleavage of the fifth base where the 3 charges are located to the 5' side of the fifth base. The MS/MS spectrum of the 3^- ion of 5'-d(CGTACG)-3', in distinct contrast to that of the doubly charged precursor ion shows preferable formation of $[a_3-B_3H-2H]^-$ ions. This observation can be rationalised on the basis that MS/MS of

the 3^- ion of 5'-d(CGTACG)-3' can yield the $[a_3-B_3H-2H]^-$ ion from fragmentation involving the loss of charged thymine which is observed to be a much more facile process compared with the neutral loss of thymine. Consistent with this are the observed intensities for the ions for charged and neutral thymine loss. The observations for (a-B) ion formation in the case of 5'-d(TCACGA)-3' yield very different results compared with all other triply charged precursor ions with the absence of (a-B) ions arising from cleavage of the fourth and fifth bases. Slightly similar results are obtained for 5'-d(CGGCCG)-3' which also show a lack of preference for the production of these ions. The highly specific nature of the formation of (a-B) ions in these cases would appear to suggest factors such as base identity and propensity for base loss are having a dominant effect upon observed fragmentation behaviour. This observed effect is further enhanced upon shifting to higher charge states which is discussed in more detail below.

The results for w ion formation in the MS/MS spectra of 4^- and 5^- ions is shown in figures 5.37 and 5.38 respectively. The MS/MS spectra of the 4^- and 5^- precursor ions show w ion charge states from 1^- up to 4^- . As would be expected for hexamer ions having four and five negative charges lying along the backbone, the relative intensities of singly charged w ions strongly decreases for w ions larger than w_1 , with a lack of preference for production of singly charged w ions larger than w_3 for 4^- precursor ions, and w_2 for 5^- precursor ions. The MS/MS spectra of $[M-4H]^{4-}$ ions of oligonucleotides show the relative intensities of doubly charged w ions to vary with each different sequence although $[w_5-H]^{2-}$ ions are not detected in any case. The production of triply charged w ions occurs for w_3 , w_4 , and w_5 cleavages, and the production of quadruply charged w ions is also detected in the MS/MS spectra of some oligonucleotides for w_4 and w_5 cleavages. In the MS/MS spectra of $[M-5H]^{5-}$ ions of oligonucleotides, triply

charged w ions for fragments larger than w_1 are observed although the relative intensities of the ions from various w cleavages is strongly sequence dependent. Quadruply charged ions are observed for w_5 cleavages in the MS/MS spectra of the majority of oligonucleotides, and in some cases for w_4 cleavages.

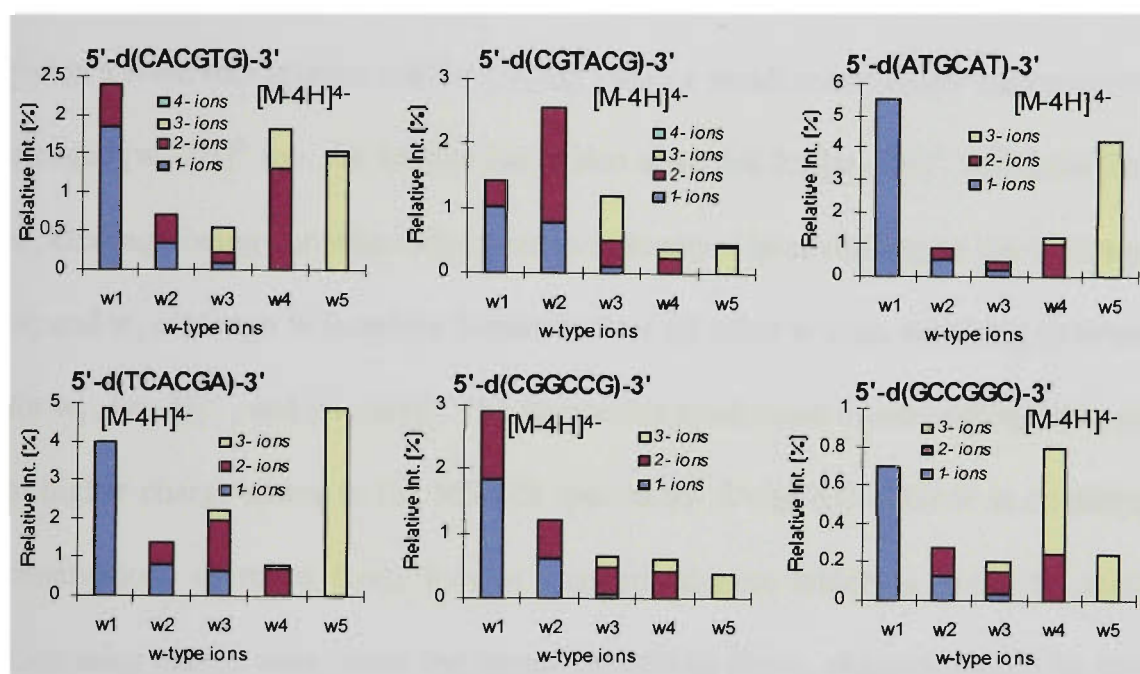


Figure 5.37: Comparison of the ion intensities of the various w ions generated in the ESI-MS/MS spectra of 4' oligonucleotide anions.

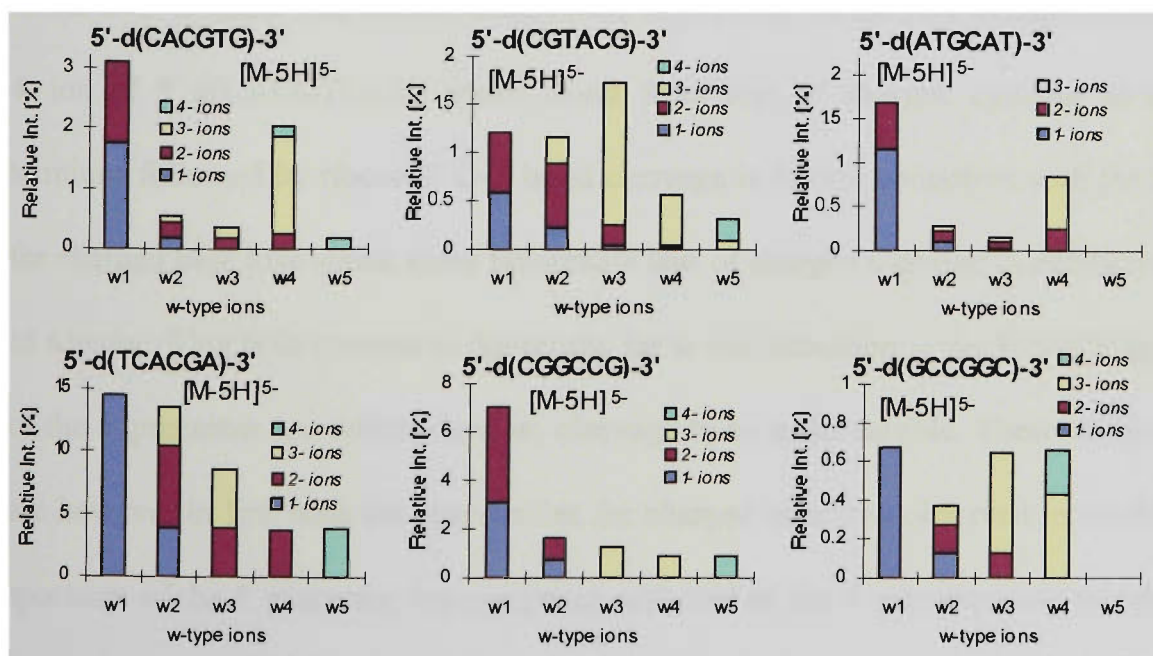


Figure 5.38: Comparison of the ion intensities of the various w ions generated in the ESI-MS/MS spectra of 5' oligonucleotide anions.

The overall trends for w ion formation in the MS/MS spectra of highly charged oligonucleotide anions yields quite different results compared with that of doubly and triply charged precursor ions, with the predominant fragmentation pathways observed varying with oligonucleotide sequence. The MS/MS spectrum of the $[M-4H]^{4-}$ ion of 5'-d(CACGTG)-3' shows w_1 cleavage to be predominant followed by w_4 cleavage which yields a relatively intense ion for $[w_4-H]^{2-}$ with a small contribution owing to the triply charged $[w_4-2H]^{3-}$ ion. An intense ion is also observed for $[w_5-2H]^{3-}$ with ions for w_2 and w_3 cleavage being considerably lower in intensity. Upon shifting to the 5^- charge state, w_1 and w_4 cleavage is found to dominate over all other w ions, resulting in intense ions for w_1^- , $[w_1-H]^{2-}$, and $[w_4-2H]^{3-}$. The favourable production of ions owing to w_4 cleavage at higher charge states in the MS/MS spectra for 5'-d(CACGTG)-3' is consistent with observations of more facile loss of charged adenine which is found to occur upon increasing charge state, since the formation of ions for w_4 cleavage would be thought to arise predominantly from the loss of adenine followed by cleavage of the adjacent ribose 3' C-O bond. The intense observation of $[w_5-2H]^{3-}$ in the MS/MS spectrum of the 4^- ion of 5'-d(CACGTG)-3', which arises from loss of charged cytosine at the 5'-terminus followed by ribose 3' C-O bond cleavage is further consistent with the results for charged base loss which show favourable loss of charged cytosine in addition to that of adenine. This is in contrast to the results for w ion formation in the MS/MS spectrum of the 5^- precursor ion which show w_5 cleavage to be unfavourable. These observations are however in line with the propensities for charged base loss observed in the MS/MS spectrum of the 5^- precursor ion compared with that of the 4^- precursor ion which show the loss of charged adenine to dominate over the loss of all other base anions. Interestingly, the intense observation of singly and doubly charged w_1 ions in the

MS/MS spectra of the 4^- and 5^- precursor ions is not reflected in the propensities for charged thymine loss, and may be a result of w_1 ion formation occurring without prior thymine loss.

The relative intensities of w ions in the MS/MS spectrum of the 4^- ion of 5'-d(CGTACG)-3' show favourable formation of singly and doubly charged ions for w_2 cleavage. Similar to the formation of w_4 ions in 5'-d(CACGTG)-3', the formation of w_2 ions in 5'-d(CGTACG)-3' from initial loss of the adjacent 5' base involves the loss of adenine. These observations are reflected in the trends for charged base loss in the MS/MS spectrum of the 4^- ion of 5'-d(CGTACG)-3' which show predominant ions for loss of charged adenine. A different propensity for w ion formation is observed in the MS/MS spectrum of the 5^- precursor ion with w_3 cleavage yielding the overall most intense signal for w ions followed by w_2 ions and w_1 ions in almost equal proportions. This change in w ion propensities upon shifting from the 4^- to 5^- charge states of 5'-d(CGTACG)-3' is also accompanied by a redistribution in the relative propensities for charged base loss with the loss of adenine no longer being favoured. The more favourable formation of w_3 ions does not produce a marked increase in the relative intensity of the ion for charged thymine loss, and once again may indicate that w_3 cleavage is not preceded by initial base loss.

The MS/MS spectrum of the 4^- ion of 5'-d(ATGCAT)-3' shows predominant w ions for w_1^- followed by $[w_5-2H]^{3-}$. The formation of both of these ions from fragmentation initiated by loss of the adjacent 5' base involves the loss of adenine, which generates the most intense ion for charged base loss observed in the MS/MS spectrum of the 4^- precursor ion. In the MS/MS spectrum of the 5^- precursor ion, no ions are observed owing to w_5 cleavage, and instead the second most predominant w

cleavage following w_1 ion formation is w_4 cleavage yielding $[w_4-H]^{2-}$ and $[w_4-2H]^{3-}$ ions. Similar trends regarding w_5 ion formation are observed upon increasing charge state for the oligonucleotide 5'-d(GCCGGC)-3'. In the MS/MS spectrum of the 4⁻ precursor ion, the production of ions from w_4 cleavage is seen to be the dominant w cleavage, with a relatively intense ion present for w_1^- . Ions owing to w_2 , w_3 and w_5 cleavage are also observed although at lower relative intensity. Upon increasing the precursor ion charge state to 5⁻, no ions for w_5 cleavage are detected. The dominant ions for w cleavage are those for w_1 , w_3 , and w_4 fragments followed by ions for w_2 cleavage.

The MS/MS spectrum of the 4⁻ charge state ion of 5'-d(TCACGA)-3' shows $[w_5-2H]^{3-}$ to be the most intense w ion generated followed by w_1^- . The formation of $[w_5-2H]^{3-}$ from fragmentation involving base loss followed by ribose 3' C-O bond cleavage of the precursor ion, arises from loss of charged thymine, which generates the most intense ion for charged base loss in the MS/MS spectrum. Although w_1^- is a relatively intense ion, no ions are observed for the loss of charged or neutral guanine and may arise from w cleavage occurring without prior base loss, or alternatively from multiple collisions of larger w ions. The relative intensities of the w ions produced upon MS/MS fragmentation of the 5⁻ ion of 5'-d(TCACGA)-3' shows a distinctly different trend to that of the 4⁻ precursor ion with a sharp decline in the propensity for ion formation from w_5 cleavage. The formation of w_1^- yields the most intense w ion in the MS/MS spectrum of the 5⁻ precursor ion followed by the production of ions from w_2 cleavage and w_3 cleavage, with formation of ions from w_4 and w_5 cleavages being the least predominant. A significant shift in the propensities for charged base loss is also observed in the MS/MS spectra of the 4⁻ and 5⁻ precursor ions which is consistent with the change in preference for ion formation from w_5 cleavage. Whereas the MS/MS

spectrum of the 4⁻ precursor ion generates intense ions for charged thymine loss, the MS/MS spectrum of the 5⁻ precursor ion shows the loss of charged adenine being favoured over that of thymine. These observations are consistent with the decreased intensity of w_5 ions formed from the loss of charged thymine.

The MS/MS spectra of the 4⁻ and 5⁻ charge states of 5'-d(CGGCCG)-3' show similar propensities regarding the production of w ions with w_1 ion formation (singly and doubly charged ions) being the most predominant w cleavage observed followed by ions for $w_2 > w_3 > w_4 > w_5$. The major difference between w -ion formation observed upon increasing precursor ion charge state above the 3⁻ charge state in MS/MS spectra of 5'-d(CGGCCG)-3' is the decrease in propensity for production of ions from w_5 cleavage.

The differences in propensities for w ion formation in the MS/MS spectra of highly charged oligonucleotide precursor ions compared with that of lower charge states may be rationalised by the variation in the relative importance of the effects of charge location and base identity in directing the formation of these ions with different charge states. In the case of doubly and triply charged hexamer anions, it would be thought that there would be a preference for charge distribution along the oligonucleotide backbone which would minimise Coulombic repulsion interactions. On this basis, the location of charges on the first and fifth phosphate groups for doubly charged ions, and the first third and fifth phosphates for triply charged ions would yield the least internal repulsion and therefore the most stable structures. The propensities for formation of w ions from base loss catalysed by a deprotonated phosphate group to the 3' side of the cleavage site would, therefore, be strongly influenced by the preference for charge location on the phosphate backbone. Upon increasing charge state to 4⁻, four of the five phosphate

groups will be charged (assuming that the charges will be located primarily on the phosphate backbone rather than the nucleobases), and there would be thought to be no significant gain in stability in terms of minimised Coulombic repulsion with different arrangements of charge location. In the 5^- charge state, the backbone is fully charged and therefore would reflect a situation whereby fragmentation would not be influenced by a preference for charge location. The results obtained for the propensities for w ion formation of highly charged oligonucleotide anions appear to be strongly influenced by oligonucleotide sequence, and in the majority of cases, the major cleavage products observed correspond with the propensities for charged base loss at the cleavage site.

The relative intensities of ions generated from (a-B) cleavage in the MS/MS spectra of the 4^- and 5^- precursor ions of all oligonucleotides is shown in figure 5.39 and 5.40 respectively. The MS/MS spectra of 4^- and 5^- oligonucleotide precursor ions show formation of singly and doubly charged (a-B) ions with triply charged (a-B) ions only observed in the MS/MS spectra of the fully deprotonated species of 5'-d(ATGCAT)-3' and 5'-d(GCCGGC)-3'. In general, the formation of multiply charged (a-B) ions is not as favourable as that observed for w ions, and only becomes feasible for oligonucleotides with highly charged backbones (ie. in the 4^- and 5^- charge states) where the probability of generating singly charged (a-B) ions of larger sizes is significantly reduced. In contrast to the observations for w ions, the formation of quadruply charged (a-B) ions were not detected in the MS/MS spectra of any oligonucleotides. These observations of less abundant production of multiply charged (a-B) ions are consistent with a mechanism for ribose 3' C-O bond cleavage whereby the driving force is the location of the charge on the phosphate group to the 3' side of the cleavage site which catalyses loss of the base on the residue to the 5' side. Following ribose 3' C-O bond

cleavage from this mechanism, at least one of the charges will be residing on the w fragment generated. Consequently the formation of doubly charged (a-B) ions from this mechanism can only occur for oligonucleotide precursor ion charge states of 3⁻ to 5⁻, and similarly the formation of triply charged (a-B) ions is only possibly for fragmentation of 4⁻ and 5⁻ oligonucleotide precursor ions.

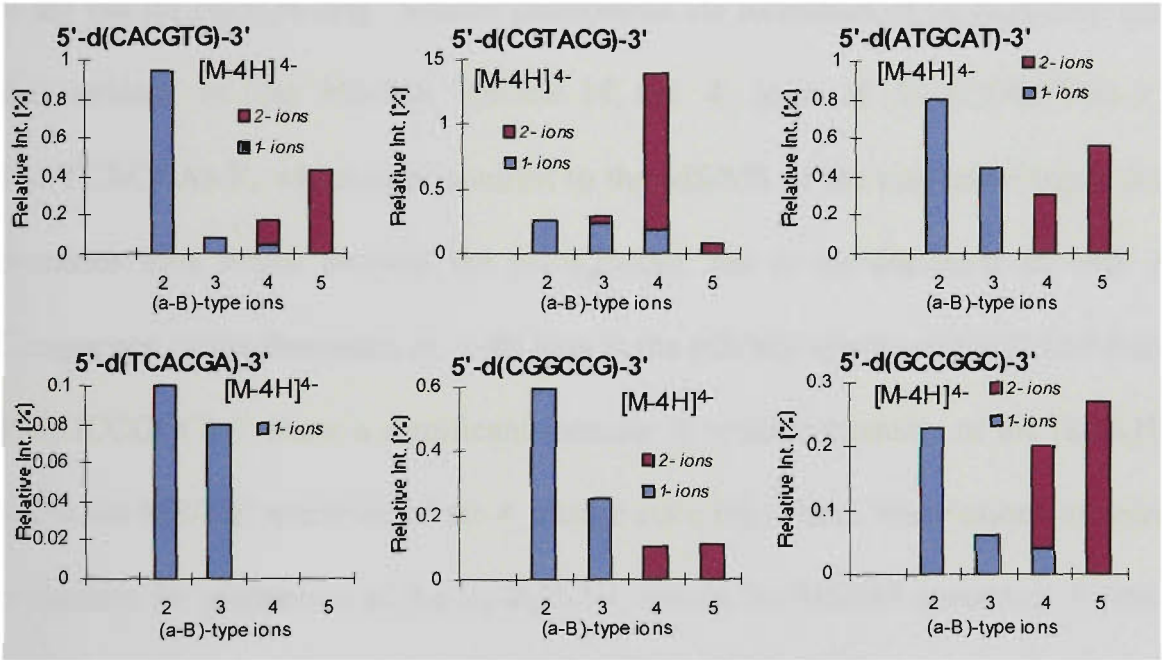


Figure 5.39: Comparison of the ion intensities of the various (a-B) ions generated in the ESI-MS/MS spectra of 4⁻ oligonucleotide anions

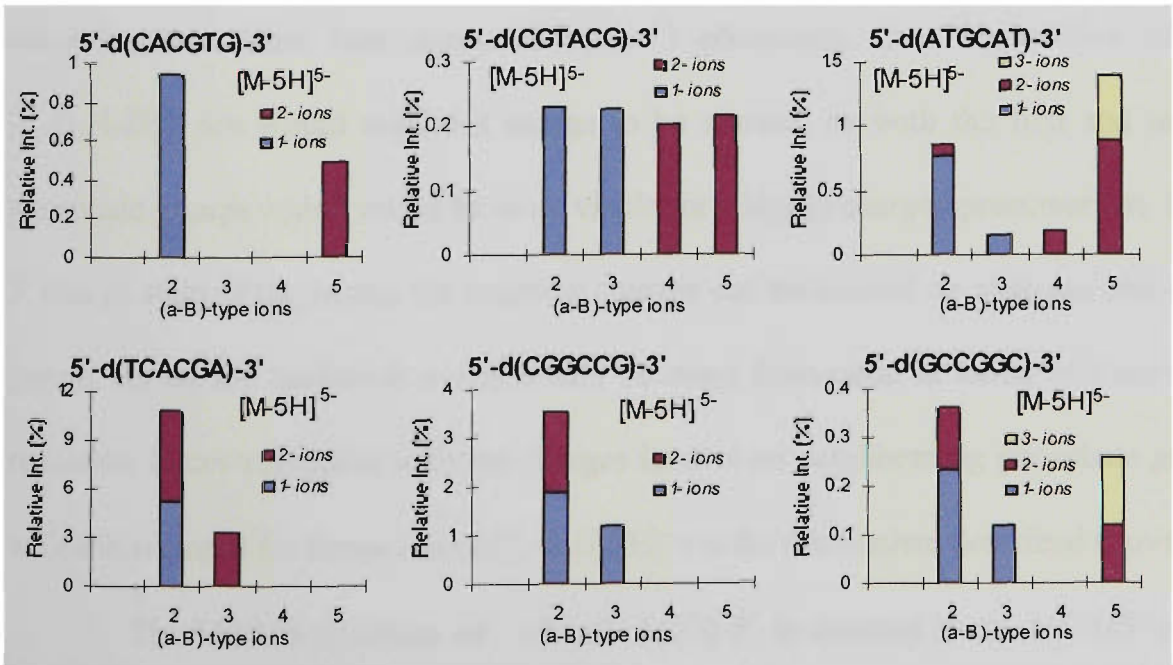


Figure 5.40: Comparison of the ion intensities of the various (a-B) ions generated in the ESI-MS/MS spectra of 5⁻ oligonucleotide anions.

The MS/MS spectra of the 4⁻ and 5⁻ charge states of all oligonucleotides show markedly different preferences for (a-B) ion intensities than that observed in the 3⁻ charge state. Whereas the MS/MS spectra of the 3⁻ ions of 5'-d(CACGTG)-3' and 5'-d(ATGCAT)-3' show predominant ions for [a₅-B₅H-3H]²⁻, the respective MS/MS spectra acquired of 4⁻ ions of these oligonucleotides show more favourable production of the ion for [a₂-B₂H-2H]⁻. Similar preferences for formation of [a₂-B₂H-2H]⁻ ions are also evident in the MS/MS spectra of the 4⁻ ions of 5'-d(CGGCCG)-3' and 5'-d(TCACGA)-3', which is in contrast to the MS/MS of the respective triply charged precursor ions which showed the [a₃-B₃H-2H]⁻ ion to be dominant in both cases. Comparison of the formation of (a-B) ions in the MS/MS spectra of the 3⁻ and 4⁻ ions of 5'-d(GCCGGC)-3' show a significant increase in relative intensity of the [a₂-B₂H-2H]⁻ ion in the MS/MS spectrum of the 4⁻ charge state ion. These observations of increased propensity for production of the [a₂-B₂H-2H]⁻ ion in the MS/MS spectra of 4⁻ precursor ions may be owing to the greater likelihood of negative charges being located on both the first and second phosphate groups. If the mechanism for (a-B) ion formation in these cases involves base loss catalysed by a 3'-phosphate, then observation of the [a₂-B₂H-2H]⁻ ion would require a charge to be situated on both the first and second phosphate groups which would be more viable for a highly charged precursor ion. In the 3⁻ charge state of hexamers, the negative charges can be located on alternate phosphate groups across the backbone which would be more favourable in terms of Coulombic repulsion forces compared with the charges located on neighbouring phosphate groups which is required for formation of [a₂-B₂H-2H]⁻ via the mechanism described above.

The MS/MS spectrum of 5'-d(CGTACG)-3', in contrast to the MS/MS spectra of the other oligonucleotides, shows the [a₄-B₄H-3H]²⁻ ion to be the most intense (a-B)

ion observed with ions for (a-B) cleavages at other positions being much lower in relative intensity. The significant relative abundance of the $[a_4-B_4H-3H]^{2-}$ ion appears to illustrate the effect of base identity. MS/MS fragmentation of the 4^- ion of 5'-d(CGTACG)-3' results in highly facile loss of charged and neutral adenine. Subsequent cleavage of the ribose C-O bond to the 3' side of adenine would result in complementary $[a_4-B_4H-3H]^{2-}$ and $[w_2-H]^{2-}$ ions for neutral base loss, or either $[a_4-B_4H-3H]^{2-}$ and w_1^- ions, or $[a_4-B_4H-2H]^-$ and $[w_2-H]^{2-}$ ions for charged base loss. Therefore, the propensities for w and (a-B) ion formation in the MS/MS spectrum of the 4^- ion of 5'-d(CGTACG)-3' which shows intense ions generated from w_2 cleavage and (a-B₄) cleavage, is consistent with fragmentation predominantly influenced by the facile loss of adenine.

The MS/MS spectrum of the 5^- ion of 5'-d(CACGTG)-3', shows highly similar propensities for formation of $[a_2-B_2H-2H]^-$ and $[a_5-B_5H-3H]^{2-}$ as observed in the MS/MS spectrum of the 4^- precursor ion, although ions for all other (a-B) ions are absent from the spectrum. The MS/MS spectrum of the 5^- ion of 5'-d(GCCGGC)-3' also shows ribose C-O bond cleavage to the 3' side of the second base yielding $[a_2-B_2H-2H]^-$ and $[a_2-B_2H-3H]^{2-}$ ions to be the most favoured, followed by the formation of the ions $[a_5-B_5H-3H]^{2-}$ and $[a_5-B_5H-4H]^{3-}$, with a small ion also present for $[a_3-B_3H-2H]^-$. The MS/MS spectra for the 5^- precursor ions of 5'-d(TCACGA)-3' and 5'-d(CGGCCG)-3' also yield the same major preference as observed in the MS/MS spectrum of the respective 4^- precursor ions of these oligonucleotides, which yield (a-B) ions predominantly from ribose C-O bond cleavage to the 3' side of the second and third bases respectively. In the case of 5'-d(CGGCCG)-3', whereas the MS/MS spectrum of the 4^- precursor ion showed all other (a-B) ions, these ions were absent in the MS/MS

spectrum of the 5⁻ precursor ion. The MS/MS spectrum of the 5⁻ ion of 5'-d(ATGCAT)-3' in comparison to that of the 4⁻ precursor ion, shows a major increase in the propensity for (a-B) ion formation from loss of the fifth base followed by ribose 3' C-O bond cleavage which gives rise to [a₅-B₅H-3H]²⁻ and [a₅-B₅H-4H]³⁻ ions. These observations may be a result of more facile loss of adenine (located at the fifth position in 5'-d(ATGCAT)-3') upon increasing the precursor ion charge state from 4⁻ to 5⁻. A strikingly different propensity for (a-B) ion formation is observed in the MS/MS spectrum of the 5⁻ ion of 5'-d(CGTACG)-3', compared with that observed in the MS/MS spectrum of the 4⁻ ion. Whereas a distinct preference for formation of [a₄-B₄H-3H]²⁻ is observed in the MS/MS spectrum of the 4⁻ charge state ion, MS/MS fragmentation of the 5⁻ precursor ion yields ions of similar intensities for all (a-B) cleavages. This observation is mirrored by the results obtained for the relative intensities of ions arising from charged base loss in the MS/MS spectrum of the 5⁻ precursor ion. In contrast to the results obtained in the MS/MS spectrum of the 4⁻ ion which show a distinct preference for loss of charged adenine, in the MS/MS spectrum of the 5⁻ precursor ion there is no significant preference for charged base loss of any particular base. Therefore, if the formation of (a-B) ions is assumed to occur predominantly via fragmentation pathways initiated by charged base loss, the lack of a preferred (a-B) cleavage can be rationalised on the basis of these observed propensities for charged base loss.

5.3.3 Summary

The charge state of the precursor ion has a significant effect on the trends observed for nucleobase loss, and the formation of w and (a-B) ions in the MS/MS spectra of oligonucleotide anions. The observed variation in the propensities for certain

fragmentation pathways with precursor ion charge state appears to reflect the extent to which there exists a preference for charge location on the backbone which can direct fragmentation behaviour analogous to the effects observed by McLuckey *et al.*¹¹. This is particularly evident when comparing the trends observed in the MS/MS spectra of doubly and triply charged oligonucleotides precursor ions with those of more highly charged species. The results obtained for the MS/MS spectra of oligonucleotide ions of lower charge states show marked similarities in the propensities for w and (a-B) formation regardless of oligonucleotide sequence. Furthermore, base location is an important factor influencing the trends for both charged and neutral nucleobase loss. The results for the fragmentation of doubly and triply charged precursor ions are, therefore, consistent with preferential charge location playing an important role in directing the major product ions types observed, since for hexamers with two and three charges, there would be expected to be a preference for charge location on the phosphodiester backbone which would yield minimum Coloumbic repulsion between charges. In the case of more highly charged species, however, there would be no overall preference for charge location in terms of minimising charge repulsion. This is reflected in the MS/MS results obtained upon shifting to higher charge states, which show the production of the various w and (a-B) ions to be highly sequence-dependent and strongly influenced by the relative propensities for base loss at the cleavage site.

REFERENCES

- (1) Rodgers, M. T.; Campbell, S.; Marzluff, M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.
- (2) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.
- (3) Bartlett, M. G.; McCloskey, J. A.; Manalili, S.; Griffey, R. H. *J. Mass Spectrom* **1996**, *31*, 1277-1283.
- (4) Habibi-Goudarzi, S.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 102-113.
- (5) Cerny, R. L.; Gross, M. L.; Grotjahn, L. *Anal. Biochem.* **1986**, *156*, 424-435.
- (6) Phillips, D. R.; McCloskey, J. A. *Int. J. Mass Spectrom. Ion Proc.* **1993**, *2*, 402-412.
- (7) McLuckey, S. A.; Vaidyanathan, G.; Habibi-Goudarzi, S. *J. Mass Spectrom* **1995**, *30*, 1222-1229.
- (8) Barry, J. P.; Vouros, P.; Van Schepdael, A.; Law, S.-J. *J. Mass Spectrom* **1995**, *30*, 993-1006.
- (9) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Chem. Soc.* **1994**, *5*, 740-747.
- (10) Iannitti, P.; Sheil, M. M.; Wickham, G. *J. Am. Chem. Soc.* **1997**, *119*, 1490-1491.
- (11) McLuckey, S. A.; Vaidyanathan, G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *162*, 1-16.

Chapter 6

**SEQUENCE ANALYSIS OF DNA ALKYLATING AGENTS BY
ELECTROSPRAY TANDEM MASS SPECTROMETRY**

6.1 INTRODUCTION

At present, the sequence selectivity of alkylating anticancer drugs, mutagens and carcinogens is commonly determined using electrophoresis-based DNA sequencing methods¹⁻⁴. While such assays are extremely sensitive and can yield a quantitative measure of the affinity of agents for particular nucleotide residues, no direct information concerning the secondary structure of ligand-DNA adducts is obtained. In contrast, methods such as NMR and X-ray crystallography can be used to derive structural information concerning such adducts, but normally do not provide information concerning sequence selectivity. As outlined in chapter one (section 1.6), mass spectrometry, on the other hand, has the potential to provide structural as well as sequence information. In particular, tandem mass spectrometry following electrospray ionisation (ESI-MS/MS) had been used previously for the location of modified bases within oligonucleotides or the analysis of oligonucleotides with methylphosphonate and phosphothiorate backbones. Most of these studies, however, were directed towards gaining an understanding of oligonucleotide fragmentation behaviour and the rationalisation of mechanisms for oligonucleotide dissociation pathways. Prior to the commencement of the work described in this thesis, there had been no reports of the application of ESI-MS and ESI-MS/MS for the sequence analysis of DNA-interactive ligands. This chapter describes the application of ESI-MS/MS for the determination of the site(s) of alkylation of three DNA-alkylating agents and the structural analysis of the resulting alkylated oligonucleotides.

6.2 DNA ALKYLATING AGENTS USED IN THIS STUDY

Three alkylating agents were used in the present sequencing study and all consist of an alkylating functional group linked to a DNA-intercalating chromophore. The ligands can be divided into two structural classes. Hedamycin and DC92-B (shown in figures 6.1(a) and (b) respectively) are natural products obtained from *Streptomyces* microorganisms and each comprise of a highly substituted anthraquinone chromophore attached to a rigid bis(epoxide) alkylating sidechain.

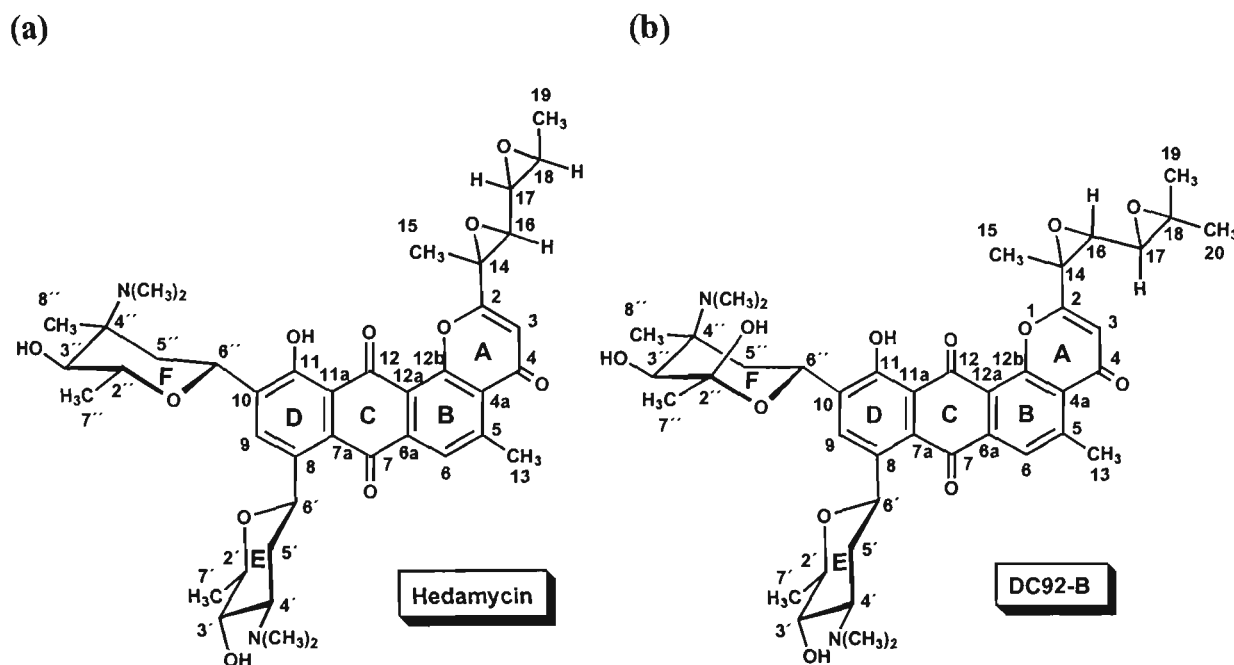


Figure 6.1: Structures of (a) hedamycin, and (b) DC92-B.

The third ligand is a synthetic derivative of phenanthridine and contains a primary alkyl bromide functional group linked by a flexible polymethylene linker chain of variable length to the nitrogen of the phenanthridine ring (shown in figure 6.2). The DNA sequence selectivity of each of these compounds has been determined previously by gel electrophoretic methods¹⁻⁴. The gel-based sequencing methods have shown that hedamycin, DC92-B and the C6-linked phenanthridinium bromide specifically alkylate guanine bases, most likely at the N-7 atom in the major groove. The highest affinity binding sites for hedamycin have been shown to be guanines in 5'-pyGT sites, with 5'-CGT sites being preferred to 5'-TGT sites.² A selectivity was found also for

alkylation of guanines in 5' G(G)_n². The sequence selectivity of DC92-B has been found to be the same as that of hedamycin². Sequencing studies of phenC₆Br have shown this compound to preferentially alkylate guanines in 5'-GT-3' sequences and in runs of guanines [(Gp)_n]¹.

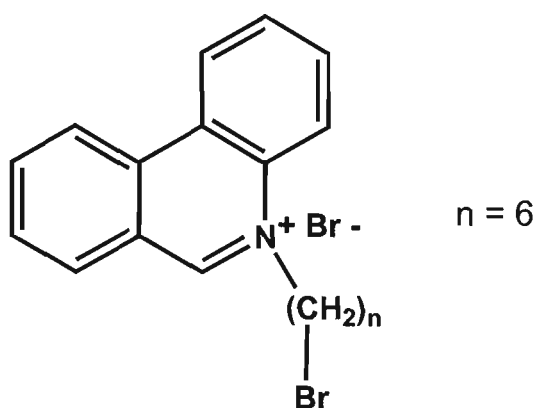


Figure 6.2: Structure of *n*-bromoalkylphenanthridinium bromide, *n*=6.

6.3 HEDAMYCIN BINDING TO DNA - RESULTS AND DISCUSSION

6.3.1 Hedamycin binding to the hexanucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(GCCGGC)-3', 5'-d(CGGCCG)-3'

The self complementary hexanucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3', were designed to incorporate the high affinity hedamycin binding sites, 5'-CGT and 5'-CGG.² In 5'-d(CACGTG)-3', a 5'-CGT site is located the middle of the sequence as well as a lower affinity 5'-TG site at the 3' terminus. In 5'-d(CGTACG)-3', the preferred 5'-CGT site is at the 5' end of the sequence and a 5'-CG site is at the 3' end.

(a) Hedamycin binding to 5'-d(CACGTG)-3'

Figure 6.3 shows the HPLC chromatogram of the 5'-d(CACGTG)-3'/hedamycin reaction mixture after reaction for 1 hour at a molar ratio of 1:1 hedamycin:duplex oligonucleotide. The HPLC chromatogram shows two main peaks, and the ESI mass spectra acquired on the VG Quattro™ mass spectrometer of a fraction of each peak is shown in figures 6.4(a) and (b) respectively.

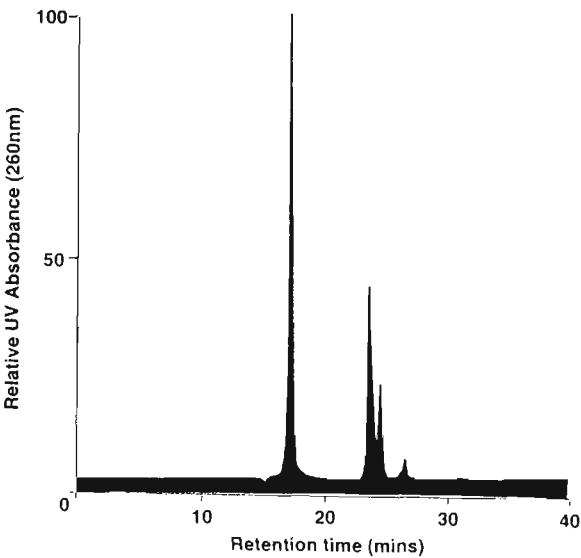


Figure 6.3: HPLC profile of the reaction between 5'-d(CACGTG)-3' and hedamycin.

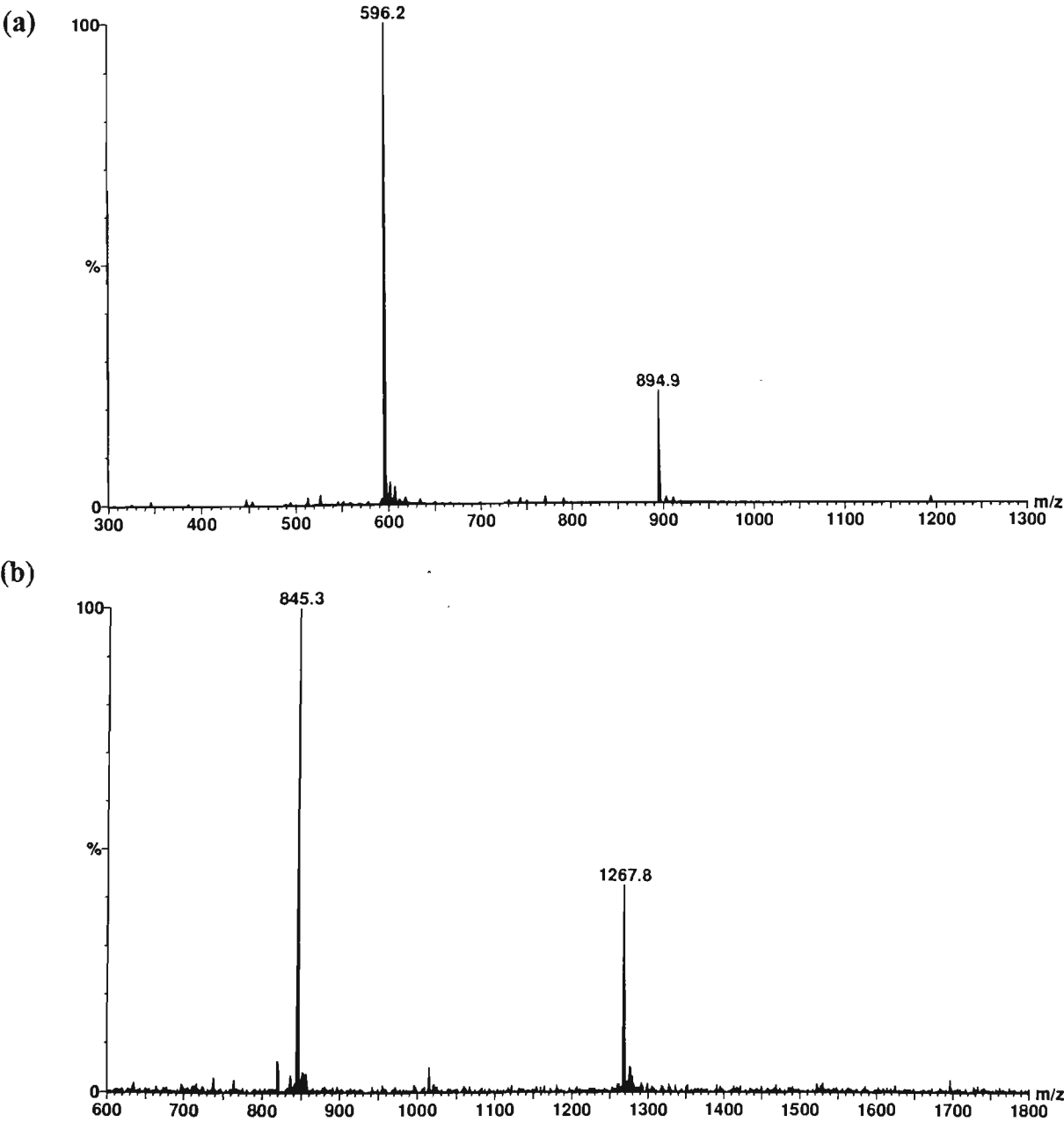


Figure 6.4: ESI mass spectrum acquired in 50% aqueous methanol of (a) the peak eluting at 17.6 minutes, and (b) the peak eluting at 23.2 minutes in the HPLC profile of the 5'-d(CACGTG)-3'/hedamycin reaction mixture.

The ESI mass spectrum of the species eluting at 17.6 minutes in the HPLC profile (figure 6.4(a)) yields peaks at m/z 894.9 and m/z 596.2 which corresponds to the $[M-2H]^{2-}$ and $[M-3H]^{3-}$ ions respectively of the free unbound oligonucleotide, $M_r(\text{meas.})=1792 \pm 1$ Da ($M_r(\text{calc.})=1791.3$ Da monoisotopic mass). Figure 6.4(b) shows the ESI mass spectrum of the species eluting at 23.2 minutes which gives rise to two peaks at 1267.8 and 845.3 owing to the $[M-2H]^{2-}$ and $[M-3H]^{3-}$ ions respectively of the single stranded hedamycin/oligonucleotide adduct, $M_r(\text{meas.})=2538 \pm 1$ Da ($M_r(\text{calc.})=2537.7$ Da monoisotopic mass). These results, which show the formation of one major hedamycin adduct, were typical of the HPLC chromatograms of all hedamycin reactions with hexanucleotides.

The MS/MS spectrum of the $[M-2H]^{2-}$ ion of the 5'-d(CACGTG)-3'/hedamycin oligonucleotide adduct acquired on the AutospecTM mass spectrometer is shown in figure 6.5. The spectrum is strikingly simple with only five major product ions observed⁵ indicative of a single major fragmentation pathway that is illustrated in scheme 6.1. Cleavage of the N-glycosidic bond of the alkylated guanine gives rise to the depurinated oligonucleotide (product ion type **II**) observed either as the doubly charged species at m/z 819.2 or the corresponding singly charged ion at m/z 1639.2. The complementary fragment for the alkylated guanine (product ion type **I**) is also detected at m/z 896.7. Following glycosidic bond cleavage at the modified site, cleavage of the ribose 3' C-O bond of the depurinated nucleotide residue yields the w_2 ion at m/z 650.0 (product ion type **IV**) and the $[a_4-B_4H-2H]^-$ ion at m/z 988.1 (product ion type **III**). In contrast, the ESI-MS/MS spectrum of the unmodified oligonucleotide (figure 6.6) under the same collision conditions is characterised by an extensive number of product ions from phosphodiester backbone cleavage, in particular loss of nucleobases followed by ribose 3' C-O bond cleavage of the residue from which the base is lost. This pathway yields a characteristic series of w -type and $(a-B)$ -type ions as indicated in the spectrum (see figure 5.3(a) and table 5.2 for product ion assignments). Unlike the fragmentation observed for the hedamycin alkylated adduct however, no any one single cleavage pathway dominates the spectrum of the free oligonucleotide.

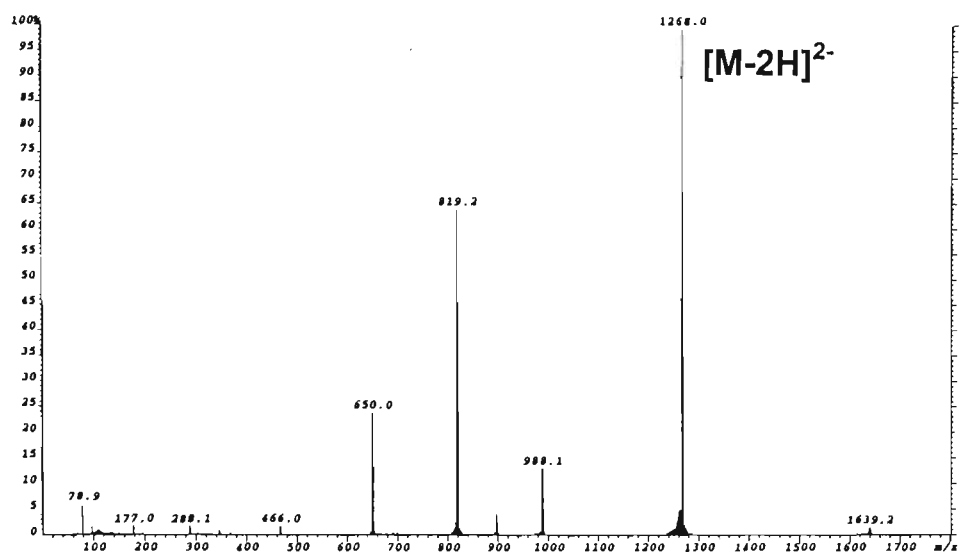
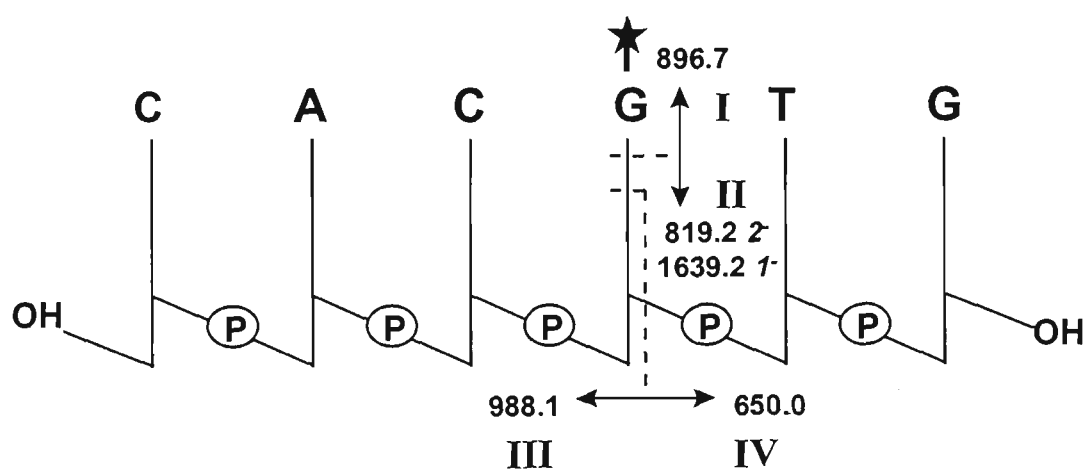


Figure 6.5: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the 5'-d(CACGTG)-3'/hedamycin adduct.



Scheme 6.1: Fragmentation pathways observed in the ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the 5'-d(CACGTG)-3'/hedamycin adduct.

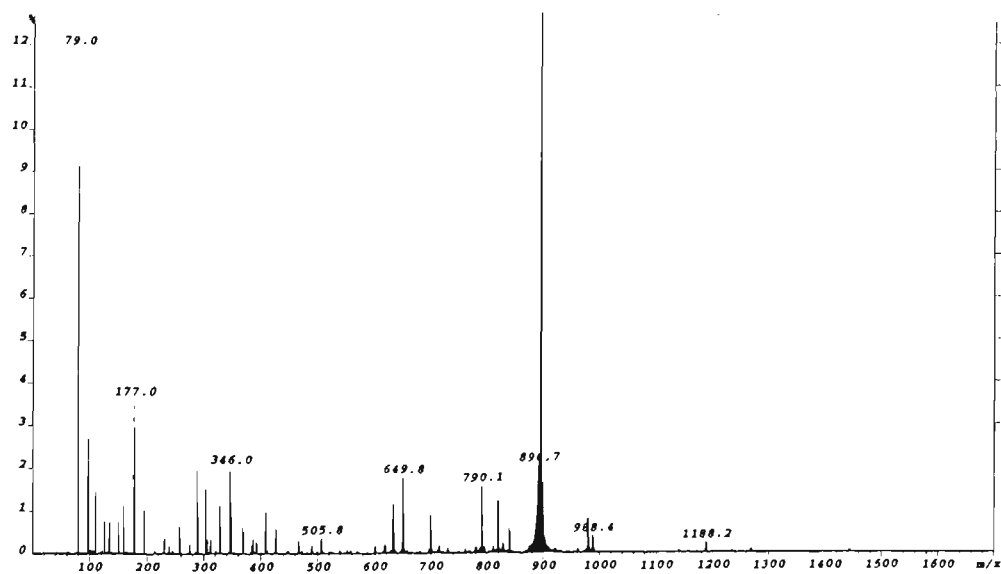
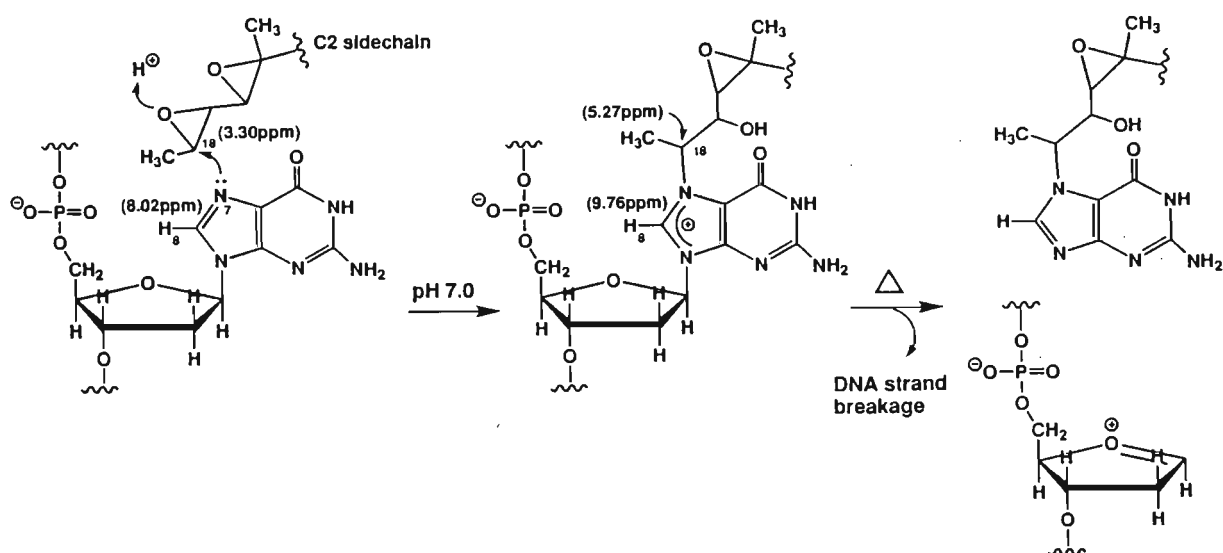


Figure 6.6: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of unmodified 5'-d(CACGTG)-3'.

The observation of a single major fragmentation pathway is consistent with labilisation of the glycosidic bond by alkylation and parallels the thermally-induced process that occurs in solution. The proposed mechanism for guanine N-7 alkylation by hedamycin inferred from NMR results⁵ is shown in scheme 6.2. Electrophilic attack of the epoxide side chain on the N-7 of guanine results in N-7 alkylation and quaternisation of the base nitrogen. This places a formal positive charge on the purine which subsequently causes weakening of the N-glycosidic bond of the alkylated residue. Therefore, when the alkylated adduct is collisionally activated, the most facile decomposition route is cleavage of the N-glycosidic bond of the alkylated base. The favourable subsequent cleavage of the ribose C-O bond to the 3' side of the residue from which the base was lost has been previously rationalised as being energetically driven by the formation of a stable furan product (see chapter one, section 1.6.2).



Scheme 6.2: Proposed mechanism for guanine N-7 alkylation by hedamycin and NMR proton resonances.

The structures of each of the four product ion types observed in the MS/MS spectrum of the $[M-2H]^{2-}$ ion of the hedamycin/5'-d(CACGTG)-3' adduct were supported by further analyses, whereby the same product ions were generated by collisional activation in the source and their MS/MS spectra analysed (in an analogous fashion to the same experiments performed on the free oligonucleotides as described in chapter four (sections 4.4 and 4.5) by multiple stage mass spectrometry to confirm their

assignments. The MS/MS spectrum of 2^- ion of the 'in-source' generated product ion type **II** appearing at m/z 819.2 in the MS/MS spectrum of the $[M-2H]^{2-}$ ion of 5'-d(CACGTG)-3'/hedamycin is shown in figure 6.7. Fragmentation of this ion is highly specific producing product ion types **III** at m/z 988.4, and **IV** at m/z 650.1. These are consistent with cleavage of the ribose 3' C-O bond following loss of the G-4 guanine. The monomeric constituents which comprise the low m/z region of the spectrum are typical of those observed for unmodified oligonucleotides as described in chapter four, section 4.4.1.

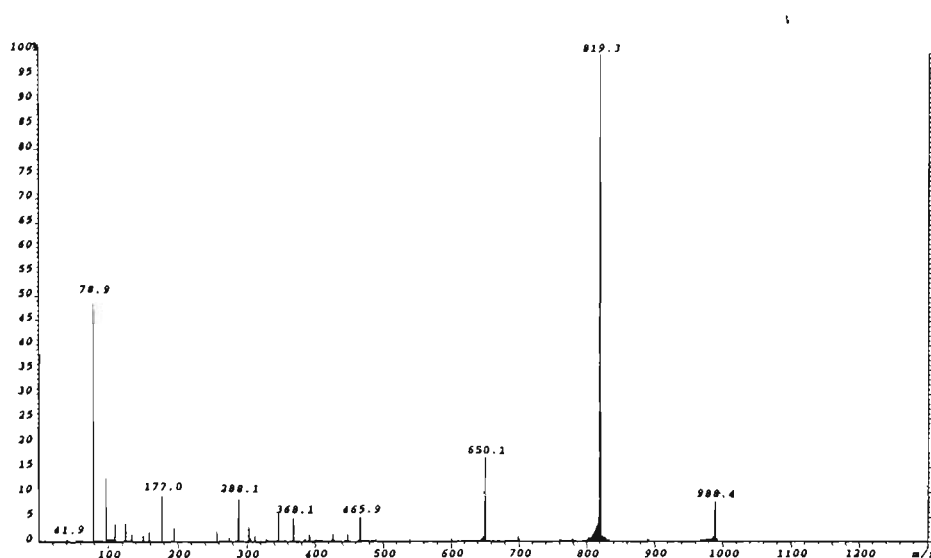
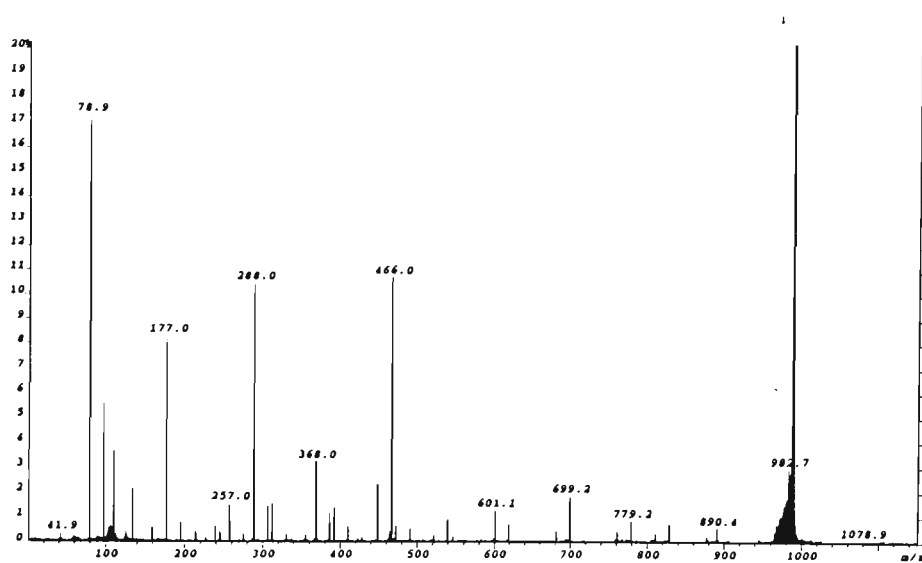


Figure 6.7: ESI-MS/MS spectrum of 2^- ion of product ion type **II** (m/z 819.3) generated by in-source fragmentation of the 5'-d(CACGTG)-3'/hedamycin adduct.

The MS/MS spectra of product ion types **III** and **IV** are shown in figures 6.8(a) and (b) respectively. Fragmentation of the ion at m/z 988.4 (**III**) yields an intense ion at m/z 466.0 for $[C_{nt}+s+p-H]^-$ and at m/z 288.0 for $[C_{nt}-H]^-$. The ion at m/z 466.0 can arise either from cleavage of the ribose 3' C-O bond to the 3' side of adenine yielding the species $[psCps-H]^-$, or alternatively this can be attributed to the 5' terminal ion, $[c_2-AH-2H]^-$. An ion is also generated at m/z 448.1 owing to the loss of H_2O from either species. The ion at m/z 288.0 can be attributed either to $[c_1-2H]^-$ or $[C_{nt}-H]^-$. The product ion corresponding to cleavage of the ribose 3' C-O bond of fragment **IV** to the 3' side of C-1 is observed at m/z 779.2. Decomposition of the m/z 988.4 product ion also yields (a-B) ions, $[a_3-B_3H-2H]^-$ (m/z 699.1) and $[a_2-B_2H-2H]^-$ (m/z 386.0), b_3^- (m/z 828.1) and

b_2^- (m/z 539.3), and $[c_3-2H]^-$ (m/z 890.4) and $[c_2-2H]^-$ (m/z 601.1). The ion at m/z 618.9 is attributed to $[C_{nt}+s+p-H_2O-H]^-$. The assignments of the various other nucleotide product ions observed in the spectrum are described in detail in chapter four, section 4.4. The MS/MS spectrum of the ion at m/z 650.2 (IV) shows predominantly fragmentation via cleavage of the ribose 3' C-O bond to the 3' side of thymine resulting in the ion at m/z 346.1 for w_1^- and the complementary ion, $[psT-H]^-$, at m/z 303.0. Loss of neutral thymine produces the nucleotide fragments $[G_{nt}+s-H]^-$ at m/z 426.1, and $[G_{nt}+s-H_2O-H]^-$ at m/z 408.0. Guanine and thymine nucleobase anions are also present at m/z 150.0 and 125.0 respectively.

(a)



(b)

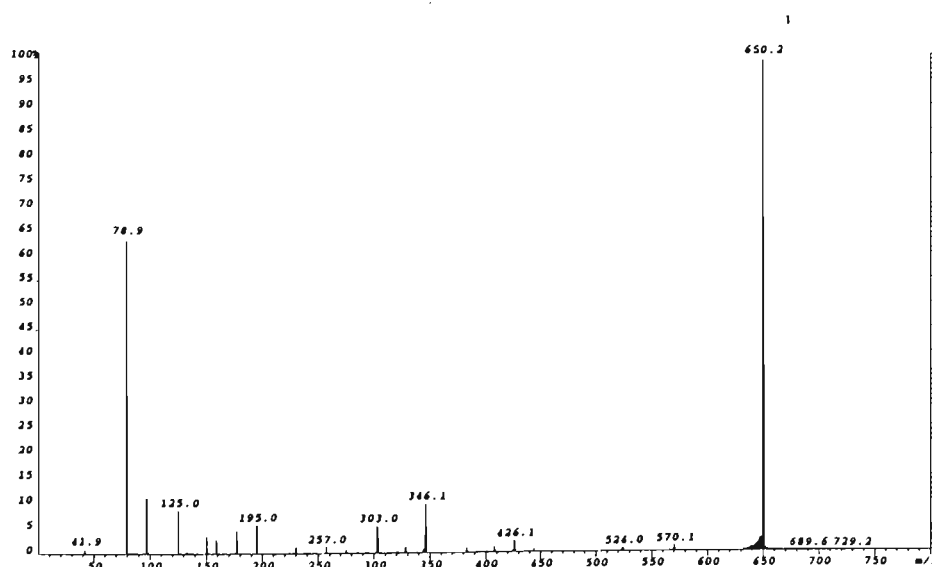


Figure 6.8: ESI-MS/MS spectrum of the product ions at (a) m/z 988.4 (III) and (b) m/z 650.2 (IV) generated by in-source fragmentation of the 5'-d(CACGTG)-3'/hedamycin adduct.

Figure 6.9 shows the MS/MS spectrum of the product ion corresponding to the [hedamycin+guanine-H]⁺ adduct (m/z 896.2) (**I**) generated by in-source collisional activation of the 5'-d(CACGTG)-3'/hedamycin adduct. Fragmentation of this ion yields an intense ion at m/z 150.0 for deprotonated guanine which indicates guanine is alkylated by hedamycin. Scheme 6.3 illustrates possible assignments of other product ions observed in the spectrum. The ion at m/z 677.2 is consistent with the mass of a

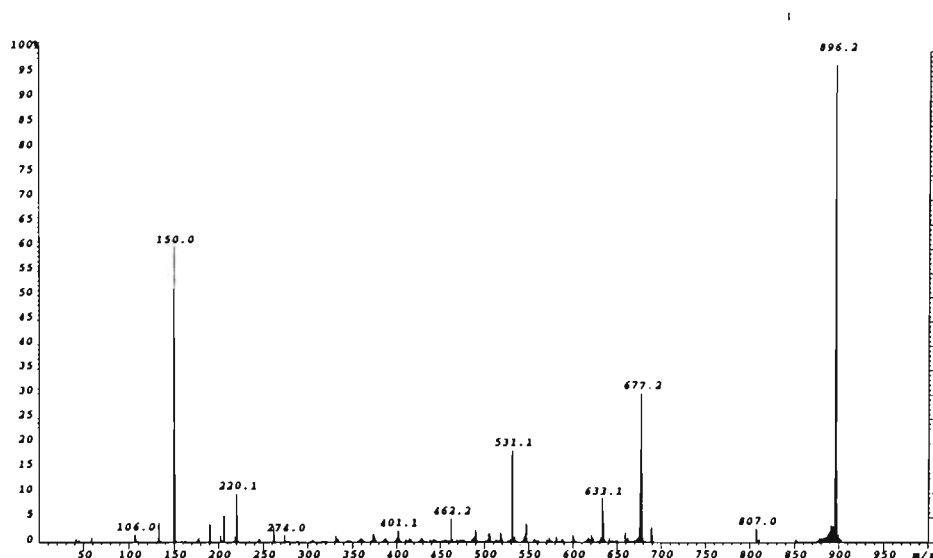
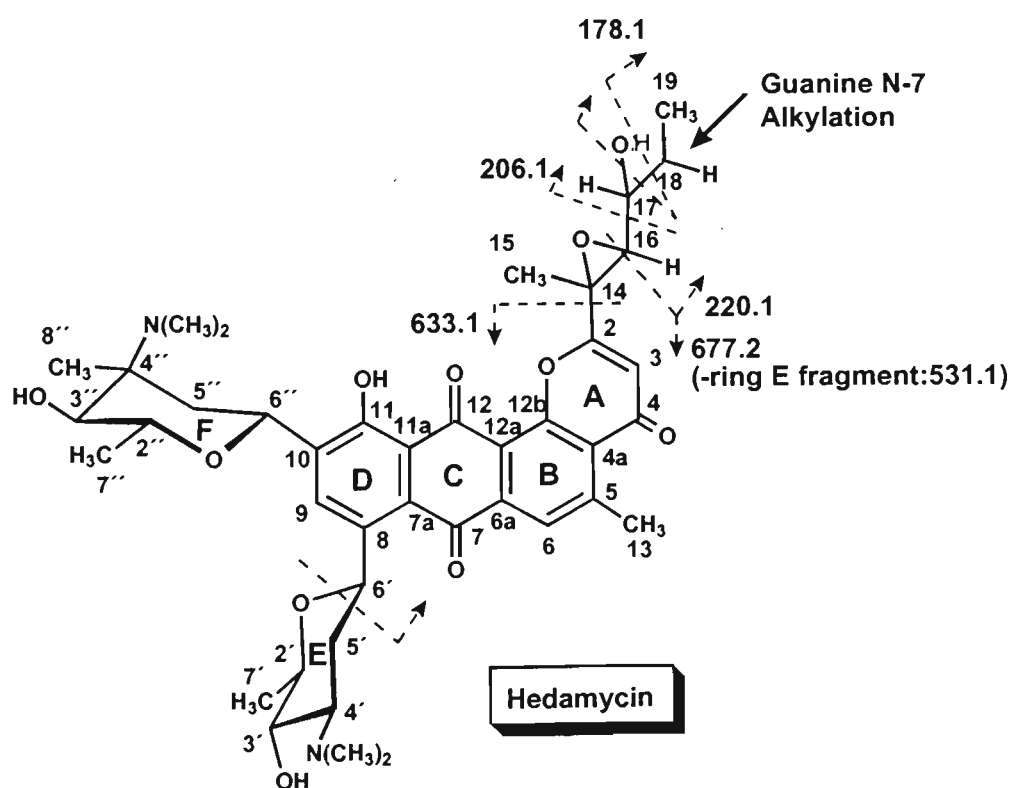


Figure 6.9: ESI-MS/MS spectrum of the product ion at m/z 896.2 (**I**) generated by in-source fragmentation of the 5'-d(CACGTG)-3'/hedamycin adduct.



Scheme 6.3: Various cleavages observed in the ESI-MS/MS spectrum of the product ion at m/z 896.2 (**I**) derived from the 5'-d(CACGTG)-3'/hedamycin adduct.

hedamycin fragment formed from cleavage of the C16-C14 and C16-O bonds of the bis-epoxide sidechain. Cleavage of the entire bis-epoxide sidechain may account for the ion at m/z 633.1. The ion at m/z 531.1 corresponds to the further loss of a residue of ring E from the species at m/z 677.2 arising from cleavage of the C6'-O and C5'-C6' bonds. The ions at m/z 178.1, m/z 206.1, m/z 220.1 and m/z 262.0 may be attributed to guanine-containing species resulting from fragmentation of the bis-epoxide at successive positions along the sidechain as indicated in scheme 6.3. The fragmentation pattern observed thus provides evidence of guanine alkylation of the bis-epoxide sidechain. It should be noted that although it is thought hedamycin alkylation of guanine occurs at the terminal carbon of the bis-epoxide side chain as indicated in the scheme, this is difficult to discern conclusively from this fragmentation pattern observed in the spectrum, since analogous product ions could be similarly generated from guanine alkylation at other positions along the bis-epoxide side chain.

(b) Hedamycin binding to 5'-d(CGTACG)-3'

The HPLC profile for the reaction between hedamycin and 5'-d(CGTACG)-3' after 1 hour is shown in figure 6.10. The chromatogram indicates the formation of one major adduct eluting at 22.5 minutes with the peak for the unbound oligonucleotide appearing at 16.9 minutes. The molecular weight of the adduct was confirmed by ESI-

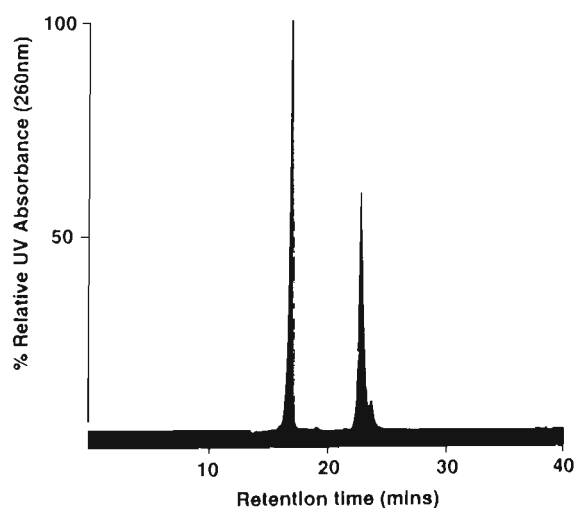


Figure 6.10: HPLC profile of the reaction between 5'-d(CGTACG)-3' and hedamycin .

MS (spectrum not shown), and corresponded to that of a single stranded 5'-d(CGTACG)-3'/hedamycin adduct ($M_{r(meas.)}=2538 \pm 1$ Da, $M_{r(calc.)}=2537.7$ Da monoisotopic mass). The ESI-MS/MS spectrum of this adduct is shown in figure 6.11. As was the case for the 5'-d(CACGTG)-3'/hedamycin adduct, a highly specific fragmentation pattern is observed. In contrast to the results obtained for the 5'-d(CACGTG)-3'/hedamycin adduct, however, only product ions owing to the loss of hedamycin+guanine, ie. product ion type **II** (m/z 819.1 and m/z 1639.3 for the 2⁻ and 1⁻ ions respectively), and the [hedamycin+guanine-H]⁻ ion, product ion type **I** (m/z 896.3) are observed. The same fragmentation is also observed upon MS/MS of the $[M-3H]^{3-}$

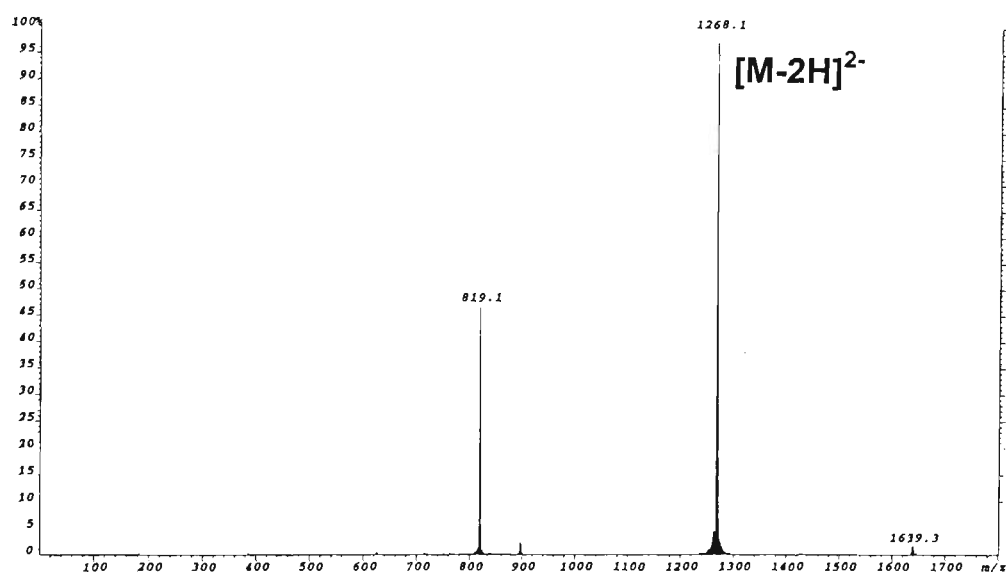
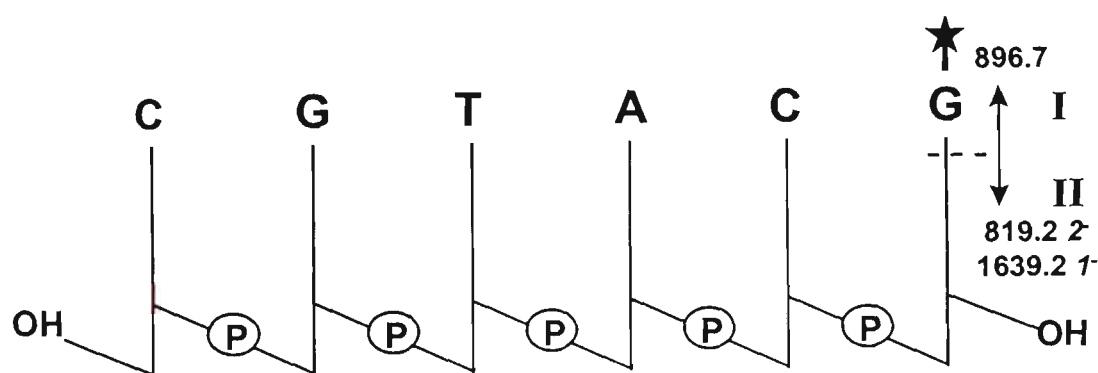


Figure 6.11: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the 5'-d(CGTACG)-3'/hedamycin adduct.



Scheme 6.4: Fragmentation pathways observed in the ESI-MS/MS spectrum of the 5'-d(CGTACG)-3'/hedamycin adduct.

precursor ion of the 5'-d(CGTACG)-3'/hedamycin adduct (figure 6.12) which only shows product ions **I** (m/z 896.7) and **II** (m/z 819.1). That there are no other fragments observed in either spectra indicative of cleavage of the ribose 3' C-O bond of the depurinated nucleotide is consistent with the site of binding of the ligand being the guanine residue at the 3' terminus as illustrated in scheme 6.4.

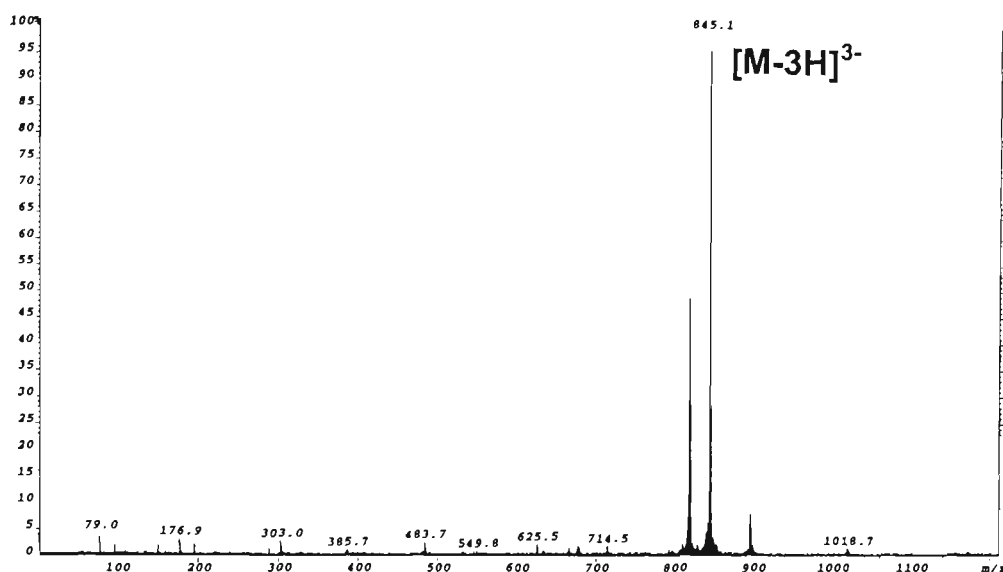


Figure 6.12: ESI-MS/MS spectrum of $[M-3H]^{3-}$ ion of the 5'-d(CGTACG)-3'/hedamycin adduct.

In order to obtain more conclusive evidence of the site of hedamycin alkylation on 5'-d(CGTACG)-3', MS/MS analysis was again conducted on the product ions observed in the ESI-MS/MS spectrum. Figure 6.13 shows the MS/MS spectrum of the ion at m/z 819.1 (**II**) generated by in-source collisional activation of the 5'-d(CGTACG)-3'/hedamycin adduct (that was observed in the spectra of both the 2' and 3' precursor ions). Unlike the MS/MS spectrum of this product ion from the MS/MS of the 5'-d(CACGTG)-3'/hedamycin adduct, fragmentation of this ion type from the 5'-d(CGTACG)-3'/hedamycin adduct does not produce a single dominant cleavage, but rather shows numerous fragments from a variety of backbone cleavages. The assignments of monomeric product ions appearing in the low m/z range (below m/z 500) have been previously given in chapter 4, section 4.4.1. Table 6.1 gives the assignments of polynucleotide product ions appearing in the MS/MS spectrum of the 2' ion of the depurinated product ion (**II**) (figure 6.13). As can be seen from the list of assignments in

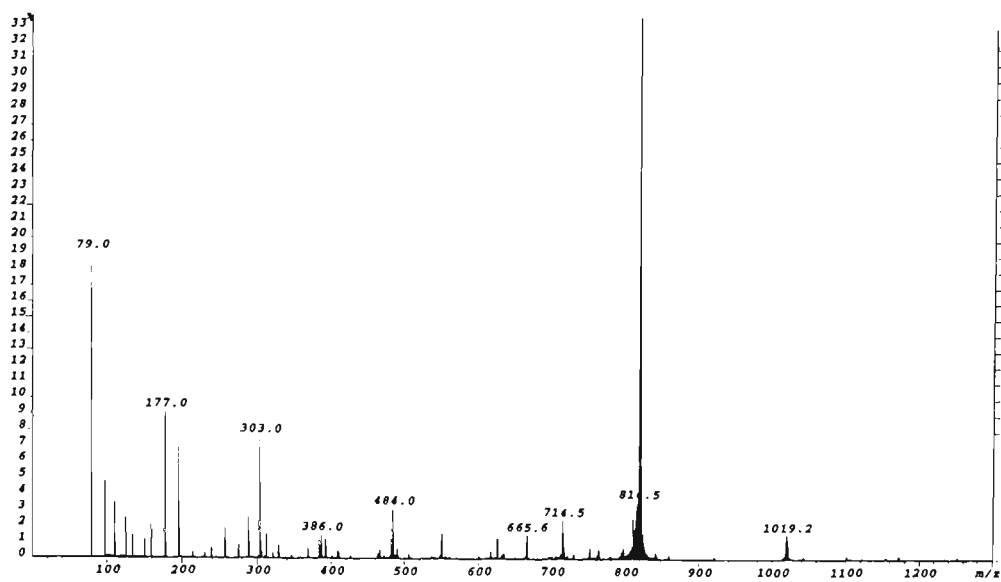


Figure 6.13: ESI-MS/MS spectrum of 2⁻ ion of product ion type II at m/z 819.1 generated by in-source collisional activation of the 5'-d(CGTACG)-3'/hedamycin adduct.

Table 6.1 : Assignment of polynucleotide product ions observed in the ESI-MS/MS spectrum of the 2⁻ ion of product ion type II (m/z 819.1) of the 5'-d(CGTACG)-3'/hedamycin adduct.

m/z	Rel. Int.(%)	m/z(calc).	Assignment
1252.3	0.2	1252.2	w ₄ ⁻ or d ₄ ⁻ or [(GTAC)+H ₂ O-H] ⁻
1172.2	0.3	1172.2	y ₄ ⁻ or b ₄ ⁻ or [(GTAC)-p+H ₂ O-H] ⁻
1101.0	0.3	1101.1	[w ₄ -GH] ⁻ or [d ₄ -GH] ⁻ or [(TAC)+s+p+H ₂ O-H] ⁻
1021.2	1.3	1021.2	[y ₄ -GH] ⁻ or [b ₄ -GH] ⁻ or [(TAC)+s+H ₂ O-H] ⁻
1019.2	1.6	1019.2	[a ₄ -B ₄ H-2H] ⁻ or [z ₄ -AH-2H] ⁻ or [(GTAC)-AH-p-H] ⁻
921.2	0.2	921.0	[c ₃ -2H] ⁻
859.2	0.3	859.2	b ₃ ⁻
841.2	0.6	841.2	[a ₃ -2H] ⁻
819.2	100	819.1	[(5'-d(CGTACG)-3')-GH-2H] ²⁻
810.2	2.7	810.1	[c ₃ -CH-2H] ⁻ or [(GT)+s+p-H] ⁻
764.3	0.7	764.1	[(5'-d(CGTACG)-3')-GH-CH-2H] ²⁻
754.3	0.8	754.6	[(5'-d(CGTACG)-3')-GH-CH-H ₂ O-2H] ²⁻
730.1	0.3	730.1	[b ₅ -H] ²⁻
		or 730.1	[(GT)+s-H] ⁻
		or 730.0	[(GT)+p+H ₂ O-H] ⁻
714.5	2.5	714.6	[w ₅ -GH-H] ²⁻ or [d ₅ -CH-H] ²⁻
665.6	1.6	665.6	[a ₅ -B ₅ H-3H] ²⁻ or [z ₅ -GH-3H] ²⁻
635.1	0.4	635.1	w ₂ ⁻ or d ₂ ⁻
625.5	1.4	625.6	[w ₄ -H] ²⁻ or [d ₄ -H] ²⁻ or [(GTAC)+H ₂ O-2H] ²⁻
617.1	0.5	617.1	[x ₂ -2H] ⁻ or [c ₂ -2H] ⁻
601.1	0.3	601.1	[(AC)-H] ⁻
549.9	1.6	550.1	[w ₄ -GH-H] ²⁻ or [d ₄ -GH-H] ²⁻ or [(TAC)+s+p+H ₂ O-2H] ²⁻
537.0	0.2	537.1	[a ₂ -2H] ⁻ or [z ₂ -2H] ⁻

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), Mr =98.0368 Da
p = PO₃H, Mr = 79.9663 Da
B_n denoted a mononucleotide which may be either psB or sBp
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

table 6.1, a significant number of ions can have more than one possible structure which complicates the task of identifying the site of depurination. For example, the ions at m/z 1019.2 and 665.6 presumably arise from (a-B)-type cleavage yielding $[a_4-B_4H-2H]^-$ and $[a_5-B_5H-3H]^{2-}$ respectively which would indicate guanine loss at the 3' terminus. As indicated in table 6.1, these ions can alternatively be 3' terminal ions produced by z-type cleavage in conjunction with nucleobase loss to give $[z_4-AH-2H]^-$ and $[z_5-GH-3H]^{2-}$ respectively which in both cases can be derived from a precursor having lost the G-2 guanine. The ion observed at m/z 625.5 can be attributed to any one of $[w_4-H]^{2-}$, $[d_4-H]^{2-}$ or $[(GTAC)+H_2O-2H]^{2-}$. The product ions at m/z 549.9 and 714.5 can correspond to the series of w-type ions, $[w_2-GH]^-$, $[w_4-GH-H]^{2-}$ and $[w_5-GH-H]^{2-}$ from which the 3'-terminal guanine has been lost, or alternatively can be assigned as $[d_2-GH]^-$, $[d_4-GH-H]^{2-}$ and $[w_5-(G-2)H-H]^{2-}$ respectively. Evidence of depurination of the 3' terminal guanine can be found from the presence of the low abundance ions at m/z 841.2, 859.2 and 921.2 which correspond to $[a_3-2H]^-$, b_3^- and $[c_3-2H]^-$ successively. These ions, which incorporate the (CGT) residue, cannot be derived from anywhere else in the oligonucleotide and thus provide evidence that the G-2 position is still attached (and therefore not alkylated). Further, evidence that the site of depurination is not the G-2 guanine can be derived from the presence of the ion at m/z 810.2 owing to either $[c_3-CH-2H]^-$ or $[(GT)+s+p-H]^-$, and the ion at m/z 730.1 which has three possible assignments all of which incorporate the G-2 guanine (see table 6.1).

To probe further the fragment ions produced from the decomposition of product ion type **II**, the MS/MS spectrum of the singly charged product ion (m/z 1639.5) was also obtained and is shown in figure 6.14. The assignments of product ions greater than mononucleotide size are given in table 6.2. The spectrum shows the same major product ion types as observed in the MS/MS spectrum of the doubly charged precursor albeit all product ions are now singly charged. As observed in the MS/MS spectrum of the 2^- ion of product ion **II** (figure 6.13), ions are present which incorporate the trinucleotide residue (5'-CGT) at m/z 841.2, 859.3 and 921.3 for $[a_3-2H]^-$, b_3^- and

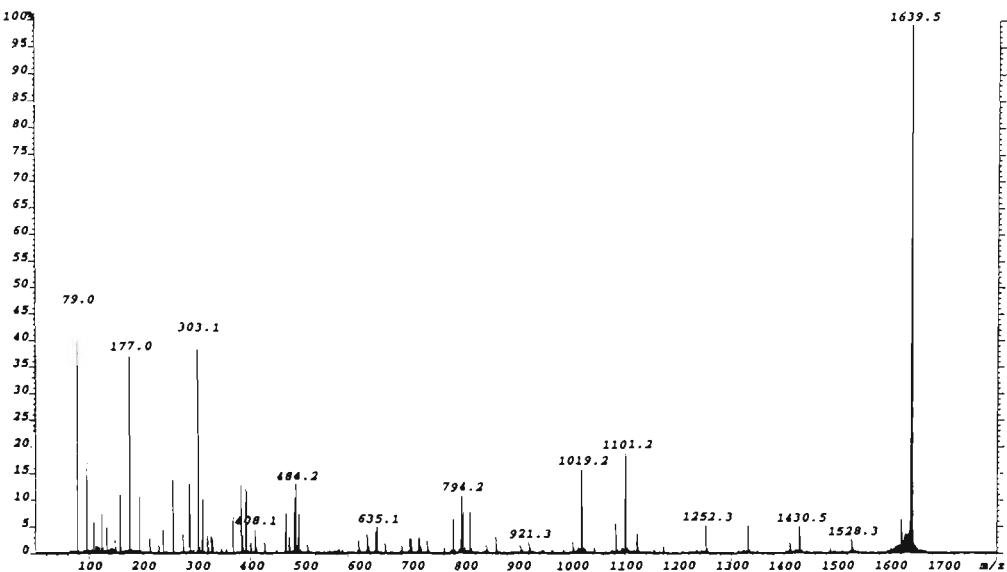


Figure 6.14: ESI-MS/MS spectrum of 1⁻ ion of product ion type II at m/z 1639.5 generated by in-source collisional activation of the 5'-d(CGTACG)-3'/hedamycin adduct.

Table 6.2: Assignment of polynucleotide product ions observed in the ESI-MS/MS spectrum of the 1⁻ ion of product ion type II (m/z 1639.5) generated by in-source collisional activation of the 5'-d(CGTACG)-3'/hedamycin adduct.

m/z	Rel. Int.(%)	m/z(calc).	Assignment
1639.5	100	1639.3	[(5'-d(CGTACG)-3')-GH-H] ⁻
1621.5	6.9	1621.3	loss of H ₂ O
1528.3	3.0	1528.3	loss of Cytosine
1488.1	1.1	1488.3	loss of Guanine
1430.5	5.6	1430.2	[w ₅ -GH] ⁻ or [d ₅ -CH] ⁻
1412.6	2.0	1412.2	[x ₅ -GH-2H] ⁻
1332.3	5.6	1332.3	[a ₅ -B ₅ H-2H] ⁻ or [z ₅ -GH-2H] ⁻
		1332.2	[(GTAC)+p+H ₂ O-H] ⁻
1252.3	5.7	1252.2	w ₄ ⁻ or d ₄ ⁻ or [(GTAC)+H ₂ O-H] ⁻
1172.4	1.5	1172.2	y ₄ ⁻ or b ₄ ⁻ or [(GTAC)-p+H ₂ O-H] ⁻
1154.2	0.9	1154.2	[a ₄ -2H] ⁻ or [z ₄ -2H] ⁻ or [(GTAC)-p-H] ⁻
1123.1	4.0	1123.1	[x ₄ -CH-2H] ⁻ or [c ₄ -CH-2H] ⁻ or [(GTA)+s+p-H] ⁻
1101.2	19.4	1101.1	[w ₄ -GH] ⁻ or [d ₄ -GH] ⁻ or [(TAC)+s+p+H ₂ O-H] ⁻
1083.1	5.8	1083.1	[x ₄ -GH-2H] ⁻ or [c ₄ -GH-2H] ⁻ or [(TAC)+s+p-H] ⁻
1043.3	1.5	1043.2	[a ₄ -CH-2H] ⁻ or [z ₄ -CH-2H] ⁻ or [(GTA)+s-H] ⁻
		1043.1	[(GTA)+p+H ₂ O-H] ⁻
1019.2	16.2	1019.2	[a ₄ -B ₄ H-2H] ⁻ or [z ₄ -AH-2H] ⁻ or [(GTAC)-AH-p-H] ⁻
1003.3	2.2	1003.2	[(TAC)+s-H] ⁻ or [z ₄ -GH-2H] ⁻
		1003.1	[(TAC)+p+H ₂ O-H] ⁻
984.7	0.9	985.2	[(TAC)+s-H ₂ O-H] ⁻
		985.1	[(TAC)+p-H] ⁻
963.3	0.9	963.1	[(GTA)+H ₂ O-H] ⁻
944.6	0.9	945.1	[(GTA)-H] ⁻
921.3	2.4	921.0	[c ₃ -2H] ⁻
905.1	1.6	905.1	[(TAC)-H] ⁻
859.3	3.3	859.2	b ₃ ⁻
841.3	1.8	841.2	[a ₃ -2H] ⁻
810.2	8.1	810.1	[c ₃ -CH-2H] ⁻ or [(GT)+s+p-H] ⁻
797.2	7.3	797.1	[(AC)+s+p+H ₂ O-H] ⁻ or [w ₃ -GH] ⁻
794.2	11.3	794.1	[(TA)+s+p-H] ⁻

Table 6.2 cont/...

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
779.1	6.9	779.1	[(AC)+s+p-H] ⁻ or [x ₃ -GH-2H] ⁻
776.1	2.0	776.1	[(TA)+s+p-H ₂ O-H] ⁻
761.4	1.2	761.1	[(AC)+s+p-H ₂ O-H] ⁻
730.1	2.5	730.1	[(GT)+s-H] ⁻
		730.0	[(GT)+p+H ₂ O-H] ⁻
715.0	1.6	715.0	[a ₃ -B ₃ H-2H] ⁻ or [z ₃ -AH-2H] ⁻
714.3	3.2	714.1	[(TA)+s-H] ⁻
		714.1	[(TA)+p+H ₂ O-H] ⁻
712.1	2.4	712.1	[(GT)+p-H] ⁻
		712.1	[(GT)+s-H ₂ O-H] ⁻
699.3	3.0	699.1	[(AC)+p+H ₂ O-H] ⁻
		699.1	[(AC)+s-H] ⁻ or [z ₃ -GH-H] ⁻
696.3	3.2	696.1	[(TA)+s-H ₂ O-H] ⁻
		696.1	[(TA)+p-H] ⁻
681.2	1.6	681.1	[(AC)+s-H ₂ O-H] ⁻
		681.1	[(AC)+p-H] ⁻
650.3	1.8	650.1	[(GT)+H ₂ O-H] ⁻
635.1	5.3	635.1	w ₂ ⁻ or d ₂ ⁻
632.1	4.1	632.1	[(GT)-H] ⁻
617.3	1.5	617.1	[x ₂ -2H] ⁻ or [c ₂ -2H] ⁻
616.2	3.9	616.1	[(TA)-H] ⁻
601.1	3.2	601.1	[(AC)-H] ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), *Mr* = 98.0368 Da

p = PO₃H, *Mr* = 79.9663 Da

B_n denoted a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

[c₃-2H]⁻ respectively. In addition, ions of very low abundance are also observed at *m/z* 963.3 and 944.6 which correspond to the trinucleotide ions [(GTA)+H₂O-H]⁻ and [(GTA)-H]⁻ respectively. Table 6.2 also shows the presence of a series of ions incorporating the (GT) dinucleotide residue at *m/z* 650.1, 712.1 and 730.1. These fragmentation patterns which indicate the G-2 guanine is still attached to the oligonucleotide are consistent with the site of guanine depurination, and thus alkylation, being that of the 3' terminal guanine.

The MS/MS spectrum of the singly charged ion of the hedamycin+guanine adduct, product ion type I at *m/z* 896.4 (figure 6.15) shows an identical fragmentation pattern to that observed for the same product ion in the MS/MS spectrum of 5'-d(CACGTG)-3'/hedamycin. The assignments of the ions observed in this spectrum are illustrated in scheme 6.3. Similar to the spectrum of this ion type given in figure 6.9, the ion for G⁻ at *m/z* 150.0 is once again found to be the most intense signal.

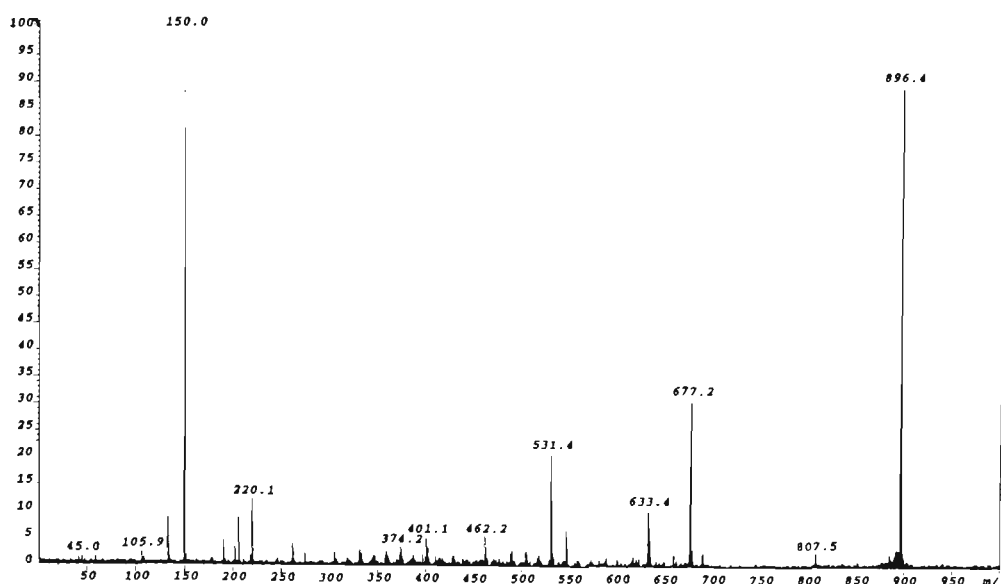
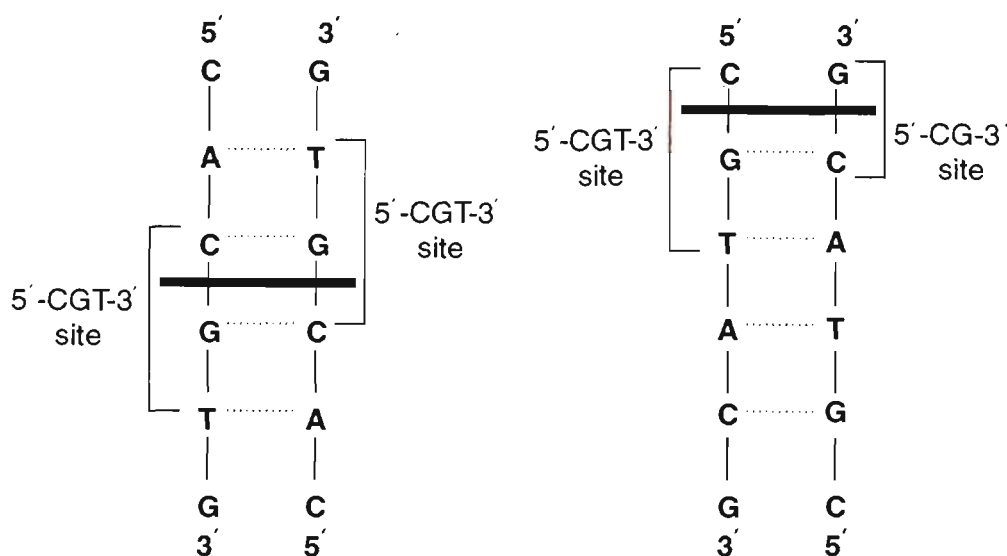


Figure 6.15: ESI-MS/MS spectrum of product ion type I at m/z 896.4 generated by in-source collisional activation of the 5'-d(CGTACG)-3'/hedamycin adduct.

(c) **Hedamycin binding to 5'-d(CACGTG)-3' versus 5'-d(CGTACG)-3':**
Discussion

Scheme 6.5 illustrates the differences in hedamycin binding between the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. In solution under the conditions used for the hedamycin binding reaction, the oligonucleotide would be expected to be predominantly in double helical form. Hedamycin binding to double helical DNA is thought to occur by initial rapid intercalation via the planar



Scheme 6.5: Comparison of hedamycin binding to the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

anthraquinone chromophore threading the DNA double helix and stacking between adjacent 5'-CG base pairs (shown by the bold line). This places the bulky amino sugar groups in the minor groove and the alkylating bis-epoxide in the major groove. In the case of 5'-d(CACGTG)-3', the location of the high affinity 5'-CGT site centrally yields two identical binding sites so that hedamycin alkylation of the guanine of either strand essentially produces the same product. This is confirmed by the results obtained for the 5'-d(CACGTG)-3'/hedamycin adduct which shows formation of a single product owing to alkylation solely of guanine at the 5'-d(CACGTG)-3' site. The fact that no adduct was formed owing to binding at the 5'-TG further supports both NMR⁵ and sequencing studies² showing the hedamycin selectivity for 5'-CG over 5'-TG. In contrast, placement of the high affinity 5'-CGT hedamycin-binding site at the end of the oligonucleotide, as in 5'-d(CGTACG)-3' yields two different possible binding sites following hedamycin intercalation between adjacent 5'-CG base pairs ie. alkylation of the guanine at the 5'-d(CGTACG)-3' site or the 5'-d(CGTACG)-3' site. The preference for alkylation by hedamycin of the 3' terminal guanine can be explained on the basis of "fraying" or unwinding of the terminal base pairs from insufficient helical constraints at the end of the duplex. Hedamycin alkylation of the 3' terminal guanine results in stabilisation of these base-pairing interactions owing to the resulting hydrophobic interactions created by the placement of two of the substituent groups of the ligand (ie. the bis-epoxide side chain and an amino sugar group) in proximity to either side of the terminal base-pair. The observed preference for guanine alkylation at the 3' terminal position in the case of 5'-d(CGTACG)-3' is also consistent with NMR data obtained for the 5'-d(CGTACG)-3' hedamycin adduct.⁵

(d) Hedamycin binding to 5'-d(GCCGGC)-3'

In order to investigate hedamycin binding to a 5'-CGG site and compare the sequence selectivity of binding to this site with that of the 5'-CGT site, reactions between hedamycin and the self complementary oligonucleotides 5'-d(GCCGGC)-3' and 5'-d(CGGCCG)-3' were also carried out. These oligonucleotides incorporate the

5'-CGG site analogous to the 5'-CGT site in the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' with the binding site located centrally for 5'-d(GCCGGC)-3' and at the 5'-terminus for 5'-d(CGGCCG)-3'. In addition, 5'-d(CGGCCG)-3' contains a 5'-CG binding site at the 3' terminus similar to 5'-d(CGTACG)-3'.

Figure 6.16 shows the HPLC profile for the reaction between hedamycin and the oligonucleotide 5'-d(GCCGGC)-3' after 1 hour, showing the formation of a major adduct at a retention time of 23.2 minutes with the peak owing to the unbound oligonucleotide at 17.4 minutes. The ESI-MS spectrum (not shown) of this adduct yielded a molecular weight of $M_{r(meas.)}=2539 \pm 1$ Da corresponding to a single-stranded 5'-d(GCCGGC)-3'/hedamycin adduct ($M_{r(calc.)}=2538.7$ Da monoisotopic mass).

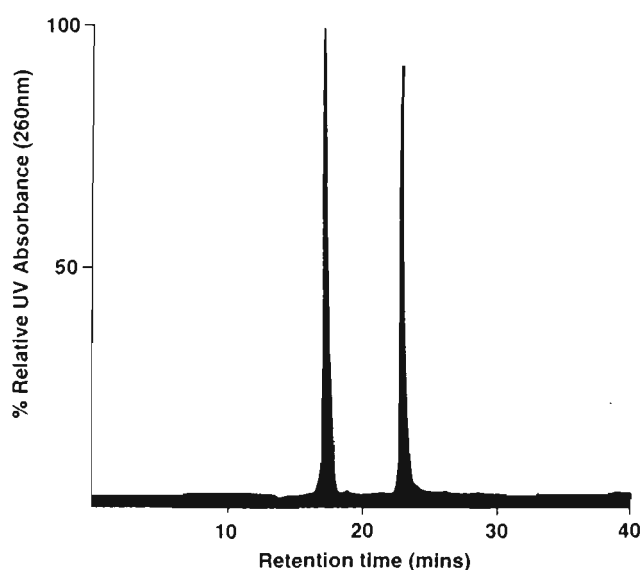


Figure 6.16: HPLC profile of the reaction between 5'-d(GCCGGC)-3' and hedamycin.

The ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of this adduct (m/z 1268.9) is shown in figure 6.17, with the specific fragmentation pathway observed represented in scheme 6.6. Cleavage of the N-glycosidic bond of the alkylated guanine yields the $[\text{hedamycin}+\text{guanine-H}]^{-}$ ion (**I**) at m/z 896.5 and the 1⁻ and 2⁻ ions of the depurinated oligonucleotide (**II**) at m/z 1640.3 and m/z 819.8, respectively. Subsequent cleavage of the ribose 3' C-O bond of the depurinated oligonucleotide results in the ions at m/z 1004.1 (**III**), and at m/z 635.1 (**IV**). The presence of these ions enables the site of alkylation on the oligonucleotide to be determined readily as the guanine at the 5'-d(GCCGGC)-3' site as illustrated in scheme 6.6. This specific fragmentation pathway

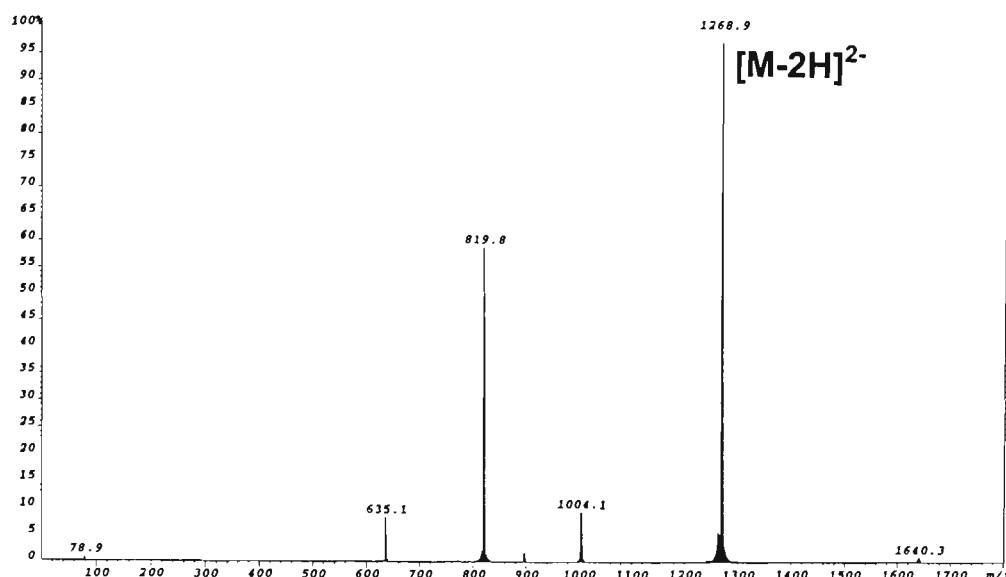
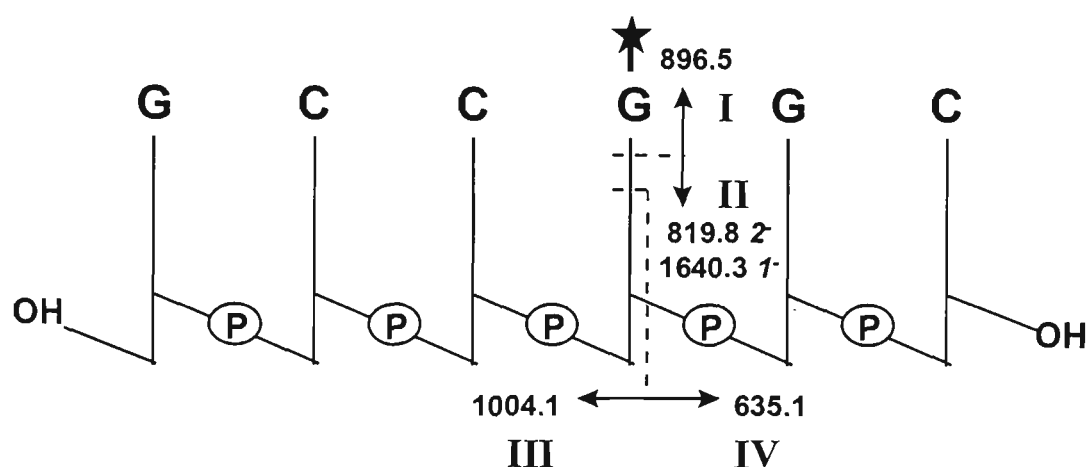


Figure 6.17: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the 5'-d(GCCGGC)-3'/hedamycin adduct.



Scheme 6.6: Fragmentation pathway generated in the ESI-MS/MS spectrum of the 5'-d(GCCGGC)-3'/hedamycin adduct.

is also evident from the MS/MS fragmentation of the 2^- and 1^- ions of the depurinated oligonucleotide product ion (II) which are shown in figures 6.18(a) and (b) respectively. In each case, fragmentation of the precursor ion occurs via facile cleavage of the ribose 3' C-O bond at the site of depurination producing the ions at m/z 1004.4 (III) and m/z 635.1 (IV). Further decomposition of the hedamycin+guanine adduct, product ion type I (not shown) yielded the G^- ion at m/z 150.0 as the most abundant product ion with the fragmentation pattern the same as that observed for this product ion when derived from

either of the 5'-d(CACGTG)-3'/hedamycin (figure 6.9) or 5'-d(CGTACG)-3'/hedamycin adducts (figure 6.15). The origin of the major product ions is illustrated in scheme 6.2.

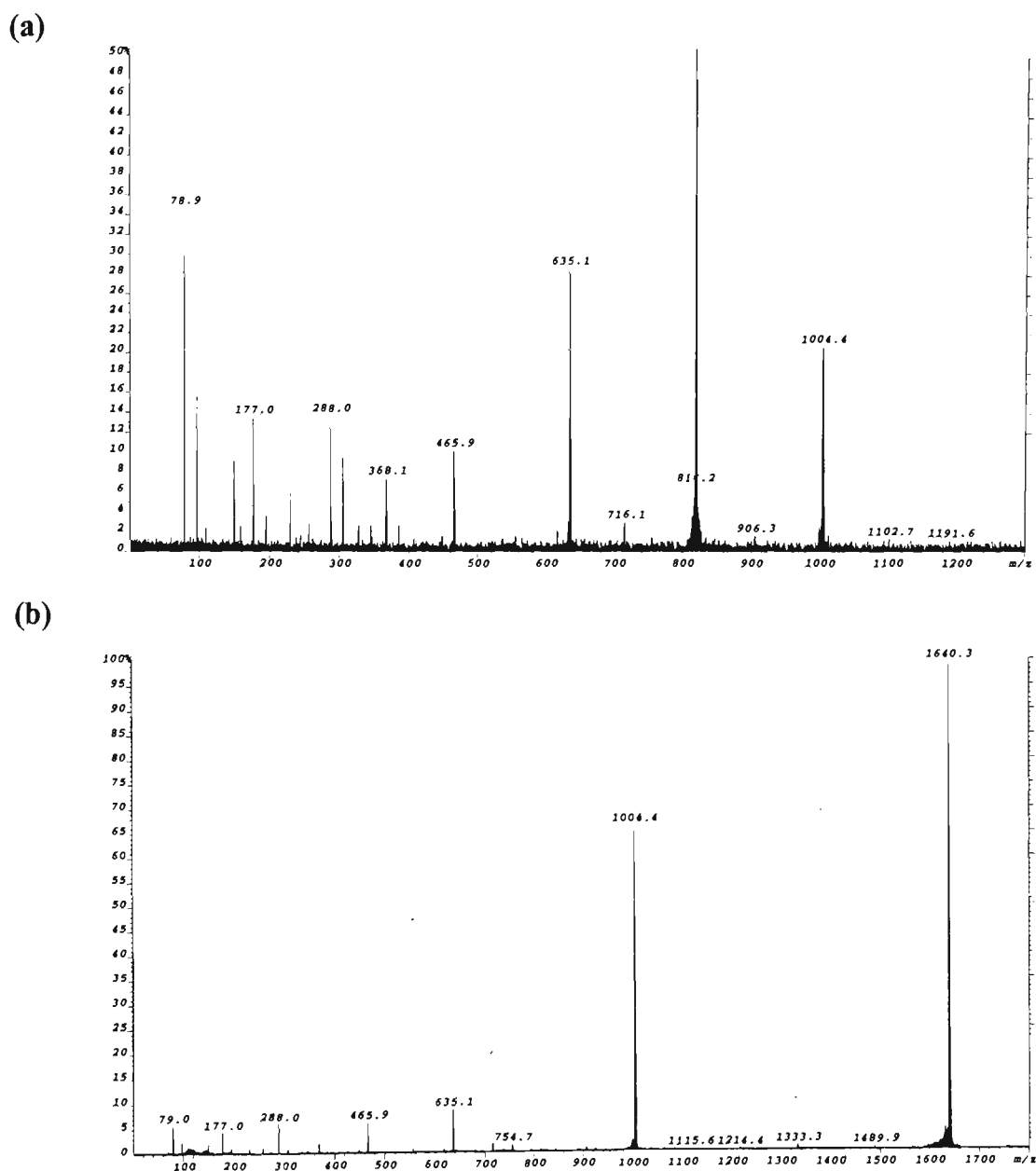


Figure 6.18: ESI-MS/MS spectrum of the (a) 2⁻ ion (*m/z* 819.8) and (b) 1⁻ ion (*m/z* 1640.3) of product ion type II generated by in-source fragmentation of the 5'-d(GCCGGC)-3'/hedamycin adduct.

(e) Hedamycin binding to 5'-d(CG GCCG)-3'

The reaction between hedamycin and 5'-d(CG GCCG)-3' also resulted in the formation of a major adduct after 1 hour of reaction as shown in the HPLC chromatogram given in figure 6.19 eluting at 23.3 minutes. The peak for the unbound oligonucleotide is present at a retention time of 17.8 minutes. The molecular weight

determined from ESI-MS analysis was consistent with hedamycin bound to single-stranded 5'-d(CGGCCG)-3' ($M_{r(meas.)}=2539 \pm 1$ Da, $M_{r(calc.)}=2538.7$ Da monoisotopic mass). Figure 6.20 shows the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the 5'-d(CGGCCG)-3'/hedamycin adduct. The spectrum shows a major fragmentation pathway (illustrated in scheme 6.7) which give rise to ions only for the loss of alkylated guanine (II) at m/z 819.6 (2^- ion) and m/z 1640.1 (1^- ion), and a [hedamycin+guanine-H] $^-$ ion (I) at m/z 896.6. The absence of ions for cleavage of the ribose 3' C-O bond at the site of nucleobase loss is consistent with alkylation of the terminal guanine analogous to the results obtained for the hedamycin/5'-d(CGTACG)-3' adduct (see above).

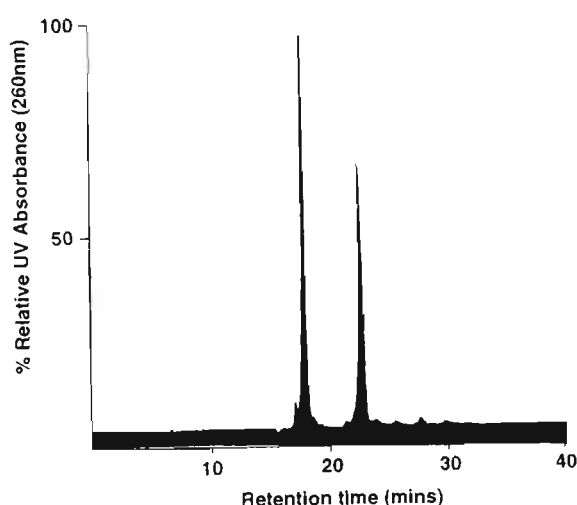


Figure 6.19: HPLC profile of the reaction between 5'-d(CGGCCG)-3' and hedamycin.

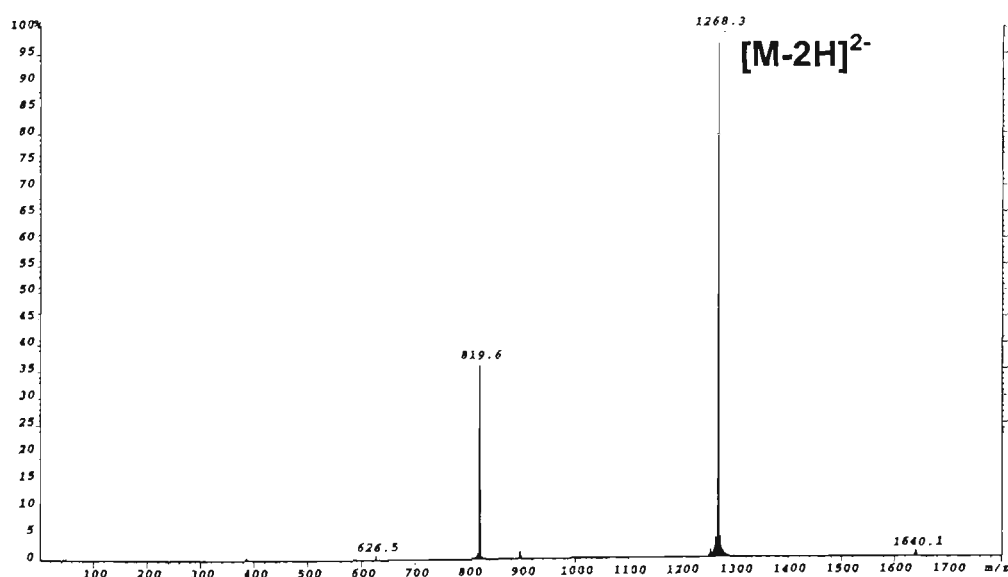
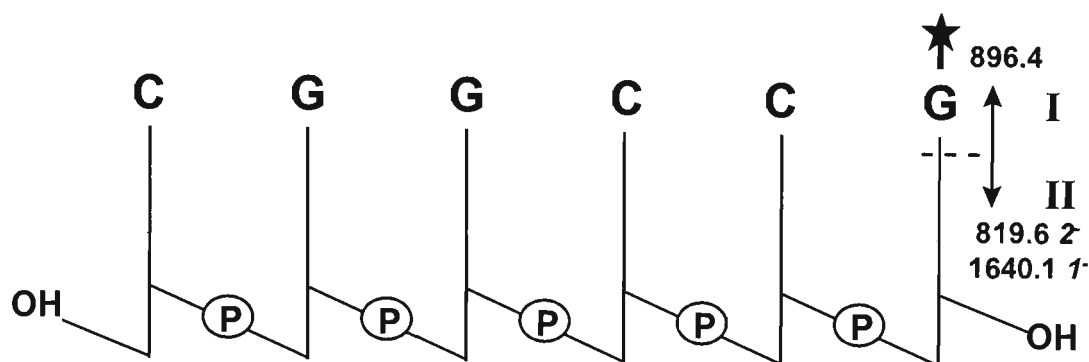


Figure 6.20: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the 5'-d(CGGCCG)-3'/hedamycin adduct.



Scheme 6.7: Fragmentation pathways observed in the ESI-MS/MS spectrum of the 5'-d(CGGCCG)-3'/hedamycin adduct.

The MS/MS spectra of the depurinated oligonucleotide product ion (**II**) generated by in-source collisional activation of the 5'-d(CGGCCG)-3'/hedamycin adduct is given in figure 6.21 for the 2⁻ ion (m/z 819.6) and figure 6.22 for the 1⁻ ion (m/z 1640.5). In both cases, fragmentation of the precursor ion results in a range of product ions with no single fragmentation pathway being dominant. These observations support the location of the site of alkylation as the terminal guanine, since, if alkylation was occurring at G-2 or G-3, decomposition of product ion type **II** would be thought to give rise to facile ribose 3' C-O bond cleavage producing product ion types **III** and **IV**. The assignments of the various product ions observed in the MS/MS spectra of the 2⁻ and 1⁻ ions of product ion type **II** are listed in tables 6.3 and 6.4 respectively. From the assignment of product ions, it can be seen that because the oligonucleotide 5'-d(CGGCCG)-3' is only composed of C and G nucleotide residues, the majority of backbone cleavages result in product ions having masses which could be attributed to more than one possible type of fragment ion. For example, a, b, c, and d backbone cleavages between residues G-1 and G-2 would result in ions with the same m/z values as ions resulting from z, y, x, and w backbone cleavages respectively between C-4 and C-5. Similar problems in the assignment of product ions also exist for fragments produced from a, b, c and d-type backbone cleavages between C-4 and C-5 which could be attributed to z, y, x, and w-type cleavages between G-1 and G-2.

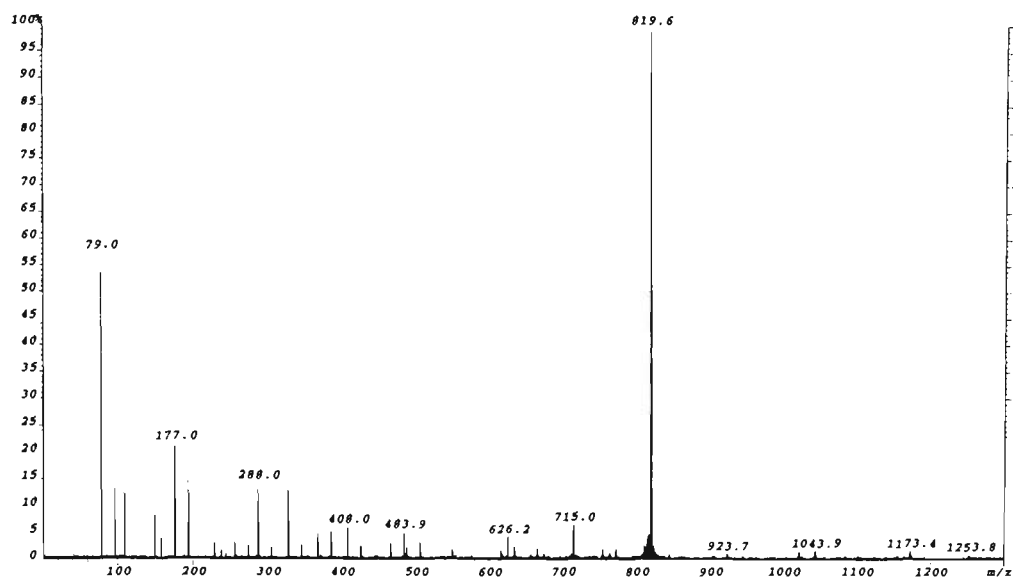


Figure 6.21: ESI-MS/MS spectrum of 2⁻ ion of product ion type II at m/z 819.6 generated by in-source collisional activation of the 5'-d(CGGCCG)-3'/hedamycin adduct.

Table 6.3 : Assignment of polynucleotide product ions observed in the ESI-MS/MS spectrum of the 2⁻ ion of product ion type II (m/z 819.6) of the 5'-d(CGGCCG)-3'/hedamycin adduct.

m/z	Rel. Int.(%)	m/z(calc.)	Assignment
1253.8	0.3	1253.2	w ₄ ⁻ or d ₄ ⁻ or [(GGCC)+H ₂ O-H] ⁻
1173.4	1.0	1173.2	y ₄ ⁻ or b ₄ ⁻ or [(GGCC)-p+H ₂ O-H] ⁻
1043.9	1.3	1044.2	[a ₄ -B ₄ H-2H] ⁻ or [z ₄ -CH-2H] ⁻ or [(GGC)+s-H] ⁻
1021.7	1.3	1021.2	[y ₄ -GH-H] ⁻ or [b ₄ -GH-H] ⁻ or [(GCC)+s+H ₂ O-H] ⁻
923.7	1.9	924.1	w ₃ ⁻ or [(GCC)+H ₂ O-H] ⁻
819.6	100	819.6	[(5'-d(CGGCCG)-3')-GH-2H] ²⁻
773.0	2.4	773.1	[w ₃ -GH] ⁻ or [(CC)+s+p+H ₂ O-H] ⁻
		or 773.1	[b ₃ -CH] ⁻ or [(GG)+s+H ₂ O-H] ⁻
755.0	2.3	755.1	[x ₃ -GH] ⁻ or [(CC)+s+p-H] ⁻
		or 755.1	[a ₃ -CH] ⁻ or [(GG)+s-H] ⁻
		or 755.1	[(GG)+p+H ₂ O-H] ⁻
715.0	7.0	715.1	[a ₃ -B ₃ H-2H] ⁻ or [z ₃ -CH-2H] ⁻ or [(GC)+s-H] ⁻
		or 715.1	[(GC)+p+H ₂ O-H] ⁻
666.5	1.2	666.1	[a ₅ -B ₅ H-3H] ²⁻ or [z ₅ -GH-3H] ²⁻ or [(GGCC)+s-2H] ²⁻
657.1	1.5	657.1	[(CC)+s-H ₂ O-H] ⁻
		or 657.1	[(CC)+p-H] ⁻
		or 657.1	[(GG)-H] ⁻
635.2	2.8	635.1	w ₂ ⁻ or d ₂ ⁻ or [(GC)+H ₂ O-H] ⁻
626.2	4.5	626.1	[w ₄ -H] ²⁻ or [d ₄ -H] ²⁻ or [(GGCC)+H ₂ O-2H] ²⁻
617.2	2.0	617.1	[c ₂ -2H] ⁻ or [x ₂ -2H] ⁻ or [(GC)-H] ⁻
550.7	2.0	550.5	[w ₄ -GH-H] ²⁻ or [d ₄ -GH-H] ²⁻ or [(CCG)+s+p+H ₂ O-2H] ²⁻

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), M_r =98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

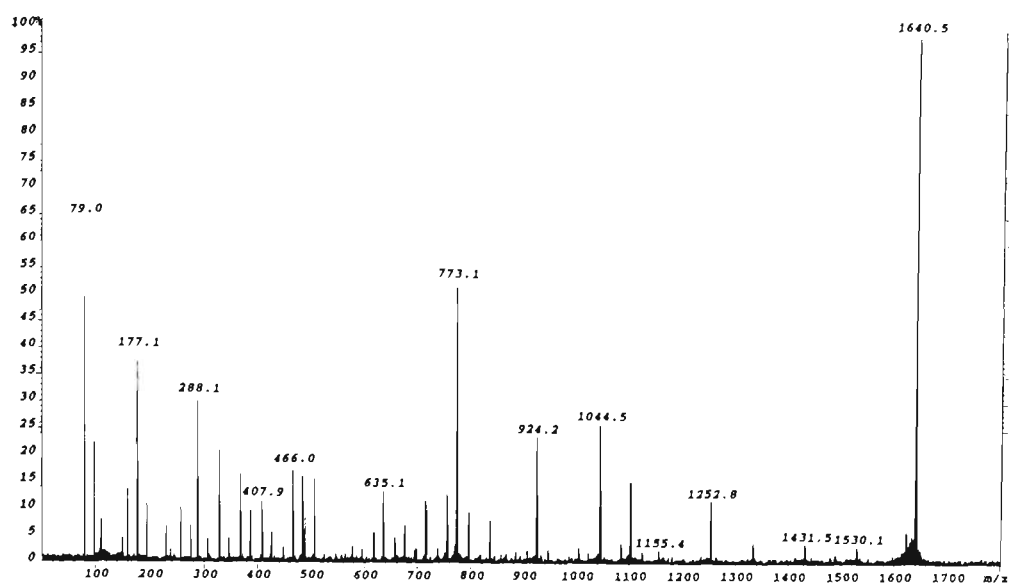


Figure 6.22: ESI-MS/MS spectrum of 1⁻ ion of product ion type II at m/z 1640.5 generated by in-source collisional activation of the 5'-d(CGGCCG)-3'/hedamycin adduct.

Table 6.4 : Assignment of polynucleotide product ions observed in the ESI-MS/MS spectrum of the 1⁻ ion of product ion type II (m/z 1640.5) of the 5'-d(CGGCCG)-3'/hedamycin adduct.

m/z	Rel. Int.(%)	m/z(calc).	Assignment
1640.5	100	1640.3	[5'-d(CGGCCG)-3'-GH-H]-
1622.4	6.5	1622.3	loss of H ₂ O
1530.1	3.7	1530.3	loss of cytosine
1431.5	4.1	1431.3	[w ₅ -GH] ⁻ or [d ₅ -CH] ⁻
1333.4	4.2	1333.2	[a ₅ -B ₅ H-2H] ⁻ or [z ₅ -GH-2H] ⁻
1252.8	12.0	1253.2	w ₄ ⁻ or d ₄ ⁻ or [(GGCC)+H ₂ O-H] ⁻
1155.4	2.6	1155.2	[z ₄ -2H] ⁻ or [a ₄ -2H] ⁻ or [(GGCC)-p-H] ⁻
1124.3	2.4	1124.1	[x ₄ -CH-2H] ⁻ or [c ₄ -CH-2H] ⁻ or [(GGC)+s+p-H] ⁻
1102.1	15.8	1102.2	[w ₄ -GH] ⁻ or [d ₄ -GH] ⁻ or [(GCC)+s+p+H ₂ O -H] ⁻
1084.0	4.1	1084.2	[x ₄ -GH-2H] ⁻ or [c ₄ -GH-2H] ⁻ or [(GCC)+s+p-H] ⁻
1044.3	26.6	1044.2	[a ₄ -B ₄ H-2H] ⁻ or [z ₄ -CH-2H] ⁻ or [(GGC)+s-H] ⁻
1003.9	3.2	1004.2	[a ₄ -GH-2H] ⁻ or [z ₄ -GH-2H] ⁻ or [(CCG)+s-H] ⁻
		1004.1	[(CCG)+p+H ₂ O-H] ⁻
946.3	2.8	946.1	[c ₃ -2H] ⁻ or [(GGC)-H] ⁻
924.2	24.1	924.1	w ₃ ⁻ or [(GCC)+H ₂ O-H] ⁻
905.6	2.3	906.1	[x ₃ -2H] ⁻ or [(GCC)-H] ⁻
884.5	2.6	884.2	b ₃ ⁻ or [(GGC)-p+H ₂ O-H] ⁻
835.1	8.2	835.1	[c ₃ -CH-2H] ⁻ or [(GG)+s+p-H] ⁻
		or 835.0	[(CC)+p+s+p-H] ⁻
		or 835.1	[(CC)+s+p+s-H ₂ O -H] ⁻
795.1	9.7	795.1	[c ₃ -GH-2H] ⁻ or [x ₃ -CH-2H] ⁻ or [(CG)+s+p-H] ⁻
773.1	52.9	773.1	[w ₃ -GH] ⁻ or [(CC)+s+p+H ₂ O -H] ⁻
		or 773.1	[b ₃ -CH] ⁻ or [(GG)+s+H ₂ O -H] ⁻
755.2	13.4	755.1	[x ₃ -GH] ⁻ or [(CC)+s+p-H] ⁻
		or 755.1	[a ₃ -CH] ⁻ or [(GG)+s-H] ⁻
		or 755.1	[(GG)+p+H ₂ O -H] ⁻
737.2	3.2	737.1	[(CC)+s+p-H ₂ O-H] ⁻
		or 737.1	[(GG)+s-H ₂ O-H] ⁻
		or 737.1	[(GG)+p-H] ⁻
715.5	12.1	715.1	[a ₃ -B ₃ H-2H] ⁻ or [z ₃ -CH-2H] ⁻ or [(GC)+s-H] ⁻
		or 715.1	[(GC)+p+H ₂ O -H] ⁻
697.0	3.1	697.1	[(GC)+p-H] ⁻
		or 697.1	[(GC)+s-H ₂ O-H] ⁻

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
675.0	7.2	675.1	[(CC)+p+ H ₂ O -H] ⁻
		or 675.1	[z ₃ -GH-2H] ⁻ or [(CC)+s-H] ⁻
657.0	4.9	657.1	[(CC)+s-H ₂ O -H] ⁻
		or 657.1	[(CC)+p-H] ⁻
		or 657.1	[(GG)-H] ⁻
635.1	13.8	635.1	w ₂ ⁻ or d ₂ ⁻ or [(GC)+H ₂ O-H] ⁻
616.7	5.8	617.1	[c ₂ -2H] ⁻ or [x ₂ -2H] ⁻ or [(GC)-H] ⁻
595.1	3.0	595.2	[(CC)+H ₂ O-H] ⁻
577.4	3.3	577.1	[(CC)-H] ⁻

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), *M_r* =98.0368 Da
p = PO₃H, *M_r* = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Moreover these ions could also be attributed to internal dinucleotide (GC) and tetranucleotide (GGCC) ion-types as described in detail in chapter four section 4.4 (see table 4.7). Confirmation of the site of guanine loss as the 3' terminal guanine could be obtained from the observation of product ions which can be unambiguously assigned as having the guanines both attached at the G-2 and G-3 positions such as: (a) dinucleotide product ions comprising the residues (GG); (b) trinucleotide product ions comprising the residues (CGG) (ie. [a₃-2H]⁻, b₃⁻, [c₃-2H]⁻, d₃⁻ ions), or (GGC) internal ions, or (c) product ions comprising the residues (CGGCC) (ie. [a₅-2H]⁻, b₅⁻, [c₅-2H]⁻, d₅⁻ ions). In the case of dinucleotide (GG) product ions, an added problem exists owing to the inability to resolve the mass difference between [(GG)_{nt}] (658.1050 Da monoisotopic mass) and [(CC)_{nt}+p] (658.0591 Da monoisotopic mass) and [(CC)_{nt}+s-H₂O] (658.1190 Da monoisotopic mass). Therefore, in many cases (GG) dinucleotide ions be alternatively assigned as (CC) dinucleotide ions with a different portion of the backbone. This problem is also apparent in the MS/MS spectrum of the 2⁻ ion of product ion type II (figure 6.20) with the ions at *m/z* 773.0, *m/z* 755.0 and *m/z* 657.1 (table 6.3). From the list of assignments given in table 6.3 it can be seen that no product ions are observed which provide conclusive evidence of depurination of the terminal guanine. The MS/MS spectrum of the 1⁻ ion of product ion type II, however, shows ions indicative of 3' terminal guanine depurination at *m/z* 946.3 for either [c₃-2H]⁻ or [(GGC)-H]⁻, and at *m/z* 884.5 for b₃⁻ or [(GGC)-p+H₂O-H]⁻. The observations of these

ions establishes the presence of guanine nucleobases at G-2 and G-3 leaving the 3' terminal guanine as the only remaining possible site of guanine depurination.

The MS/MS spectrum of the hedamycin+guanine adduct (not shown) was essentially the same as that obtained for this product ion generated from all other hedamycin/oligonucleotide adducts with the G⁻ anion being the most intense ion observed in the spectrum and the remaining fragment ions as illustrated in scheme 6.2.

(f) Hedamycin binding to 5'-d(GCCGGC)-3' versus 5'-d(CGGCCG)-3': Discussion

The results obtained for hedamycin binding to the oligonucleotides 5'-d(GCCGGC)-3' and 5'-d(CGGCCG)-3' again show the location of the high affinity binding site to be an important factor in the sequence selective binding of hedamycin to these relatively short oligonucleotides. The oligonucleotide 5'-d(GCCGGC)-3' incorporates three potential guanine alkylation sites with the high affinity 5'-CGG site located in the middle of the sequence. The MS/MS spectrum of the only adduct formed upon hedamycin binding was found to be consistent with alkylation of the 5'-G at 5'-CGG site. That alkylation was not observed at the other two guanines not only confirms the strong affinity of hedamycin for the 5'-CGG binding site, but in addition, the lack of selectivity at 5'-GG sequences, which was also shown in previous investigations using other sequencing techniques². The effect of placing the high affinity 5'-CGG binding site at the end of the oligonucleotide as in 5'-d(CGGCCG)-3' results in a competition between two guanine binding sites (ie. a 5'-CGG and a 5'-CG site) upon hedamycin binding to the oligonucleotide duplex. This situation is analogous to placement of the 5'-CGT at the end of the oligonucleotide 5'-d(CGTACG)-3' illustrated in scheme 6.4. The preference for guanine alkylation at the 3' terminal guanine in the case of 5'-d(CGGCCG)-3' is consistent with the preference found for the oligonucleotide 5'-d(CGTACG)-3'. Once again, this can be rationalised upon the basis of the lack of helical constraints at the end of these relatively short duplexes and the stabilisation of base-pairing from hydrophobic interactions resulting from alkylation of the guanine of the terminal base-pair (explained above).

6.3.2 *Hedamycin binding to the octanucleotides 5'-d(GCGTACGC)-3', 5'-d(ACGTACGT)-3', and 5'-d(TCGTACGA)-3'*

In order to investigate further the features of hedamycin sequence selectivity observed for the hexanucleotides, reactions between hedamycin and the self complementary octanucleotide, 5'-d(GCGTACGC)-3' were carried out. These oligonucleotides were designed to incorporate the core 5'-d(CGTACG)-3' sequence with an additional G-C base-pair at each termini so as to provide additional helical constraints (the G-C base pair was selected rather than the A-T base pair for this study owing to its greater stability). The sequence selective binding of hedamycin to this oligonucleotide was examined to provide insight into the basis of the observed change in sequence selectivity of hedamycin binding from the high affinity 5'-CGT site to the 5'-CG site observed in the case of 5'-d(CGTACG)-3'. In addition, hedamycin binding studies were conducted on the octanucleotides 5'-d(TCGTACGA)-3' and 5'-d(ACGTACGT)-3' which also incorporate the core 5'-d(CGTACG)-3' sequence but with 5'T-A and 5'A-T base pairs respectively at each terminus. Of interest was how selectivity of hedamycin binding to the core 5'-d(CGTACG)-3' sequence would be affected by the helical constraints imposed by the addition of A-T (or T-A) base pairs to each terminus which are less stable in comparison to G-C base pairs (owing to their being three hydrogen bonds between G-C base pairs whereas only two are possible between A-T (see chapter one, section 1.1.3). The oligonucleotides 5'-d(GCGTACGC)-3' and 5'-d(TCGTACGA)-3' were designed to enable direct comparison between the binding of hedamycin to the 5'-CGT site versus the 5'-CG site when the binding site is not terminal in the oligonucleotide, whereas the oligonucleotide 5'-d(ACGTACGT)-3' comprises two high affinity 5'-CGT sites.

(a) *Hedamycin binding to 5'-d(GCGTACGC)-3'*

The HPLC profile of the reaction between hedamycin and the octanucleotide 5'-d(GCGTACGC)-3' at a molar ratio of 1:1 hedamycin:duplex oligonucleotide obtained after 1 hour is shown in figure 6.23. Figures 6.24(a), (b) and (c) shows the ESI

mass spectra of the peaks eluting at 21.3, 23.5 and 25.0 minutes respectively. The ESI mass spectrum of the species eluting at 21.3 minutes (figure 6.24(a)) yields ions at m/z 601.5, 802.2 and 1203.9 corresponding to the $[M-4H]^4$, $[M-3H]^3$ and $[M-2H]^2$ ions, respectively of the unbound oligonucleotide 5'-d(GCGTACGC)-3' ($M_r(\text{meas.})=2410 \pm 1$ Da, $M_r(\text{calc.})=2409.4$ Da monoisotopic mass). The species eluting at 23.5 minutes and 25.0 minutes produced very similar ESI-MS spectra (shown in figures 6.24(b) and (c) respectively) each producing ions at m/z 788.1/788.3 and at m/z 1051.3/1051.4 owing to the $[M-4H]^4$ and $[M-3H]^3$ ions of a single-stranded 5'-d(GCGTACGC)-3'/hedamycin adduct ($M_r(\text{meas.})=3157 \pm 1$ Da, $M_r(\text{calc.})=3155.8$ Da monoisotopic mass). The ions present at m/z 900.6/901.2 and m/z 1261.2/1261.6 in both spectra are owing to the $[M-7H]^7$ and $[M-5H]^5$ ions of the duplex adduct (5'-d(GCGTACGC)-3'/hedamycin)₂.

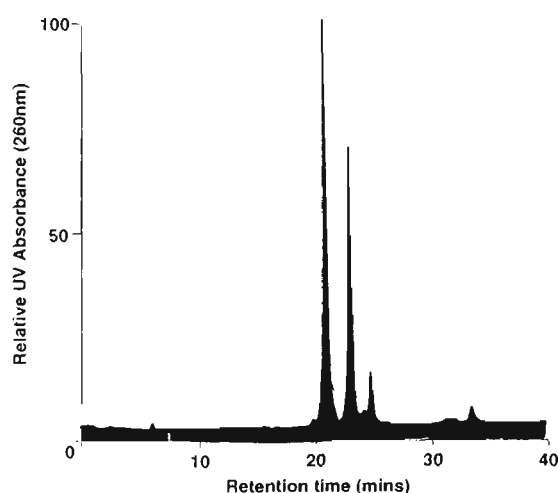
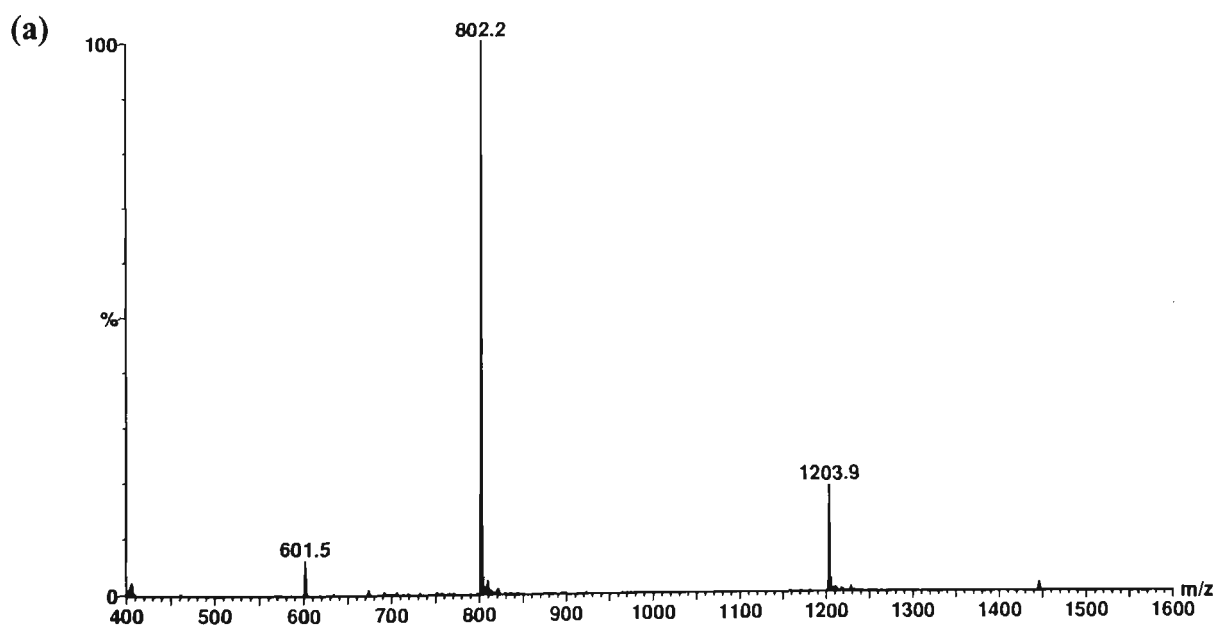


Figure 6.23: HPLC profile of the reaction between 5'-d(GCGTACGC)-3' and hedamycin.



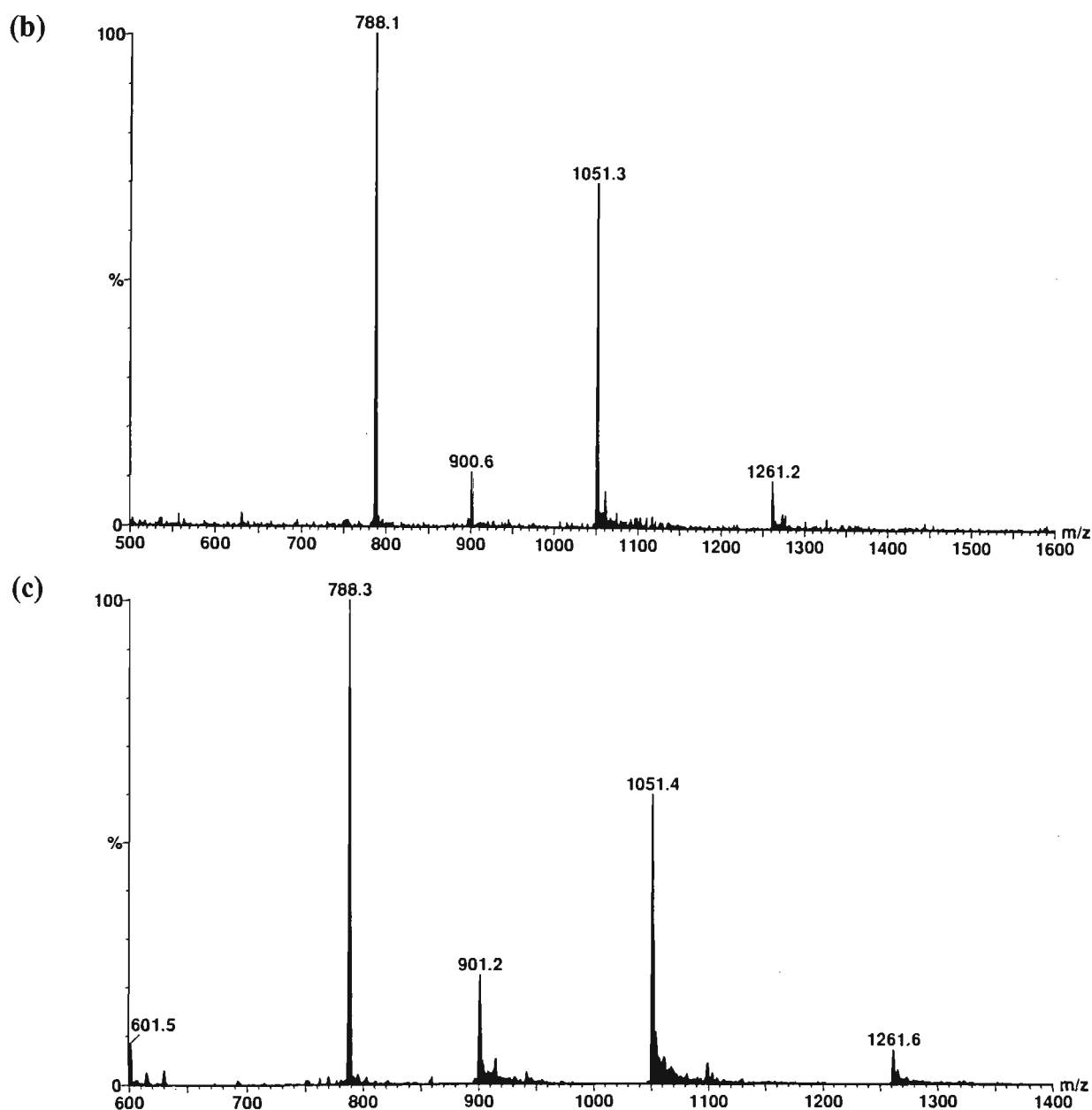


Figure 6.24: ESI mass spectra of (a) the peak eluting at 21.3 minutes, (b) the peak eluting at 23.5 minutes and (c) the peak eluting at 25.0 minutes in the HPLC profile of the 5'-d(GCGTACGC)-3'/hedamycin reaction mixture.

The ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major adduct formed in the 5'-d(GCGTACGC)-3'/hedamycin reaction mixture is given in figure 6.25. As observed for the hedamycin adducts of the hexanucleotides, the spectrum shows a highly specific fragmentation pathway initiated by the loss of alkylated guanine at m/z 1128.2 (**II**) followed by cleavage of the ribose C-O bond to the 3' side of the depurinated G-3 yielding ions at m/z 715.1 (**III**) and m/z 1541.2 (**IV**). In contrast to the MS/MS spectra of the $[M-2H]^{2-}$ ions of hedamycin adducts of hexanucleotides, no peak is detected for the hedamycin+guanine adduct (**I**), and therefore analysis of the $[M-3H]^{3-}$ ions of hedamycin/octanucleotide adducts was undertaken. Shown in figure 6.26 is the

ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of major adduct formed at m/z 1051.3. N-glycosidic bond cleavage of the alkylated guanine results in observation of both the 2⁻ ion of the depurinated oligonucleotide (II) at m/z 1128.2 and the [hedamycin+guanine-H]⁻ ion (I) at m/z 896.5. Cleavage of the ribose 3' C-O bond at the site of depurination gives rise to the ions at m/z 715.1 for product ion type III, and at m/z 1541.3 and m/z 770.1 for the 1⁻ and 2⁻ ions of product ion type IV. This fragmentation pattern is illustrated in scheme 6.8 and clearly establishes the site of hedamycin alkylation of the major adduct at the G-3 guanine in 5'-d(GCGTACGC)-3'.

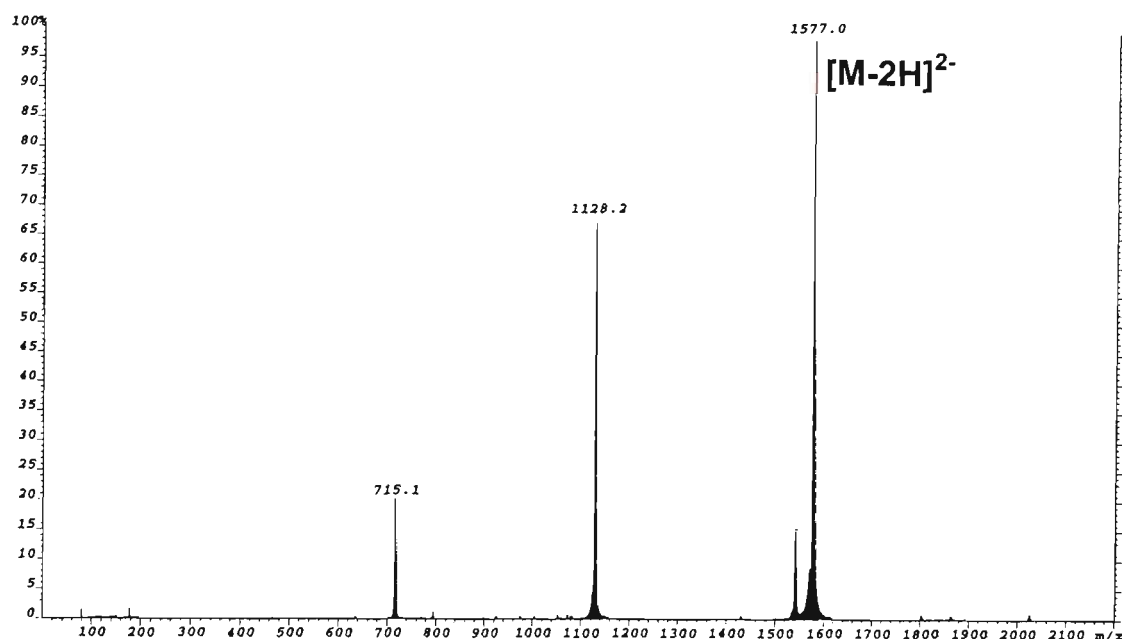


Figure 6.25: ESI-MS/MS spectrum of $[M-2H]^{2-}$ ion of the major adduct formed between 5'-d(GCGTACGC)-3' and hedamycin.

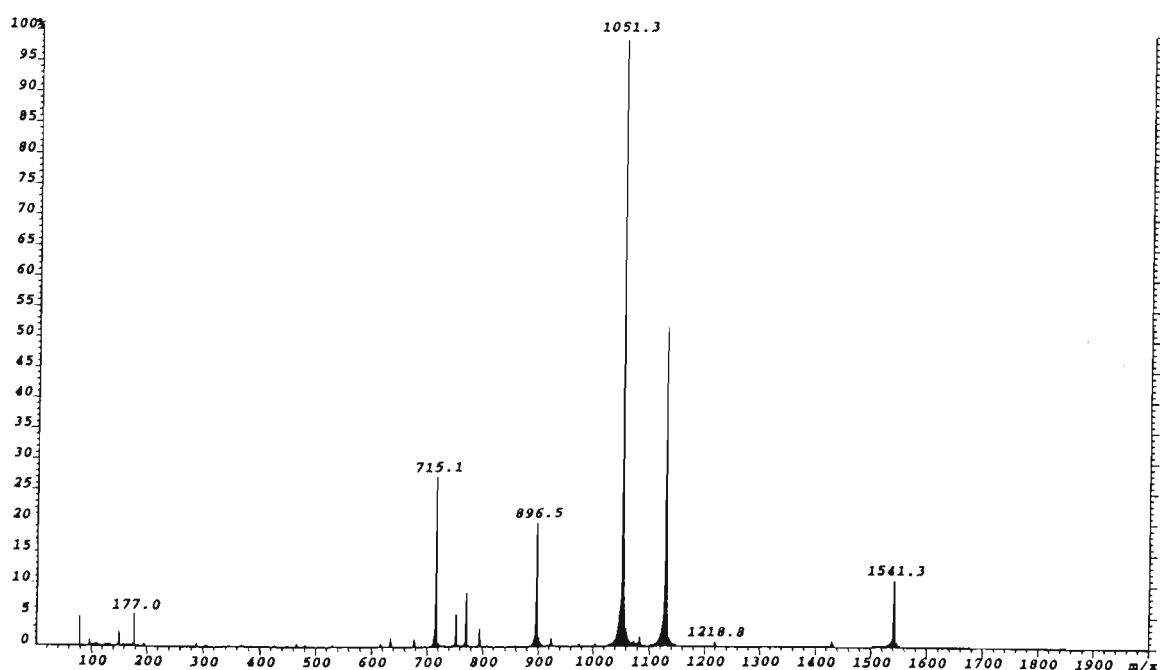
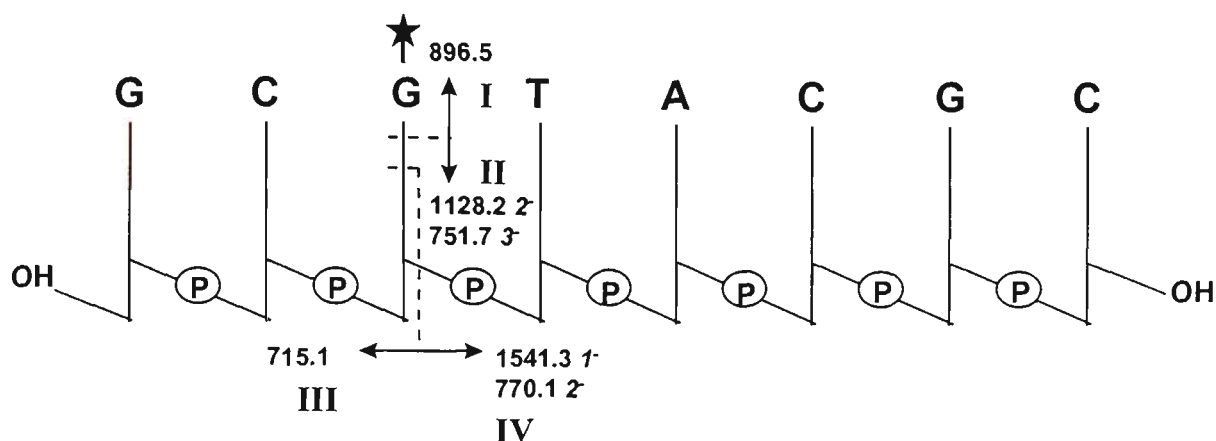


Figure 6.26: ESI-MS/MS spectrum of $[M-3H]^{3-}$ ion of the major adduct formed between 5'-d(GCGTACGC)-3' and hedamycin.



Scheme 6.8: Fragmentation pathways observed in the ESI-MS/MS spectra of the major adduct formed between hedamycin and 5'-d(GCGTACGC)-3'.

Figure 6.27 shows the ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the minor adduct (m/z 1050.9) formed upon hedamycin binding to 5'-d(GCGTACGC)-3'. Loss of alkylated guanine results in the ion at m/z 896.2 for the singly charged hedamycin+guanine adduct (I), and an ion at m/z 1128.2 for the 2⁻ ion of product ion type II. The ions at m/z 974.4 and m/z 306.1 correspond to product ion types III (2⁻ ion) and IV (1⁻ ion) respectively, arising from cleavage of the ribose 3' C-O bond following N-glycosidic bond cleavage of the depurinated residue. This dissociation is consistent with guanine alkylation at the 5'-CGC site as shown in scheme 6.9.

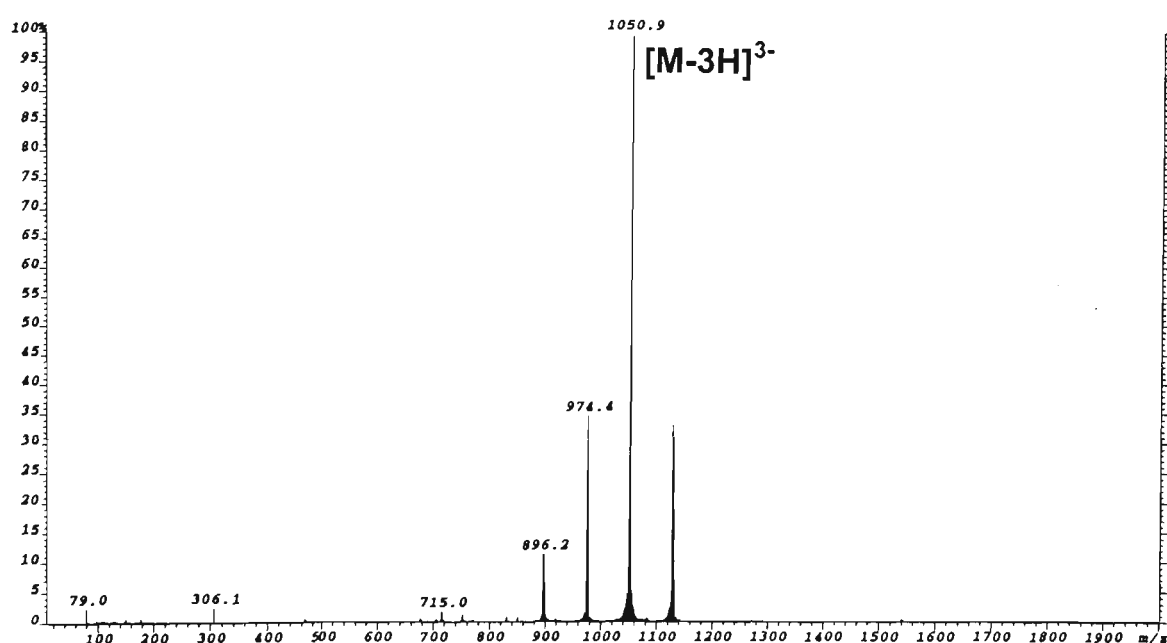
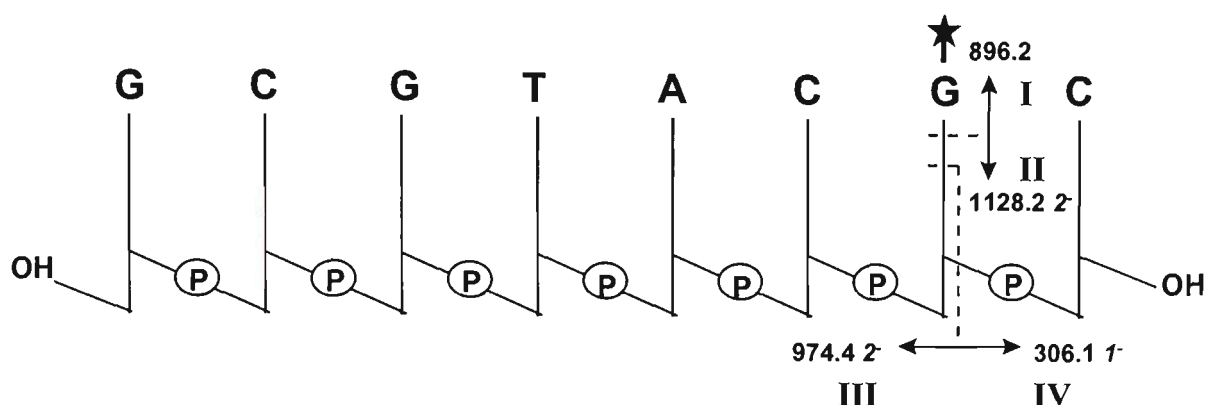


Figure 6.27: ESI-MS/MS spectrum of $[M-3H]^{3-}$ ion of the minor adduct formed between hedamycin and 5'-d(GCGTACGC)-3'.



Scheme 6.9: Fragmentation pathways observed in the ESI-MS/MS spectrum of the minor adduct formed between hedamycin and 5'-d(GCGTACGC)-3'.

(b) Hedamycin binding to 5'-d(ACGTACGT)-3'

The reaction between hedamycin and the oligonucleotide 5'-d(ACGTACGT)-3' also produces two predominant adducts in the HPLC profile after 1 hour which is shown in figure 6.28. The peak for the unbound oligonucleotide is observed at a retention time of 22.4 minutes. ESI-MS spectra obtained of the peaks eluting at 26.7 and 24.7 (not shown) yielded a molecular weight in each case of a single-stranded 5'-d(ACGTACGT)-3'/hedamycin adduct ($M_{r(meas.)}=3155 \pm 1$ Da, $M_{r(calc.)}=3154.8$ Da monoisotopic mass). Figure 6.29 shows the ESI-MS/MS spectrum of the $[M-3H]^3-$ ion (m/z 1050.7) of the major 5'-d(ACGTACGT)-3'/hedamycin adduct which had a retention time of 24.7 minutes in the HPLC chromatogram. The spectrum shows an ion at m/z 1127.6 owing to the doubly charged ion for the loss of alkylated guanine (**II**), with the respective triply charged ion present at m/z 751.5. The ions at m/z 698.9 and 1555.8 correspond to product ion types **III** and **IV** respectively from cleavage of the ribose C-O bond to the 3' side of depurinated G-3. This fragmentation pathway, which is shown in scheme 6.10, confirms the site of hedamycin alkylation as the guanine at the 5'-CGT site: 5'-d(ACGTACGT)-3'.

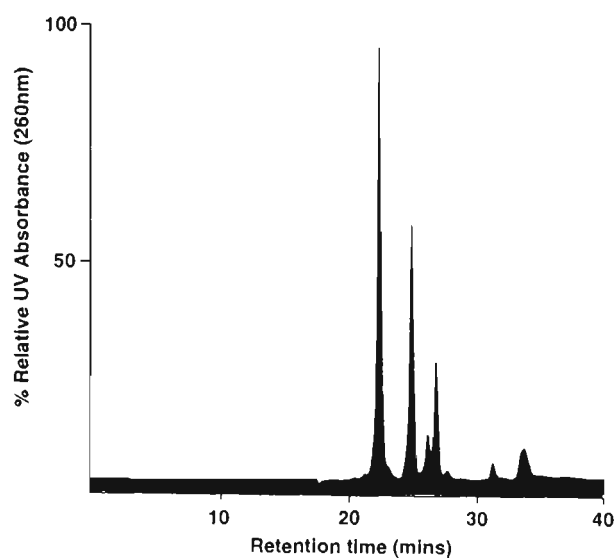


Figure 6.28: HPLC profile of the reaction between 5'-d(ACGTACGT)-3' and hedamycin.

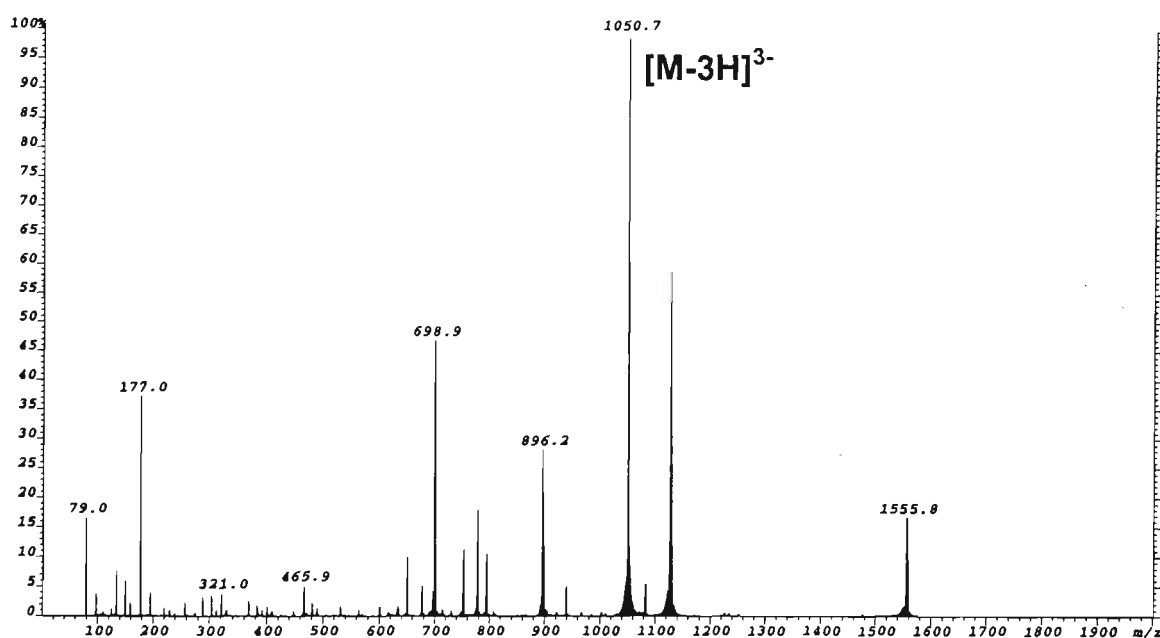
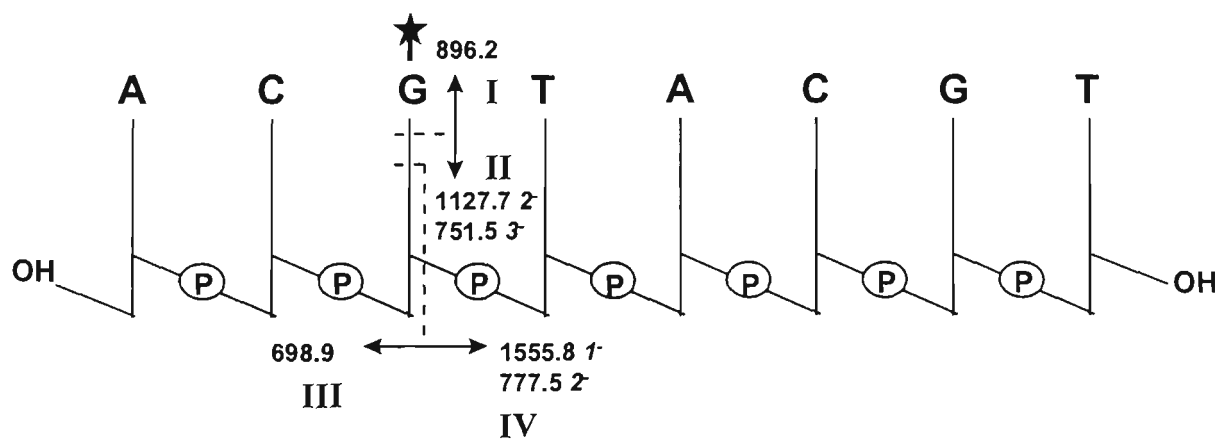


Figure 6.29: ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the major adduct formed between hedamycin and 5'-d(ACGTACGT)-3'.



Scheme 6.10: Fragmentation pathways observed in the ESI-MS/MS spectrum of the major adduct formed between hedamycin and 5'-d(ACGTACGT)-3'.

The MS/MS spectrum of the $[M-3H]^{3-}$ ion of the minor adduct (m/z 1050.6) is given in figure 6.30. Fragmentation of this ion yields an ion for the loss of alkylated guanine (**II**) at m/z 1128.0 and the [hedamycin+guanine-H] $^-$ ion (**I**) at m/z 896.5. Subsequent cleavage of the ribose 3' C-O bond of the depurinated base results in product ion type **III** at m/z 966.7 for the doubly charged ion, and at m/z 1934.3 for the singly charged ion. The complementary product ion type **IV** is present at m/z 321.0. As shown in scheme 6.11, this is indicative of guanine alkylation at the 5'-d(ACGTACGT)-3' site.

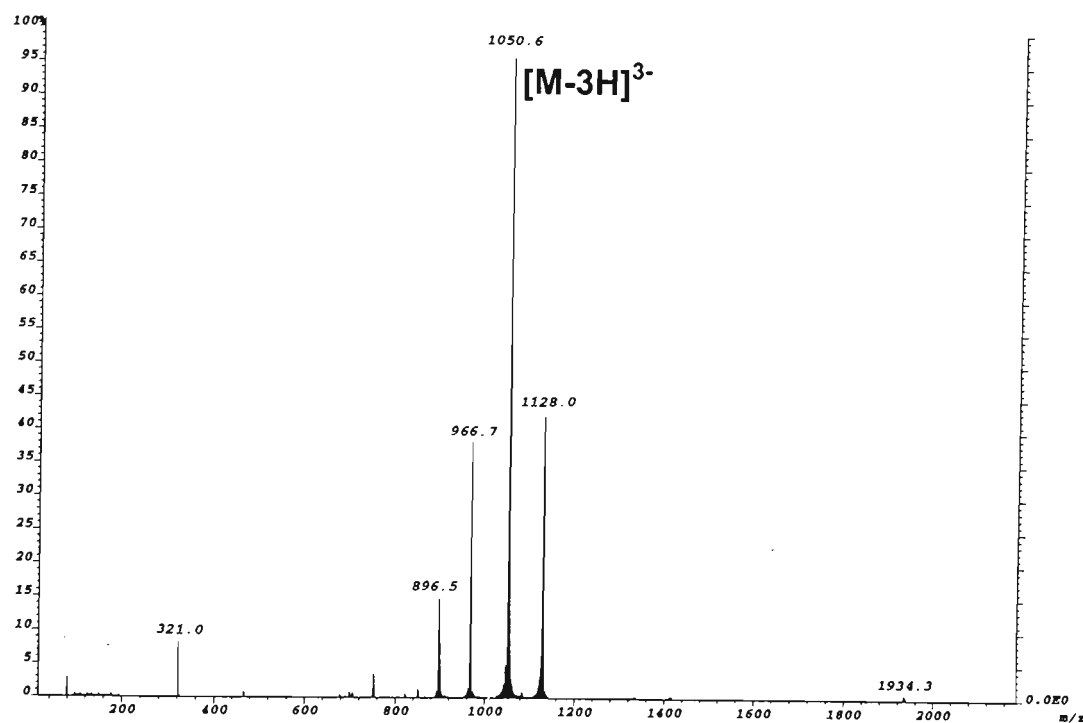
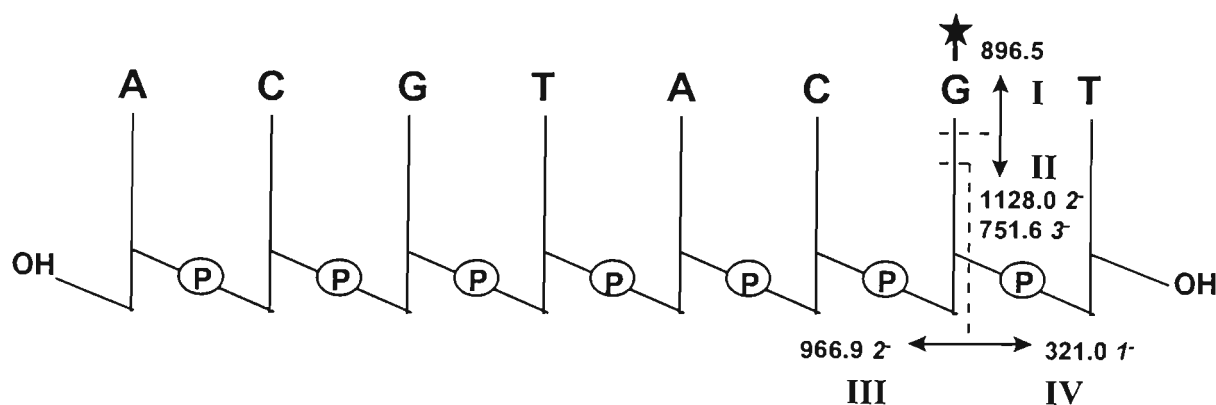


Figure 6.30: ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the minor adduct formed between hedamycin and 5'-d(ACGTACGT)-3'.



Scheme 6.11: Fragmentation pathways observed in the ESI-MS/MS spectrum of the minor adduct formed between hedamycin and 5'-d(ACGTACGT)-3'.

(c) **Hedamycin binding to 5'-d(TCGTACGA)-3'**

Figure 6.31 shows the HPLC chromatogram obtained upon reaction of hedamycin with 5'-d(TCGTACGA)-3' after 1 hour. ESI-MS analysis of the fractions collected from the HPLC purification identified the largest peak eluting at 22.7 minutes to be the unbound oligonucleotide ($M_{r(meas.)}=2409 \pm 1$ Da, $M_{r(calc.)}=2408.5$ Da monoisotopic mass). The species eluting at retention times of 26.3 minutes and 28.1 minutes each yielded molecular weights ($M_{r(meas.)}$) of 3155 ± 1 Da in their ESI-MS spectra. This is consistent, in each case, with a single-stranded 5'-d(TCGTACGA)-3'/hedamycin adduct ($M_{r(calc.)}=3154.8$ Da monoisotopic mass). The ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion (m/z 1050.8) of the most abundant adduct peak at 28.1 minutes in the HPLC profile is given in figure 6.32. Cleavage of the glycosidic bond of the alkylated guanine produces the 2⁻ and 3⁻ ions of the depurinated oligonucleotide (II) at m/z 1127.7 and m/z 751.5 respectively, and a singly charged ion for the hedamycin+guanine adduct at m/z 896.2. The ions at m/z 962.0 and m/z 330.1 correspond to product ion type III (2⁻ ion), and product ion IV, respectively from cleavage of the ribose C-O bond to the 3' side of the G-7 guanine following depurination. As illustrated in scheme 6.12, this enables identification of the site of alkylation of the predominant adduct as the guanine at the 5'-CGA-3' site (ie. 5'-d(TCGTACGA)-3').

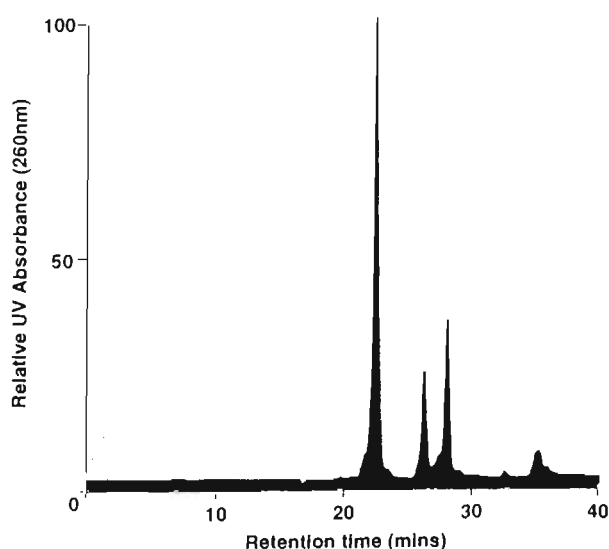


Figure 6.31: HPLC profile of the reaction between 5'-d(TCGTACGA)-3' and hedamycin.

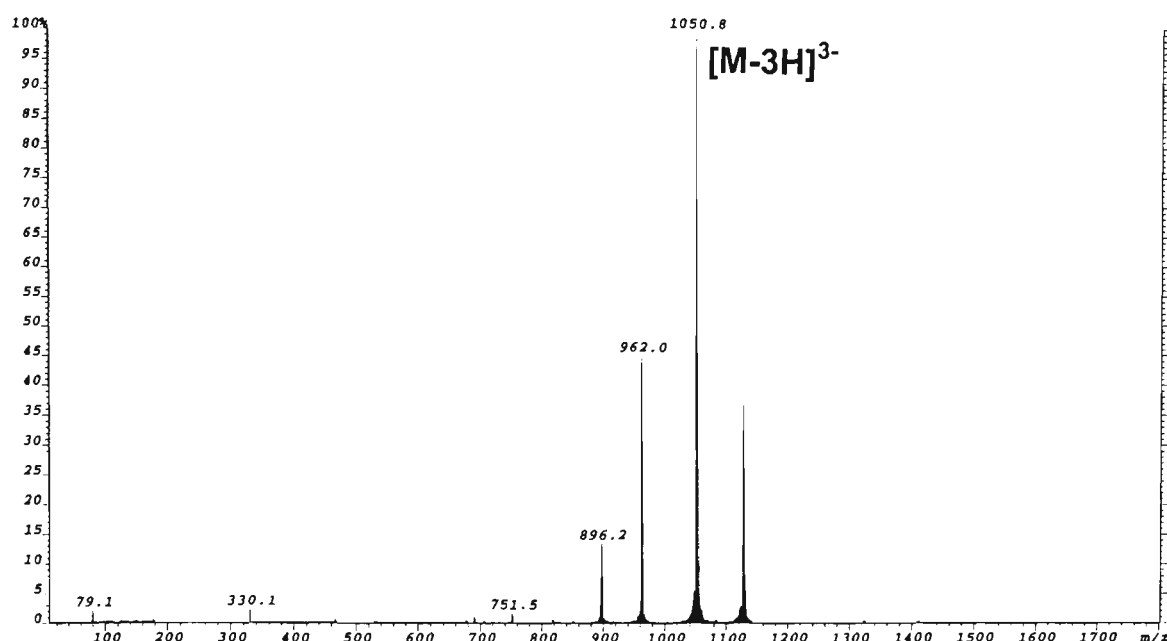
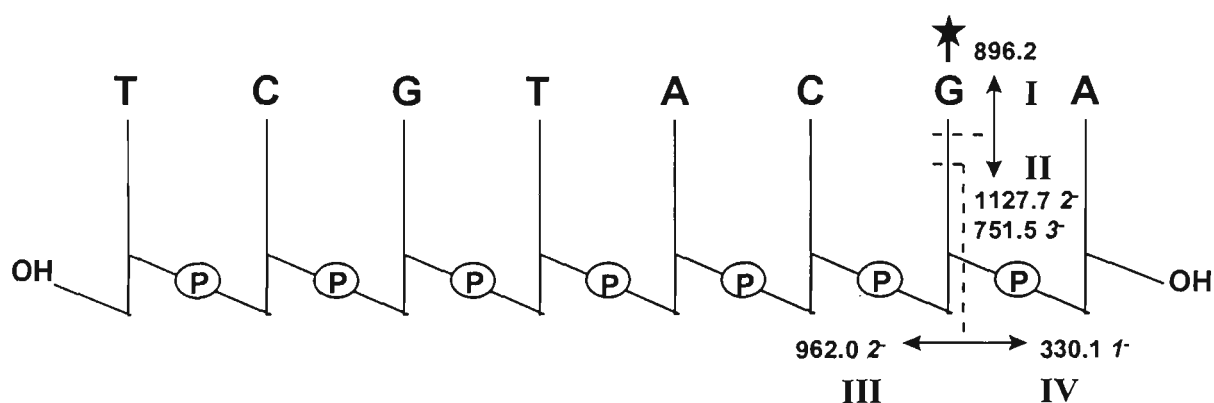


Figure 6.32: ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the major adduct formed between hedamycin and 5'-d(TCGTACGA)-3'.



Scheme 6.12: Fragmentation pathways observed in the ESI-MS/MS spectrum of the major adduct formed between hedamycin and 5'-d(TCGTACGA)-3'.

The ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion (m/z 1050.6) of the minor adduct eluting at a retention time of 26.3 minutes on the HPLC profile is shown in figure 6.33. Loss of alkylated guanine yields the 2⁻ and 3⁻ ions of the depurinated oligonucleotide (**II**) at m/z 1127.7 and 751.5 respectively, and the singly charged hedamycin+guanine adduct at m/z 896.2. The ions at m/z 689.9 and m/z 1564.2 arise from loss of the guanine at the 5'-d(TCGTACGA)-3' site followed by cleavage of the ribose 3' C-O bond giving rise to product ion types **III** and **IV** respectively. This fragmentation confirms the site of hedamycin alkylation of the minor adduct as that of the 5'-d(TCGTACGA)-3' site as illustrated in scheme 6.13.

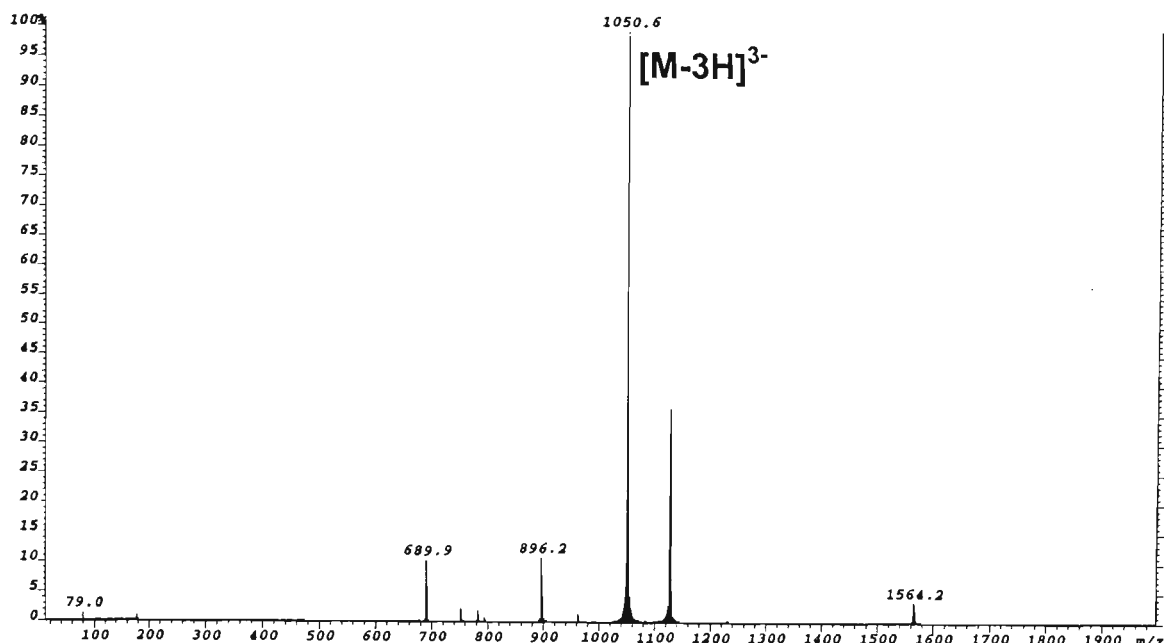
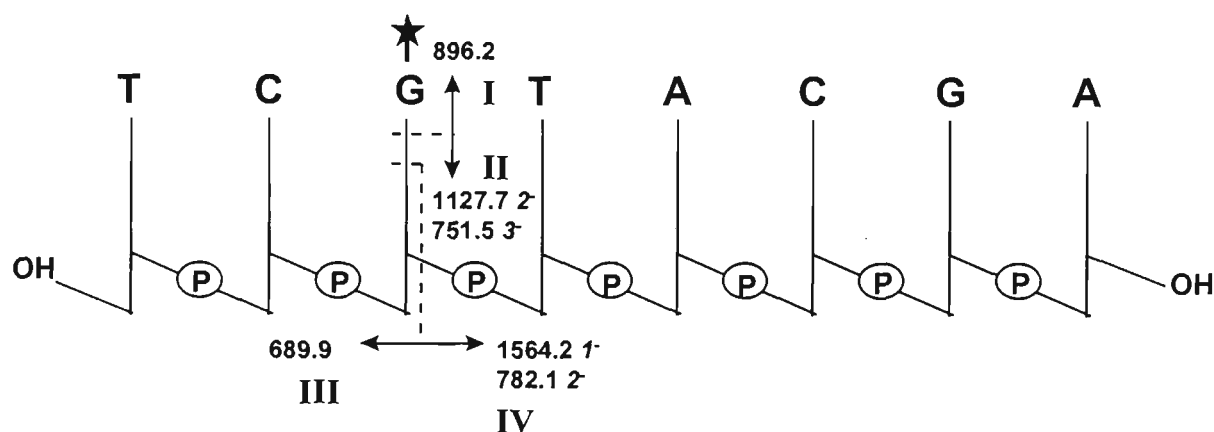


Figure 6.33: ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the minor adduct formed between hedamycin and 5'-d(TCGTACGA)-3'.



Scheme 6.13: Fragmentation pathways observed in the ESI-MS/MS spectrum of the minor adduct formed between hedamycin and 5'-d(TCGTACGA)-3'.

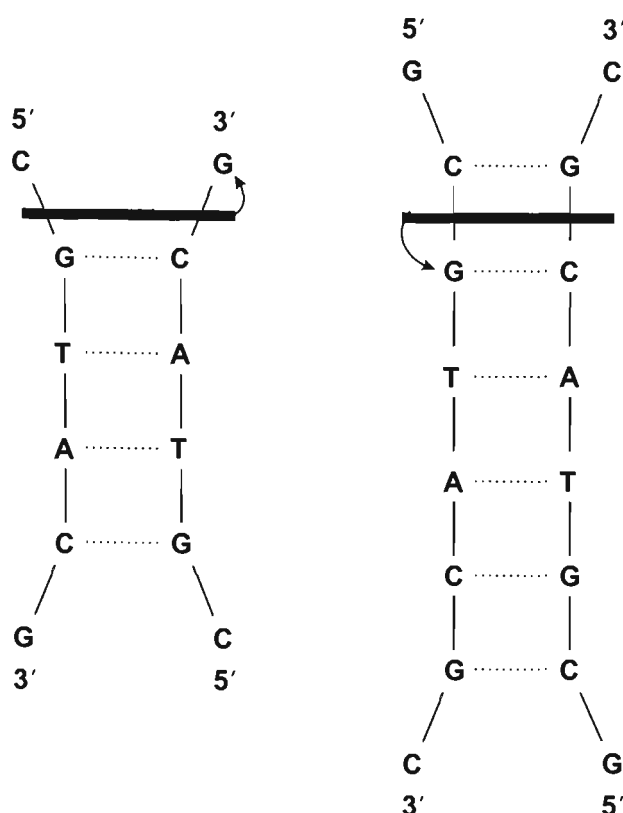
(d) Hedamycin binding to octanucleotides : Discussion

Table 6.5 summarises the results obtained for the binding of hedamycin to each of the octanucleotides 5'-d(GCGTACGC)-3', 5'-d(ACGTACGT)-3' and 5'-d(TCGTACGA)-3'. Each oligonucleotide comprises a 5'-CGT binding site in the same sequence location, however they differ in the base preceding the binding site. This enables a comparison to be made of sequence selectivity of hedamycin binding to the larger sequences 5'-GCGT, 5'-ACGT and 5'-TCGT.

Table 6.5: Results for the binding of hedamycin to octanucleotides.

Oligonucleotide Adducts	HPLC retention time	% Relative adduct formation	Structure
<i>5'-d(GCGTACGC)-3'</i>			
Major Adduct	23.4 mins	36.9%	5'-d(GC <u>G</u> *TACGC)-3'
Minor Adduct	25.1 mins	7.1%	5'-d(GCGTAC <u>G</u> *C)-3'
<i>5'-d(ACGTACGT)-3'</i>			
Major Adduct	24.7 mins	30.1%	5'-d(AC <u>G</u> *TACGT)-3'
Minor Adduct	26.7 mins	13.9%	5'-d(ACGTAC <u>G</u> *T)-3'
<i>5'-d(TCGTACGA)-3'</i>			
Major Adduct	28.1 mins	21.6%	5'-d(TCGTAC <u>G</u> *C)-3'
Minor Adduct	26.3 mins	14.7%	5'-d(TC <u>G</u> *TACGC)-3'

These data concerning the relative abundance of different adducts formed between all the octanucleotides show that the 5'-CGT binding site having the greatest hedamycin binding affinity is that within the sequence 5'-d(GCGTACGC)-3' which yielded 36.9% on the HPLC profile for formation of the 5'-d(GCG*TACGC)-3'/hedamycin adduct. Moreover, this sequence showed the lowest affinity of hedamycin binding to the 5'-CG site of all octanucleotides yielding only 7.1% formation of the 5'-d(GCGTACG*C)-3'/hedamycin adduct. These results show a change in preference for hedamycin binding from the 5'-CG site to the 5'-CGT site in the core -(5'-d(CGTACG)-3')- sequence which is consistent with sequencing studies using other techniques which have shown that hedamycin has a much higher affinity for isolated guanines at 5'-CGT sites compared with 5'-CGC sites². These results can be rationalised by the extra helical constraints introduced by the additional G-C base pair at both ends of the double helix as illustrated in scheme 6.14. When the C-G hedamycin intercalation site is not terminal in the sequence, as in 5'-d(GCGTACGC)-3', guanine alkylation is no longer directed by the lack of helical stability present about the C-G intercalation site as was the case with 5'-d(CGTACG)-3'. Consequently, a change in preference for hedamycin alkylation at the 5'-CGT site over the 5'-CG site in the core -(5'-d(CGTACG)-3')- sequence is observed when the C-G base-pair intercalation site is no longer terminal.



Scheme 6.14: Binding of hedamycin to 5'-d(CGTACG)-3' and 5'-d(GCGTACGC)-3'.

The octanucleotide 5'-d(ACGTACGT)-3' which contains two 5'-CGT binding sites, ie. one at each terminus, showed binding to the 5'-CGT site at the 5' terminus to be favoured over that of the 5'-CGT site at the 3' terminus with the 5'-d(ACGTACGT)-3'/hedamycin adduct formed in an abundance of 30.1%, compared with 13.9% for the 5'-d(ACGTACGT)-3'/hedamycin adduct. Therefore, even when a high affinity site is located at the 3' terminus by placing an additional A-T rather than a G-C base pair at each terminus to provide for the helical constraints, the binding selectivity of hedamycin switches to the 5'-CGT site at the 5' terminus. This is further consistent with the rationalisation for the loss of hedamycin binding selectivity for the 5'-CGT site in the case of 5'-d(CGTACG)-3' being a result of the lack of helical constraints about the hedamycin intercalation site.

Hedamycin binding to the octanucleotide 5'-d(TCGTACGA)-3' produces markedly different results from that observed for both 5'-d(GCGTACGC)-3' and 5'-d(ACGTACGT)-3' with the HPLC data showing the greatest preference for hedamycin binding to the 5'-CG site with the 5'-d(TCGTACGA)-3'/hedamycin adduct

being formed in an abundance of 21.6%. In contrast, the binding of hedamycin to the 5'-CGT site produces the 5'-d(TCG*TACGA)-3'/hedamycin adduct in only 14.7% relative abundance. These results may indicate that the helical constraints provided by the presence of the T-A base pairs about the C-G intercalation site in the core -(5'-d(CGTACG)-3')- sequence do not provide sufficient stability at the terminus of the oligonucleotide to change the sequence selectivity of hedamycin binding such that the guanine at the high affinity 5'-CGT site is favoured. However, in light of the results obtained for hedamycin binding to 5'-d(ACGTACGT)-3', this would seem unlikely. An alternative explanation for these observations may be that there is an additional sequence selectivity for hedamycin determined by the identity of the base preceding the binding site (ie. a lower affinity for hedamycin binding to 5'-CGT sites where the 5'-CGT site is preceded by T) based on the interactions between the ligand and neighbouring bases.

6.4 DC92-B BINDING TO DNA - RESULTS AND DISCUSSION

The antitumour agent DC92-B, is structurally similar to hedamycin, differing only by the presence of a hydroxyl group attached to the C-2'' position of the ring F aminosugar, and the bis(epoxide) sidechain which contains a *cis*-epoxide and terminal dimethyl groups (as opposed to a *trans*-epoxide and a terminal methyl group in hedamycin). Even though the substitution and stereochemistry of the bis(epoxide) sidechain is very different to that of hedamycin, previous studies have established that both ligands exhibit essentially the same sequence selectivity. That is a preference for alkylation of isolated guanines in 5'-pyGT (where py is one of the pyrimidine bases C or T), with 5'-CGT sites being preferred to 5'-TGT sites². In the present study, DC92-B binding reactions were conducted on the two hexamers 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' under the identical conditions to those in which the binding reactions to hedamycin were carried out in order to compare the results obtained to those of hedamycin.

6.4.1 DC92-B binding to 5'-d(CACGTG)-3'

Figure 6.34 shows the HPLC chromatogram of the DC92-B/5'-d(CACGTG)-3' reaction mixture after 1 hour. The peak eluting at 17.6 minutes was identified by ESI-MS to be the unbound oligonucleotide. The ESI-MS spectrum of the peak eluting at 23.7 minutes is given in figure 6.35. The spectrum shows two main peaks at m/z 855.1 and 1283.1 which correspond to the $[M-3H]^{3-}$ and $[M-2H]^{2-}$ ions respectively of a single-stranded 5'-d(CACGTG)-3'/DC92-B adduct ($M_{r(meas.)}=2568 \pm 1$ Da, $M_{r(calc.)}=2567.7$ Da monoisotopic mass).

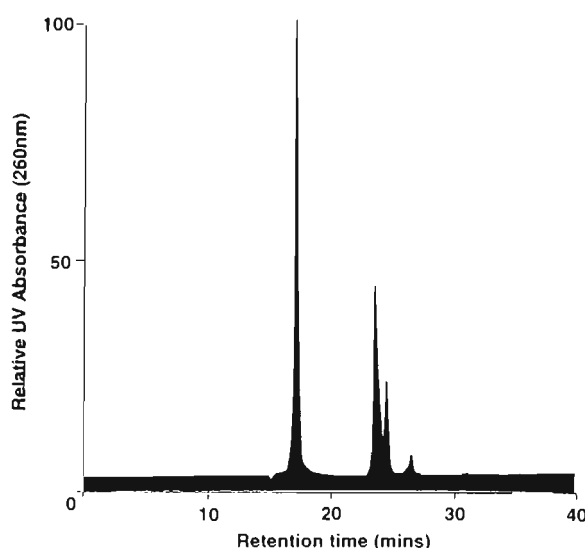


Figure 6.34: HPLC profile of the reaction between 5'-d(CACGTG)-3' and DC92-B.

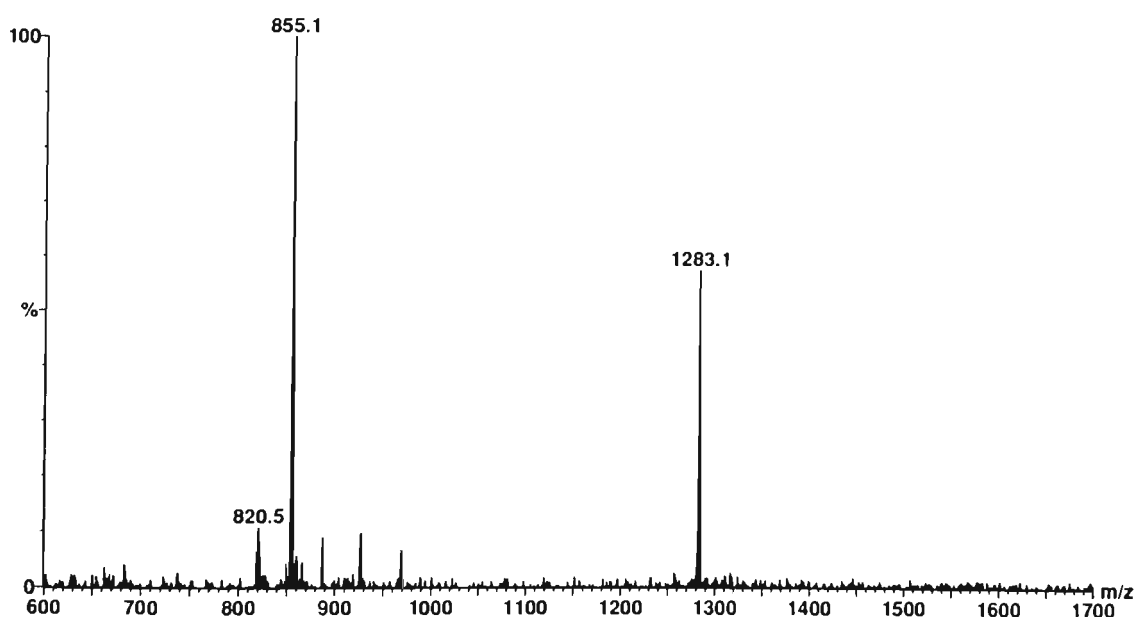


Figure 6.35: ESI mass spectrum of the peak eluting at 23.7 minutes in the HPLC profile of the 5'-d(CACGTG)-3'/DC92-B reaction mixture.

Figure 6.36 shows the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the adduct, and the fragmentation observed is illustrated in scheme 6.15. Consistent with the results obtained for the hedamycin/DNA adducts, the MS/MS spectrum of the 5'-d(CACGTG)-3'/DC92-B adduct is very simple, with the observation of a single fragmentation pathway about the site of alkylation. Loss of the alkylated nucleobase via N-glycosidic bond cleavage gives rise to the ions at m/z 819.2 and 1639.7 for the 2^- and 1^- ions respectively of the depurinated oligonucleotide (II). Cleavage of the ribose 3' C-O bond of the depurinated residue yields the 5' product ions at m/z 988.2 and the 3' fragment at m/z 650.0. The singly charged ion of the DC92-B+guanine adduct (I) is observed at m/z 926.2.

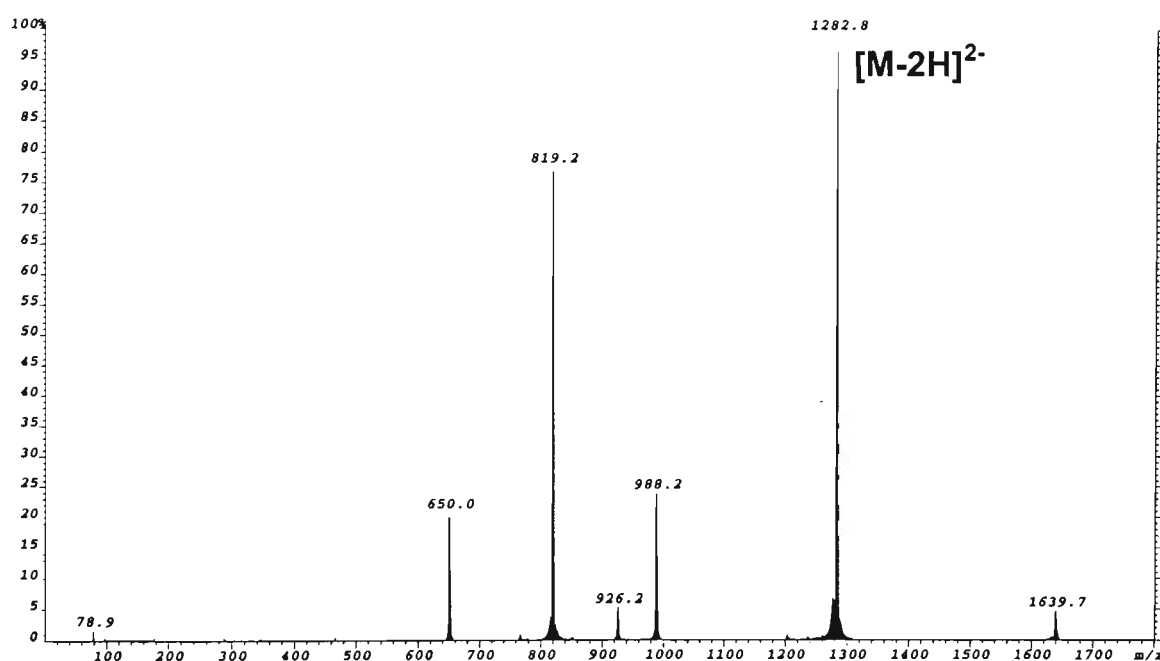
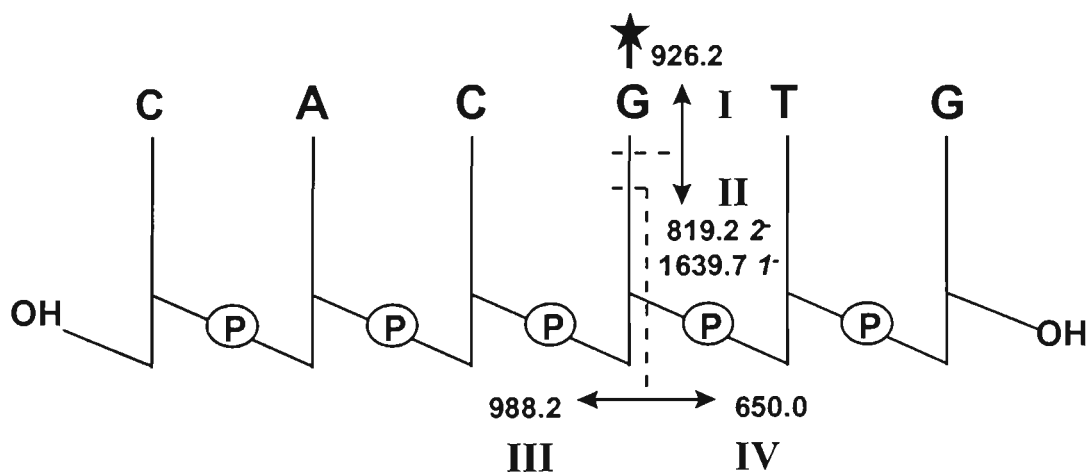


Figure 6.36: ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major adduct formed between DC92-B and 5'-d(CACGTG)-3'.



Scheme 6.15: Fragmentation pathway generated in the ESI-MS/MS spectrum of the major adduct formed between DC92-B and 5'-d(CACGTG)-3'.

The HPLC profile of the DC92-B/5'-d(CACGTG)-3' reaction (figure 6.34) also showed another peak alongside the peak for the major adduct at a retention time of 24.7 minutes. The ESI mass spectrum of a fraction of this peak is shown in figure 6.37. Two main ions are observed in the spectrum at m/z 854.9 and m/z 1282.7 corresponding to the $[M-3H]^{3-}$ and $[M-2H]^{2-}$ ions of a single-stranded 5'-d(CACGTG)-3'/DC92-B adduct ($M_{r(meas.)}=2568 \pm 1$ Da , $M_{r(calc.)}=2567.7$ Da monoisotopic mass). Figure 6.38 is the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of this adduct, and, as can be seen, the fragmentation in the spectrum is very similar to that of the major adduct. In addition to

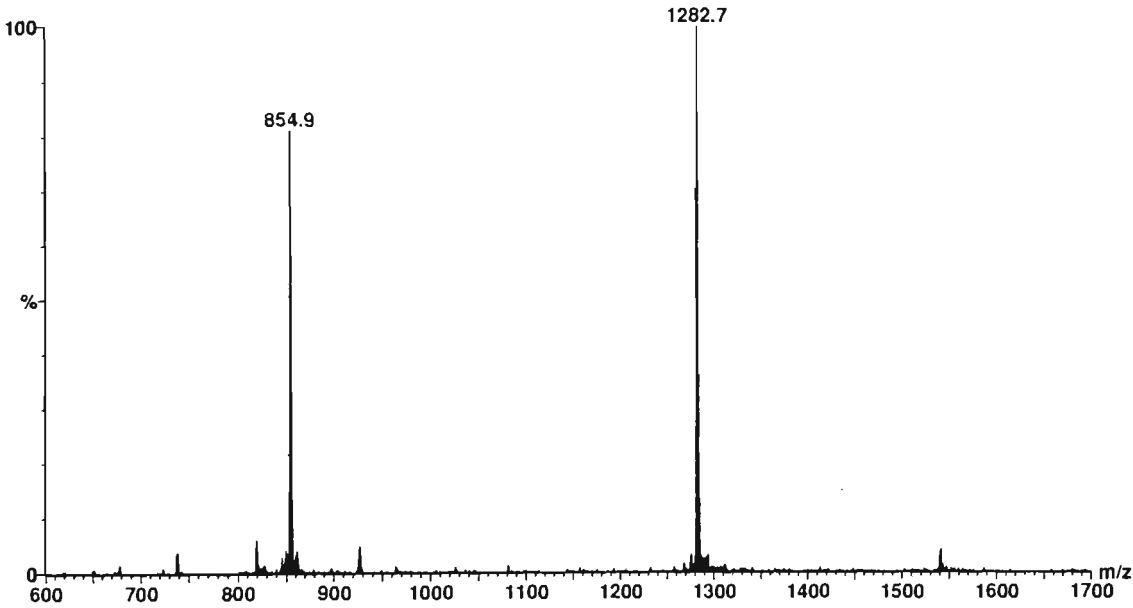


Figure 6.37: ESI mass spectrum of the peak eluting at 24.7 minutes in the HPLC profile of the 5'-d(CACGTG)-3'/DC92-B reaction mixture.

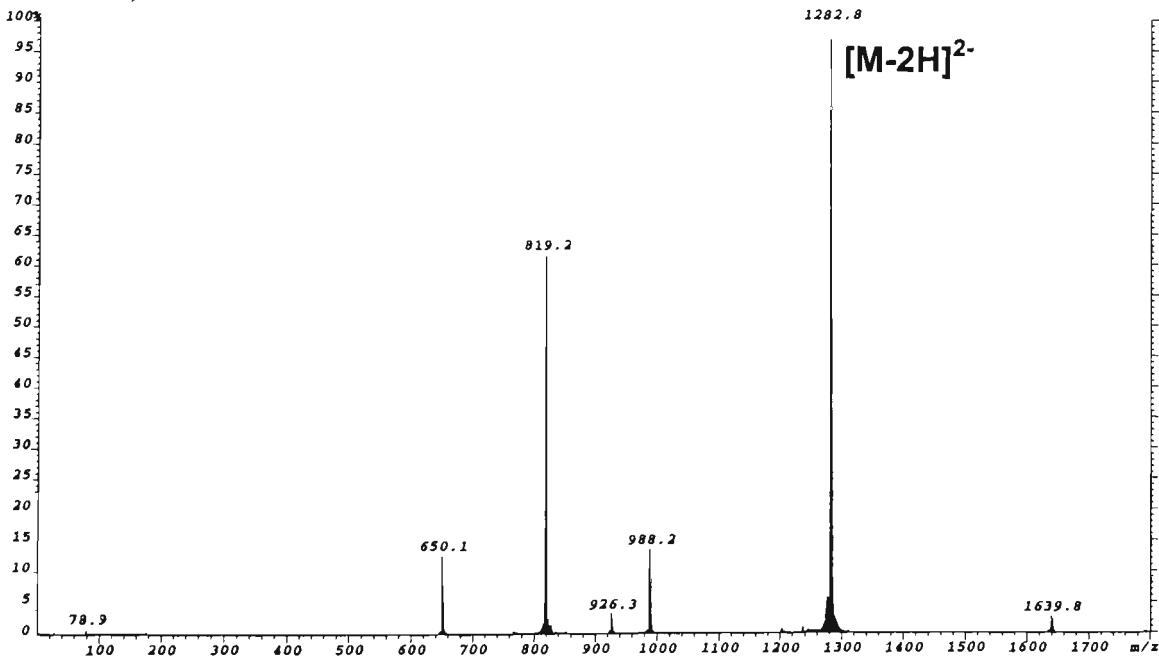


Figure 6.38: ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the minor adduct formed between DC92-B and 5'-d(CACGTG)-3' eluting at a retention time of 24.7 minutes on the HPLC profile of the reaction mixture.

the ions observed for the loss of alkylated guanine via N-glycosidic bond cleavage at the site of alkylation (at m/z 819.2 (2^- ion) and 1639.8 (1^- ion) for product ion type **II**, and at m/z 926.3 for product ion type **I**), cleavage of the ribose 3' C-O bond of the depurinated residue yields the same fragments at m/z 650.1 (**III**) and m/z 988.2 (**IV**). This fragmentation indicates guanine alkylation by DC92-B at the 5'-CGT site, identical to that observed for the major adduct illustrated in scheme 6.15. To investigate further the possible structural differences between the two adducts, ESI-MS/MS of both adducts was also carried out in the positive ion mode. The ESI-MS/MS spectra of the $[M+2H]^{2+}$ ions of the major and minor adducts are shown in figures 6.39(a) and (b) respectively.

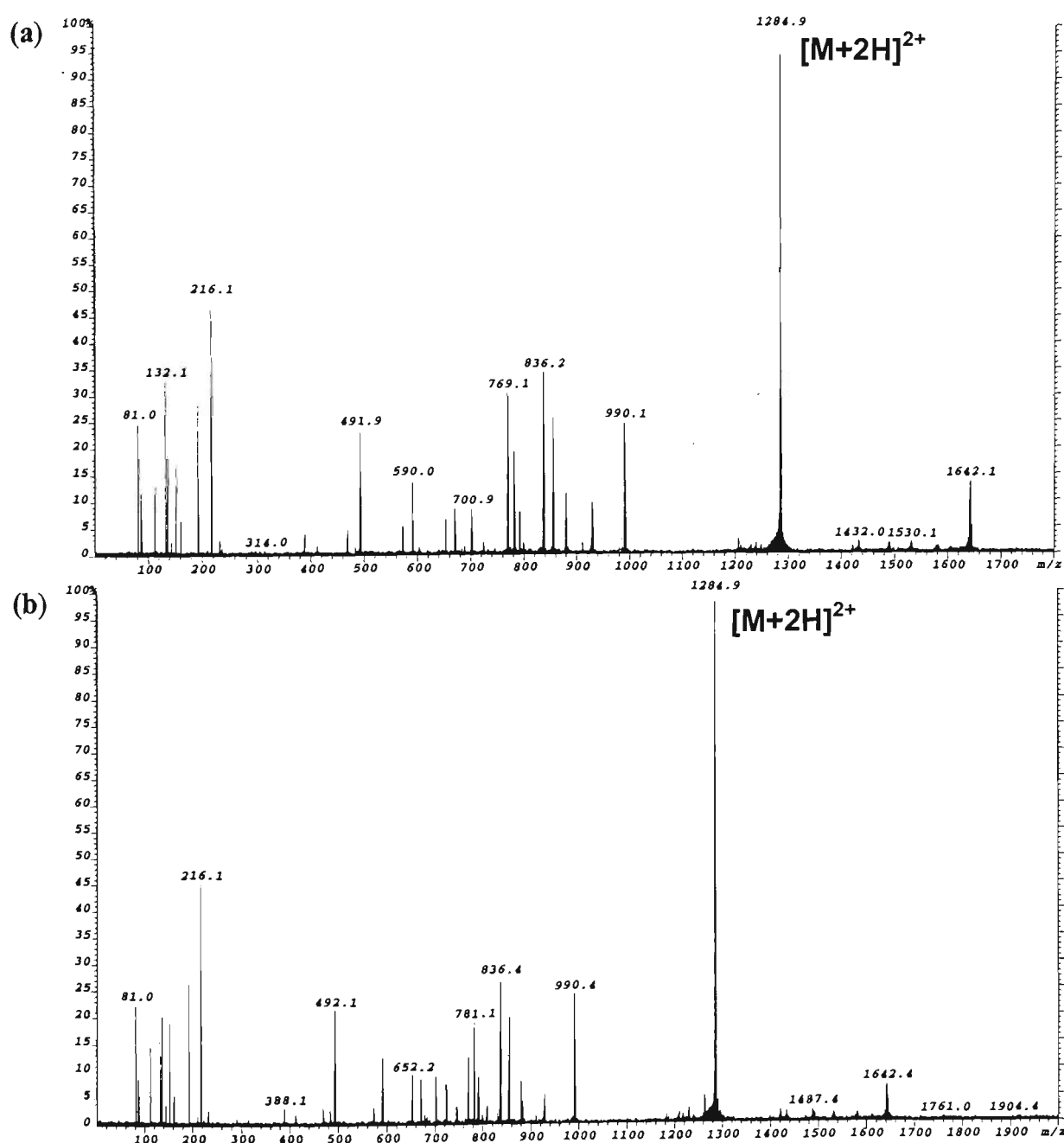


Figure 6.39: ESI-MS/MS spectra of the $[M+2H]^{2+}$ ions of the (a) major adduct and (b) minor adduct formed in the reaction between DC92-B and 5'-d(CACGTG)-3'.

Table 6.6: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+2H]^{2+}$ ion of the major adduct formed between DC92-B and 5'-d(CACGTG)-3' (m/z 1284.9).

m/z	Rel. Int.(%)	m/z (calc).	Assignment
1530.1	2.3	1530.3	loss of both cytosine and DC92-B+guanine
1490.1	2.2	1490.2	loss of both guanine and DC92-B+guanine
1432.0	2.4	1432.2	$[w_5\text{-GH}(4)+2H]^+$
879.0	11.8	879.1	$[a_4\text{-B}_4\text{H-CH}]^+$
780.9	20.1	781.1	$[c_3\text{-CH}]^+$ or $[(AC)+p+s+H]^+$
700.9	8.8	701.1	$[(AC)+p+H_2O+H]^+$
		or 701.1	$[a_3\text{-CH}]^+$
669.8	9.5	670.1	$[c_3\text{-CH-CH}]^+$ or $[A_{nt}+p+s+p+s+H]^+$
590.0	14.4	590.1	$[a_3\text{-CH-CH}]^+$ or $[A_{nt}+s+p+s+H]^+$
		or 590.1	$[A_{nt}+p+s+p+H_2O+H]^+$
572.0	5.7	572.1	$[A_{nt}+s+p+s-H_2O+H]^+$
		or 572.0	$[A_{nt}+p+s+p+H]^+$
491.9	24.0	492.1	$[c_2\text{-CH}]^+$ or $[A_{nt}+p+s+H]^+$
467.9	4.6	468.1	$[c_2\text{-AH}]^+$ or $[C_{nt}+p+s+H]^+$
412.0	1.9	412.1	$[a_2\text{-CH}]^+$ or $[A_{nt}+s+H]^+$
		or 412.0	$[A_{nt}+p-H_2O+H]^+$
387.9	4.3	388.1	$[a_2\text{-B}_2\text{H}]^+$
		or 388.0	$[C_{nt}+p+H_2O+H]^+$
232.0	3.5	232.1	$[sG\text{-H}_2O+H]^+$
216.1	48.9	216.1	$[sA\text{-H}_2O+H]^+$
192.0	29.3	192.1	$[sC\text{-H}_2O+H]^+$
161.0	7.1	161.0	$[s+p-H_2O+H]^+$
152.1	17.3	152.1	$[GH+H]^+$
136.0	19.0	136.1	$[AH+H]^+$
112.1	13.4	112.1	$[CH+H]^+$
81.0	25.3	81.0	$[s\text{-H}_2O+H]^+$

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), M_r =98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

The positive ion spectra are also very similar. These spectra comprise a more extensive range of product ions than was observed in the respective negative ion MS/MS spectra which are dominated by the observation of the oligonucleotide backbone fragments arising from cleavage about the site of modification. The formation of these ions is clearly not as favourable in corresponding positive ion MS/MS spectra, presumably as a consequence of there containing the negatively charged phosphate backbone. The singly charged ion for loss of alkylated guanine (**II**) is present at m/z 1642.1 although the corresponding doubly charged ion is absent. Product ions for cleavage of the ribose 3' C-O bond of the alkylated residue are present at m/z 990.1 (**III**) and m/z 652.1 (**IV**). In addition to these ions, numerous other oligonucleotide fragment ions are also observed. Identical types of product ions are present in the MS/MS spectrum of the

minor adduct (figure 6.39(b)) and so the assignments of these ions are not repeated here. The assignments of the major product ions appearing in the MS/MS spectrum of the major DC92-B/5'-d(CACGTG)-3' adduct (figure 6.39(a)) are given in table 6.6. The identity of these ions were confirmed from multiple stage mass spectrometry experiments conducted on the $[M+H]^+$ ion of unmodified 5'-d(CACGTG)-3' which was discussed in detail in chapter four, section 4.5.1. The MS/MS spectra of both of the DC92-B/5'-d(CACGTG)-3' adducts also show intense ions present at m/z 132, 769, 791, 836, 854. From analysis of the MS/MS spectrum of the $[DC92-B+guanine+H]^+$ product ion at m/z 928.1 generated by in-source collisional activation of the major and minor DC92-B/5'-d(CACGTG)-3' adducts (shown below in figures 6.40 and 6.42 respectively), it can be seen that these ions derive from further fragmentation of the $[DC92-B+guanine+H]^+$ ion. The absence of these product in negative ion MS/MS spectra is consistent with the presence of a formal positive charge on the alkylated purine owing to quaternisation of the base nitrogen upon electrophilic attack of the epoxide side chain of DC92-B on the N-7 of guanine. Therefore, upon positive ion MS/MS of the alkylated adduct, fragments of the alkylated adduct would be favourably generated owing to the presence of this formal positive charge on the alkylated residue. Comparing the positive ion MS/MS spectra acquired for the two adducts formed upon reaction of DC92-B with 5'-d(CACGTG)-3' (figure 6.39(a) and (b)), each spectrum shows an almost identical fragmentation pattern with no clear distinction evident in the types of product ions formed which may be indicative of a difference in the binding of DC92-B.

To probe further whether any difference in the two adducts could be determined, MS/MS was conducted on the singly charged product ion of the DC92-B+guanine adduct (m/z 928.1). Figure 6.40 shows the MS/MS spectrum of this product ion generated by in-source collisional activation of the major DC92-B/5'-d(CACGTG)-3' adduct thereby corresponding to this ion in the MS/MS spectrum of major adduct (figure 6.39(a)). The spectrum shows an ion at m/z 152.0 indicative of

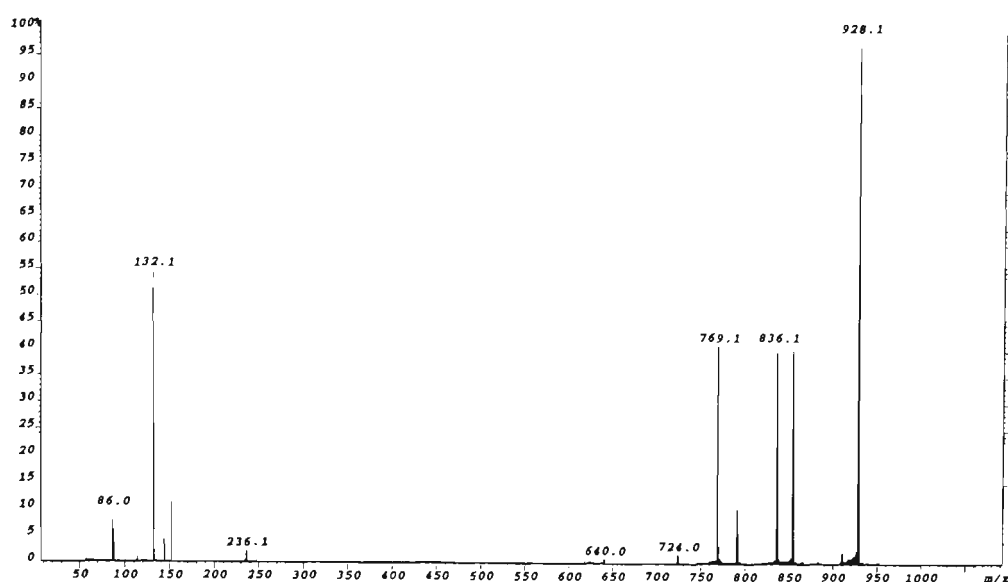
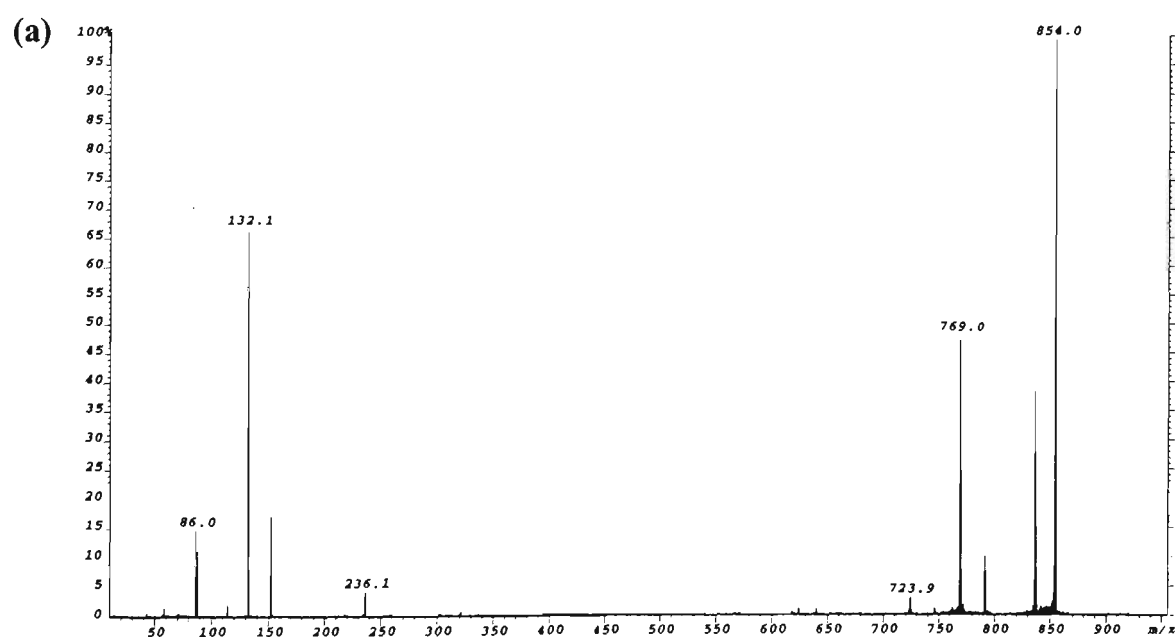


Figure 6.40: ESI-MS/MS spectrum of the $[DC92-B+guanine+H]^+$ adduct at m/z 928.1 generated by in-source collisional activation in the positive ion mode of the DC92-B/5'-d(CACGTG)-3' major adduct.

protonated guanine, GH_2^+ . The intense ion at m/z 132.1 may be attributed to the fragment formed from cleavage across the C6'-O bond and the C5'-C4' bond of ring E. This ion observed at m/z 236.1 corresponds to the mass of a guanine-bound fragment formed from cleavage of the bis-epoxide sidechain at the C14-C16 bond. MS/MS spectra were also acquired on the product ions at m/z 854.1, 836.1, 791.1, and 769.1 generated by in-source collisional activation of the major DC92-B/5'-d(CACGTG)-3' adduct in order to establish their identities. These are shown in figures 6.41(a) to (d) respectively.



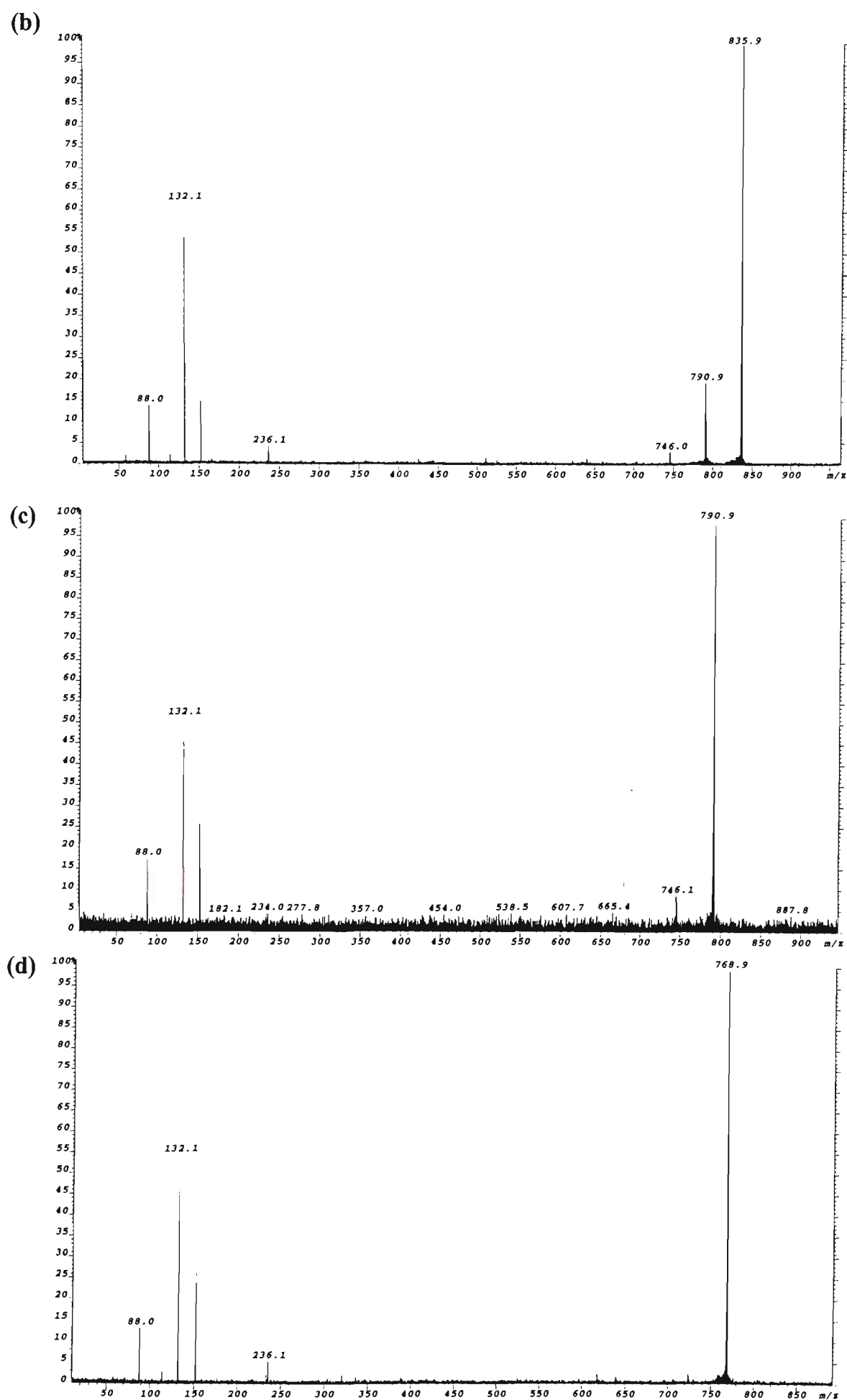
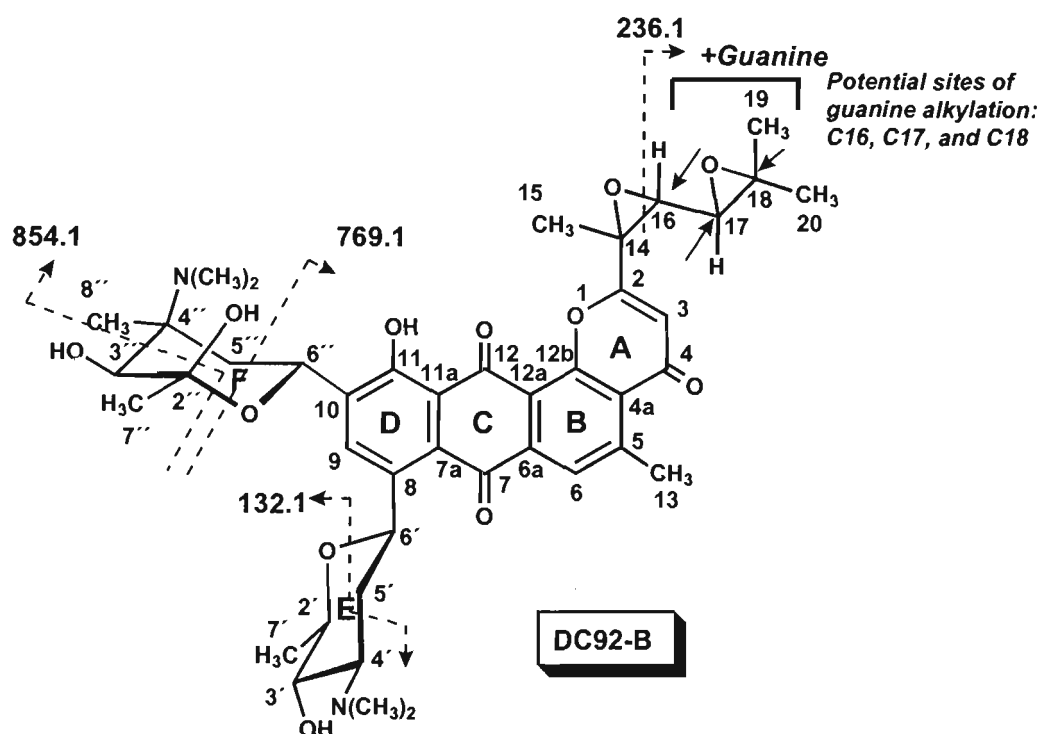


Figure 6.41: ESI-MS/MS spectra of the product ions at (a) m/z 854.0, (b) m/z 835.9, (c) m/z 790.9 and (d) m/z 768.9 generated by in-source collisional activation of the DC92-B/5'-d(CACGTG)-3' major adduct in the positive ion mode.

Each MS/MS spectrum shows an ion present at m/z 152.1 for protonated guanine (GH_2^+) in addition to the fragment at m/z 132.1. The ion at m/z 236.1 in the MS/MS spectrum of the $[\text{DC92-B}+\text{guanine}+\text{H}]^+$ ion is also observed in the MS/MS spectra of the product ions at m/z 854.1, 836.1, and 769.1. Fragmentation of the ion at m/z 854.0 (figure 6.41(a)) can be seen to give rise to the product ions at m/z 836.1, 791.1 and 769.1. This ion, which arises from the loss of 74 Da from the DC92-B+guanine product ion, may be owing to loss of a residue from ring F via cleavage across the $\text{C3''}-\text{C4''}$ and the $\text{C2''}-\text{O}$ bonds. The ion at m/z 835.9 corresponds to the loss of H_2O from the species at m/z 854.0, and the MS/MS spectrum of this species (figure 6.41(b)) gives rise to ions at m/z 790.9 and m/z 746.0. This corresponds to the consecutive loss of 45 Da in each case and may arise from loss of the dimethylamino groups from rings E and F. It can be seen that the MS/MS spectrum of the ion at m/z 790.9 also results in loss of 45 Da to give the ion at m/z 746.1. The ion at m/z 768.9, corresponds to the additional loss of 85 Da from the species at m/z 854.1 and may arise from the further cleavage of the $\text{C5''}-\text{C6''}$ bond of ring F. The fragmentation observed is illustrated in scheme 6.16.



Scheme 6.16: Fragmentation pathways observed in the ESI-MS/MS spectrum the $[\text{DC92-B}+\text{guanine}+\text{H}]^+$ product ion generated by in-source collisional activation of the major adduct formed between DC92-B and 5'-d(CACGTG)-3' in the positive ion mode.

Figure 6.42 shows the MS/MS spectrum of the $[\text{DC92-B+guanine+H}]^+$ product ion generated by in-source collisional activation of the minor adduct formed between DC92-B and 5'-d(CACGTG)-3' (figure 6.39(b)). Comparing this spectrum with that of the corresponding $[\text{DC92-B+guanine+H}]^+$ product ion from the major DC92-B/5'-d(CACGTG)-3' adduct (figure 6.40) it can be seen that both spectra show an almost identical fragmentation pattern with the exception of the presence of an ion at m/z 222.0 in the MS/MS spectrum of DC92-B+guanine product ion of the minor adduct (figure 6.42) in contrast to the presence of an ion at m/z 236.1 in the MS/MS spectrum of the $[\text{DC92-B+guanine+H}]^+$ product ion of the major adduct (figure 6.40). The ion at m/z 222.0 may be attributed to a guanine-bound fragment formed from cleavage of the bis-epoxide sidechain at the C16-C17 bond, as illustrated in scheme 6.17. This observation, is the only difference detected between the two DC92-B/5'-d(CACGTG)-3' adducts from MS/MS analysis, and could be evidence that guanine alkylation is occurring at a different position on the bis-epoxide sidechain. For example, the ion at m/z 236.1 in the MS/MS spectrum of the $[\text{DC92-B+guanine+H}]^+$ product ion of the major adduct may result from guanine alkylation at the C16 of the bis-epoxide sidechain whereas the fact that an ion at m/z 222.0 is detected in the case of the minor adduct may be indicative of guanine alkylation at the C17 carbon.

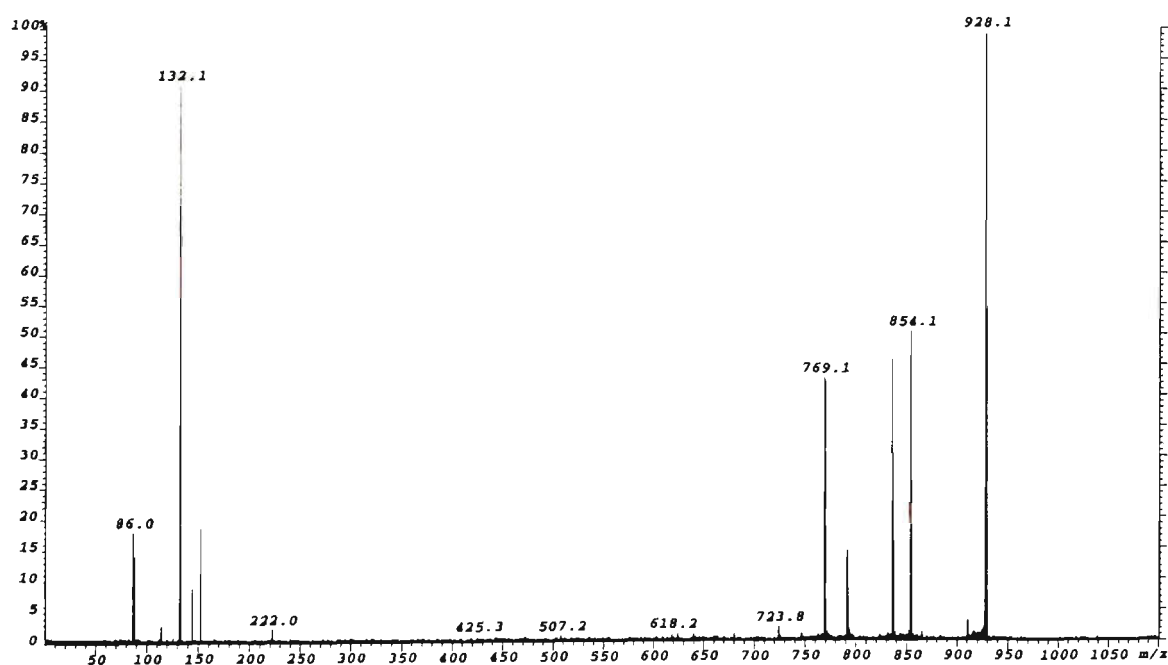
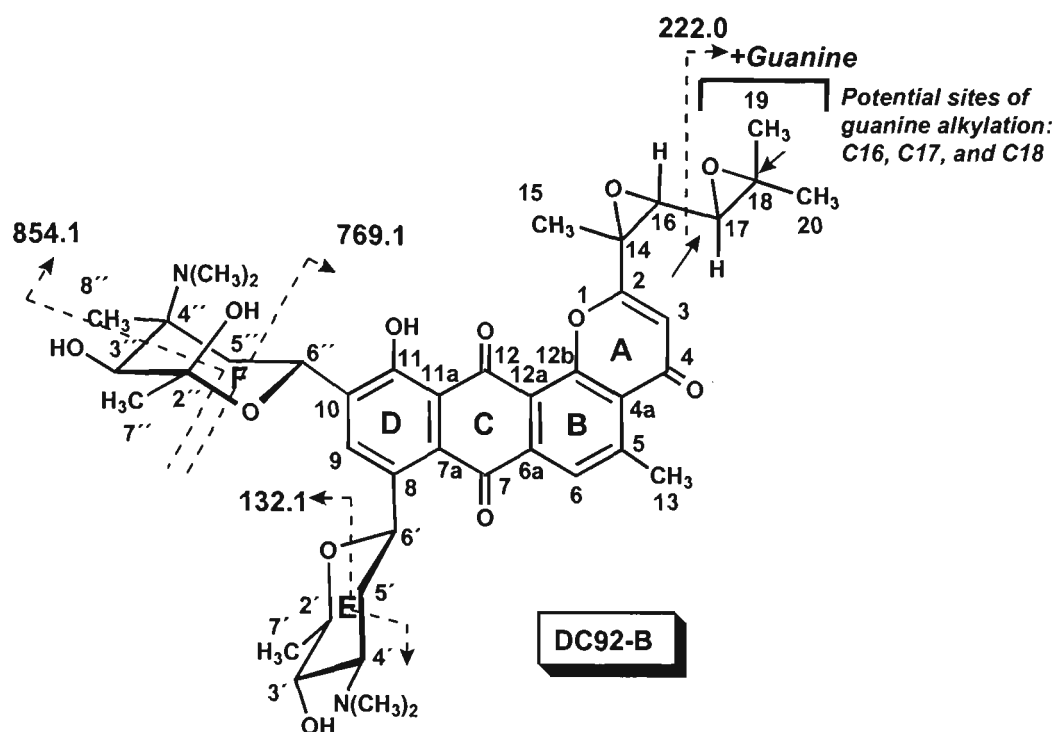


Figure 6.42: ESI-MS/MS spectrum of the $[\text{DC92-B+guanine+H}]^+$ ion generated by in-source collisional activation of the DC92-B/5'-d(CACGTG)-3' minor adduct.



Scheme 6.17: Fragmentation pathways observed in the ESI-MS/MS spectrum of the $[\text{DC92-B} + \text{guanine} + \text{H}]^+$ product ion generated by in-source collisional activation of the minor adduct formed between DC92-B and 5'-d(CACGTG)-3' in the positive ion mode.

6.4.2 DC92-B binding to 5'-d(CGTACG)-3'

Figure 6.43 shows the HPLC profile for the reaction between DC92-B and 5'-d(CGTACG)-3' obtained after the reaction was allowed to proceed for 1 hour. The peak eluting after 16.9 minutes is the unbound oligonucleotide. A fraction was collected for the major peak eluting at a retention time of 21.2 minutes and analysed by ESI-MS. The ESI-MS spectrum of this species is shown in figure 6.44, and yields two ions at m/z 855.3 and m/z 1283.2 corresponding to the $[\text{M}-3\text{H}]^{3-}$ and $[\text{M}-2\text{H}]^{2-}$ ions of a single-stranded DC92-B/5'-d(CGTACG)-3' adduct ($M_{\text{r}}(\text{meas.}) = 2569 \pm 1$ Da, $M_{\text{r}}(\text{calc.}) = 2567.7$ Da monoisotopic mass). Figure 6.45 is the ESI-MS/MS spectrum of the $[\text{M}-2\text{H}]^{2-}$ ion of this adduct (m/z 1282.8). Fragmentation of this ion produces the 2⁻ and 1⁻ ions for the loss of alkylated guanine (**II**) at m/z 819.1 and m/z 1639.2 respectively. The $[\text{DC92-B} + \text{guanine} + \text{H}]^+$ product ion is observed at m/z 926.2. Product ions indicative of cleavage of the ribose 3' C-O bond at the site of alkylation are the ions at m/z 386.0

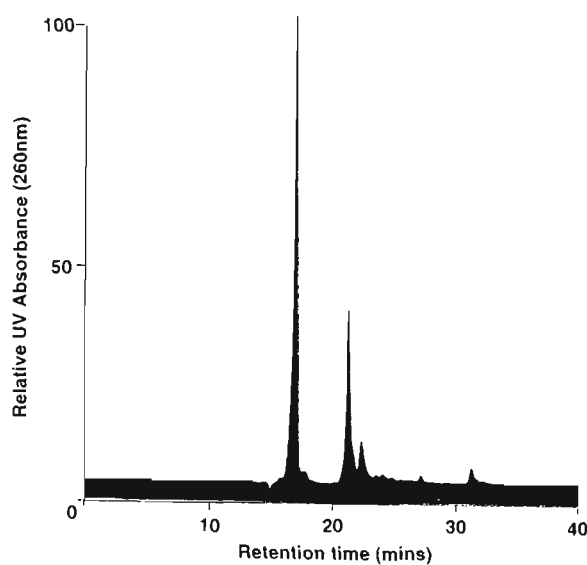


Figure 6.43: HPLC profile of the reaction between 5'-d(CGTACG)-3' and DC92-B.

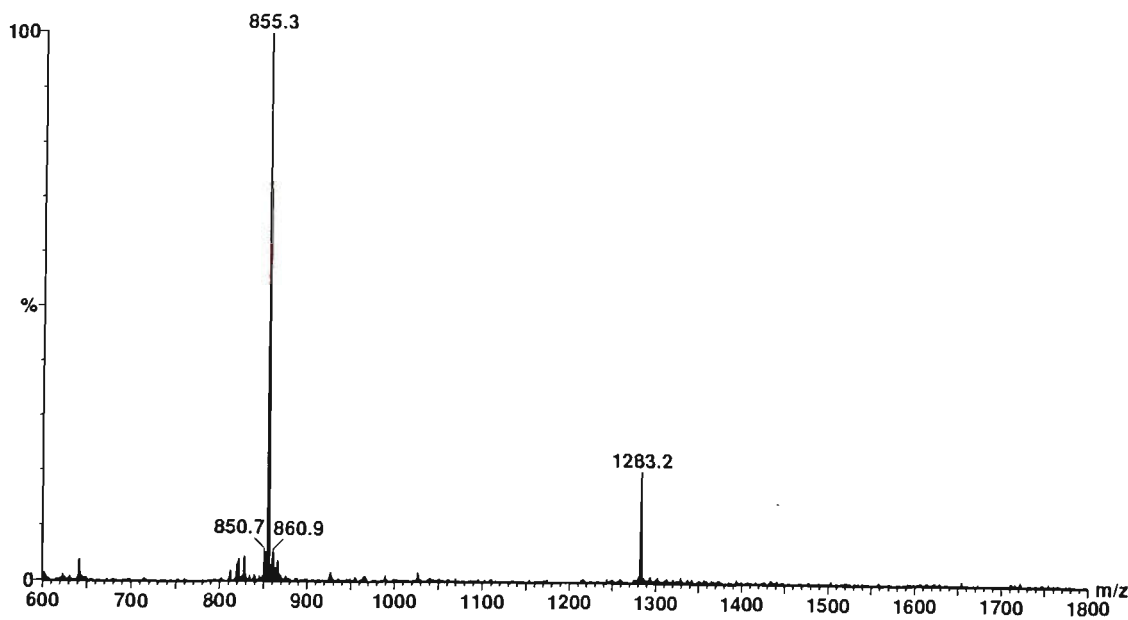


Figure 6.44: ESI mass spectrum of the peak eluting at 21.2 minutes in the HPLC profile of the 5'-d(CGTACG)-3'/DC92-B reaction mixture

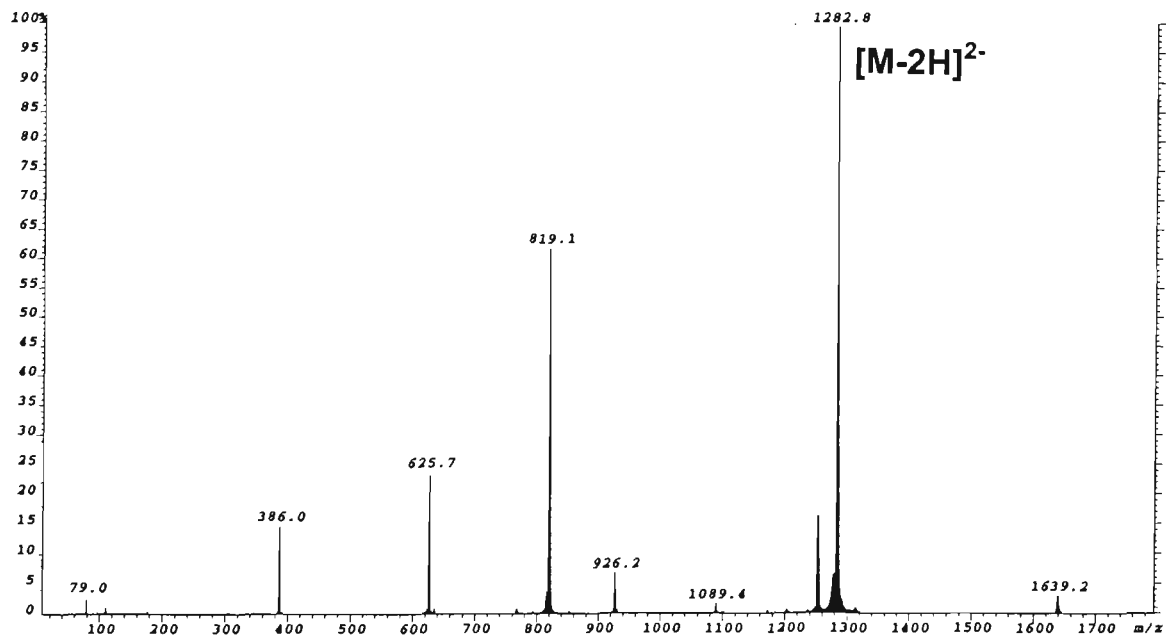
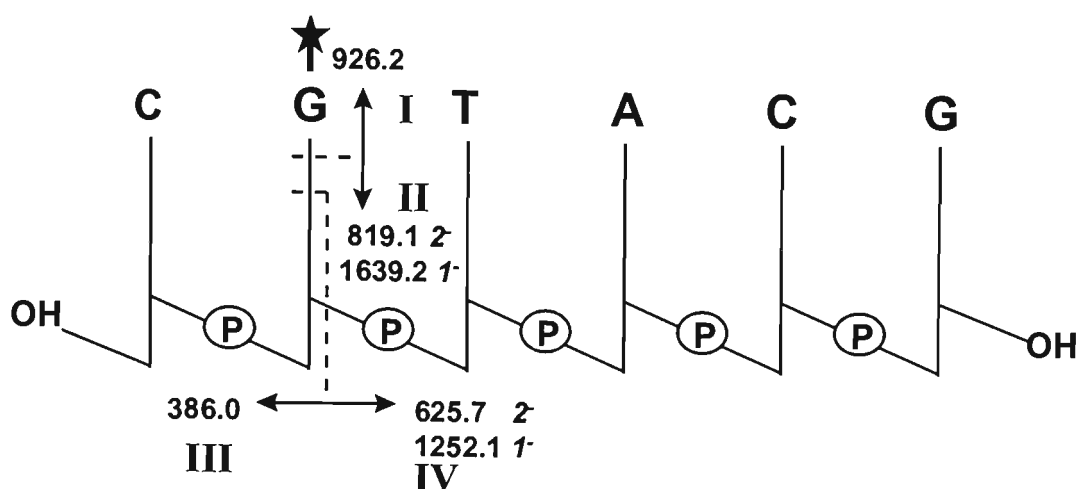


Figure 6.45: ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major adduct formed between DC92-B and 5'-d(CGTACG)-3' eluting at a retention time of 21.2 minutes on the HPLC profile of the reaction mixture.

(III), and at m/z 1252.1 and m/z 625.7 for the 1⁺ and 2⁺ ions, respectively, of product ion type IV. This dissociation pathway clearly establishes DC92-B alkylation of the guanine at the 5'-CGT site and is illustrated in scheme 6.18.



Scheme 6.18: Fragmentation pathways observed in the ESI-MS/MS spectrum of the major adduct formed between DC92-B and 5'-d(CGTACG)-3'.

6.4.3 DC92-B binding to 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3': Discussion

The results for DC92-B binding to the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' in each case showed alkylation to be highly sequence selective for the 5'-CGT site. These observations are consistent with sequencing studies of Prakash *et al.*² which showed DC92-B to exhibit a high sequence selectivity for alkylation of isolated guanines in 5'-pyGT (where py is one of the pyrimidine bases C or T), with 5'-CGT sites being preferred to 5'-TGT sites. Their results with DC92-B yielded the same sequence selectivity as for hedamycin^{2,3}, despite the structural differences present in the substitution and stereochemistry of the bis(epoxide) sidechains of these two ligands. It is noteworthy that the results obtained for the binding of DC92-B to the oligonucleotide 5'-d(CGTACG)-3' in the present work showed selectivity for the 5'-CGT site, when the results obtained for hedamycin binding to this oligonucleotide showed selectivity for the 5'-CG site rather than the 5'-CGT site. One possible explanation for the observed differences in selectivity of DC92-B may arise from a

difference in the location of guanine alkylation on the bis-epoxide sidechain compared with that of hedamycin. Whereas hedamycin has a *trans*-epoxide and a terminal methyl group, DC92-B, in contrast, contains a *cis*-epoxide and terminal dimethyl groups. Consequently, it may be that guanine N-7 alkylation by DC92-B would be expected to be more preferable at C16 or C17 rather than the terminal C18 owing to the steric hindrance associated with the presence of two methyl groups attached to C18. In comparison, guanine alkylation of the C18 of hedamycin, which contains only one bound methyl group, would be expected to be much more sterically favourable. Some support for this rationalisation may be found in the results obtained whereby two adducts are formed upon DC92-B binding to 5'-d(CACGTG)-3'. Positive ion MS/MS analysis of the [DC92-B+guanine+H]⁺ product ion generated by in-source collisional activation of each of the adducts, yielded fragmentation suggesting that the distinguishing feature between the two adducts is the site of guanine alkylation on the bis-epoxide sidechain. Although the precise site of binding of the ligand to guanine can not be deduced from the MS/MS data obtained in the present work, the results suggest that this difference may arise from guanine alkylation at the C16 carbon in formation of the major DC92-B/5'-d(CACGTG)-3' adduct, with the minor adduct corresponding to alkylation of the C17 carbon. Further work in this area is however required in order to conclusively establish the structural basis of these observed differences in the binding of hedamycin and DC92-B.

6.5 *PhenC₆Br* - RESULTS AND DISCUSSION

The synthetic ligand *n*-bromohexylphenanthridinium bromide (phenC₆Br) has been shown by previous sequencing studies to preferentially alkylate guanines at 5'-GT-3' sequences and in runs of guanines, [(Gp)_n]¹. The similarity in sequence selectivity of this synthetic derivative to the naturally occurring antitumour antibiotics, hedamycin and DC92-B was of interest in the present study owing to the structural differences between the two classes of compounds. That is the ligands, hedamycin and

DC92-B are anthraquinone-based natural products that contain bis-epoxide alkylating side-chains, and phenC₆Br is a synthetic derivative of phenanthridine and contains an alkylating moiety linked by a polymethylene chain to a quaternized phenanthridine ring. For the purposes of comparing the sequence selectivity of phenC₆Br binding to that of hedamycin and DC92-B, binding studies were conducted on the two oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

6.5.1 PhenC₆ Binding to 5'-d(CACGTG)-3'

The reaction between phenC₆Br and 5'-d(CACGTG)-3' was monitored by HPLC and was judged to be complete after approximately 18 hours. The chromatogram taken at this time point is shown in figure 6.46. Fractions collected from the HPLC profile were initially analysed by ESI-MS in order to ascertain the identities of each species. The peak eluting at 19.1 minutes was found to be the unbound oligonucleotide and the peak at 46.2 minutes the unreacted ligand (spectra not shown). Note that the retention time for the unbound 5'-d(CACGTG)-3' in this case is slightly higher than that observed previously since a shallower gradient was employed for purification of phenC₆Br/oligonucleotide reaction mixtures. The peaks that eluted between 29 and 37 minutes correspond to ligand by-products formed during sample incubation. The formation of one major adduct is observed at a retention time of 23.3 minutes on the HPLC profile and the ESI mass spectrum of the fraction collected for this peak is shown in figure 6.47. The spectrum shows ions present at m/z 683.3 and m/z 1025.5 consistent with the $[M-3H]^{3-}$ and $[M-2H]^{2-}$ ions of a single-stranded phenC₆/5'-d(CACGTG)-3' adduct ($M_{r(meas.)}=2053 \pm 1$ Da, $M_{r(calc.)}=2052.5$ Da monoisotopic mass). The molecular weight of the neutral adduct measured from the ions observed in the ESI-MS spectrum corresponds to the species $[\{\text{phenC}_6/5'\text{-d(CACGTG)-3'}\}^{2+}\text{-2H}]^0$ in which the phenC₆ ligand is bound minus the two bromines. The overall charge of 2⁺ on the adduct arises from a positive charge present on the ligand owing to the quaternary nitrogen, and a formal positive charge located on the alkylated purine.

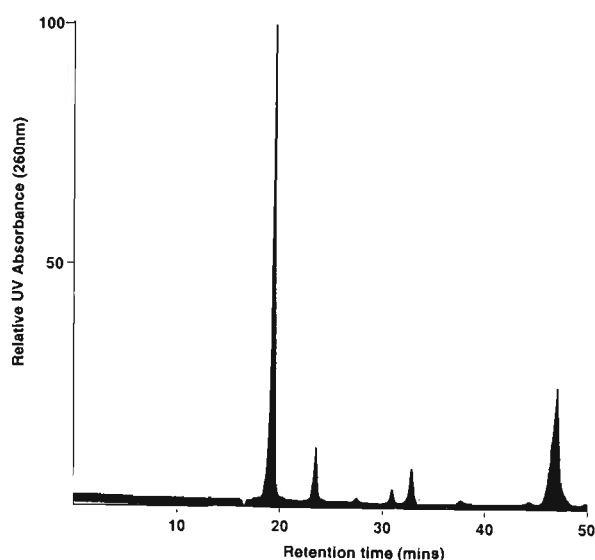


Figure 6.46: HPLC profile of the reaction between 5'-d(CACGTG)-3' and phenC₆Br.

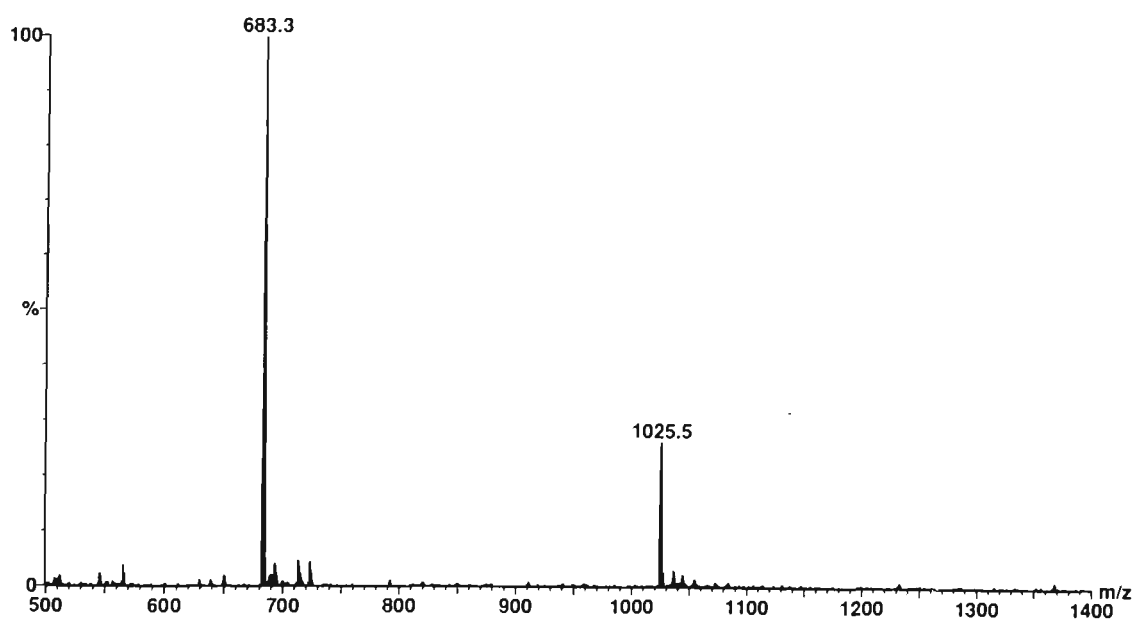


Figure 6.47: ESI mass spectrum of the peak eluting at 23.3 minutes in the HPLC profile of the phenC₆/5'-d(CACGTG)-3' reaction mixture

The ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the phenC₆/5'-d(CACGTG)-3' adduct (m/z 1025.5) is shown in figure 6.48. A single fragmentation pathway is observed which yields three main product ions in the MS/MS spectrum. Cleavage of the N-glycosidic bond of the alkylated guanine yields the 2⁻ ion of the depurinated oligonucleotide at m/z 819.2 (II). Subsequent cleavage of the ribose 3' C-O bond of the alkylated residue gives rise to the 5' terminal ion at m/z 988.2 (III) and the 3' terminal ion at m/z 650.1, therefore enabling facile identification of the site of alkylation as the guanine at the 5'-CGT site (shown in scheme 6.19).

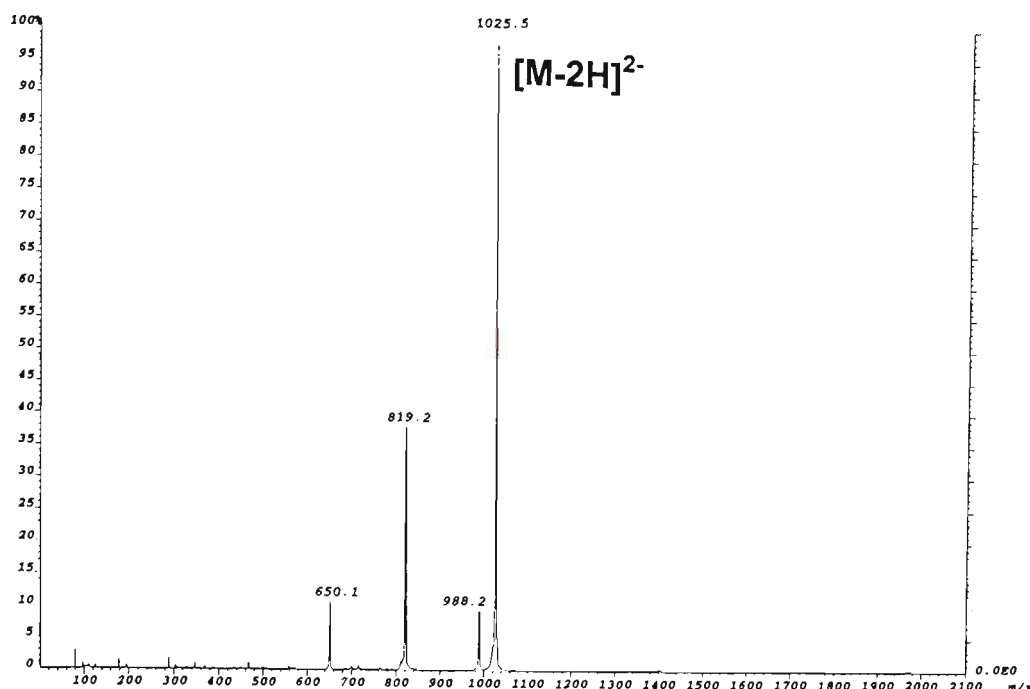
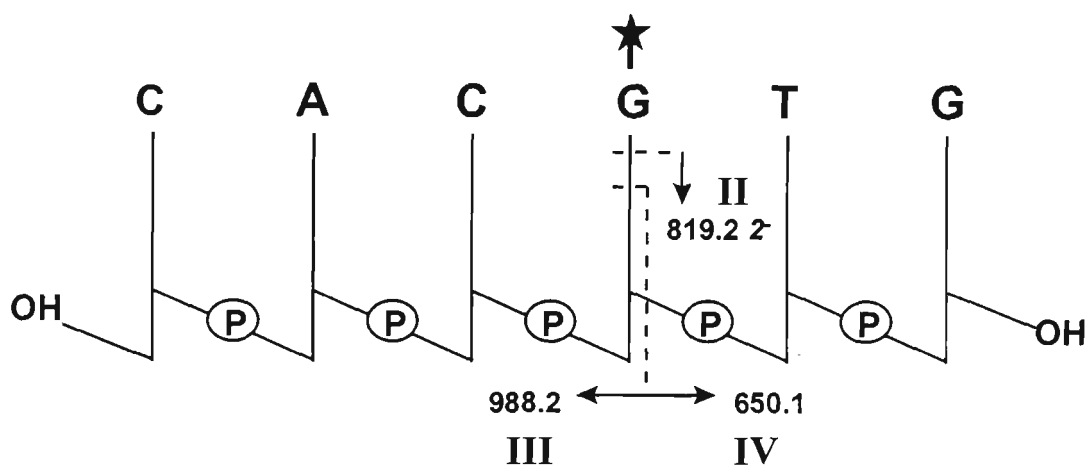


Figure 6.48: ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major adduct formed between phenC₆Br and 5'-d(CACGTG)-3' (m/z 1025.5).



Scheme 6.19: Fragmentation pathways observed in the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major adduct formed between phenC₆Br and 5'-d(CACGTG)-3'.

The absence of the anion for the ligand+guanine adduct (I) is understandable given the presence of a positive charge on the phenC₆ ligand from the quaternised nitrogen and the formal positive charge on the alkylated guanine. Figure 6.49 shows the MS/MS spectrum of the 2^+ ion of the phenC₆/5'-d(CACGTG)-3' adduct. As observed in the positive ion MS/MS spectra of the DC92-B/5'-d(CACGTG)-3' adducts, the MS/MS spectrum acquired in the positive ion mode yields a markedly different fragmentation pattern to that observed from the negatively charged precursor with a more extensive range of ions generated. The most abundant ion in the spectrum at m/z 413.3 corresponds to the mass of a protonated $[(\text{phenC}_6 + \text{guanine})^{2+} - \text{H}]^+$ (I) adduct.

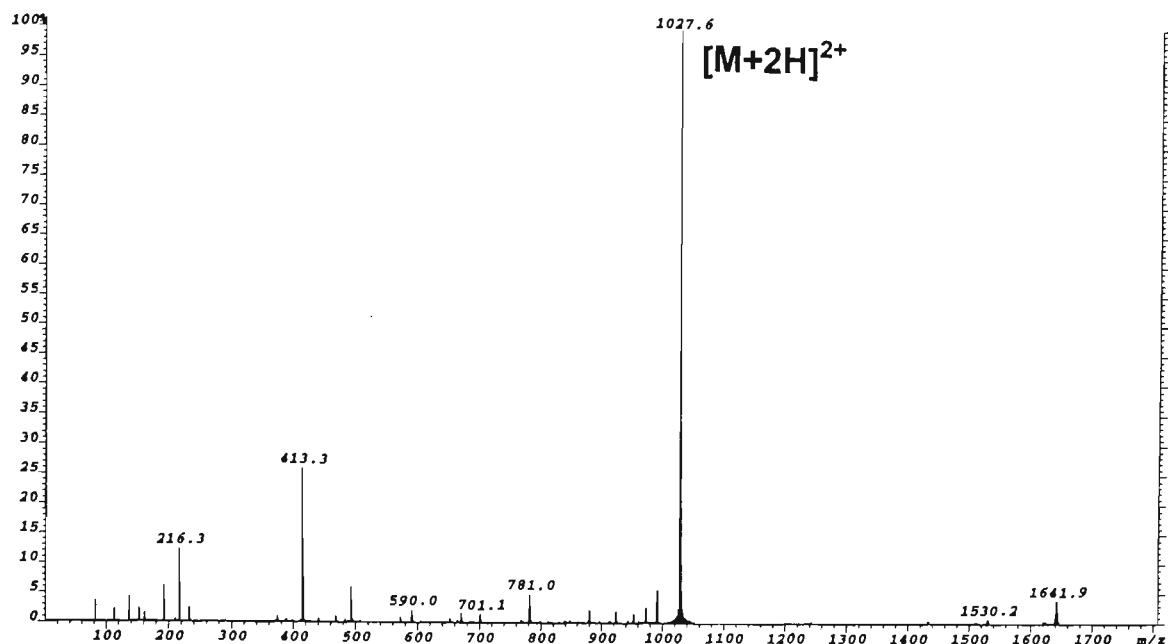


Figure 6.49: ESI-MS/MS spectrum of the $[M+2H]^{2+}$ ion of the major adduct formed between phenC₆Br and 5'-d(CACGTG)-3' (m/z 1027.6).

Table 6.7: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the 2⁺ ion of the major adduct formed between phenC₆Br and 5'-d(CACGTG)-3'.

m/z	Rel. Int.(%)	m/z(calc).	Assignment
1530.2	0.9	1530.3	loss of both cytosine and DC92-B+guanine
1432.0	0.8	1432.2	$[w_5\text{-GH}(4)+2H]^+$
972.2	3.4	972.1	$[(TG)+p+s+p+s\text{-H}_2O+H]^+$
879.3	3.3	879.1	$[a_4\text{-B}_4H\text{-CH}]^+$
781.0	5.0	781.1	$[c_3\text{-CH}]^+$ or $[(AC)+p+s+H]^+$
701.1	1.6	701.1	$[(AC)+p+H_2O+H]^+$
		or 701.1	$[a_3\text{-CH}]^+$
670.1	1.7	670.1	$[c_3\text{-CH-CH}]^+$ or $[A_{nt}+p+s+p+s+H]^+$
590.0	2.5	590.1	$[a_3\text{-CH-CH}]^+$ or $[A_{nt}+s+p+s+H]^+$
		or 590.1	$[A_{nt}+p+s+p+H_2O+H]^+$
572.0	1.0	572.1	$[A_{nt}+s+p+s\text{-H}_2O+H]^+$
		or 572.0	$[A_{nt}+p+s+p+H]^+$
492.2	6.2	492.1	$[c_2\text{-CH}]^+$ or $[A_{nt}+p+s+H]^+$
468.2	1.3	468.1	$[c_2\text{-AH}]^+$ or $[C_{nt}+p+s+H]^+$
440.3	0.7	440.1	$[a_4\text{-B}_4H\text{-CH+H}]^{2+}$
388.1	1.2	388.1	$[a_2\text{-B}_2H]^+$
		or 388.0	$[C_{nt}+p+H_2O+H]^+$
232.2	2.6	232.1	$[sG\text{-H}_2O+H]^+$
216.3	12.5	216.1	$[sA\text{-H}_2O+H]^+$
192.2	6.4	192.1	$[sC\text{-H}_2O+H]^+$
161.2	1.6	161.0	$[s+p\text{-H}_2O+H]^+$
152.1	2.4	152.1	$[GH+H]^+$
136.1	4.5	136.1	$[AH+H]^+$
112.1	2.3	112.1	$[CH+H]^+$
81.0	4.0	81.0	$[s\text{-H}_2O+H]^+$

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), M_r =98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Figure 6.50 shows the MS/MS spectrum of this product ion generated by in-source collision of the adduct. Fragmentation of this product ion can be seen to give rise to an ion at m/z 152.1 which is indicative of protonated guanine, GH_2^+ . The intensity of the ion at m/z 180.1 suggests this to arise from another aromatic system, presumably from cleavage of the N-C bond of the ligand yielding protonated phenanthridine. The ion at m/z 234.2 corresponds to the mass of the complementary fragment ion formed from the loss of the phenanthridine ring from the phenC_6 +guanine adduct. These observations are consistent with the structure of the ion at m/z 413.3 being that of the phenC_6 +guanine adduct. The complementary ion arising from loss of the alkylated base from the $\text{phenC}_6/5'\text{-d(CACGTG)-3'}$ adduct, ie. the 1^+ ion of the depurinated oligonucleotide, is also observed in the MS/MS spectrum of the 2^+ ion of the $\text{phenC}_6/5'\text{-d(CACGTG)-3'}$ adduct (figure 6.49) at m/z 1641.9. Subsequent cleavage of the ribose C-O bond to the 3' side of the alkylated base only gives rise to the 5' terminal product ion **III** at m/z 990.4. The assignments of the various other oligonucleotide fragments observed in the spectrum are given in table 6.7.

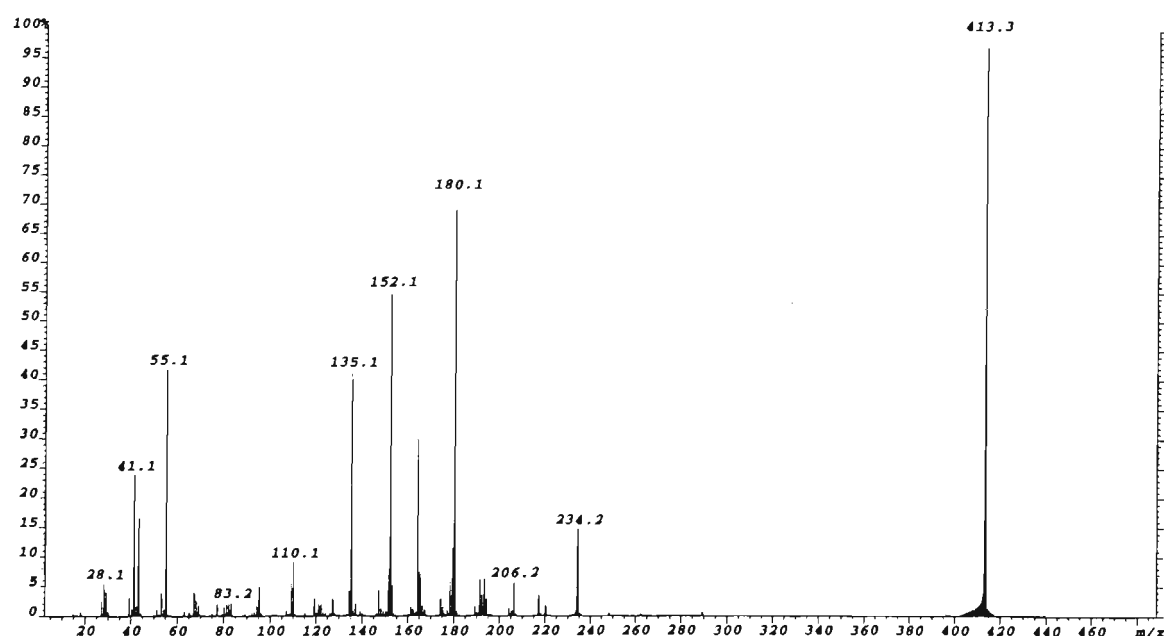


Figure 6.50: ESI-MS/MS spectrum of the ion at m/z 413.3 generated by in-source collisional activation of the major adduct formed between phenC_6Br and $5'\text{-d(CACGTG)-3'}$ in the positive ion mode.

Comparing the negative and positive ion MS/MS spectra of the doubly charged ion of the $\text{phenC}_6/5'\text{-d(CACGTG)-3'}$ adduct, it can be seen that both spectra show the

same major fragmentation pathway which is the loss of the alkylated base. However, whereas the negative ion MS/MS spectrum yields ions solely for the oligonucleotide cleavage products resulting from this fragmentation, the positive ion MS/MS spectrum, in contrast, is dominated by the ion for the alkylated base with a range of other oligonucleotide fragments present. These observations can be attributed to the location of the positive charge on the nitrogen of the phenC₆ ligand which makes the production of ligand-bound fragments in the positive ion mode more favourable, and consequently accounts for the absence of such ions in the negative ion mode.

6.5.2 *PhenC₆ Binding to 5'-d(CGTACG)-3'*

The HPLC profile obtained for the reaction between phenC₆Br and 5'-d(CGTACG)-3' after the reaction was allowed to proceed for 18 hours is shown in figure 6.51. ESI-MS analysis of fractions collected for peaks present in HPLC profile revealed the peak eluting at 18.3 minutes to be that of the unbound oligonucleotide, with the unreacted ligand eluting after 46.2 minutes. As observed in the HPLC profile of the reaction between phenC₆Br and 5'-d(CACGTG)-3', species owing to side reactions of the ligand are once again observed to be eluted between 29 and 37 minutes. The HPLC profile shows a peak present at a retention time of 23.3 minutes and another peak which is approximately 45% of the intensity of the larger peak at a retention time of 22.8 minutes. The ESI mass spectrum of the more abundant peak (23.3 minutes) is shown in figure 6.52. The spectrum shows the presence of two main ions at m/z 683.3 and 1025.3 which correspond to the $[M-3H]^{3-}$ and $[M-2H]^{2-}$ ions of a phenC₆/5'-d(CGTACG)-3' adduct. ($M_{r(meas.)}=2053 \pm 1$ Da, $M_{r(calc.)}=2052.5$ Da monoisotopic mass for the neutral species $[\{\text{phenC}_6/5'\text{-d(CGTACG)-3'}\}^{2+}\text{-2H}]^0$). ESI-MS analysis of the fraction collected for the minor species eluting at a retention time of 22.8 minutes (not shown) also yielded a molecular weight of $M_{r(meas.)}=2053 \pm 1$ Da indicative of the formation of a phenC₆/5'-d(CGTACG)-3' adduct.

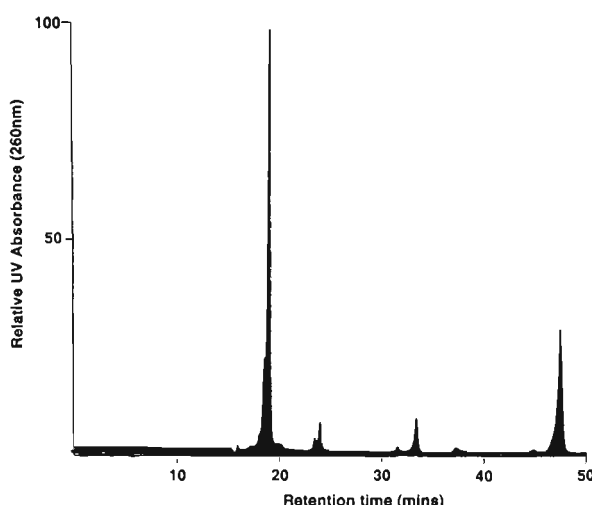


Figure 6.51: HPLC profile of the reaction between 5'-d(CGTACG)-3' and phenC₆Br.

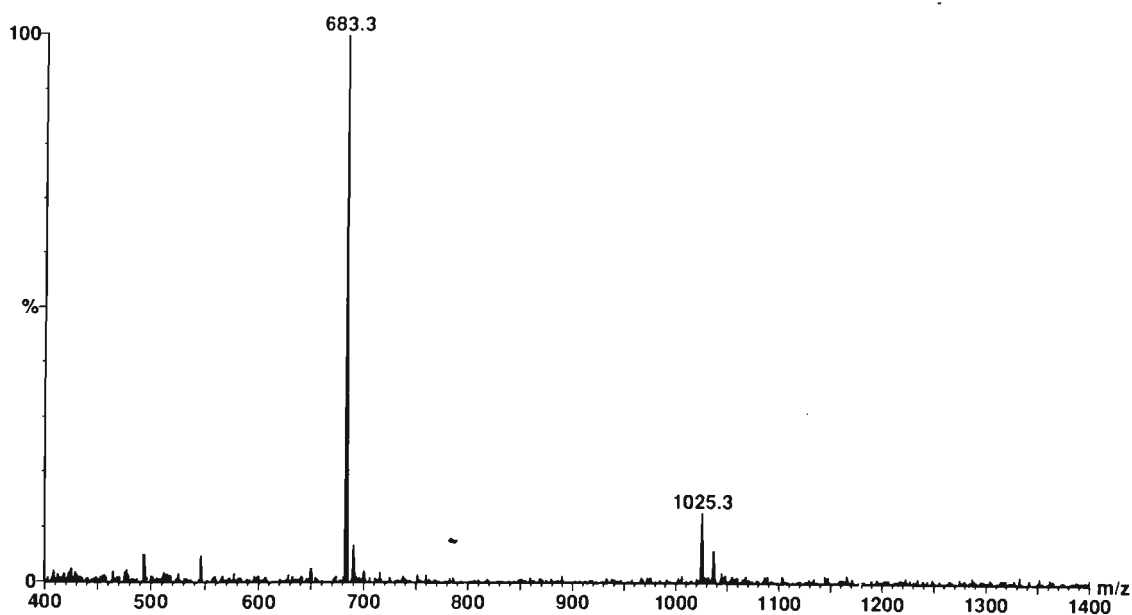


Figure 6.52: ESI mass spectrum of the peak eluting at 23.3 minutes in the HPLC profile of the phenC₆/5'-d(CGTACG)-3' reaction mixture.

Figure 6.53 shows the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the major phenC₆/5'-d(CGTACG)-3' adduct formed (m/z 1025.4) which elutes at a retention time of 23.3 minutes in the HPLC profile of the reaction mixture (figure 6.51). Fragmentation of this adduct gives rise to an ion at m/z 819.3 for the 2^- ion for the oligonucleotide fragment (II) resulting from the loss of alkylated guanine. Cleavage of the ribose 3' C-O bond at the depurinated site produces the ion at m/z 386.2 for the 5' terminal fragment (III), and the ions at m/z 625.6 and 1252.2 for the 2^- and 1^- ions respectively of the 3' terminal fragment (IV). The formation of these product ions from MS/MS fragmentation of the adduct is illustrated in scheme 6.20. These ions establish

the site of phenC₆ alkylation on 5'-d(CGTACG)-3' as the guanine at the 5'-CGT site. As observed in the negative ion MS/MS spectrum of the phenC₆/5'-d(CACGTG)-3' adduct, no product ion is detected for the alkylated base in the MS/MS spectrum of the [M-2H]²⁻ ion of phenC₆/5'-d(CGTACG)-3'.

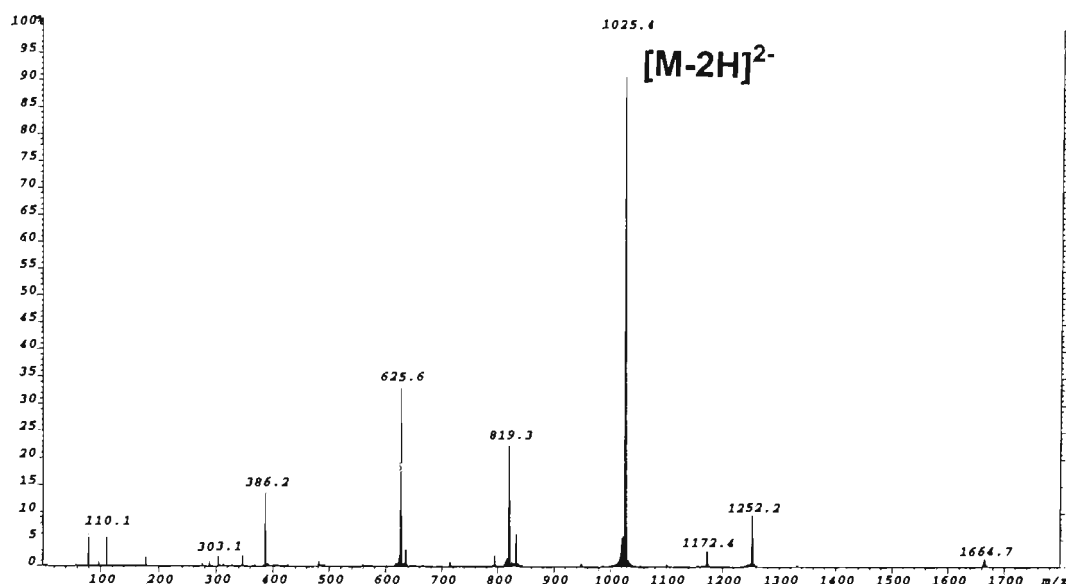
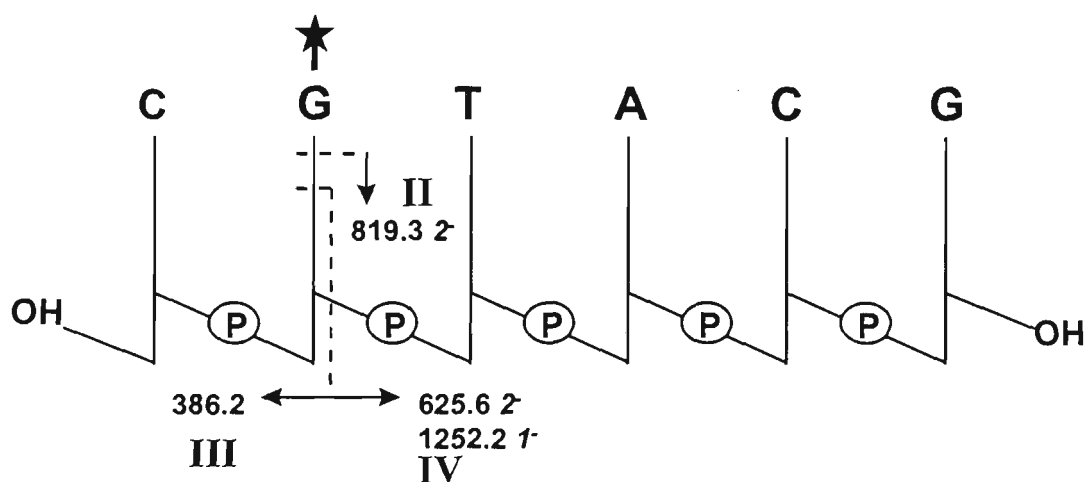


Figure 6.53: ESI-MS/MS spectrum of the [M-2H]²⁻ ion of the major adduct formed between phenC₆Br and 5'-d(CGTACG)-3' (*m/z* 1025.4).



Scheme 6.20: Fragmentation pathways observed in the ESI-MS/MS spectrum of the major adduct formed between phenC₆Br and 5'-d(CGTACG)-3'.

The positive ion MS/MS spectrum of the [M+2H]²⁺ ion of the major phenC₆/5'-d(CGTACG)-3' adduct (*m/z* 1027.6) is given in figure 6.54. The ion at *m/z* 413.3 corresponds to the [(phenC₆+guanine)²⁺-H]⁺ adduct. The MS/MS spectrum of this product ion produced the same results as that observed in the MS/MS spectrum of the analogous product ion at *m/z* 413.3 generated from the fragmentation of the phenC₆/

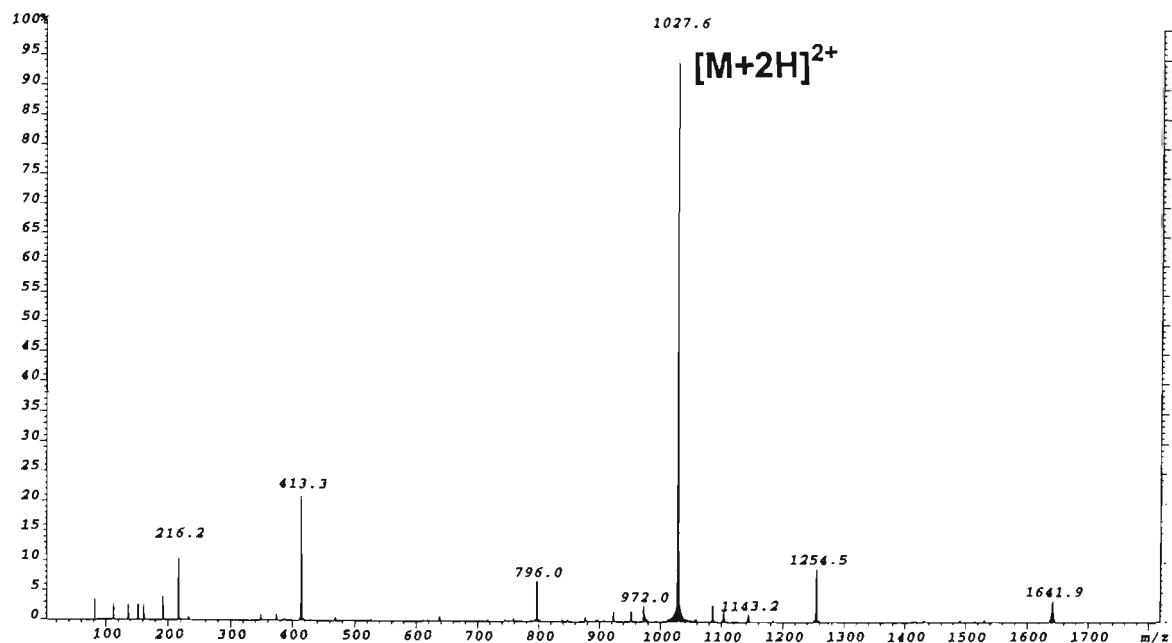


Figure 6.54: ESI-MS/MS spectrum of the $[M+2H]^{2+}$ ion of the major adduct formed between phenC6Br and 5'-d(CGTACG)-3' (m/z 1027.6).

Table 6.8: Assignment of oligonucleotide product ions observed in the ESI-MS/MS spectrum of the 2^{+} ion of the major adduct formed between phenC6Br and 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1530.3	0.5	1530.3	loss of both cytosine and DC92-B+guanine
1490.1	0.5	1432.2	$[w_5\text{-GH}(4)+2H]^+$
1183.5	0.4	1183.2	$[z_5\text{-GH-GH}]^+$ or $[a_5\text{-CH-GH}]^+$
		or 1183.1	$[(TAC)+p+s+p+H_2O+H]^+$
1143.2	1.6	1143.2	$[w_4\text{-CH}+2H]^+$ or $[d_4\text{-CH}+2H]^+$ or $[(GTA)+p+s+H_2O+H]^+$
1103.3	2.0	1103.2	$[w_4\text{-GH}+2H]^+$ or $[d_4\text{-GH}+2H]^+$ or $[(TAC)+p+s+H_2O+H]^+$
1085.3	2.9	1085.2	$[x_4\text{-GH}]^+$ or $[c_4\text{-GH}]^+$ or $[(TAC)+p+s+H]^+$
972.0	3.0	972.1	$[(TG)+p+s+p+s-H_2O+H]^+$
876.2	0.7	876.1	$[(TA)+s+p+s-H_2O+H]^+$
		or 876.1	$[(TA)+p+s+p+H]^+$
637.1	0.7	637.1	$[w_2+2H]^+$ or $[d_2+2H]^+$
468.1	0.7	468.1	$[c_2\text{-AH}]^+$ or $[C_{nt}+p+s+H]^+$
232.1	0.7	232.1	$[sG\text{-H}_2O+H]^+$
216.2	13.1	216.1	$[sA\text{-H}_2O+H]^+$
192.2	4.2	192.1	$[sC\text{-H}_2O+H]^+$
161.1	2.6	161.0	$[s+p\text{-H}_2O+H]^+$
152.1	2.8	152.1	$[GH+H]^+$
136.1	3.0	136.1	$[AH+H]^+$
112.1	2.9	112.1	$[CH+H]^+$
81.0	3.4	81.0	$[s\text{-H}_2O+H]^+$

*calculated monoisotopic mass
s = deoxyribose-H₂O (C₅H₆O₂), *M_r* =98.0368 Da
p = PO₃H, *M_r* = 79.9663 Da
B_{*nt*} denotes a mononucleotide which may be either psB or sBp
[(B₁...B_{*n*})] denotes a polynucleotide which may be either (psB₁...+psB_{*n*}) or (sB₁p...+sB_{*n*}p)

5'-d(CACGTG)-3' adduct shown in figure 6.50. The spectrum also shows the 1^+ ion of the depurinated oligonucleotide (**II**) at m/z 1641.9 resulting from loss of the alkylated base, and the 1^+ ion of the 3' terminal fragment (**IV**) at m/z 1254.5 arising from subsequent cleavage of the ribose C-O bond to the 3' side of the depurinated nucleotide. The relatively abundant ion at m/z 796.0 corresponds to the species, $[(TA)+p+s+H]^+$, and may be attributed to the further decomposition of the m/z 1254.5 fragment via loss of the cytosine C(6) followed by ribose C-O bond cleavage to the 3' side of this residue. The assignments of other oligonucleotide product ions generated in the spectrum are given in table 6.8.

Figure 6.55 shows the MS/MS spectrum of the $[M-2H]^{2-}$ ion of the minor adduct formed between phenC₆Br and 5'-d(CGTACG)-3' which had a retention time of 22.8 minutes in the HPLC profile of the reaction mixture (figure 6.51). In contrast to the MS/MS spectrum of the 2^- ion of the major adduct, the corresponding spectrum of the minor adduct gives rise to only a single product ion at m/z 819.4 owing to the 2^- ion of the oligonucleotide fragment resulting from the loss of alkylated guanine. The fact that no ions are observed for ribose 3' C-O bond cleavage of the depurinated residue is consistent with location of the ligand at the 3' terminus as illustrated in scheme 6.21.

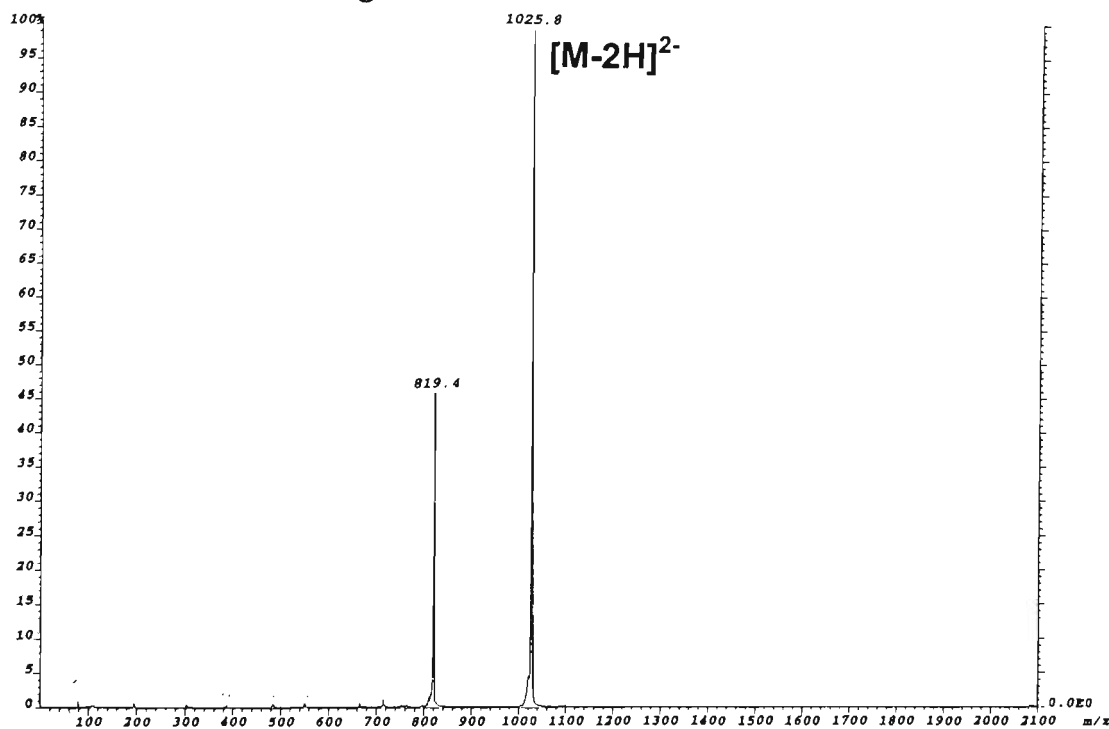
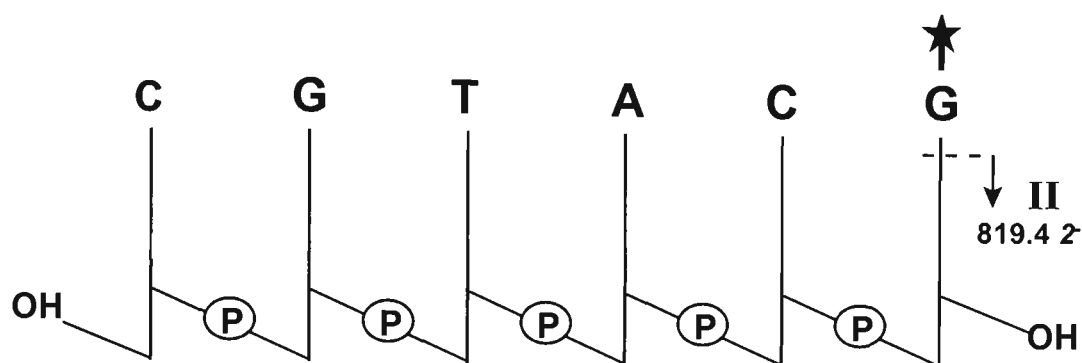


Figure 6.55: ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the minor adduct formed between phenC₆Br and 5'-d(CGTACG)-3' (m/z 1025.8).



Scheme 6.21: Fragmentation pathway observed in the ESI-MS/MS spectrum of the minor adduct formed between phenC₆Br and 5'-d(CGTACG)-3'.

6.5.3 PhenC₆ Binding to 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3': Discussion

The binding of phenC₆Br to the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' suggests that this ligand has a strong selectivity for guanine alkylation at the 5'-CGT site in each case. These observations are consistent with the sequence selectivity of this ligand determined by gel electrophoretic sequencing studies, which showed preferential alkylation guanines at 5'-GT-3' sequences. In the case of 5'-d(CACGTG)-3', reaction with phenC₆Br gives rise to only one adduct in the HPLC profile owing to binding at the 5'-CGT site. On the other hand, binding of phenC₆ to 5'-d(CGTACG)-3' produces two adducts arising from alkylation of each of the guanines. Alkylation at the 5'-CGT site is responsible for approximately 70% of total adduct formation with the remaining 30% of adduct formation from alkylation at the 5'-CG site (these calculations are based on a comparison of the peak heights of each adduct in the HPLC profile of the reaction mixture). It is interesting that these results show a greater selectivity for phenC₆ alkylation at the 5'-CGT site in 5'-d(CGTACG)-3' compared with that of hedamycin which prefers the 5'-CG site. For all three alkylating agents examined, however, the strongest selectivity for guanine alkylation at the 5'-CGT site of 5'-d(CGTACG)-3' is displayed by DC92-B.

6.6 SUMMARY OF RESULTS FOR ALKYLATING AGENTS

The three alkylating agents studied in this work, hedamycin, DC92-B and phenC₆Br, each comprise an alkylating functional group linked to a DNA-intercalating chromophore. Binding of these compounds to double stranded DNA is thought to occur initially via intercalation followed by alkylation at guanine bases. Sequencing studies of each of these ligands have shown that they bind to DNA in a highly sequence selective manner. The fragmentation observed in the negative ion ESI-MS/MS spectra of all alkylated DNA adducts results in highly simple spectra comprising product ions owing to loss of the alkylated nucleobase followed by subsequent cleavage of the ribose 3' C-O bond of the alkylated residue. The simplicity of resulting spectra are in marked contrast to the fragmentation observed in the MS/MS spectra of unmodified oligonucleotides which show product ions for an extensive range of backbone cleavages with no single pathway being predominant (as shown in chapters four and five). As a consequence of the observation of a single major fragmentation pathway in the case of alkylated adducts, a clear indication of the site of alkylation of the ligand on the oligonucleotide can be obtained from the MS/MS spectra. The results obtained for the sequence analysis of hedamycin binding to hexamers and octamers showed alkylation exclusively at guanine bases. Hedamycin reactions with all of the hexamers showed highly sequence selective binding resulting in the formation of only one adduct. ESI-MS/MS spectra of the hedamycin adducts formed with the oligonucleotides, 5'-d(CACGTG)-3' and 5'-d(GCCGGC)-3' establish alkylation of the guanine at the 5'-CGT-3' sites and 5'-CGG-3' sites respectively which is consistent with the sequence selectivity of hedamycin determined from previous studies using other techniques. In contrast, MS/MS spectra of the hedamycin adducts formed with both 5'-d(CGTACG)-3' and 5'-d(CGGCCG)-3' show alkylation to occur at the 5'-CG-3' site rather than at the guanine located at the 5'-CGT-3' binding site (for 5'-d(CGTACG)-3') and the 5'-CGG-3' binding site (for 5'-d(CGGCCG)-3'). The preferential alkylation of the terminal guanine may arise from the hydrophobic stabilisation of the terminal base-pair from the interactions between hedamycin and the terminal base-pair upon alkylation of

the terminal guanine. Support for this explanation can be derived from the results obtained for binding of hedamycin to the octanucleotide 5'-d(GCGTACGC)-3' which was designed to incorporate the 5'-d(CGTACG)-3' sequence with the addition of G-C base pairs at each terminus to impose helical constraints (G-C base pairs being chosen over A-T base pairs owing to their greater stability). ESI-MS/MS analysis of the major adduct formed upon hedamycin binding to 5'-d(GCGTACGC)-3' (constituting 70% of all adducts) show guanine alkylation at the 5'-CGT-3' site. Hence there is a change in the sequence selectivity of hedamycin binding from the 5'-CG-3' site to the 5'-CGT-3' site in the core 5'-d(CGTACG)-3' sequence. The remaining 30% of adduct formed can be attributed to guanine alkylation at the 5'-CG-3' site. Similar results were observed for hedamycin binding to 5'-d(ACGTACGT)-3'. On the other hand, the results obtained for hedamycin binding to 5'-d(TCGTACGA)-3' show a greater preference for alkylation at the 5'-CG-3' site rather than the 5'-CGT-3' site. The structural basis for this observed selectivity of hedamycin is uncertain, and may be a consequence of the presence of an additional sequence selectivity being conferred by the presence or absence of T preceding the binding site in this case.

DNA sequencing studies of the structurally-related compound, DC92-B yielded the same sequence selectivity as that observed for hedamycin i.e. preferential binding to isolated guanines at 5'-CGT-3' sites. The observed similarity in DNA binding properties of these compounds is intriguing considering the structural differences of their alkylating bis-epoxide sidechains in both their substitution and stereochemistry. The binding of DC92-B to the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' in each case is consistent with these results showing preferential alkylation of the guanine at the 5'-CGT-3' site. In contrast to hedamycin, DC92-B binding to these oligonucleotides was found to give rise to two adducts. The ESI-MS/MS of each of the adducts formed with the oligonucleotide 5'-d(CACGTG)-3' yields very similar spectra in both positive and negative ion mode, indicative of guanine alkylation at the 5'-CGT-3' site. The results obtained from MS/MS analyses of the product ions corresponding to the alkylated base generated by in-source collisional activation of the

adducts, however, suggests the distinguishing feature between the two adducts may be in the location of guanine alkylation on the bis-epoxide sidechain. The difference in selectivity of DC92-B binding to 5'-d(CGTACG)-3' (5'-CGT-3' site) compared with that of hedamycin (5'-CG-3'), is the only example indicative of the structural differences between the two ligands affecting their binding properties. In light of the results obtained for DC92-B binding to 5'-d(CACGTG)-3', a possible explanation could be that the differences in the substitution and stereochemistry of their bis-epoxide sidechains gives rise to differences in the location of guanine alkylation on the bis-epoxide sidechain for each ligand. Guanine alkylation is thought to occur from electrophilic attack of the terminal carbon of the bis-epoxide in the binding of hedamycin. The steric hindrance presented by the dimethylamino group attached at the terminal carbon on the bis-epoxide sidechain of DC92-B may result in preferable guanine alkylation by a different carbon on the sidechain. Such differences may account for the observed differences in the sequence selectivity of binding of both compounds.

The synthetic compound, phenC₆Br was included in the present study since, although it does not display significant antitumour activity, it is similar to hedamycin and DC92-B, in that it both intercalates and alkylates DNA with preferential alkylation of guanines occurring in 5'-GT-3' sequences as well as in runs of guanines [(Gp)_n]. However, whereas hedamycin and DC92-B both comprise of intercalating anthraquinone chromophores attached to an alkylating bis-epoxide sidechain, phenC₆Br contains an alkylating moiety linked by a polymethylene chain to a quaternized phenanthridine ring. ESI-MS/MS results of adducts formed between phenC₆Br and the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' confirmed a preference for adduct formation from alkylation of the guanine at the 5'-CGT-3' site as predicted from the results of previous DNA sequencing studies. In the case of 5'-d(CACGTG)-3', the binding of phenC₆Br was found to give rise exclusively to the formation of the adduct for 5'-CGT-3' alkylation. In contrast, phenC₆Br binding to 5'-d(CGTACG)-3' also results in the formation of a small amount of the adduct arising from alkylation of the guanine at the 5'-CG-3' site. The similarity of the sequence selectivity of phenC₆Br

binding compared with that of the naturally-based products, hedamycin and DC92-B is a significant observation considering the marked differences in structure between these two classes of compounds.

Overall, the results obtained here show that ESI-MS/MS has the potential to be used as an important tool when trying to understand the structural basis of sequence selectivity of alkylating ligands when they bind to DNA.

REFERENCES

- (1) Wickham, G.; Prakash, A. S.; Wakelin, L. P. G.; McFadyen, W. D. *Biochim. Biophys. Acta* **1991**, *1073*, 528-537.
- (2) Prakash, A. S.; Moore, A. G.; Murray, V.; Matias, C.; McFadyen, W. D.; Wickham, G. *Chemico-Biological Interactions* **1995**, *95*, 17-28.
- (3) Murray, V.; Moore, A. G.; Matias, C.; Wickham, G. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1995**, *1261*, 195-200.
- (4) Murray, V.; Matias, C.; McFadyen, W. D.; Wickham, G. *Biochimica et Biophysica Acta-Gene Structure and Expression* **1996**, *1305*, 79-86.
- (5) Iannitti, P.; Sheil, M. M.; Wickham, G. *J. Am. Chem. Soc.* **1997**, *119*, 1490-1491.
- (6) Pavlopoulos, S.; Bicknell, W.; Craik, D. J.; Wickham, G. *Biochemistry* **1996**, *35*, 9314-9324.

Chapter 7

CONCLUSIONS

The aim of the work described in this thesis was to analyse the fragmentation observed in electrospray tandem mass spectra (ESI-MS/MS) of oligonucleotides on a hybrid-sector-time-of-flight instrument, and ultimately apply this information to the identification and characterisation of the ligand-DNA binding sites of a range of novel alkylating agents.

In the first stage of this study (described in chapter three), preliminary experiments were undertaken to optimise the intensity of precursor ions of oligonucleotides with different numbers of charges in negative ion ESI-MS spectra for subsequent MS/MS analyses. This involved a determination of optimum analyte concentration, and an examination of the effects of solvent composition (such as percentage of organic solvent, pH and the use of chemical additives such as organic bases). In addition, the effects of varying experimental parameters in the electrospray interface (such as the potential applied to the sampling cone and skimmer lens) were examined and it was found that the charge state distribution of hexanucleotide ions could be readily altered by adjusting the sampling cone and skimmer lens voltages.

The second stage of this work (described in chapter four) involved the detailed characterisation of the product ions observed in the MS/MS of oligonucleotides in the negative and positive ion modes. The effect a number factors on the MS/MS fragmentation of oligonucleotide anions were then explored. In particular, the focus of this work was to assess and compare the degree to which oligonucleotide charge state, sequence, base identity and base location influence observed fragmentation patterns. In the light of the knowledge gained from these studies, the binding properties and sequence selectivity of three novel intercalating, alkylating agents hedamycin, DC92-B,

and *n*-bromohexylphenanthridinium bromide to constructed sequences of DNA were then examined. Prior to this study there had been no reports of the use of mass spectrometry for the sequence analysis of ligand binding to DNA, and so this work aimed to explore the extent to which ESI-MS/MS could be usefully applied to the determination of the sequence selectivity of DNA-alkylating agents and ultimately, to determine the potential of ESI-MS/MS to assist in the design of DNA-interactive antitumour agents.

The types of product ions observed in the negative and positive ion MS/MS spectra of singly charged unmodified oligonucleotides on the hybrid-sector-TOF instrument were characterised in detail by MS/MS, and, in some cases, high resolution MS, of product ions generated by in-source collisional activation. The negative ion MS/MS spectra of unmodified oligonucleotides showed extensive backbone fragmentation, with sequence ions arising from *w* and (*a*-B) cleavages being predominant. The low *m/z* region of these spectra were found to comprise monomeric oligonucleotide constituents (backbone phosphate and sugar fragments, all constituent bases (except in some cases thymine)) and a characteristic series of mononucleotide ions for each nucleobase present in the oligonucleotide sequence. From high resolution MS, it was found that two different structures were generally present for the backbone and mononucleotide product ion types comprising the addition of 80 Da and 98 Da on a simple nucleotide unit, B_{nt} (or [s+p]). The region above *m/z* 500 was found to comprise numerous polynucleotide product ions including both terminal ions (sequence ions) and internal ions. As observed with mononucleotide ions, polynucleotide ions with additions of 80 Da, 98 Da, 258 Da and 276 Da could arise from two possible structures which differ in the addition of a phosphate or a sugar group provided that both such structures are possible from the sequence of the oligonucleotide. In cases where product ions could be assigned as more than one structure of different base composition, the identity of the ion could be discerned from MS/MS spectra of the source-generated product ion. Although in a majority of cases, it was not possible to distinguish terminal sequence

ions from internal ions, the assignment of *w* and (a-B) sequence ions was made on the basis of the generally greater relative intensity of these ions.

Analysis of fragmentation pathways observed upon MS/MS of the major product ions observed in the MS/MS spectra of both 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' showed that the majority of internal ions present in MS/MS spectra originate from the further decomposition of the major sequence ions which would be thought to arise from multiple collisions. MS/MS of source-generated *w* ions showed these ions to give rise to smaller *w* ions and internal ions of the type $[(B..B_n)+p+s-H]^-$ which correspond to the complementary 3' and 5' ions arising from fragmentation of the precursor *w* ion via nucleobase loss and cleavage of the ribose 3' C-O bond of the residue from which the base is lost. The (a-B) ions were also found to fragment via the same fragmentation pathway, although the production of smaller (a-B) ions was observed to be significantly less favoured than production of ions for the 3' fragments arising from this pathway *i.e.* $[(B..B_n)+p+s-H]^-$. This would seem to suggest a preference for location of the negative charge on the phosphate group to the 3' side of the cleavage site. In all cases a lack of preference was observed for the production of ions from loss of neutral thymine, and cleavage of the ribose 3' C-O bond to the 3' side of thymine-containing residues.

The positive ion MS/MS spectra of singly charged precursors showed some significant differences in fragmentation patterns compared with those observed in the negative ion mode, although intense *w* and (a-B) sequence ions were also generated in some cases. Characterisation of product ions in the low *m/z* region of positive ion MS/MS spectra showed the formation of ions comprising sugar residues, *i.e.* ions for protonated sugar $[C_5H_4O+H]^+$, sugar+phosphate $[C_5H_4O+PO_3H+H]^+$, protonated bases (apart from thymine), and sugar+nucleobase $[C_5H_4O+BH+H]^+$ residues were particularly abundant. One key finding was that the MS/MS spectrum of the ion at *m/z* 81.0 observed in positive ion MS/MS spectra was consistent with this ion arising from the sugar group, $[s-H_2O+H]^+$, and not a protonated phosphate group as has been reported in a previous study¹. The absence of ions for protonated thymine, and for loss of neutral

thymine from the precursor ion, was consistent with the low proton affinity of thymine in comparison to the other nucleobases and with the results of earlier studies of positively charged oligonucleotides². A characteristic feature in the m/z region above 500 of positive ion MS/MS spectra of singly charged oligonucleotides was the presence of series of polynucleotide ions arising from multiple losses of nucleobases (predominantly cytosine and guanine). MS/MS analysis of the major product ion types observed in the positive MS/MS spectra of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' showed dissociation to occur predominantly via nucleobase loss followed by cleavage of the adjacent ribose 3' C-O bond from which the base is lost, analogous to the major fragmentation pathways observed in the negative ion mode. However, the observation of w and (a-B) sequence ions as intact species was not found to be as prevalent owing to the facile loss of neutral cytosine and guanine from these fragments. The facile loss of cytosine and guanine as neutral species in the positive ion mode can be rationalised by their greater relative stabilities as leaving groups compared with adenine and thymine following protonation and glycosidic bond cleavage².

The differences in fragmentation pathways observed for a range of singly charged oligonucleotides in negative versus positive ion MS/MS spectra were consistent with the location of the charge directing fragmentation pathways. In the case of singly charged oligonucleotide anions, the dominant observation of w and (a-B) sequence ions is consistent with the charge being located on the phosphate groups catalysing nucleobase loss. Subsequent cleavage of the ribose 3' C-O bond of the residue from which the base is lost is thought to be driven by the formation of a stable furan product³. The MS/MS spectra of singly charged protonated oligonucleotides, on the other hand, show nucleobase loss from the precursor ion, and from fragments formed from backbone cleavage, to dominate the types of ions observed. The characteristic series of product ion types generated were found to be consistent with the relative proton affinities and leaving group stabilities of the nucleobases as determined by Rodgers *et*

*al.*². This, in turn, is consistent with the positive charge being located on the nucleobase and facilitating N-glycosidic bond cleavage.

In chapter five, an analysis of the major ion types generated from the MS/MS of the doubly charged anions of a set of hexamers with both unibase and heterogenous base composition was carried out. It was found that base location and the site(s) of deprotonation of the precursors were important factors determining the propensity for formation of different sequence ions. This was a key observation in the case of *w* ion formation, with a preference exhibited for formation of w_1^- ions whereas the w_3^- ions were generally low in abundance. These results could be explained by a preference for the distribution of the two charges on the terminal phosphate groups such that minimum Coulombic repulsion is experienced. The MS/MS spectra of heterogeneous oligonucleotides, showed base identity to play an important role in influencing the formation of *w* and (*a*-B) ions. A distinct lack of preference was observed for production of *w* and (*a*-B) sequence ions where thymine was located at the cleavage site. The effects of base identity, however, were not found to overcome the trends arising from the effects of charge location in directing formation of the major ions observed.

Chapter five also describes the comparison of the MS/MS spectra of all charge states (1^- to 5^-) of a series of hexamers. These data showed that precursor ion charge state strongly dictated the propensities for formation of the major product ion types arising from nucleobase loss, and cleavage of the ribose 3' C-O bond of the residue from which the base is lost to give *w* and (*a*-B) sequence ions. In the case of nucleobase loss, it was found that increasing precursor ion charge state not only resulted in an increased preference for bases to be lost as charged rather than neutral species, but also influenced the relative propensities for both charged and neutral base loss of the four bases. With respect to neutral base loss, the fragmentation of singly and doubly charged oligonucleotide anions showed similar trends with favourable loss of the 3' terminal base as a neutral followed by that of the 5' terminal base. In each case, these trends were overcome by the strong lack of preference for the loss of neutral thymine. MS/MS

spectra of 3⁻ anions showed predominant ions for neutral base loss of the 5' terminal base except when thymine was located at the 5' terminus. Fragmentation of 4⁻ precursor ions showed neutral base loss of the 5' terminal base to be highly favoured with a strong lack of preference for neutral base loss from the 3'-terminus. In addition to these effects of base location, less favourable production of ions formed via the loss of GH was generally observed in the MS/MS spectra of higher charge state precursor ions, with the reverse effect found for AH loss which appeared to become more favourable. The results obtained for base anion loss as a function of oligonucleotide charge state show base loss to be directed by both charge location and base identity, although the extent of the contribution of each of these factors was also found to be a function of precursor ion charge state. The MS/MS spectra of multiply charged oligonucleotide anions in almost all cases yielded markedly different preferences for base anion intensities from that observed for the corresponding singly charged precursor ions. In the MS/MS spectra of singly charged deprotonated oligonucleotides, the relative propensities for base anion formation appear to be influenced by the relative acidities of the bases with the loss of charged thymine observed to be highly favoured and a lack of preference for loss of charged cytosine. These results are consistent with calculated values of the relative acidities of the nucleobases determined by Rodgers *et al.*⁴ of T > G > A > C. Base location is also observed to have an effect on the propensities for base anion formation in the MS/MS spectra of singly charged precursor ions, which is particularly evident from the increased preference for formation of anions of cytosine and adenine when these bases are located at the 5'-terminal position. Base location was also seen to have an effect in the case of charged adenine loss, which is observed to be highly favourable when it is located at the 5'-terminal position. In the case of doubly charged oligonucleotide anions, all MS/MS spectra show the 5' terminal base anion to yield the most intense signal followed by the 3' terminal base. Upon increasing the number of charges on the oligonucleotide backbone, the influence of charge location on the propensity for base anion formation is seen to diminish, and the propensity for base loss

was found to become more strongly dictated by base identity. Although the MS/MS spectra of precursors from 2⁻ to 5⁻ of oligonucleotides showed an overall preference for the loss of the 5'-terminal base anion (with only a few exceptions), the subsequent order of base anion intensities in MS/MS spectra was observed to vary significantly upon shifting to higher charge state precursor anions. The most striking observation was the substantial increase in propensity for loss of A⁻ in the MS/MS spectra of highly charged anions of the oligonucleotides 5'-d(CACGTG)-3', 5'-d(CGTACG)-3' and 5'-d(TCACGA)-3'. These findings would appear to be consistent with those obtained by McLuckey *et al.*³ on ion trap instrumentation which show the preferential loss of A⁻ in the decomposition of highly charged species. The relative propensities for formation of ions owing to loss of a neutral nucleobase (with the oligonucleotide remaining intact), appeared to be strongly influenced by base location, with base identity also having a significant effect in the case of neutral thymine loss. In all MS/MS spectra, except where thymine was located at the 3' terminal position, the most intense ion arising from neutral base loss corresponded to loss of the 3' terminal base. This preference may arise from the stability of this ion to further decomposition via ribose 3' C-O bond cleavage. The lack of preference for loss of neutral thymine was also a strong trend observed in the MS/MS spectra of all oligonucleotides in which thymine was not a terminal base. In these cases, no ions for loss of neutral thymine were observed, which appears to suggest that the location of thymine at the 5' or 3' terminus is integral to its ability to be lost as a neutral species. These data are consistent with results obtained from other studies which showed that neutral thymine is a significantly less stable leaving group in comparison to the other nucleobases⁴. The preference for neutral base loss of the 3' terminal base appears to be followed in most cases by loss of the 5' terminal base, although this was observed to be dependent on oligonucleotide sequence composition and the identity of the 5' terminal base. The effect of increasing charge state on the relative intensities of ions owing to charged nucleobase loss, *i.e.* [M-B⁻-nH]⁽ⁿ⁺¹⁾⁻ ions, in the MS/MS spectra of the 3⁻ and 4⁻ precursor ions corresponded reasonably well with the effects noted for the

relative intensities of the base anions. The propensities for charged base loss observed in the MS/MS spectra of each of the 3^- precursors showed production of ions for the charged loss of the 5'-terminal base to be favoured. The order for charged base loss observed in the MS/MS spectra of the 3^- ions of each oligonucleotide was in most cases consistent with the relative intensities generated from the formation of base anions in each respective spectrum, and furthermore, are markedly similar to the propensities for ion formation from neutral base loss. The major difference in the case of neutral base loss compared with charged base loss for 3^- precursor ions arose from the loss of neutral thymine being unfavourable. Furthermore, the trends for charged base loss in the 3^- charge state showed the loss of charged cytosine to be unfavourable in cases where cytosine was not the 5' base. Upon considering precursor ions in the 4^- charge state, the propensities for charged base loss were found to be strongly influenced by base identity and sequence. In particular, enhanced loss of adenine was observed for pathways involving charged and neutral base loss for the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'. This trend was consistent with the observed relative abundances of base anions, and is likewise reflected in the ion abundances from fragmentation owing to neutral base loss. The propensities for ions formed from charged base loss in the MS/MS spectra of 5^- precursor ions did not correspond to the base anion intensities which may be a result of the effects of multiple collisions of product ions for charged base loss and the further decomposition of product ions contributing to relative intensities of base anions, however, it is not clear why this would particularly be the case with the 5^- precursor ions.

The trends observed for the formation of w and (a-B) ions in the MS/MS spectra of singly charged oligonucleotide precursor anions showed relatively intense w_3^- ions and $[a_4-B_4H-2H]^-$ ions, except when thymine is located to the 5' side of the cleavage site. These observations may arise from a preference for location of the negative charge in the case of singly charged hexamers on the phosphate group of the third residue which could catalyse loss of the adjacent bases to the 5' and 3' side. These results are

interesting because it is generally assumed that *w* and *a-B* ions are formed via a common pathway (with the underlying assumption that there is then a statistical distribution of charge between the two possible fragments). The results obtained here, however, suggest these ions are likely to be formed via two different mechanisms. In the case of the *a-B* ions, this may involve cleavage remote to the site of the charge. The MS/MS spectra of doubly charged oligonucleotides anions, on the other hand, showed formation of w_1^- ions to be particularly favourable, with a distinct lack of preference for formation of *w* ions from ribose 3' C-O bond cleavage at the third residue. As discussed previously, the favourable formation of w_1^- ions may arise from a combination of two effects. First, multiple collisions may yield the smallest possible *w* fragments and, secondly a preference for charge location in the case of doubly charged anions on the first and fifth phosphate groups would be consistent with the least charge repulsion, which catalyses loss of the base on the adjacent 5' residue and subsequent cleavage of the ribose 3' C-O bond. The lack of propensity for formation of w_3^- ions is also consistent with the preference for the charge being located at the terminal phosphate residues. The MS/MS spectra of triply charged precursor ions showed the production of the singly charged ions w_1^- and w_2^- to be prevalent, with a strong lack of preference for formation of singly charged *w* ions larger than w_3^- , consistent with the 3 negative charges being preferentially located on the phosphate backbone so as to minimise Coulombic repulsions. Doubly charged $[w_5-H]^{2-}$ ions were generally found to be the most abundant of the *w* ions, consistent with the trends for charged base loss observed in the MS/MS spectra of triply charged precursor ions whereby loss of the 5' terminal base as an anion was found to be particularly favourable. Considering the precursors with higher charges (4⁻ and 5⁻), the propensities for formation of the various *w* ions was found to be highly dependent on oligonucleotide sequence.

In the case of (*a-B*) ions, the MS/MS spectra of singly and doubly charged oligonucleotide anions showed a preference for specific cleavage types, with the formation of $[a_4-B_4H-2H]^-$ ions being favourable in the MS/MS spectra of singly

charged anions and the formation of $[a_3-B_3H-2H]^-$ ions being generally favoured for doubly charged anions. In contrast, upon considering higher charge states, the formation of (a-B) ions is observed to be highly influenced by factors such as base identity and propensity for base loss. The observed variation in the propensities for certain fragmentation pathways with precursor ion charge state appears to reflect the extent to which there exists a preference for charge location on the backbone which can direct fragmentation behaviour. This was particularly evident when comparing the trends observed in the MS/MS spectra of doubly and triply charged oligonucleotide precursor ions with those of more highly charged species. The results obtained for the MS/MS spectra of oligonucleotide ions of lower charge states show marked similarities in the propensities for w and (a-B) formation regardless of oligonucleotide sequence. Furthermore, base location was found to be an important factor influencing the trends for both charged and neutral nucleobase loss. The results for the fragmentation of doubly and triply charged precursor ions are, therefore, consistent with preferential charge location playing an important role in directing the major product ions types observed, since for hexamers with two and three charges, there would be expected to be a preference for charge location on the phosphodiester backbone which would yield minimum Coloumbic repulsion between charges. In the case of more highly charged species, however, there would be no overall preference for charge location in terms of minimising charge repulsion. This is reflected in the MS/MS results obtained upon considering higher charge states, which showed the production of the various w and (a-B) ions to be highly sequence-dependent and strongly influenced by the relative propensities for base loss at the cleavage site.

The final part of the work (described in chapter six) conducted for this thesis involved the application of ESI-MS/MS to characterising the DNA-binding properties of three novel alkylating agents. The three alkylating agents studied in this work, hedamycin, DC92-B and phenC₆Br, each comprise an alkylating functional group linked to a DNA-intercalating chromophore. DNA sequencing studies of each of these ligands

by other techniques had shown that they bind to double-stranded DNA in a highly sequence selective manner which is thought to involve initial rapid intercalation followed by alkylation⁵. In each case, negative ion ESI-MS/MS fragmentation of alkylated-DNA adducts showed highly specific fragmentation pathways indicative of the site of alkylation. The MS/MS spectra were found to be relatively simple, comprising product ions for loss of the alkylated nucleobase, the oligonucleotide minus the alkylated nucleobase, and w and (a-B) ions owing to subsequent cleavage of the ribose 3' C-O bond of the alkylated residue from which the base had been lost. This was in marked contrast to the fragmentation observed in the MS/MS of unmodified oligonucleotides which showed product ions arising from an extensive range of backbone cleavages, with no one single pathway being predominant.

The results obtained for the sequence analysis of all three alkylating agents examined in this work showed alkylation occurred exclusively at guanine bases, consistent with earlier sequencing studies of these ligands⁵. Hedamycin reactions with all of the hexamers showed highly sequence selective binding resulting in the formation of only one adduct. ESI-MS/MS spectra of the hedamycin adducts formed with the oligonucleotides, 5'-d(CACGTG)-3' and 5'-d(GCCGGC)-3' established alkylation of the guanine at the 5'-CGT-3' sites and 5'-CGG-3' sites. In contrast, the MS/MS spectra of the hedamycin adducts formed with both 5'-d(CGTACG)-3' and 5'-d(CGGCCG)-3' showed alkylation occurred at the 5'-CG-3' site (rather than at the guanine located at the 5'-CGT-3' binding site for 5'-d(CGTACG)-3' and the 5'-CGG-3' binding site for 5'-d(CGGCCG)-3'). The observation that alkylation of the terminal guanine was preferred may arise from the hydrophobic stabilisation of the terminal base-pair from the interactions between hedamycin and the terminal base-pair upon alkylation of the terminal guanine. Support for this explanation can be derived from the results obtained for binding of hedamycin to the octanucleotide 5'-d(GCGTACGC)-3', which was designed to incorporate the 5'-d(CGTACG)-3' sequence with the addition of G-C base pairs at each terminus to impose helical constraints (with G-C base pairs being chosen

over A-T base pairs owing to their greater stability). ESI-MS/MS analysis of the major adduct formed upon hedamycin binding to 5'-d(GCGTACGC)-3' (which constituted 70% of all adducts) showed guanine alkylation at the 5'-CGT-3' site, which was a switch in sequence selectivity of hedamycin binding from the 5'-CG-3' site to the 5'-CGT-3' site in the core 5'-d(CGTACG)-3' sequence. The remaining 30% of adduct formation was accounted for by alkylation at the 5'-CG-3' site. Similar results were observed for hedamycin binding to 5'-d(ACGTACGT)-3'. On the other hand, the results obtained for hedamycin binding to 5'-d(TCGTACGA)-3' showed a greater preference for alkylation at the 5'-CG-3' site rather than the 5'-CGT-3' site. The structural basis for this observed selectivity of hedamycin is uncertain, however, it may be a consequence of the base preceding the preferred site (in this case T) also affecting the sequence selectivity of hedamycin.

DNA sequencing studies of the structurally related ligand, DC92-B yielded the same sequence selectivity as that observed for hedamycin *i.e.* preferential binding to isolated guanines at 5'-CGT-3' sites. The observed similarity in DNA binding properties of these ligands is somewhat surprising given the structural differences between their alkylating bis-epoxide sidechains in terms of both substitution and stereochemistry. The binding of DC92-B to the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' in each case was consistent with these results, showing preferential alkylation of the guanine at the 5'-CGT-3' site. In contrast to hedamycin, DC92-B binding to these oligonucleotides was found to give rise to two adducts. ESI-MS/MS analysis of each of the adducts formed with the oligonucleotide 5'-d(CACGTG)-3' yielded very similar spectra in both positive and negative ion mode, which were indicative of guanine alkylation at the 5'-CGT-3' site. The results obtained from MS/MS analysis of the product ions from the alkylated base, generated by in-source collisional activation of these adducts, however, suggested the distinguishing feature between the two adducts may be in the location of guanine alkylation on the bis-epoxide sidechain. The different selectivity of DC92-B binding to 5'-d(CGTACG)-3' (5'-CGT-3' site) compared with

that of hedamycin (5'-CG-3' site), is a notable result since it is an example of the effects of the structural differences between the two compounds on their binding properties. In light of the results obtained for DC92-B binding to 5'-d(CACGTG)-3', a possible explanation could be that the differences in the substitution and stereochemistry of their bis-epoxide sidechains gives rise to differences in the location of guanine alkylation on the bis-epoxide sidechain of each ligand. Guanine alkylation is thought to occur via electrophilic attack of the terminal carbon of the bis-epoxide in the binding of hedamycin. The steric hindrance presented by the dimethylamino group attached at the terminal carbon on the bis-epoxide sidechain of DC92-B may result in preferable guanine alkylation via a different carbon on the sidechain. These structural differences may account for the observed differences in the sequence selectivity of binding of both compounds.

The synthetic compound, phenC₆Br, was included in the present study since, although it does not display significant antitumour activity, it has a similar mode of action to hedamycin and DC92-B, in that it both intercalates and alkylates DNA with preferential alkylation of guanines occurring in 5'-GT-3' sequences as well as in runs of guanines [(Gp)_n]. However, whereas hedamycin and DC92-B both comprise of intercalating, anthraquinone chromophores attached to an alkylating bis-epoxide sidechain, phenC₆Br contains an alkylating moiety linked by a polymethylene chain to a quaternized phenanthridine ring. ESI-MS/MS of the adducts formed between phenC₆Br and the oligonucleotides 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3' confirmed a preference for adduct formation from alkylation of the guanine at the 5'-CGT-3' site as predicted by the results of previous DNA sequencing studies. In the case of 5'-d(CACGTG)-3', the binding of phenC₆Br was found to give rise exclusively to the adduct for alkylation at 5'-CGT-3'. In contrast, phenC₆Br binding to 5'-d(CGTACG)-3' also resulted in the formation of a small amount of the adduct arising from alkylation of the guanine at the 5'-CG-3' site. The similarity of the sequence selectivity of phenC₆Br binding compared with that of the natural products, hedamycin and DC92-B is a

significant observation considering the marked differences in structure between these two classes of compounds.

In conclusion, therefore, the results obtained in this thesis indicate the enormous potential of ESI-MS/MS for use as a highly sensitive and specific technique for the characterisation of ligand-DNA adducts.

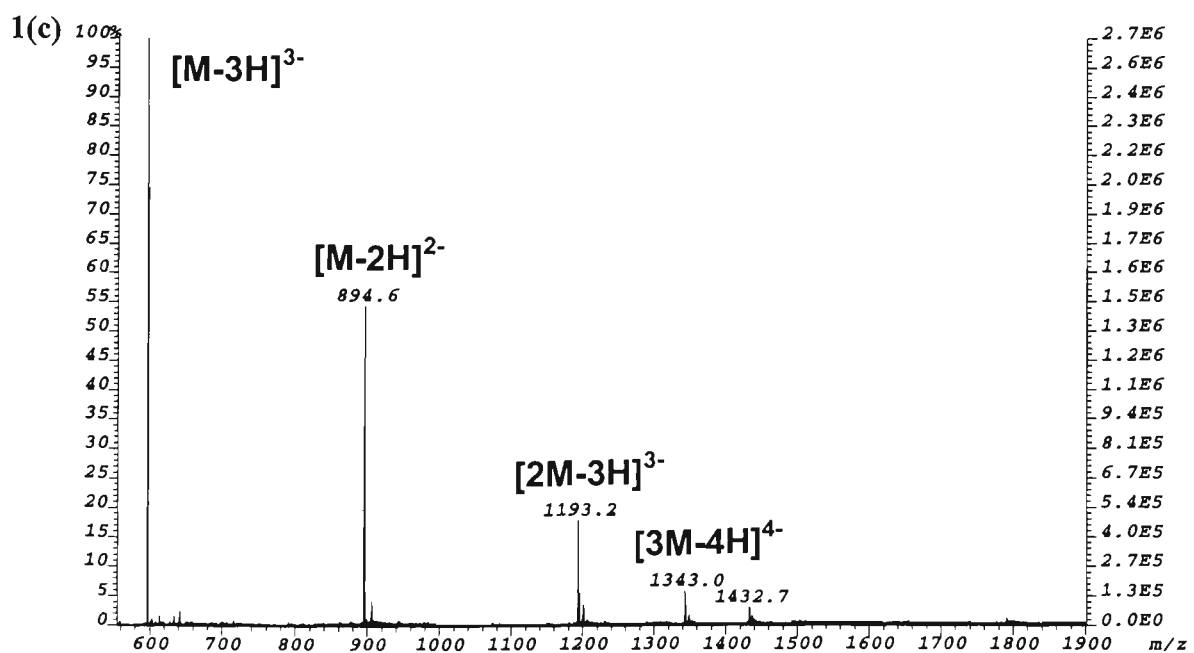
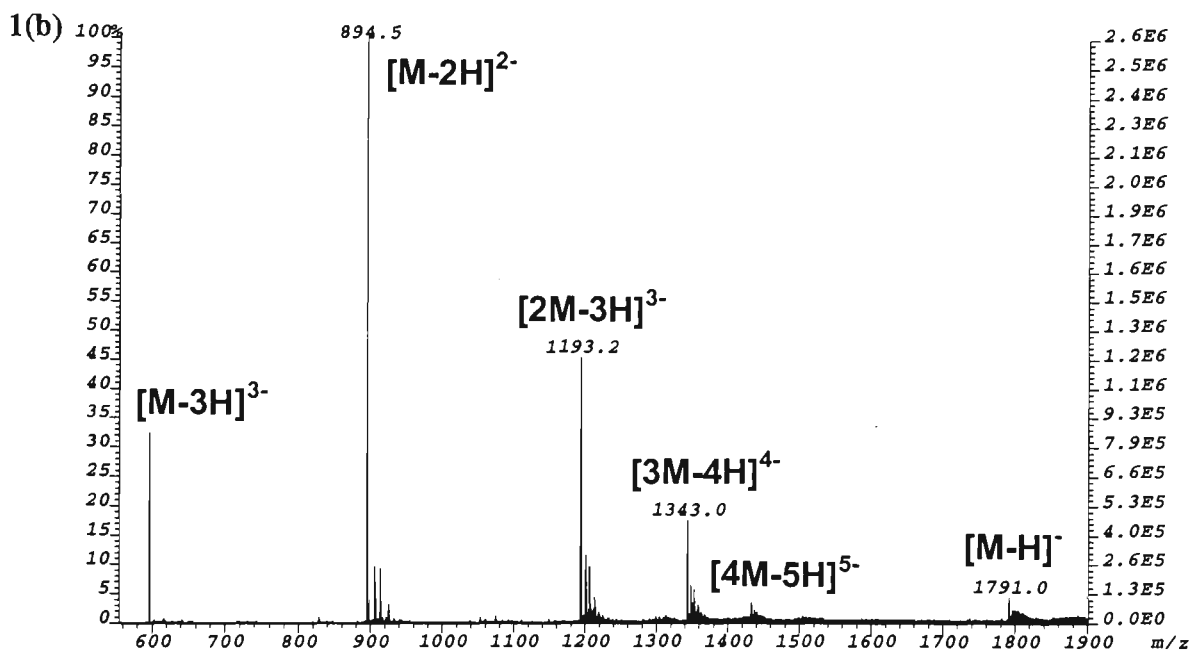
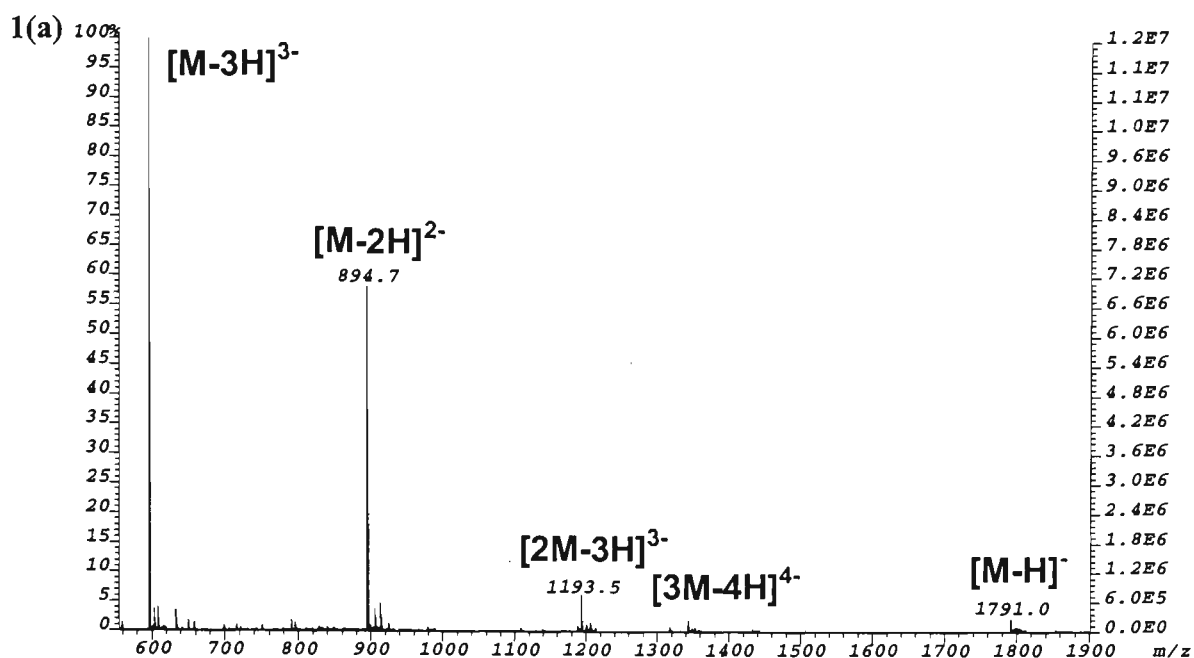
REFERENCES

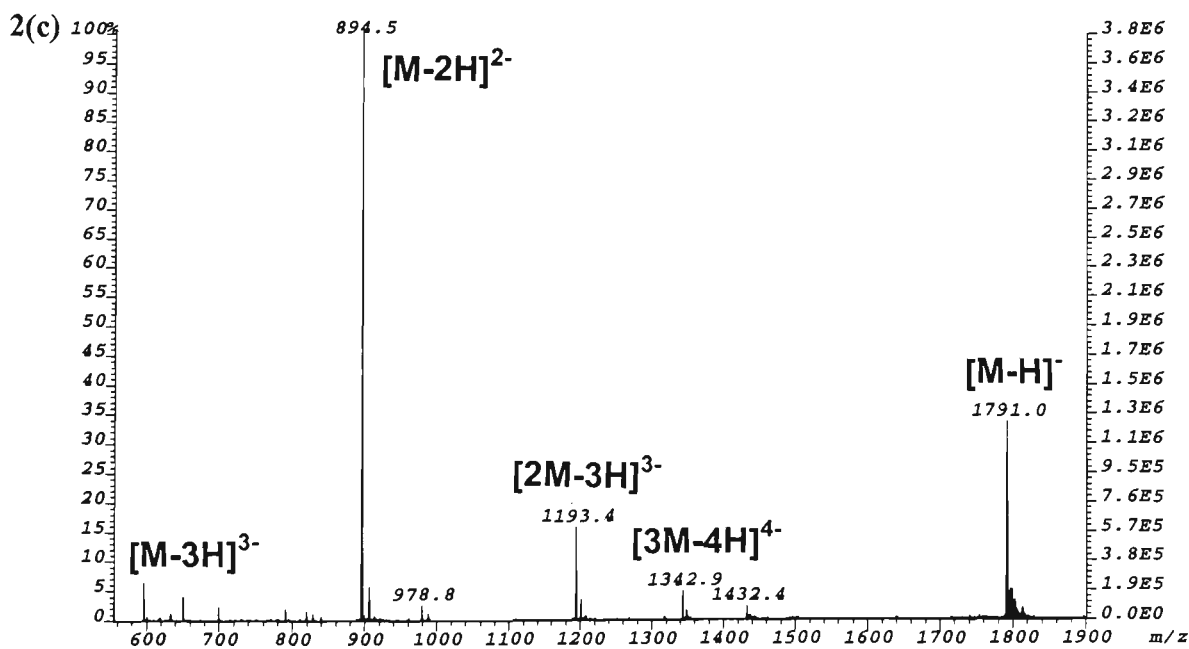
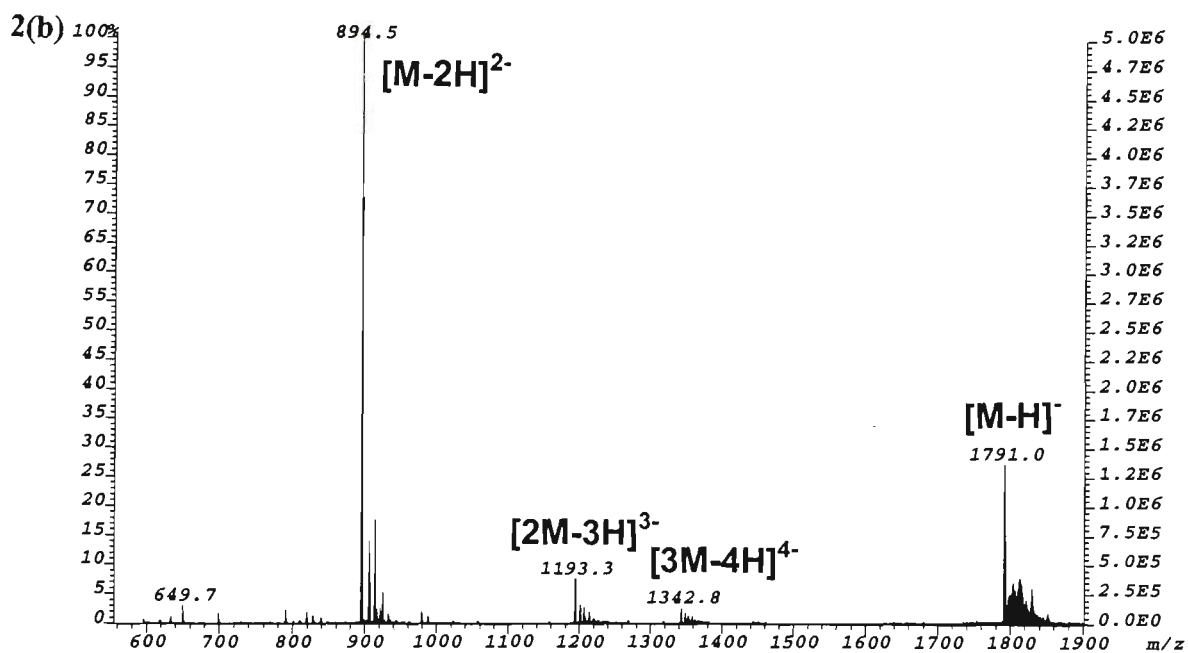
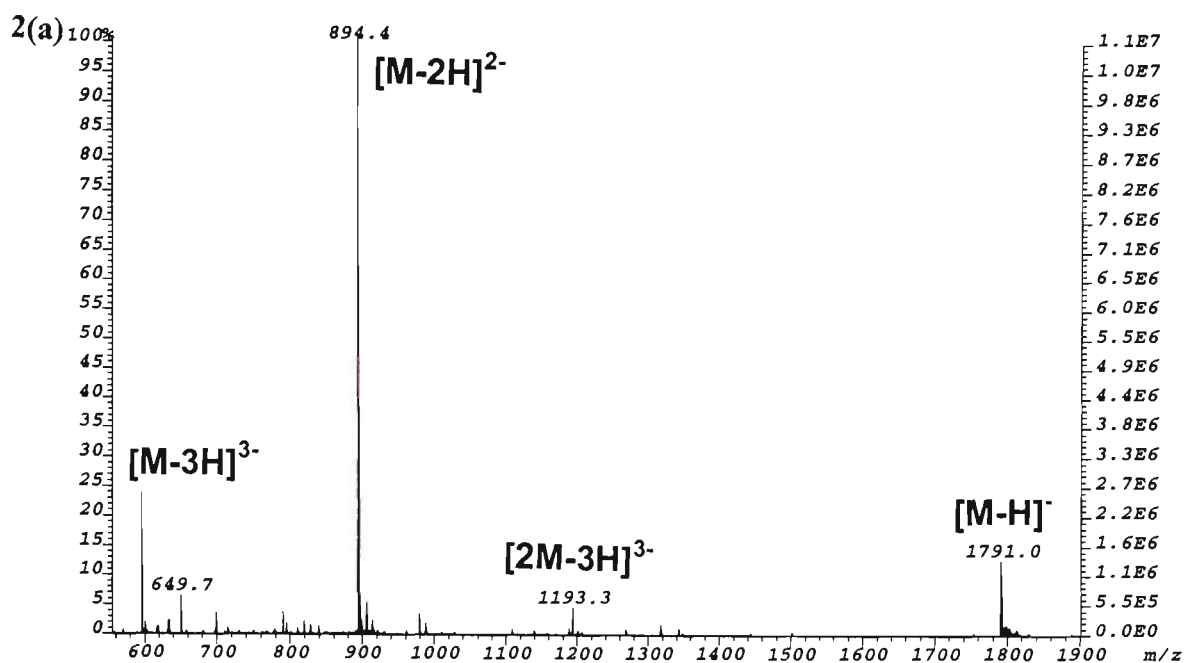
- (1) Wang, P.; Bartlett, M. G.; Martin, L. B. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 846-856.
- (2) Rodgers, M. T.; Campbell, S.; Marzluff, E. M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1995**, *148*, 1-23.
- (3) McLuckey, S. A.; Habibi-Gourdarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085-12095.
- (4) Rodgers, M. T.; Campbell, S.; Marzluff, M.; Beauchamp, J. L. *Int. J. Mass Spectrom. Ion Proc.* **1994**, *137*, 121-149.
- (5) Prakash, A. S.; Moore, A. G.; Murray, V.; Matias, C.; McFadyen, W. D.; Wickham, G. *Chemico-Biological Interactions* **1995**, *95*, 17-28.

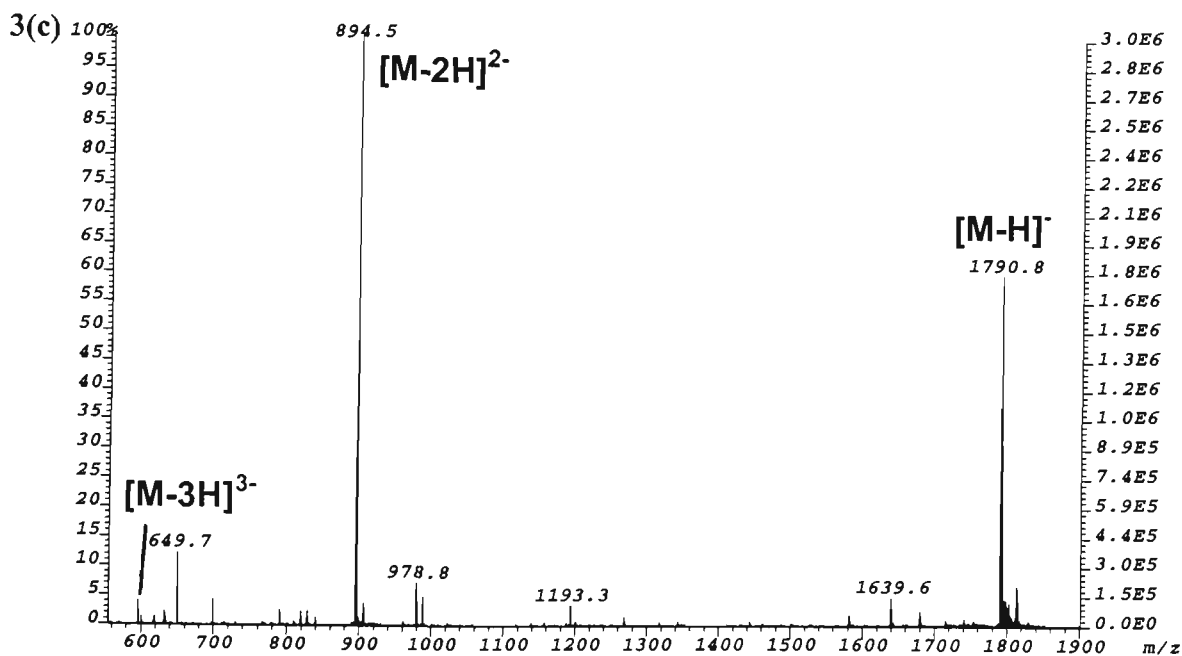
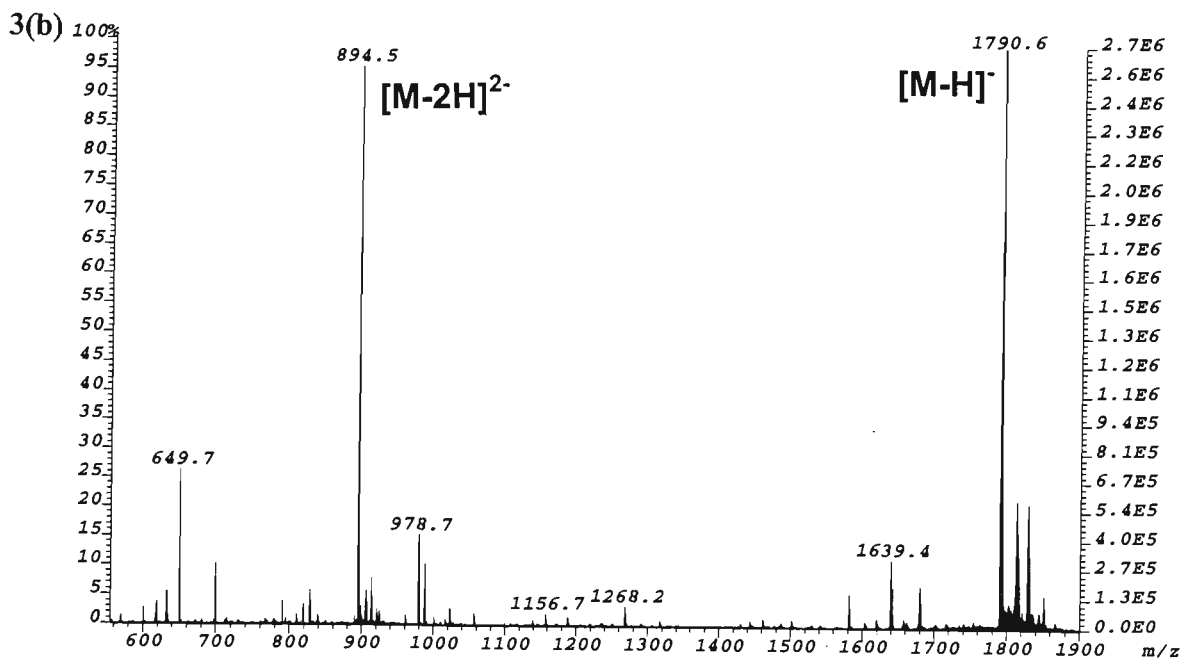
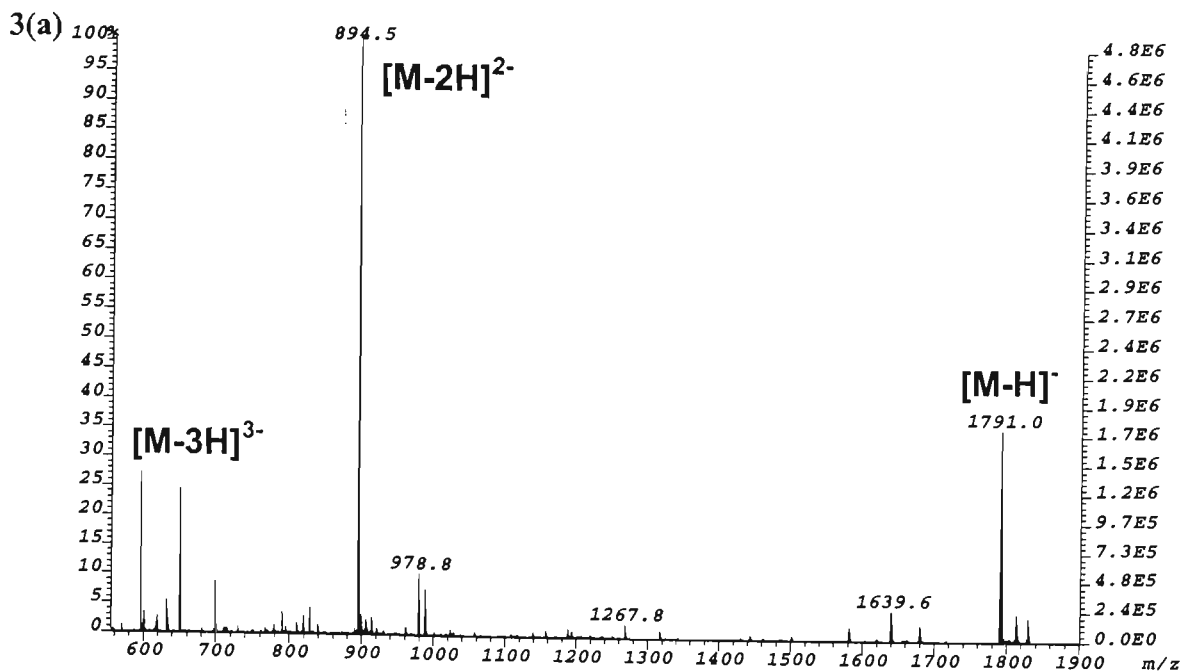
Appendix 3.1

The effect of acid and base addition on the negative ion ESI-MS/MS spectra of 5'-d(CACGTG)-3'.

Figure No.	Cone Voltage (V)	Solvent Composition
1(a)	20.8	50 % Aqueous Isopropanol
1(b)	20.8	50 % Aqueous Isopropanol + 1% Acetic Acid
1(c)	20.8	50 % Aqueous Isopropanol + 1% Ammonium Hydroxide
2(a)	44.0	50 % Aqueous Isopropanol
2(b)	44.0	50 % Aqueous Isopropanol + 1% Acetic Acid
2(c)	44.0	50 % Aqueous Isopropanol + 1% Ammonium Hydroxide
3(a)	64.0	50 % Aqueous Isopropanol
3(b)	64.0	50 % Aqueous Isopropanol + 1% Acetic Acid
3(c)	64.0	50 % Aqueous Isopropanol + 1% Ammonium Hydroxide







Appendix 4.1

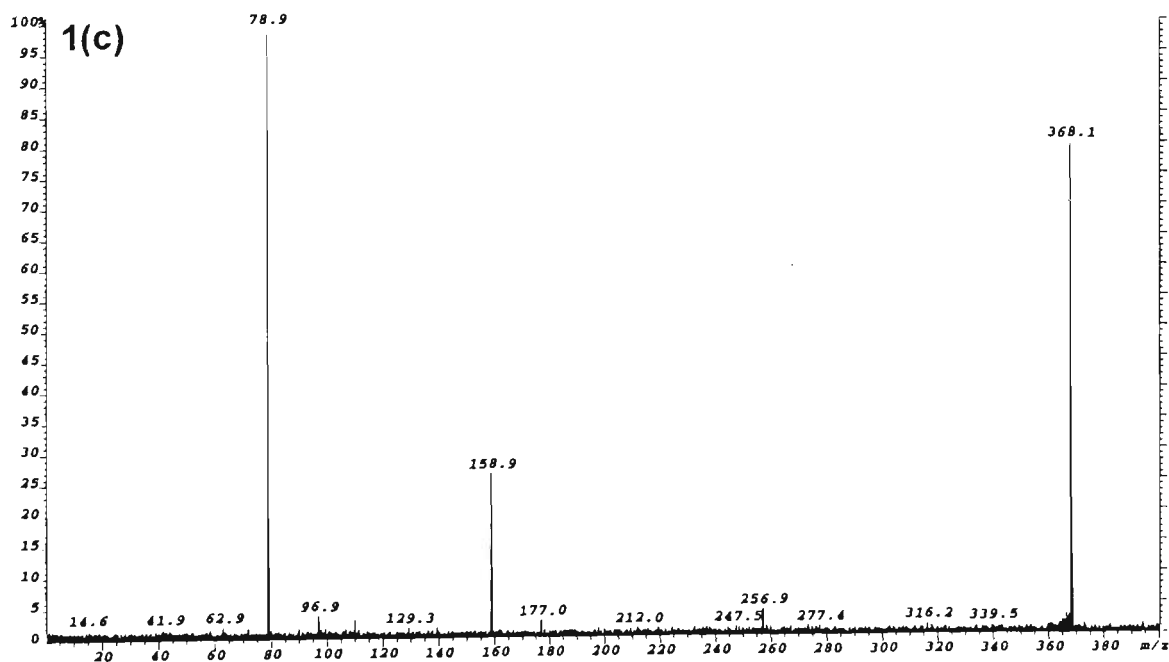
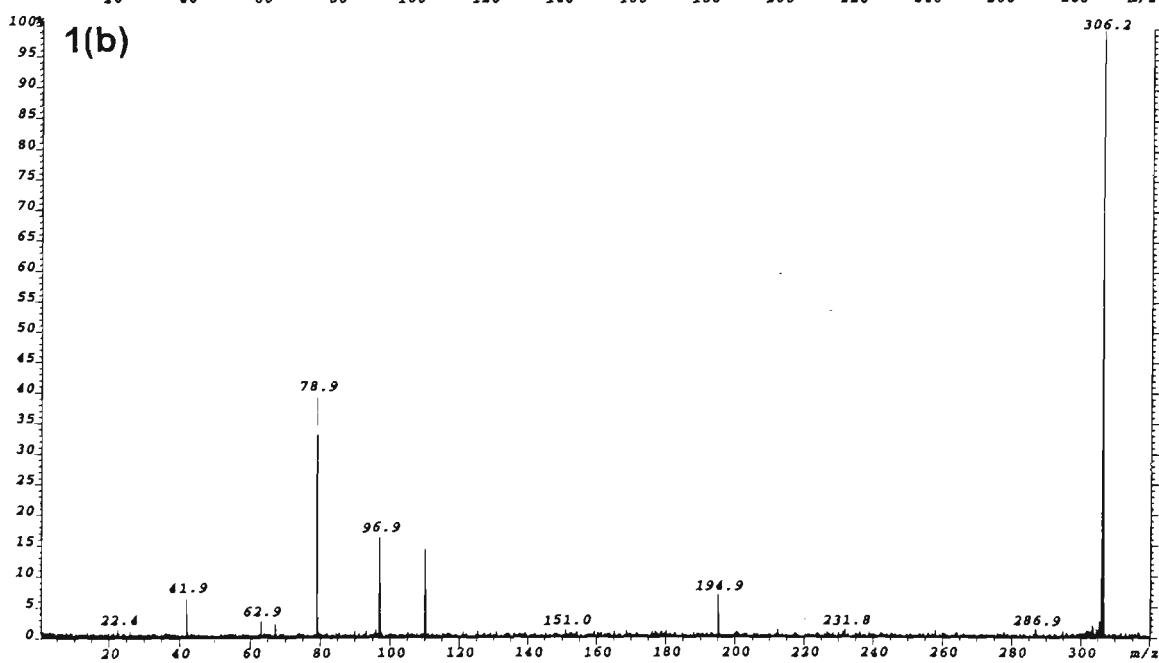
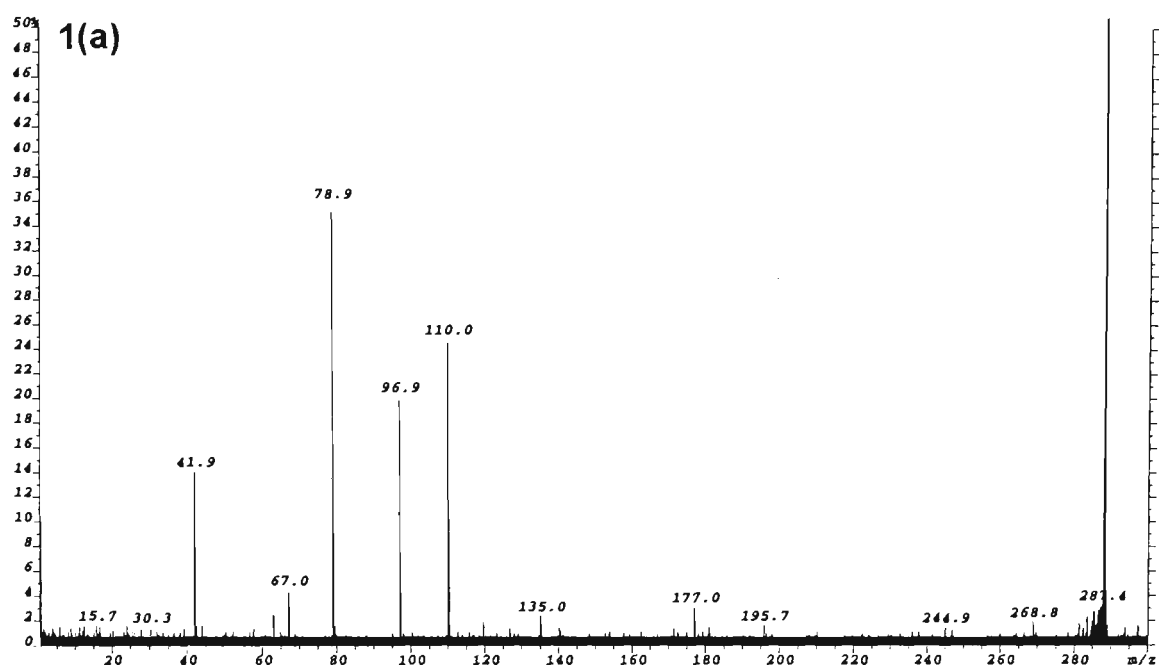
ESI-MS/MS spectra of the ‘source-generated’ series of nucleotide ion types of cytosine, adenine and thymine observed in the ESI-MS/MS spectra of the [M-H]⁻ ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

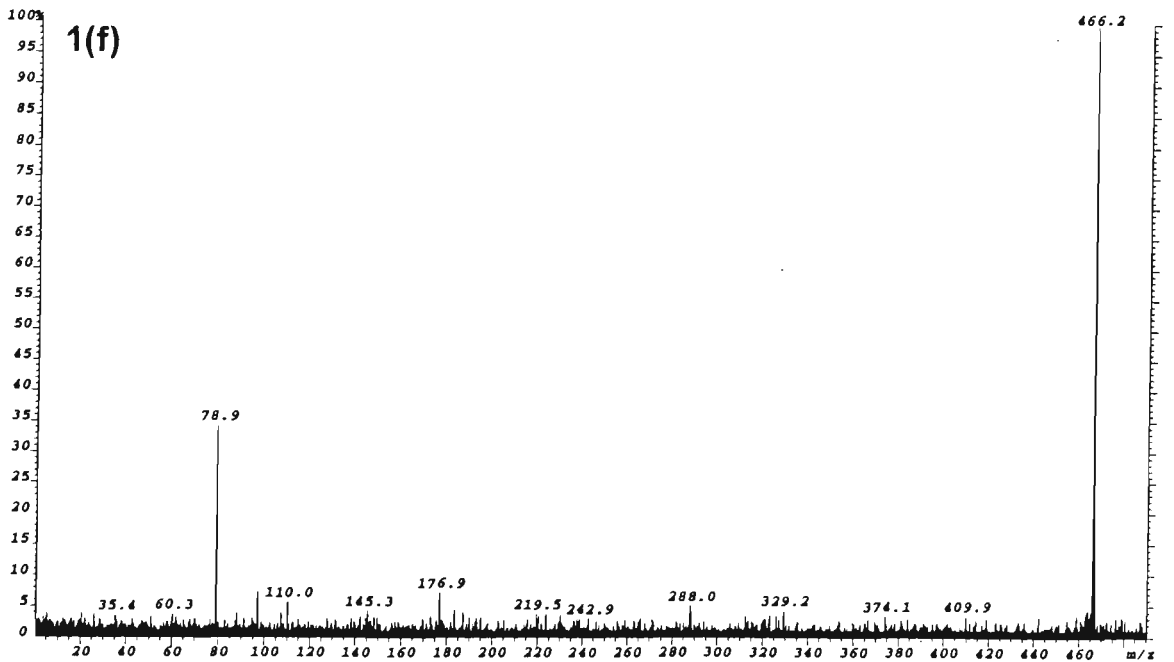
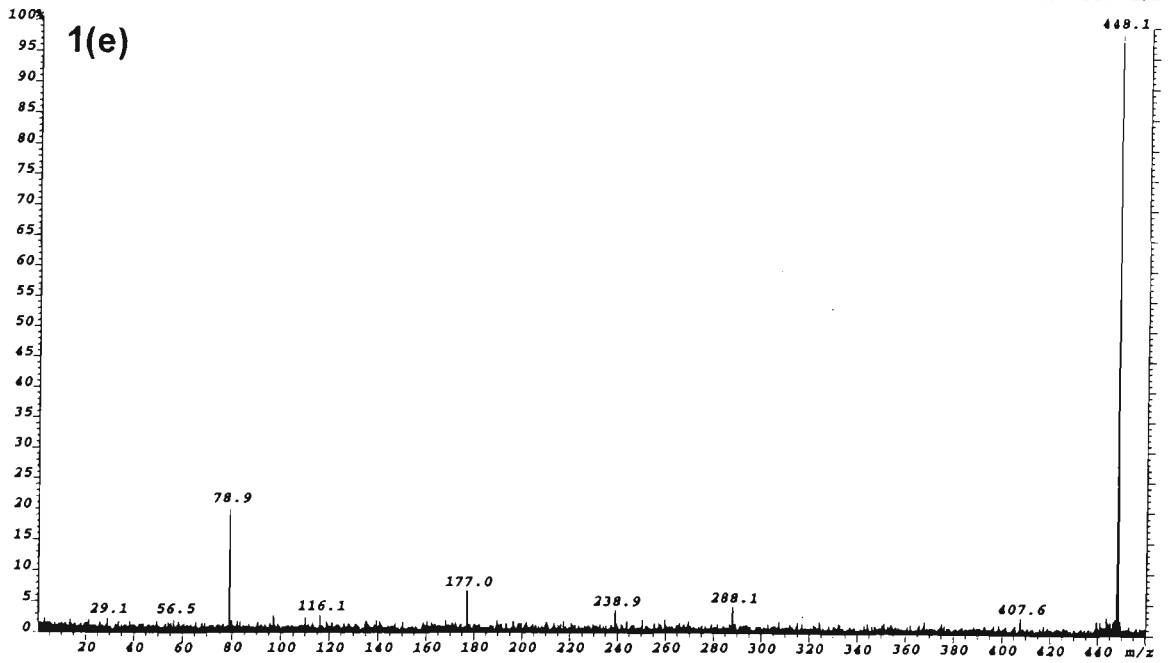
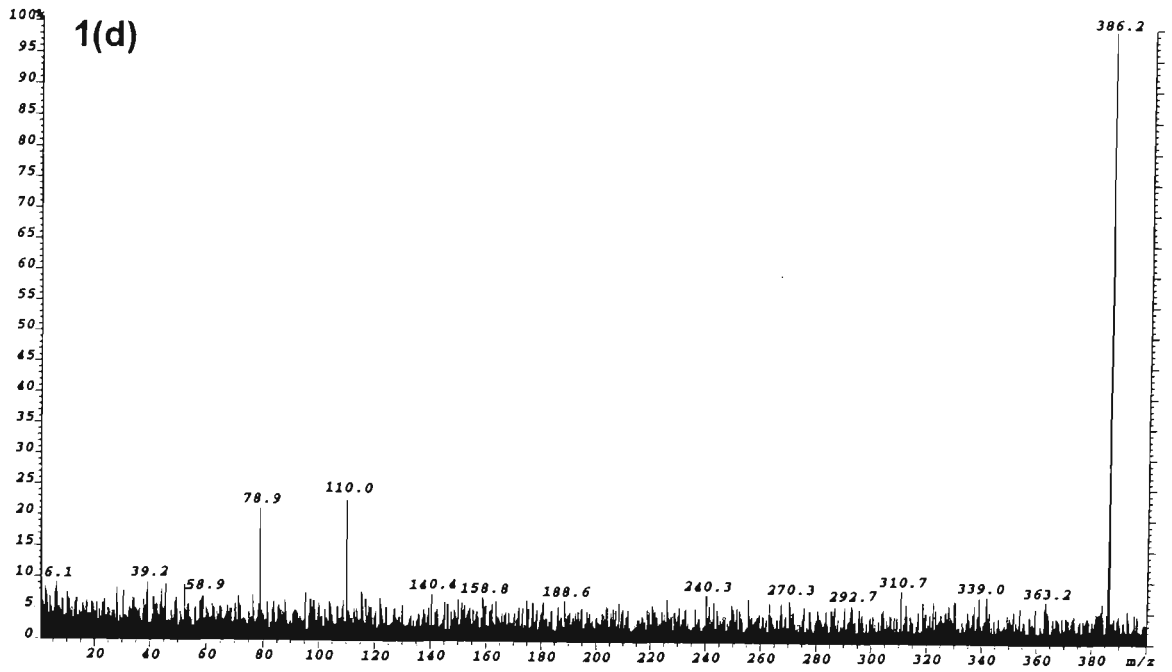
Figure No.	Assignment	Precursor Ion <i>m/z</i>
1(a)	[C _{nt} -H] ⁻	288.1
1(b)	[C _{nt} +H ₂ O-H] ⁻	306.2
1(c)	[C _{nt} +(p or s-H ₂ O)-H] ⁻	368.1
1(d)	[C _{nt} +(s or p+H ₂ O)-H] ⁻	386.2
1(e)	[C _{nt} +p+s-H ₂ O -H] ⁻	448.1
1(f)	[C _{nt} +p+s-H] ⁻	466.2
2(a)	[A _{nt} -H] ⁻	312.1
2(b)	[A _{nt} +(p or s-H ₂ O)-H] ⁻	392.1
2(c)	[A _{nt} +(s or p+H ₂ O)-H] ⁻	410.2
2(d)	[A _{nt} +p+s-H] ⁻	490.2
3(a)	[T _{nt} +H ₂ O-H] ⁻	321.1
3(b)	[T _{nt} +(p or s-H ₂ O)-H] ⁻	383.1
3(c)	[T _{nt} +(s or p+H ₂ O)-H] ⁻	401.3
3(d)	[T _{nt} +p+s-H] ⁻	481.2
3(e)	[T _{nt} +s+p+H ₂ O)-H] ⁻	499.2

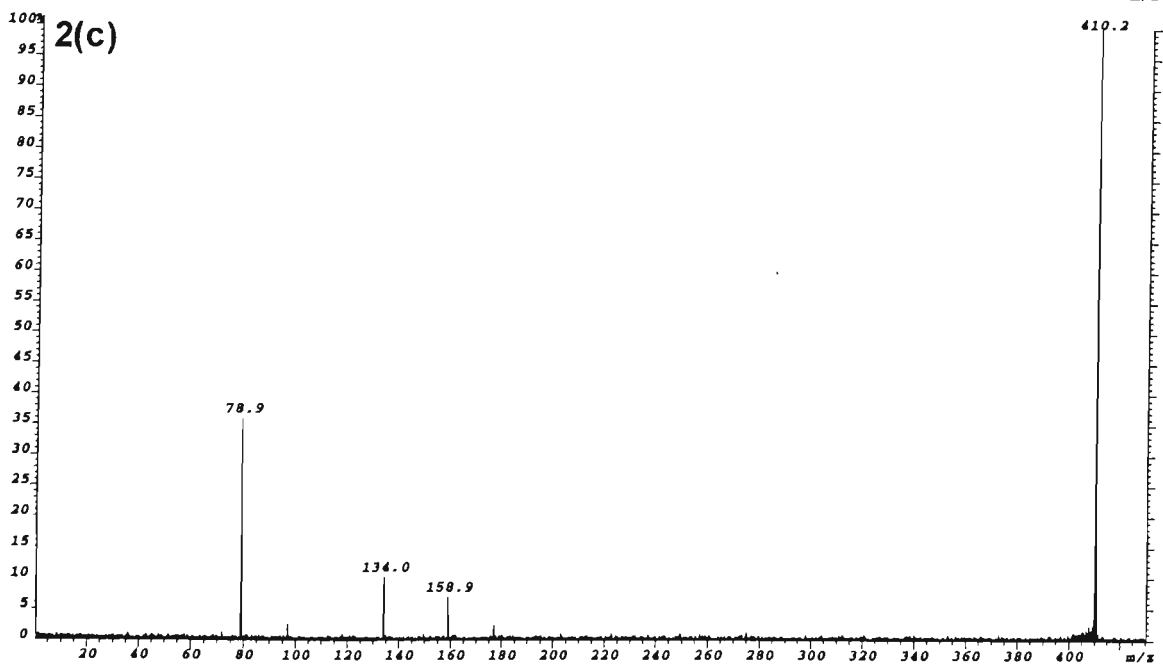
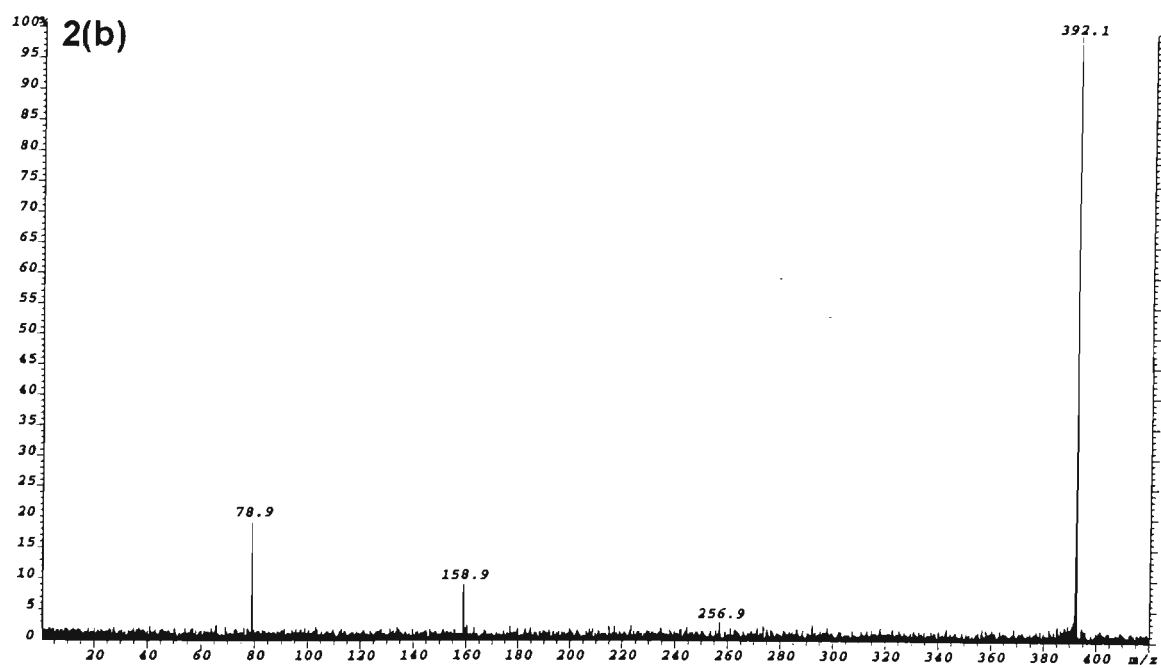
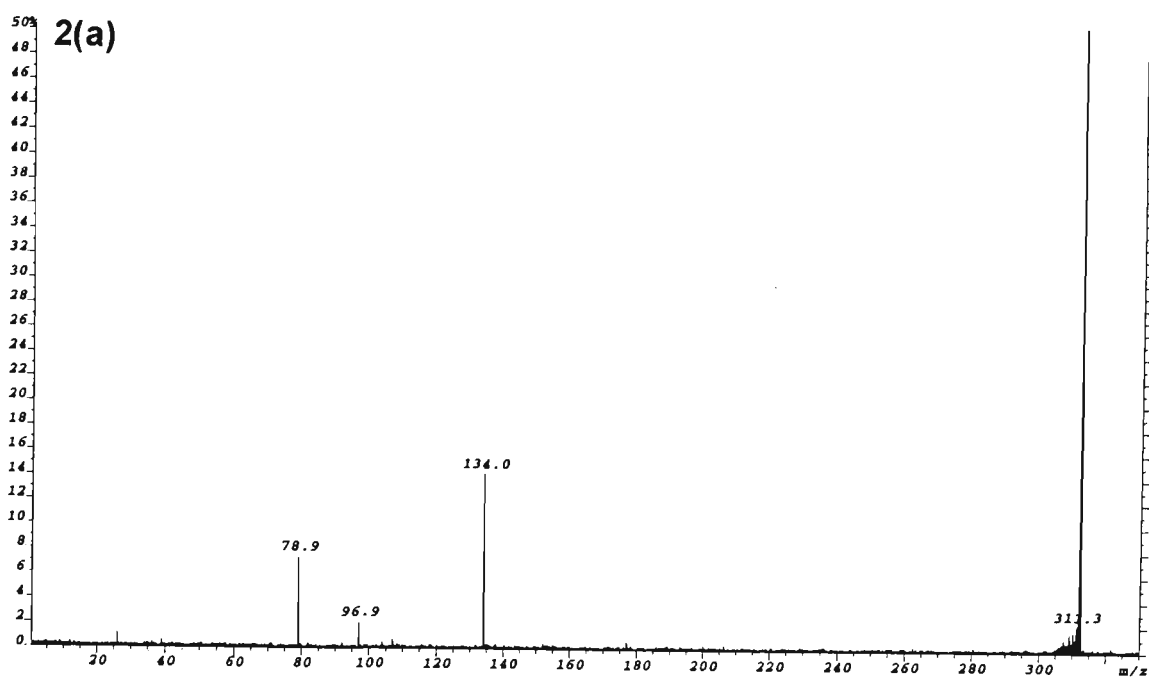
s = deoxyribose-H₂O (C₅H₆O₂), *Mr* =98.0368 Da

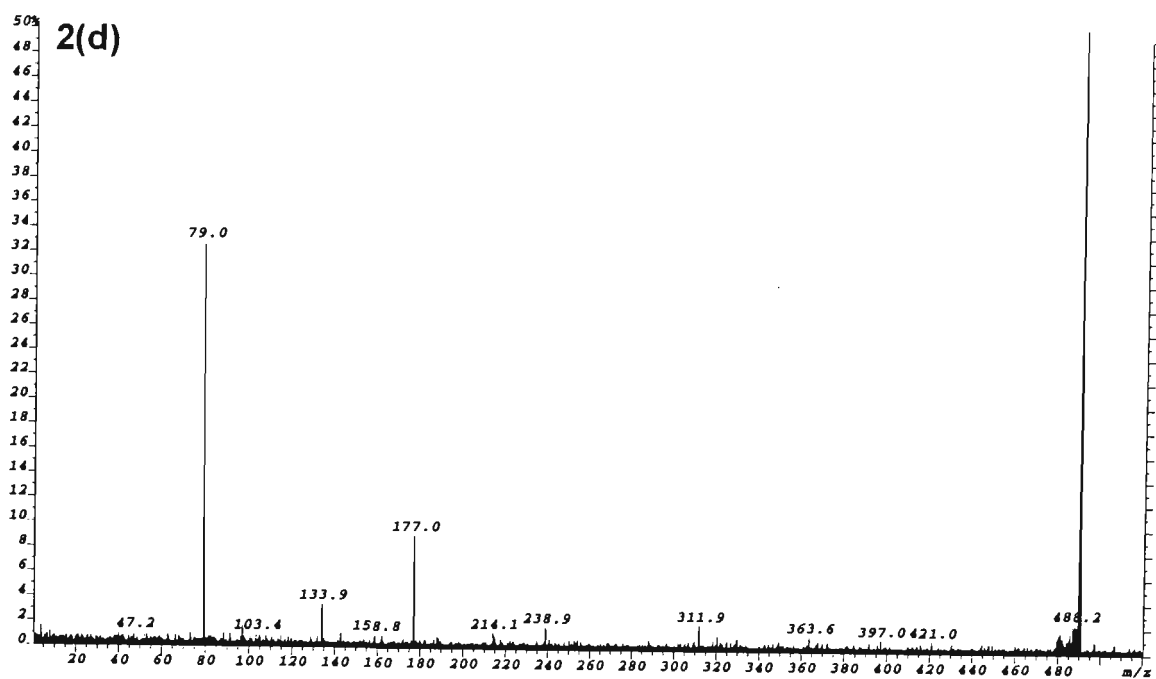
p = PO₃H, *Mr* = 79.9663 Da

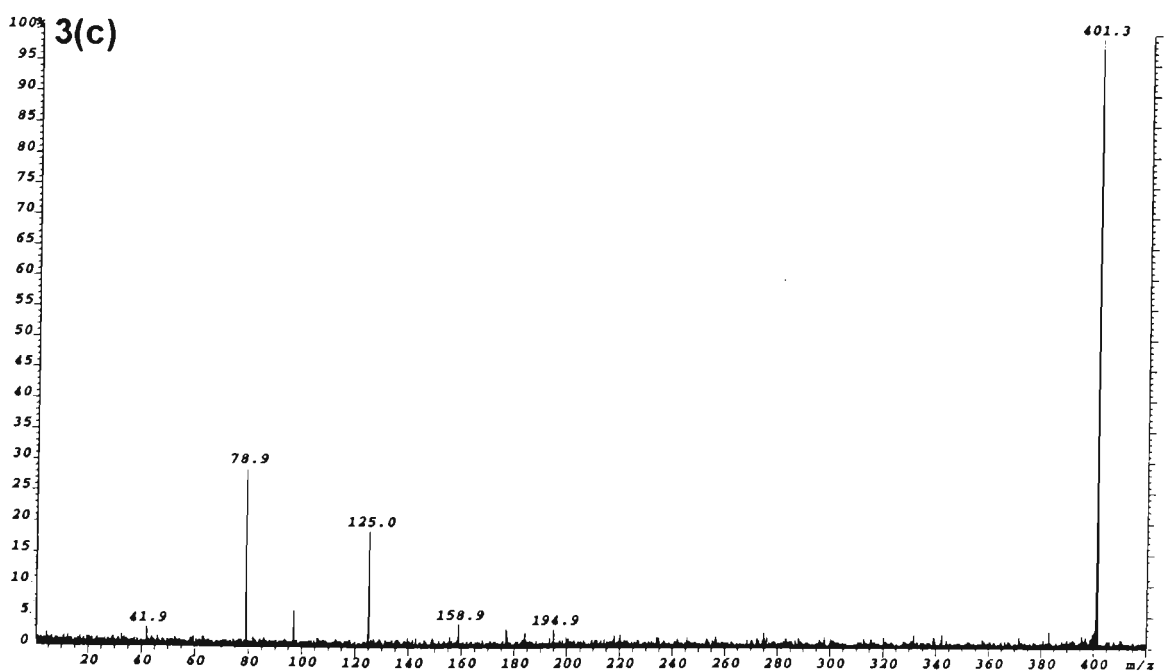
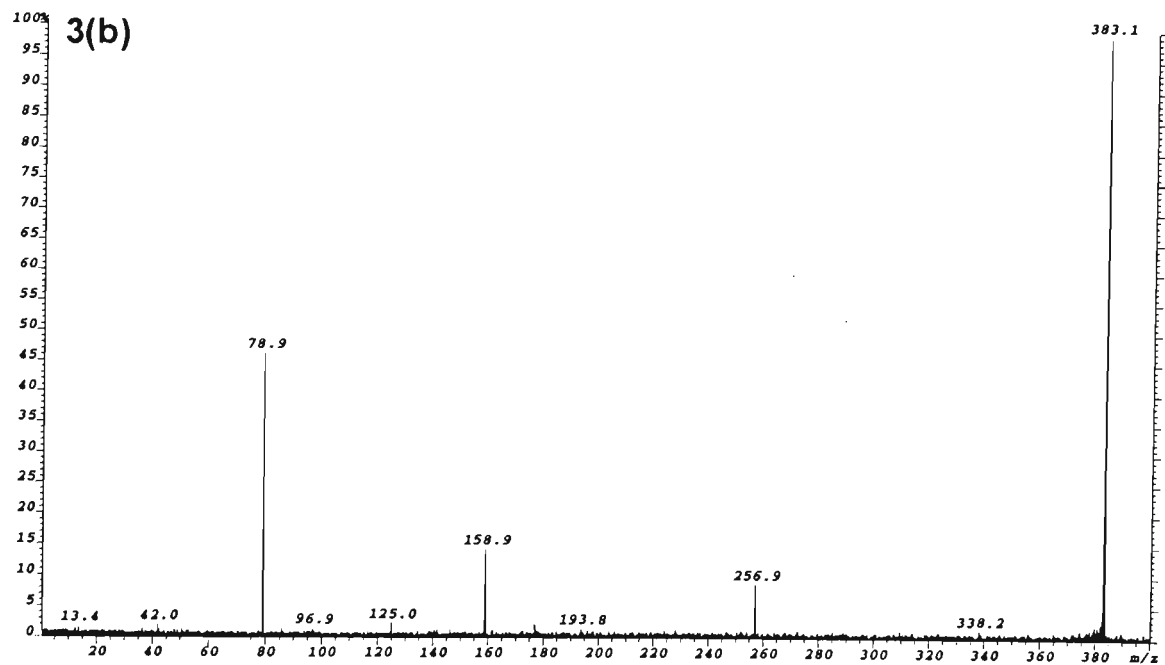
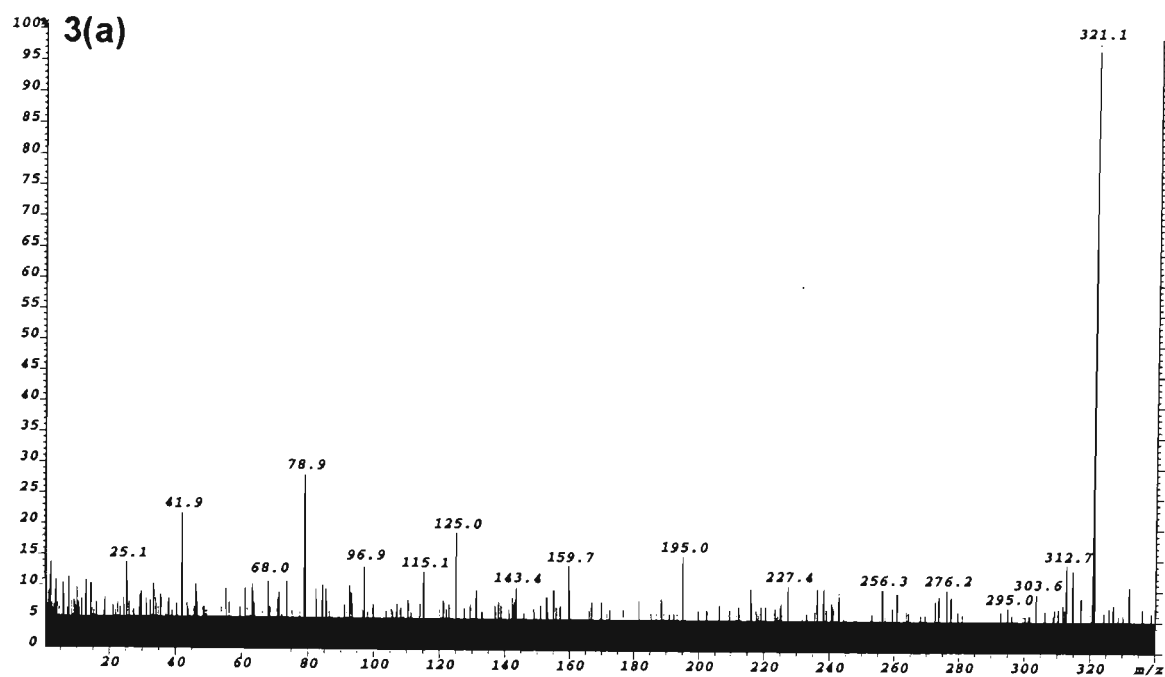
B_{nt} denotes a mononucleotide which may be either psB or sBp

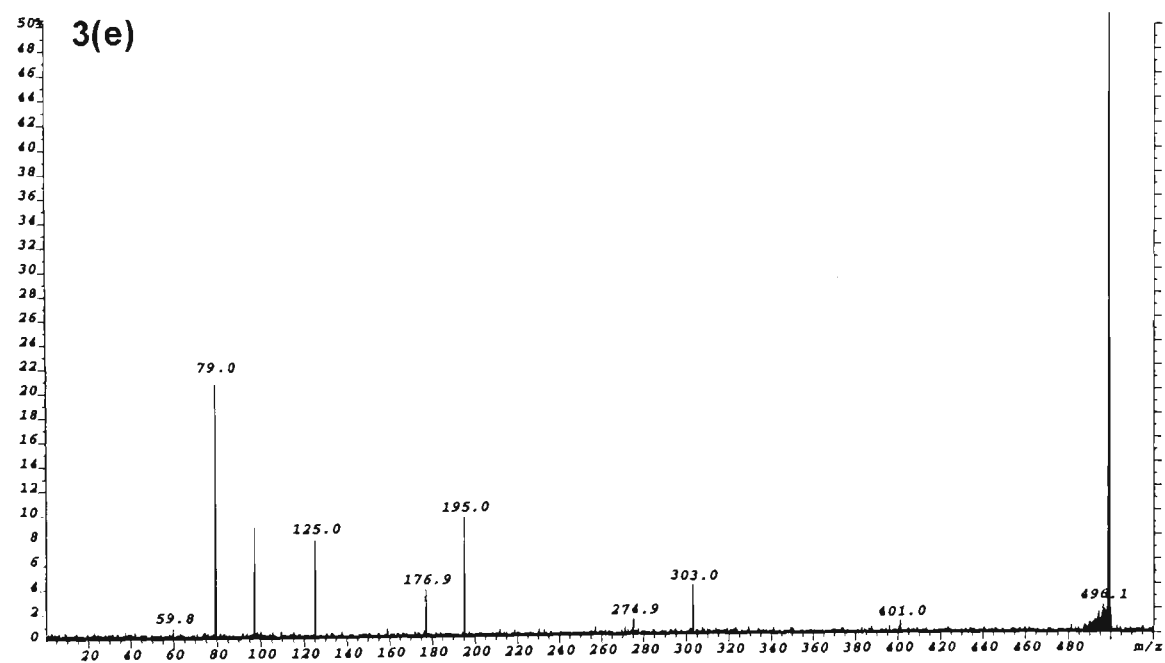
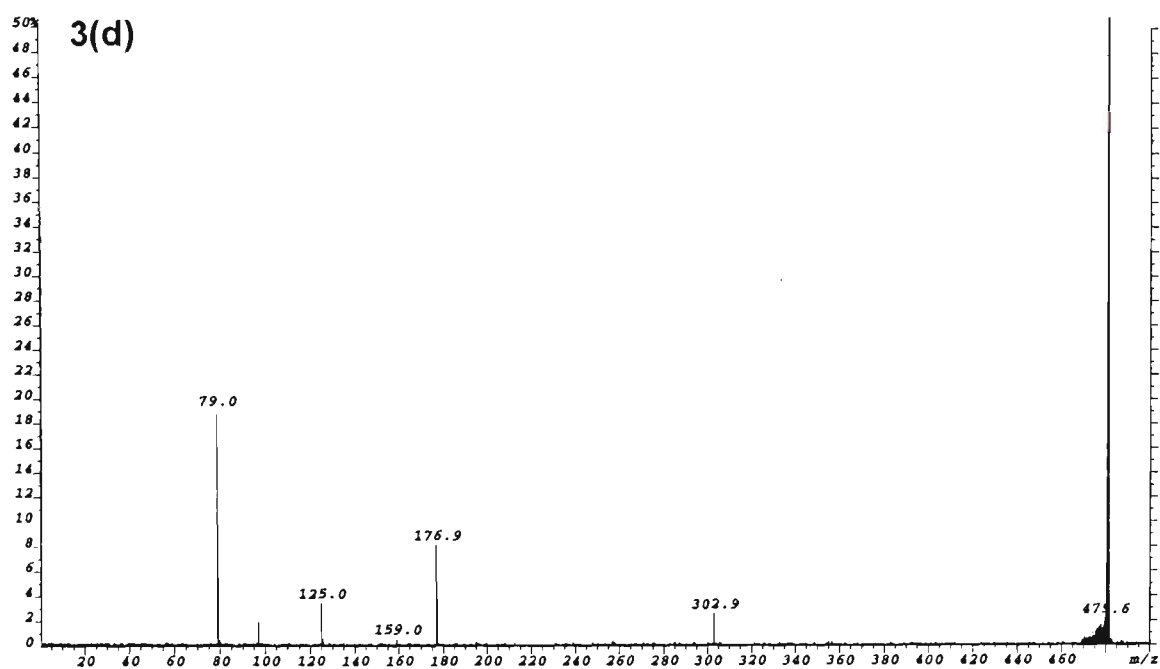










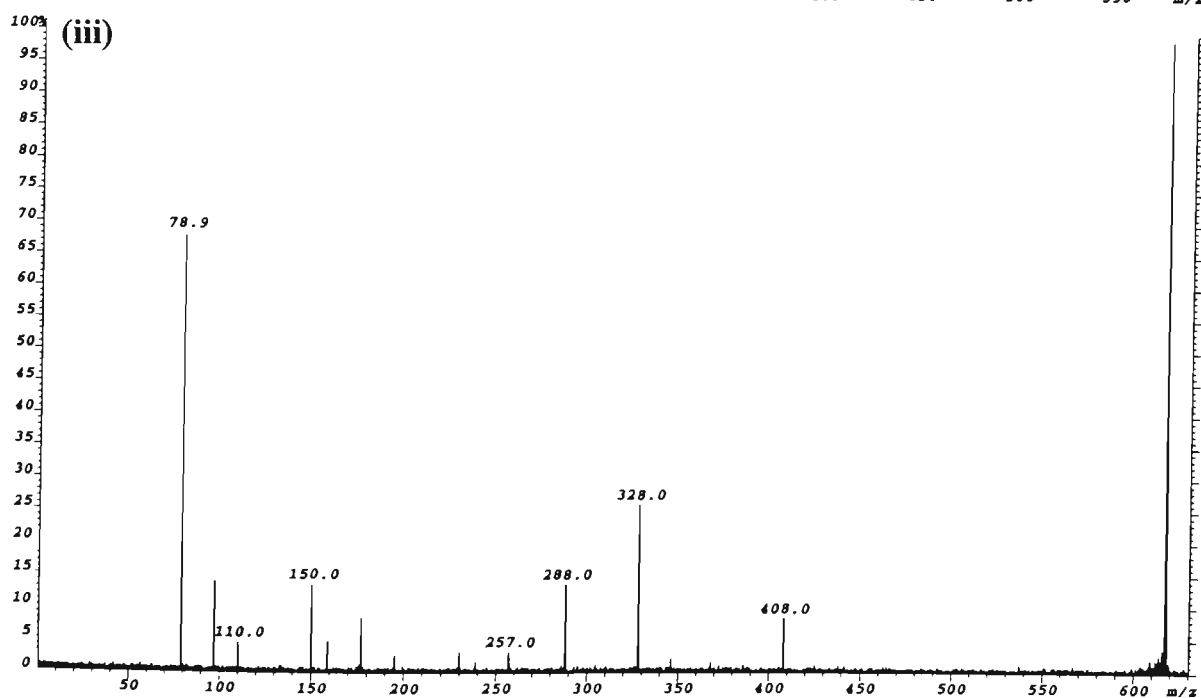
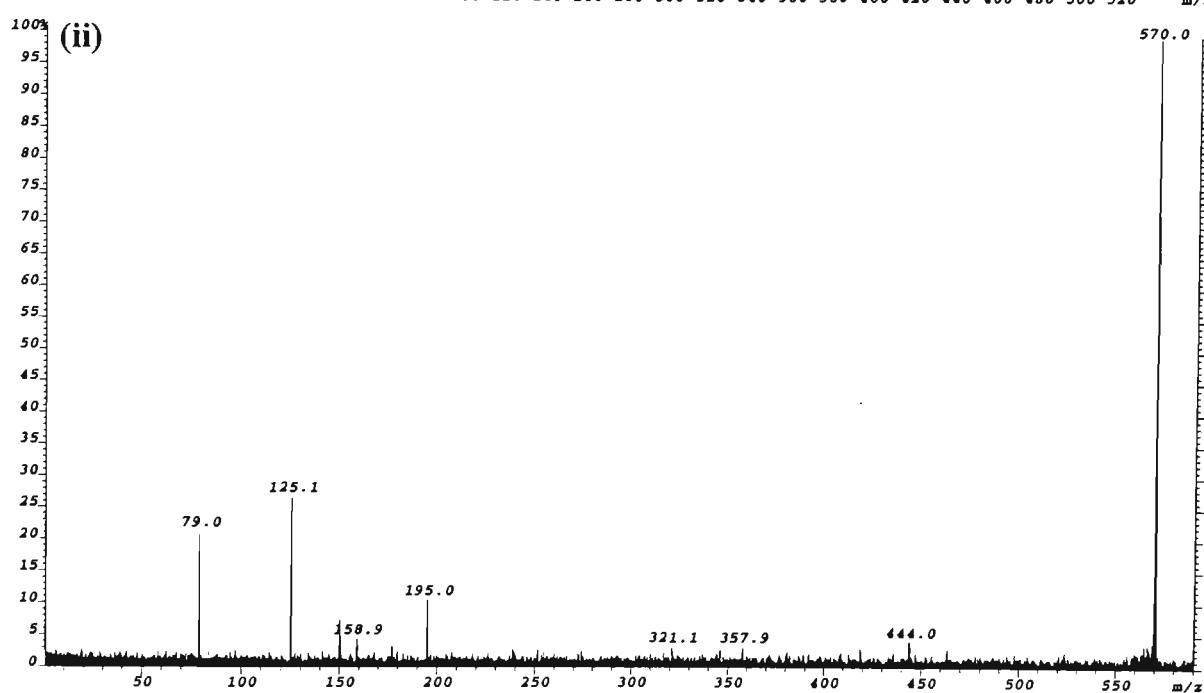
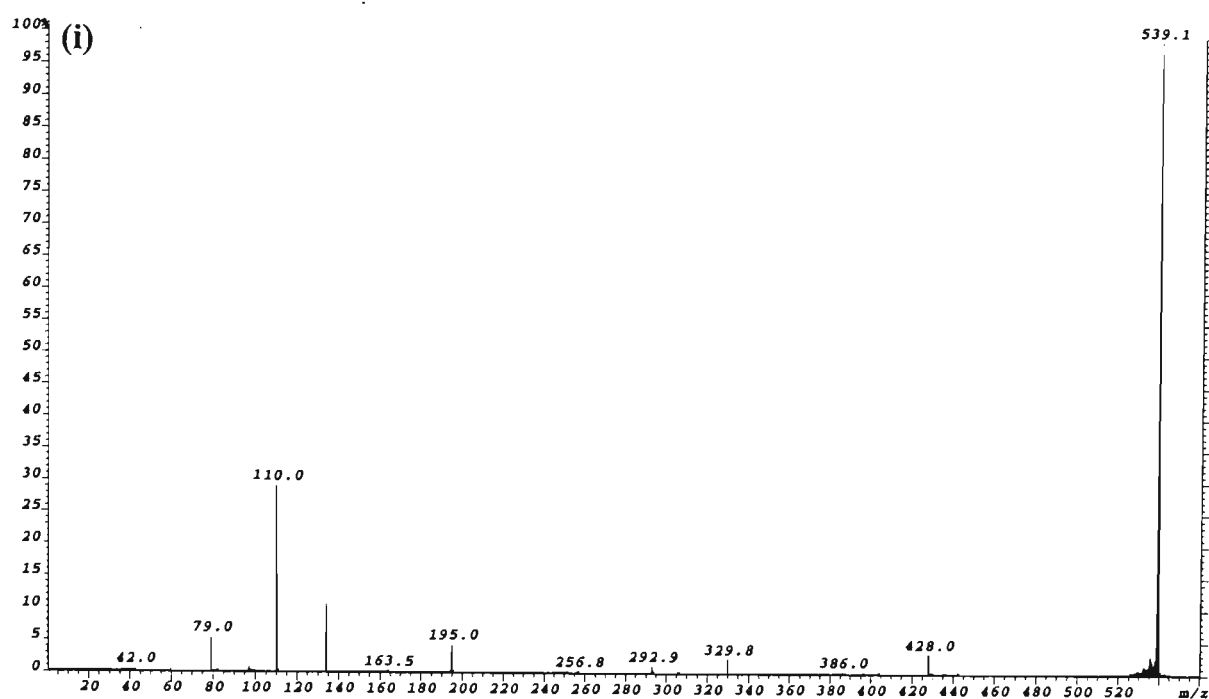


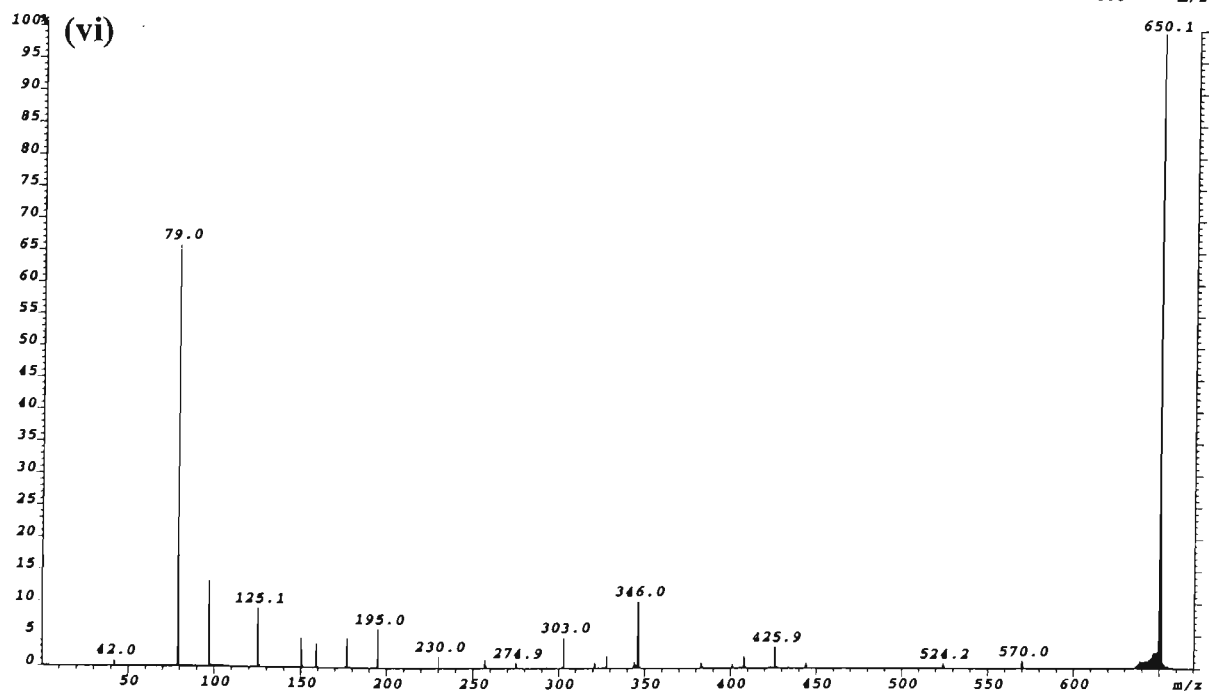
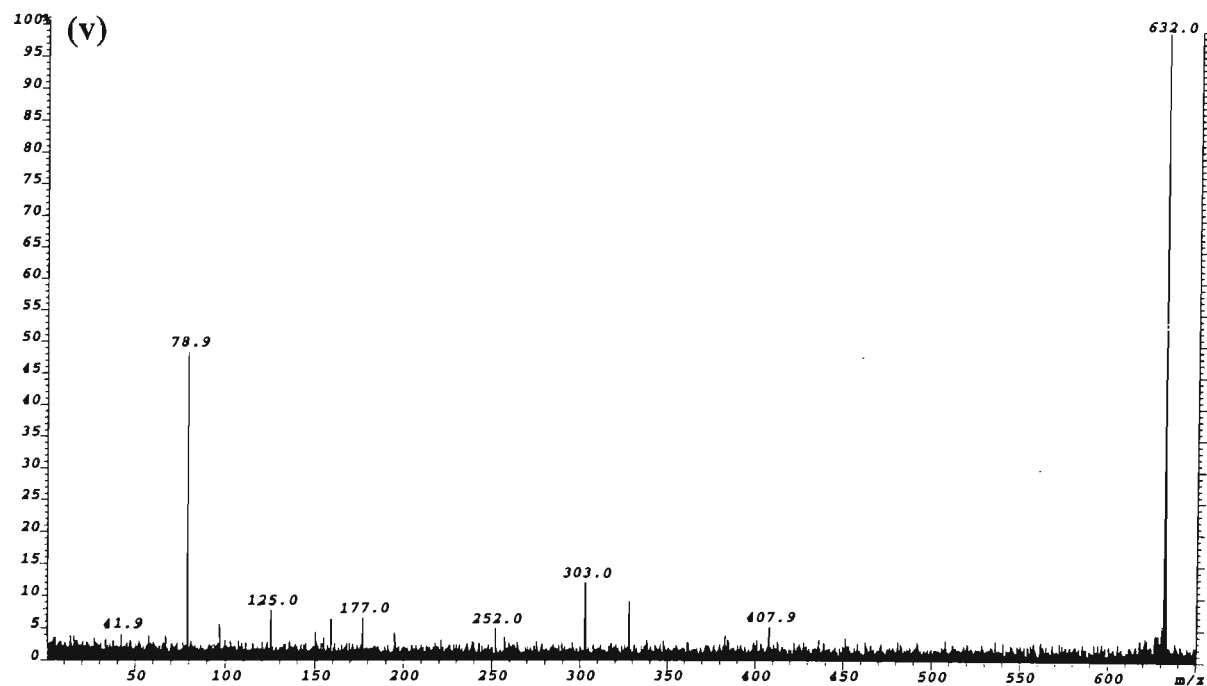
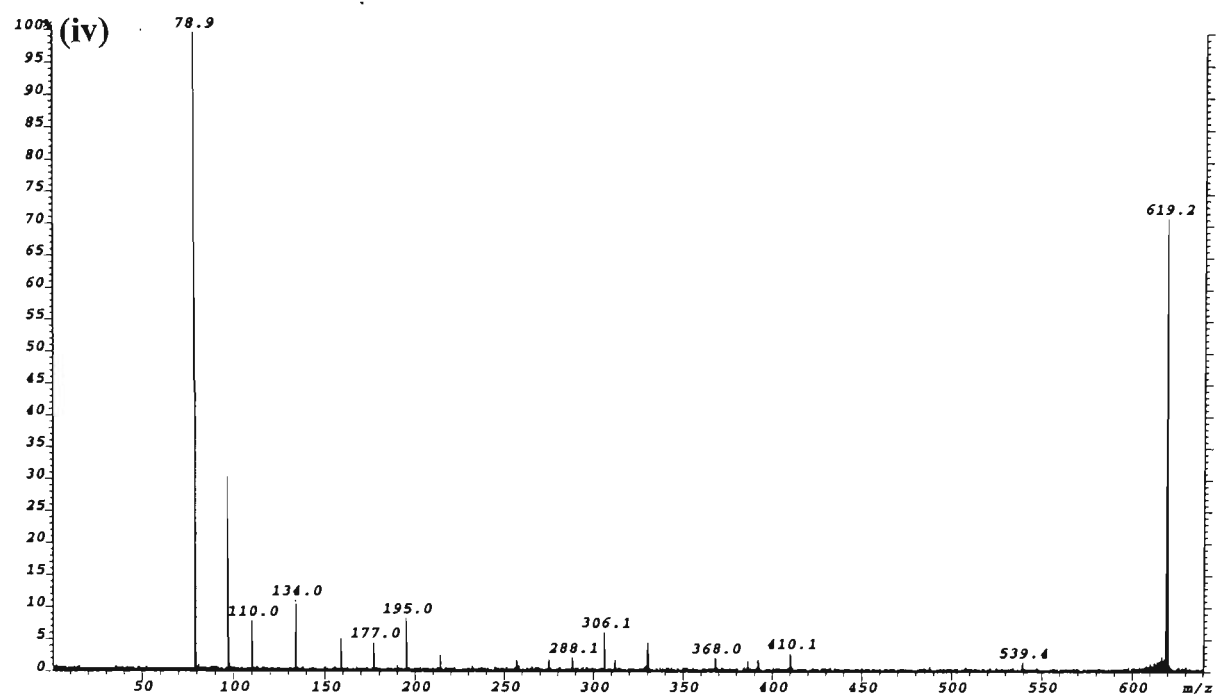
Appendix 4.2

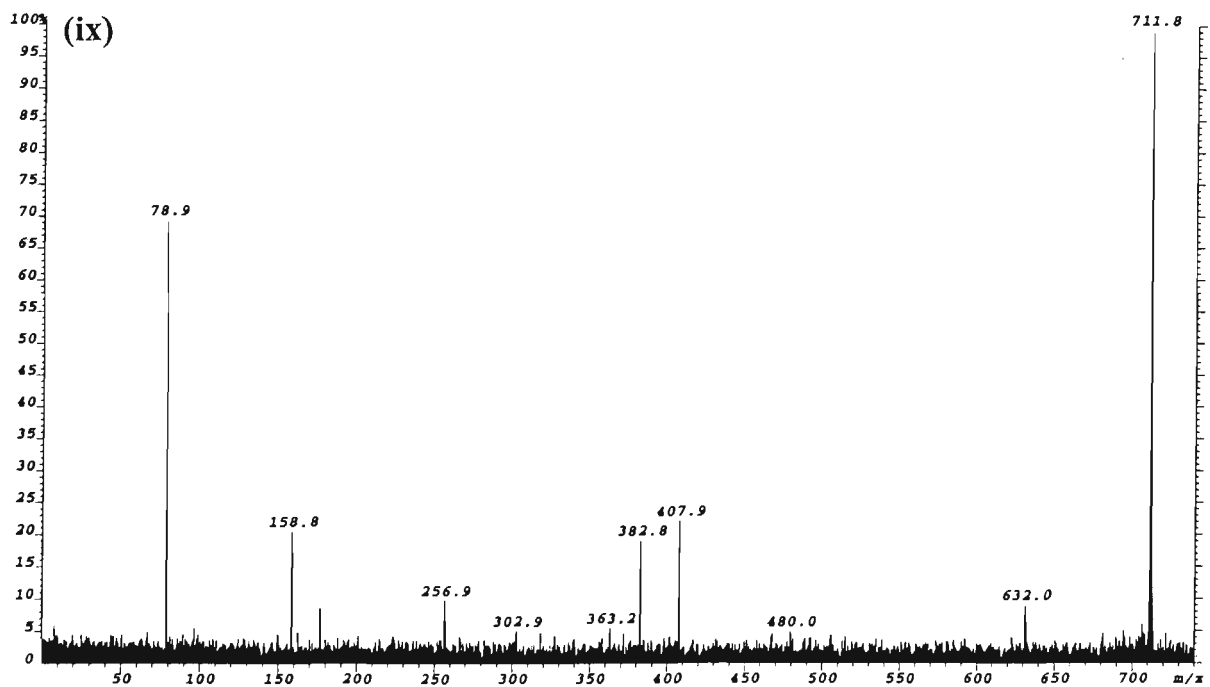
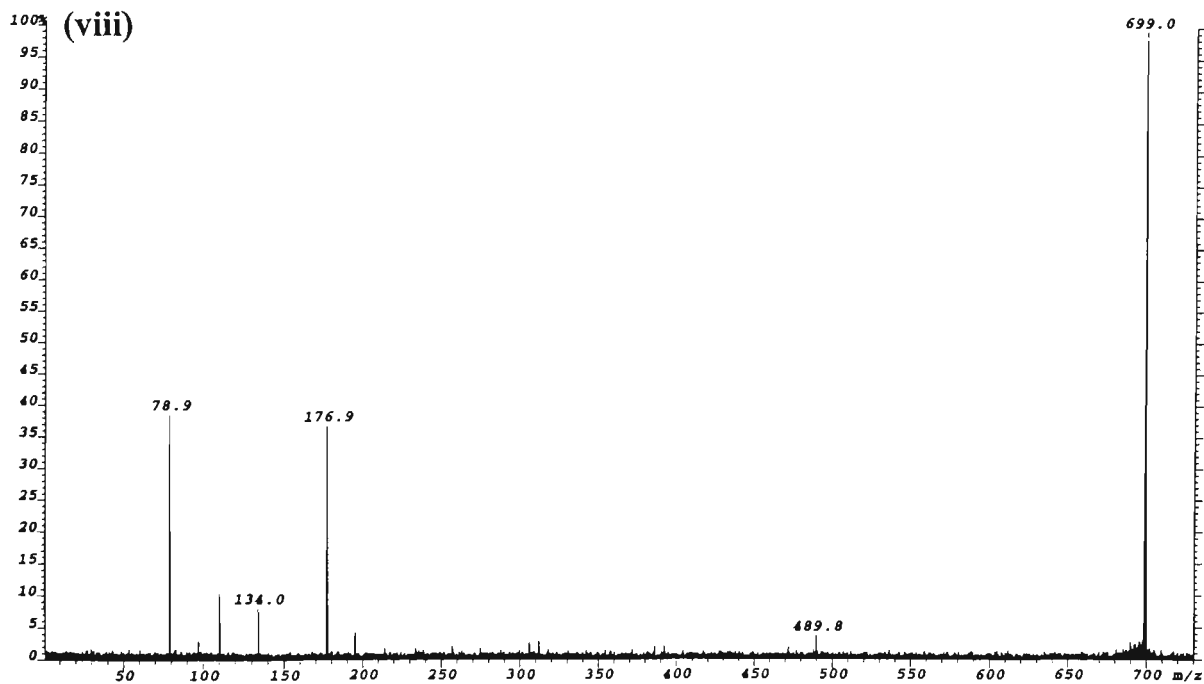
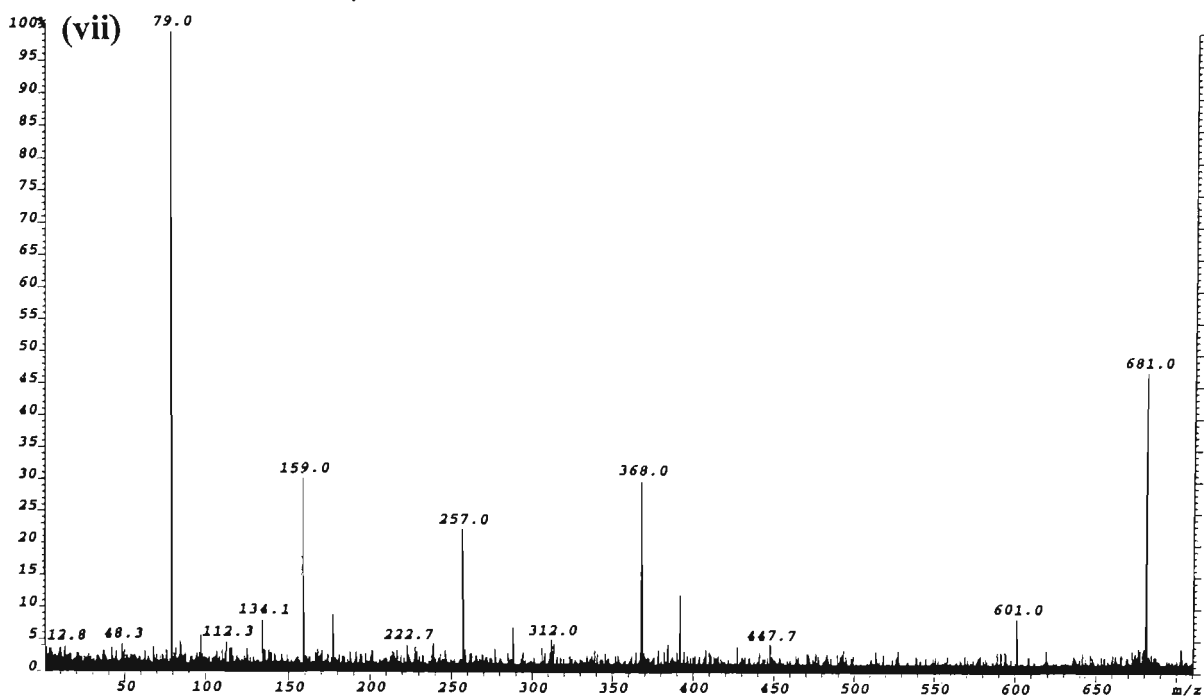
ESI-MS/MS spectra of the ‘source-generated’ polynucleotide product ions observed in the ESI-MS/MS spectrum of the [M-H]⁻ ion of 5'-d(CACGTG)-3'.

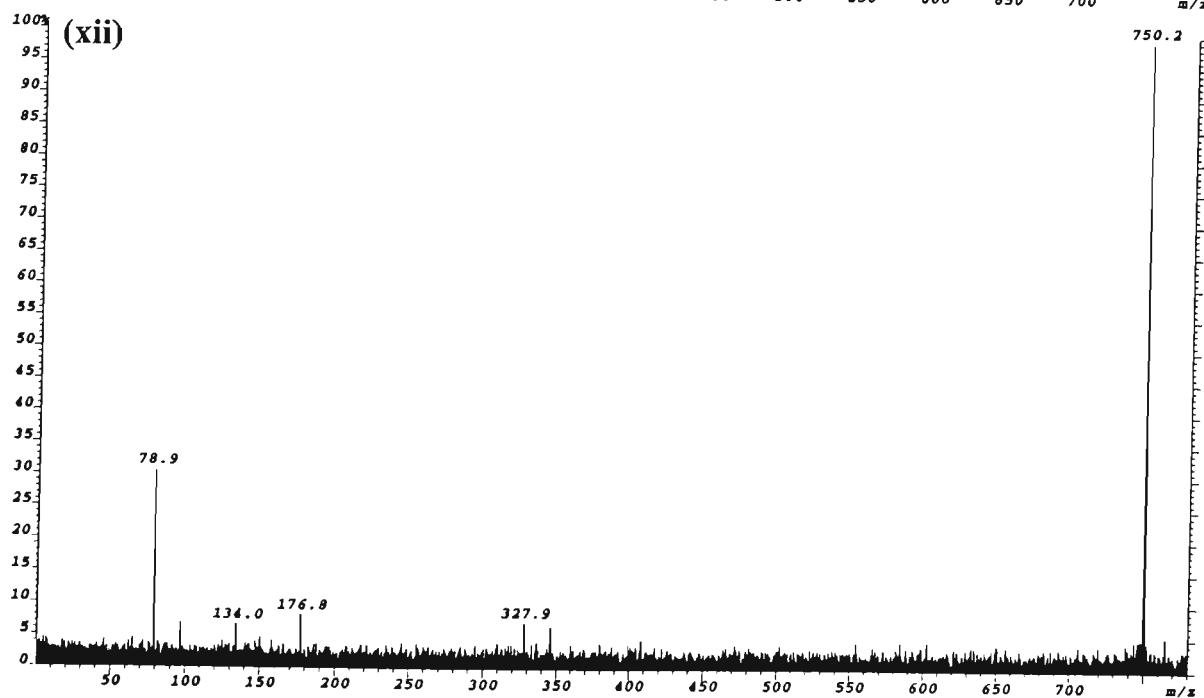
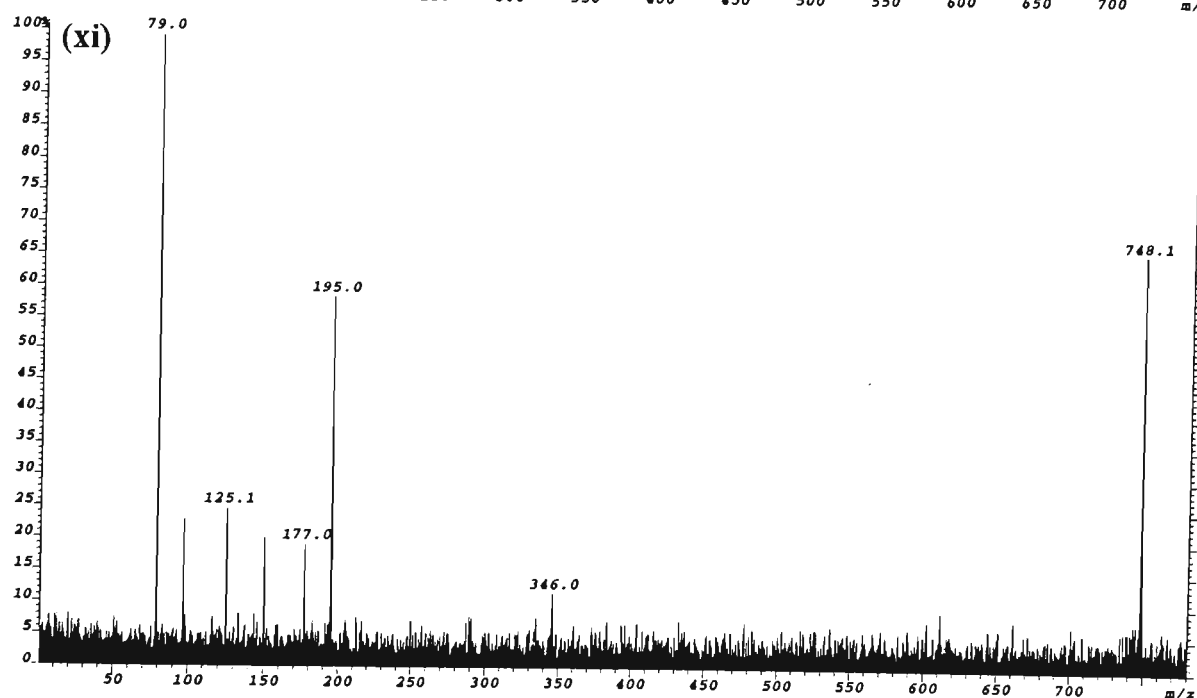
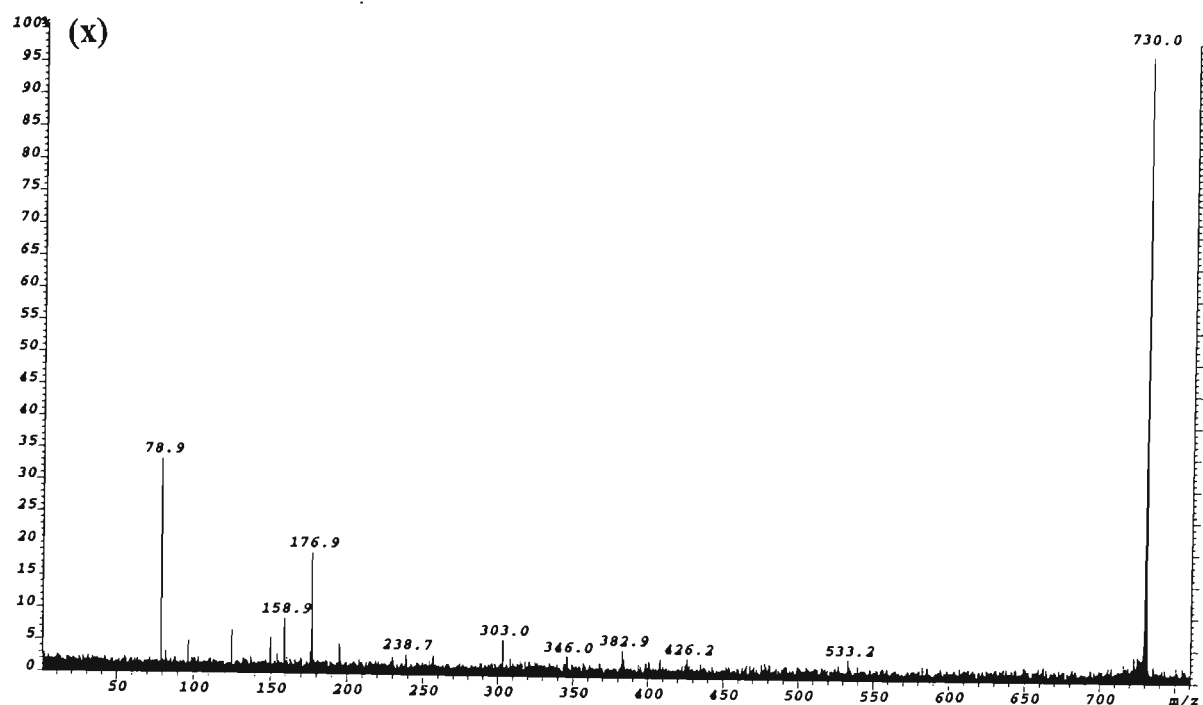
Figure No.	Precursor Ion	
	m/z	Assignment
(i)	539.1	b ₂ ⁻ and/or [(AC)-p+H ₂ O-H] ⁻
(ii)	570.0	y ₂ ⁻ and/or [(GT)-H ₂ O-H] ⁻
(iii)	617.1	[(CG)-H] ⁻
(iv)	619.2	d ₂ ⁻ and/or [(AC)+H ₂ O-H] ⁻
(v)	632.0	[x ₂ -2H] ⁻ and/or [(GT)-H] ⁻
(vi)	650.1	w ₂ ⁻ and/or [(GT)+H ₂ O-H] ⁻
(vii)	681.0	[(AC)+p-H] ⁻ and/or [(AC)+s-H ₂ O-H] ⁻
(viii)	699.0	[a ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s-H] ⁻ and/or [(AC)+p+H ₂ O-H] ⁻
(ix)	711.8	[(GT)+s-H ₂ O-H] ⁻ and/or [(GT)+p-H] ⁻
(x)	730.0	[z ₃ -GH-2H] ⁻ and/or [(GT)+s-H] ⁻
(xi)	748.0	[y ₃ -GH] ⁻ and/or [(GT)+s+H ₂ O-H] ⁻
(xii)	750.2	
(xiii)	761.1	[(CA and/or AC)+s+p-H ₂ O-H] ⁻
(xiv)	779.1	[c ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s+p-H] ⁻
(xv)	795.1	[(CG)+s+p-H] ⁻
(xvi)	809.9	[a ₃ -2H] ⁻ and/or [x ₃ -GH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
(xvii)	828.0	[w ₃ -GH] ⁻ and/or [(GT)+s+p+H ₂ O-H] ⁻
(xviii)	877.0	[a ₄ -GH-CH-2H] ⁻ and/or [(AC)+s+p+s-H] ⁻ or [(AC)+p+s+p+H ₂ O-H] ⁻
(xix)	881.0	[z ₃ -2H] ⁻
(xx)	890.1	[c ₃ -2H] ⁻
(xxi)	899.1	[c ₃ -2H] ⁻
(xxii)	921.1	[(CGT)-H] ⁻
(xxiii)	930.1	[(ACG)-H] ⁻
(xxiv)	961.1	[x ₃ -2H] ⁻
(xxv)	979.3	w ₃ ⁻
(xxvi)	988.1	[a ₄ -B ₄ H-2H] ⁻
(xxvii)	1019.4	[z ₄ -GH-2H] ⁻ and/or [(CGT)+s-H] ⁻ or [(CGT)+p+H ₂ O-H] ⁻
(xxviii)	1028.1	[a ₄ -CH-2H] ⁻ and/or [(ACG)+s-H] ⁻
(xxix)	1099.2	[x ₄ -GH-2H] ⁻ and/or [(CGT)+s+p-H] ⁻
(xxx)	1117.1	[w ₄ -GH] ⁻ and/or [(CGT)+s+p+H ₂ O-H] ⁻
(xxxi)	1139.4	[a ₄ -2H] ⁻
(xxxii)	1250.2	[x ₄ -2H] ⁻
(xxxiii)	1268.2	w ₄ ⁻
(xxxiv)	1430.4	[w ₅ -GH] ⁻ and/or [(ACGT)+s+p+H ₂ O-H] ⁻
(xxxv)	1581.3	w ₅ ⁻
(xxxvi)	1639.1	[M-GH-H] ⁻

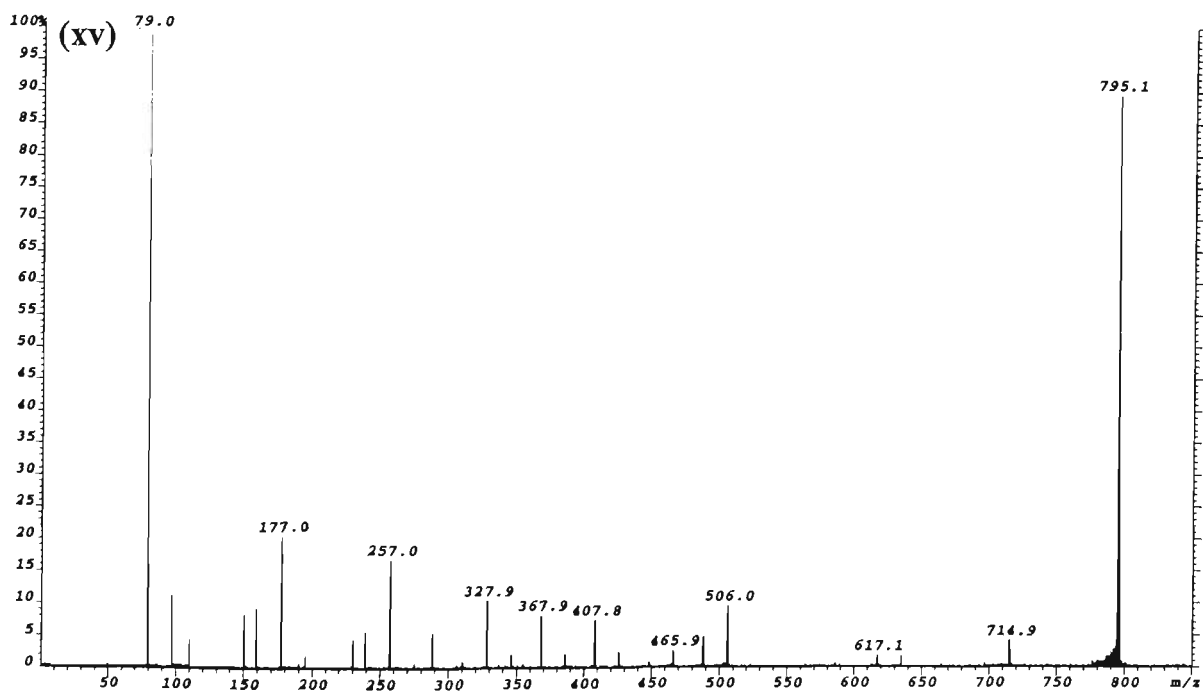
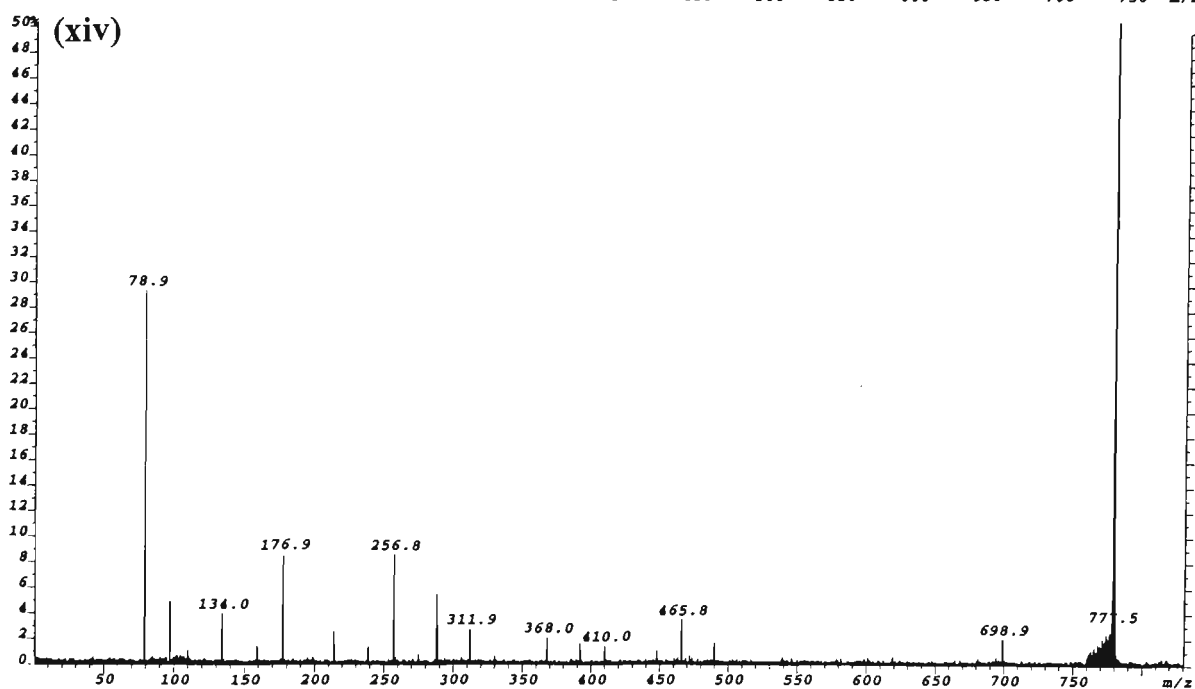
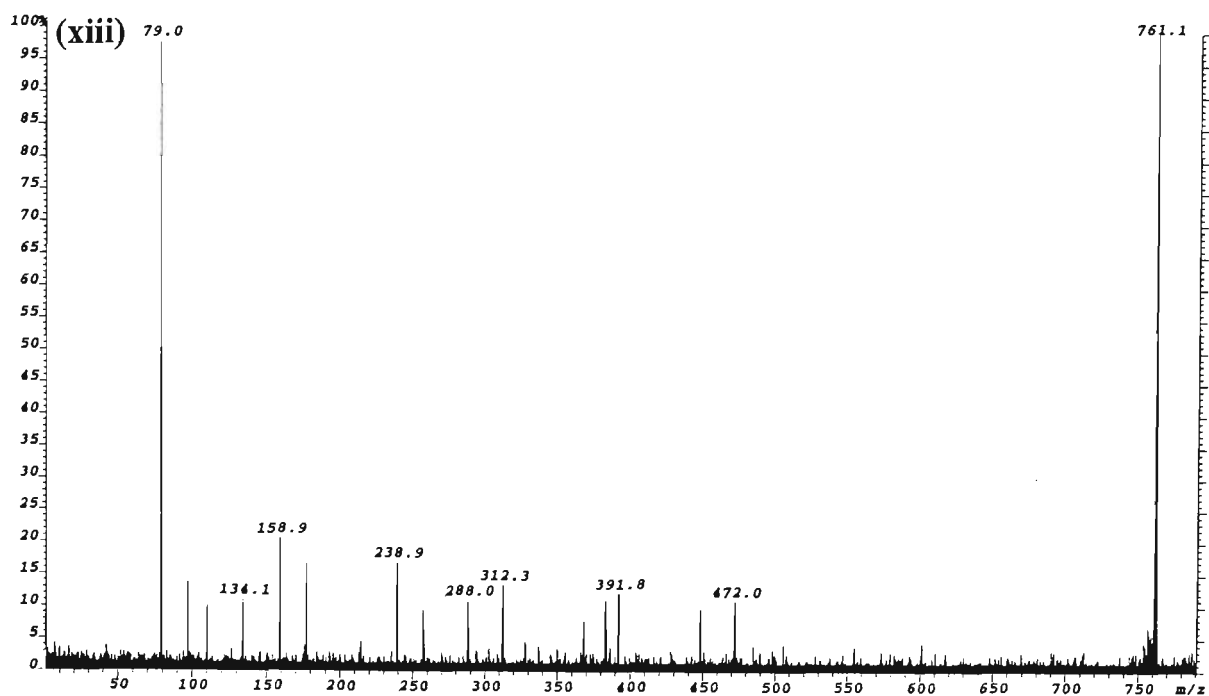
s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da
p = PO₃H, M_r = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

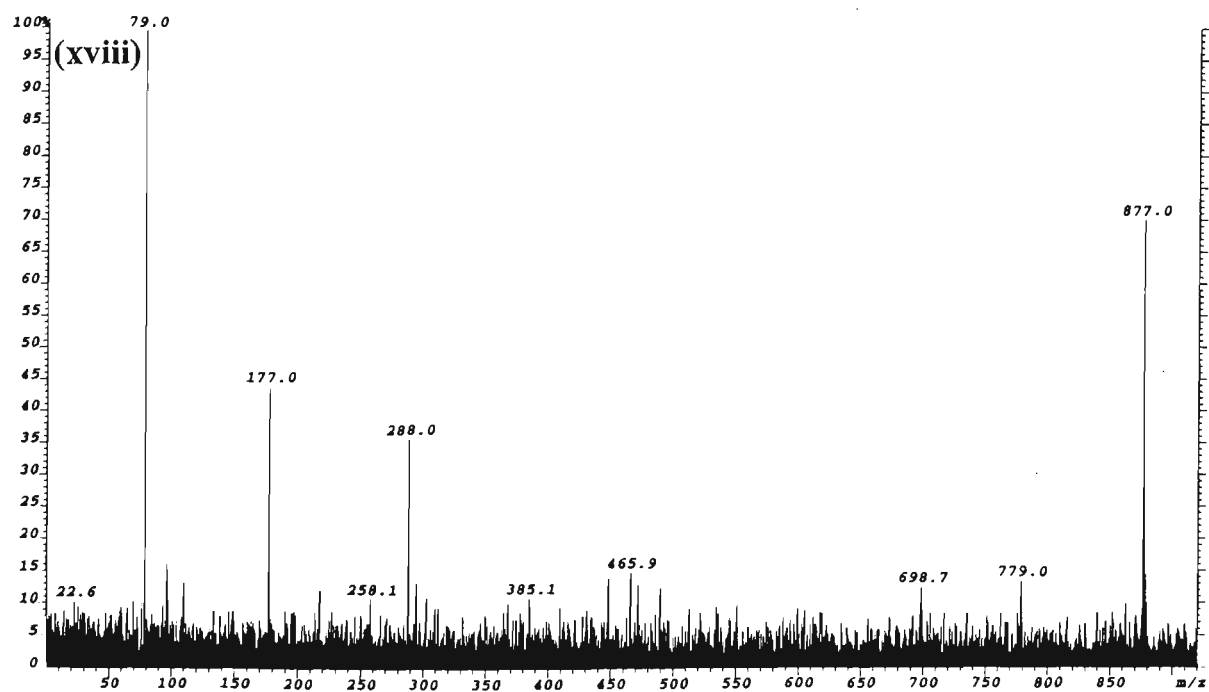
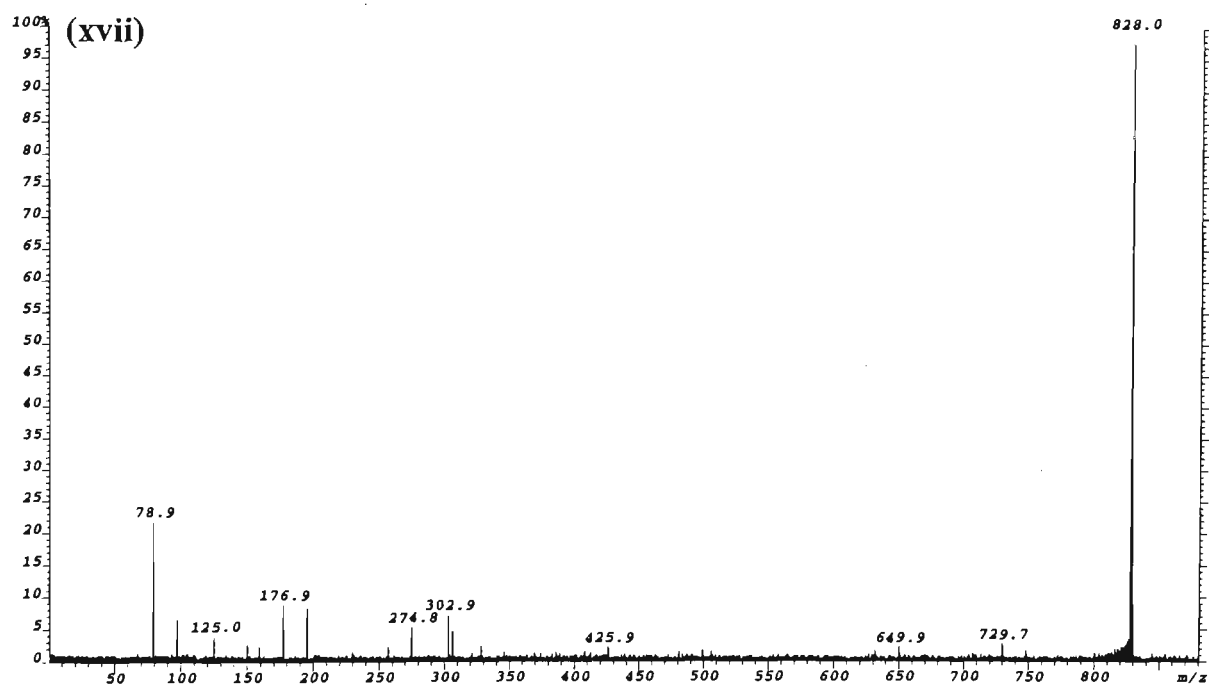
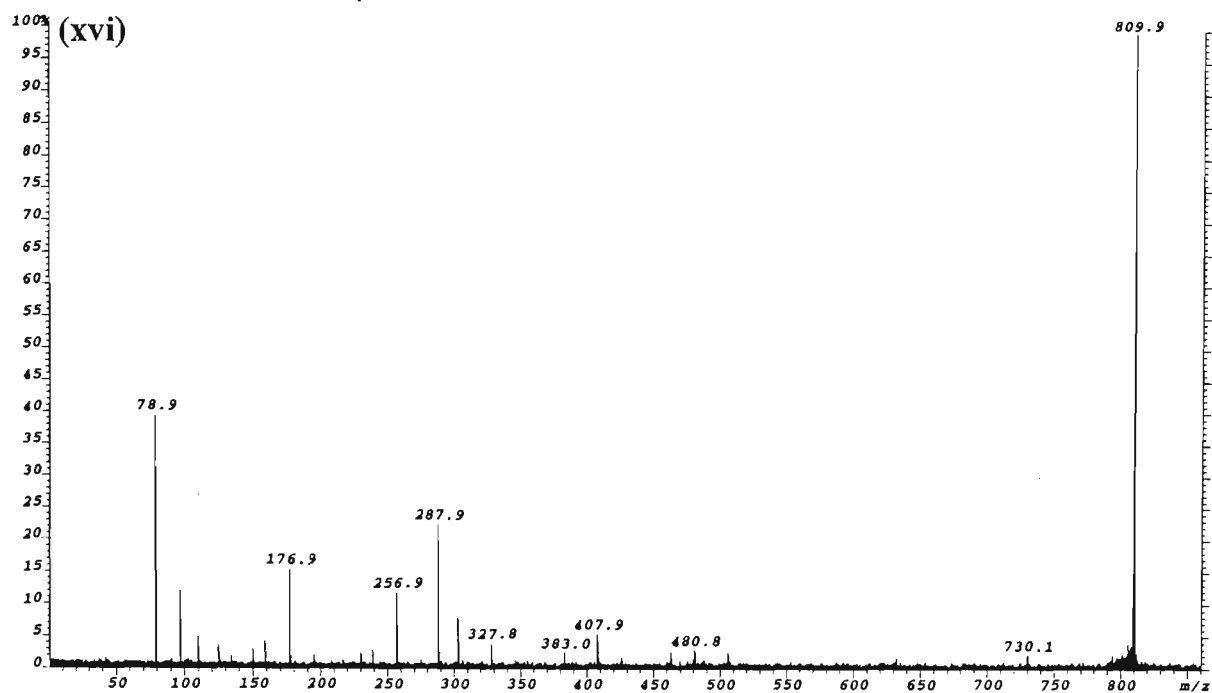


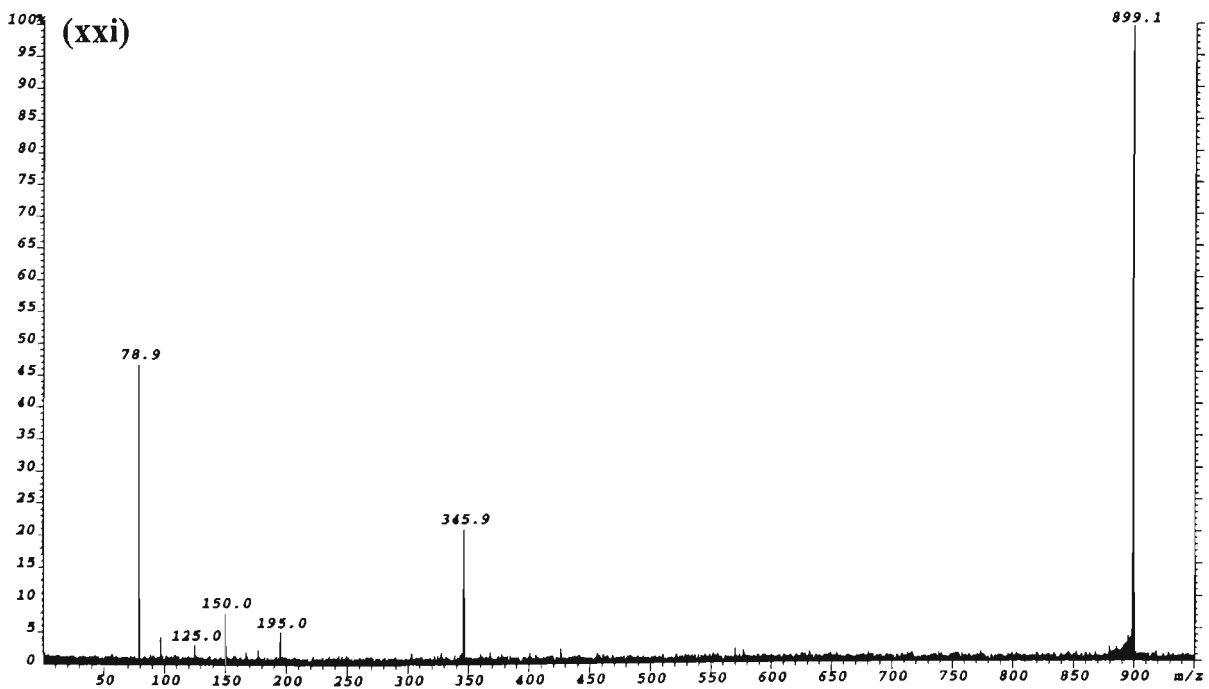
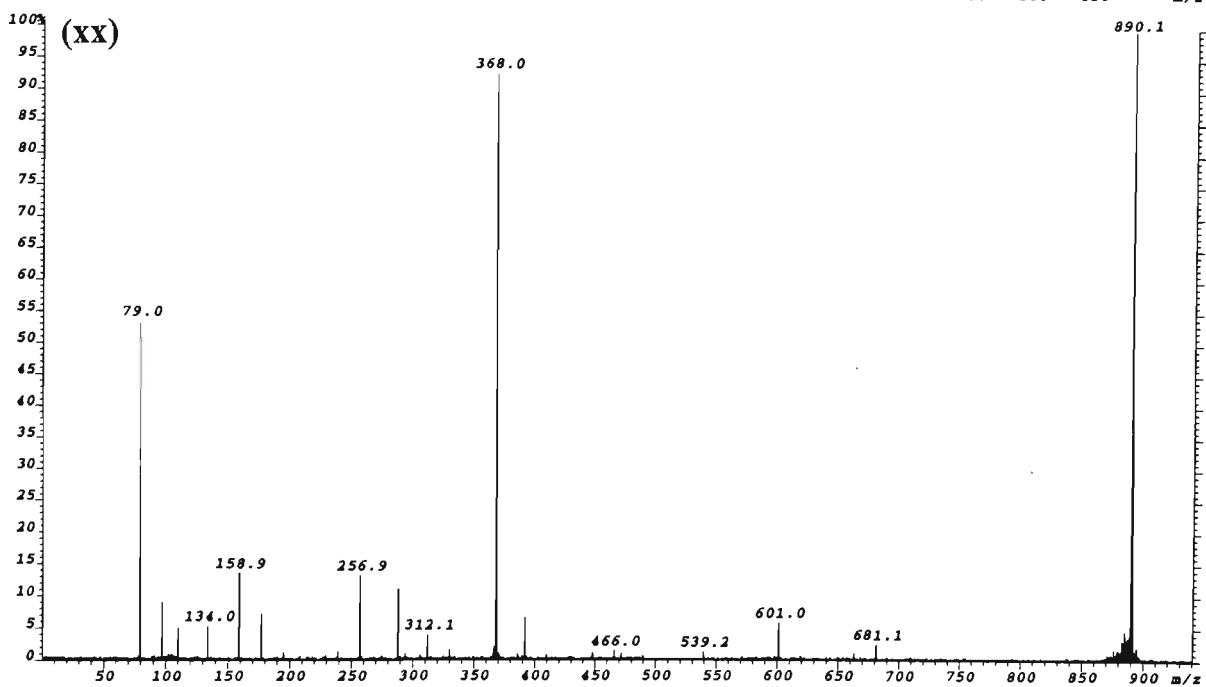
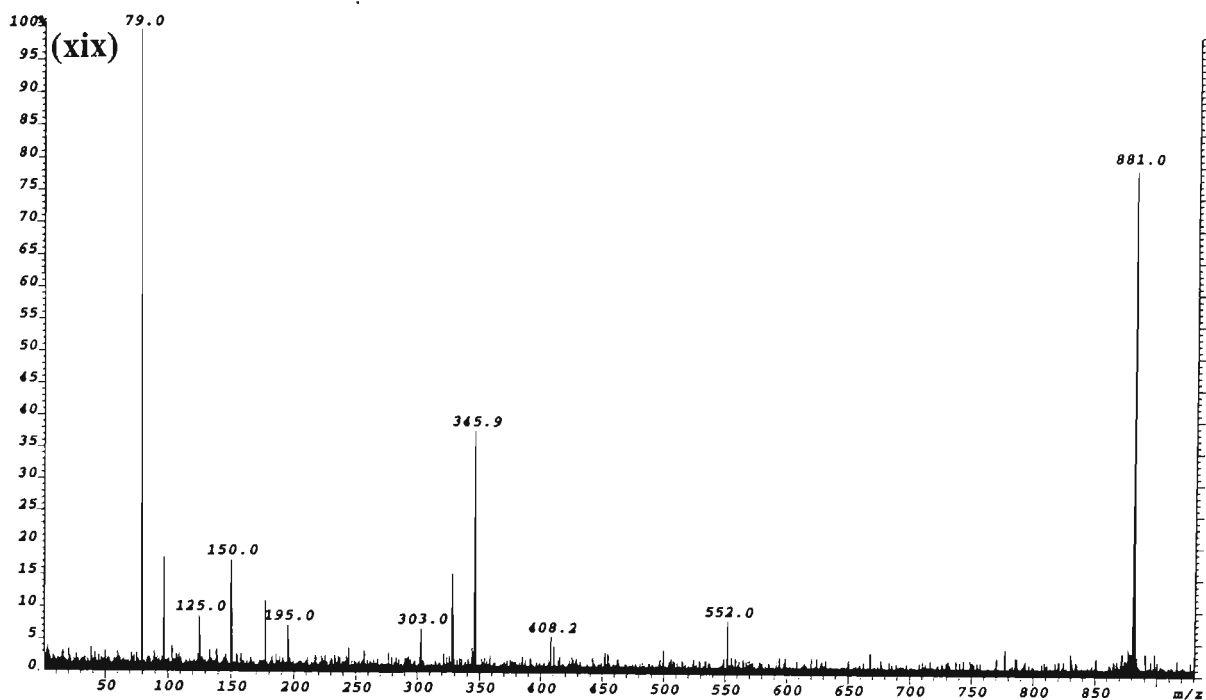


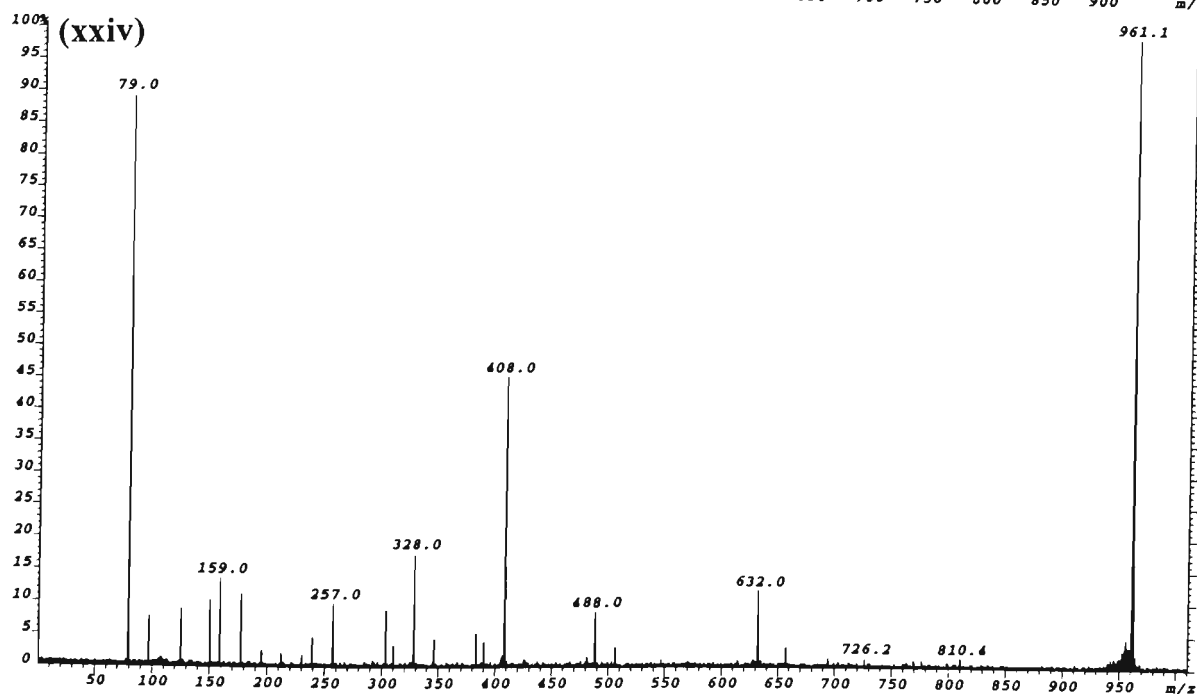
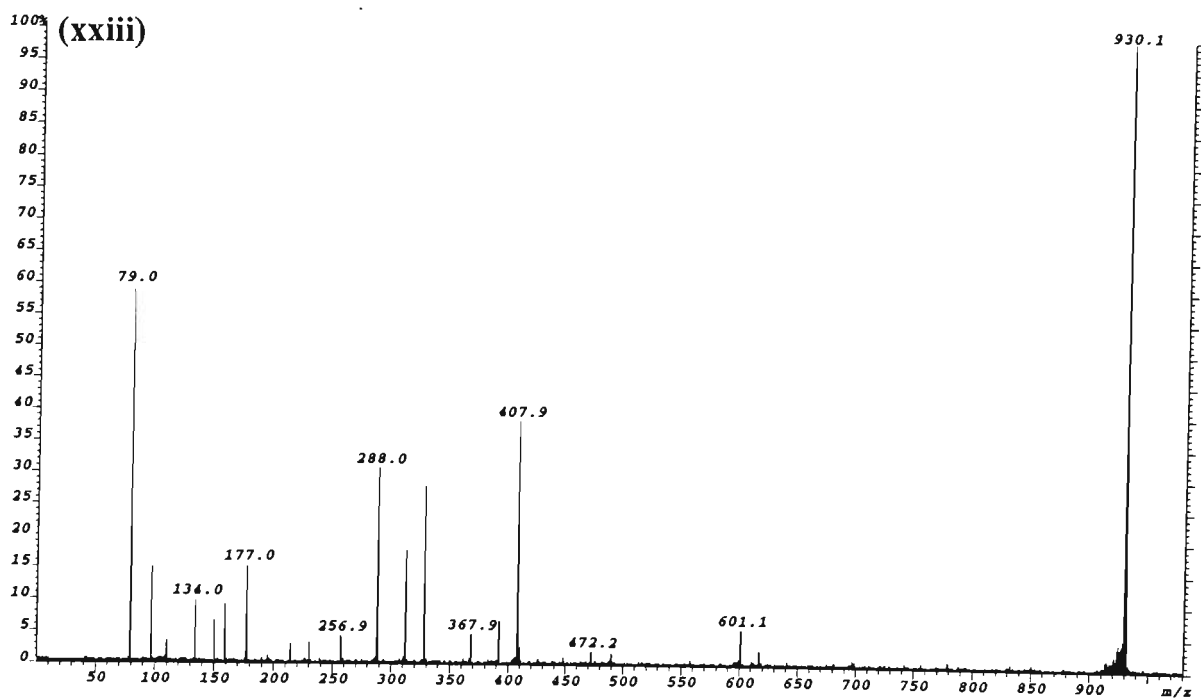
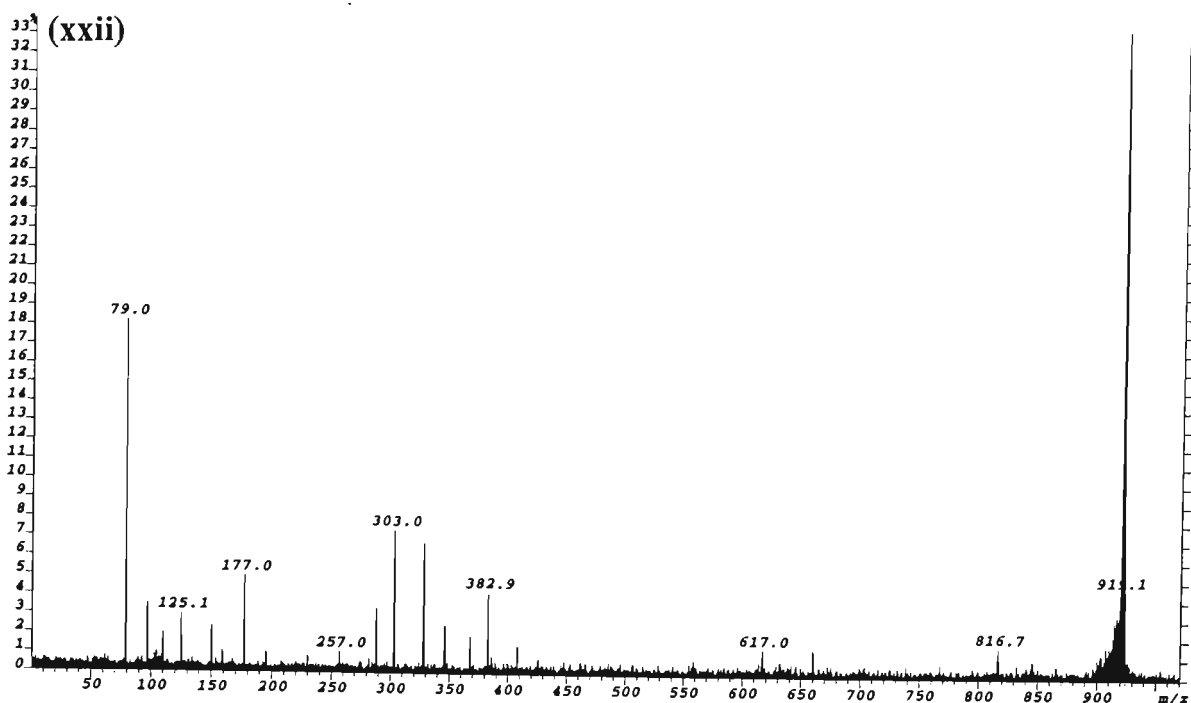


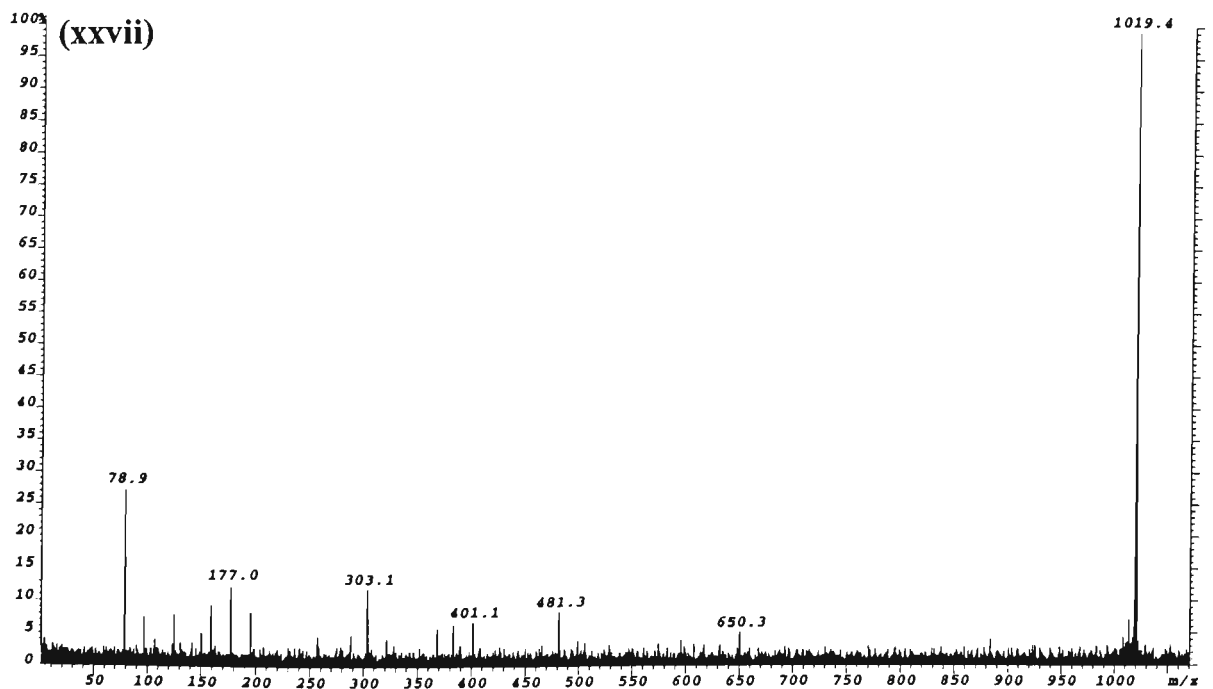
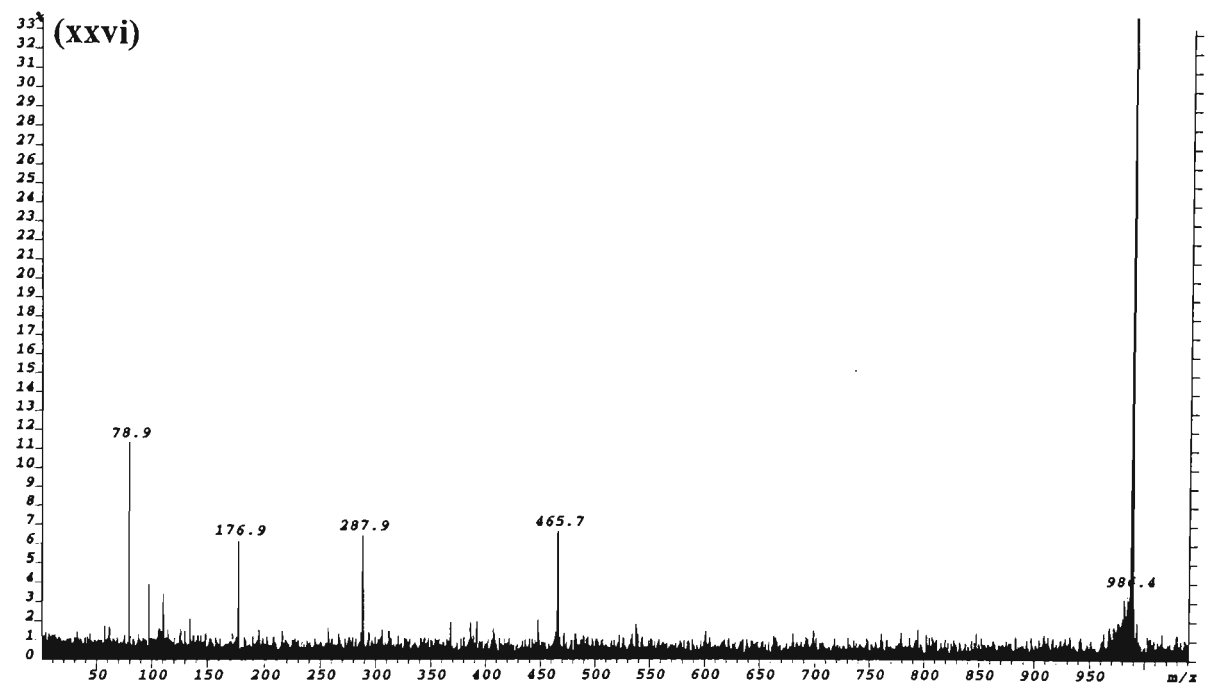
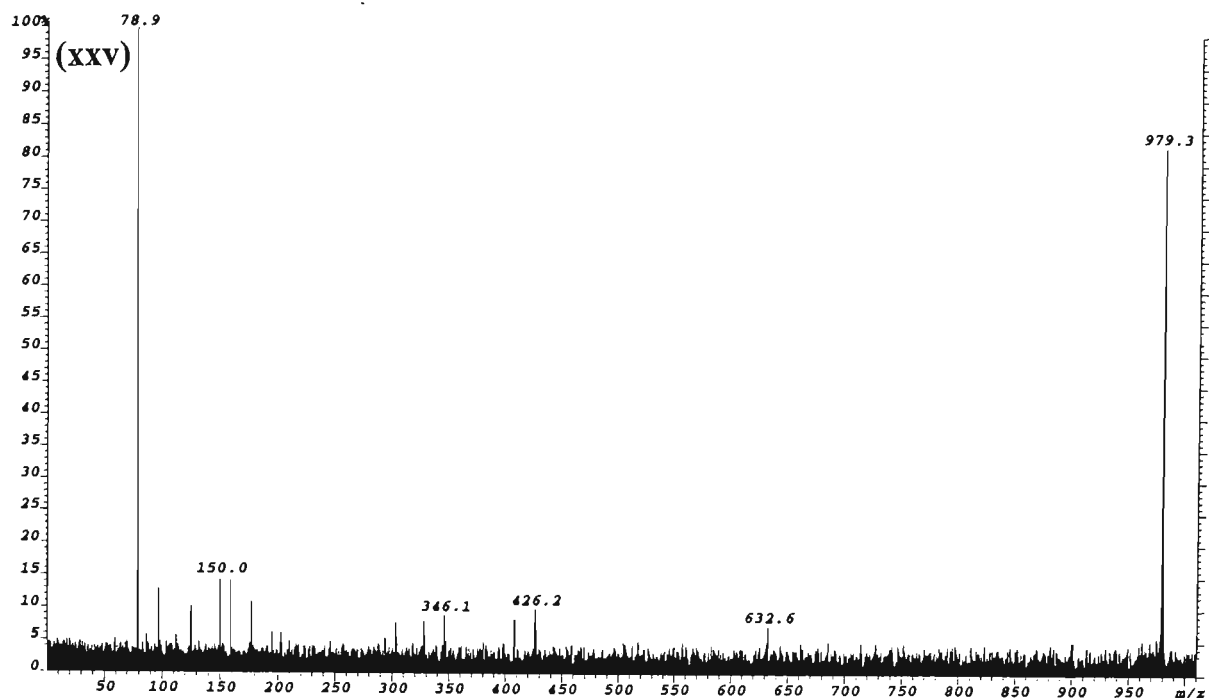


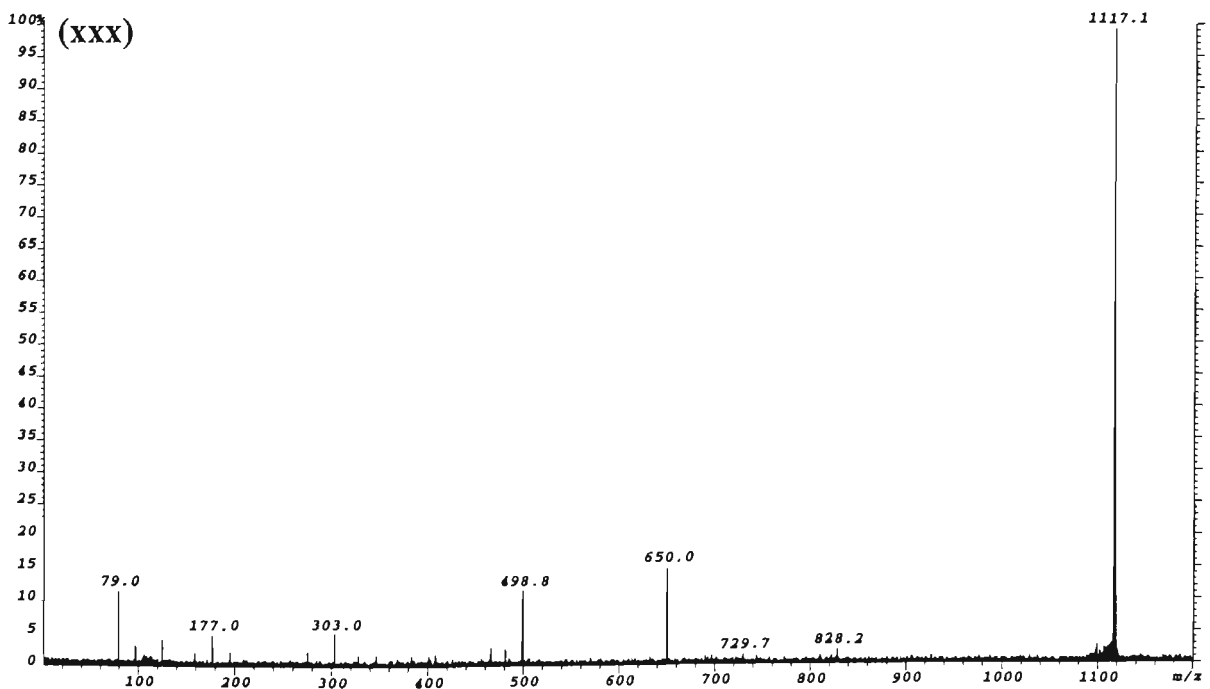
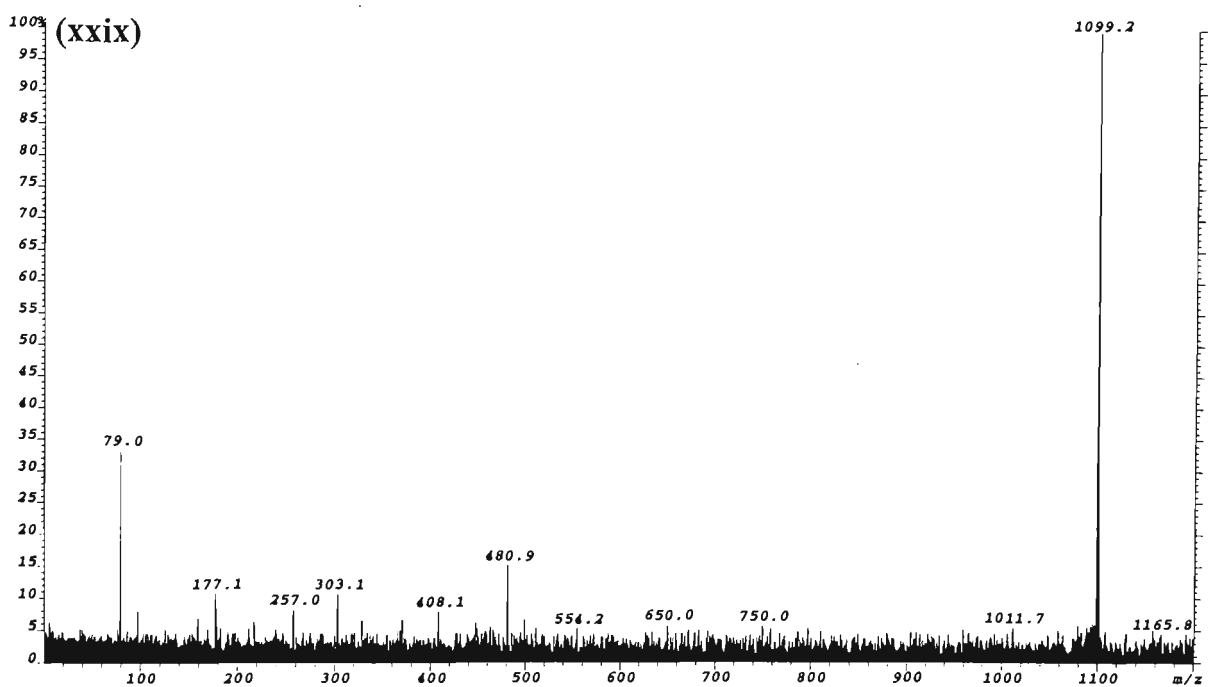
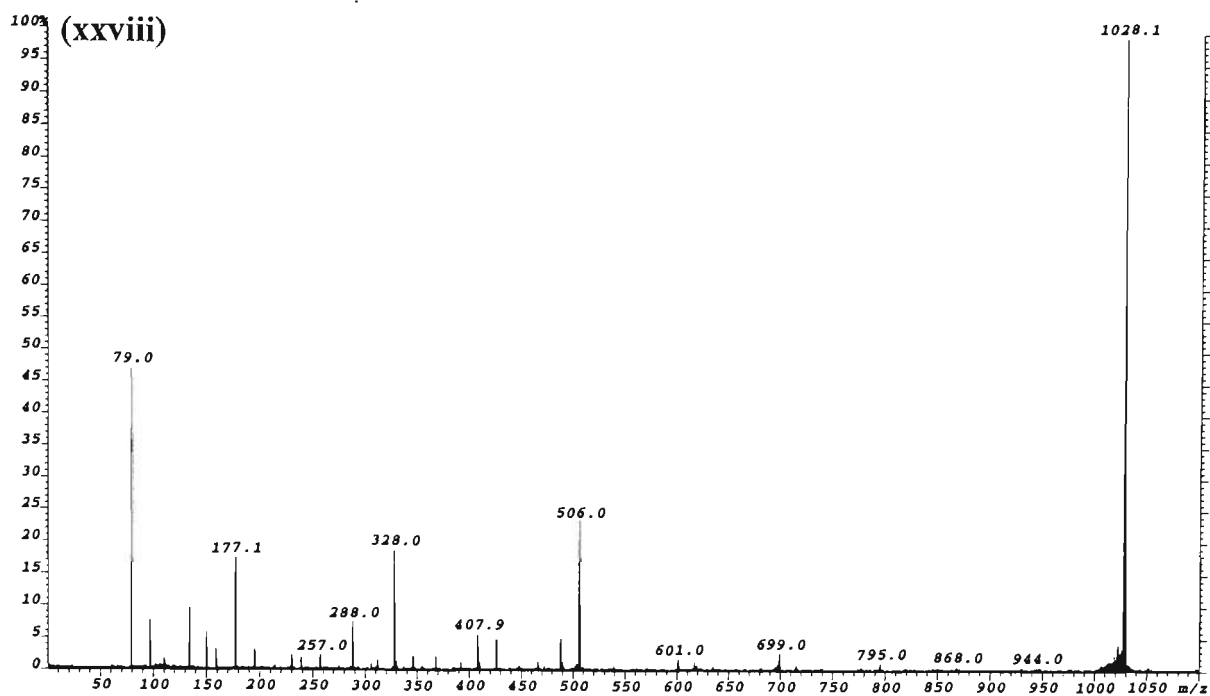


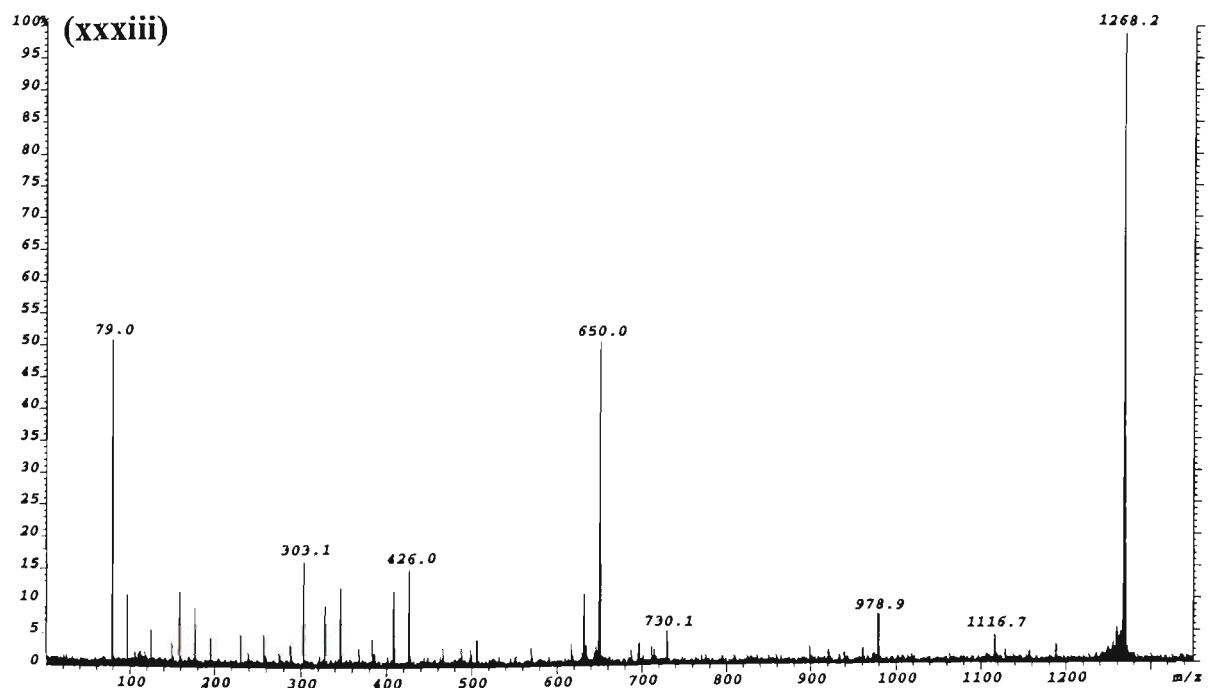
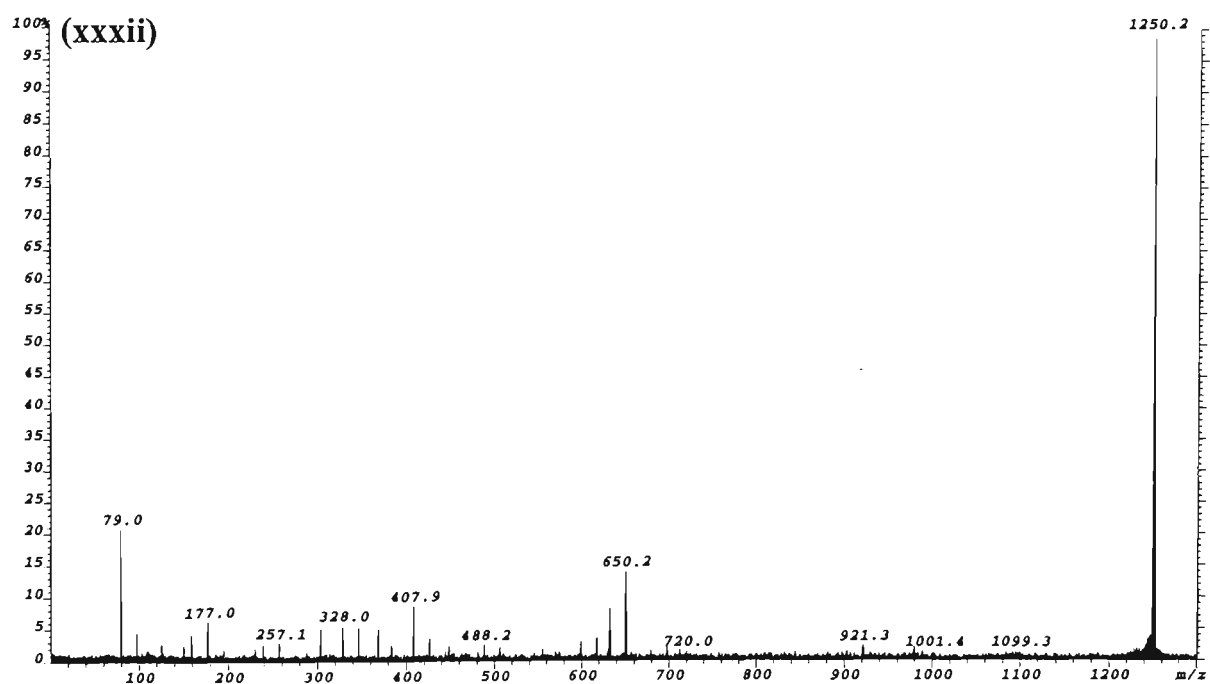
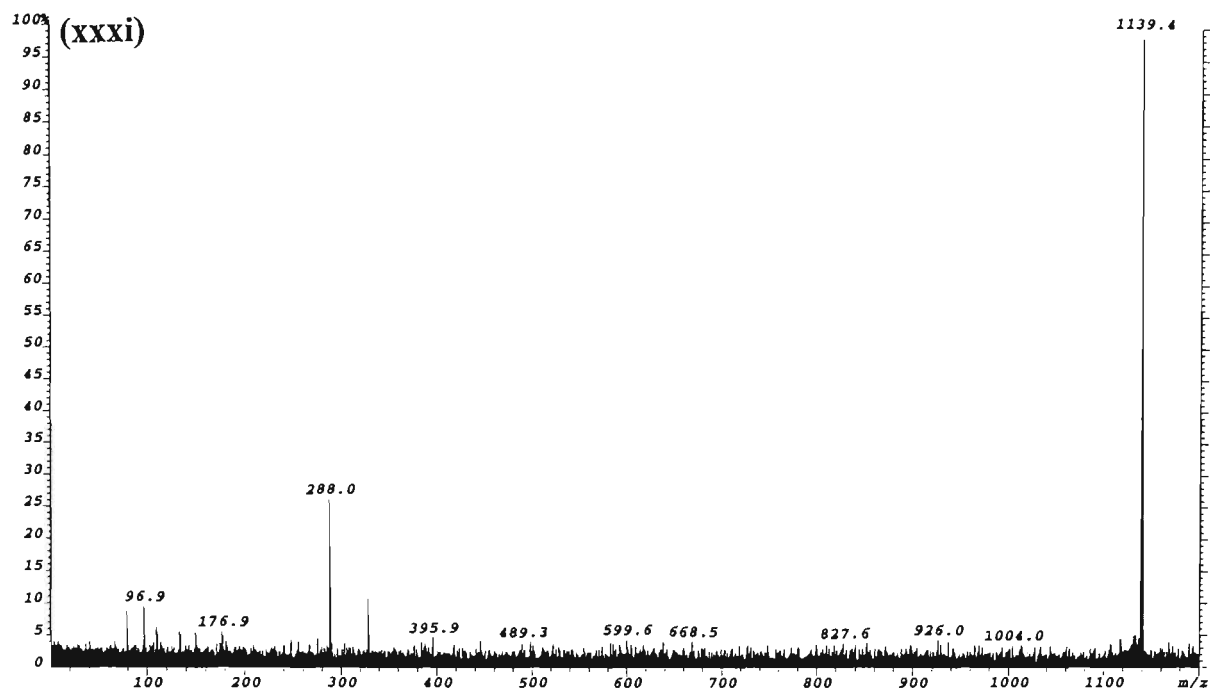


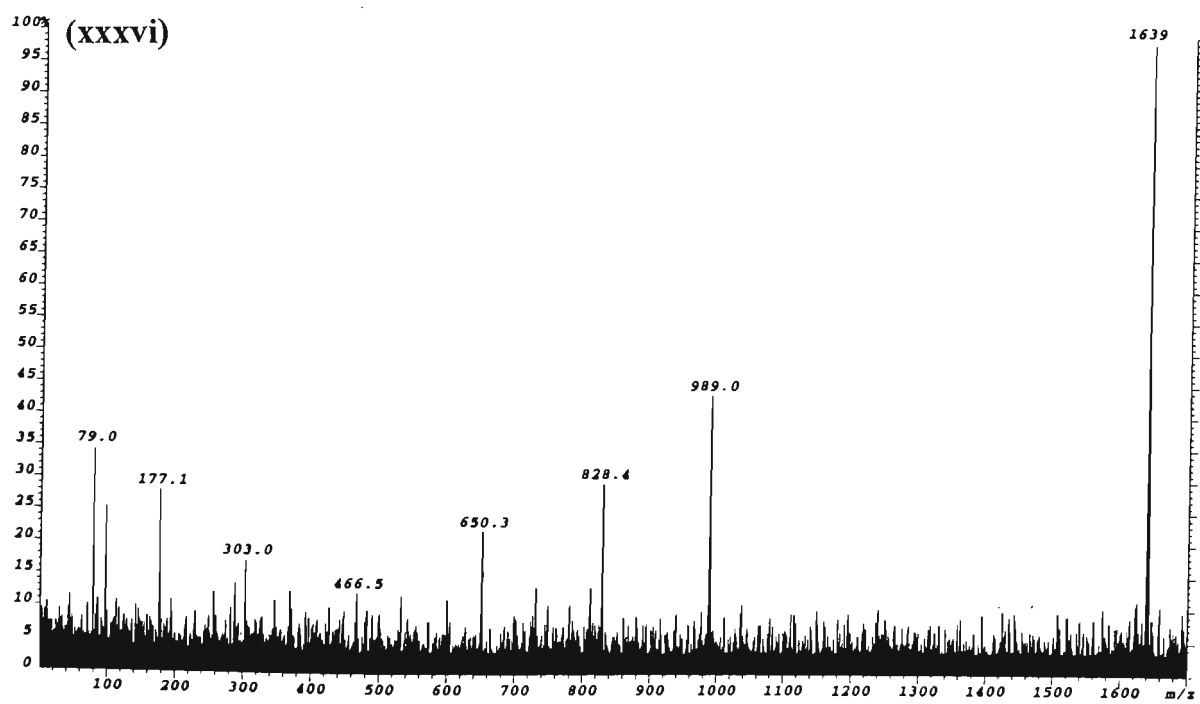
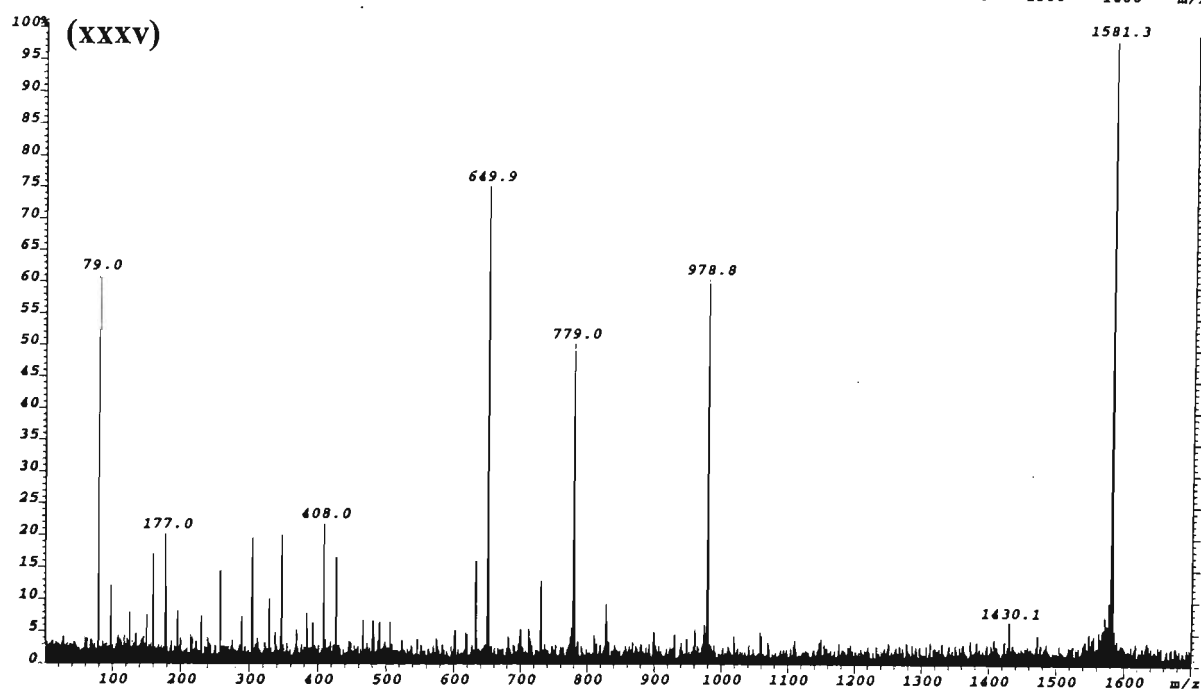
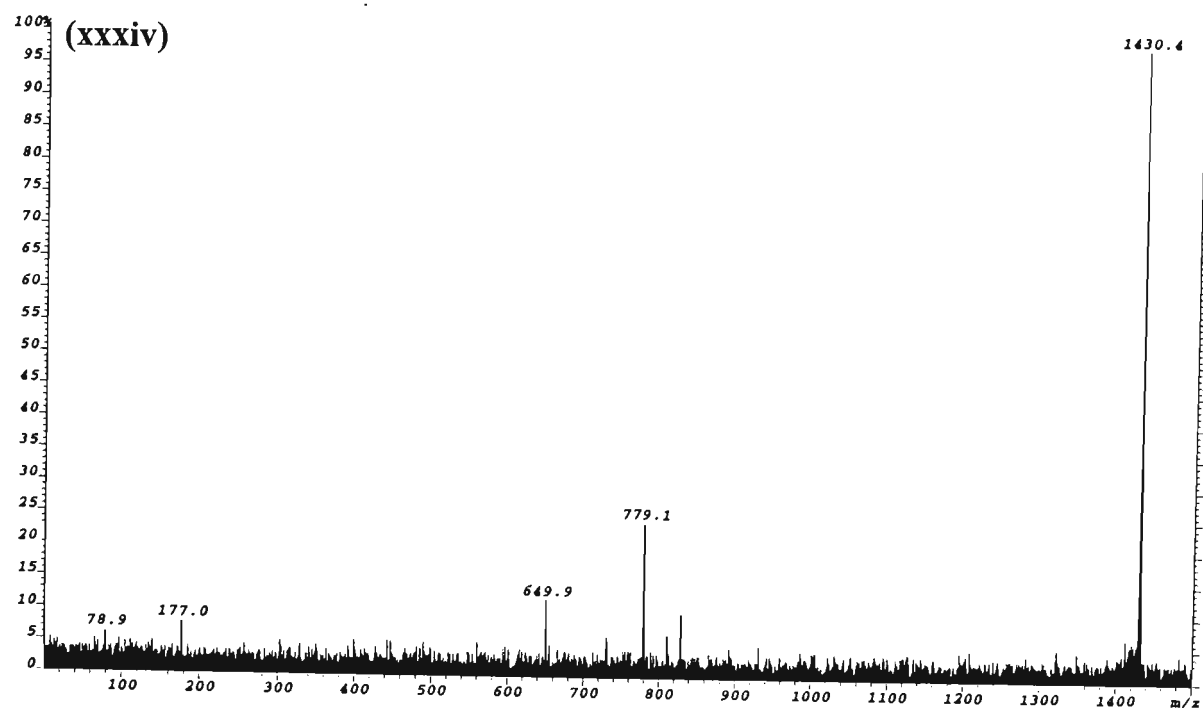












Appendix 4.3

ESI-MS/MS spectra of the ‘source-generated’ polynucleotide product ions observed in the ESI-MS/MS spectrum of the [M-H]⁻ ion of 5'-d(CGTACG)-3'.

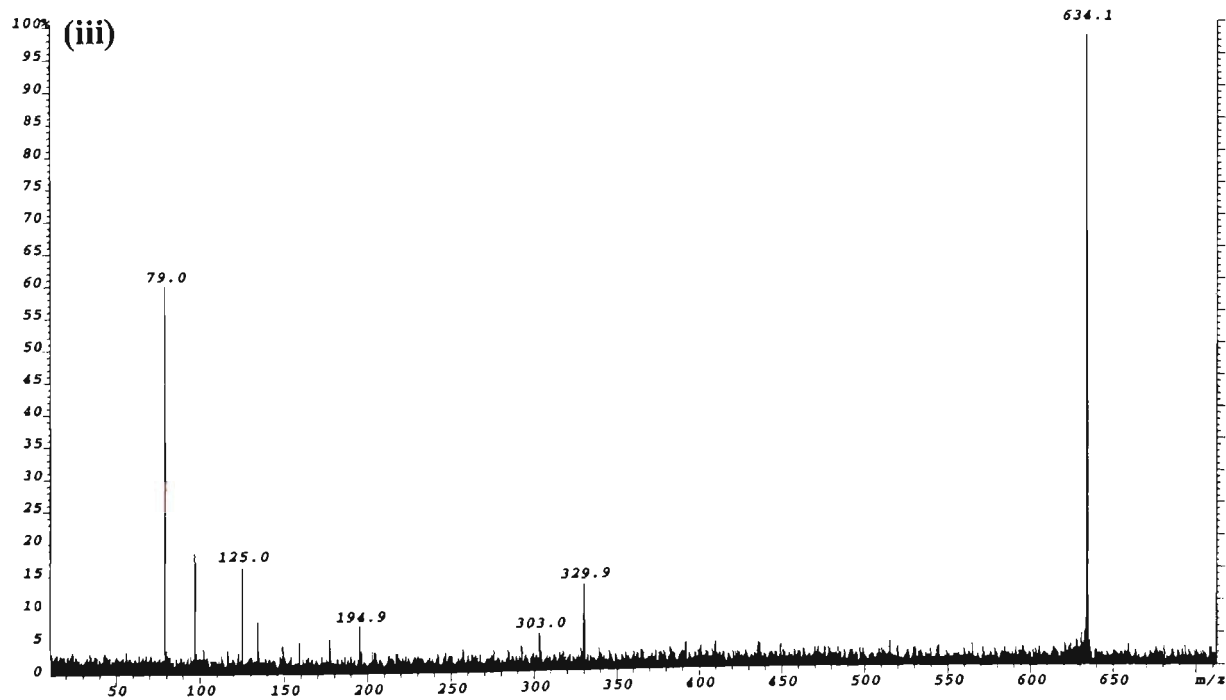
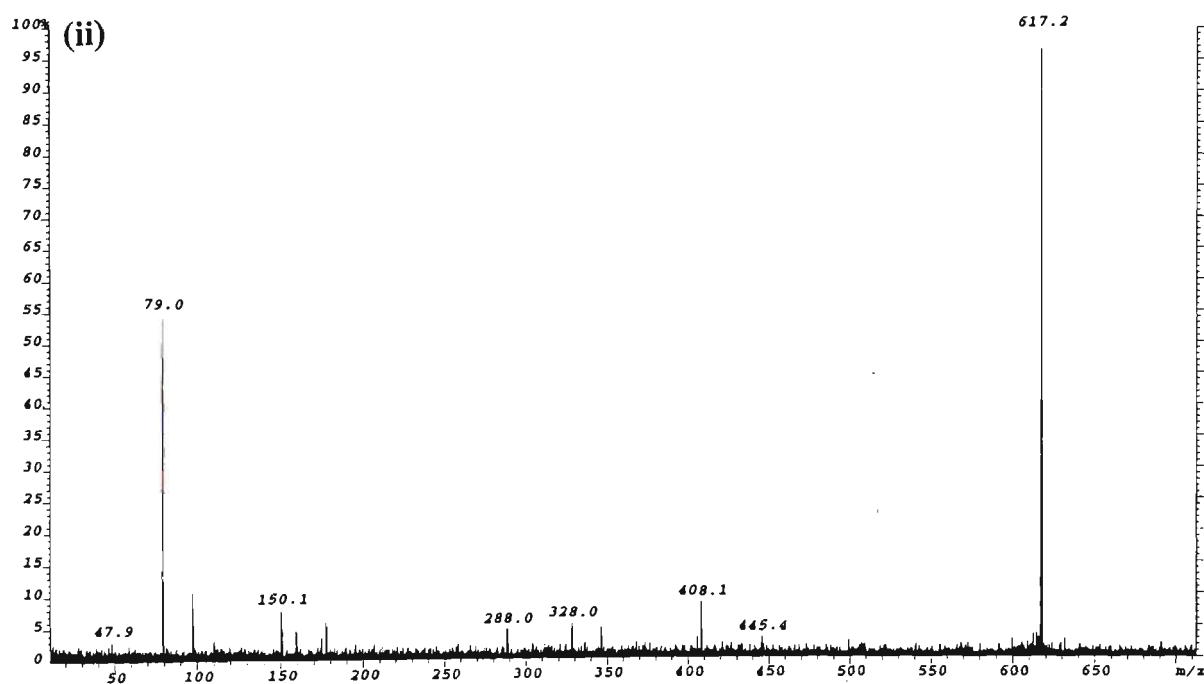
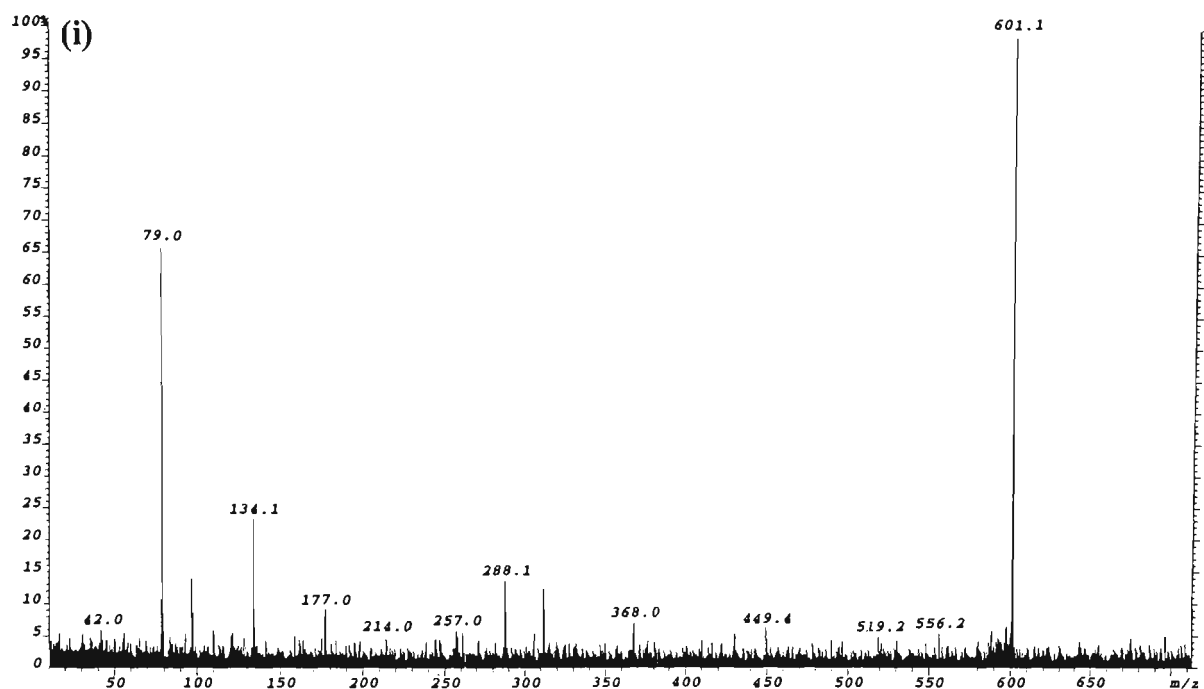
Figure No.	Precursor Ion	
	<i>m/z</i>	Assignment
(i)	601.1	[(AC)-H] ⁻
(ii)	617.2	[x ₂ -2H] ⁻ and/or [c ₂ -2H] ⁻
(iii)	634.1	[(TA)+H ₂ O-H] ⁻
(iv)	635.1	w ₂ ⁻ and/or d ₂ ⁻
(v)	650.0	[(GT)+H ₂ O-H] ⁻
(vi)	681.0	[(AC)+s-H ₂ O-H] ⁻ and/or [(AC)+p-H] ⁻
(vii)	696.0	[(TA)+s-H ₂ O-H] ⁻ and/or [(TA)+p-H] ⁻
(viii)	714.0	[(TA)+s-H] ⁻ and/or [(TA)+p+H ₂ O-H] ⁻
(ix)	715.1	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -AH-2H] ⁻
(x)	730.1	[(GT)+s-H] ⁻ and/or [(GT)+p+H ₂ O-H] ⁻
(xi)	776.1	[(TA)+s+p-H ₂ O-H] ⁻
(xii)	794.1	[(TA)+s+p-H] ⁻
(xiii)	810.0	[c ₃ -CH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
(xiv)	841.2	[a ₃ -2H] ⁻
(xv)	859.2	b ₃ ⁻
(xvi)	868.2	y ₃ ⁻
(xvii)	905.1	[(TAC)-H] ⁻
(xviii)	921.1	[c ₃ -2H] ⁻
(xix)	930.1	[x ₃ -2H] ⁻
(xx)	948.1	w ₃ ⁻
(xxi)	985.0	[(TAC)+s-H ₂ O-H] ⁻ and/or [(TAC)+p-H] ⁻
(xxii)	1019.1	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -AH-2H] ⁻ and/or [(GTAC)-AH-p-H] ⁻
(xxiii)	1043.2	[a ₄ -CH-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(GTA)+s-H] ⁻
(xxiv)	1083.1	[x ₄ -GH-2H] ⁻ and/or [c ₄ -GH-2H] ⁻ and/or [(TAC)+s+p-H] ⁻
(xxv)	1101.1	[w ₄ -GH] ⁻ and/or [d ₄ -GH] ⁻ and/or [(TAC)+s+p+H ₂ O-H] ⁻
(xxvi)	1123.1	[x ₄ -CH-2H] ⁻ and/or [c ₄ -CH-2H] ⁻ and/or [(GTA)+s+p-H] ⁻
(xxvii)	1154.3	[a ₄ -2H] ⁻ and/or [z ₄ -2H] ⁻ and/or [(GTAC)-p-H] ⁻
(xxviii)	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GTAC)-p+H ₂ O-H] ⁻
(xxix)	1234.1	[x ₄ -2H] ⁻ and/or [c ₄ -2H] ⁻ and/or [(GTAC)-H] ⁻
(xxx)	1252.2	w ₄ ⁻ and/or d ₄ ⁻ and/or [(GTAC)+H ₂ O-H] ⁻
(xxxi)	1332.2	[a ₅ -B ₅ H-2H] ⁻ and/or [z ₅ -GH-2H] ⁻ and/or [(GTAC)+p+H ₂ O-H] ⁻
(xxxii)	1430.2	[w ₅ -GH] ⁻ and/or [d ₅ -GH] ⁻
(xxxiii)	1581.3	w ₅ ⁻ and/or d ₅ ⁻
(xxxiv)	1639.3	[M-GH-H] ⁻
(xxxv)	1680.5	[M-CH-H] ⁻

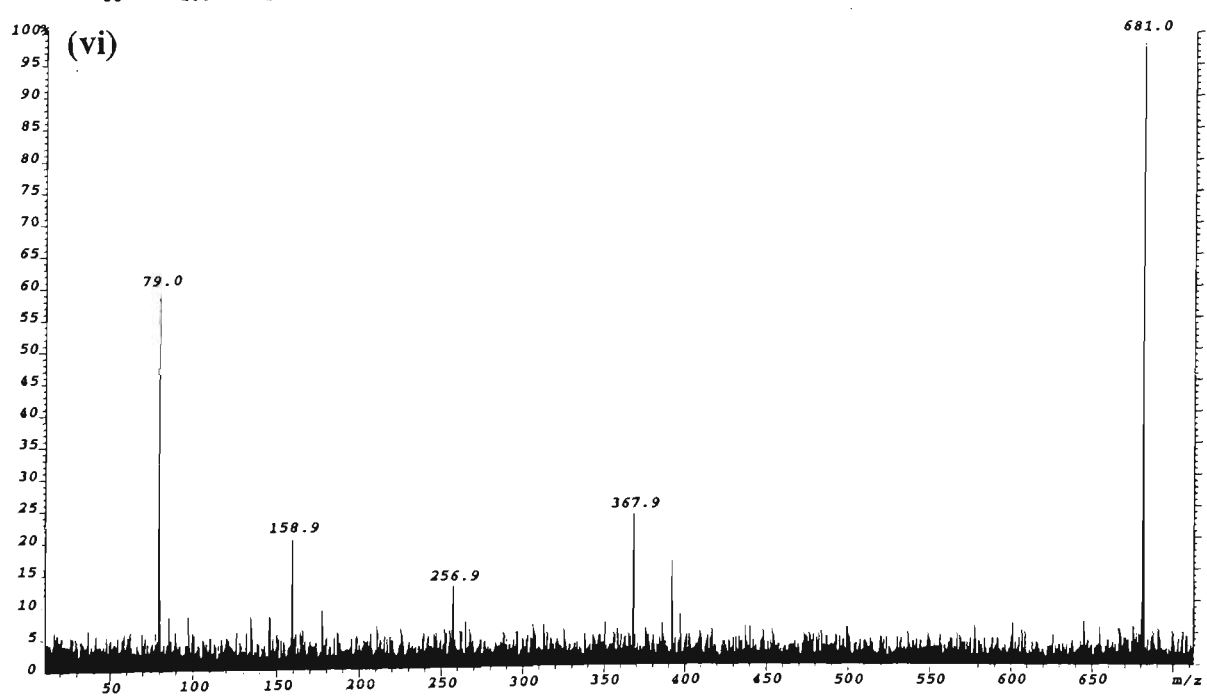
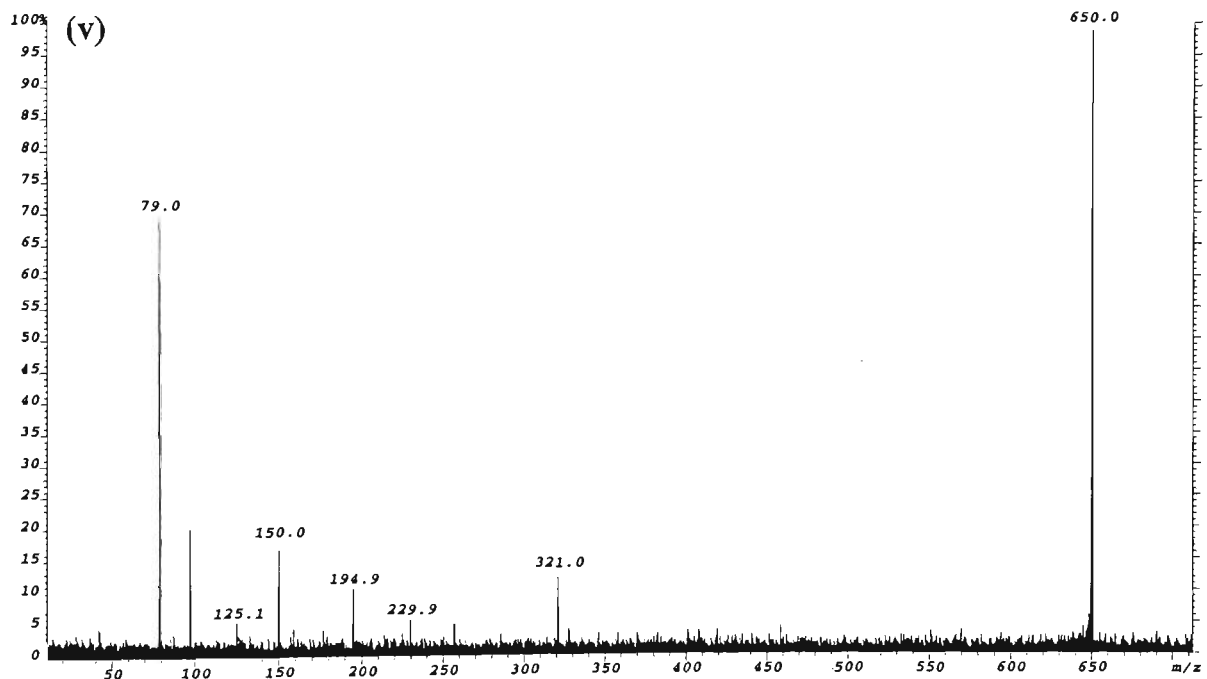
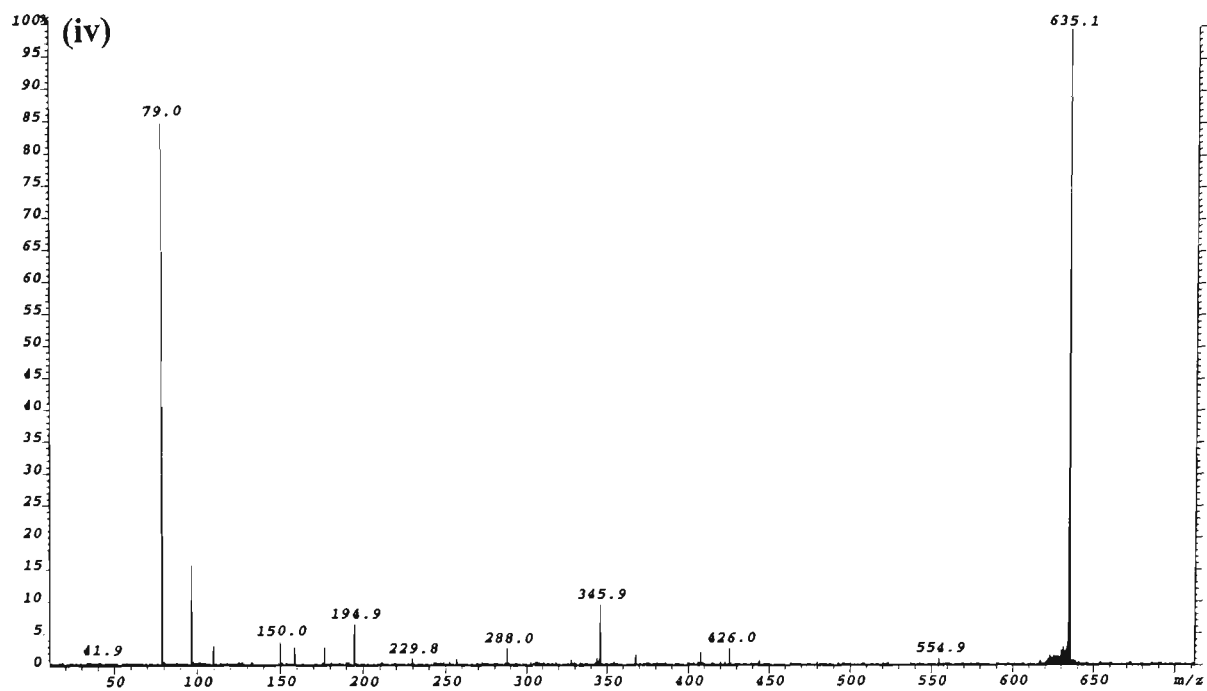
s = deoxyribose-H₂O (C₅H₆O₂), *M_r* = 98.0368 Da

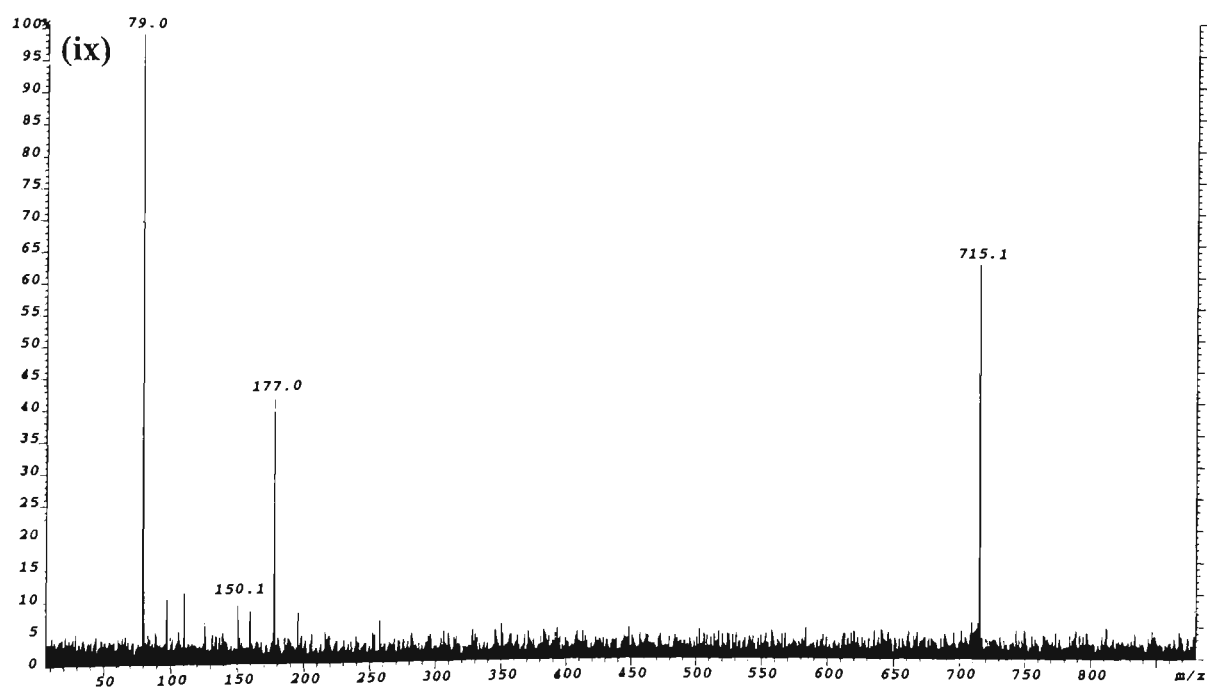
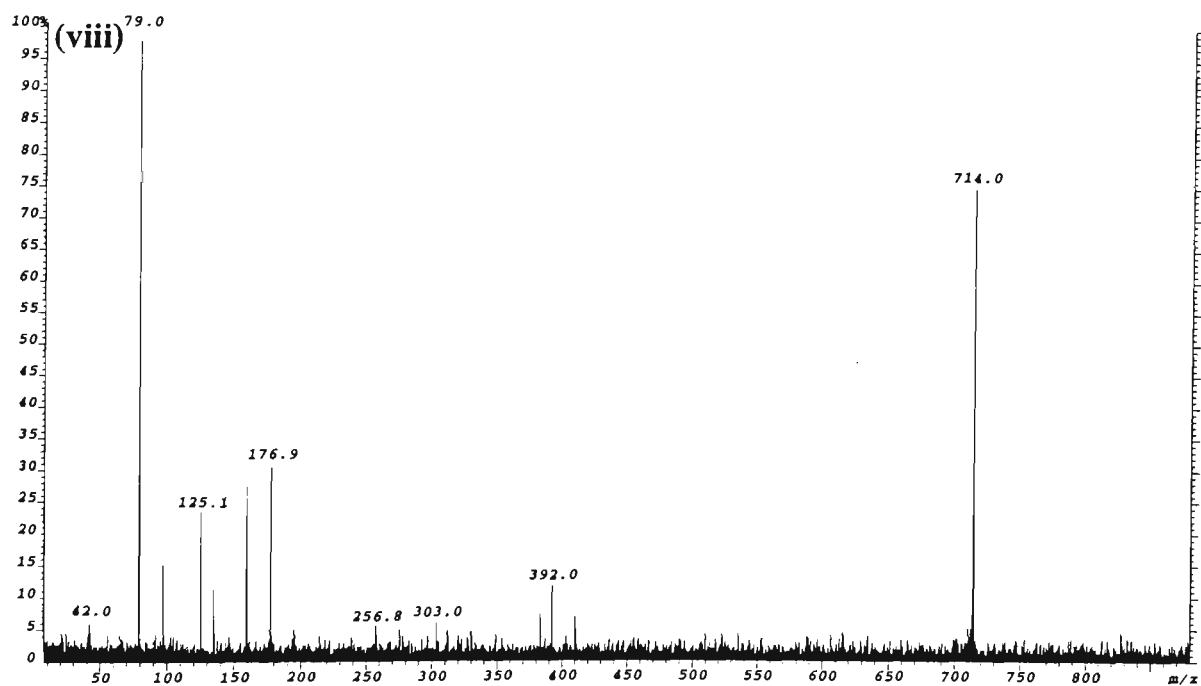
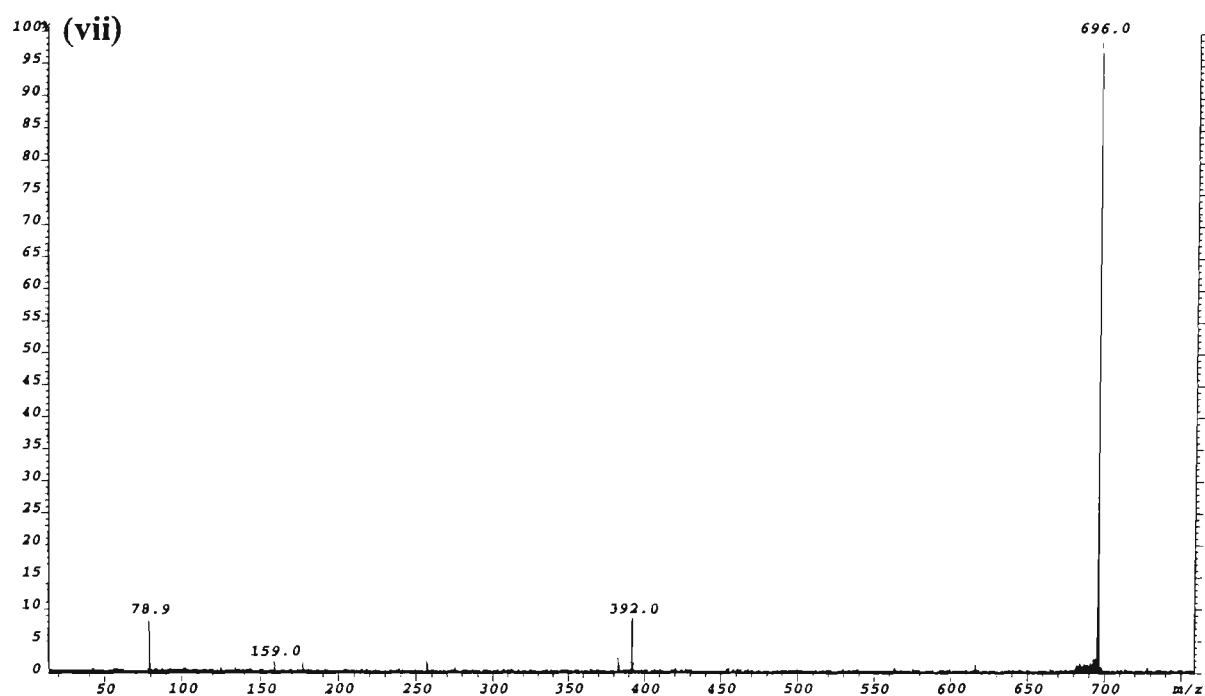
p = PO₃H, *M_r* = 79.9663 Da

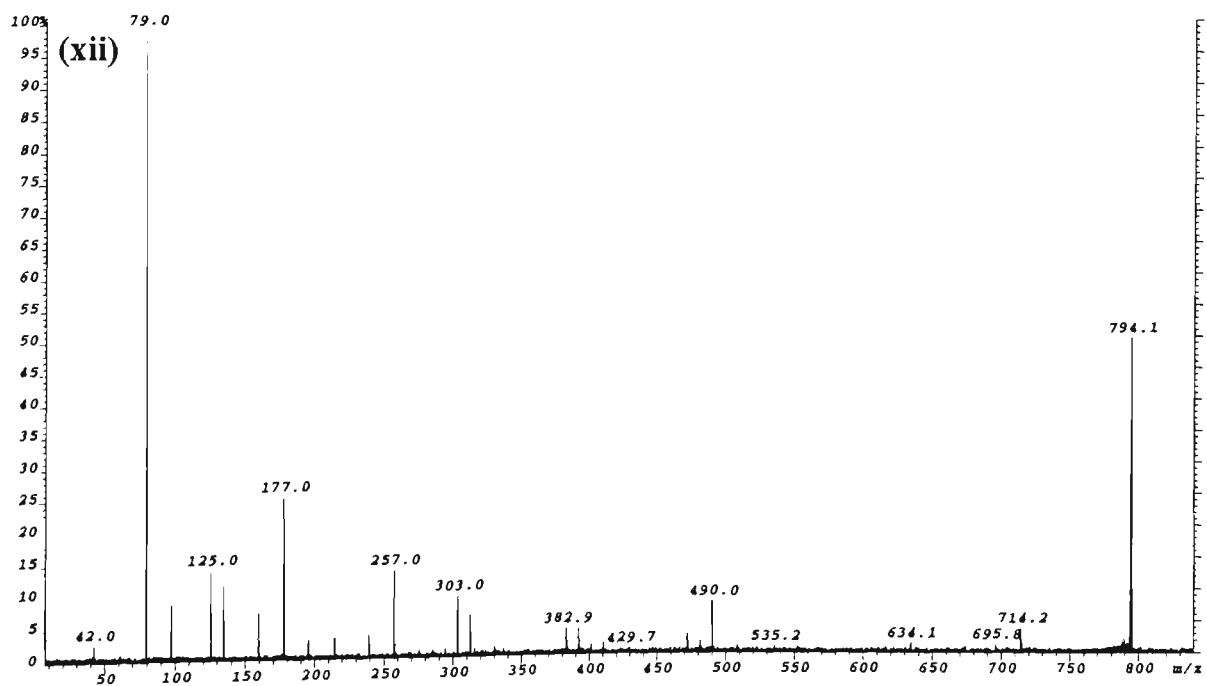
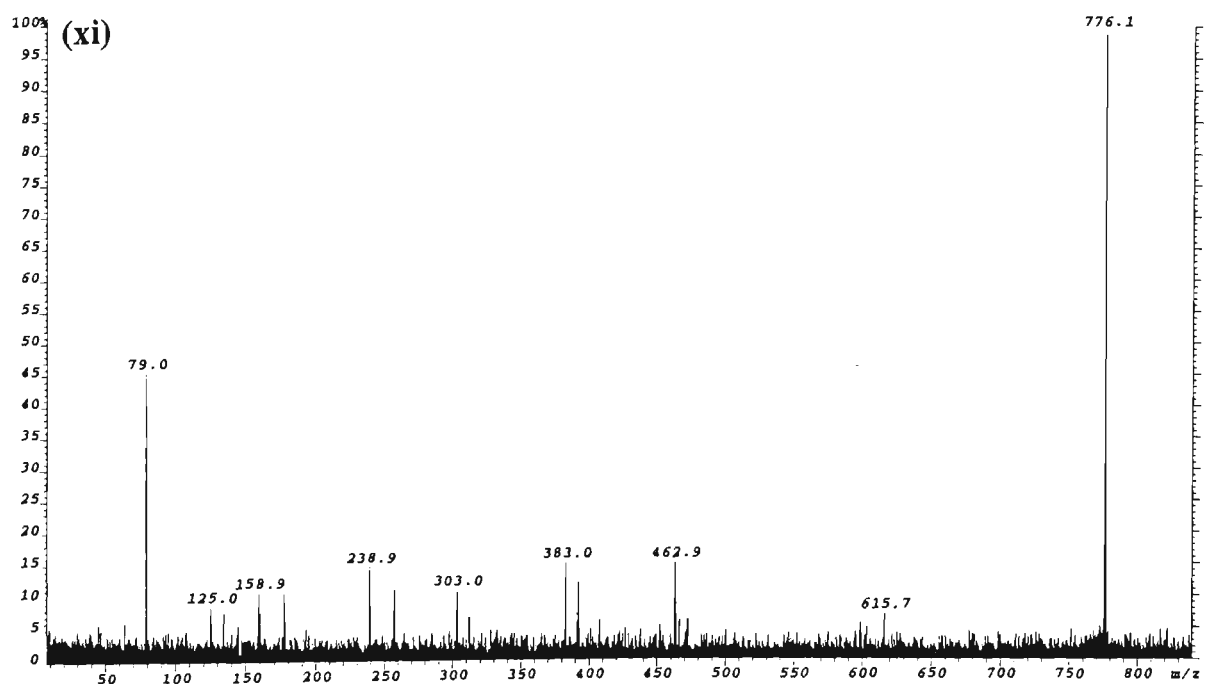
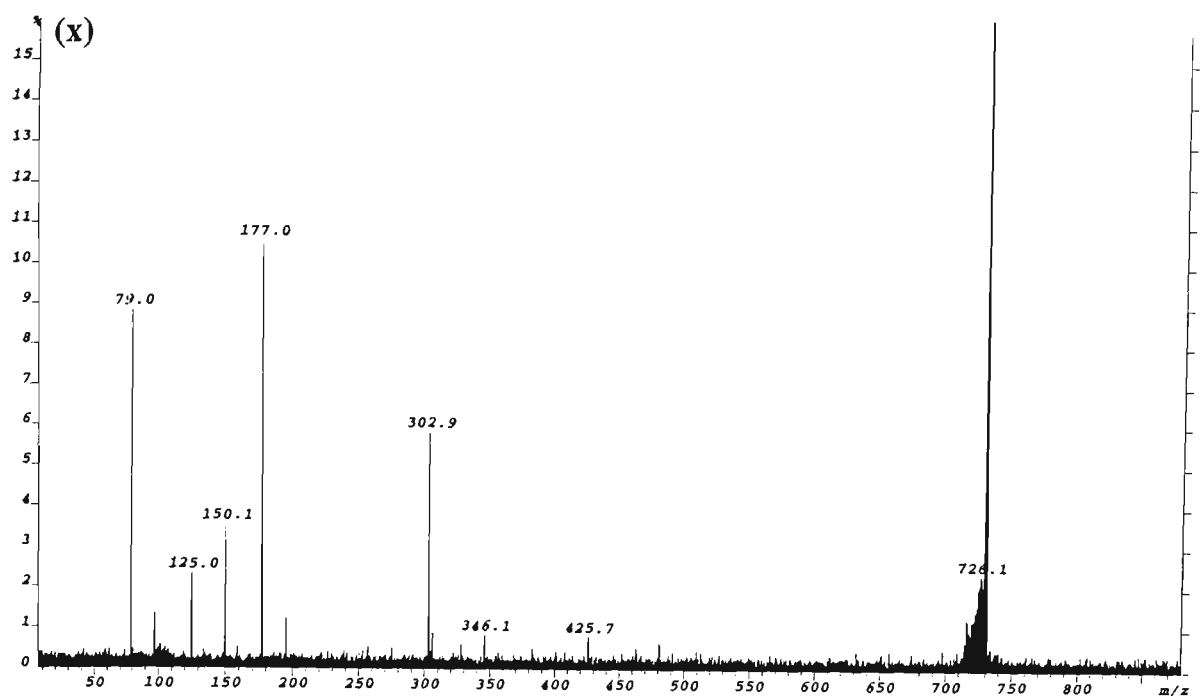
B_{nt} denotes a mononucleotide which may be either psB or sBp

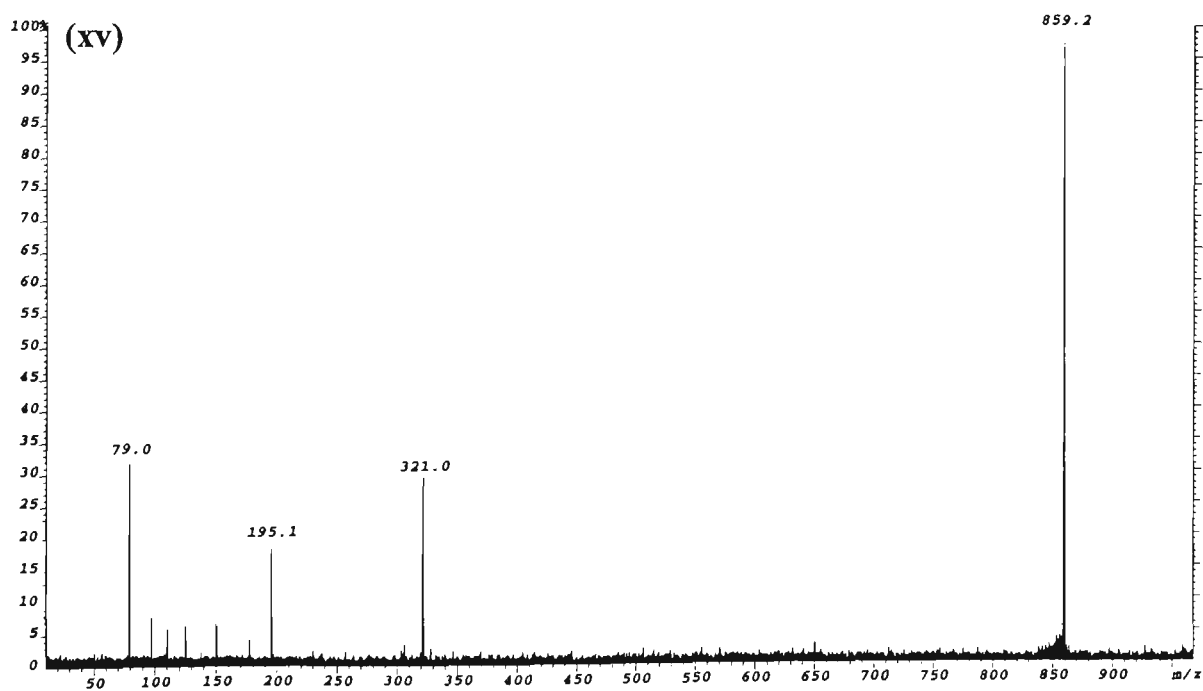
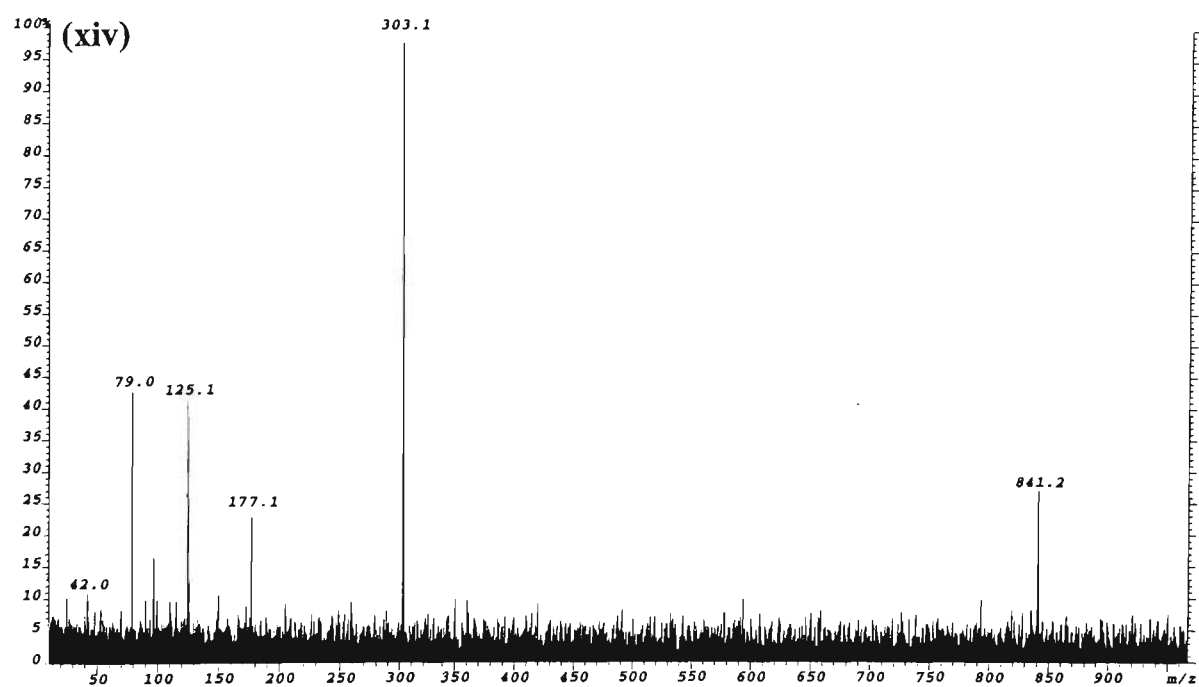
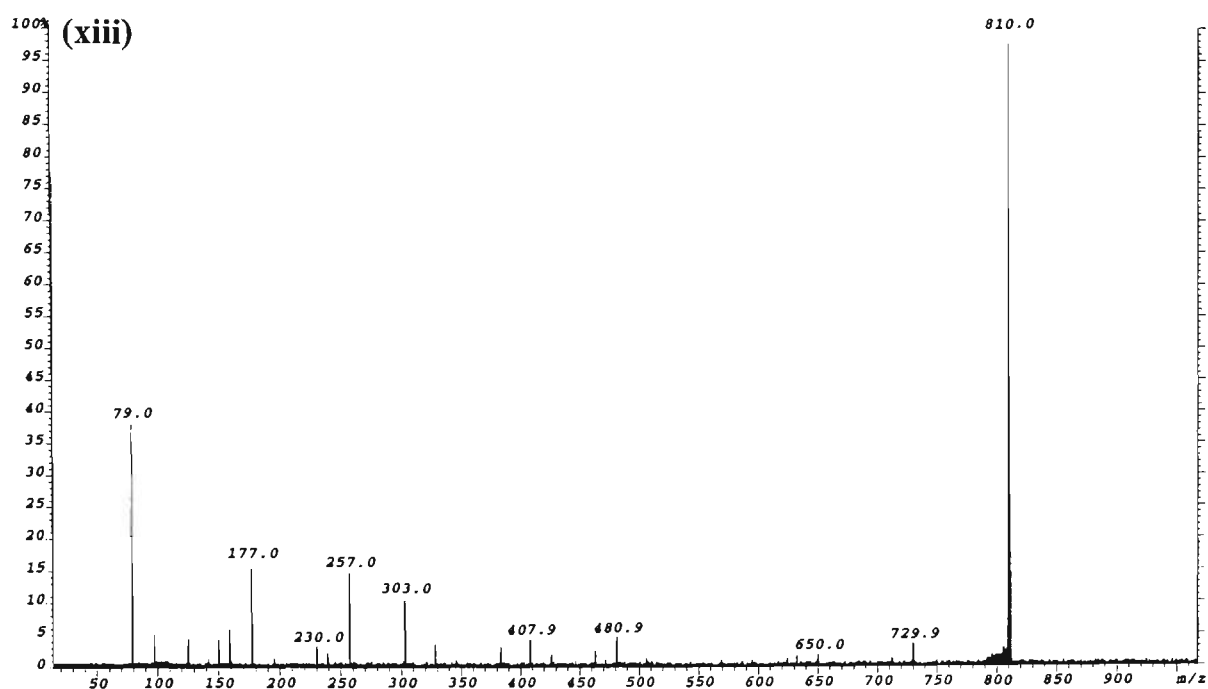
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

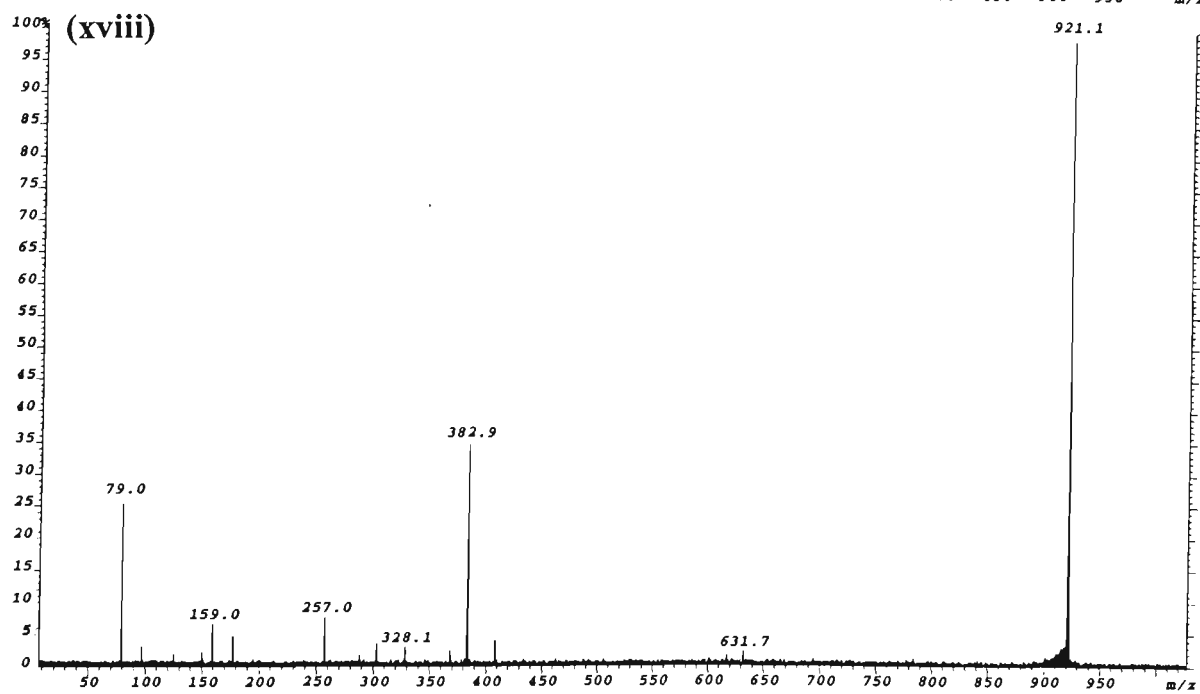
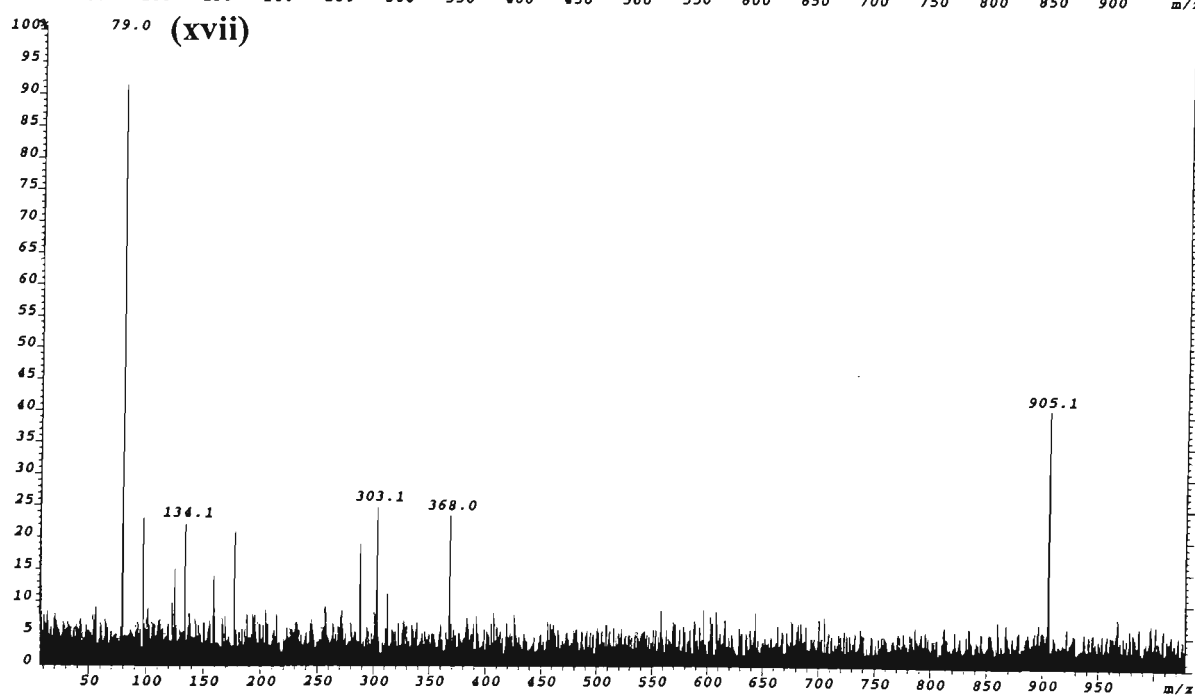


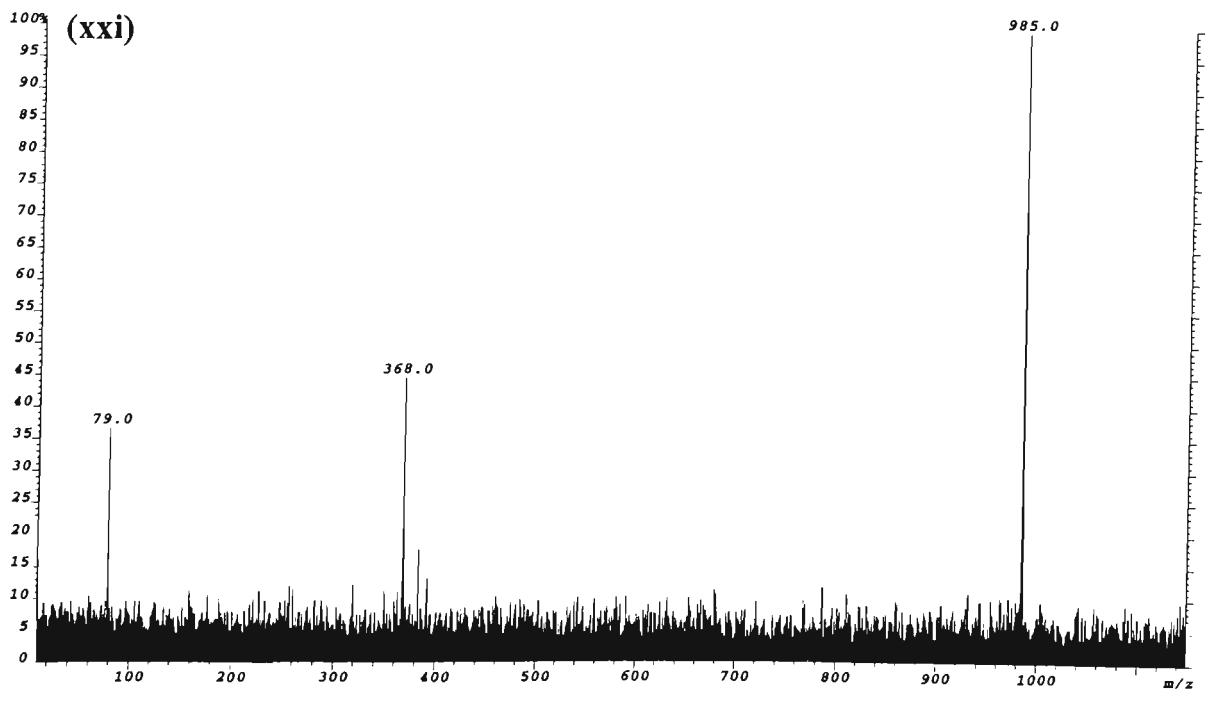
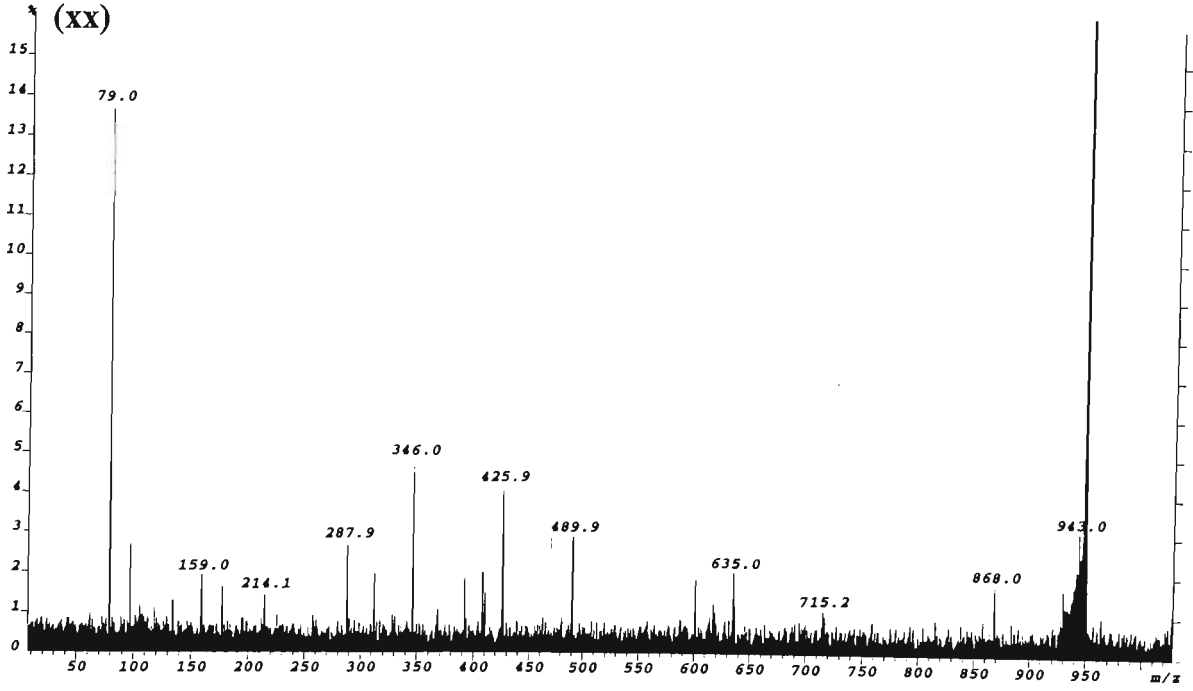
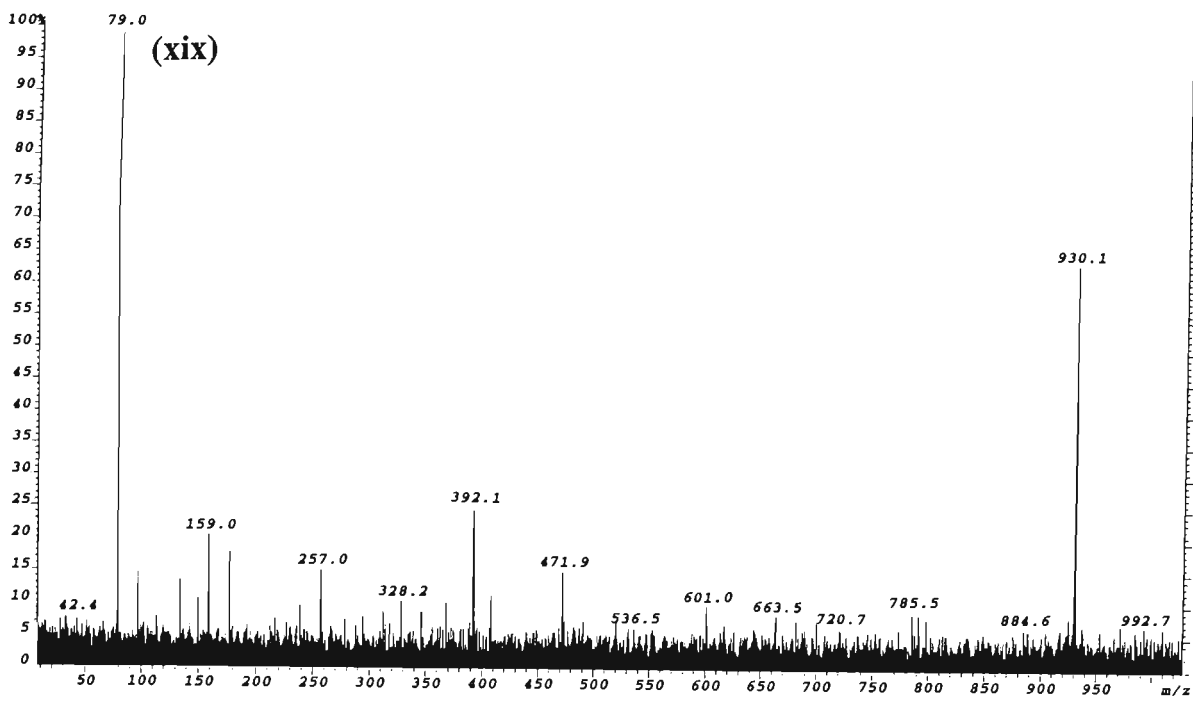


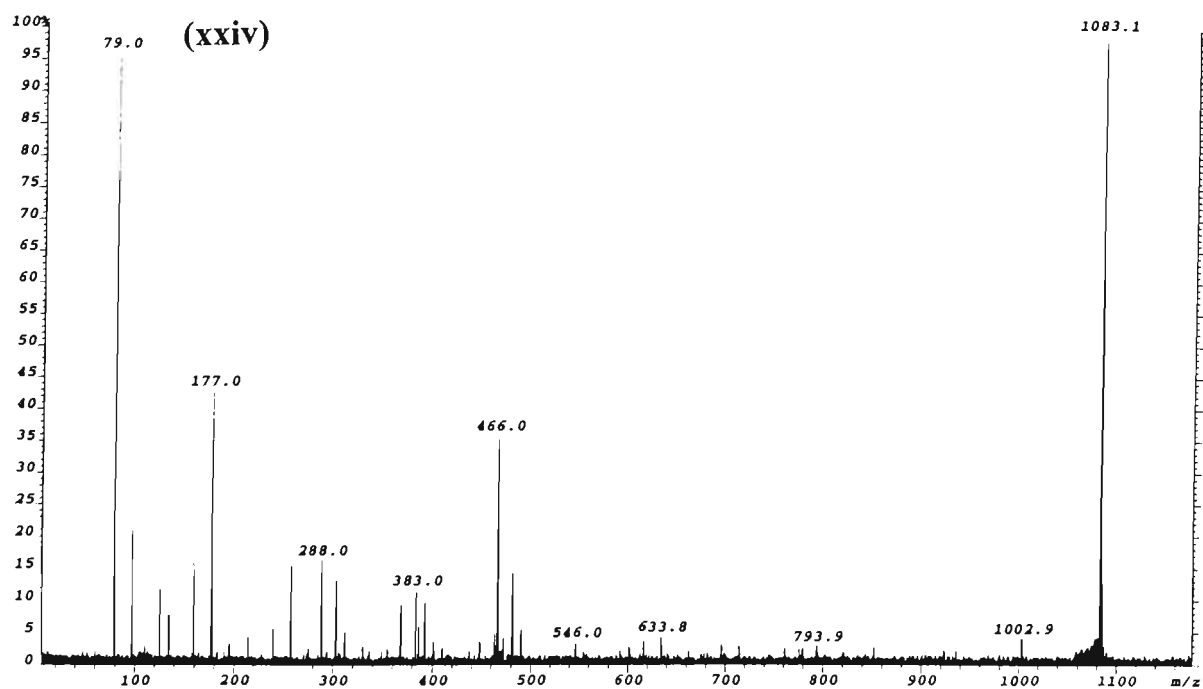
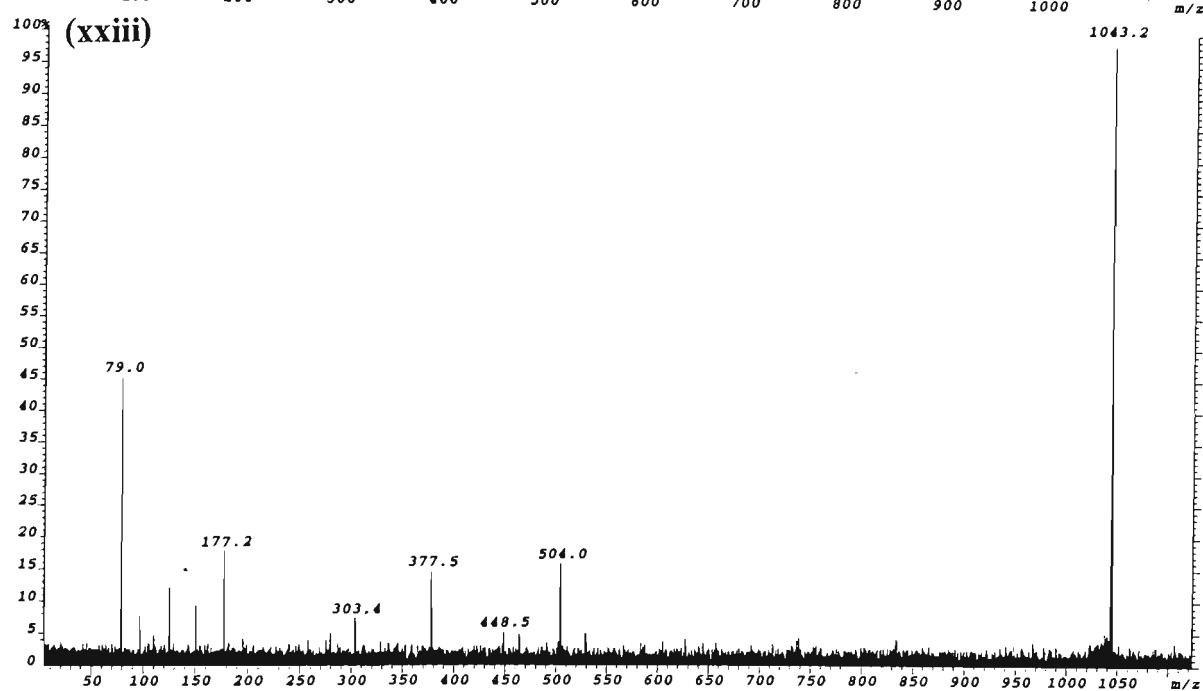
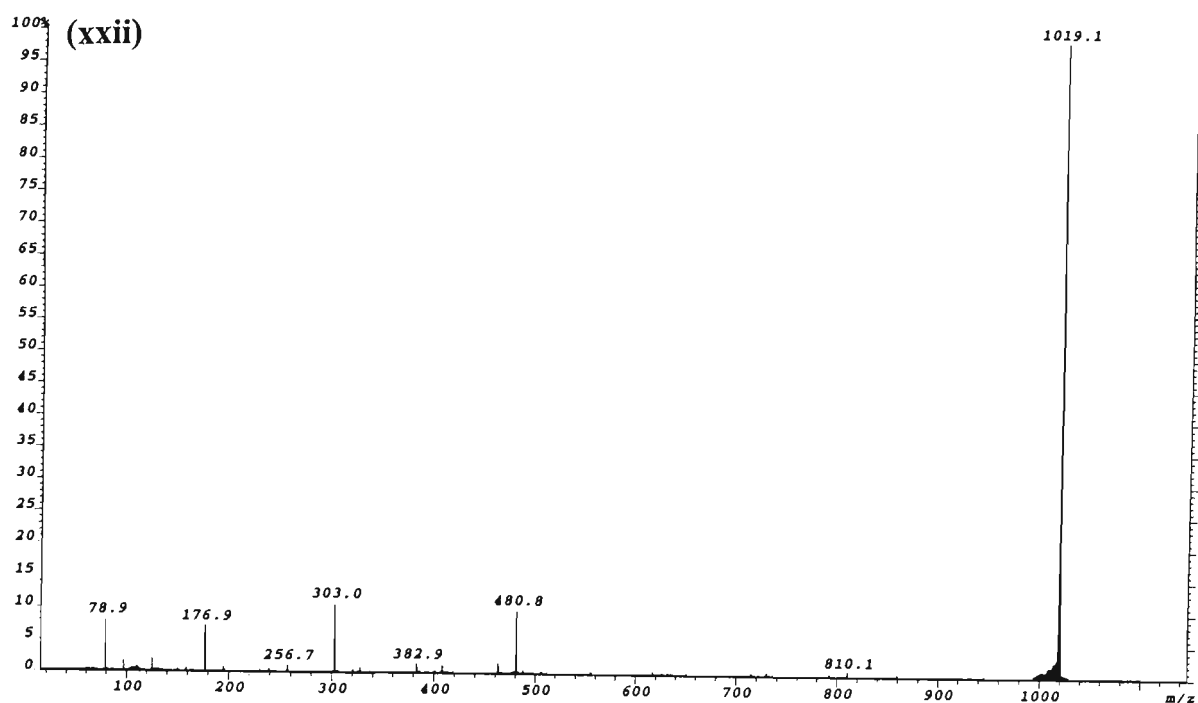


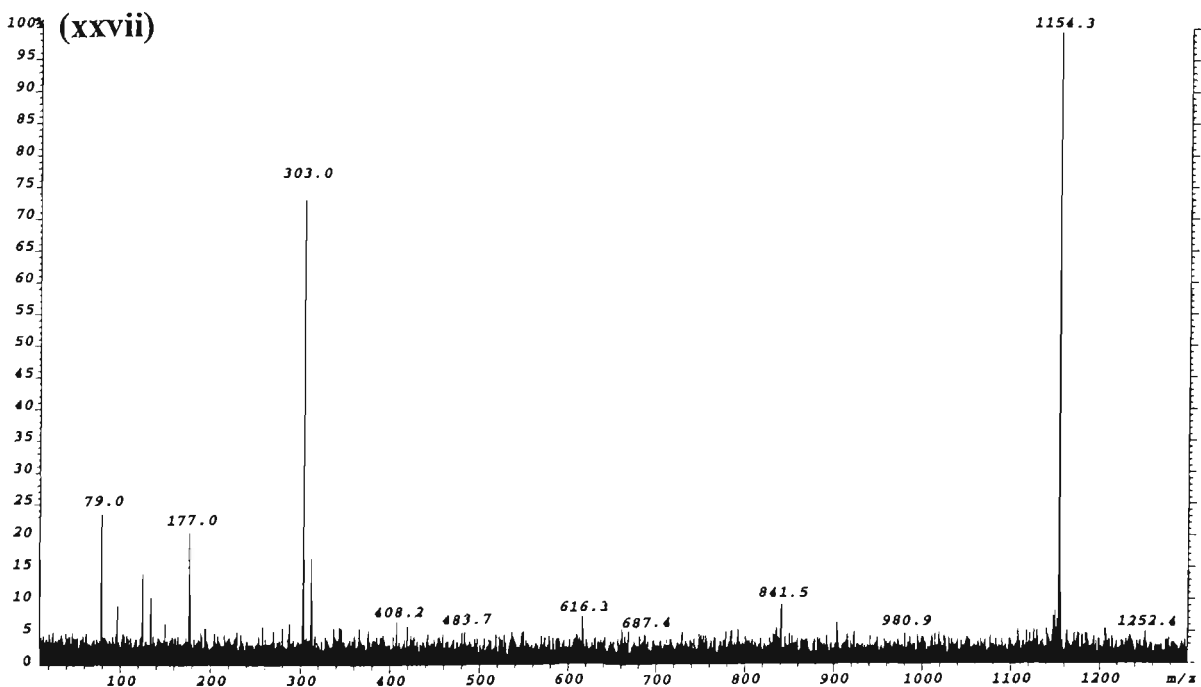
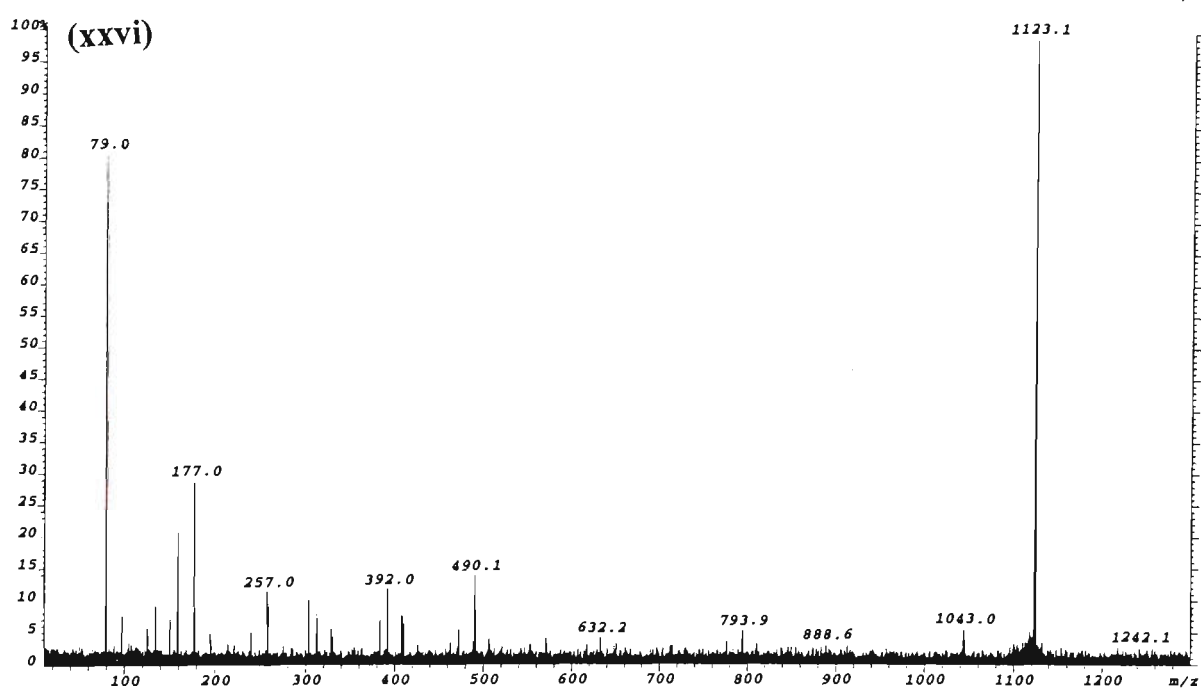
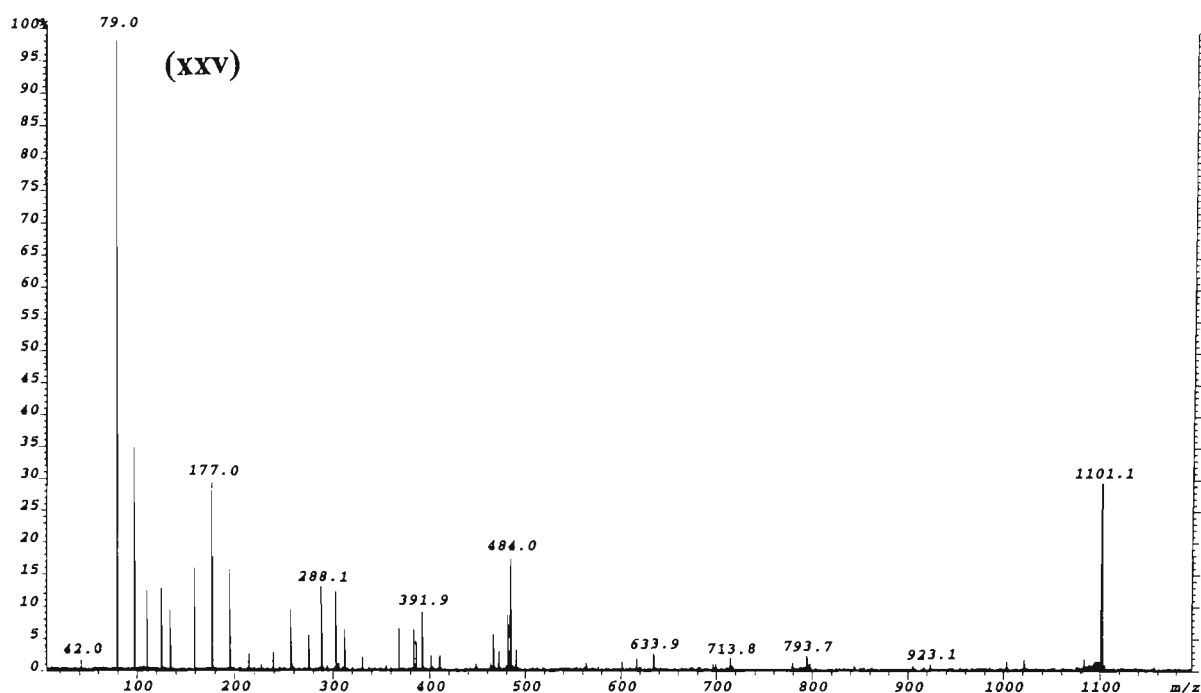


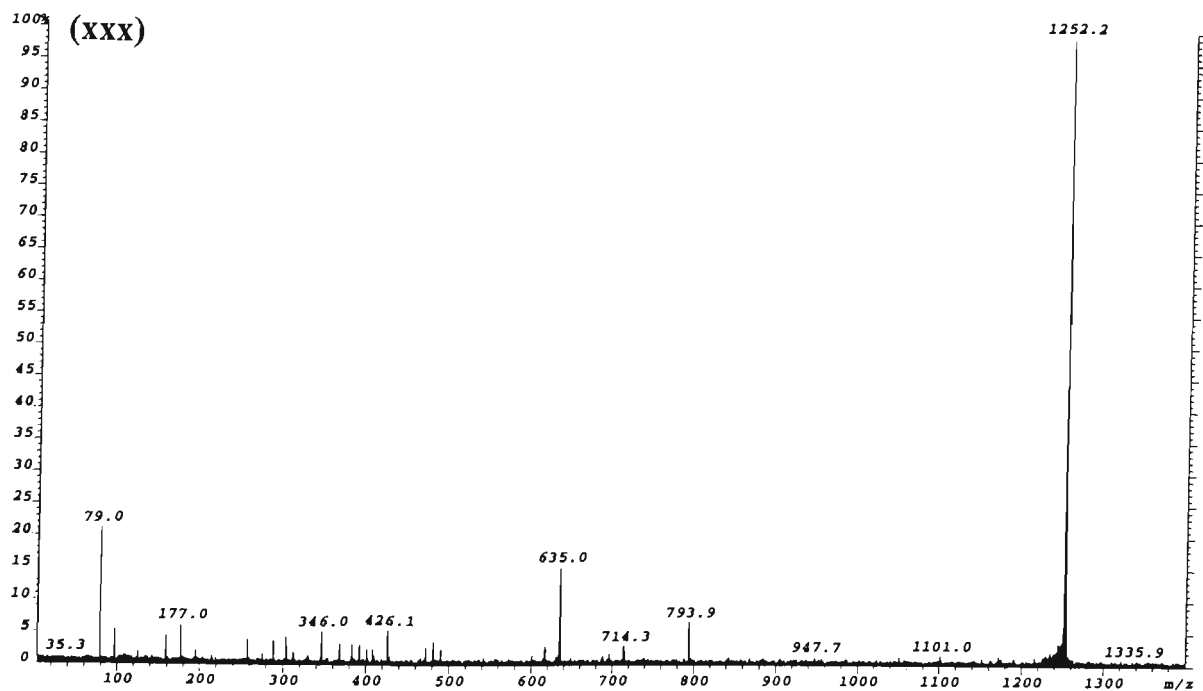
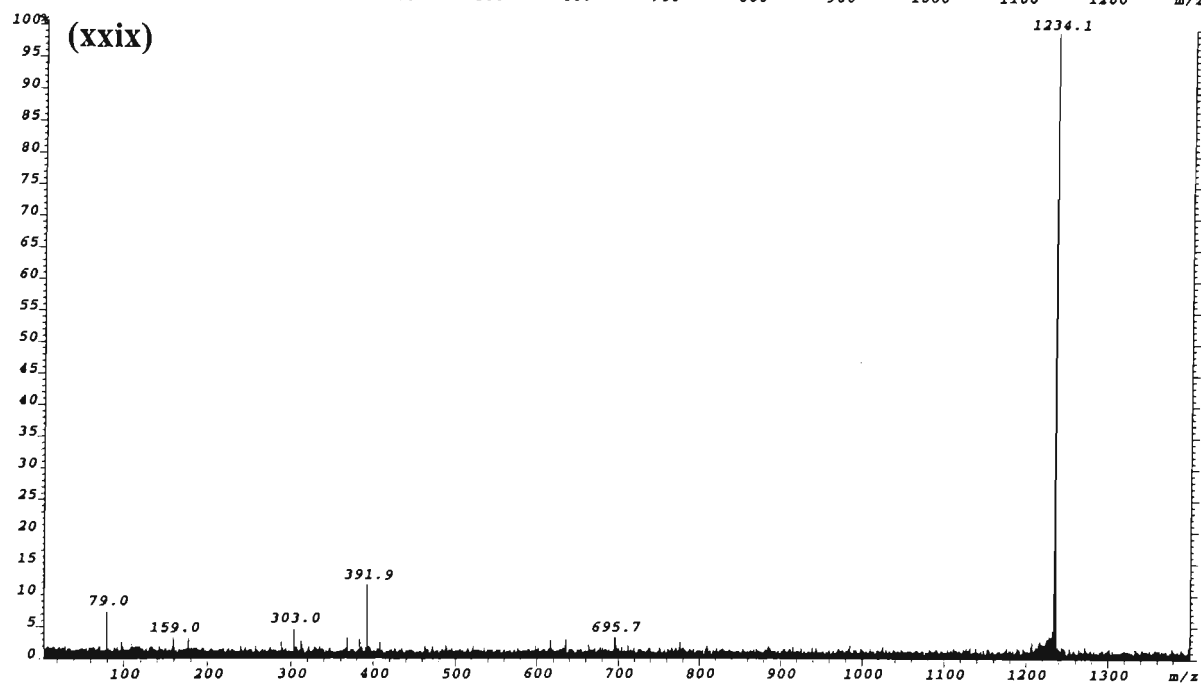
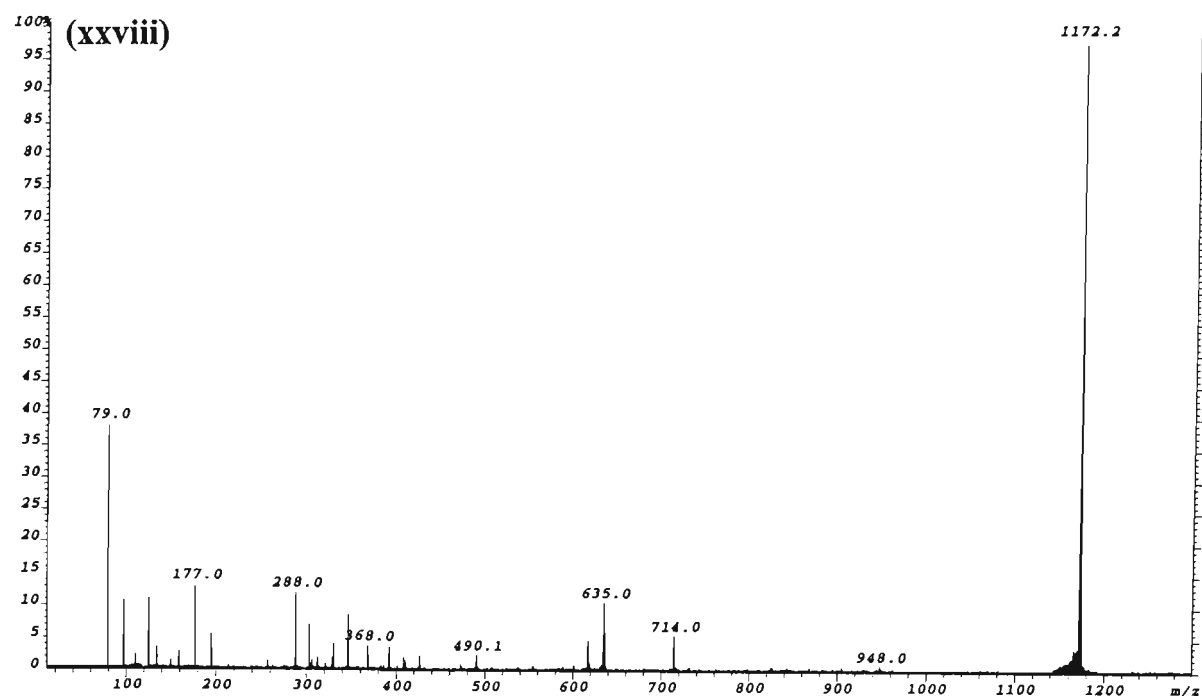


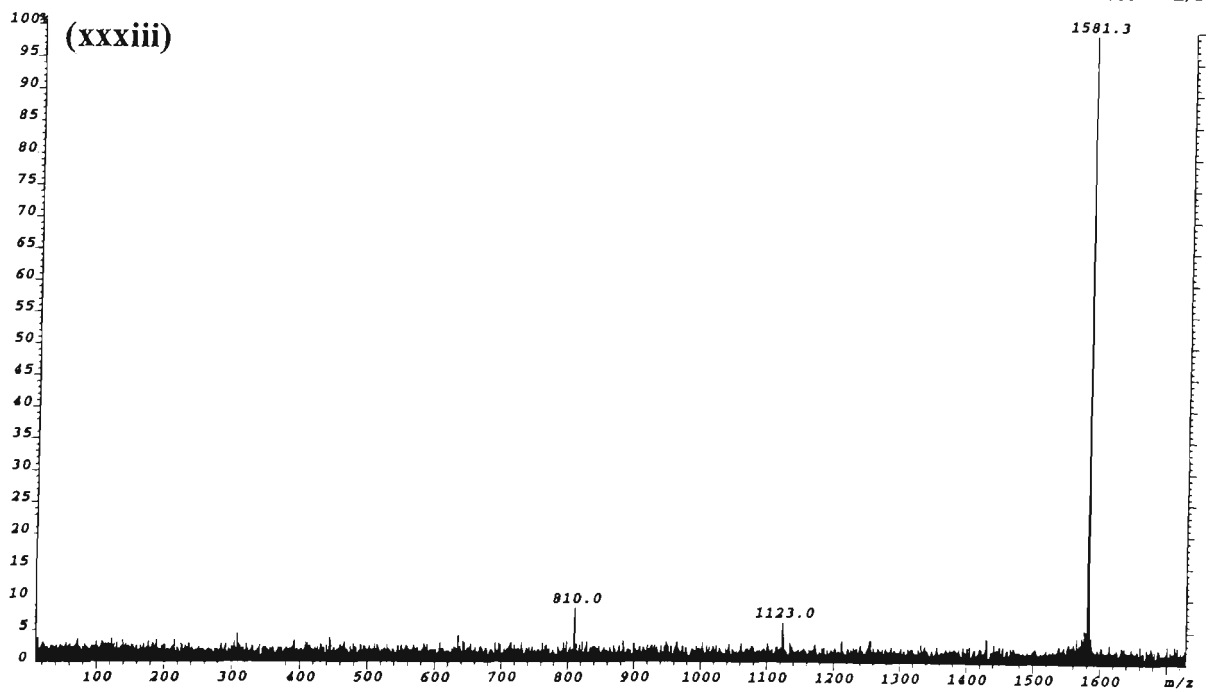
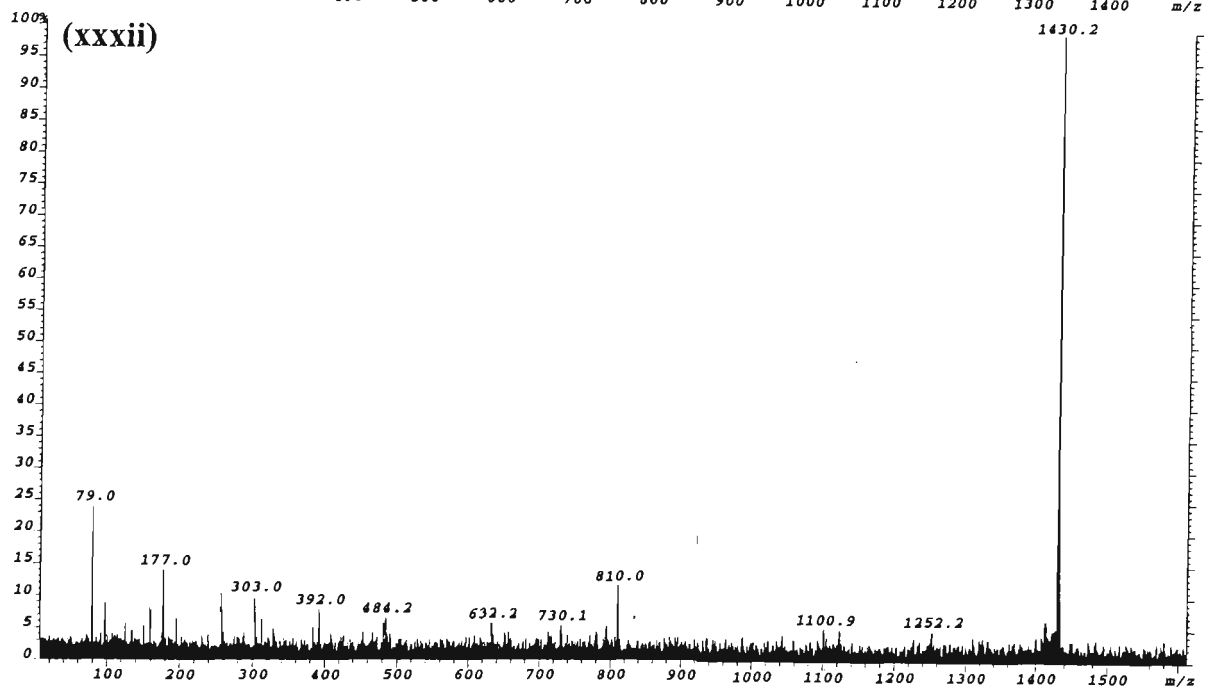
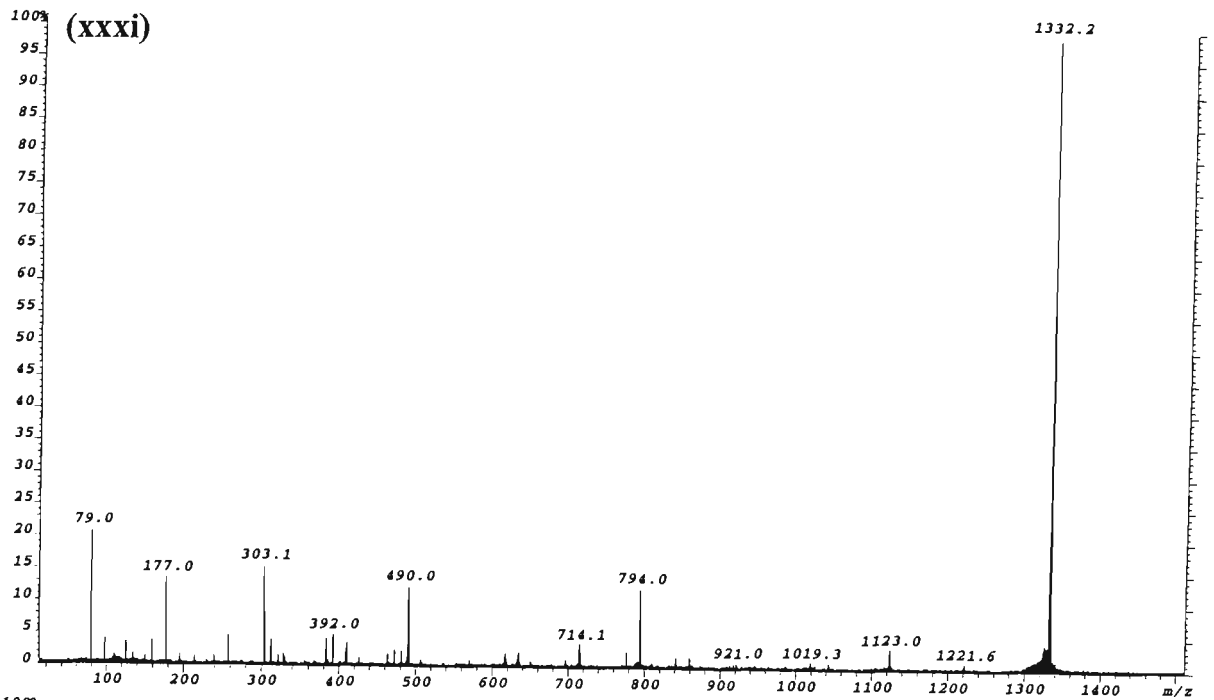


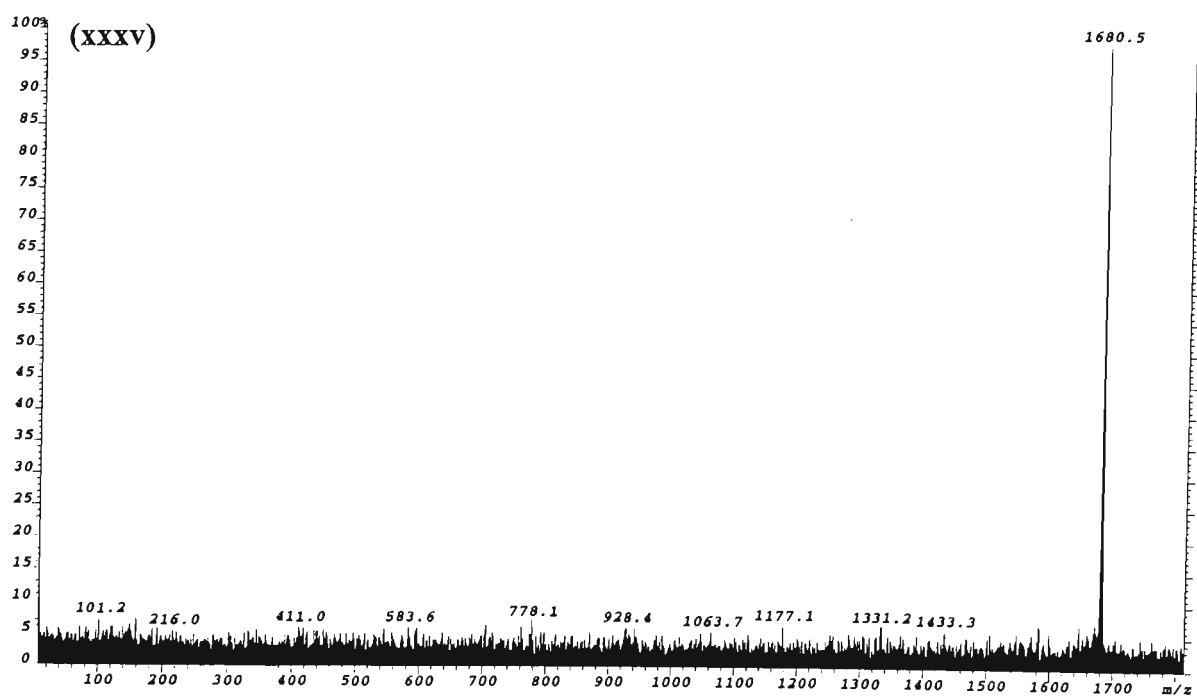
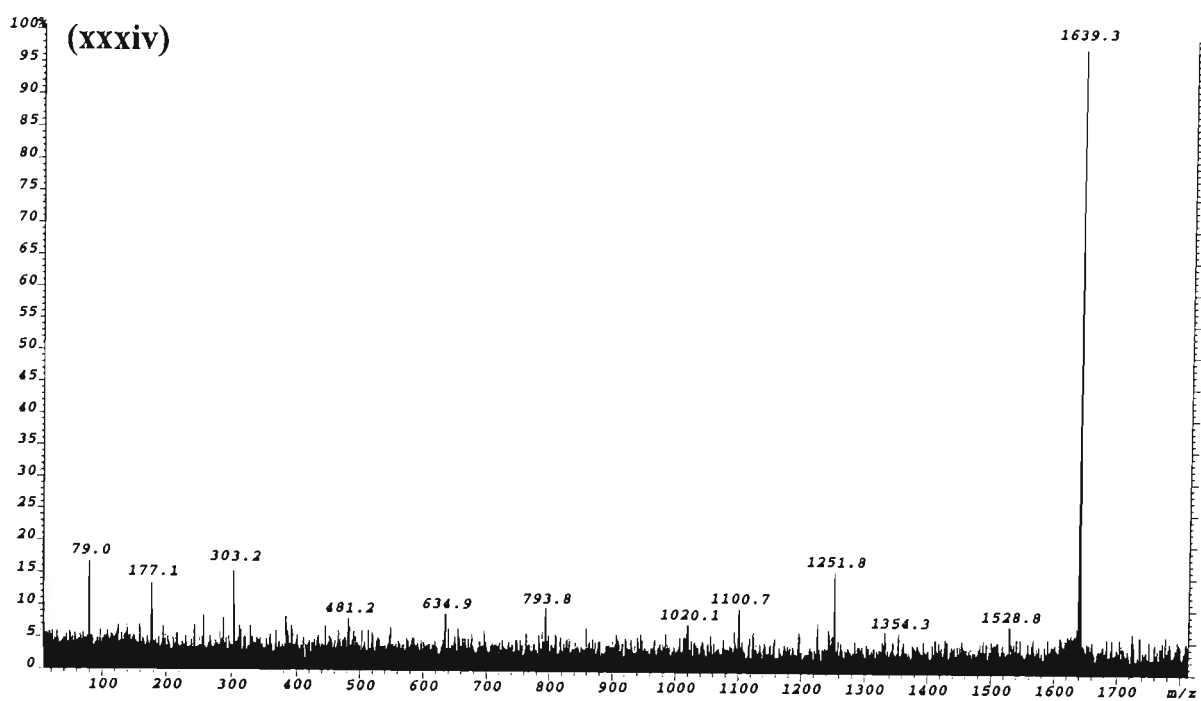






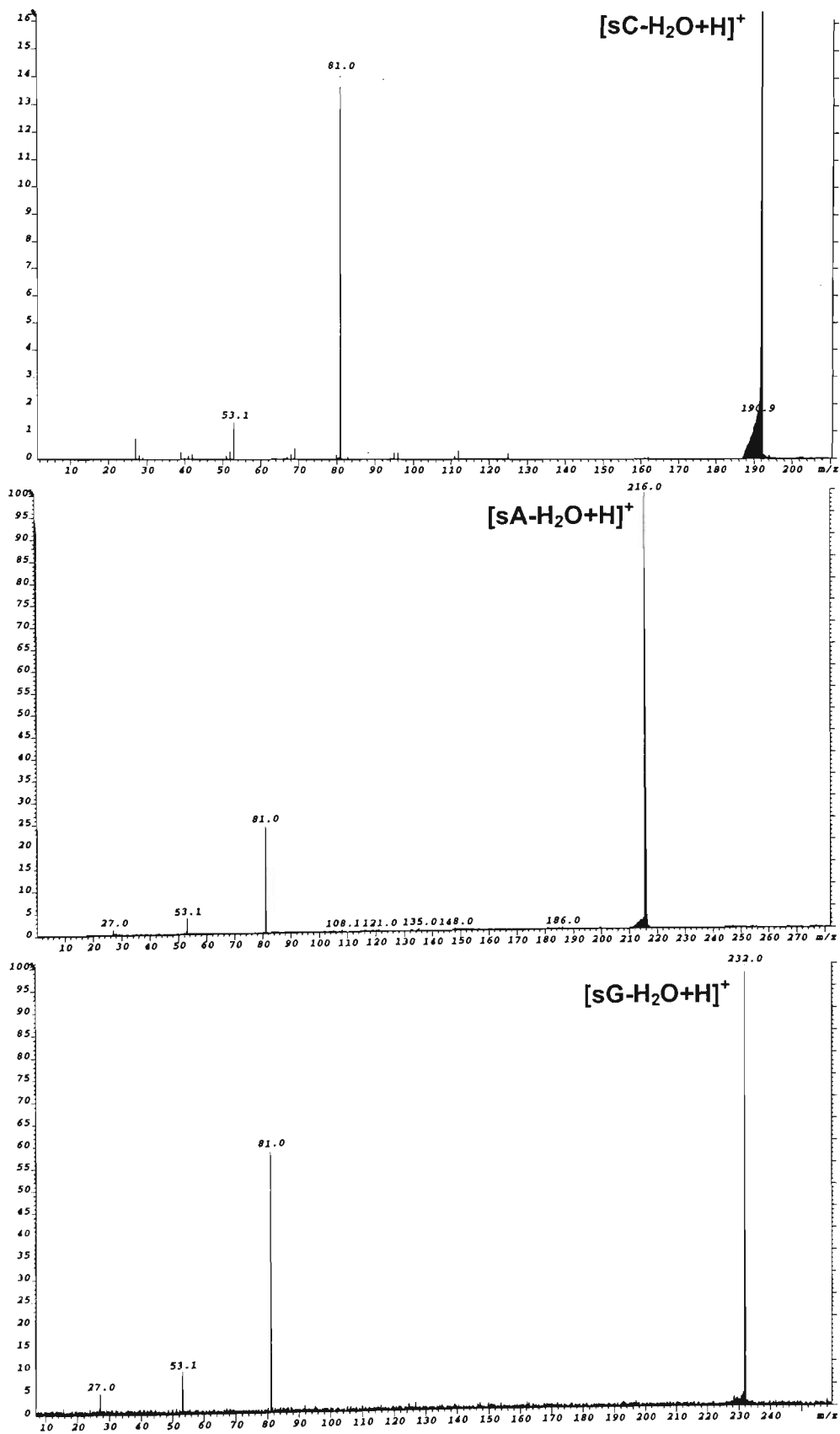






Appendix 4.4

ESI-MS/MS spectra of the 'source-generated' product ion types $[sB-H_2O+H]^+$ observed in the ESI-MS/MS spectra of the $[M+H]^+$ ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

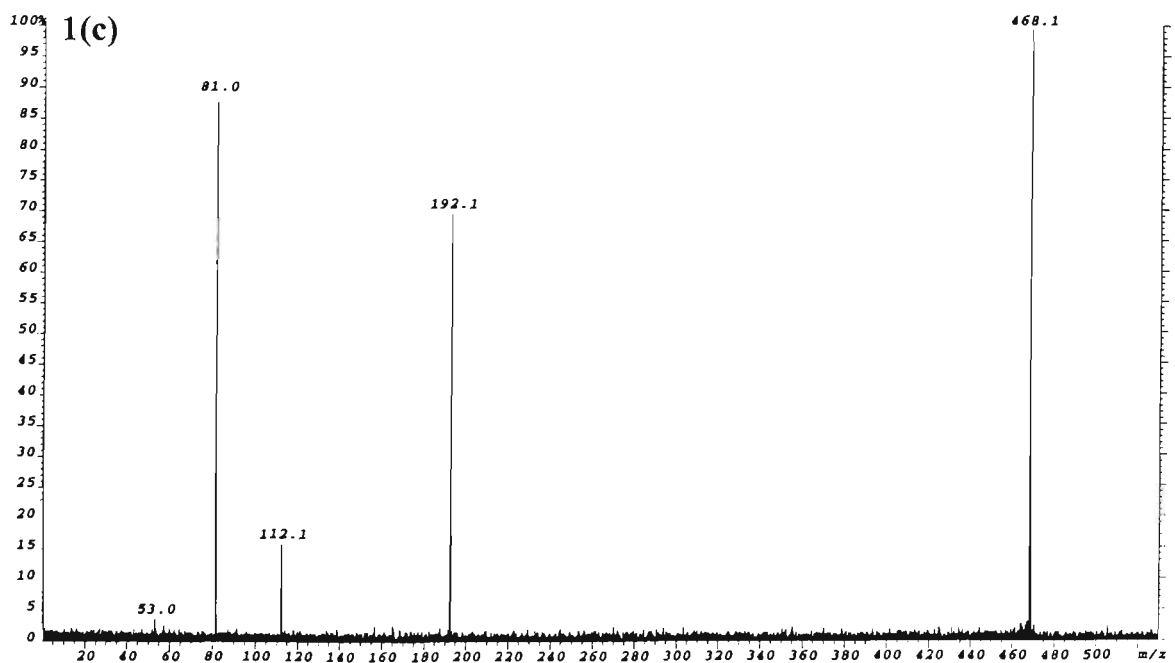
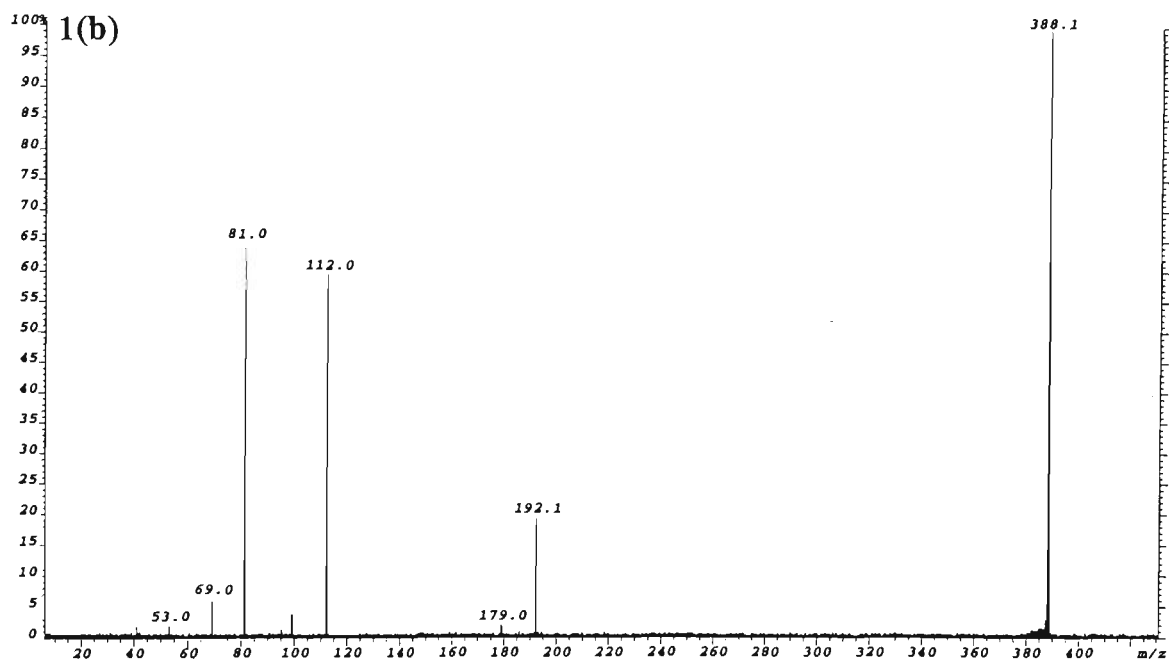
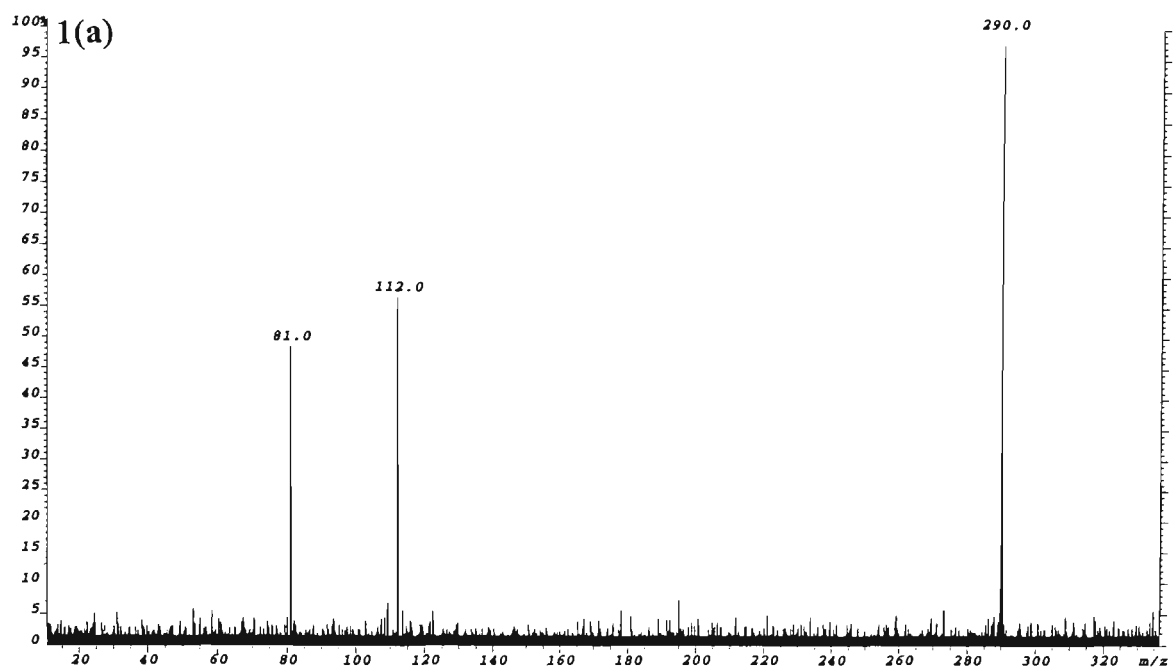


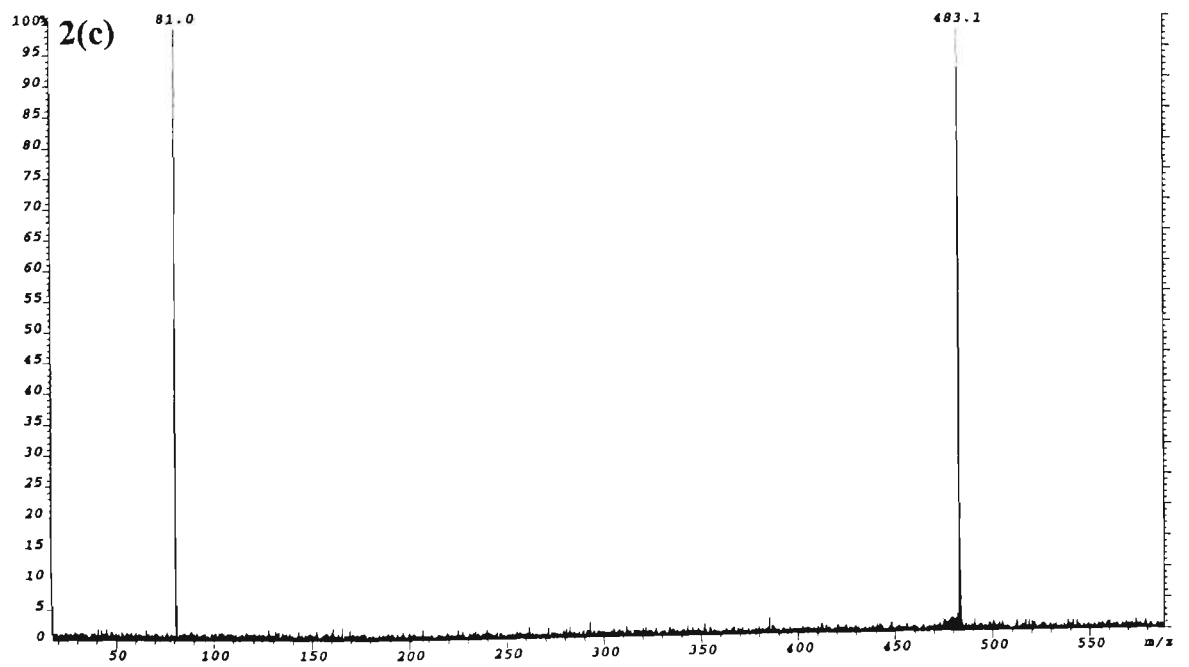
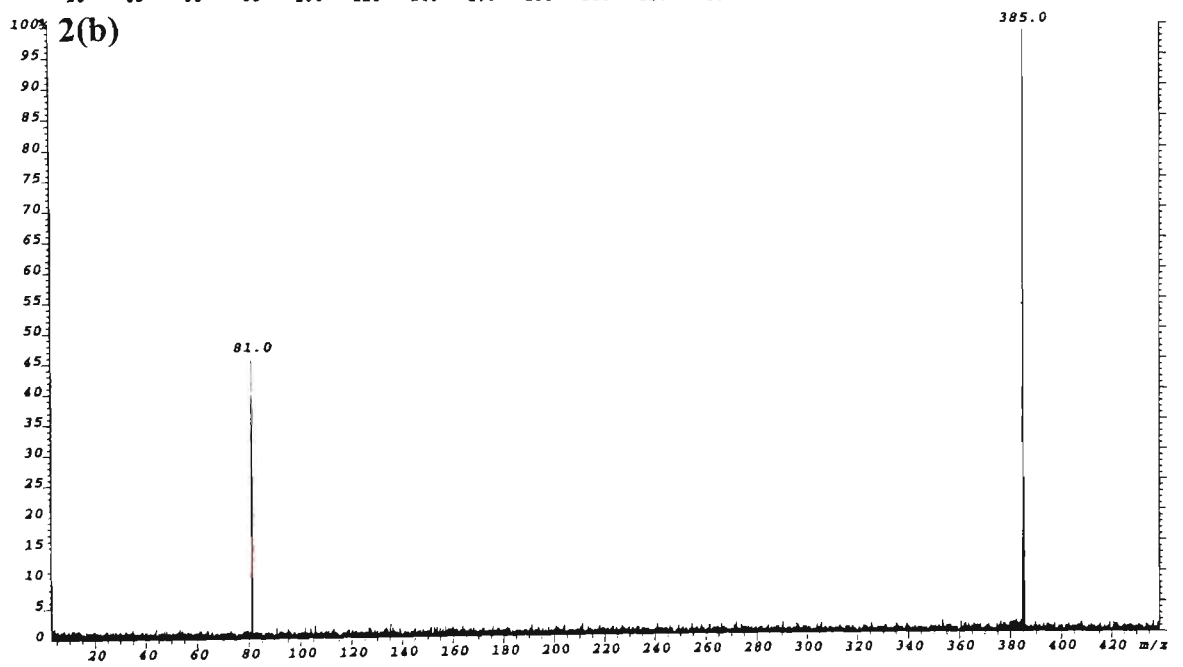
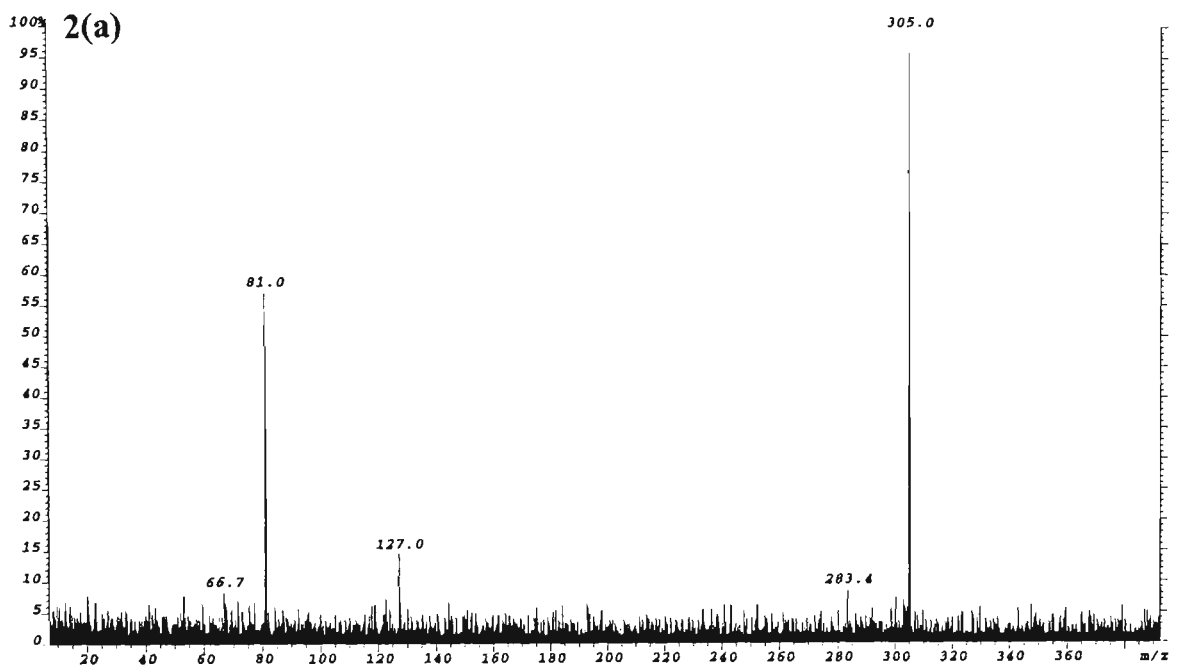
Appendix 4.5

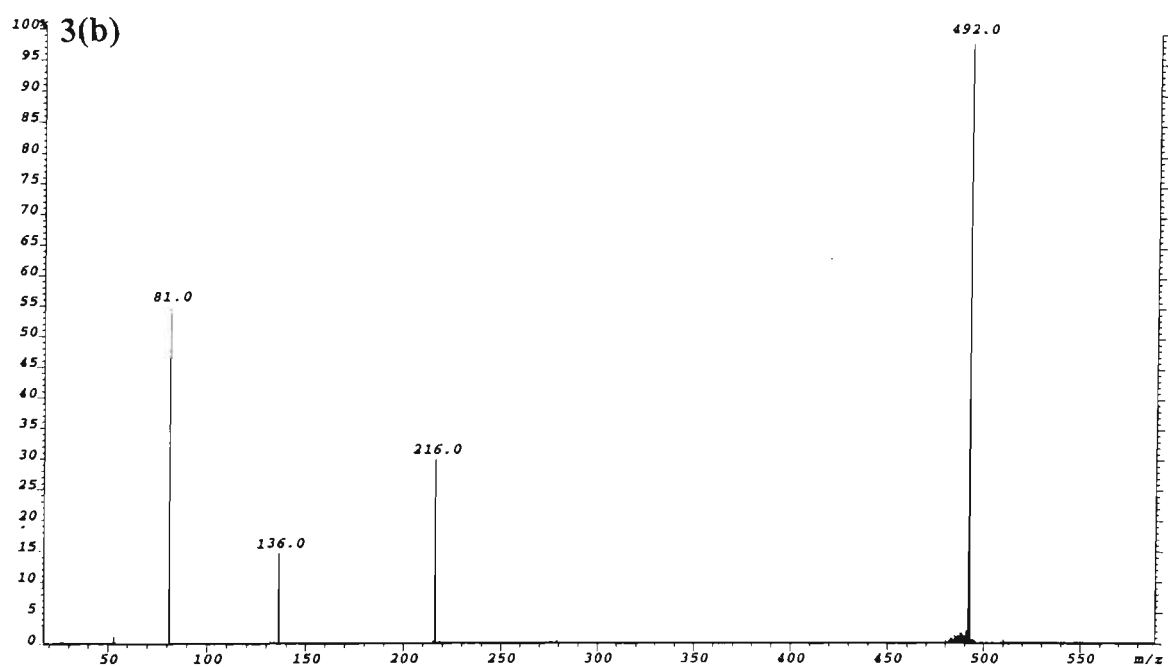
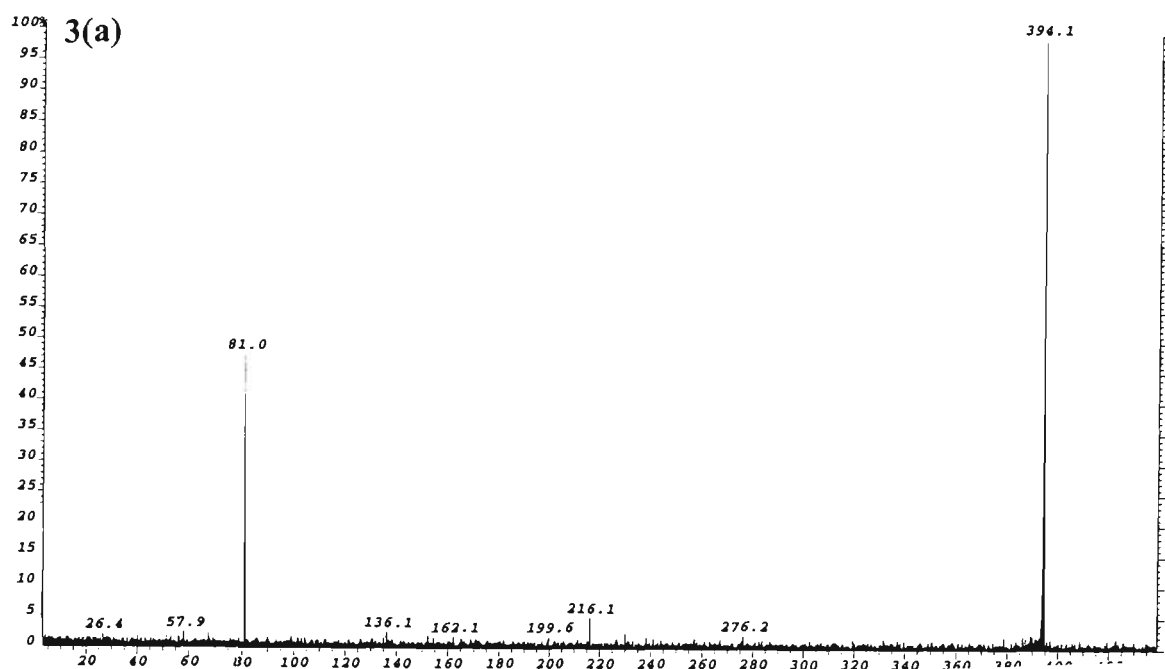
ESI-MS/MS spectra of the ‘source-generated’ nucleotide product ion types observed in the ESI-MS/MS spectra of the $[M+H]^+$ ions of 5'-d(CACGTG)-3' and 5'-d(CGTACG)-3'.

Figure No.	Precursor Ion	
	Assignment	m/z
1(a)	$[C_{nt}+H]^+$	290.0
1(b)	$[C_{nt}+(s \text{ or } p+H_2O)+H]^+$	388.1
1(c)	$[C_{nt}+p+s+H]^+$	468.1
2(a)	$[T_{nt}+H]^+$	305.0
2(b)	$[T_{nt}+(p \text{ or } s-H_2O)+H]^+$	385.0
2(c)	$[T_{nt}+p+s+H]^+$	483.1
3(a)	$[A_{nt}+(p \text{ or } s-H_2O)+H]^+$	394.1
3(b)	$[A_{nt}+p+s+H]^+$	492.0

s = deoxyribose-H₂O (C₅H₆O₂), Mr =98.0368 Da
p = PO₃H, Mr = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp





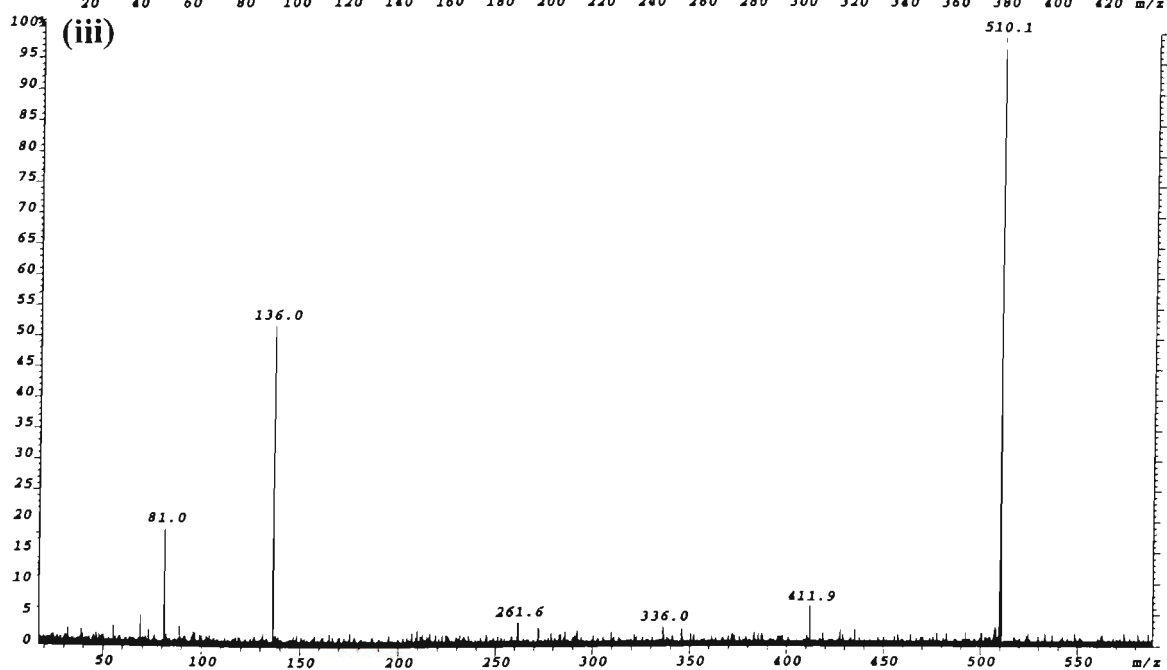
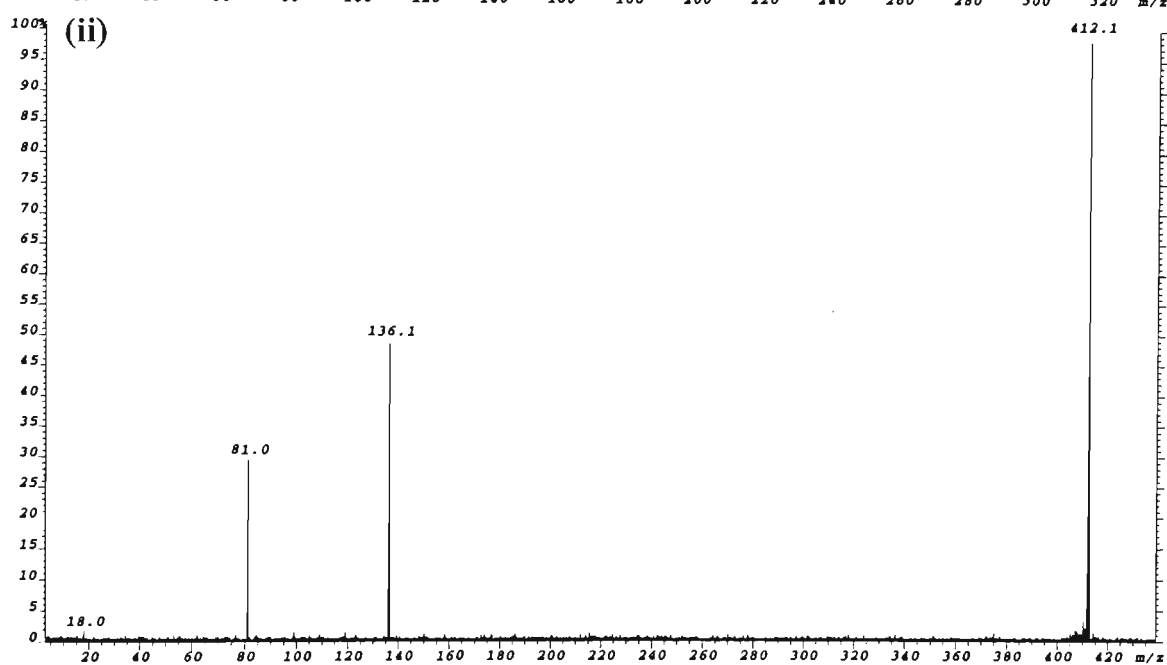
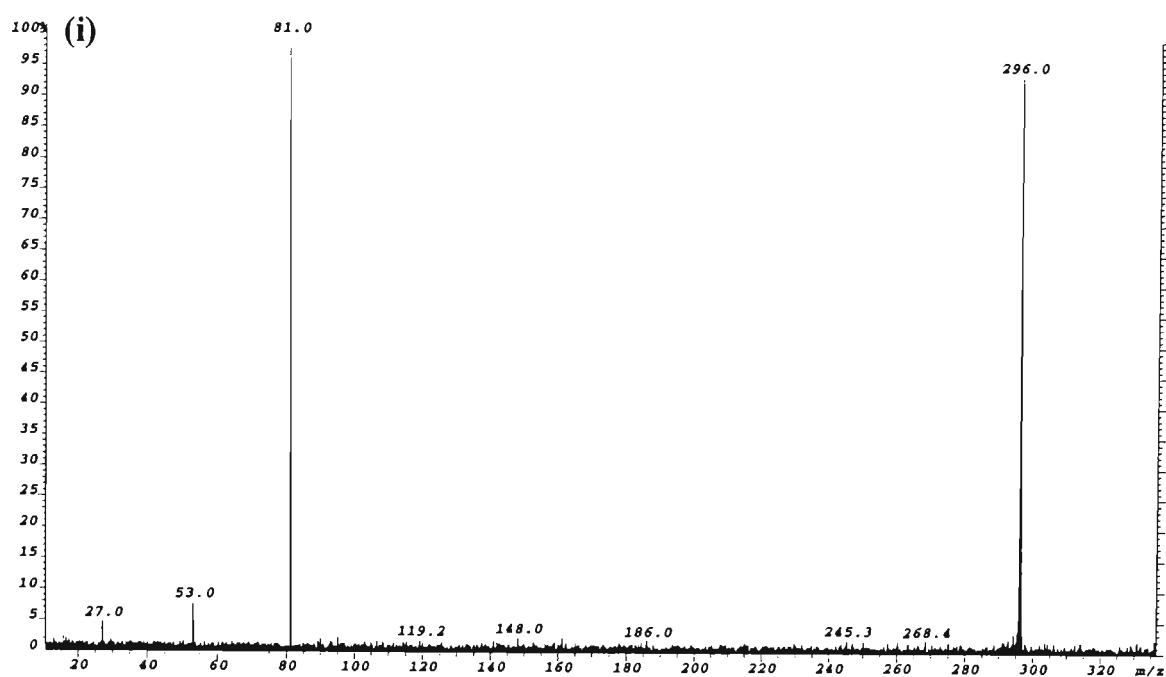


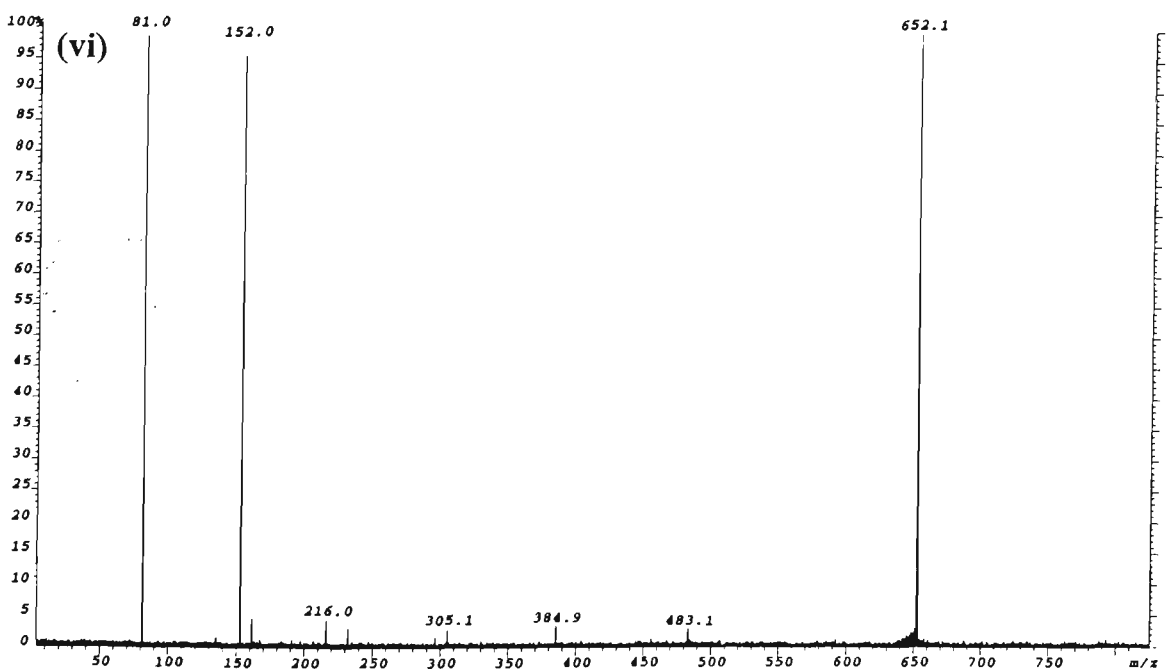
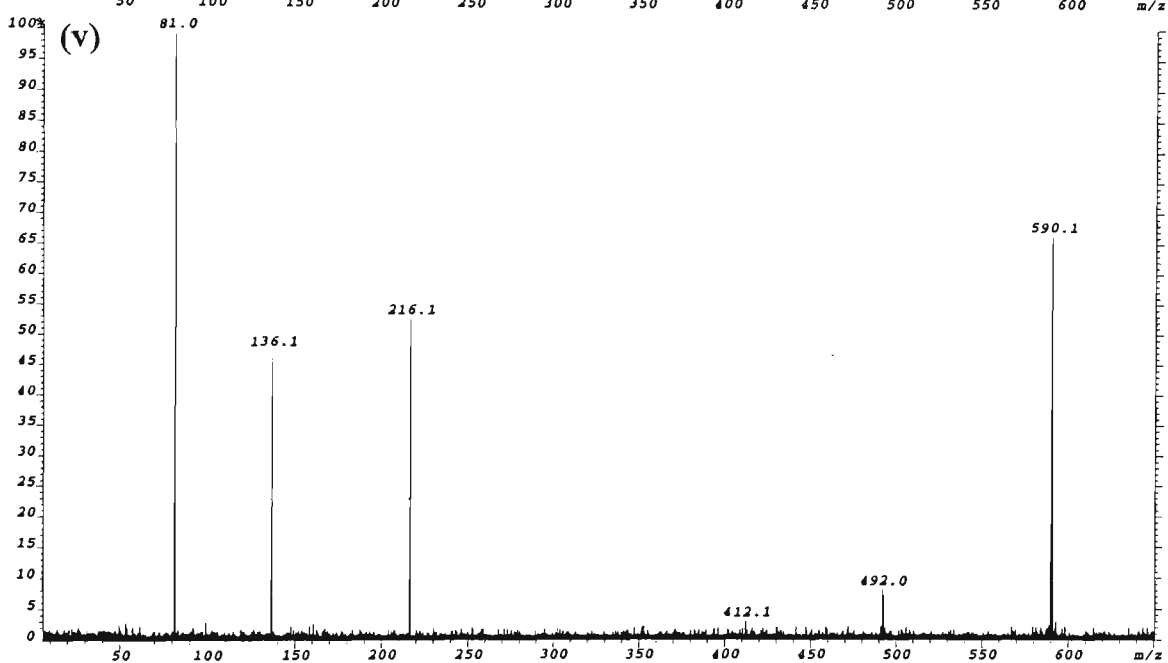
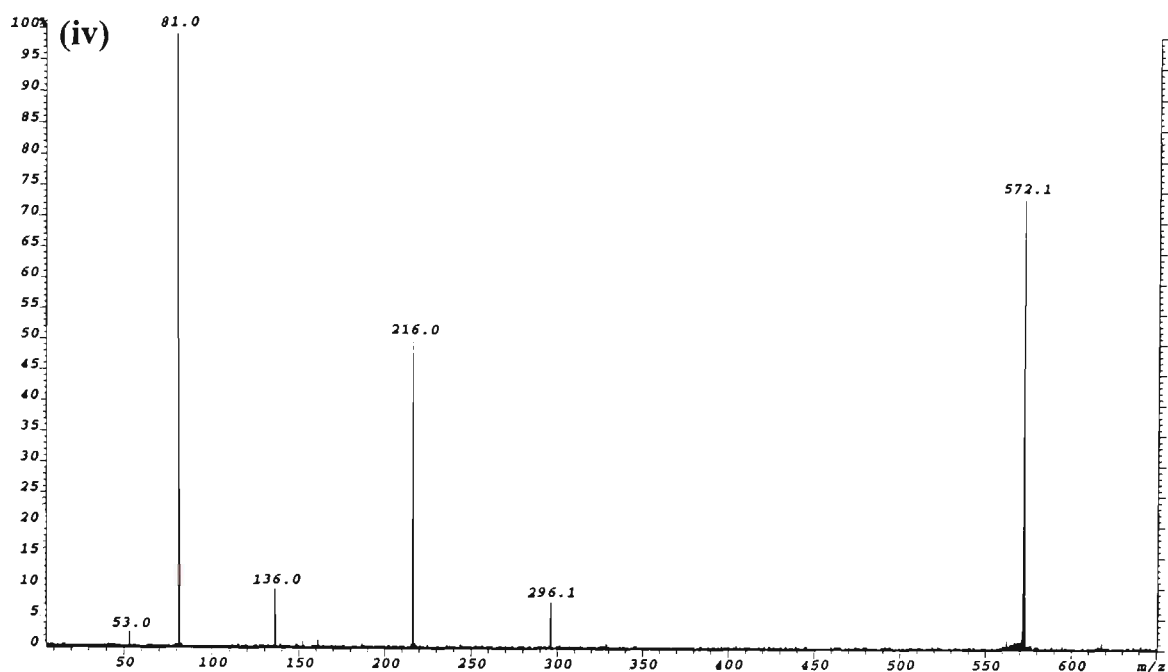
Appendix 4.6

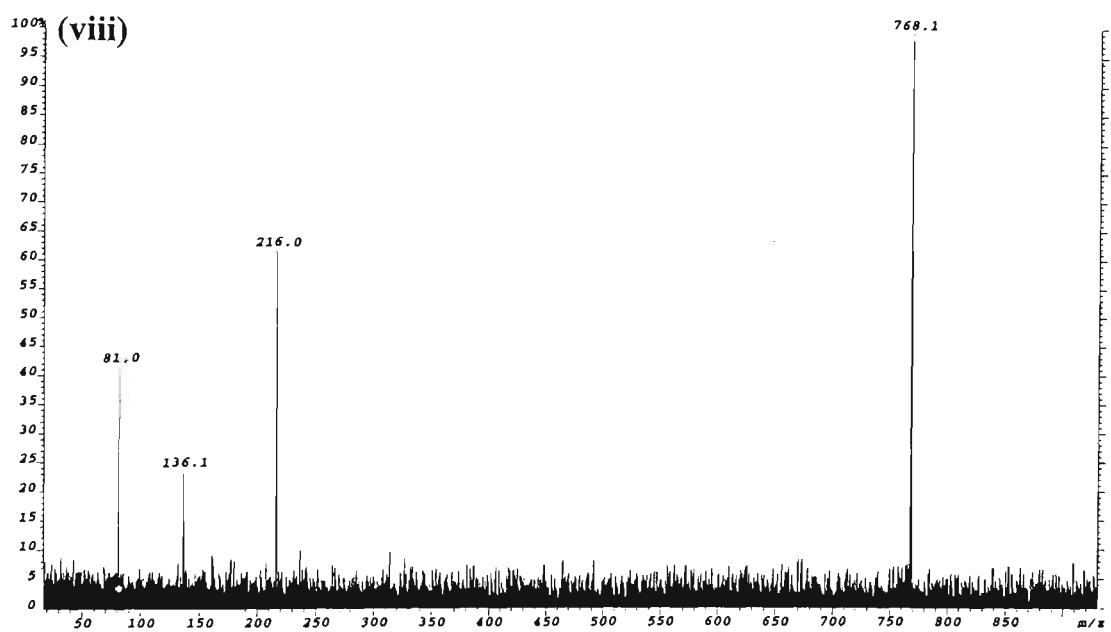
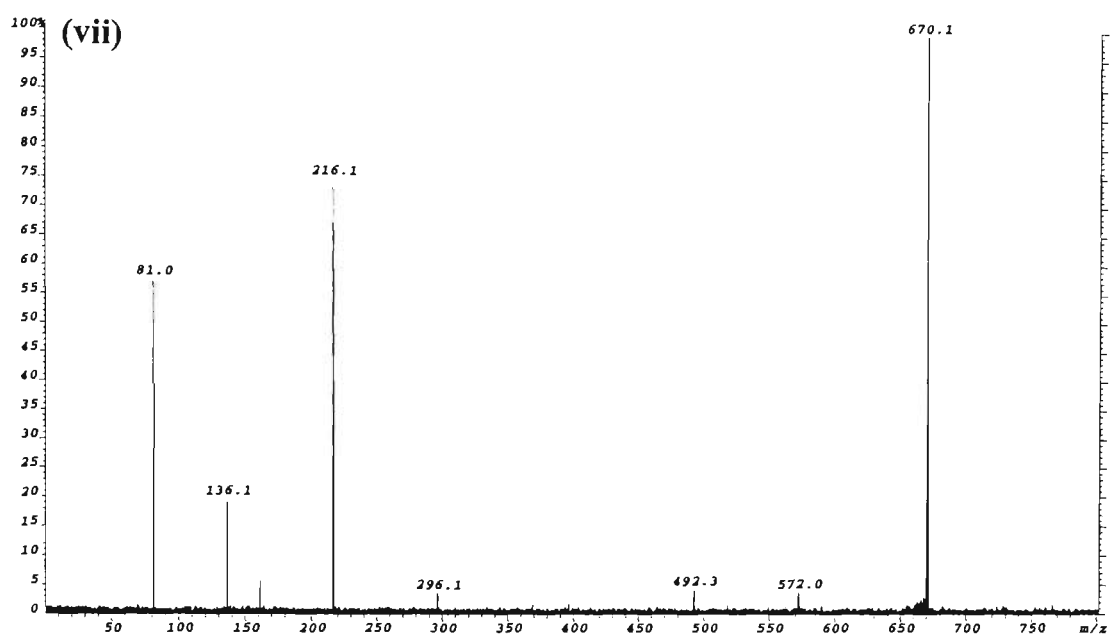
ESI-MS/MS spectra of the ‘source-generated’ nucleotide product ion types of adenine observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CACGTG)-3'.

Figure No.	Assignment	Precursor Ion	m/z
(i)	$[A_{nt}-H_2O+H]^+$		296.0
(ii)	$[A_{nt}+(s \text{ or } p+H_2O)+H]^+$		412.1
(iii)	$[A_{nt}+p+s+H_2O+H]^+$		510.1
(iv)	$[A_{nt}+p+s+p+H]^+$ and/or $[A_{nt}+s+p+s-H_2O+H]^+$		572.1
(v)	$[A_{nt}+p+s+p+H_2O+H]^+$ and/or $[A_{nt}+s+p+s+H]^+$		590.1
(vi)	$[A_{nt}+p+s+p+s-H_2O+H]^+$		652.1
(vii)	$[A_{nt}+s+p+s+p+H]^+$		670.1
(viii)	$[A_{nt}+p+s+p+s+p+H_2O+H]^+$ and/or $[A_{nt}+s+p+s+p+s+p+H]^+$		768.1

s = deoxyribose-H₂O (C₅H₆O₂), Mr =98.0368 Da
p = PO₃H, Mr = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp

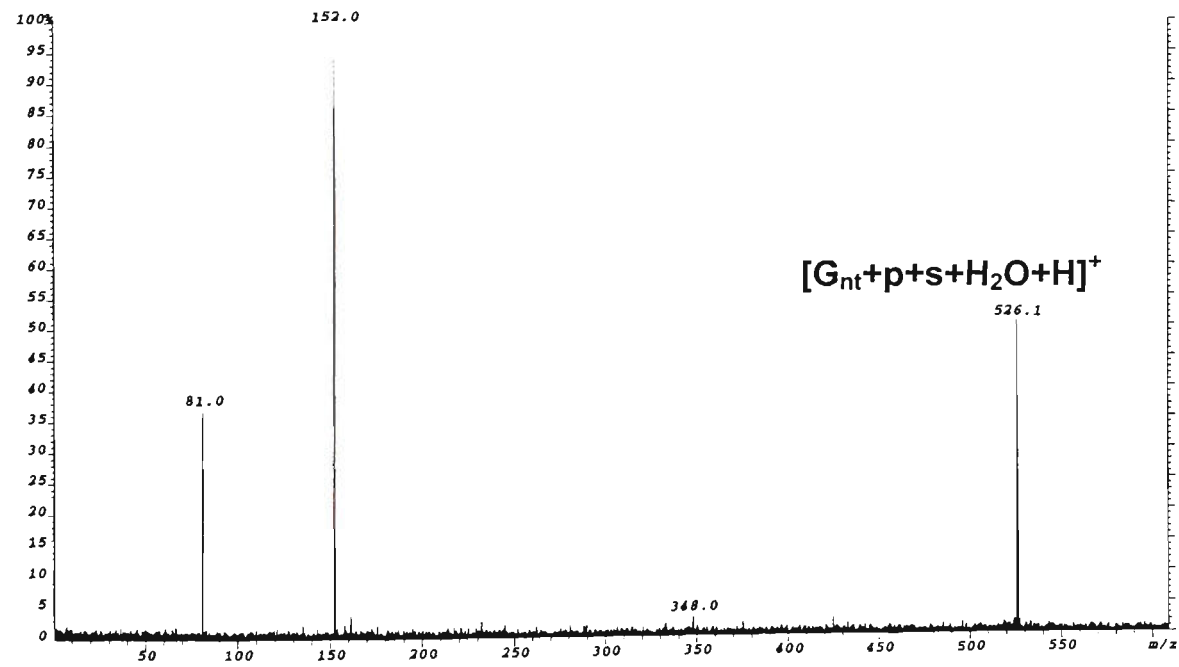
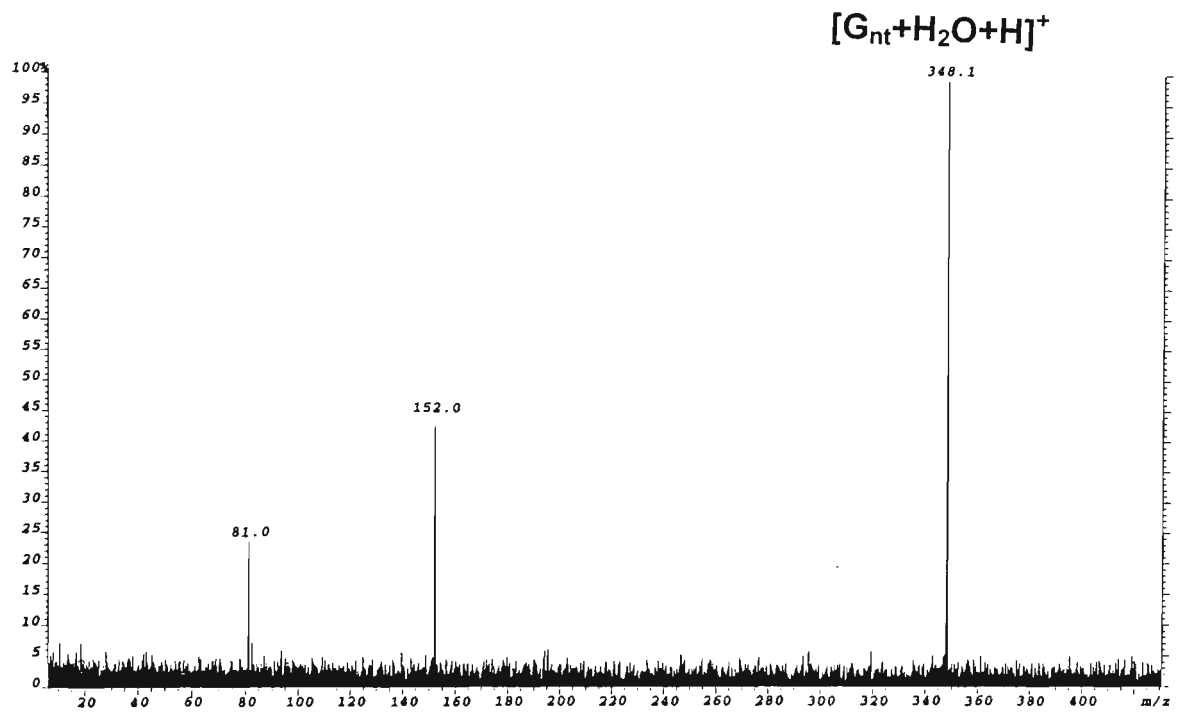






Appendix 4.7

ESI-MS/MS spectra of the 'source-generated' nucleotide product ion types of guanine observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGTACG)-3'



Appendix 4.8

ESI-MS/MS spectra of the ‘source-generated’ polynucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CACGTG)-3'.

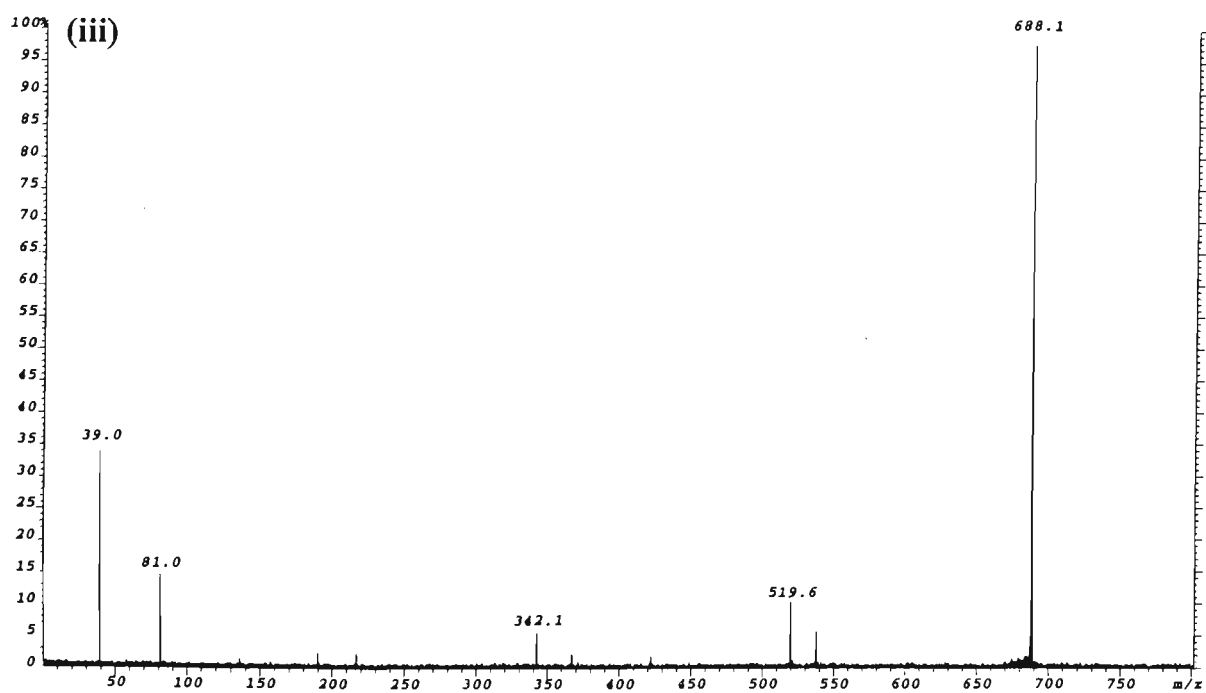
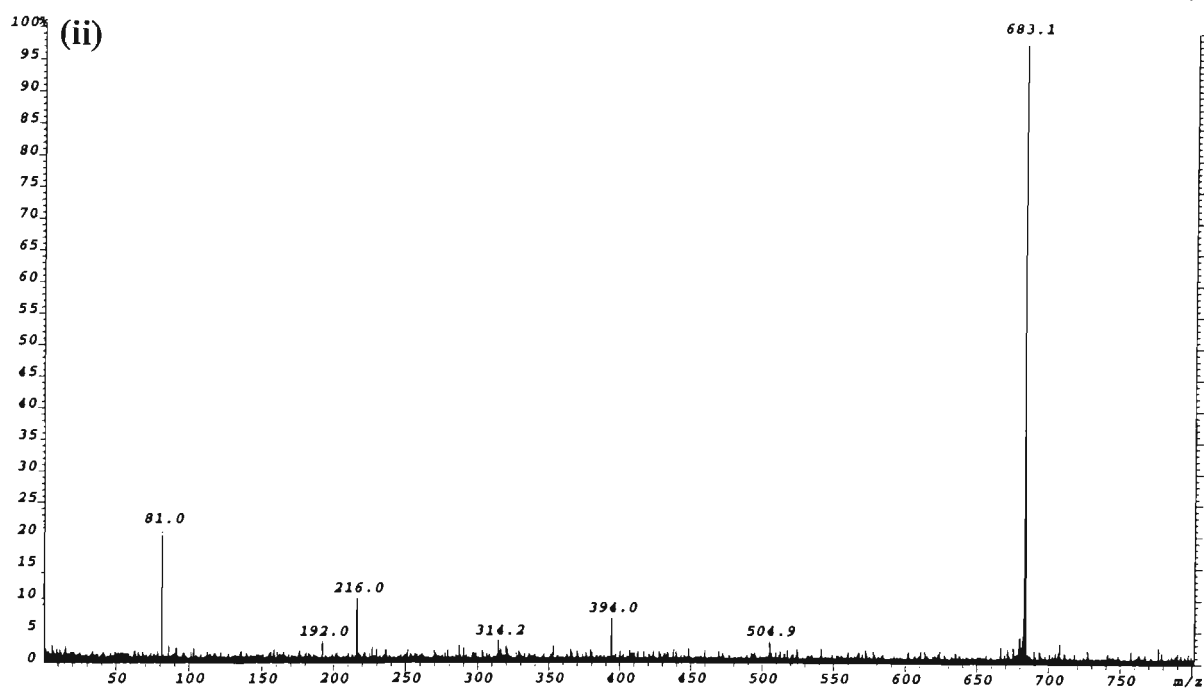
Figure No.	Precursor Ion	
	<i>m/z</i>	Assignment
(i)	603.1	c_2^+ and/or $[(AC)+H]^+$
(ii)	683.1	$[(AC \text{ or } CA)+s-H_2O+H]^+$ and/or $[(AC)+p+H]^+$
(iii)	688.1	$[d_3-CH-CH+2H]^+$ and/or $[A_n+s+p+s+p+H_2O+H]^+$
(iv)	701.1	$[a_3-B_3H]^+$ and/or $[(AC)+p+H_2O+H]^+$
(v)	732.1	$[z_3-GH]^+$ and/or $[(GT)+s+H]^+$ and/or $[(GT)+p+H_2O+H]^+$
(vi)	781.1	$[c_3-CH]^+$ and/or $[(AC)+p+s+H]^+$
(vii)	799.1	$[d_3-CH+2H]^+$ and/or $[(AC)+p+s+H_2O+H]^+$
(viii)	812.1	$[x_3-GH+2H]^+$ and/or $[(GT)+p+s+H]^+$
(ix)	830.1	$[w_3-GH+H]^+$ and/or $[(GT)+p+s+H_2O+H]^+$
(x)	855.1	$[a_4-B_4H-AH]^+$
(xi)	861.1	$[(CA \text{ or } CA)+s+p+s-H_2O+H]^+$ and/or $[(AC)+p+s+p+H]^+$
(xii)	879.1	$[a_4-B_4H-CH]^+$ and/or $[(AC)+s+p+s+H]^+$ and/or $[(AC)+p+s+p+H_2O+H]^+$
(xiii)	981.1	$[w_3+2H]^+$
(xiv)	990.2	$[a_4-B_4H]^+$
(xv)	1270.2	$[w_4+2H]^+$
(xvi)	1414.1	$[c_5-CH]^+$ and/or $[x_5-GH]^+$
(xvii)	1472.2	$[w_5-CH+2H]^+$
(xviii)	1531.0	$[M-GH-CH+H]^+$
(xix)	1623.3	$[M-GH-H_2O+H]^+$
(xx)	1641.3	$[M-GH+H]^+$
(xxi)	1681.3	$[M-CH+H]^+$
(xxii)	1718.2	$[2M-GH+2H]^{2+}$
(xxiii)	1737.1	$[2M-CH+2H]^{2+}$

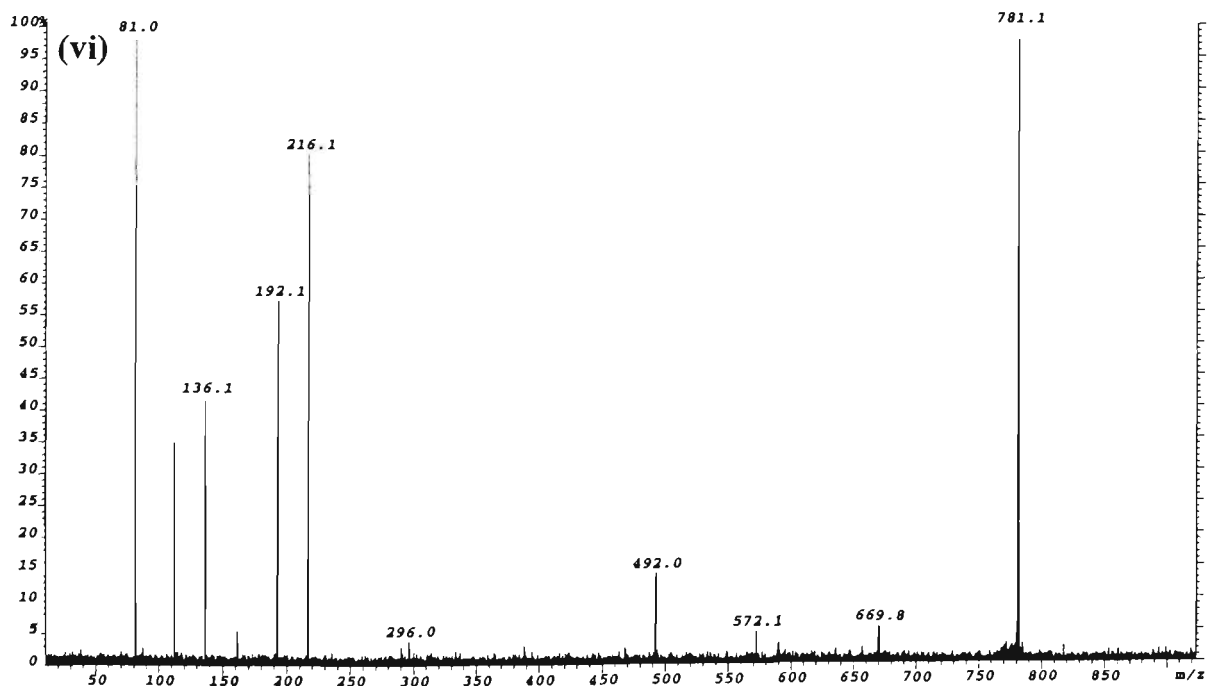
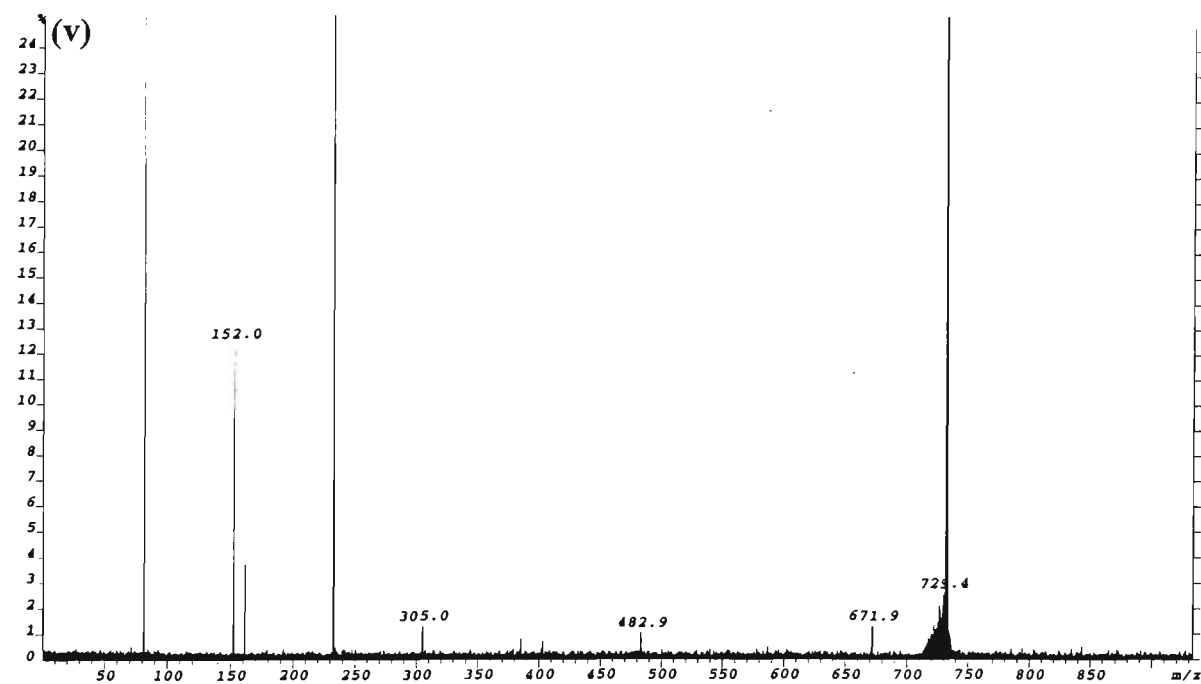
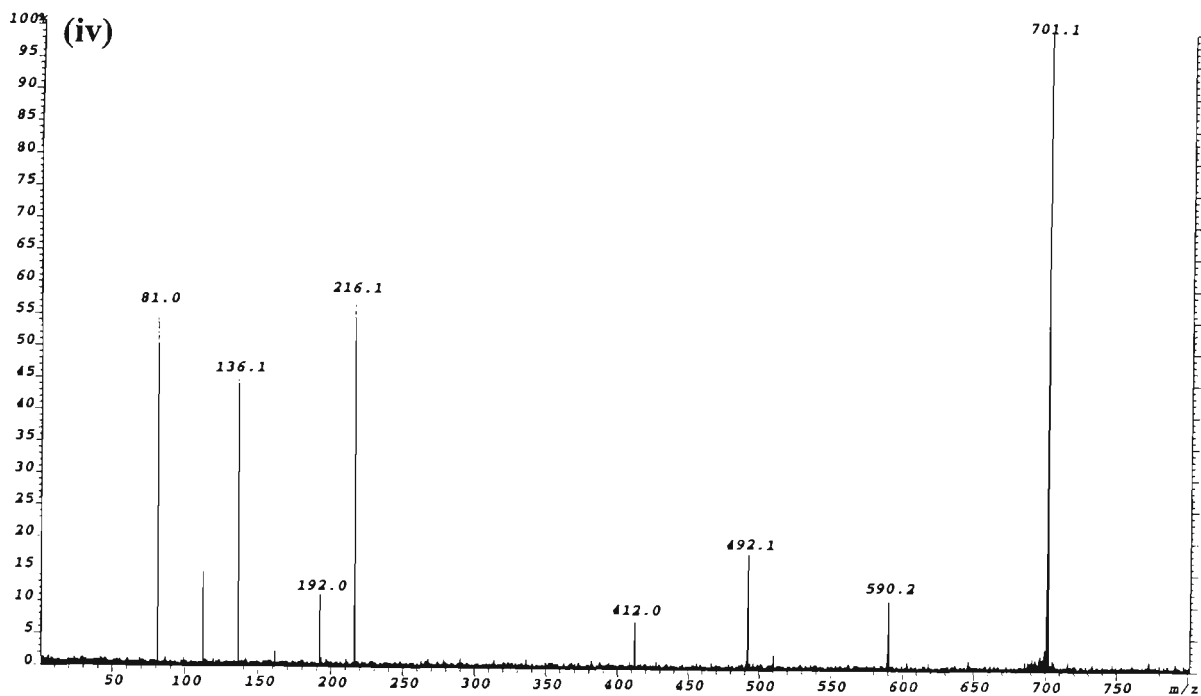
s = deoxyribose-H₂O (C₅H₆O₂), *M_r* = 98.0368 Da

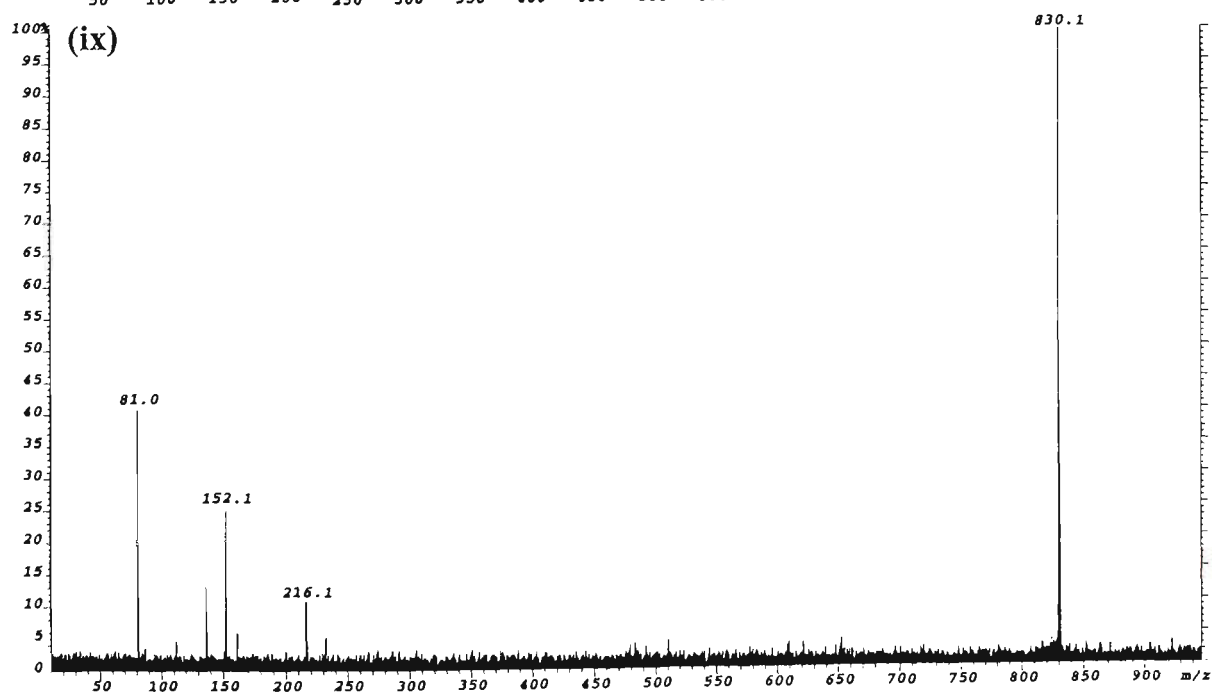
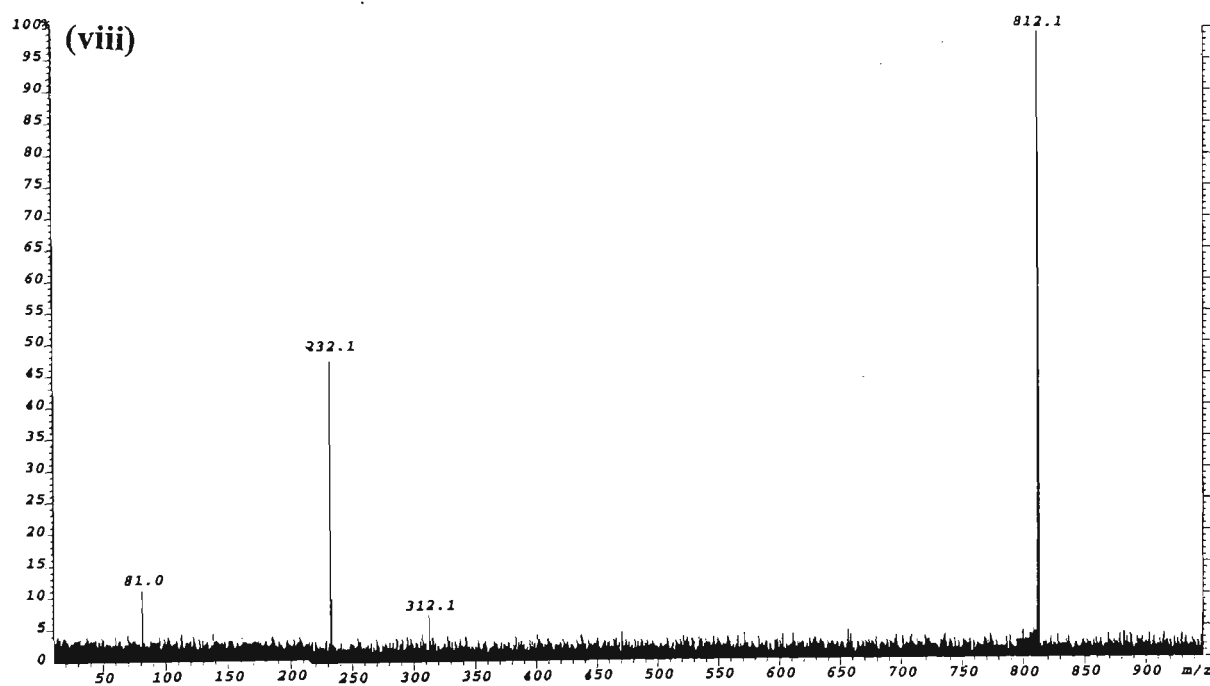
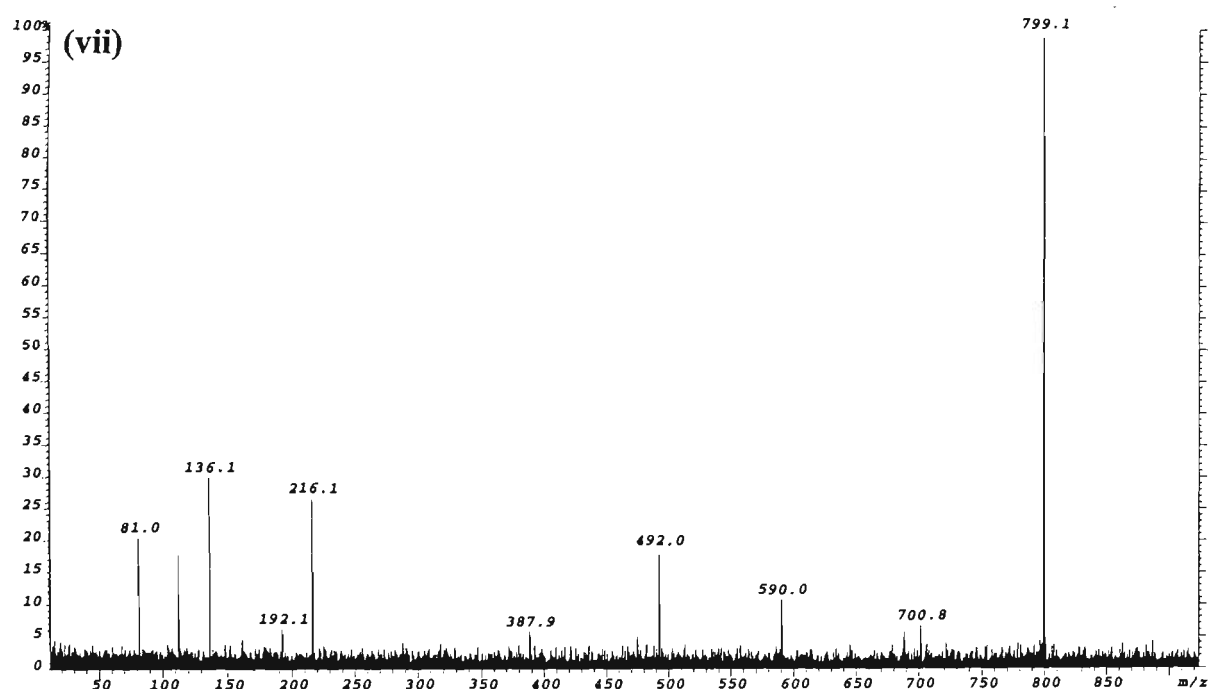
p = PO₃H, *M_r* = 79.9663 Da

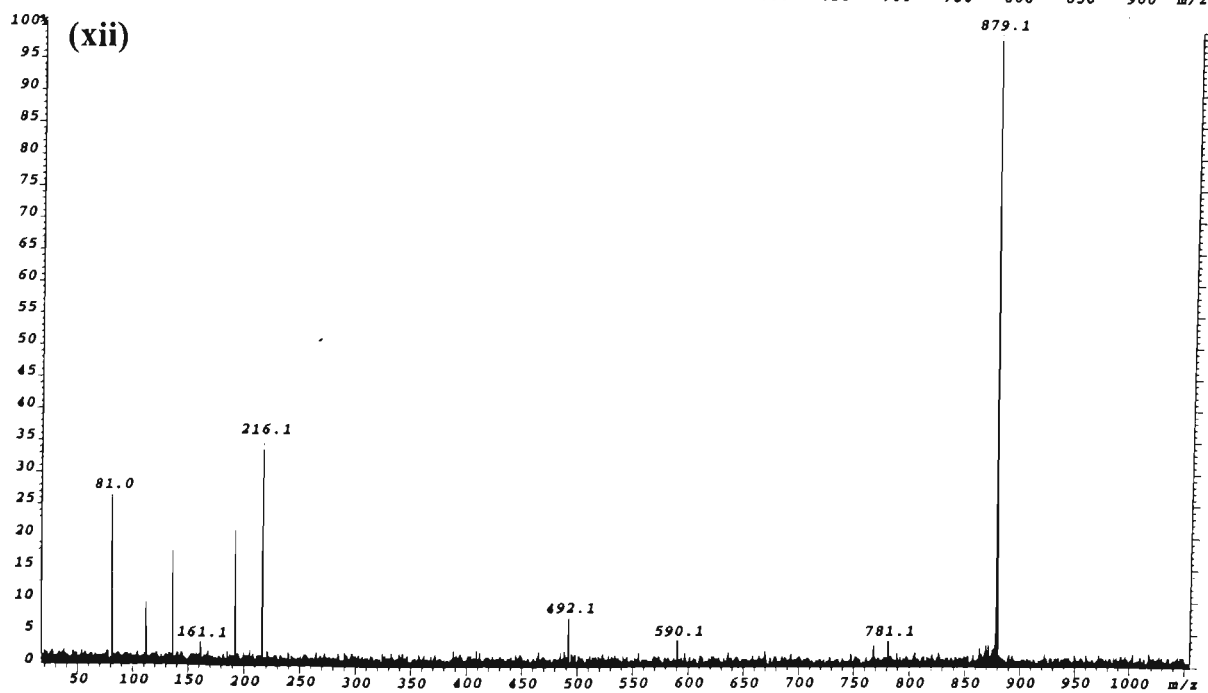
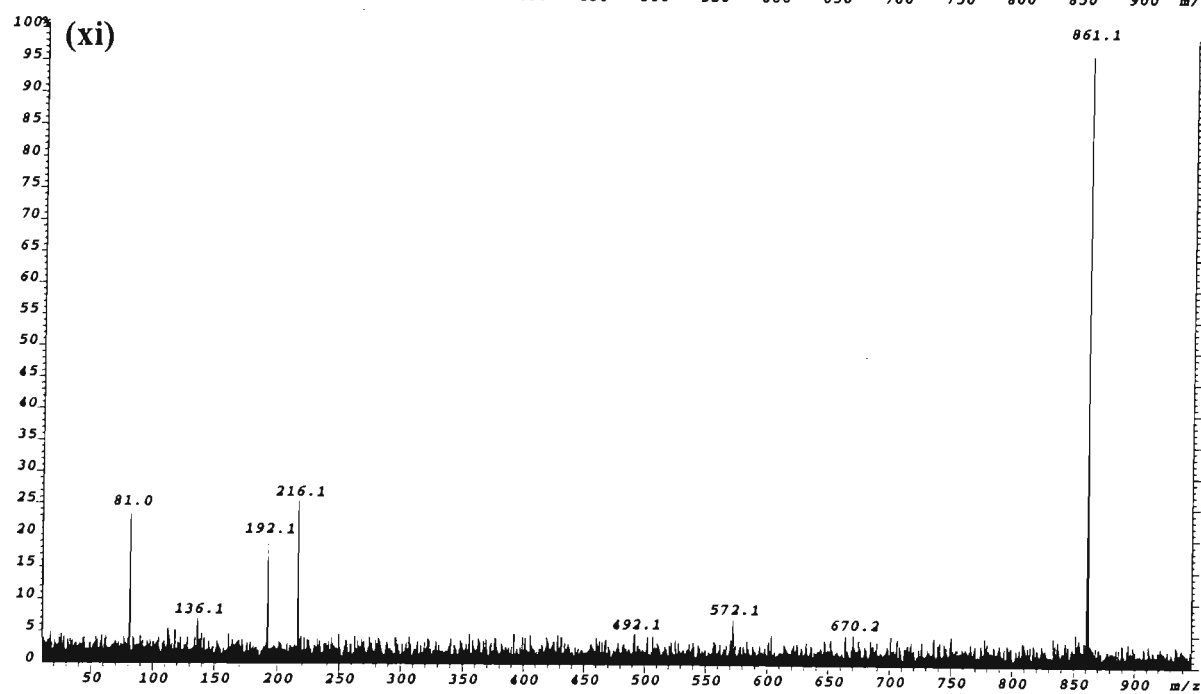
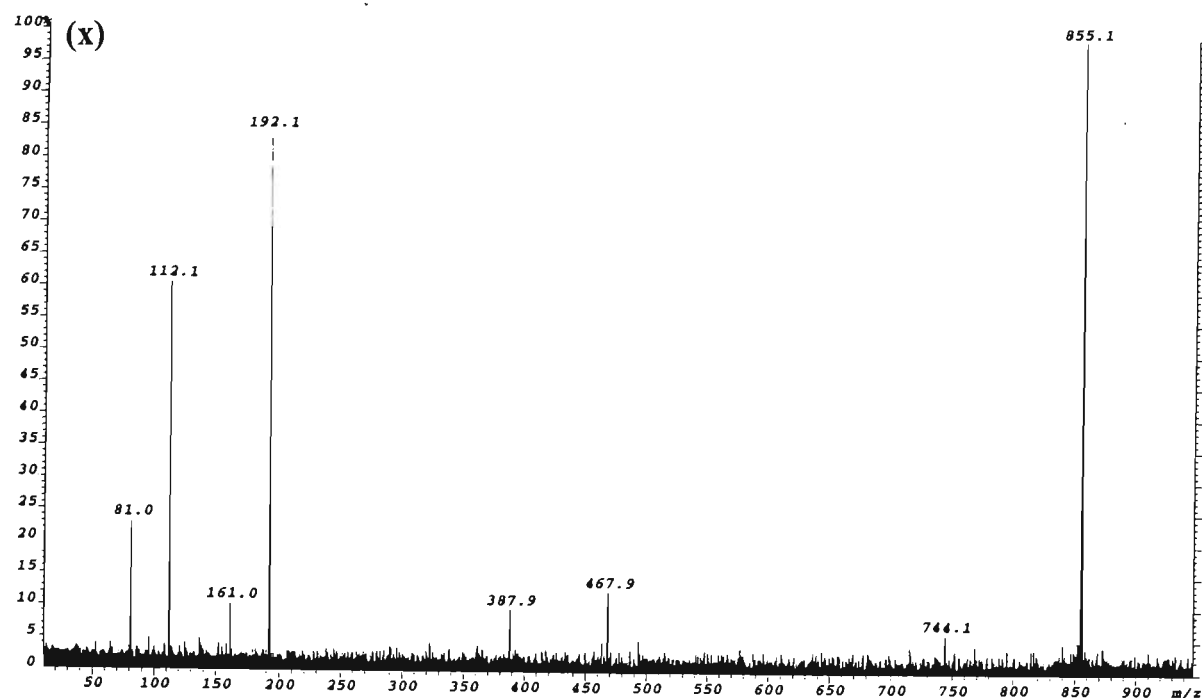
B_{nt} denotes a mononucleotide which may be either psB or sBp

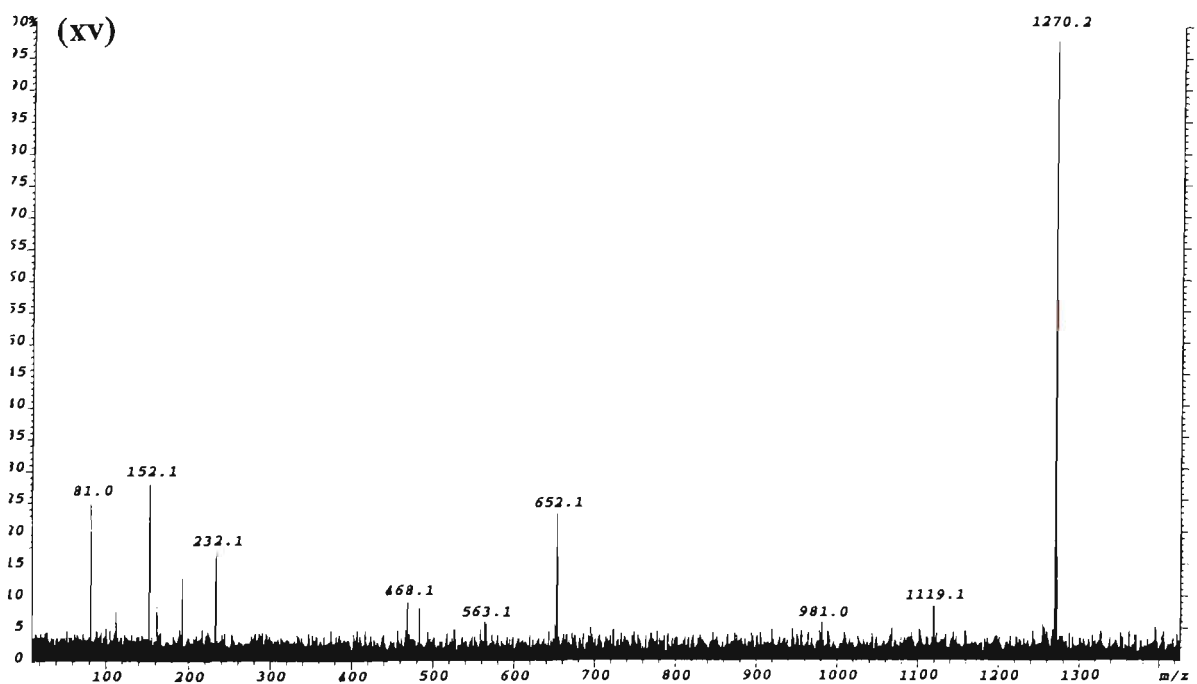
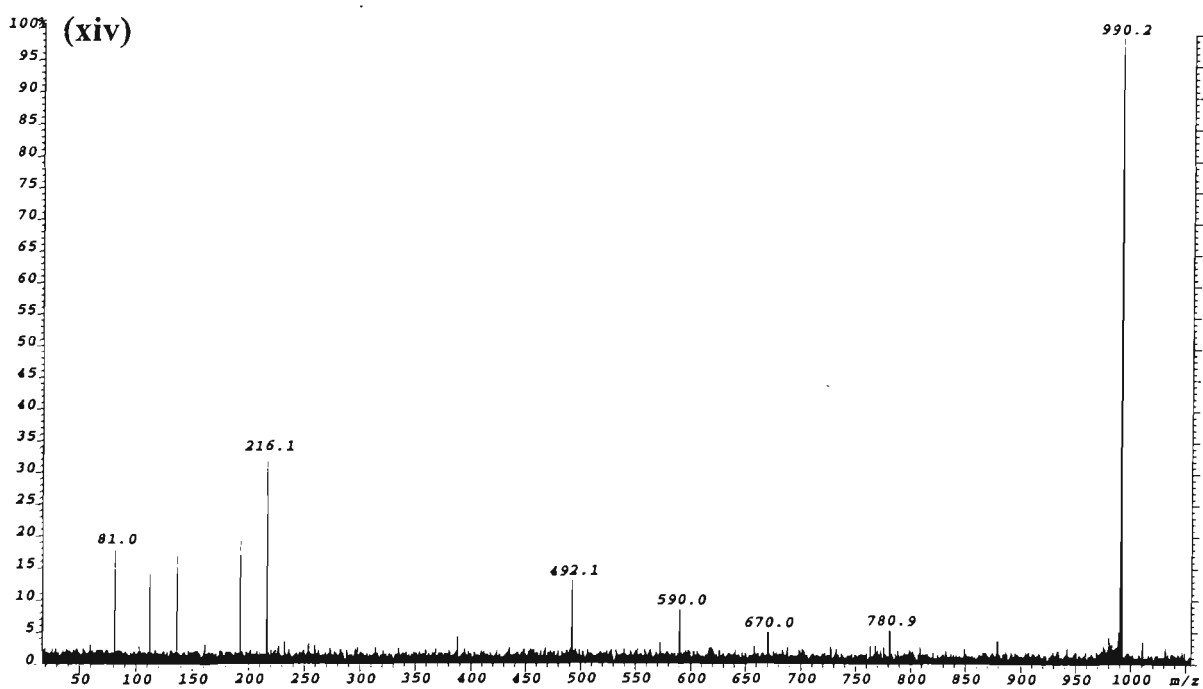
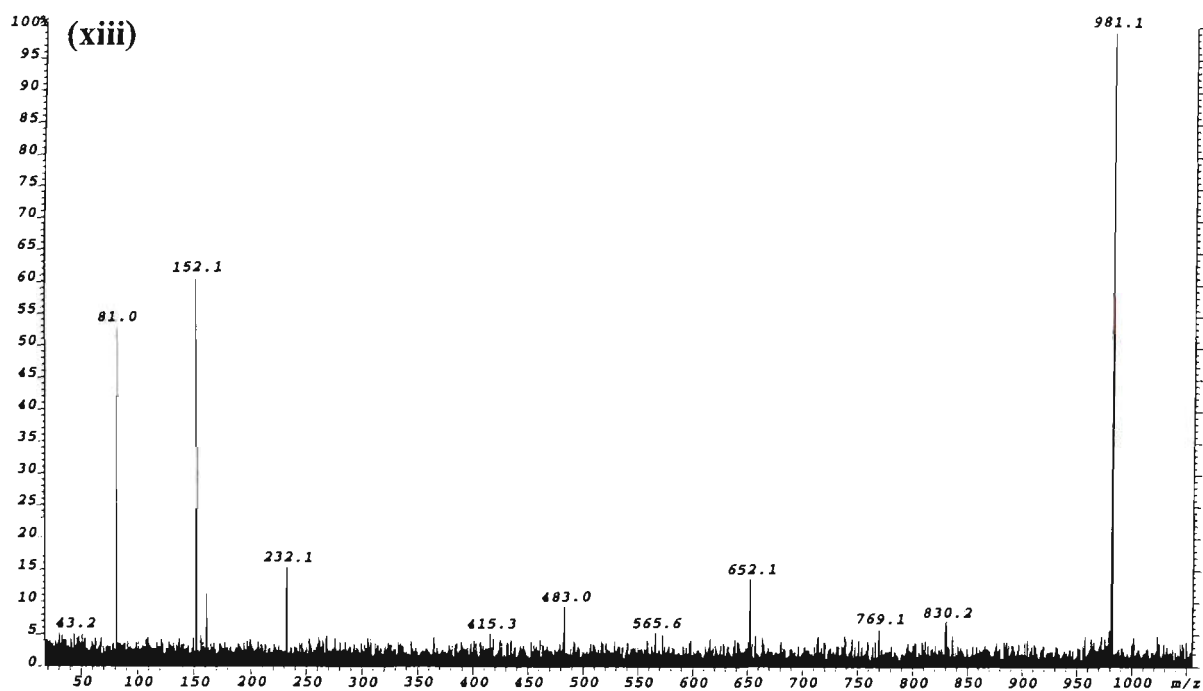
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

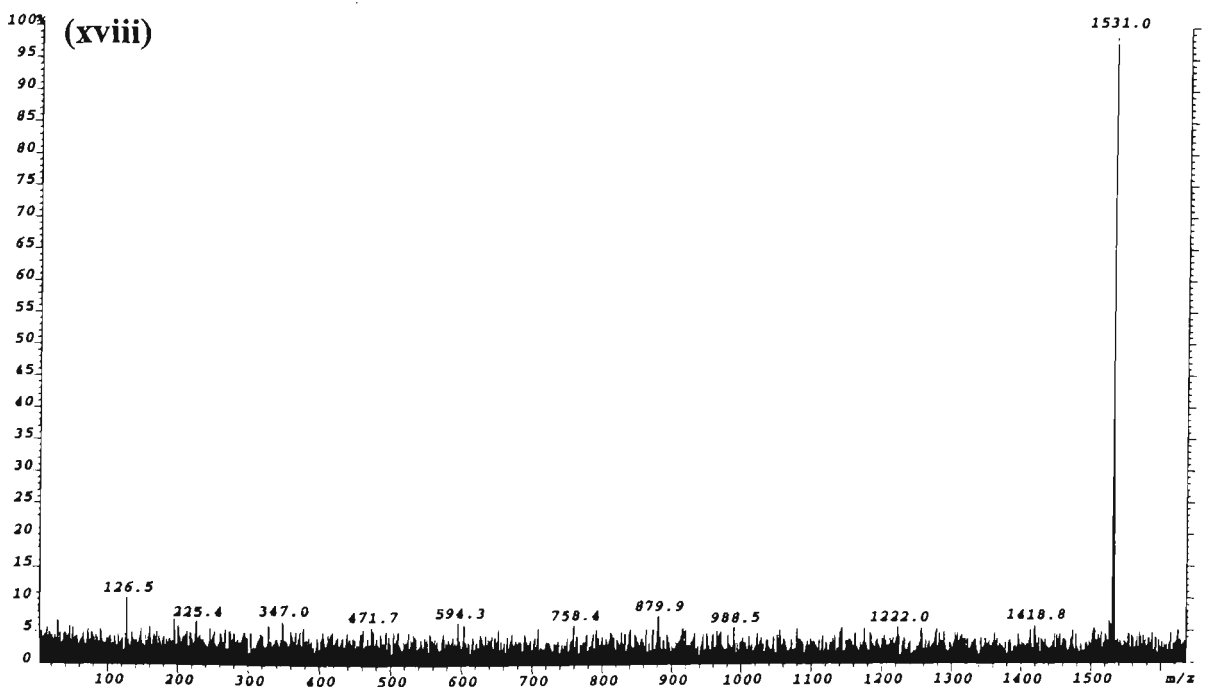
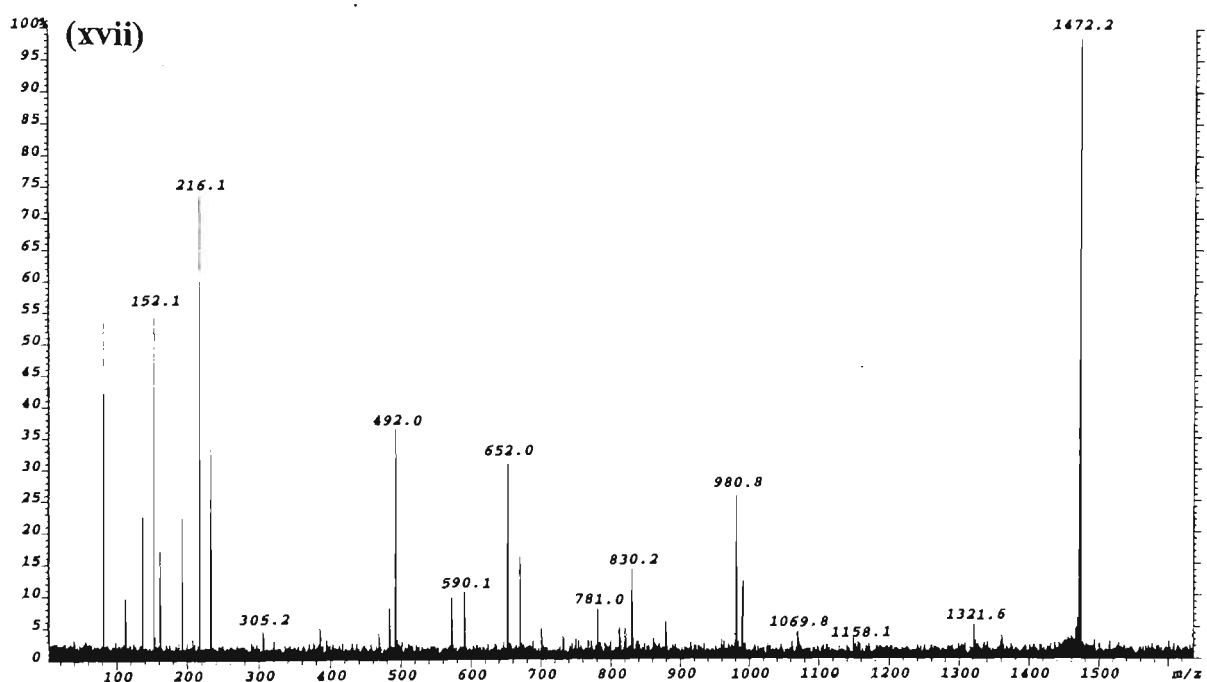
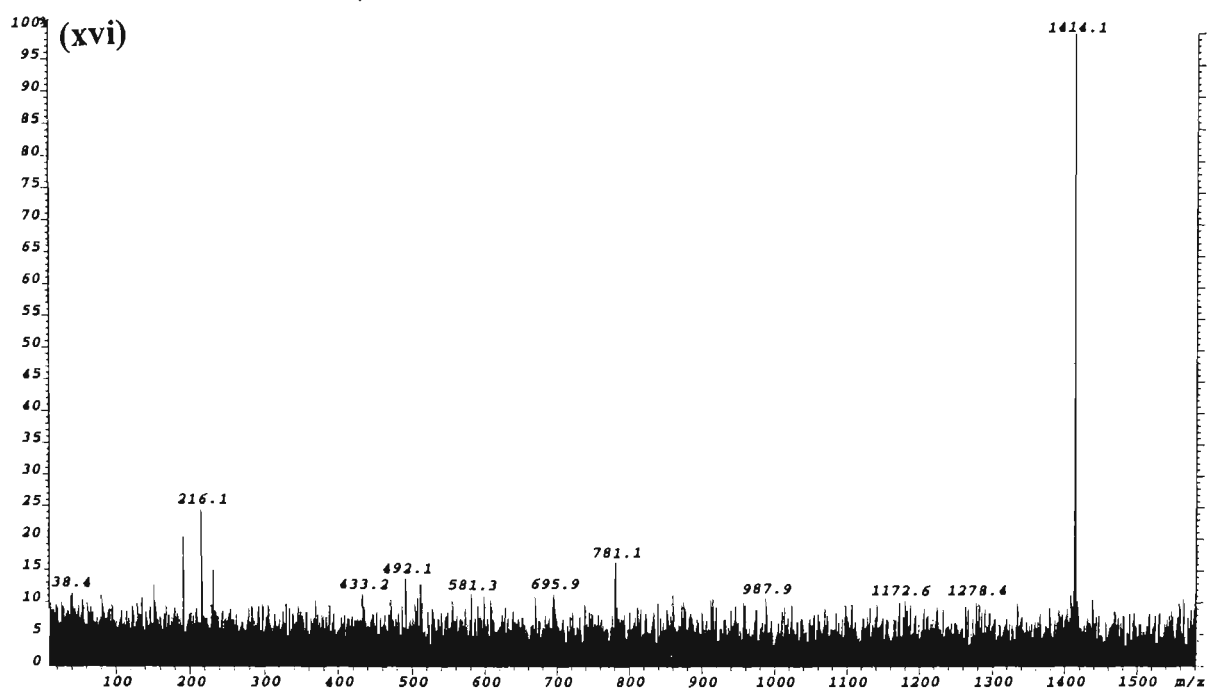


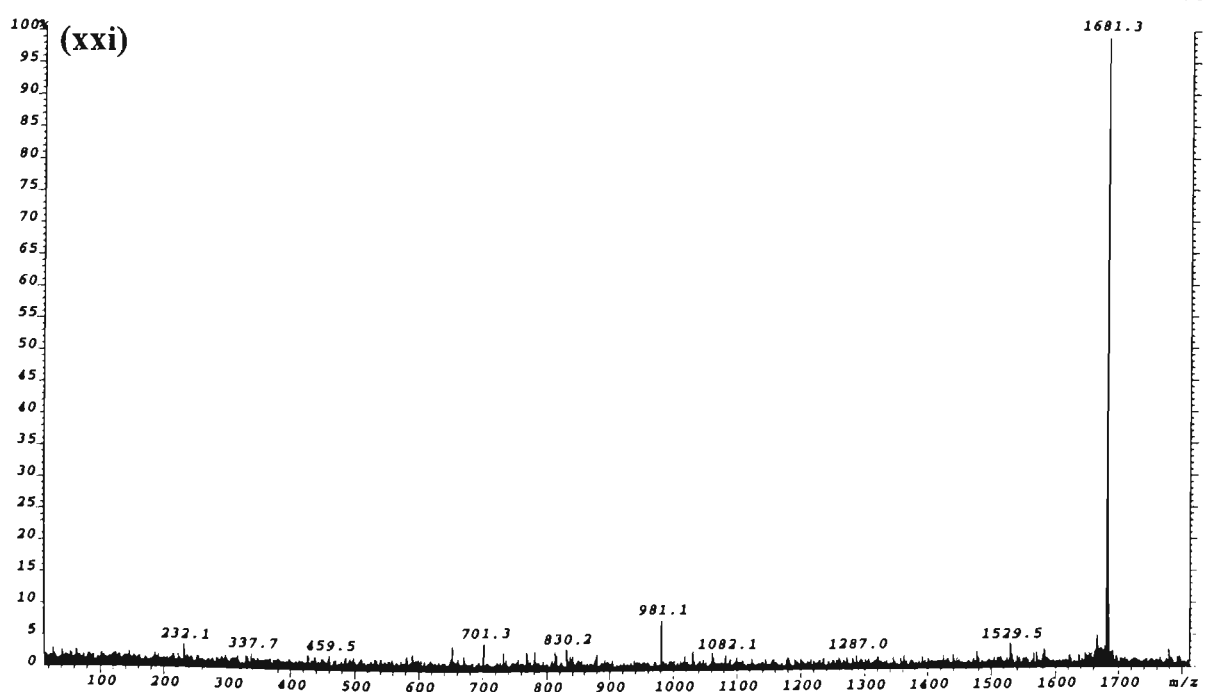
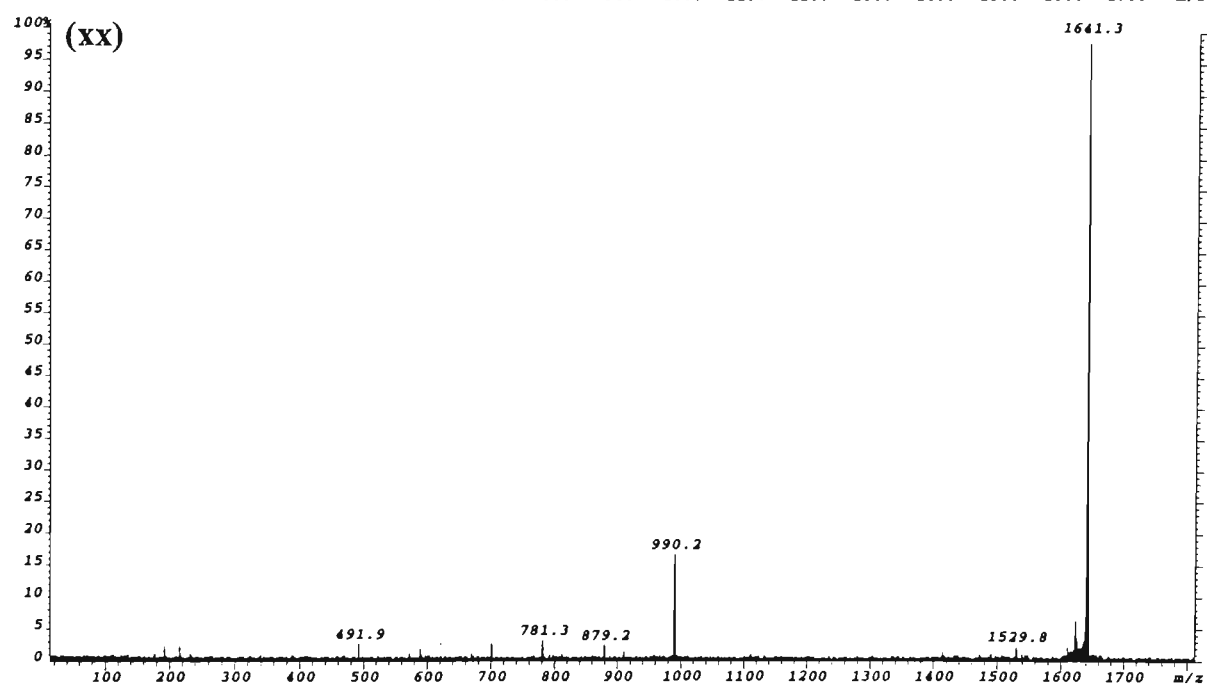
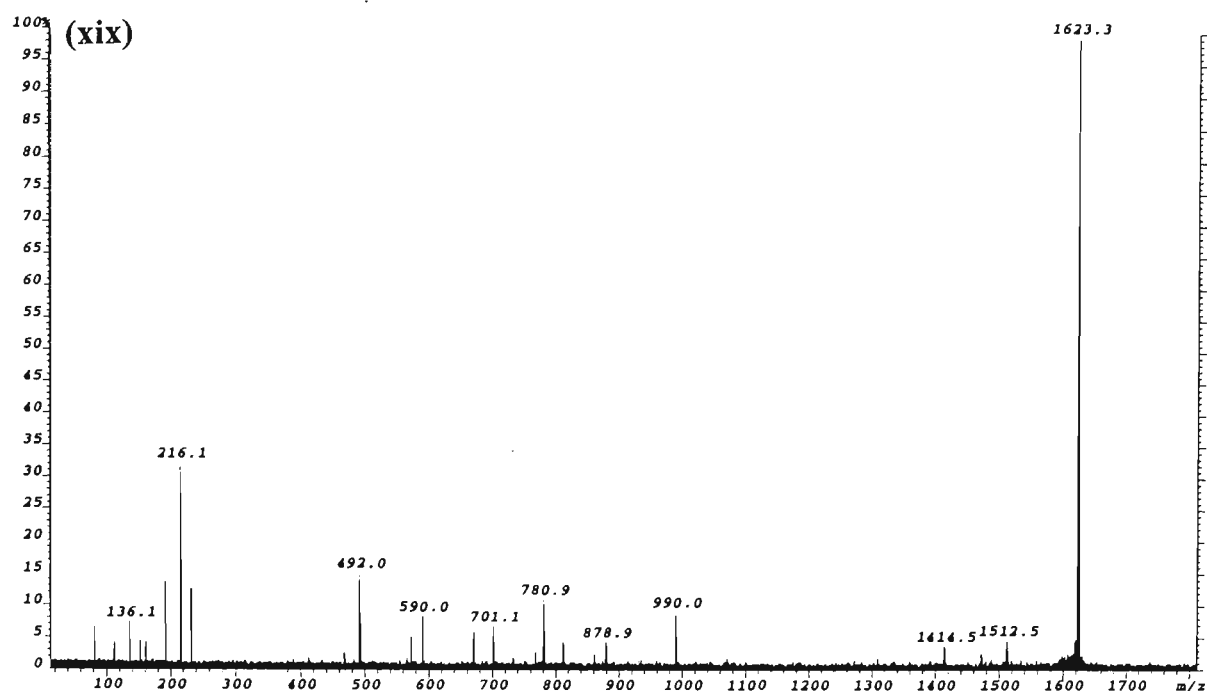


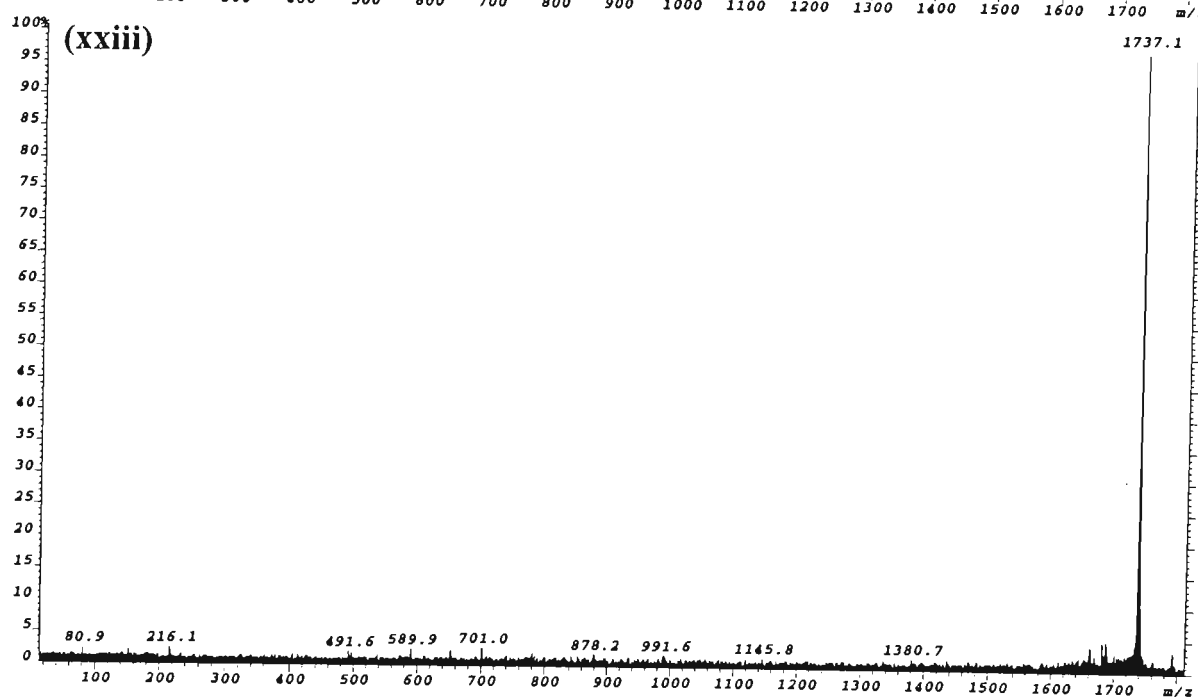
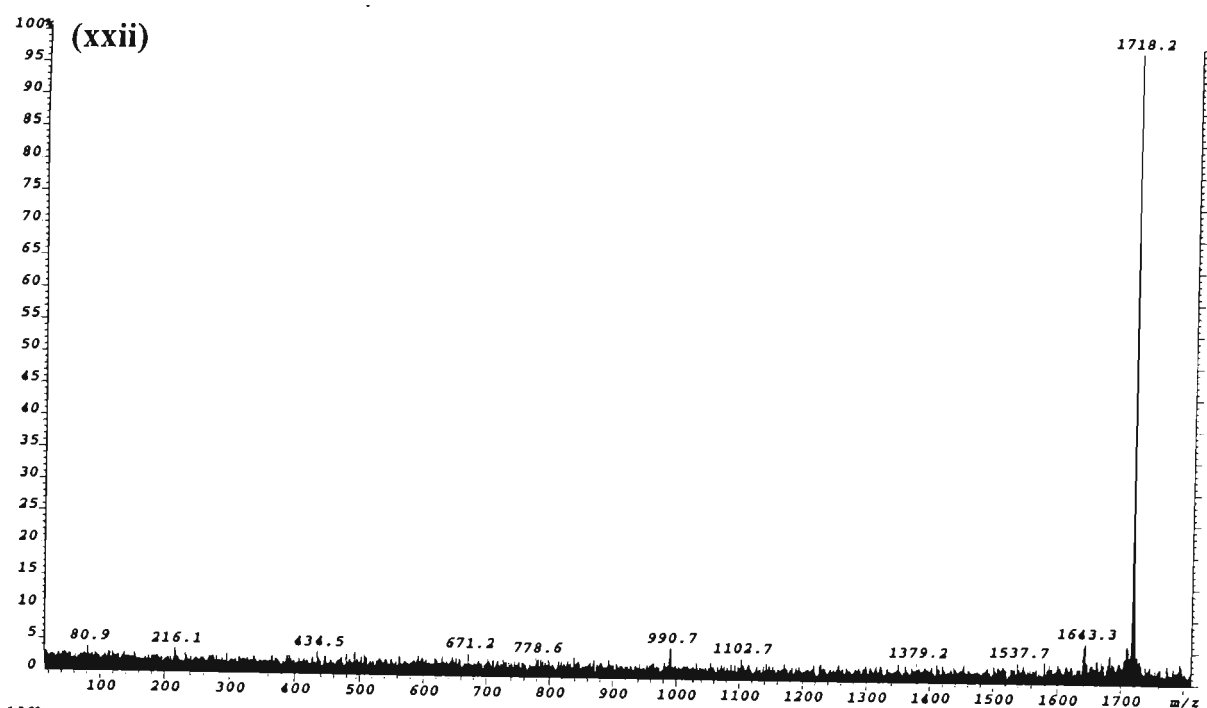










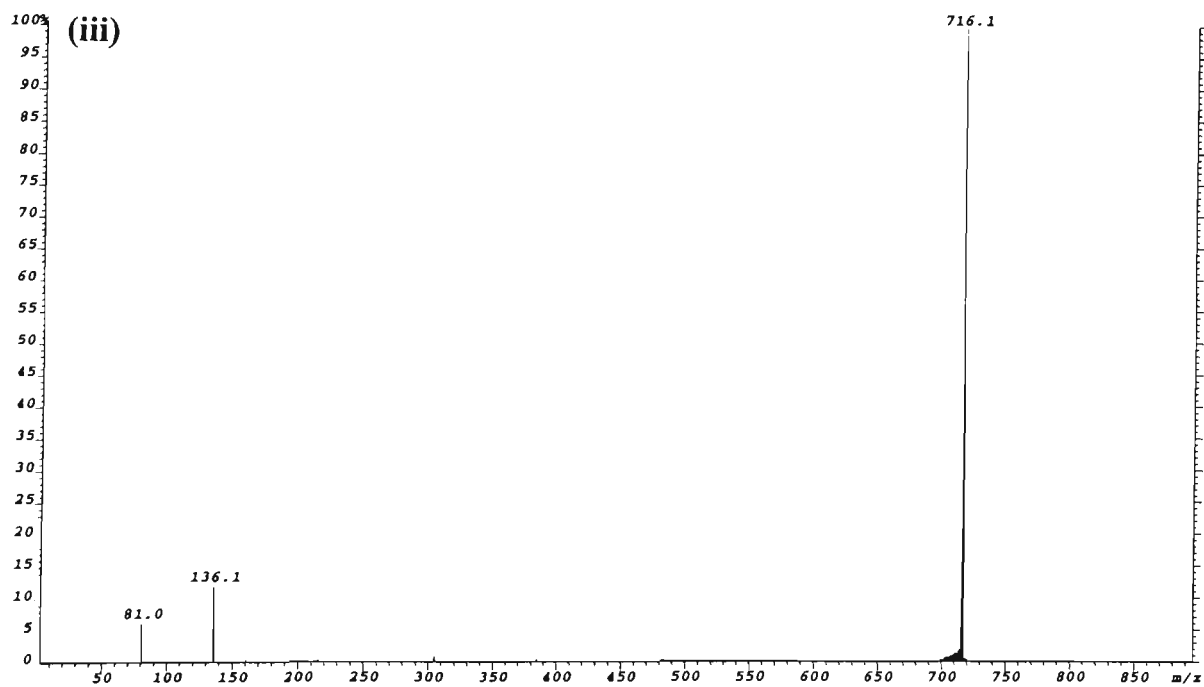
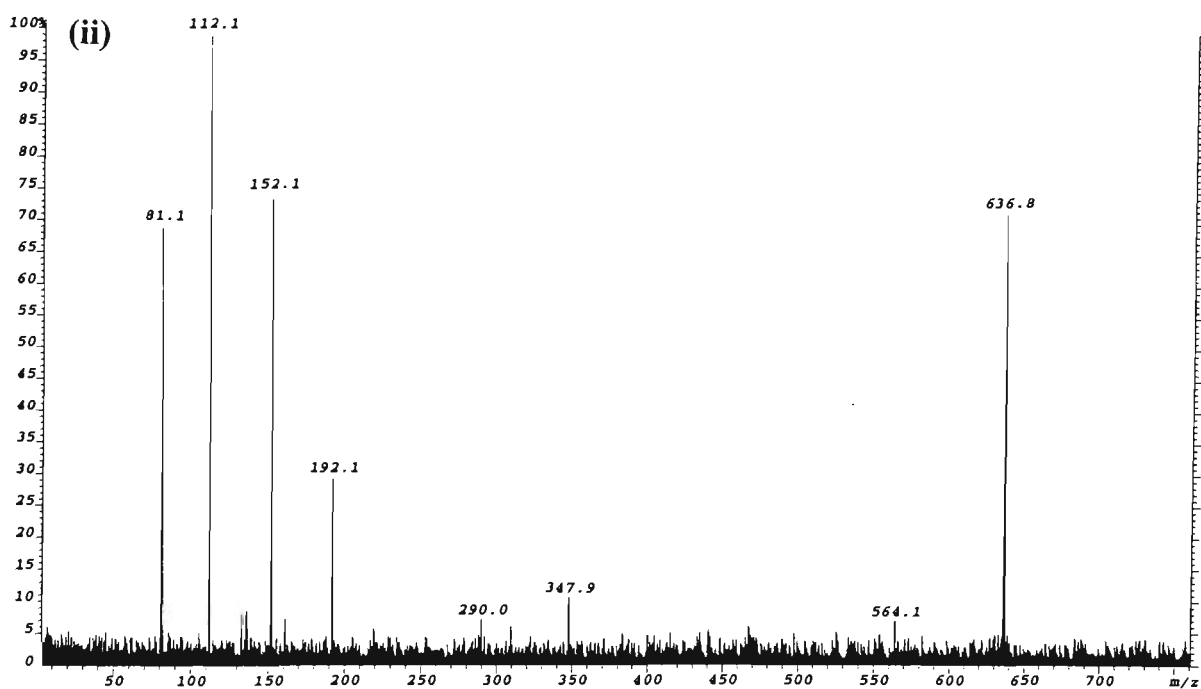
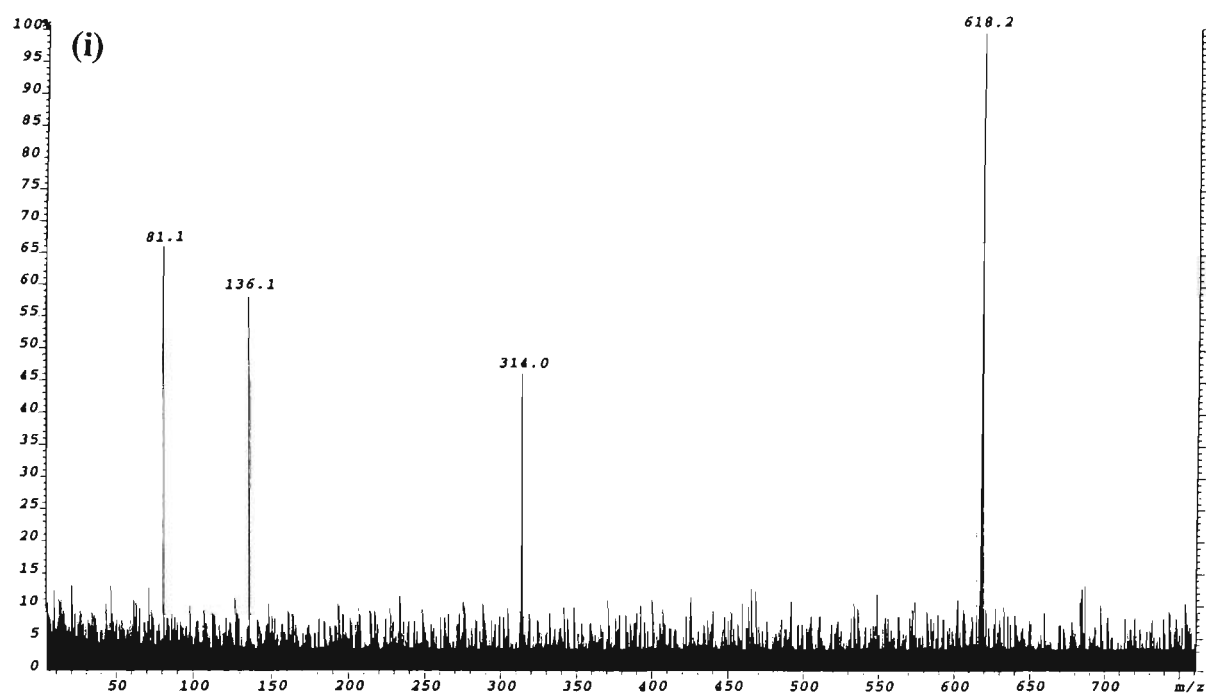


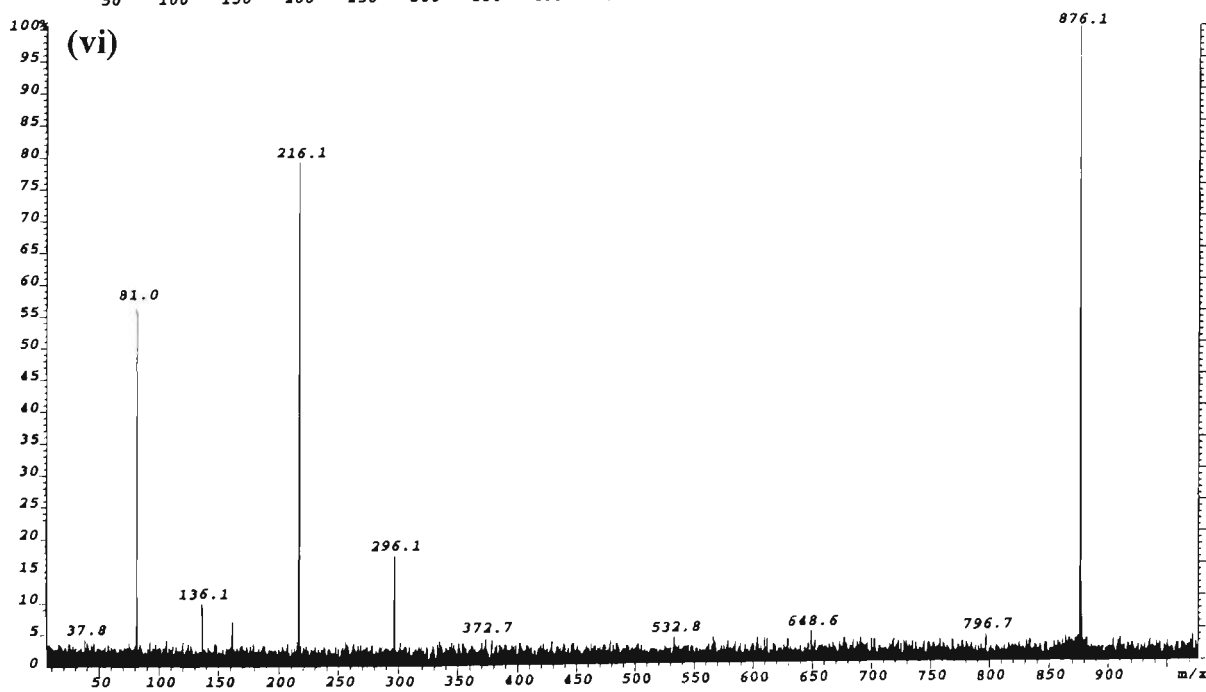
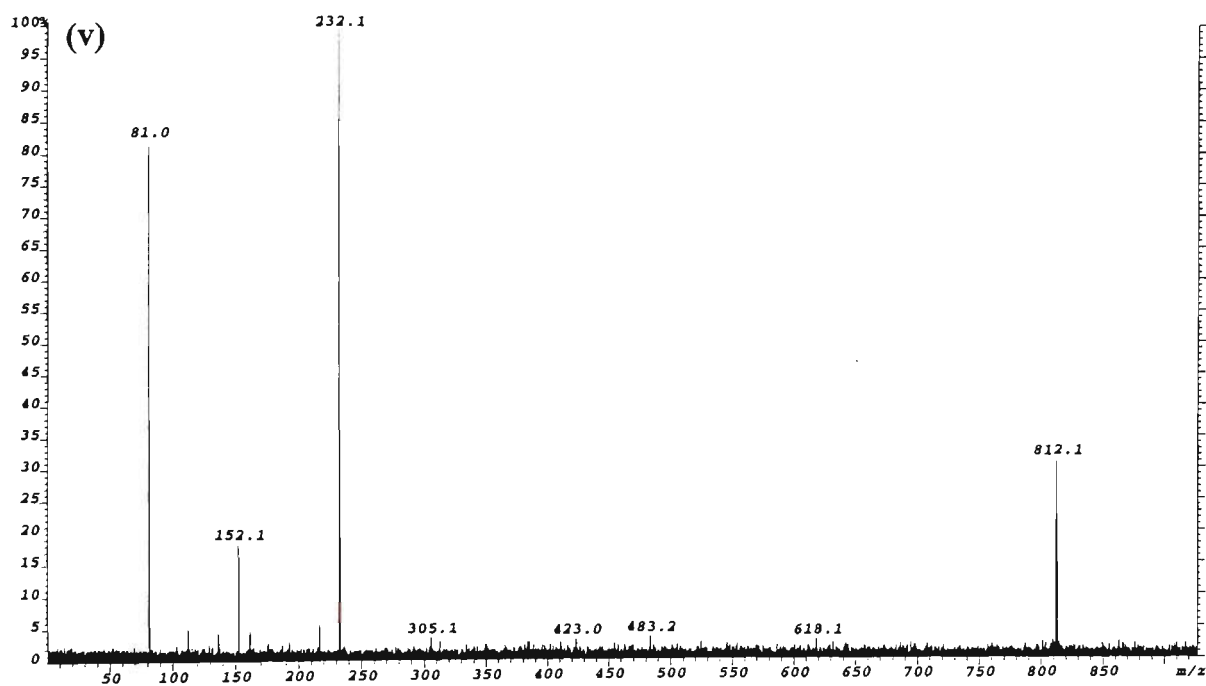
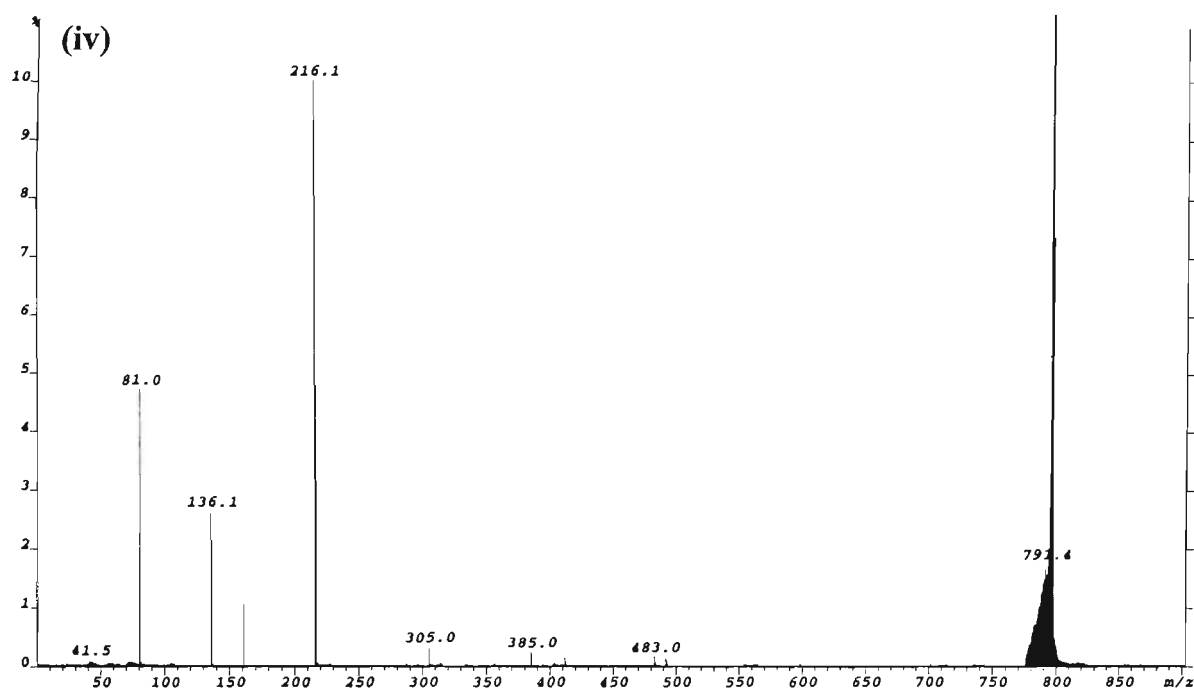
Appendix 4.9

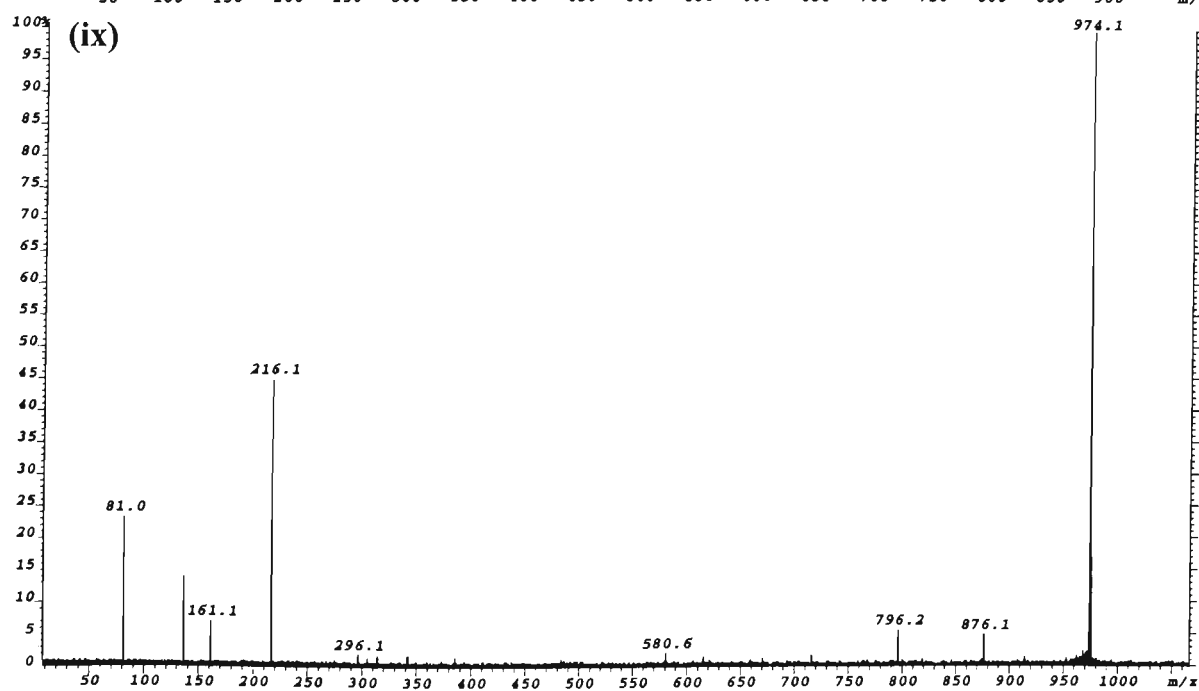
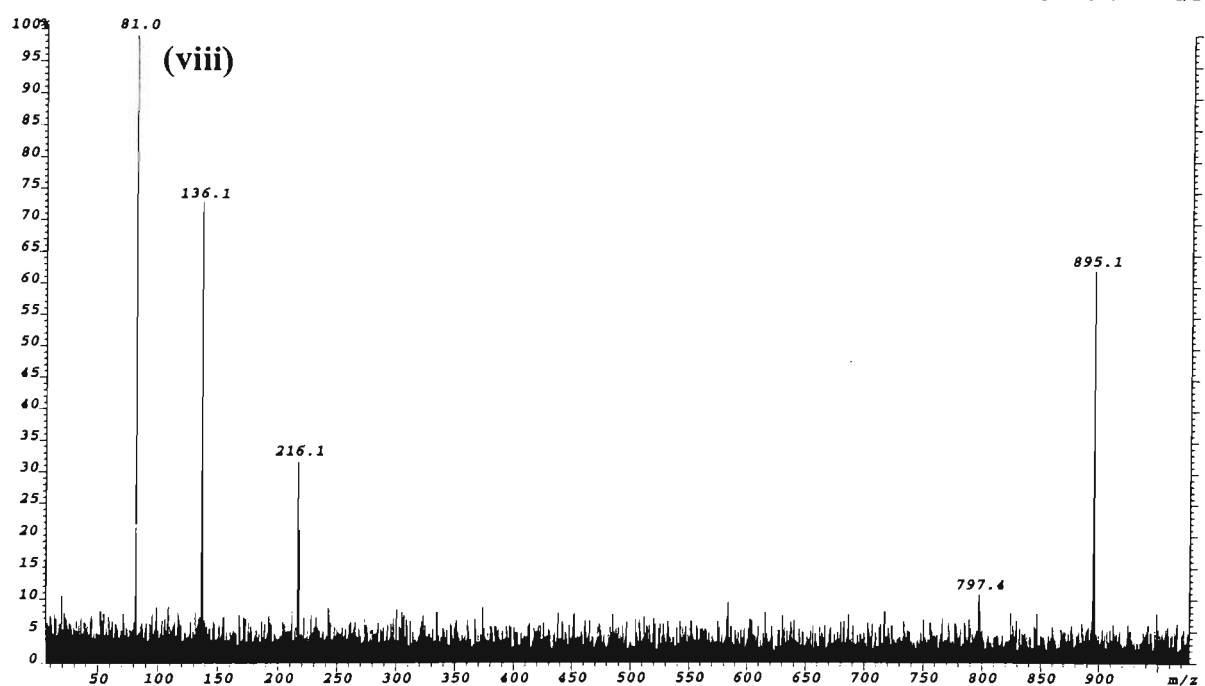
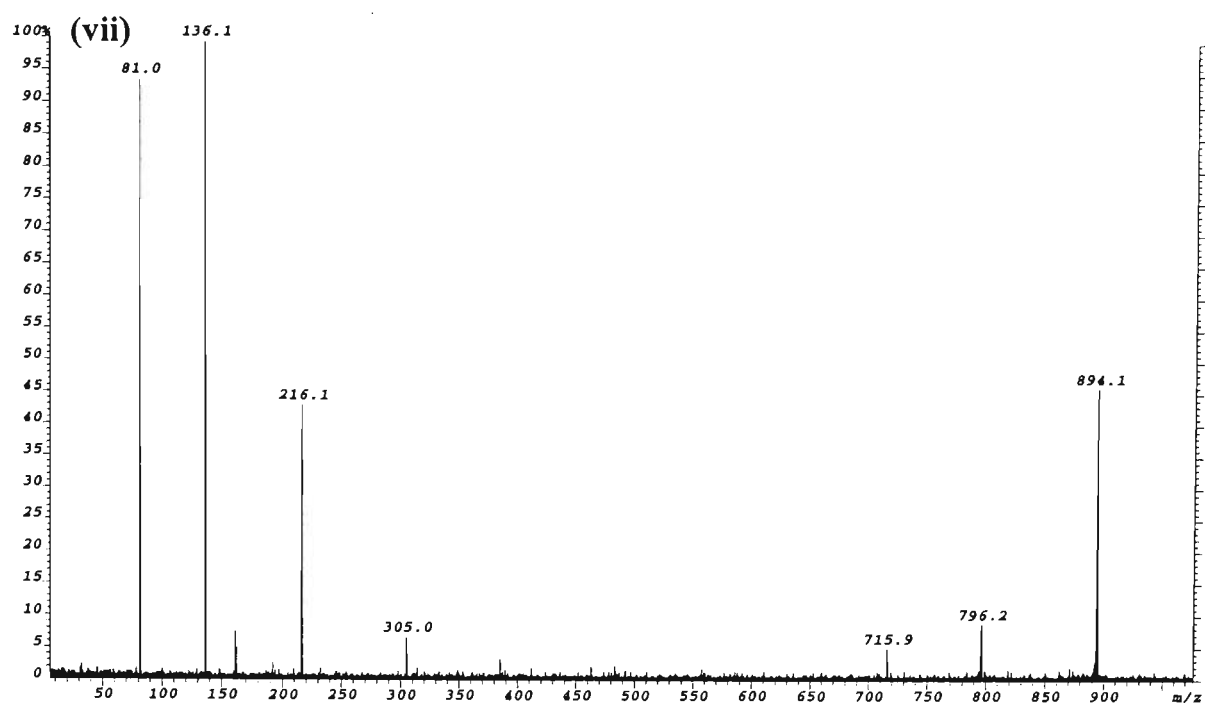
ESI-MS/MS spectra of the ‘source-generated’ polynucleotide product ions observed in the ESI-MS/MS spectrum of the $[M+H]^+$ ion of 5'-d(CGTACG)-3'.

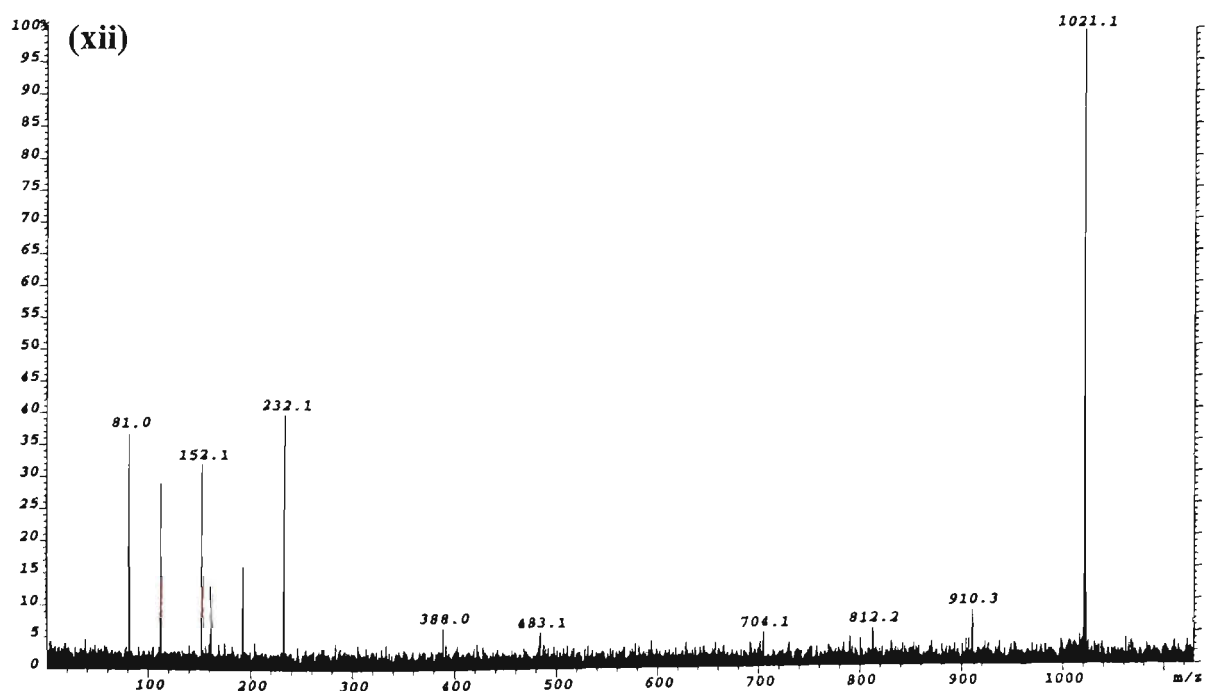
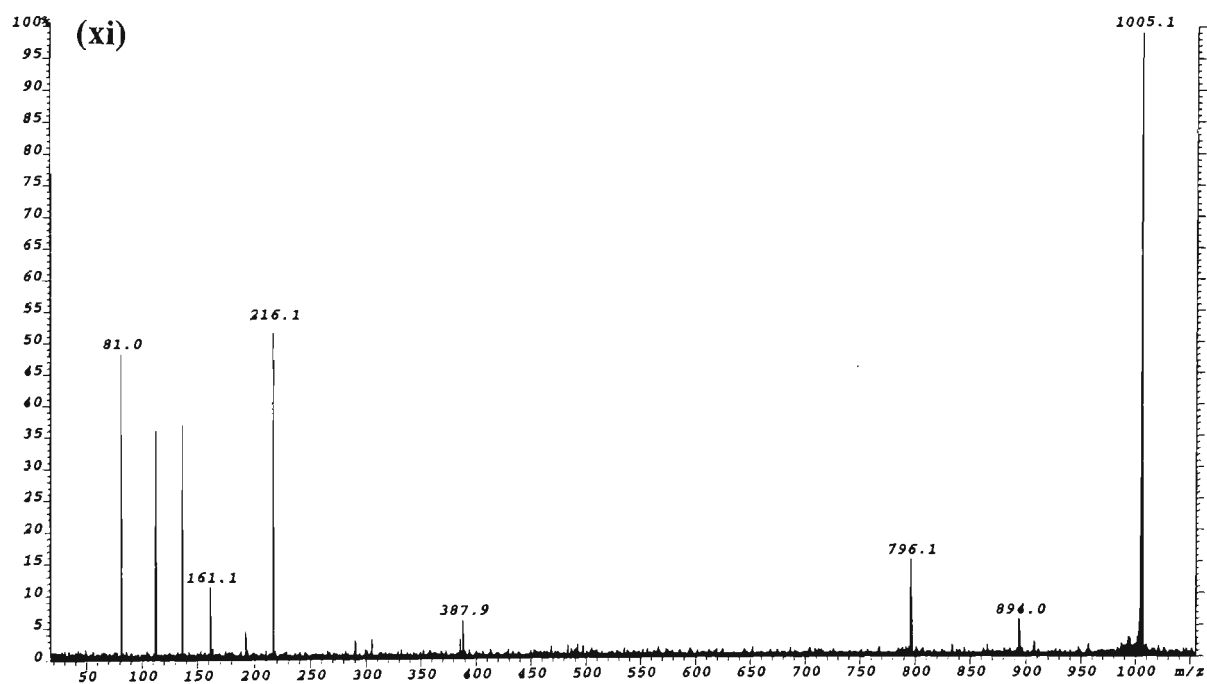
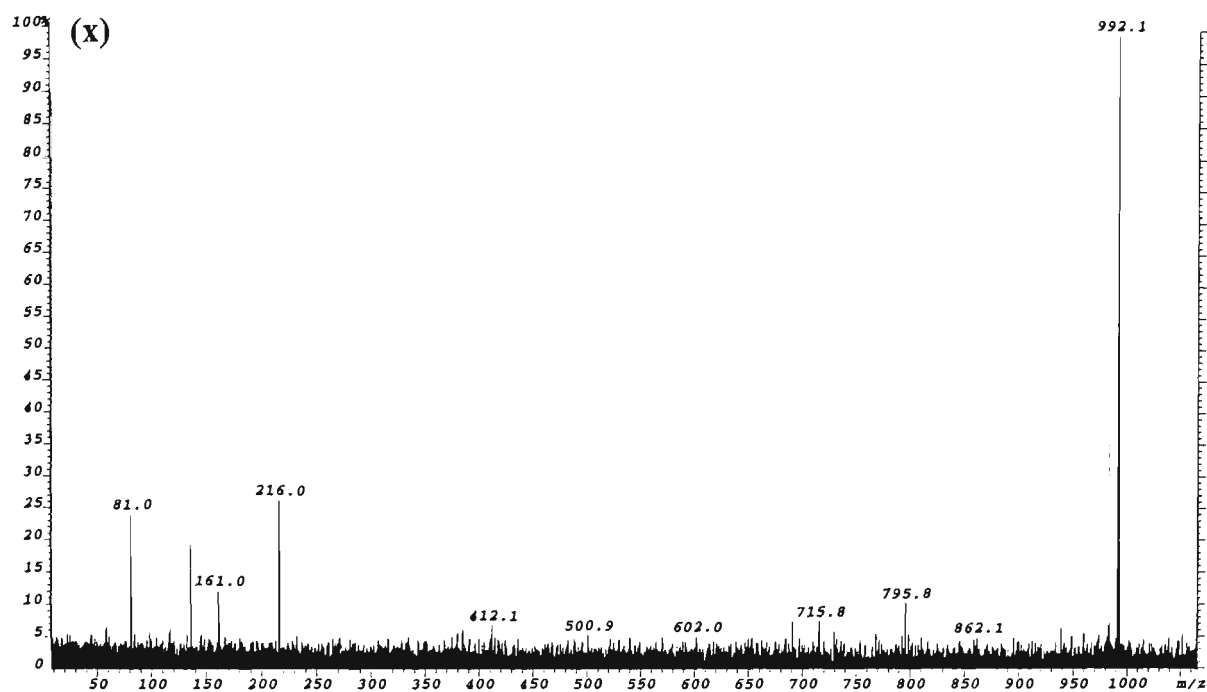
Figure No.	Precursor Ion	
	<i>m/z</i>	Assignment
(i)	618.2	$[(TA)+H]^+$
(ii)	636.8	$[w_2+2H]^+$ and/or $[d_2+2H]^+$
(iii)	716.1	$[(TA)+s+H]^+$
(iv)	796.2	$[(TA)+p+s+H]^+$ and/or $[(TA)+p+H_2O+H]^+$
(v)	812.1	$[c_3-CH+2H]^+$ and/or $[(GT)+p+s+H]^+$
(vi)	876.1	$[(TA)+s+p+s-H_2O+H]^+$ and/or $[(TA)+p+s+p+H]^+$
(vii)	894.1	$[z_4-GH-CH]^+$ and/or $[a_4-GH-CH]^+$ and/or $[(TA)+s+p+s+H]^+$ and/or $[(TA)+p+s+p+H_2O+H]^+$ and/or $[(GT)+p+s+p+H]^+$
(viii)	895.1	$[z_4-GH-CH]^+$ and/or $[a_4-GH-CH]^+$ and/or $[(TA)+s+p+s+H]^+$ and/or $[(TA)+p+s+p+H_2O+H]^+$
(ix)	974.1	$[x_4-GH-CH]^+$ and/or $[c_4-GH-CH]^+$ and/or $[(TA)+p+s+p+s+H]^+$
(x)	992.1	$[w_4-GH-CH+2H]^+$ and/or $[d_4-GH-CH+2H]^+$ and/or $[(TA)+p+s+p+s+H_2O+H]^+$
(xi)	1005.1	$[a_4-GH]^+$ and/or $[z_4-GH]^+$ and/or $[(TAC)+s+H]^+$ and/or $[(TAC)+p+H_2O+H]^+$
(xii)	1021.1	$[a_4-B_4H]^+$ and/or $[z_4-AH]^+$ and/or $[(GT)+C_{nt}+s+H]^+$
(xiii)	1085.0	$[x_4-GH]^+$ and/or $[c_4-GH]^+$ and/or $[(TAC)+s+p+H]^+$
(xiv)	1103.2	$[w_4-GH+2H]^+$ and/or $[d_4-GH+2H]^+$ and/or $[(TAC)+p+s+H_2O+H]^+$
(xv)	1107.3	$[(GTA)+p+s-H_2O+H]^+$
(xvi)	1106.7	$[(GTA)+p+s-H_2O+H]^+$
(xvii)	1125.2	$[x_4-CH]^+$ and/or $[c_4-CH]^+$ and/or $[(GTA)+p+s+H]^+$
(xviii)	1143.2	$[w_4-CH+2H]^+$ and/or $[d_4-CH+2H]^+$ and/or $[(GTA)+p+s+H_2O+H]^+$
(xix)	1183.3	$[z_5-GH-GH]^+$ and/or $[a_5-CH-GH]^+$ and/or $[(TAC)+p+s+p+H_2O+H]^+$
(xx)	1223.3	$[z_5-CH-GH]^+$ and/or $[a_5-CH-CH]^+$ and/or $[(GTA)+p+s+p+H_2O+H]^+$
(xxi)	1254.4	$[w_4+2H]^+$ and/or $[d_4+2H]^+$ and/or $[(GTAC)+H_2O+H]^+$
(xxii)	1281.3	$[w_5-GH-GH+2H]^+$ and/or $[d_5-CH-GH+2H]^+$
(xxiii)	1334.2	$[a_5-B_5H]^+$ and/or $[z_5-GH]^+$ and/or $[(GTAC)+p+H_2O+H]^+$
(xxiv)	1414.2	$[c_5-CH]^+$ and/or $[x_5-GH]^+$
(xxv)	1432.2	$[w_5-GH+2H]^+$ and/or $[d_5-CH+2H]^+$
(xxvi)	1472.4	$[M-GH-GH-H_2O+2H]^+$ and/or $[w_5-CH+2H]^+$
(xxvii)	1531.0	$[M-GH-CH+2H]^+$
(xxviii)	1584.1	$[w_5+H]^+$

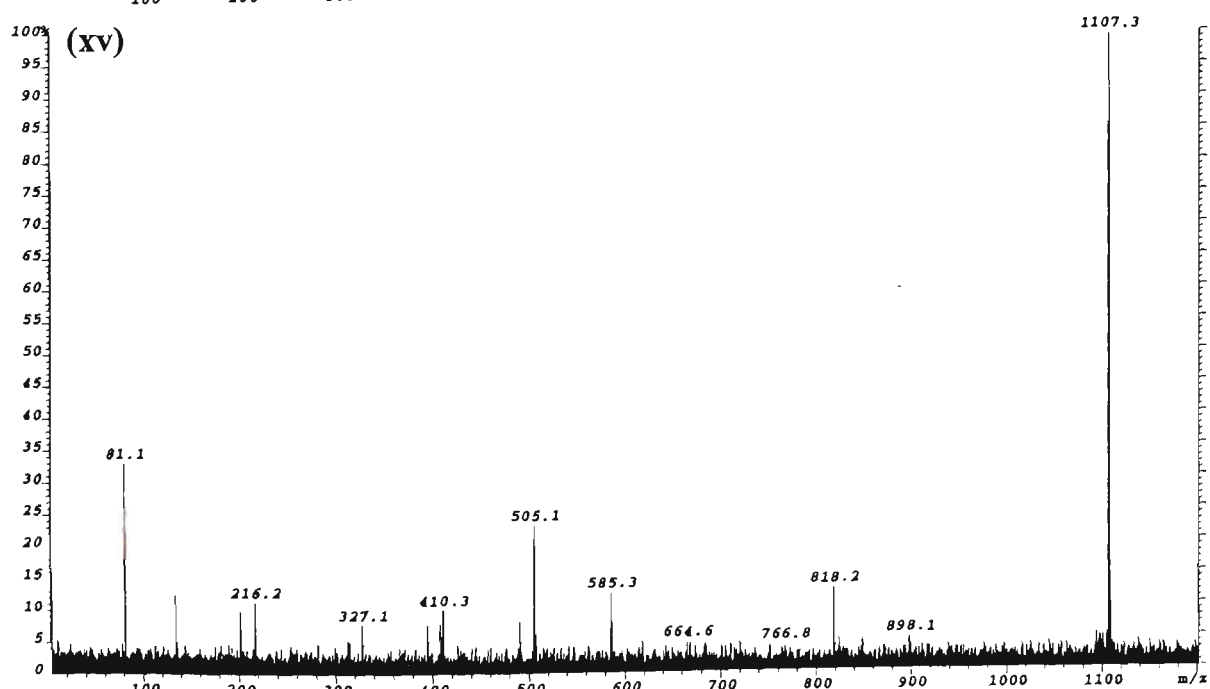
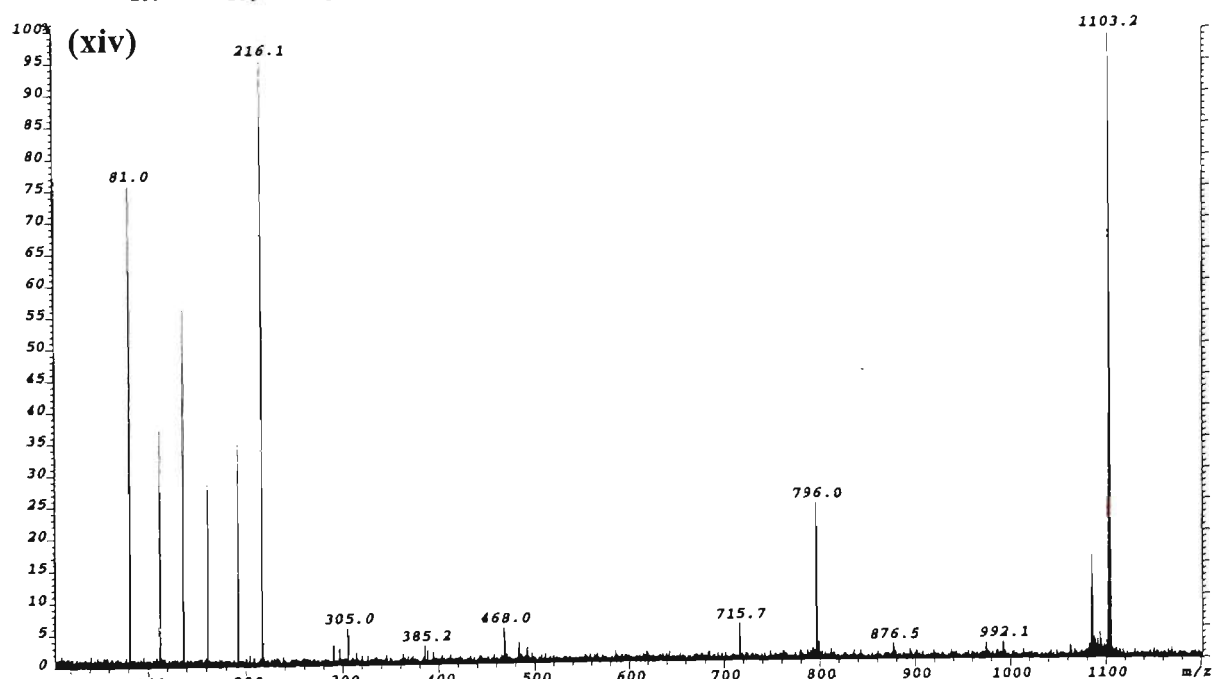
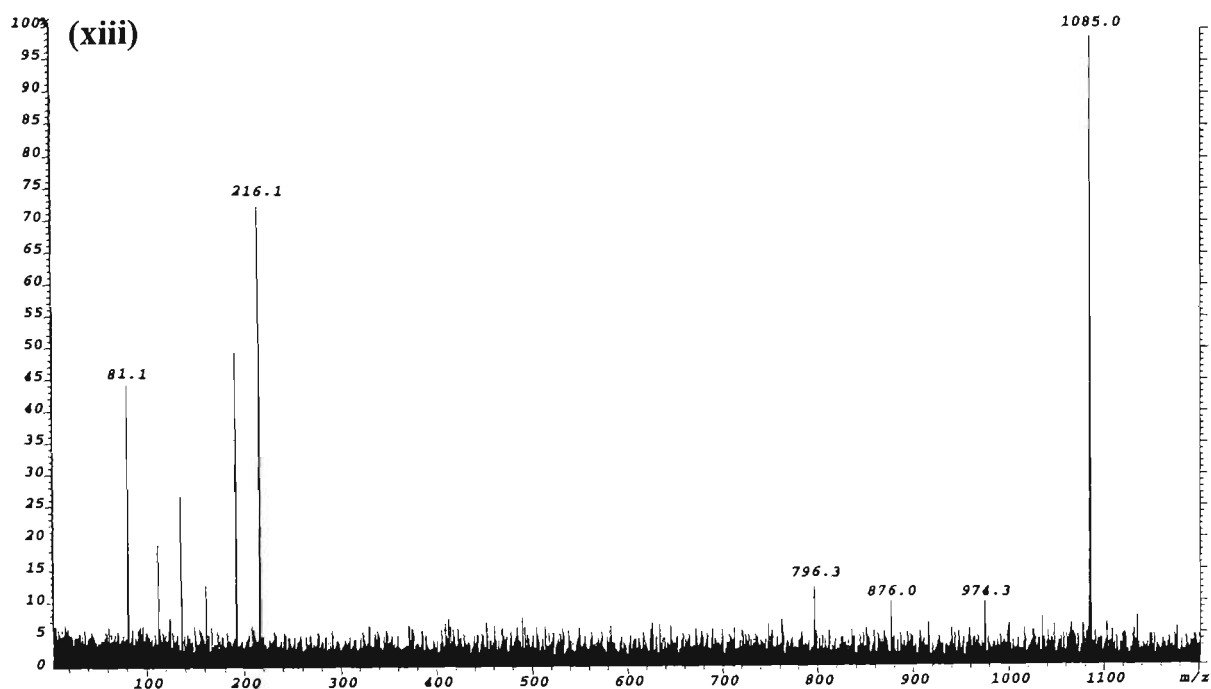
s = deoxyribose-H₂O (C₅H₆O₂), *M_r* = 98.0368 Da
p = PO₃H, *M_r* = 79.9663 Da
B_{nt} denotes a mononucleotide which may be either psB or sBp
[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

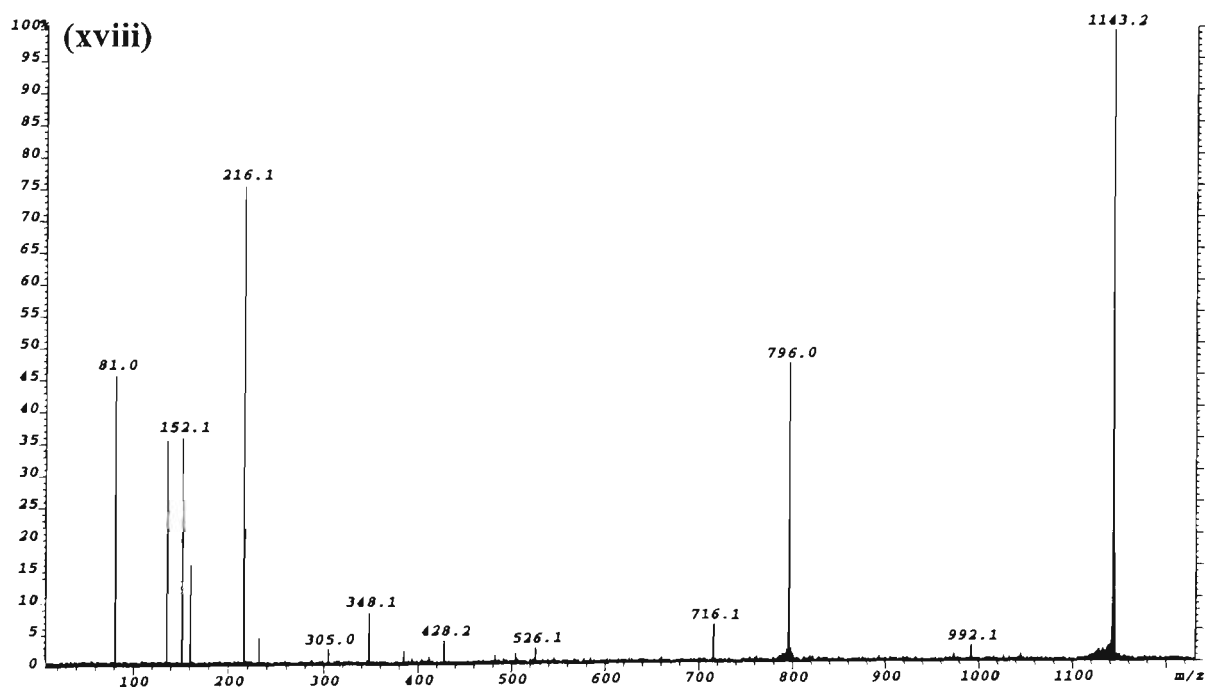
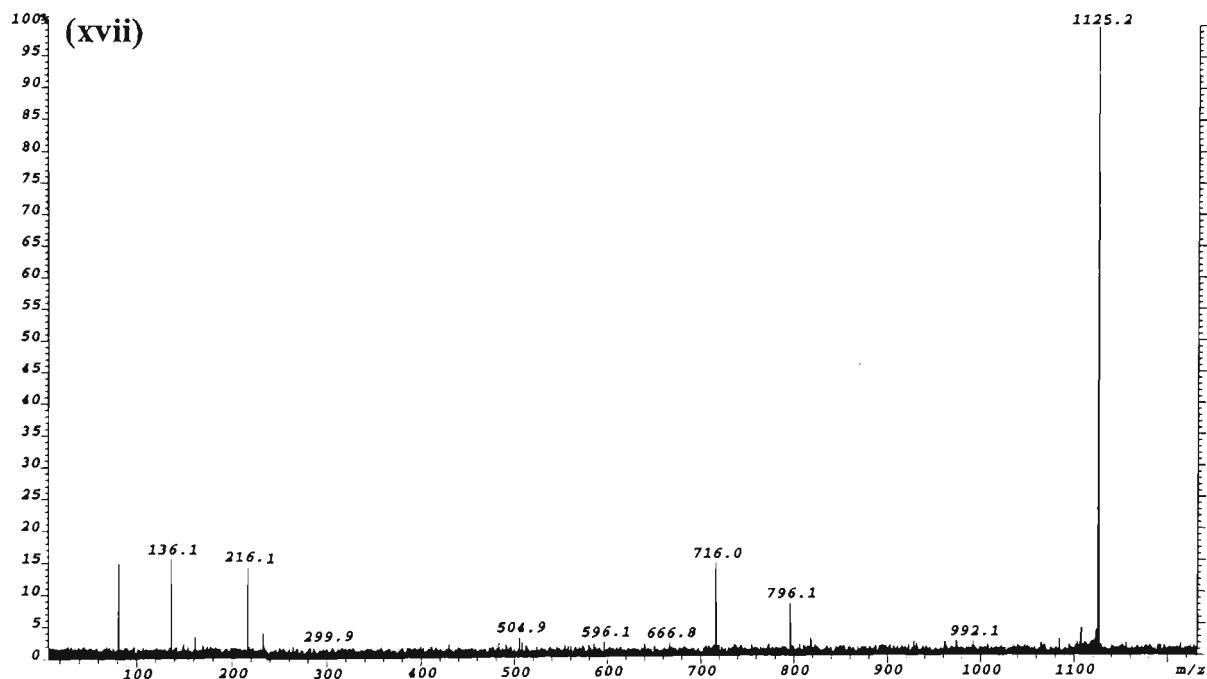
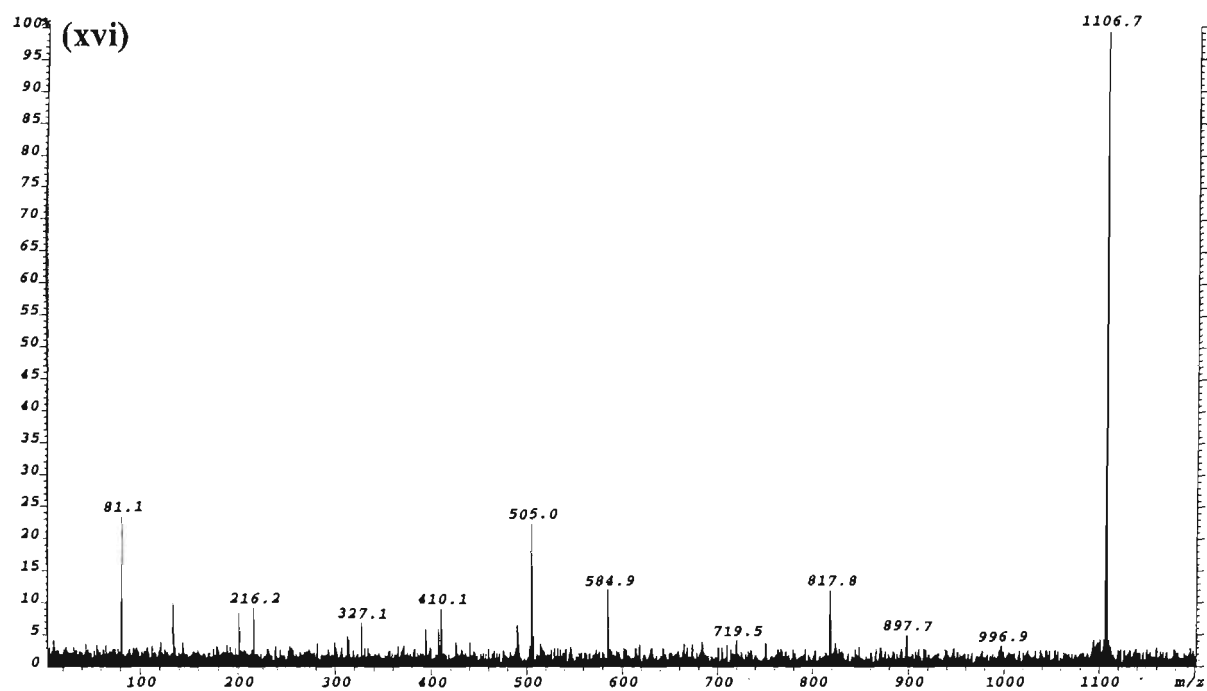


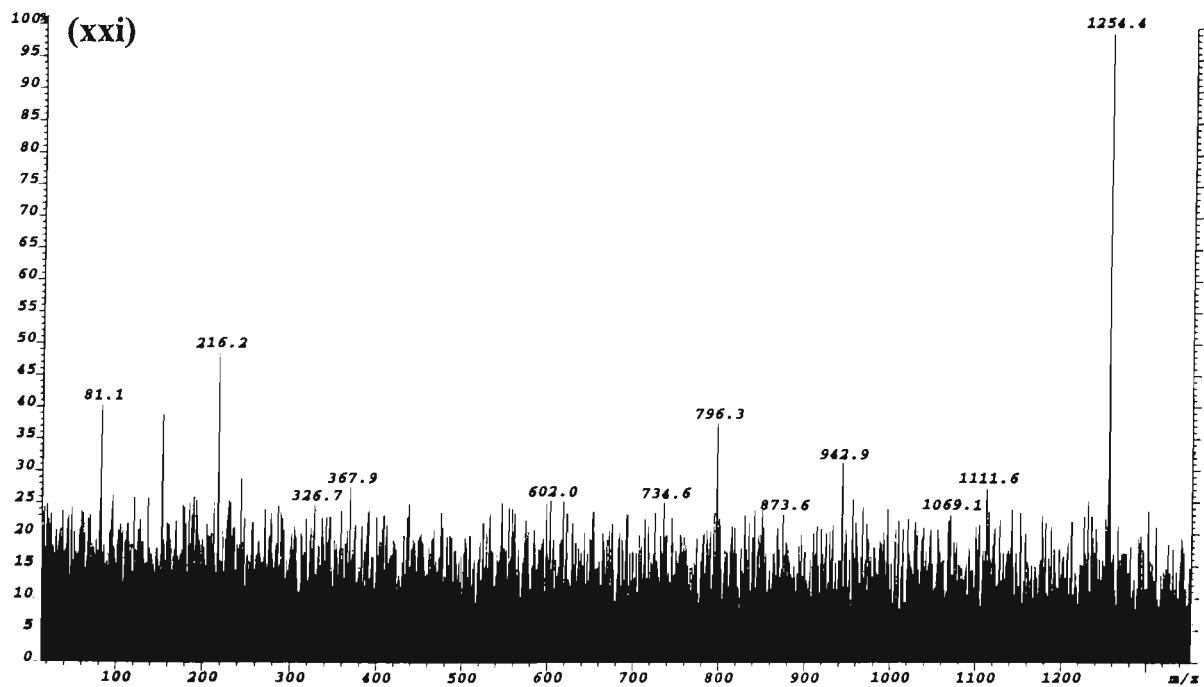
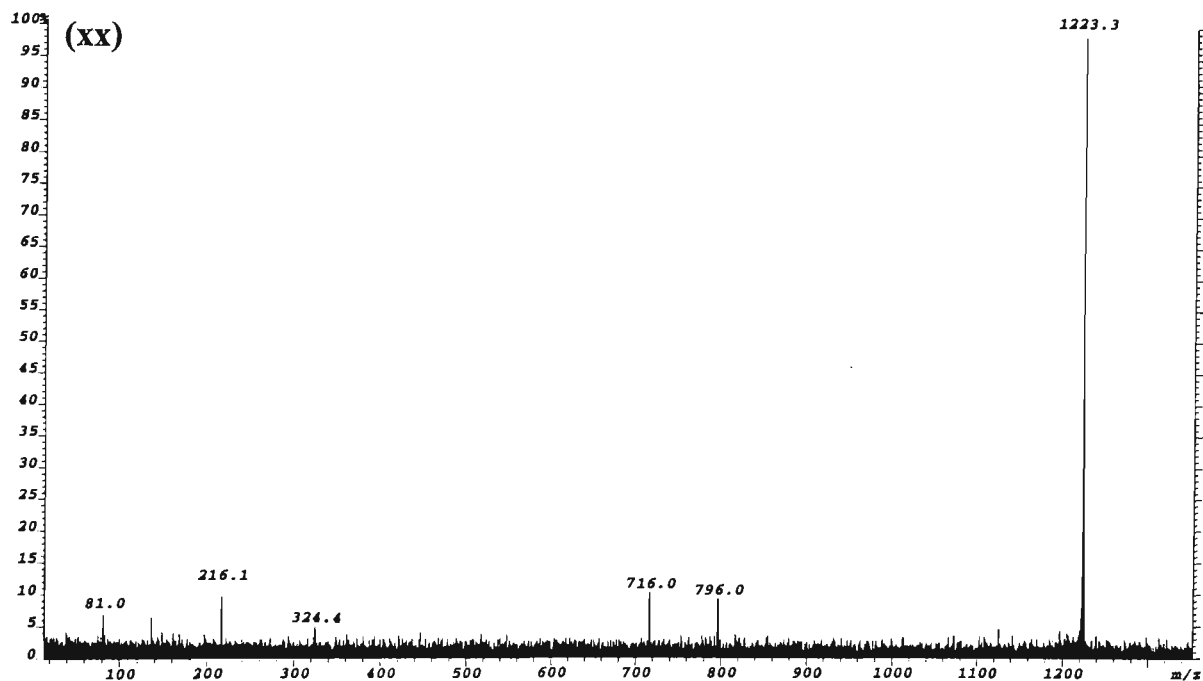
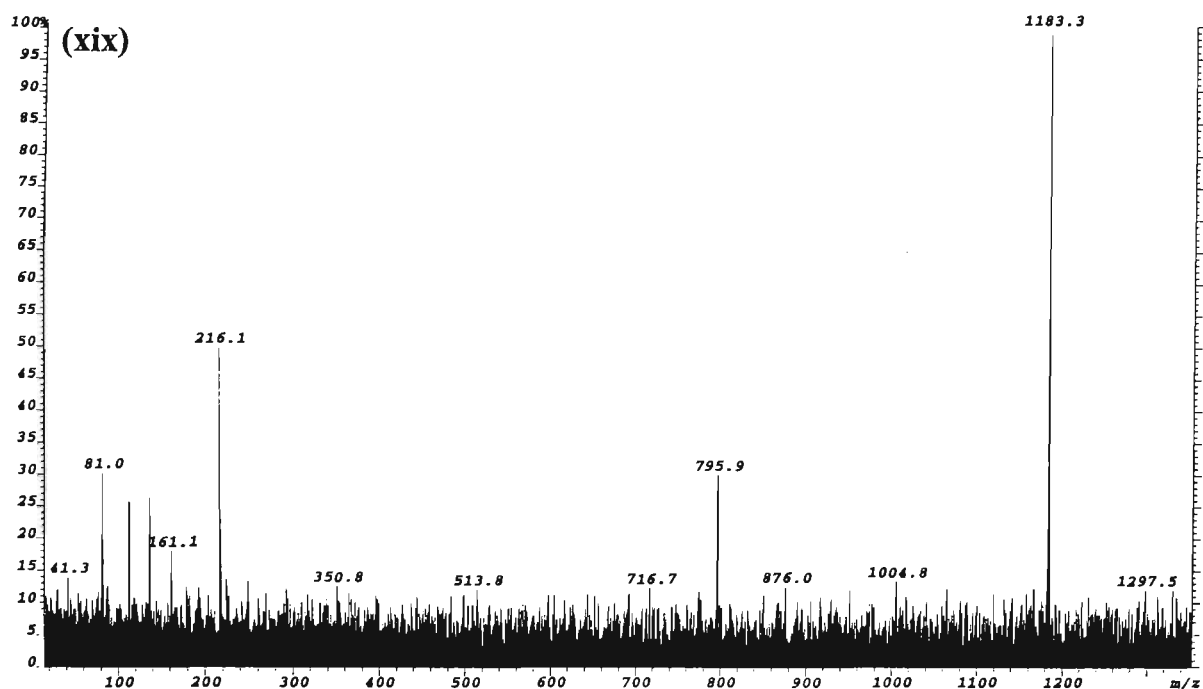


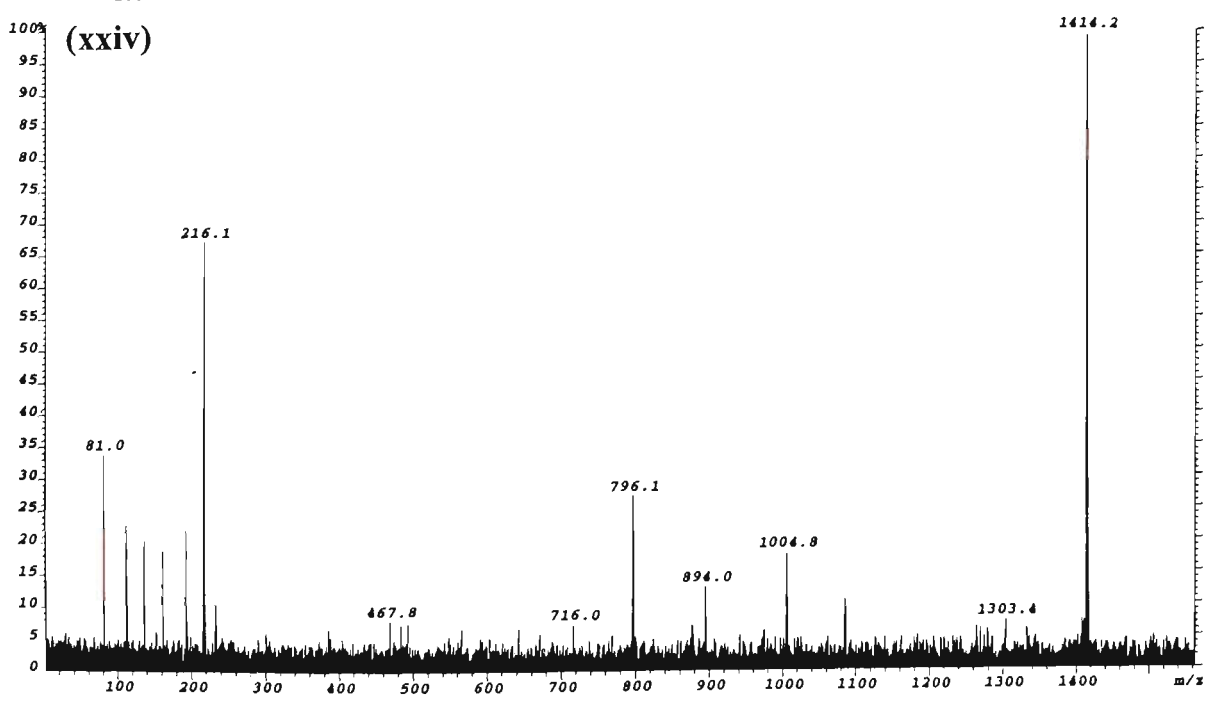
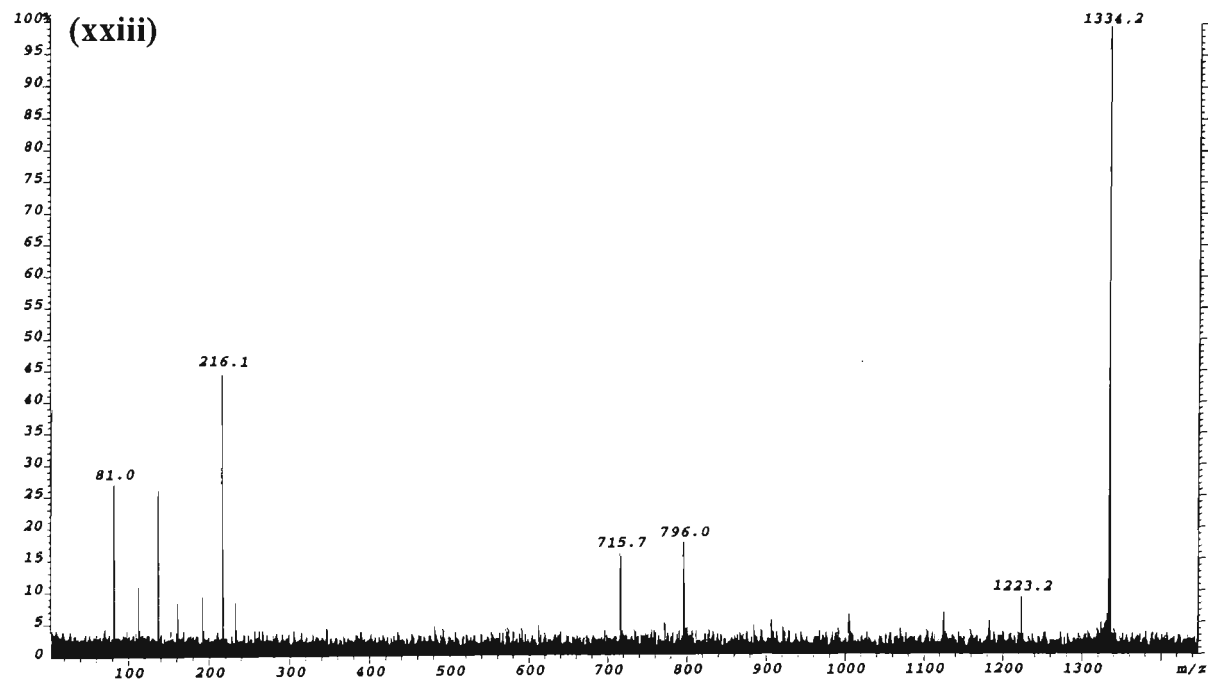
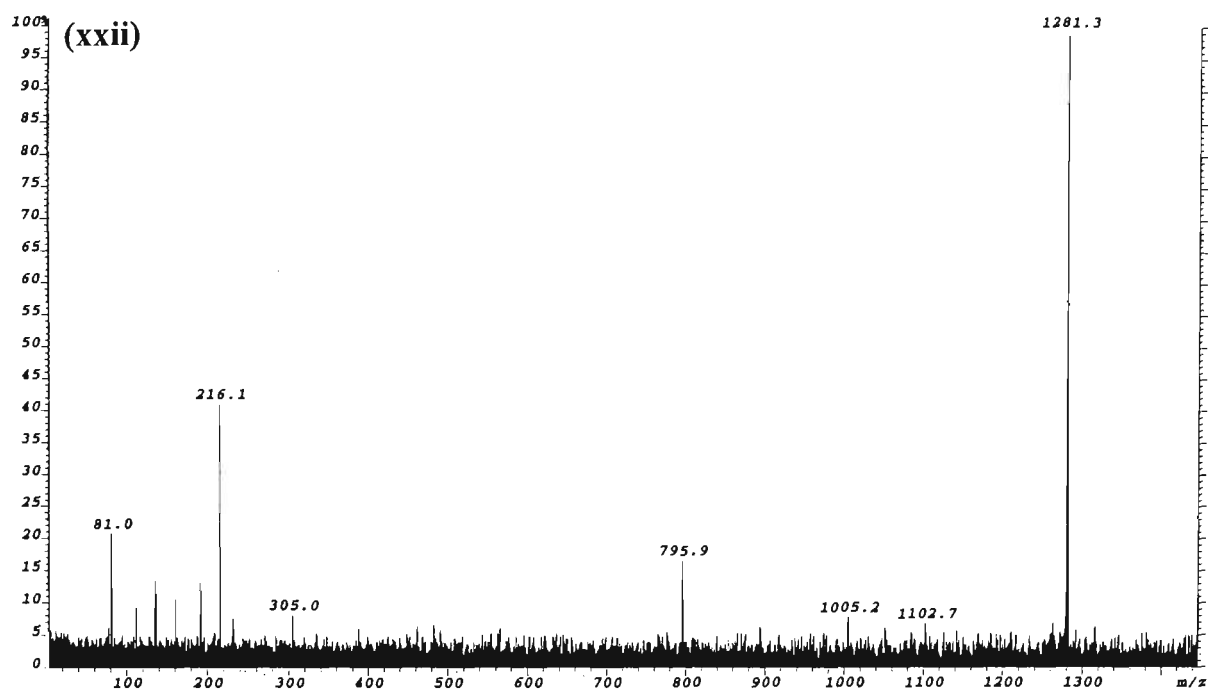


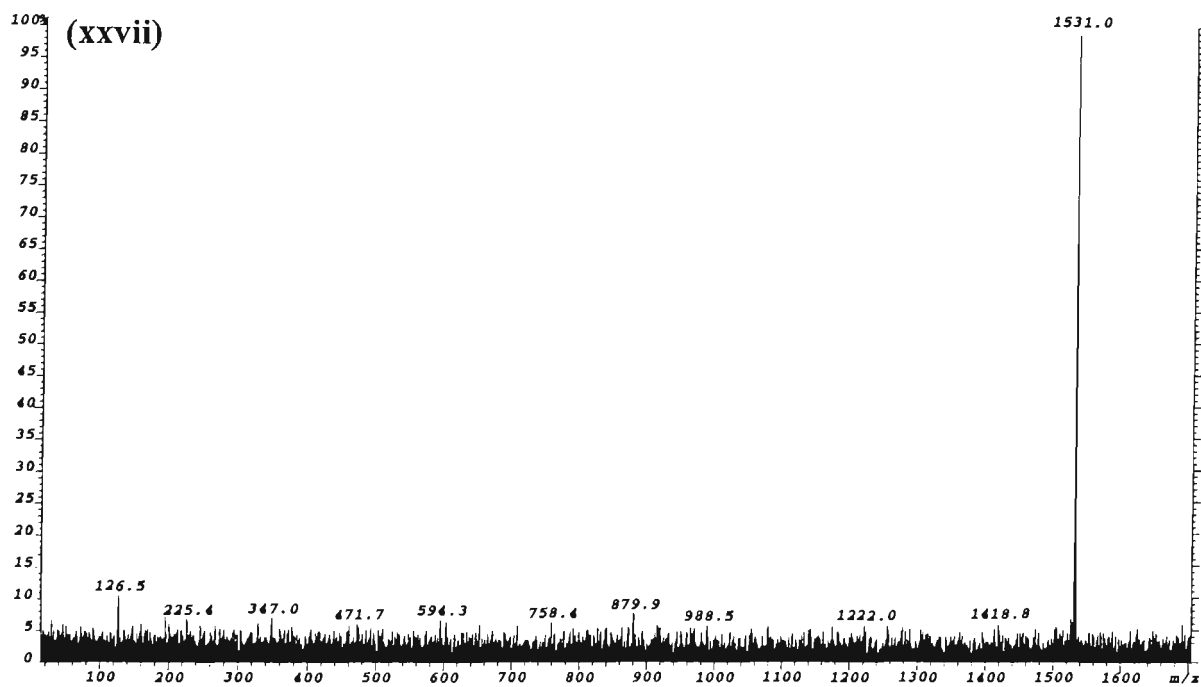
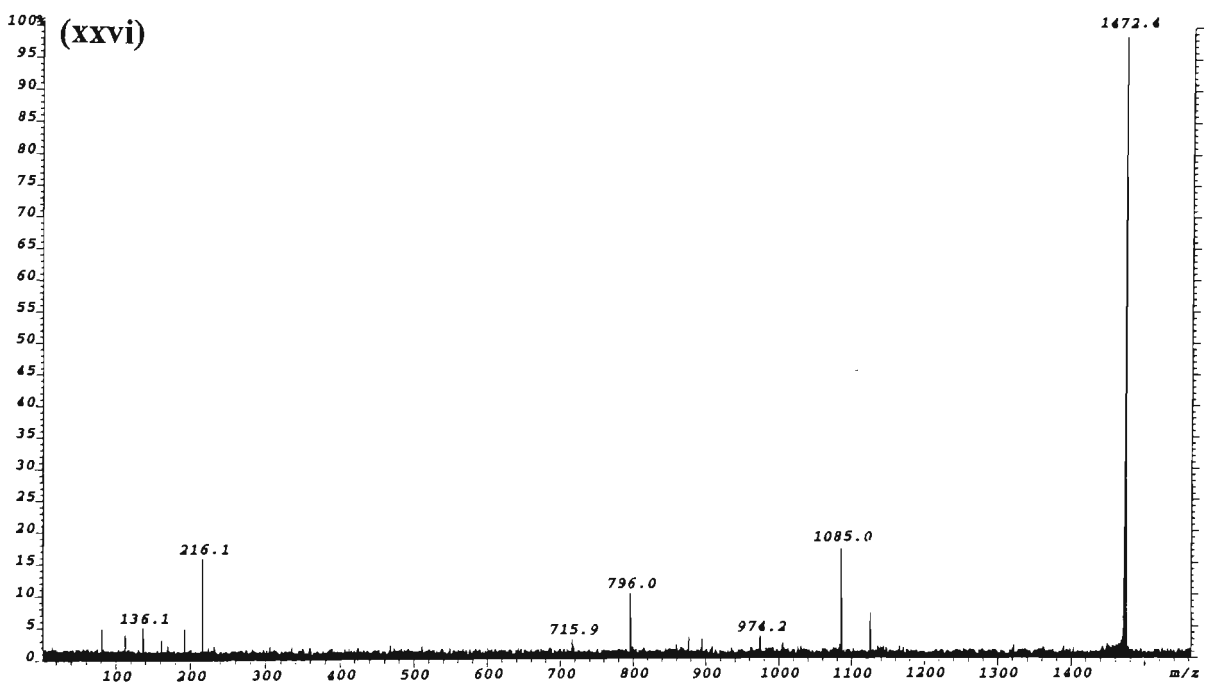
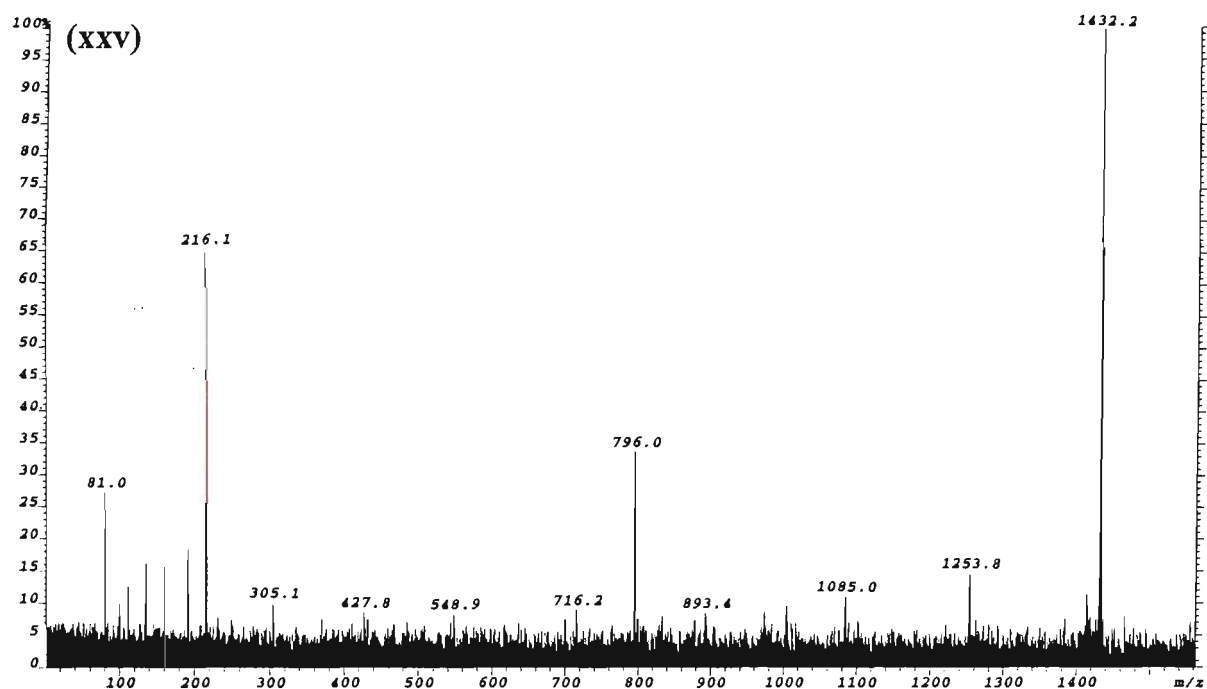


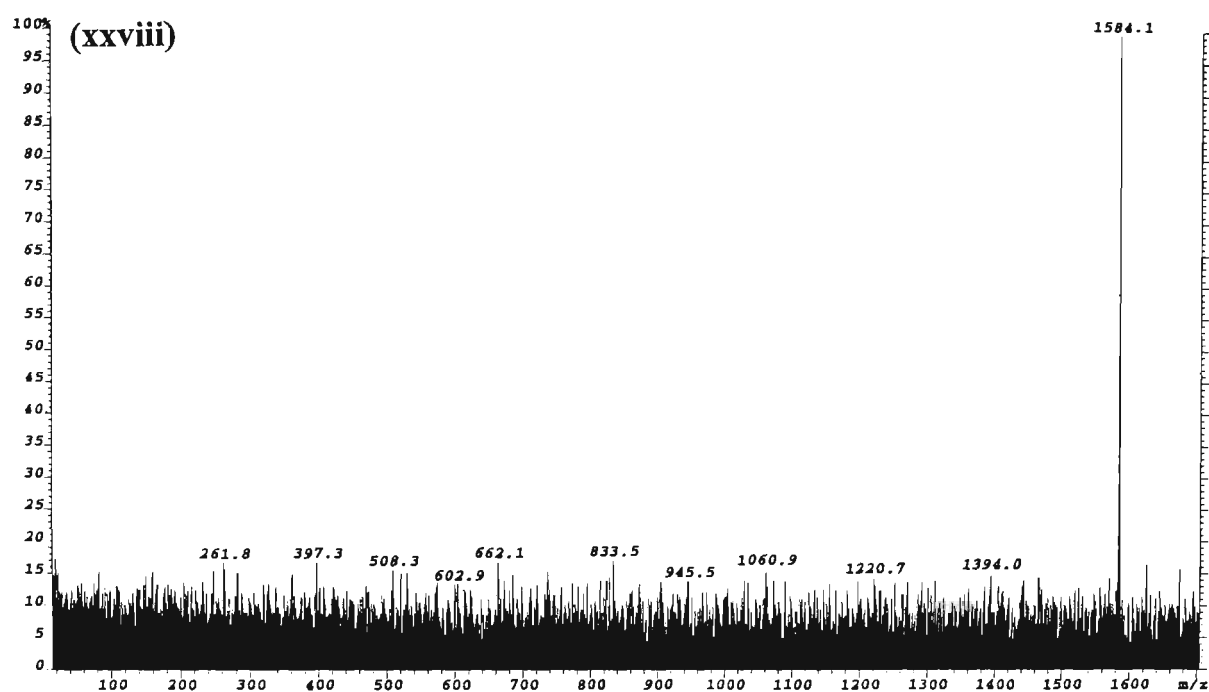












Appendix 5.1

Assignment of product ions observed in the ESI-MS/MS spectra of the $[M-2H]^{2-}$ ions of 5'-d(AAAAAA)-3', 5'-d(CCCCCC)-3', and 5'-d(TTTTTT)-3'.

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the oligonucleotide 5'-d(AAAAAA)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1582.3	0.06	1582.3	w_5^- and/or d_5^-
1502.3	0.10	1502.3	y_5^- and/or b_5^-
1484.3	0.11	1484.3	$[a_5-2H]^-$ and/or $[z_5-2H]^-$
1349.3	0.68	1349.3	$[a_5-B_5H-2H]^-$
1269.3	0.17	1269.2	w_4^- and/or d_4^- and/or $[(AAAA)+H_2O-H]^-$
1251.3	0.11	1251.2	$[c_4-2H]^-$ and/or $[x_4-2H]^-$ and/or $[(AAAA)-H]^-$
1189.4	0.60	1189.3	y_4^- and/or b_4^- and/or $[(AAAA)-p+H_2O-H]^-$
1171.3	0.10	1171.2	$[a_4-2H]^-$ and/or $[z_4-2H]^-$ and/or $[(AAAA)-p-H]^-$
1116.2	0.24	1116.2	$[c_4-AH-2H]^-$ and/or $[x_4-AH-2H]^-$ and/or $[(AAA)+s+p-H]^-$
1054.3	0.11	1054.2	$[y_4-AH]^-$ and/or $[b_4-AH]^-$ and/or $[(AAA)+s+H_2O-H]^-$
1036.2	1.24	1036.2	$[a_4-B_4H-2H]^-$ and/or $[z_4-AH-2H]^-$ and/or $[(AAA)+s-H]^-$
		and/or 1036.2	$[(AAA)+p+H_2O-H]^-$
1018.1	0.08	1018.1	$[(AAA)+p-H]^-$
		and/or 1018.2	$[(AAA)+s-H_2O-H]^-$
956.2	1.24	956.2	w_3^- and/or d_3^- and/or $[(AAA)+H_2O-H]^-$
938.1	0.51	938.2	$[c_3-2H]^-$ and/or $[x_3-2H]^-$ and/or $[(AAA)-H]^-$
907.2	100	907.2	$[M-2H]^{2-}$
876.2	0.79	876.2	y_3^- and/or b_3^- and/or $[(AAA)-p+H_2O-H]^-$
858.2	0.24	858.2	$[a_3-2H]^-$ and/or $[z_3-2H]^-$ and/or $[(AAA)-p-H]^-$
839.7	3.60	839.7	$[M-AH-2H]^{2-}$
821.2	0.14	821.1	$[w_3-AH]^-$ and/or $[d_3-AH]^-$ and/or $[(AA)+s+p+H_2O-H]^-$
803.1	1.18	803.1	$[c_3-AH-2H]^-$ and/or $[x_3-AH-2H]^-$ and/or $[(AA)+s+p-H]^-$
790.6	1.82	790.6	$[w_5-H]^{2-}$ and/or $[d_5-H]^{2-}$
785.1	0.45	785.1	$[(AA)+s+p-H_2O-H]^-$
781.6	0.20	781.6	$[c_5-3H]^{2-}$ and/or $[x_5-3H]^{2-}$
741.6	0.11	741.6	$[a_5-3H]^{2-}$ and/or $[z_5-3H]^{2-}$
723.1	2.47	723.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-AH-2H]^-$ and/or $[(AA)+s-H]^-$
		and/or 732.1	$[(AA)+p+H_2O-H]^-$
705.0	0.44	705.1	$[(AA)+p-H]^-$ and/or $[(AA)+s-H_2O-H]^-$
674.1	0.47	674.1	$[a_5-B_5H-3H]^{2-}$ and/or $[z_5-AH-3H]^{2-}$
643.0	4.51	643.1	w_2^- and/or d_2^- and/or $[(AA)+H_2O-H]^-$
634.1	1.69	634.1	$[w_4-H]^{2-}$ and/or $[d_4-H]^{2-}$ and/or $[(AAAA)+H_2O-2H]^{2-}$
625.0	1.86	625.1	$[c_2-3H]^{2-}$ and/or $[x_2-3H]^{2-}$ and/or $[(AA)-2H]^{2-}$
606.5	0.17	606.6	$[a_5-AH-AH-3H]^{2-}$ and/or $[z_5-AH-AH-3H]^{2-}$ and/or $[(AAA)+s+p+s-2H]^{2-}$
588.1	0.10	588.1	$[A_{nt}+s+p+s-H]^-$
		and/or 588.0	$[A_{nt}+p+s+p+H_2O-H]^-$
570.1	0.11	570.1	$[A_{nt}+p+s+p-H]^-$
		and/or 570.0	$[A_{nt}+s+p+s-H_2O-H]^-$
566.6	0.22	566.6	$[w_5-AH-H]^{2-}$ and/or $[d_5-AH-H]^{2-}$ and/or $[(AAA)+s+p+H_2O-2H]^{2-}$
563.1	0.37	563.1	y_2^- and/or b_2^- and/or $[(AA)-p+H_2O-H]^-$

545.0	0.23	545.1	$[a_2-2H]^-$ and/or $[z_2-2H]^-$ and/or $[(AA)-p-H]^-$
508.6	0.28	508.6	$[(AAA)+s-H_2O-2H]^{2-}$ and/or $[(AAA)+p-2H]^{2-}$
490.1	3.10	490.1	$[c_2-AH-2H]^-$ and/or $[x_2-AH-2H]^-$ and/or $[A_{nt}+s+p-H]^-$
472.1	1.24	472.1	$[A_{nt}+s+p-H_2O-H]^-$
410.0	2.09	410.1	$[a_2-B_2H-2H]^-$ and/or $[z_2-AH-2H]^-$ and/or $[A_{nt}+s-H]^-$
		410.0	$[A_{nt}+p+H_2O-H]^-$
392.0	5.36	392.1	$[A_{nt}+s-H_2O-H]^-$
		and/or 392.0	$[A_{nt}+p-H]^-$
330.1	10.5	330.1	w_1^- and/or d_1^- and/or $[A_{nt}+H_2O-H]^-$
312.0	10.3	312.0	$[c_1-2H]^-$ and/or $[x_1-2H]^-$ and/or $[A_{nt}-H]^-$
294.0	0.52	294.0	$[A_{nt}-H_2O-H]^-$
275.0	0.37	275.1	$[sps-H]^-$
		and/or 275.0	$[psp+H_2O-H]^-$
257.0	1.27	257.1	$[sps-H_2O-H]^-$
		and/or 257.0	$[psp-H]^-$
239.0	0.47	239.0	$[psp-H_2O-H]^-$
214.0	2.29	214.0	$[p+AH-H]^-$
195.0	3.27	195.0	$[ps+H_2O-H]^-$
177.0	8.15	177.0	$[ps-H]^-$
159.0	1.07	159.0	$[ps-H_2O-H]^-$
134.1	13.7	134.0	A^-
97.0	2.13	97.0	$[a_1-B_1H-2H]^-$ and/or $[s-H]^-$
		and/or 97.0	$[p+H_2O-H]^-$
79.0	10.7	79.0	$[p-H]^-$ (PO_3^-)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the oligonucleotide 5'-d(CCCCCC)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1253.7	0.52	1253.2	$[a_5-B_5H-2H]^-$
1173.3	0.44	1173.2	w_4^- and/or d_4^- and/or $[(CCCC)+H_2O-H]^-$
1093.1	0.94	1093.2	y_4^- and/or b_4^- and/or $[(CCCC)-p+H_2O-H]^-$
1075.2	0.34	1075.2	$[a_4-2H]^-$ and/or $[z_4-2H]^-$ and/or $[(CCCC)-p-H]^-$
1044.1	0.33	1044.1	$[c_4-CH-2H]^-$ and/or $[x_4-CH-2H]^-$ and/or $[(CCC)+s+p-H]^-$
982.1	0.18	982.2	$[y_4-CH]^-$ and/or $[b_4-CH]^-$ and/or $[(CCC)+s+H_2O-H]^-$
963.9	0.86	964.2	$[a_4-B_4H-2H]^-$ and/or $[z_4-CH-2H]^-$ and/or $[(CCC)+s-H]^-$
		and/or 964.1	$[(CCC)+p+H_2O-H]^-$
884.1	1.45	884.1	w_3^- and/or d_3^- and/or $[(CCC)+H_2O-H]^-$
866.1	0.75	866.1	$[c_3-2H]^-$ and/or $[x_3-2H]^-$ and/or $[(CCC)-H]^-$
835.1	100	835.1	$[M-2H]^{2-}$
804.2	0.62	804.2	y_3^- and/or b_3^- and/or $[(CCC)-p+H_2O-H]^-$
786.2	0.45	786.2	$[a_3-2H]^-$ and/or $[z_3-2H]^-$ and/or $[(CCC)-p-H]^-$
779.6	1.94	779.6	$[M-CH-2H]^{2-}$
755.1	0.97	755.1	$[c_3-CH-2H]^-$ and/or $[x_3-CH-2H]^-$ and/or $[(CC)+s+p-H]^-$
737.1	0.37	737.1	$[(CC)+s+p-H_2O-H]^-$
730.6	3.06	730.6	$[w_5-H]^{2-}$ and/or $[d_5-H]^{2-}$
721.6	0.40	721.6	$[c_5-3H]^{2-}$ and/or $[x_5-3H]^{2-}$
690.6	0.23	690.6	$[a_5-3H]^{2-}$ and/or $[z_5-3H]^{2-}$
675.1	1.86	675.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-CH-2H]^-$ and/or $[(CC)+s-H]^-$
		and/or 675.1	$[(CC)+p+H_2O-H]^-$
657.0	0.48	657.1	$[(CC)+p-H]^-$ and/or $[(CC)+s-H_2O-H]^-$

626.1	0.45	626.1	$[a_5-B_5H-3H]^{2-}$ and/or $[z_5-CH-3H]^{2-}$
594.9	4.46	595.1	w_2^- and/or d_2^- and/or $[(CC)+H_2O-H]^-$
586.1	2.23	586.1	$[w_4-H]^{2-}$ and/or $[d_4-H]^{2-}$ and/or $[(CCCC)+H_2O-2H]^{2-}$
577.1	1.32	577.1	$[c_2-3H]^{2-}$ and/or $[x_2-3H]^{2-}$ and/or $[(CC)-2H]^{2-}$
534.0	0.29	534.0	$[C_{nt}+p+s+p-H_2O-H]^-$
515.1	0.45	515.1	y_2^- and/or b_2^- and/or $[(CC)-p+H_2O-H]^-$
497.1	0.37	497.1	$[a_2-2H]^-$ and/or $[z_2-2H]^-$ and/or $[(CC)-p-H]^-$
484.0	0.42	484.0	$[C_{nt}+s+p+H_2O-H]^-$
466.0	2.12	466.0	$[c_2-CH-2H]^-$ and/or $[x_2-CH-2H]^-$ and/or $[C_{nt}+s+p-H]^-$
448.0	0.46	448.0	$[C_{nt}+s+p-H_2O-H]^-$
404.0	0.31	404.1	$[C_{nt}+s+H_2O-H]^-$
386.1	2.43	386.1	$[a_2-B_2H-2H]^-$ and/or $[z_2-CH-2H]^-$ and/or $[C_{nt}+s-H]^-$
		386.0	$[C_{nt}+p+H_2O-H]^-$
368.1	3.38	368.1	$[C_{nt}+s-H_2O-H]^-$
		and/or 368.0	$[C_{nt}+p-H]^-$
306.1	10.1	306.0	w_1^- and/or d_1^- and/or $[C_{nt}+H_2O-H]^-$
288.1	9.02	288.0	$[c_1-2H]^-$ and/or $[x_1-2H]^-$ and/or $[C_{nt}-H]^-$
275.0	0.41	275.1	$[sps-H]^-$
		and/or 275.0	$[psp+H_2O-H]^-$
270.0	0.26	270.0	$[C_{nt}-H_2O-H]^-$
257.0	0.92	257.1	$[sps-H_2O-H]^-$
		and/or 257.0	$[psp-H]^-$
239.0	0.24	239.0	$[psp-H_2O-H]^-$
195.0	1.65	195.0	$[ps+H_2O-H]^-$
177.0	4.27	177.0	$[ps-H]^-$
159.0	0.85	159.0	$[ps-H_2O-H]^-$
110.0	2.87	110.0	C^-
97.0	4.46	97.0	$[a_1-B_1H-2H]^-$ and/or $[s-H]^-$
		and/or 97.0	$[p+H_2O-H]^-$
79.0	9.10	79.0	$[p-H]^-$ (PO_3^-)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-2H]^{2-}$ ion of the oligonucleotide 5'-d(TTTTTT)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1537.3	0.28	1537.2	w_5^- and/or d_5^-
1457.2	0.22	1457.3	y_5^- and/or b_5^-
1438.8	0.35	1439.2	$[a_5-2H]^-$ and/or $[z_5-2H]^-$
1312.5	0.31	1312.2	$[a_5-B_5H-2H]^-$
1233.3	0.38	1233.2	w_4^- and/or d_4^- and/or $[(TTTT)+H_2O-H]^-$
1215.3	0.35	1215.2	$[c_4-2H]^-$ and/or $[x_4-2H]^-$ and/or $[(TTTT)-H]^-$
1153.2	2.14	1153.2	y_4^- and/or b_4^- and/or $[(TTTT)-p+H_2O-H]^-$
1135.2	0.35	1135.2	$[a_4-2H]^-$ and/or $[z_4-2H]^-$ and/or $[(TTTT)-p-H]^-$
1089.1	0.19	1089.1	$[c_4-TH-2H]^-$ and/or $[x_4-TH-2H]^-$ and/or $[(TTT)+s+p-H]^-$
1027.1	0.16	1027.2	$[y_4-TH]^-$ and/or $[b_4-TH]^-$ and/or $[(TTT)+s+H_2O-H]^-$
1009.1	0.60	1009.1	$[a_4-B_4H-2H]^-$ and/or $[z_4-TH-2H]^-$ and/or $[(TTT)+s-H]^-$
		and/or 1009.1	$[(TTT)+p+H_2O-H]^-$
929.1	1.92	929.2	w_3^- and/or d_3^- and/or $[(TTT)+H_2O-H]^-$
911.1	1.23	911.2	$[c_3-2H]^-$ and/or $[x_3-2H]^-$ and/or $[(TTT)-H]^-$
880.2	100	880.2	$[M-2H]^{2-}$
849.2	1.26	849.2	y_3^- and/or b_3^- and/or $[(TTT)-p+H_2O-H]^-$
831.1	0.50	831.1	$[a_3-2H]^-$ and/or $[z_3-2H]^-$ and/or $[(TTT)-p-H]^-$
818.1	1.54	818.1	$[M-TH-2H]^{2-}$

785.1	0.95	785.1	$[c_3\text{-TH-2H}]^-$ and/or $[x_3\text{-TH-2H}]^-$ and/or $[(\text{TT})+\text{s+p-H}]^-$
768.1	6.03	768.1	$[w_5\text{-H}]^{2-}$ and/or $[d_5\text{-H}]^{2-}$
759.1	0.54	759.1	$[(\text{TT})+\text{s+p-H}_2\text{O-H}]^-$
728.1	0.35	728.1	$[c_5\text{-3H}]^{2-}$ and/or $[x_5\text{-3H}]^{2-}$
719.1	0.16	719.1	$[a_5\text{-3H}]^{2-}$ and/or $[z_5\text{-3H}]^{2-}$
705.1	1.45	705.1	$[a_3\text{-B}_3\text{H-2H}]^-$ and/or $[z_3\text{-TH-2H}]^-$ and/or $[(\text{TT})+\text{s-H}]^-$
		and/or 705.1	$[(\text{TT})+\text{p+H}_2\text{O-H}]^-$
687.1	0.79	687.1	$[(\text{TT})+\text{p-H}]^-$ and/or $[(\text{TT})+\text{s-H}_2\text{O-H}]^-$
656.1	0.35	656.1	$[a_5\text{-B}_5\text{H-3H}]^{2-}$ and/or $[z_5\text{-TH-3H}]^{2-}$
625.0	4.72	625.1	w_2^- and/or d_2^- and/or $[(\text{TT})+\text{H}_2\text{O-H}]^-$
616.0	5.07	616.1	$[w_4\text{-H}]^{2-}$ and/or $[d_4\text{-H}]^{2-}$ and/or $[(\text{TTTT})+\text{H}_2\text{O-2H}]^{2-}$
607.1	4.13	607.1	$[c_2\text{-3H}]^{2-}$ and/or $[x_2\text{-3H}]^{2-}$ and/or $[(\text{TT})\text{-2H}]^{2-}$
593.1	0.25	593.1	$[a_5\text{-TH-TH-3H}]^{2-}$ and/or $[z_5\text{-TH-TH-3H}]^{2-}$ and/or $[(\text{TTT})+\text{s+p+s-2H}]^{2-}$
561.1	0.10	561.1	$[\text{T}_{\text{nt}}+\text{p+s+p-H}]^-$
		and/or 561.0	$[\text{T}_{\text{nt}}+\text{s+p+s-H}_2\text{O-H}]^-$
553.1	0.20	553.1	$[w_5\text{-TH-H}]^{2-}$ and/or $[d_5\text{-TH-H}]^{2-}$ and/or $[(\text{TTT})+\text{s+p+H}_2\text{O-2H}]^{2-}$
545.0	0.63	545.1	y_2^- and/or b_2^- and/or $[(\text{TT})\text{-p+H}_2\text{O-H}]^-$
527.0	0.50	527.1	$[a_2\text{-2H}]^-$ and/or $[z_2\text{-2H}]^-$ and/or $[(\text{TT})\text{-p-H}]^-$
499.0	0.38	499.0	$[\text{T}_{\text{nt}}+\text{s+p+H}_2\text{O-H}]^-$
481.0	2.90	481.0	$[c_2\text{-TH-2H}]^-$ and/or $[x_2\text{-TH-2H}]^-$ and/or $[\text{T}_{\text{nt}}+\text{s+p-H}]^-$
463.0	0.98	463.0	$[\text{T}_{\text{nt}}+\text{s+p-H}_2\text{O-H}]^-$
418.0	0.28	418.1	$[\text{T}_{\text{nt}}+\text{s+H}_2\text{O-H}]^-$
401.0	3.25	401.1	$[a_2\text{-B}_2\text{H-2H}]^-$ and/or $[z_2\text{-TH-2H}]^-$ and/or $[\text{T}_{\text{nt}}+\text{s-H}]^-$
		401.0	$[\text{T}_{\text{nt}}+\text{p+H}_2\text{O-H}]^-$
383.1	9.14	383.1	$[\text{T}_{\text{nt}}+\text{s-H}_2\text{O-H}]^-$
		and/or 383.0	$[\text{T}_{\text{nt}}+\text{p-H}]^-$
321.1	15.4	321.1	w_1^- and/or d_1^- and/or $[\text{T}_{\text{nt}}+\text{H}_2\text{O-H}]^-$
303.1	37.3	303.0	$[c_1\text{-2H}]^-$ and/or $[x_1\text{-2H}]^-$ and/or $[\text{T}_{\text{nt}}\text{-H}]^-$
285.0	0.32	285.0	$[\text{T}_{\text{nt}}\text{-H}_2\text{O-H}]^-$
275.0	0.95	275.1	$[\text{sps-H}]^-$
		and/or 275.0	$[\text{psp+H}_2\text{O-H}]^-$
257.0	2.84	257.1	$[\text{sps-H}_2\text{O-H}]^-$
		and/or 257.0	$[\text{psp-H}]^-$
239.0	0.47	239.0	$[\text{psp-H}_2\text{O-H}]^-$
195.0	7.00	195.0	$[\text{ps+H}_2\text{O-H}]^-$
177.0	12.3	177.0	$[\text{ps-H}]^-$
159.0	1.77	159.0	$[\text{ps-H}_2\text{O-H}]^-$
125.1	23.3	125.0	T^-
97.0	4.7	97.0	$[a_1\text{-B}_1\text{H-2H}]^-$ and/or $[\text{s-H}]^-$
		and/or 97.0	$[\text{p+H}_2\text{O-H}]^-$
79.0	15.9	79.0	$[\text{p-H}]^- (\text{PO}_3^-)$

*calculated monoisotopic mass

s = deoxyribose- H_2O ($\text{C}_5\text{H}_6\text{O}_2$), $M_r = 98.0368$ Da

p = PO_3H , $M_r = 79.9663$ Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

$[(\text{B}_1\text{..B}_n)]$ denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Appendix 5.2

Assignment of product ions observed in the ESI-MS/MS spectra of the [M-3H]³⁻, [M-4H]⁴⁻ and [M-5H]⁵⁻ ions of 5'-d(CACGTG)-3', 5'-d(CGTACG)-3', 5'-d(ATGCAT)-3', 5'-d(TCACGA)-3', 5'-d(CGGCCG)-3', and 5'-d(GCCGGC)-3'.

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-3H]³⁻ ion of the oligonucleotide 5'-d(CACGTG)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1317.2	0.04	1317.2	[a ₅ -B ₅ H-2H] ⁻
1268.2	0.01	1268.2	w ₄ ⁻
1188.2	0.02	1188.2	y ₄ ⁻
1139.0	0.02	1139.1	[a ₄ -2H] ⁻
1108.2	0.02	1108.2	[c ₄ -CH-2H] ⁻ and/or [(ACG)+s+p-H] ⁻
988.2	0.02	988.2	[a ₄ -B ₄ H-2H] ⁻
979.2	0.02	979.2	w ₃ ⁻
930.1	0.01	930.1	[(ACG)-H] ⁻
899.1	0.04	899.2	[y ₃ -2H] ⁻
875.1	0.02	875.1	[(CG)+p+s+p-H] ⁻
839.2	0.1	839.1	[M-C-3H] ²⁻
831.5	0.03	831.6	[M-T-3H] ²⁻
828.5	0.03	828.1	[w ₃ -GH] ⁻ and/or [(GT)+s+p+H ₂ O-H] ⁻
827.5	0.03	827.1	[M-A-3H] ²⁻
819.2	0.04	819.1	[M-G-3H] ²⁻
810.0	0.02	810.2	[a ₃ -2H] ⁻
	and/or	810.1	[x ₃ -GH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
795.2	0.04	795.1	[(CG)+s+p-H] ⁻
790.3	0.41	790.1	[w ₅ -H] ²⁻
779.0	0.04	779.1	[c ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s+p-H] ⁻
750.0	0.06	750.1	[y ₅ -H] ²⁻
730.0	0.02	730.1	[z ₃ -GH-2H] ⁻ and/or [(GT)+s-H] ⁻
721.1	0.04	721.1	[a ₅ -3H] ²⁻
714.4	0.04	714.6	[w ₅ -GH-H] ²⁻ and/or [(ACGT)+p+s+H ₂ O-2H] ²⁻
699.0	0.08	699.1	[a ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s-H] ⁻
	and/or	699.1	[(AC)+p+H ₂ O-H] ⁻
698.0	0.03	698.1	[c ₅ -B ₅ H-3H] ²⁻ and/or [(CACG)+s+p-2H] ²⁻
658.1	0.07	658.1	[a ₅ -B ₅ H-3H] ²⁻
650.3	0.14	650.1	w ₂ ⁻ and/or [(GT)+H ₂ O-H] ⁻
633.6	0.21	633.6	[w ₄ -H] ²⁻
619.0	0.08	619.1	d ₂ ⁻ and/or [(AC)+H ₂ O-H] ⁻
616.6	0.09	616.5	[(CGTA)-2H] ²⁻
596.1	100	596.1	[M-3H]³⁻
570.0	0.08	570.1	y ₂ ⁻ and/or [(GT)-H ₂ O-H] ⁻
559.1	0.2	559.1	[M-CH-3H] ³⁻
551.1	0.04	551.1	[M-AH-3H] ³⁻
546.1	0.09	546.1	[M-GH-3H] ³⁻
526.5	0.21	526.4	[w ₅ -2H] ³⁻
513.5	0.02	513.6	[a ₄ -CH-3H] ²⁻ and/or [(ACG)+s-2H] ²⁻
	and/or	513.6	[(ACG)+p+H ₂ O-2H] ²⁻
506.0	0.08	506.0	[x ₂ -TH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
493.5	0.04	493.6	[a ₄ -B ₄ H-3H] ²⁻
490.0	0.08	490.1	[c ₂ -CH-2H] ⁻ and/or [A _{nt} +s+p-H] ⁻
489.1	0.13	489.1	[w ₃ -H] ²⁻
466.0	0.02	466.0	[c ₂ -AH-2H] ⁻ and/or [C _{nt} +s+p-H] ⁻
426.0	0.02	426.1	[z ₂ -TH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
422.0	0.02	422.1	[w ₄ -2H] ³⁻

410.0	0.01	410.1	[a ₂ -CH-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	0.03	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
392.0	0.01	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.0	0.04	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.0	0.3	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
368.0	0.02	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.1	0.44	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.0	0.05	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
324.6	0.1	324.5	[w ₂ -H] ²⁻ and/or [(GT)+H ₂ O-2H] ²⁻
312.0	0.03	312.0	[A _{nt} -H] ⁻
306.0	0.01	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.0	0.03	303.0	[T _{nt} -H] ⁻
288.1	0.14	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.0	0.01	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.0	0.01	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.01	239.0	[psp-H ₂ O-H] ⁻
230.0	0.02	230.1	[sG-H ₂ O-H] ⁻
195.0	0.03	195.0	[ps+H ₂ O-H] ⁻
177.0	0.13	177.0	[ps-H] ⁻
158.9	0.02	159.0	[ps-H ₂ O-H] ⁻
150.0	0.1	150.0	G ⁻
134.1	0.14	134.0	A ⁻
125.0	0.30	125.0	T ⁻
110.0	0.41	110.0	C ⁻
97.0	0.17	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	0.83	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-4H]⁴⁻ ion of the oligonucleotide 5'-d(CACGTG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc). *</i>	<i>Assignment</i>
1188.3	0.18	1188.2	y ₄ ⁻
1139.0	0.03	1139.1	[a ₄ -2H] ⁻
1108.3	0.06	1108.2	[c ₄ -CH-2H] ⁻ and/or [(ACG)+s+p-H] ⁻
1028.2	0.05	1028.2	[a ₄ -CH-2H] ⁻ and/or [(ACG)+s-H] ⁻
	and/or	1028.1	[(ACG)+p+H ₂ O-H] ⁻
988.2	0.04	988.2	[a ₄ -B ₄ H-2H] ⁻
979.2	0.09	979.2	w ₃ ⁻
961.1	0.02	961.1	[x ₃ -2H] ⁻
930.1	0.03	930.1	[(ACG)-H] ⁻
921.1	0.03	921.1	[(CGT)-H] ⁻
899.1	0.1	899.2	[y ₃ -2H] ⁻
810.0	0.03	810.2	[a ₃ -2H] ⁻
	and/or	810.1	[x ₃ -GH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
795.0	0.23	795.1	[(CG)+s+p-H] ⁻
779.0	0.04	779.1	[c ₃ -B ₃ H-2H] ⁻ and/or [(AC)+s+p-H] ⁻

750.1	0.49	750.1	[y5-H] ²⁻
730.0	0.02	730.1	[z3-GH-2H] ⁻ and/or [(GT)+s-H] ⁻
721.1	0.18	721.1	[a5-3H] ²⁻
715.1	0.09	715.1	[(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p-H ₂ O-H] ⁻
699.0	0.08	699.1	[a3-B3H-2H] ⁻ and/or [(AC)+s-H] ⁻
	and/or	699.1	[(AC)+p+H ₂ O-H] ⁻
698.0	0.05	698.1	[c5-B5H-3H] ²⁻ and/or [(CACG)+s+p-2H] ²⁻
681.1	0.02	681.1	[(AC)+p-H] ⁻
	and/or	681.1	[(AC)+s-H ₂ O-H] ⁻
658.0	0.43	658.1	[a5-B5H-3H] ²⁻
650.1	0.35	650.1	w ₂ ⁻ and/or [(GT)+H ₂ O-H] ⁻
633.5	1.32	633.6	[w4-H] ²⁻
619.0	0.05	619.1	d ₂ ⁻ and/or [(AC)+H ₂ O-H] ⁻
617.0	0.10	617.1	[(CG)-H] ⁻
609.0	0.03	609.1	[c4-3H] ²⁻
601.0	0.06	601.1	[c2-2H] ⁻ and/or [(AC)-H] ⁻
593.6	0.06	593.6	[y4-H] ²⁻
570.0	0.08	570.1	y ₂ ⁻ and/or [(GT)-H ₂ O-H] ⁻
559.1	1.08	559.1	[M-C ⁻ -4H] ³⁻
554.1	0.49	554.1	[M-T ⁻ -4H] ³⁻
551.0	1.81	551.1	[M-A ⁻ -4H] ³⁻
545.9	0.27	546.1	[M-G ⁻ -4H] ³⁻
526.4	1.47	526.4	[w5-2H] ³⁻
506.0	0.81	506.0	[x2-TH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
493.5	0.13	493.6	[a4-B4H-3H] ²⁻
490.0	0.19	490.1	[c2-CH-2H] ⁻ and/or [A _{nt} +s+p-H] ⁻
489.1	0.13	489.1	[w3-H] ²⁻
488.1	0.2	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
466.1	0.35	466.0	[c2-AH-2H] ⁻ and/or [C _{nt} +s+p-H] ⁻
446.8	100	446.8	[M-4H]⁴⁻
426.0	0.18	426.1	[z2-TH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
422.0	0.50	422.1	[w4-2H] ³⁻
419.0	0.23	419.1	[M-CH-4H] ⁴⁻
413.1	0.03	413.1	[M-AH-4H] ⁴⁻
410.0	0.09	410.1	[a2-CH-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	0.30	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
394.5	0.07	394.5	[w5-3H] ⁴⁻
392.0	0.03	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.1	0.94	386.1	[a2-B2H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.0	0.18	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.0	1.86	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	0.66	328.0	[x1-2H] ⁻ and/or [G _{nt} -H] ⁻
325.6	0.33	325.7	[w3-2H] ³⁻
324.5	0.36	324.5	[w2-H] ²⁻ and/or [(GT)+H ₂ O-2H] ²⁻
312.0	0.13	312.0	[A _{nt} -H] ⁻
306.0	0.07	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.0	0.64	303.0	[T _{nt} -H] ⁻
288.0	1.03	288.0	[c1-2H] ⁻ and/or [C _{nt} -H] ⁻
275.0	0.18	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.0	0.22	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.04	239.0	[psp-H ₂ O-H] ⁻
230.0	0.11	230.1	[sG-H ₂ O-H] ⁻
195.0	0.27	195.0	[ps+H ₂ O-H] ⁻
177.0	0.96	177.0	[ps-H] ⁻

172.5	0.53	172.5	[w ₂ -2H] ³⁻
159.0	0.20	159.0	[ps-H ₂ O-H] ⁻
150.1	0.59	150.0	G ⁻
134.0	1.39	134.0	A ⁻
125.0	1.12	125.0	T ⁻
110.0	2.33	110.0	C ⁻
97.0	0.84	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	5.13	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.23	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-5H]⁵⁻ ion of the oligonucleotide 5'-d(CACGTG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc). *</i>	<i>Assignment</i>
899.3	0.09	899.2	[y ₃ -2H] ⁻
795.1	0.29	795.1	[(CG)+s+p-H] ⁻
715.0	0.16	715.1	[(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p-H ₂ O-H] ⁻
696.9	0.12	697.1	[(CG)+s-H ₂ O-H] ⁻
	and/or	697.1	[(CG)+p-H] ⁻
657.9	0.49	658.1	[a ₅ -B ₅ H-3H] ²⁻
650.0	0.19	650.1	w ₂ ⁻ and/or [(GT)+H ₂ O-H] ⁻
633.7	0.24	633.6	[w ₄ -H] ²⁻
617.0	0.07	617.1	[(CG)-H] ⁻
593.7	0.34	593.6	[y ₄ -H] ²⁻
506.0	0.34	506.0	[x ₂ -TH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
489.1	0.17	489.1	[w ₃ -H] ²⁻
487.9	0.13	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
465.9	0.27	466.0	[c ₂ -AH-2H] ⁻ and/or [C _{nt} +s+p-H] ⁻
426.0	0.17	426.1	[z ₂ -TH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
422.0	1.61	422.1	[w ₄ -2H] ³⁻
419.0	1.22	419.1	[M-C ⁻ -5H] ⁴⁻
413.0	8.37	413.1	[M-A ⁻ -5H] ⁴⁻
415.3	0.40	415.3	[M-T ⁻ -5H] ⁴⁻
409.0	0.57	409.1	[M-G ⁻ -5H] ⁴⁻
408.0	0.52	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
394.3	0.20	394.5	[w ₅ -3H] ⁴⁻
385.9	0.94	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
367.9	0.34	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
357.3	100	357.3	[M-5H]⁵⁻
346.0	1.75	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	0.67	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
325.6	0.20	325.7	[w ₃ -2H] ³⁻
324.5	0.24	324.5	[w ₂ -H] ²⁻ and/or [(GT)+H ₂ O-2H] ²⁻
316.1	0.18	316.3	[w ₄ -3H] ⁴⁻
312.0	0.17	312.0	[A _{nt} -H] ⁻
306.0	0.10	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.1	0.46	303.0	[T _{nt} -H] ⁻
288.0	0.97	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
283.3	0.92	283.1	[G _{nt} +p+H ₂ O-H] ⁻
274.9	0.44	275.1	[sps-H] ⁻

	and/or	275.0	[psp+ H ₂ O-H] ⁻
257.0	0.28	257.1	[sps- H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
241.4	0.15	241.1	[T _{nt} -p+H ₂ O-H] ⁻
239.0	0.08	239.0	[psp- H ₂ O-H] ⁻
230.0	0.15	230.1	[sG-H ₂ O-H] ⁻
216.0	0.10	216.0	[w ₂ -2H] ³⁻ and/or [(GT)+H ₂ O-3H] ³⁻
195.0	0.32	195.0	[ps+ H ₂ O-H] ⁻
177.0	1.97	177.0	[ps-H] ⁻
172.5	1.34	172.5	[w ₂ -2H] ³⁻
158.9	0.15	159.0	[ps-H ₂ O-H] ⁻
150.1	0.75	150.0	G ⁻
134.0	1.92	134.0	A ⁻
125.0	1.38	125.0	T ⁻
110.0	2.76	110.0	C ⁻
97.0	1.20	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	7.35	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.54	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-3H]³⁻ ion of the oligonucleotide 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc). *</i>	<i>Assignment</i>
1501.5	0.02	1501.3	y ₅ ⁻
1332.3	0.02	1332.3	[a ₅ -B ₅ H-2H] ⁻ and/or [z ₅ -GH-2H] ⁻
		1332.2	[(GTAC)+p+H ₂ O-H] ⁻
1252.3	0.02	1252.2	w ₄ ⁻ and/or d ₄ ⁻ and/or [(GTAC)+H ₂ O-H] ⁻
1234.3	0.01	1234.2	[x ₄ -2H] ⁻ and/or [c ₄ -2H] ⁻ and/or [(GTAC)-H] ⁻
1221.3	0.01	1221.2	[a ₅ -CH-CH-2H] ⁻ and/or [z ₅ -CH-GH-2H] ⁻
1172.0	0.06	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GTAC)-p+H ₂ O-H] ⁻
1154.4	0.06	1154.2	[a ₄ -2H] ⁻ and/or [z ₄ -2H] ⁻ and/or [(GTAC)-p-H] ⁻
1123.4	0.02	1123.1	[x ₄ -CH-2H] ⁻ and/or [c ₄ -CH-2H] ⁻ and/or [(GTA)+s+p-H] ⁻
1043.2	0.04	1043.2	[a ₄ -CH-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(GTA)+s-H] ⁻
		1043.1	[(GTA)+p+H ₂ O-H] ⁻
1019.4	0.14	1019.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -AH-2H] ⁻ and/or [(GTAC)-AH-p-H] ⁻
963.1	0.01	963.1	[(GTA)+H ₂ O-H] ⁻
948.1	0.33	948.2	w ₃ ⁻
868.0	0.14	868.2	y ₃ ⁻
859.2	0.02	859.2	b ₃ ⁻
841.2	0.14	841.2	[a ₃ -2H] ⁻
839.2	0.18	839.1	[M-C ⁻ -3H] ²⁻
832.6	0.14	832.7	[M-T ⁻ -3H] ²⁻
827.2	0.07	827.1	[M-A ⁻ -3H] ²⁻
819.2	0.08	819.1	[M-G ⁻ -3H] ²⁻
810.2	0.10	810.1	[c ₃ -CH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
794.1	0.11	794.1	[(TA)+s+p-H] ⁻
790.1	0.62	790.1	[w ₅ -H] ²⁻
750.1	0.30	750.1	[y ₅ -H] ⁻
730.1	0.10	730.1	[b ₅ -H] ²⁻
	and/or	730.1	[(GT)+s-H] ⁻
	and/or	730.0	[(GT)+p+H ₂ O-H] ⁻
715.1	0.72	715.0	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -AH-2H] ⁻
699.1	0.05	699.1	[(AC)+s -H] ⁻

	and/or	681.1	[(AC)+p+H ₂ O-H] ⁻
696.1	0.05	696.1	[(TA)+s-H ₂ O-H] ⁻
		696.1	[(TA)+p-H] ⁻
681.1	0.04	681.1	[(AC)+s-H ₂ O-H] ⁻
	and/or	681.1	[(AC)+p-H] ⁻
665.6	0.40	665.6	[a ₅ -B ₅ H-3H] ²⁻ and/or [z ₅ -GH-3H] ²⁻
650.1	0.06	650.1	[(GT)+H ₂ O-H] ⁻
635.1	0.83	635.1	w ₂ ⁻ and/or d ₂ ⁻
632.2	0.18	632.0	[(GT)-H] ⁻
625.6	0.52	625.6	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(GTAC)+H ₂ O-2H] ²⁻
617.1	0.38	617.1	[x ₂ -2H] ⁻ and/or [c ₂ -2H] ⁻
596.1	100	596.1	[M-3H]³⁻
559.1	0.44	559.1	[M-CH-3H] ³⁻
555.1	0.18	555.1	y ₂ ⁻ and/or b ₂ ⁻
551.1	0.14	551.1	[M-AH-3H] ³⁻
546.1	0.15	546.1	[M-GH-3H] ³⁻
537.1	0.12	537.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻
526.4	0.31	526.4	[w ₅ -2H] ³⁻
520.5	0.05	520.4	[x ₅ -4H] ³⁻
512.9	0.05	512.9	[w ₅ -2H] ³⁻
509.1	0.23	509.1	[a ₄ -B ₄ H-3H] ²⁻
506.1	0.17	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
490.1	0.18	490.1	[A _{nt} +s+p-H] ⁻
488.0	0.05	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.1	0.14	481.0	[T _{nt} +s+p-H] ⁻
473.5	0.44	473.6	[w ₃ -H] ²⁻
472.1	0.22	472.0	[A _{nt} +s+p-H ₂ O-H] ⁻
463.1	0.04	463.0	[T _{nt} +s+p-H ₂ O-H] ⁻
426.1	0.16	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
410.1	0.06	410.1	[A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.1	0.12	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
401.1	0.02	401.1	[T _{nt} +s-H] ⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-H] ⁻
392.1	0.11	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.1	0.17	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.1	0.06	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
368.1	0.05	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
357.1	0.02	357.1	[a ₃ -B ₃ H-3H] ²⁻
346.1	0.83	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
330.1	0.02	330.2	[A _{nt} +H ₂ O-H] ⁻
328.1	0.17	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.1	0.34	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻
312.1	0.16	312.0	[A _{nt} -H] ⁻
306.0	0.05	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.1	0.74	303.0	[T _{nt} -H] ⁻
288.1	0.20	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.1	0.04	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.1	0.08	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.02	239.0	[psp-H ₂ O-H] ⁻
230.1	0.05	230.1	[sG-H ₂ O-H] ⁻
214.1	0.16	214.0	[sA-H ₂ O-H] ⁻
195.0	0.21	195.0	[ps+H ₂ O-H] ⁻
177.0	0.84	177.0	[ps-H] ⁻
159.0	0.11	159.0	[ps-H ₂ O-H] ⁻

150.0	0.32	150.0	G ⁻
134.0	0.37	134.0	A ⁻
125.0	0.58	125.0	T ⁻
110.0	0.83	110.0	C ⁻
97.0	0.50	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	2.18	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), Mr =98.0368 Da

p = PO₃H, Mr = 79.9663 Da

B_{nt} denoted a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-4H]⁴⁻ ion of the oligonucleotide 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1172.0	0.04	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GTAC)-p+H ₂ O-H] ⁻
1154.0	0.02	1154.2	[a ₄ -2H] ⁻ and/or [z ₄ -2H] ⁻ and/or [(GTAC)-p-H] ⁻
1043.0	0.03	1043.2	[a ₄ -CH-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(GTA)+s-H] ⁻
		1043.1	[(GTA)+p+H ₂ O-H] ⁻
1018.9	0.18	1019.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -AH-2H] ⁻
	and/or	1019.1	[(GTAC)-AH-p-H] ⁻
948.0	0.09	948.2	w ₃ ⁻
868.0	0.10	868.2	y ₃ ⁻
841.2	0.07	841.2	[a ₃ -2H] ⁻
810.2	0.07	810.1	[c ₃ -CH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
750.0	0.30	750.1	[y ₅ -H] ⁻
730.1	0.05	730.1	[b ₅ -H] ²⁻
	and/or	730.1	[(GT)+s-H] ⁻
	and/or	730.0	[(GT)+p+H ₂ O-H] ⁻
715.2	0.24	715.0	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -AH-2H] ⁻
665.6	0.09	665.6	[a ₅ -B ₅ H-3H] ²⁻ and/or [z ₅ -GH-3H] ²⁻
635.1	0.79	635.1	w ₂ ⁻ and/or d ₂ ⁻
632.2	0.09	632.0	[(GT)-H] ⁻
625.4	0.23	625.6	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(GTAC)+H ₂ O-2H] ²⁻
617.0	0.13	617.1	[x ₂ -2H] ⁻ and/or [c ₂ -2H] ⁻
559.1	0.35	559.1	[M-C-4H] ³⁻
554.1	0.24	554.1	[M-T-4H] ³⁻
551.1	1.21	551.1	[M-A-4H] ³⁻
546.1	0.26	546.1	[M-G-4H] ³⁻
537.1	0.11	537.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻
526.4	0.42	526.4	[w ₅ -2H] ³⁻
512.8	0.08	512.9	[d ₅ -2H] ³⁻
509.0	1.22	509.1	[a ₄ -B ₄ H-3H] ²⁻ and/or [z ₄ -AH-3H] ²⁻
	and/or	509.0	[(GTAC)-AH-p-2H] ²⁻
506.1	0.28	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
490.1	0.17	490.1	[A _{nt} +s+p-H] ⁻
481.1	0.27	481.0	[T _{nt} +s+p-H] ⁻
473.6	0.41	473.6	[w ₃ -H] ²⁻
463.1	0.17	463.0	[T _{nt} +s+p-H ₂ O-H] ⁻
446.8	100	446.8	[M-4H]⁴⁻
426.0	0.13	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
419.1	0.08	419.1	[M-CH-4H] ⁴⁻
416.8	0.16	416.7	[w ₄ -2H] ³⁻
413.1	0.05	413.1	[M-AH-4H] ⁴⁻
410.1	0.09	410.1	[A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.1	0.16	408.1	[G _{nt} +s-H ₂ O-H] ⁻

	and/or	408.0	[G _{nt} +p-H] ⁻
401.1	0.03	401.1	[T _{nt} +s-H] ⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-H] ⁻
394.6	0.04	394.6	[w ₅ -3H] ⁴⁻
392.1	0.10	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.1	0.25	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.1	0.12	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
368.1	0.05	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
357.1	0.05	357.1	[a ₃ -B ₃ H-3H] ²⁻
346.1	1.03	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
330.1	0.04	330.2	[A _{nt} +H ₂ O-H] ⁻
328.1	0.25	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.0	1.77	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻
315.4	0.69	315.4	[w ₃ -2H] ³⁻
312.1	0.31	312.0	[A _{nt} -H] ⁻
306.0	0.06	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.1	1.48	303.0	[T _{nt} -H] ⁻
288.1	0.51	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.1	0.09	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
266.0	0.04	266.0	y ₁ ⁻ and/or [sG+H ₂ O-H] ⁻
257.0	0.12	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.03	239.0	[psp-H ₂ O-H] ⁻
230.1	0.08	230.1	[sG-H ₂ O-H] ⁻
214.1	0.03	214.0	[sA-H ₂ O-H] ⁻
195.0	0.22	195.0	[ps+H ₂ O-H] ⁻
177.0	1.50	177.0	[ps-H] ⁻
172.6	0.40	172.5	[w ₁ -H] ²⁻
159.0	0.12	159.0	[ps-H ₂ O-H] ⁻
150.1	0.76	150.0	G ⁻
134.1	1.79	134.0	A ⁻
125.1	0.88	125.0	T ⁻
110.1	1.73	110.0	C ⁻
97.0	0.96	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	4.93	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.24	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), Mr =98.0368 Da

p = PO₃H, Mr = 79.9663 Da

B_{nt} denoted a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-5H]⁵⁻ ion of the oligonucleotide 5'-d(CGTACG)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
949.0	0.05	948.2	w ₃ ⁻
869.1	0.06	868.2	y ₃ ⁻
842.4	0.05	841.2	[a ₃ -2H] ⁻
810.7	0.05	810.1	[c ₃ -CH-2H] ⁻ and/or [(GT)+s+p-H] ⁻
795.2	0.04	795.1	[(CG)+s+p-H] ⁻
751.0	0.04	750.1	[y ₅ -H] ²⁻
715.7	0.23	715.0	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -AH-2H] ⁻
666.4	0.22	665.6	[a ₅ -B ₅ H-3H] ²⁻ and/or [z ₅ -GH-3H] ²⁻

635.6	0.23	635.1	w ₂ ⁻ and/or d ₂ ⁻
626.2	0.05	625.6	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(GTAC)+H ₂ O-2H] ²⁻
617.5	0.07	617.1	[x ₂ -2H] ⁻ and/or [c ₂ -2H] ⁻
537.5	0.06	537.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻
526.9	0.08	526.4	[w ₅ -2H] ³⁻
509.6	0.20	509.1	[a ₄ -B ₄ H-3H] ²⁻ and/or [z ₄ -AH-3H] ²⁻
	and/or	509.0	[(GTAC)-AH-p-2H] ²⁻
506.1	0.14	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
490.3	0.08	490.1	[A _{nt} +s+p-H] ⁻
481.3	0.18	481.0	[T _{nt} +s+p-H] ⁻
474.0	0.18	473.6	[w ₃ -H] ²⁻
426.0	0.11	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
419.5	0.74	419.1	[M-C ⁻ -5H] ⁴⁻
416.8	0.51	416.7	[w ₄ -2H] ³⁻
415.3	0.46	415.3	[M-T ⁻ -5H] ⁴⁻
413.1	0.43	413.1	[M-A ⁻ -5H] ⁴⁻
409.5	0.86	409.1	[M-G ⁻ -5H] ⁴⁻
394.9	0.23	394.6	[w ₅ -3H] ⁴⁻
386.2	0.23	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
357.3	100	357.3	[M-5H]⁵⁻
346.3	0.59	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.3	0.17	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.0	0.66	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻
315.7	1.47	315.4	[w ₃ -2H] ³⁻
312.3	0.47	312.0	[A _{nt} -H] ⁻
306.0	0.06	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
303.3	0.59	303.0	[T _{nt} -H] ⁻
288.3	0.23	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
283.4	0.23	283.5	[G _{nt} +p+s+p-H ₂ O-2H] ²⁻
275.3	0.12	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.2	0.10	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
230.2	0.07	230.1	[sG-H ₂ O-H] ⁻
211.3	0.28	211.0	[w ₂ -2H] ³⁻
195.2	0.16	195.0	[ps+H ₂ O-H] ⁻
177.2	1.49	177.0	[ps-H] ⁻
172.7	0.61	172.5	[w ₁ -H] ²⁻
159.1	0.09	159.0	[ps-H ₂ O-H] ⁻
150.2	0.58	150.0	G ⁻
134.2	0.76	134.0	A ⁻
125.2	0.58	125.0	T ⁻
110.2	1.08	110.0	C ⁻
97.1	0.60	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	3.19	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.34	42.0	CNO ⁻
26.0	0.12	26.0	CN ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), Mr = 98.0368 Da

p = PO₃H, Mr = 79.9663 Da

B_{nt} denoted a mononucleotide which may be either psB and/or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) and/or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-3H]^3-$ ion of the oligonucleotide 5'-d(ATGCAT)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1476.3	0.10	1476.3	y_5^-
1332.4	0.16	1332.2	$[a_5-B_5H-2H]^-$ and/or $[z_5-AH-2H]^-$ and/or $[(TCGA)+s-H]^-$
	and/or	1332.2	$[(TCGA)+p+H_2O-H]^-$
1252.0	0.12	1252.2	w_4^-
1234.0	0.07	1234.2	$[x_4-2H]^-$ and/or $[c_4-2H]^-$ and/or $[(TCGA)-H]^-$
1172.1	0.39	1172.2	y_4^- and/or b_4^- and/or $[(TCGA)-p+H_2O-H]^-$
1154.4	0.11	1154.2	$[a_4-2H]^-$ and/or $[z_4-2H]^-$ and/or $[(TCGA)-p-H]^-$
1098.9	0.22	1099.1	$[(TGC)+p+s-H]^-$
1043.2	0.10	1043.2	$[a_4-B_4H-2H]^-$ and/or $[z_4-CH-2H]^-$ and/or $[(TCGA)-CH-H]^-$
1019.2	0.16	1019.2	$[(TGC)+s-H]^-$
	and/or	1019.1	$[(TGC)+p+H_2O-H]^-$
1001.2	0.08	1001.2	$[(TGC)+s-H_2O-H]^-$
	and/or	1001.1	$[(TGC)+p-H]^-$
939.2	0.08	939.1	$[(TGC)+H_2O-H]^-$
923.0	0.54	923.2	w_3^-
905.0	0.11	905.1	$[x_3-2H]^-$
883.2	0.07	883.2	b_3^-
865.1	0.08	865.2	$[a_3-2H]^-$
843.0	0.40	843.2	y_3^-
831.2	0.35	831.1	$[M-T--2H]^2-$
826.8	0.98	826.6	$[M-A--2H]^2-$
818.8	0.15	818.6	$[M-G--2H]^2-$
810.1	0.28	810.1	$[c_3-AH-2H]^-$ and/or $[(TG)+s+p-H]^-$
795.1	0.49	795.1	$[(CG)+p+s-H]^-$
777.6	6.06	777.6	$[w_5-H]^2-$
737.7	0.22	737.6	$[y_5-H]^2-$
730.1	0.38	730.1	$[(TG)+s-H]^-$
	and/or	730.1	$[(TG)+p+H_2O-H]^-$
714.3	1.02	714.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-CH-2H]^-$
696.9	0.22	697.1	$[(CG)+s-H_2O-H]^-$
	and/or	697.1	$[(CG)+p-H]^-$
665.6	1.44	665.6	$[a_5-B_5H-3H]^2-$ and/or $[z_5-TH-3H]^2-$
650.0	0.30	650.1	$[(TG)+H_2O-H]^-$
634.1	2.93	634.1	w_2^- and/or d_2^-
625.6	1.08	625.6	$[w_4-H]^2-$
616.1	1.02	616.1	$[c_2-2H]^-$ and/or $[x_2-2H]^-$
601.1	0.73	601.1	$[(CA)-H]^-$
595.8	100	595.8	$[M-3H]^3-$
558.7	0.16	558.8	$[M-CH-3H]^3-$
553.7	0.77	553.8	$[M-TH-3H]^3-$
550.8	3.15	550.8	$[M-AH-3H]^3-$
545.4	0.24	545.4	$[M-GH-3H]^3-$
536.1	0.16	536.1	$[a_2-2H]^-$ and/or $[z_2-2H]^-$
521.1	0.14	521.1	$[a_4-B_4H-3H]^2-$ and/or $[z_4-CH-3H]^2-$
518.1	0.73	518.1	$[w_5-2H]^3-$
506.1	0.64	506.0	$[G_{nt}+s+p-H]^-$
490.0	0.18	490.1	$[A_{nt}+s+p-H]^-$ and/or $[c_2-TH-2H]^-$ and/or $[x_2-TH-2H]^-$
488.0	0.40	488.0	$[G_{nt}+s+p-H_2O-H]^-$
481.0	0.89	481.0	$[T_{nt}+s+p-H]^-$ and/or $[c_2-AH-2H]^-$ and/or $[x_2-AH-2H]^-$
466.1	1.77	466.0	$[C_{nt}+s+p-H]^-$
461.1	0.65	461.1	$[w_3-H]^2-$
448.1	0.29	448.0	$[C_{nt}+s+p-H_2O-H]^-$
426.0	0.40	426.1	$[G_{nt}+s-H]^-$
	and/or	426.1	$[G_{nt}+p+H_2O-H]^-$
410.0	0.73	410.1	$[a_2-B_2H-2H]^-$ and/or $[A_{nt}+s-H]^-$
		410.0	$[A_{nt}+p+H_2O-H]^-$

408.0	1.26	408.1	[G _{nt} +s- H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
401.0	0.56	401.1	[T _{nt} +s-H] ⁻ and/or [a ₂ -AH-2H] ⁻ and/or [z ₂ -AH-2H] ⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-H] ⁻
392.0	0.32	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.0	0.29	386.1	[C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.0	0.65	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
368.0	1.45	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.0	0.27	346.1	[G _{nt} +H ₂ O-H] ⁻
330.1	0.29	330.1	[A _{nt} +H ₂ O-H] ⁻
328.1	3.87	328.0	[G _{nt} -H] ⁻
321.1	7.66	321.0	w ₁ ⁻ and/or [T _{nt} + H ₂ O-H] ⁻
312.1	0.82	312.0	[c ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
303.0	1.73	303.0	[x ₁ -2H] ⁻ and/or [T _{nt} -H] ⁻
288.1	1.94	288.0	[C _{nt} -H] ⁻
275.0	0.47	275.1	[sps-H] ⁻
	and/or	275.0	[psp+ H ₂ O-H] ⁻
257.0	1.16	257.1	[sps- H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.29	239.0	[psp- H ₂ O-H] ⁻
230.0	0.37	230.1	[sG-H ₂ O-H] ⁻
214.1	0.40	214.0	[sA-H ₂ O-H] ⁻
195.1	3.93	195.0	[ps+ H ₂ O-H] ⁻
177.1	8.65	177.0	[ps-H] ⁻
159.0	1.44	159.0	[ps-H ₂ O-H] ⁻
150.1	1.79	150.0	G ⁻
134.1	16.1	134.0	A ⁻
125.1	4.68	125.0	T ⁻
110.0	0.32	110.0	C ⁻
97.0	4.52	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	18.68	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.35	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B)₁..B_n] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-4H]⁴⁻ ion of the oligonucleotide 5'-d(ATGCAT)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1172.4	0.36	1172.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(TCGA)-p+H ₂ O -H] ⁻
1098.9	0.17	1099.1	[(TGC)+p+s-H] ⁻
1019.5	0.16	1019.2	[(TGC)+s-H] ⁻
	and/or	1019.1	[(TGC)+p+H ₂ O-H] ⁻
923.5	0.20	923.2	w ₃ ⁻
843.1	0.11	843.2	y ₃ ⁻
810.3	0.12	810.1	[c ₃ -AH-2H] ⁻ and/or [(TG)+s+p-H] ⁻
795.3	0.31	795.1	[(CG)+p+s-H] ⁻
737.6	0.63	737.6	[y ₅ -H] ²⁻
715.0	0.44	715.1	[(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p+H ₂ O -H] ⁻
697.1	0.13	697.1	[(CG)+s-H ₂ O-H] ⁻
	and/or	697.1	[(CG)+p-H] ⁻
665.8	0.56	665.6	[a ₅ -B ₅ H-3H] ²⁻ and/or [z ₅ -TH-3H] ²⁻
650.0	0.09	650.1	[(TG)+H ₂ O-H] ⁻
633.9	0.57	634.1	w ₂ ⁻ and/or d ₂ ⁻

625.7	0.99	625.6	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻
617.1	0.37	616.6	[c ₂ -3H] ²⁻ and/or [x ₂ -3H] ²⁻
601.1	0.73	601.1	[(CA)-H] ⁻
558.8	0.15	558.8	[M-C ⁻ -3H] ³⁻
553.7	3.38	553.8	[M-T ⁻ -3H] ³⁻
550.9	5.73	550.8	[M-A ⁻ -3H] ³⁻
521.1	0.31	521.1	[a ₄ -B ₄ H-3H] ²⁻ and/or [z ₄ -CH-3H] ²⁻
518.1	4.26	518.1	[w ₅ -2H] ³⁻
506.1	0.57	506.0	[G _{nt} +s+p-H] ⁻
490.0	0.05	490.1	[A _{nt} +s+p-H] ⁻ and/or [c ₂ -TH-2H] ⁻ and/or [x ₂ -TH-2H] ⁻
488.1	0.06	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.0	0.59	481.0	[T _{nt} +s+p-H] ⁻ and/or [c ₂ -AH-2H] ⁻ and/or [x ₂ -AH-2H] ⁻
466.1	1.03	466.0	[C _{nt} +s+p-H] ⁻
461.1	0.31	461.1	[w ₃ -H] ²⁻
446.6	100	446.6	[M-4H]⁴⁻
426.1	0.31	426.1	[G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
416.7	0.24	416.7	[w ₄ -2H] ³⁻ and/or [d ₄ -2H] ³⁻
412.7	0.12	412.8	[M-AH-4H] ⁴⁻
410.1	0.79	410.1	[a ₂ -B ₂ H-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	0.59	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
401.0	0.26	401.1	[T _{nt} +s-H] ⁻ and/or [a ₂ -AH-2H] ⁻ and/or [z ₂ -AH-2H] ⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-H] ⁻
392.0	0.19	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.0	0.15	386.1	[C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
383.0	0.37	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
368.0	0.73	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.0	0.12	346.1	[G _{nt} +H ₂ O-H] ⁻
330.1	0.15	330.1	[A _{nt} +H ₂ O-H] ⁻
328.1	2.20	328.0	[G _{nt} -H] ⁻
321.1	5.55	321.0	w ₁ ⁻ and/or [T _{nt} +H ₂ O-H] ⁻
316.5	0.35	316.5	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻
312.1	0.62	312.0	[c ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
303.0	1.18	303.0	[x ₁ -2H] ⁻ and/or [T _{nt} -H] ⁻
288.1	1.54	288.0	[C _{nt} -H] ⁻
275.0	0.44	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.0	0.62	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.18	239.0	[psp-H ₂ O-H] ⁻
230.0	0.22	230.1	[sG-H ₂ O-H] ⁻
214.1	0.10	214.0	[sA-H ₂ O-H] ⁻
195.1	2.20	195.0	[ps+H ₂ O-H] ⁻
177.1	4.56	177.0	[ps-H] ⁻
159.0	0.73	159.0	[ps-H ₂ O-H] ⁻
150.1	1.20	150.0	G ⁻
134.1	12.93	134.0	A ⁻
125.1	4.42	125.0	T ⁻
110.1	0.44	110.0	C ⁻
97.0	3.31	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	16.90	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.66	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da; p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-5H]^{5-}$ ion of the oligonucleotide 5'-d(ATGCAT)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1098.4	0.03	1099.1	[(TGC)+p+s-H] ⁻
1046.1	0.03	1046.2	[(TGC)+s+H ₂ O -H] ⁻
1018.8	0.07	1019.2	[(TGC)+s-H] ⁻
	and/or	1019.1	[(TGC)+p+H ₂ O-H] ⁻
948.0	0.05	948.1	[(TGC)+H ₂ O-H] ⁻
923.5	0.03	923.2	w ₃ ⁻
921.1	0.04	921.1	[(TGC)-H] ⁻
843.1	0.02	843.2	y ₃ ⁻
809.7	0.05	810.1	[c ₃ -AH-2H] ⁻ and/or [(TG)+s+p-H] ⁻
794.9	0.23	795.1	[(CG)+p+s-H] ⁻
777.1	0.02	777.1	[(CG)+p+s-H ₂ O-H] ⁻
742.0	0.03	742.0	[b ₅ -H] ²⁻
733.1	0.02	733.1	[a ₅ -3H] ²⁻
729.9	0.04	730.1	[(TG)+s-H] ⁻
	and/or	730.1	[(TG)+p+H ₂ O -H] ⁻
714.8	0.15	715.1	[(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p+H ₂ O -H] ⁻
696.8	0.06	697.1	[(CG)+s-H ₂ O-H] ⁻
	and/or	697.1	[(CG)+p-H] ⁻
674.4	0.04	674.4	[b ₅ -B ₅ H-H] ²⁻ and/or [y ₅ -TH-H] ²⁻
665.4	0.89	665.6	[a ₅ -B ₅ H-3H] ²⁻ and/or [z ₅ -TH-3H] ²⁻
649.9	0.02	650.1	[(TG)+H ₂ O-H] ⁻
633.9	0.09	634.1	w ₂ ⁻ and/or d ₂ ⁻
625.4	0.25	625.6	[w ₄ -H] ²⁻
616.8	0.12	616.6	[c ₄ -3H] ²⁻ and/or [x ₄ -3H] ²⁻
585.4	0.19	585.6	[y ₄ -H] ²⁻ and/or [b ₄ -H] ²⁻
576.4	0.04	576.4	[a ₄ -3H] ²⁻ and/or [z ₄ -3H] ²⁻
548.9	0.24	549.1	[(TGC)+p+s-2H] ²⁻
539.6	0.24	539.6	[(TGC)+p+s-H ₂ O-2H] ²⁻
536.3	0.58	536.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻
521.0	0.18	521.1	[a ₄ -B ₄ H-3H] ²⁻ and/or [z ₄ -CH-3H] ²⁻
508.6	0.49	508.7	[M-A ⁻ -T ⁻ -5H] ²⁻
506.1	0.39	506.0	[G _{nt} +s+p-H] ⁻
503.6	0.31	503.7	[M-A ⁻ -A ⁻ -5H] ²⁻
488.1	0.07	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.0	0.21	481.0	[T _{nt} +s+p-H] ⁻ and/or [c ₂ -AH-2H] ⁻ and/or [x ₂ -AH-2H] ⁻
475.9	0.41	476.1	[(TGCA)+p+s-H ₂ O-2H] ²⁻
465.9	0.38	466.0	[C _{nt} +s+p-H] ⁻
461.1	0.07	461.1	[w ₃ -H] ²⁻
443.2	0.50	443.4	[a ₅ -B ₅ H-4H] ³⁻ and/or [z ₅ -TH-4H] ³⁻
	and/or	443.4	[(TGCA)+p+H ₂ O-2H] ²⁻
435.7	1.75	435.6	[(TG)+p+s+p-H ₂ O-2H] ²⁻
425.9	0.17	426.1	[G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
418.8	0.59	418.8	[M-C ⁻ -5H] ⁴⁻
415.1	1.55	415.1	[M-T ⁻ -5H] ⁴⁻
412.7	3.17	412.8	[M-A ⁻ -5H] ⁴⁻
410.1	0.76	410.1	[a ₂ -B ₂ H-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.8	0.40	408.8	[M-GH-4H] ⁴⁻
390.5	0.73	390.1	[y ₄ -2H] ³⁻
383.0	0.12	383.1	[T _{nt} +s-H ₂ O-H] ⁻
	and/or	383.0	[T _{nt} +p-H] ⁻
367.9	0.35	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
357.1	100	357.1	[M-5H]⁵⁻
330.1	0.12	330.1	[A _{nt} +H ₂ O-H] ⁻
328.1	0.90	328.0	[G _{nt} -H] ⁻

321.0	1.51	321.0	w ₁ ⁻ and/or [T _{nt} + H ₂ O-H] ⁻
316.5	0.14	316.5	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻
312.0	0.25	312.0	[c ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
307.5	0.05	307.5	[w ₃ -2H] ³⁻
303.0	0.39	303.0	[x ₁ -2H] ⁻ and/or [T _{nt} -H] ⁻
288.0	0.77	288.0	[C _{nt} -H] ⁻
275.0	0.43	275.1	[sps-H] ⁻
	and/or	275.0	[psp+ H ₂ O-H] ⁻
257.0	0.31	257.1	[sps- H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.07	239.0	[psp- H ₂ O-H] ⁻
230.0	0.14	230.1	[sG-H ₂ O-H] ⁻
214.0	0.05	214.0	[sA-H ₂ O-H] ⁻
211.3	0.04	211.3	[w ₂ -2H] ³⁻ and/or [d ₂ -2H] ³⁻
204.5	0.09	204.5	[a ₂ -B ₂ H-3H] ²⁻
195.0	0.66	195.0	[ps+ H ₂ O-H] ⁻
177.0	1.67	177.0	[ps-H] ⁻
160.0	0.54	160.0	[w ₁ -H] ²⁻
159.0	0.73	159.0	[ps-H ₂ O-H] ⁻
151.0	0.19	151.0	[x ₁ -3H] ²⁻
150.1	0.51	150.0	G ⁻
134.1	4.75	134.0	A ⁻
125.1	1.60	125.0	T ⁻
110.1	0.21	110.0	C ⁻
97.0	1.58	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	4.86	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	1.16	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-3H]³⁻ ion of the oligonucleotide 5'-d(TCACGA)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).</i>	<i>Assignment</i>
1470.4	0.17	1470.2	y ₅ ⁻
1068.1	0.14	1068.1	[(CAC)+s+p-H] ⁻
989.2	0.13	988.2	[(CAC)+s-H] ⁻
	and/or	988.1	[(CAC)+p+H ₂ O-H] ⁻
948.7	0.81	948.2	w ₃ ⁻ and/or [(ACG)+H ₂ O-H] ⁻
868.3	0.60	868.2	y ₃ ⁻ and/or [(ACG)-p+H ₂ O-H] ⁻
823.7	1.06	823.6	[M-T ⁻ -3H] ²⁻
819.1	0.51	819.1	[M-A ⁻ -3H] ²⁻
811.1	0.17	811.1	[M-G ⁻ -3H] ²⁻
779.0	0.79	779.1	[c ₃ -TH-2H] ⁻ and/or [x ₃ -GH-2H] ⁻ and/or [(CA)+s+p-H] ⁻ and/or [(AC)+s+p-H] ⁻
774.7	6.70	774.6	[w ₅ -H] ²⁻
734.6	0.34	734.6	[y ₅ -H] ²⁻
707.2	0.39	707.1	[(CACG)+s+p+H ₂ O-2H] ²⁻
699.6	0.64	699.1	[a ₃ -TH-2H] ⁻ and/or [z ₃ -GH-2H] ⁻ and/or [(CA)+s-H] ⁻ and/or [(AC)+s-H] ⁻
	and/or	699.1	[(CA)+p+H ₂ O -H] ⁻ and/or [(AC)+p+H ₂ O -H] ⁻
690.3	1.01	690.1	[a ₃ -B ₃ H-2H] ⁻
681.2	0.27	681.1	[(CA)+s-H ₂ O-H] ⁻
	and/or	681.1	[(CA)+p-H] ⁻
659.3	4.04	659.1	w ₂ ⁻
641.1	0.86	641.1	[x ₂ -2H] ⁻

630.4	0.54	630.1	[w ₄ -H] ²⁻
619.2	0.54	619.1	[(CA)+H ₂ O-H] ⁻
617.1	0.40	617.1	[(CG)-H] ⁻
601.1	0.67	601.1	[(CA)-H] ⁻
590.7	100	590.8	[M-3H]³⁻
579.2	0.81	579.1	y ₂ ⁻
561.2	0.27	561.1	[z ₂ -2H] ⁻
553.7	0.18	553.8	[M-CH-3H] ³⁻
548.7	0.36	547.8	[M-TH-3H] ³⁻
545.8	0.90	545.8	[M-AH-3H] ³⁻
540.4	0.12	540.4	[M-GH-3H] ³⁻
524.2	0.27	524.4	[(CACG)+p+s+p+s-H] ⁻
516.0	0.36	516.1	[w ₅ -2H] ³⁻
506.2	0.25	506.0	[G _{nt} +s+p-H] ⁻
490.2	0.45	490.1	[A _{nt} +s+p-H] ⁻ and/or [x ₂ -GH-2H] ⁻
473.6	1.80	473.6	[w ₃ -H] ²⁻
472.2	0.68	472.0	[A _{nt} +s+p-H ₂ O-H] ⁻
466.1	3.38	466.0	[C _{nt} +s+p-H] ⁻ and/or [c ₂ -TH-2H] ⁻
448.2	0.45	448.0	[C _{nt} +s+p-H ₂ O-H] ⁻
426.1	0.40	426.1	[G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
410.0	1.35	410.1	[z ₂ -GH-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	1.21	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
401.1	0.14	401.1	[a ₂ -B ₂ H-2H] ⁻
392.0	1.94	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.2	0.58	386.1	[a ₂ -TH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.1	3.92	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
345.9	0.09	346.1	[G _{nt} +H ₂ O-H] ⁻
330.1	6.58	330.1	w ₁ ⁻ and/or [A _{nt} +H ₂ O-H] ⁻
328.1	2.25	328.0	[G _{nt} -H] ⁻
321.0	0.10	321.0	[T _{nt} +H ₂ O-H] ⁻
312.1	3.38	312.0	[x ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
306.1	0.68	306.0	[C _{nt} +H ₂ O-H] ⁻
288.1	5.31	288.0	[C _{nt} -H] ⁻
275.1	0.81	275.1	[sps-H] ⁻
	and/or	275.0	[psp+H ₂ O-H] ⁻
257.1	2.57	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.90	239.0	[psp-H ₂ O-H] ⁻
230.1	1.31	230.1	[sG-H ₂ O-H] ⁻
214.1	0.68	214.0	[sA-H ₂ O-H] ⁻
195.1	6.76	195.0	[ps+H ₂ O-H] ⁻
177.0	15.32	177.0	[ps-H] ⁻
159.0	4.50	159.0	[ps-H ₂ O-H] ⁻
150.1	7.21	150.0	G ⁻
134.1	11.22	134.0	A ⁻
125.0	17.57	125.0	T ⁻
110.0	2.70	110.0	C ⁻
97.0	11.94	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	48.20	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	2.25	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-4H]^{4-}$ ion of the oligonucleotide 5'-d(TCACGA)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1183.0	0.12	1184.1	[(CAC)+s+p+H ₂ O -H] ⁻
1046.0	0.06	1046.2	[(ACG)+s+H ₂ O -H] ⁻
948.3	0.63	948.2	w ₃ ⁻ and/or [(ACG)+H ₂ O-H] ⁻
868.4	0.23	868.2	y ₃ ⁻ and/or [(ACG)-p+H ₂ O-H] ⁻
779.2	0.12	779.1	[c ₃ -TH-2H] ⁻ and/or [x ₃ -GH-2H] ⁻ and/or [(CA)+s+p-H] ⁻ and/or [(AC)+s+p-H] ⁻
734.6	1.42	734.6	[y ₅ -H] ²⁻
707.2	0.25	707.1	[(CACG)+s+p+H ₂ O-2H] ²⁻
699.1	0.20	699.1	[a ₃ -TH-2H] ⁻ and/or [z ₃ -GH-2H] ⁻ and/or [(CA)+s-H] ⁻ and/or [(AC)+s-H] ⁻
	and/or	699.1	[(CA)+p+H ₂ O -H] ⁻ and/or [(AC)+p+H ₂ O -H] ⁻
690.1	0.08	690.1	[a ₃ -B ₃ H-2H] ⁻
659.4	0.80	659.1	w ₂ ⁻
641.1	0.19	641.1	[x ₂ -2H] ⁻
630.4	0.70	630.1	[w ₄ -H] ²⁻
619.1	0.12	619.1	[(CA)+H ₂ O -H] ⁻
601.1	0.13	601.1	[(CA)-H] ⁻
579.1	0.32	579.1	y ₂ ⁻
561.2	0.27	561.1	[z ₂ -2H] ⁻
553.9	0.56	553.8	[M-C ⁻ -4H] ³⁻
549.0	4.16	547.8	[M-T ⁻ -4H] ³⁻
545.8	0.64	545.8	[M-A ⁻ -4H] ³⁻
524.2	0.29	524.4	[(CACG)+p+s+p+s-H] ⁻
516.1	5.73	516.1	[w ₅ -2H] ³⁻
490.1	0.23	490.1	[A _{nt} +s+p-H] ⁻ and/or [x ₂ -GH-2H] ⁻
473.8	1.29	473.6	[w ₃ -H] ²⁻ and/or [(ACG)+H ₂ O-2H] ²⁻
466.1	2.73	466.0	[C _{nt} +s+p-H] ⁻ and/or [c ₂ -TH-2H] ⁻
448.2	0.35	448.0	[C _{nt} +s+p- H ₂ O-H] ⁻
442.8	100	442.8	[M-4H]⁴⁻
426.0	0.06	426.1	[G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
419.6	0.08	419.7	[w ₄ -2H] ³⁻
411.2	0.06	411.3	[M-TH-4H] ⁴⁻
410.0	0.35	410.1	[z ₂ -GH-2H] ⁻ and/or [A _{nt} +s-H] ⁻
		410.0	[A _{nt} +p+H ₂ O-H] ⁻
408.0	0.10	408.1	[G _{nt} +s- H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
401.1	0.10	401.1	[a ₂ -B ₂ H-2H] ⁻ and/or [T _{nt} +s-2H] ²⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-2H] ²⁻
392.0	0.34	392.1	[A _{nt} +s-H ₂ O-H] ⁻
	and/or	392.0	[A _{nt} +p-H] ⁻
386.2	0.40	386.1	[a ₂ -TH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.0	1.25	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
330.1	3.98	330.1	w ₁ ⁻ and/or [A _{nt} +H ₂ O-H] ⁻
329.1	0.56	329.1	[w ₂ -H] ²⁻
328.1	0.88	328.0	[G _{nt} -H] ⁻
315.4	0.27	315.4	[w ₃ -2H] ³⁻ and/or [(ACG)+H ₂ O-3H] ³⁻
312.1	1.78	312.0	[x ₁ -2H] ⁻ and/or [A _{nt} -H] ⁻
306.1	0.13	306.0	[C _{nt} +H ₂ O-H] ⁻
288.1	2.96	288.0	[C _{nt} -H] ⁻
275.1	0.29	275.1	[sps-H] ⁻
	and/or	275.0	[psp+ H ₂ O-H] ⁻
257.1	0.48	257.1	[sps- H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.16	239.0	[psp- H ₂ O-H] ⁻
230.1	0.19	230.1	[sG-H ₂ O-H] ⁻

214.0	0.13	214.0	[sA-H ₂ O-H] ⁻
195.0	1.61	195.0	[ps+ H ₂ O-H] ⁻
177.1	4.64	177.0	[ps-H] ⁻
159.0	0.56	159.0	[ps-H ₂ O-H] ⁻
150.1	2.41	150.0	G ⁻
134.1	6.27	134.0	A ⁻
125.1	8.37	125.0	T ⁻
110.0	1.82	110.0	C ⁻
97.0	3.69	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	16.47	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	1.20	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-5H]⁵⁻ ion of the oligonucleotide 5'-d(TCACGA)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
775.7	3.93	776.1	[(TA)+s+p+H ₂ O-H] ⁻
770.1	3.28	770.1	[d ₅ -H] ²⁻
761.1	2.75	761.1	[(CA)+s+p+H ₂ O-H] ⁻ and/or [(AC)+s+p+H ₂ O-H] ⁻
721.7	3.34	721.1	[a ₅ -3H] ²⁻
715.1	3.28	715.1	[(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p+H ₂ O-H] ⁻
686.1	4.59	685.6	[c ₅ -GH-3H] ²⁻
666.8	4.59	667.1	[(CACG)+s+H ₂ O-H] ²⁻
659.1	3.94	659.1	w ₂ ⁻
619.3	3.08	619.1	[(CA)+H ₂ O-H] ⁻ and/or [(AC)+H ₂ O-H] ⁻
594.9	10.10	593.5	[(ACG)+s+p+s-H ₂ O-2H] ²⁻
	and/or	593.5	[(ACG)+p+s+p-2H] ²⁻
411.3	3.70	411.3	[M-T--5H] ⁴⁻
409.5	7.21	409.1	[M-A--5H] ⁴⁻
405.1	2.41	405.1	[M-G--5H] ⁴⁻
401.1	5.25	401.1	[a ₂ -B ₂ H-2H] ²⁻ and/or [T _{nt} +s-2H] ²⁻
	and/or	401.0	[T _{nt} +p+H ₂ O-2H] ²⁻
389.1	5.26	389.0	[(CA)+s+p-2H] ²⁻ and/or [(AC)+s+p-2H] ²⁻
354.1	59.44	354.1	[M-5H]⁵⁻
330.1	14.52	330.1	w ₁ ⁻ and/or [A _{nt} +H ₂ O-H] ⁻
329.1	6.45	329.1	[w ₂ -H] ²⁻
315.5	4.84	315.4	[w ₃ -2H] ³⁻ and/or [(ACG)+H ₂ O-3H] ³⁻
303.0	7.50	303.0	[T _{nt} -H] ⁻
309.1	4.84	309.0	[(CA)+H ₂ O-2H] ²⁻
300.0	3.06	300.0	[(CA)-2H] ²⁻
288.1	10.89	288.0	[C _{nt} -H] ⁻
275.1	5.81	275.1	[sps-H] ⁻
	and/or	275.0	[psp+ H ₂ O-H] ⁻
257.1	5.65	257.1	[sps- H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	7.66	239.0	[psp- H ₂ O-H] ⁻
219.1	3.87	219.0	[w ₂ -2H] ³⁻
210.1	6.45	210.0	[(CG)+H ₂ O-3H] ³⁻
200.1	5.56	200.1	[a ₂ -B ₂ H-3H] ²⁻ and/or
177.0	22.58	177.0	[ps-H] ⁻
150.2	11.13	150.0	G ⁻
134.1	28.39	134.0	A ⁻
125.1	8.47	125.0	T ⁻

110.1	17.34	110.0	C ⁻
97.0	18.15	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	100	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	14.35	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_n denotes a mononucleotide which may be either psB or sBp

[(B₁..B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-3H]³⁻ ion of the oligonucleotide 5'-d(CGGCCG)-3'.

<i>m/z</i>	<i>Rel. Int.(%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1253.1	0.04	1253.2	w ₄ ⁻ and/or d ₄ ⁻ and/or [(GGCC)+H ₂ O-H] ⁻
1173.4	0.26	1173.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GGCC)-p+H ₂ O-H] ⁻
1155.3	0.06	1155.2	[z ₄ -2H] ⁻ and/or [a ₄ -2H] ⁻ and/or [(GGCC)-p-H] ⁻
1044.2	0.21	1044.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -CH-2H] ⁻ and/or [(GGC)+s-H] ⁻
964.2	0.07	964.2	d ₃ ⁻ and/or [(GGC)+H ₂ O-H] ⁻
946.1	0.08	946.1	[c ₃ -2H] ⁻ and/or [(GGC)-H] ⁻
924.6	0.93	924.1	w ₃ ⁻ and/or [(GCC)+H ₂ O-H] ⁻
906.1	0.10	906.1	[x ₃ -2H] ⁻ and/or [(GCC)-H] ⁻
884.0	0.12	884.2	b ₃ ⁻ and/or [(GGc)-p+H ₂ O-H] ⁻
866.2	0.22	866.2	[a ₃ -2H] ⁻ and/or [(GGC)-p-H] ⁻
844.5	0.62	844.2	y ₃ ⁻ and/or [(GCC)-p+H ₂ O-H] ⁻
839.6	0.28	839.7	[M-C-3H] ²⁻
835.1	0.20	835.1	[c ₃ -CH-2H] ⁻ and/or [(GG)+s+p-H] ⁻
	and/or	835.0	[(CC)+p+s+p-H] ⁻
	and/or	835.1	[(CC)+s+p+s-H ₂ O -H] ⁻
819.4	0.27	819.7	[M-G-2H] ³⁻
795.1	0.31	795.1	[c ₃ -GH-2H] ⁻ and/or [x ₃ -CH-2H] ⁻ and/or [(CG)+s+p-H] ⁻
790.7	1.31	790.6	[w ₅ -H] ²⁻
777.0	0.07	777.1	[(GC)+s+p-H ₂ O-H] ⁻
773.3	0.10	773.1	[w ₃ -GH] ⁻ and/or [(CC)+s+p+H ₂ O -H] ⁻
	and/or	773.1	[b ₃ -CH] ⁻ and/or [(GG)+s+H ₂ O -H] ⁻
770.6	0.05	770.6	[d ₅ -H] ²⁻
755.2	0.44	755.1	[x ₃ -GH] ⁻ and/or [(CC)+s+p-H] ⁻
	and/or	755.1	[a ₃ -CH] ⁻ and/or [(GG)+s-H] ⁻
	and/or	755.1	[(GG)+p+H ₂ O -H] ⁻
750.9	1.24	750.6	[y ₅ -H] ²⁻
737.1	0.09	737.1	[(CC)+s+p-H ₂ O-H] ⁻
	and/or	737.1	[(GG)+s-H ₂ O-H] ⁻
	and/or	737.1	[(GG)+p-H] ⁻
715.3	2.07	715.1	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -CH-2H] ⁻ and/or [(GC)+s-H] ⁻
	and/or	715.1	[(GC)+p+H ₂ O -H] ⁻
697.2	0.13	697.1	[(GC)+p-H] ⁻
	and/or	697.1	[(GC)+s-H ₂ O-H] ⁻
675.2	0.41	675.1	[(CC)+p+ H ₂ O -H] ⁻
	and/or	675.1	[z ₃ -GH-2H] ⁻ and/or [(CC)+s-H] ⁻
666.5	0.55	666.1	[a ₅ -B ₅ H-3H] ²⁻
657.2	0.48	657.1	[(CC)+s-H ₂ O -H] ⁻
	and/or	657.1	[(CC)+p-H] ⁻
	and/or	657.1	[(GG)-H] ⁻
635.2	3.35	635.1	w ₂ ⁻ and/or d ₂ ⁻ and/or [(GC)+H ₂ O-H] ⁻
626.1	1.03	626.1	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(GGCC)+H ₂ O-2H] ²⁻
617.2	1.76	617.1	[c ₂ -2H] ⁻ and/or [x ₂ -2H] ⁻ and/or [(GC)-H] ⁻
596.4	100	596.4	[M-3H]³⁻
577.4	0.54	577.1	[(CC)-H] ⁻

559.4	0.90	559.4	[M-CH-3H] ³⁻
555.2	0.84	555.1	y ₂ ⁻ and/or b ₂ ⁻ and/or [(CG)-p+H ₂ O-H] ⁻
546.1	0.45	546.1	[M-GH-3H] ³⁻
537.3	0.54	537.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻ and/or [(GC)-p-H] ⁻
526.8	0.57	526.8	[w ₅ -2H] ³⁻
506.1	1.86	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
488.1	0.63	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
484.1	0.18	484.0	[C _{nt} +s+p+H ₂ O-H] ⁻
466.1	1.47	466.0	[C _{nt} +s+p-H] ⁻
461.5	1.17	461.6	[w ₃ -H] ²⁻ and/or [(GCC)+H ₂ O-2H] ²⁻
448.2	0.22	448.0	[C _{nt} +s+p-H ₂ O-H] ⁻
444.2	0.21	444.1	[G _{nt} +s+H ₂ O-H] ⁻
426.2	1.83	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
408.1	2.13	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
386.2	1.65	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.1	1.83	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.2	5.31	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	5.97	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.2	0.70	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻ and/or [(GC)+H ₂ O-2H] ²⁻
310.1	0.06	310.0	[G _{nt} -H ₂ O-H] ⁻
306.2	0.72	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
288.1	4.50	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.2	0.54	275.1	[sps-H] ⁻
	and/or	275.1	[psp+H ₂ O-H] ⁻
257.1	1.05	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.24	239.1	[psp-H ₂ O-H] ⁻
230.2	1.14	230.1	[sG-H ₂ O-H] ⁻
195.1	2.85	195.0	[ps+H ₂ O-H] ⁻
190.1	0.18	190.1	[sC-H ₂ O-H] ⁻
177.1	9.54	177.0	[ps-H] ⁻
159.0	1.80	159.0	[ps-H ₂ O-H] ⁻
150.2	8.67	150.0	G ⁻
110.1	8.31	110.0	C ⁻
97.0	6.33	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	32.10	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.48	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-4H]⁴⁻ ion of the oligonucleotide 5'-d(CGGCCG)-3'.

m/z	Rel. Int.(%)	m/z(calc).*	Assignment
1173.2	0.15	1173.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(GGCC)-p+H ₂ O-H] ⁻
924.6	0.08	924.1	w ₃ ⁻ and/or [(GCC)+H ₂ O-H] ⁻
866.3	0.05	866.2	[a ₃ -2H] ⁻ and/or [(GGC)-p-H] ⁻
843.8	0.09	844.2	y ₃ ⁻ and/or [(GCC)-p+H ₂ O-H] ⁻
795.5	0.05	795.1	[c ₃ -GH-2H] ⁻ and/or [x ₃ -CH-2H] ⁻ and/or [(CG)+s+p-H] ⁻
755.2	0.07	755.1	[x ₃ -GH] ⁻ and/or [(CC)+s+p-H] ⁻
	and/or	755.1	[a ₃ -CH] ⁻ and/or [(GG)+s-H] ⁻
	and/or	755.1	[(GG)+p+H ₂ O-H] ⁻
750.9	0.73	750.6	[y ₅ -H] ²⁻

715.5	0.25	715.1	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -CH-2H] ⁻ and/or [(GC)+s-H] ⁻
	and/or	715.1	[(GC)+p+H ₂ O-H] ⁻
675.2	0.10	675.1	[(CC)+p+H ₂ O-H] ⁻
	and/or	675.1	[z ₃ -GH-2H] ⁻ and/or [(CC)+s-H] ⁻
665.8	0.11	666.1	[a ₅ -B ₅ H-3H] ²⁻
635.3	0.61	635.1	w ₂ ⁻ and/or d ₂ ⁻ and/or [(GC)+H ₂ O-H] ⁻
626.1	0.42	626.1	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(GGCC)+H ₂ O-2H] ²⁻
617.3	0.34	617.1	[c ₂ -2H] ⁻ and/or [x ₂ -2H] ⁻ and/or [(GC)-H] ⁻
586.3	0.15	586.1	[y ₄ -H] ²⁻ and/or [b ₄ -H] ²⁻ and/or [(GGCC)-p+H ₂ O-2H] ²⁻
577.0	0.09	577.1	[(CC)-H] ⁻
559.8	0.50	559.4	[M-C-4H] ³⁻
555.2	0.20	555.1	y ₂ ⁻ and/or b ₂ ⁻ and/or [(CG)-p+H ₂ O-H] ⁻
546.3	0.39	546.1	[M-G-4H] ³⁻
537.3	0.13	537.1	[a ₂ -2H] ⁻ and/or [z ₂ -2H] ⁻ and/or [(GC)-p-H] ⁻
526.9	0.38	526.8	[w ₅ -2H] ³⁻
521.6	0.10	521.6	[a ₄ -B ₄ H-3H] ²⁻ and/or [(GGC)+s-2H] ²⁻
	and/or	521.6	[(GGC)+p+H ₂ O-2H] ²⁻
506.2	0.52	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
488.1	0.11	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.8	0.13	481.4	[w ₃ -H] ²⁻ and/or [(CGG)+H ₂ O-2H] ²⁻
466.1	0.38	466.0	[C _{nt} +s+p-H] ⁻
461.7	0.39	461.6	[w ₃ -H] ²⁻ and/or [(GCC)+H ₂ O-2H] ²⁻
447.1	100	447.1	[M-4H]⁴⁻
426.2	0.43	426.1	[a ₂ -CH-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
419.3	0.28	419.3	[M-CH-4H] ⁴⁻
417.0	0.20	417.1	[w ₄ -2H] ³⁻ and/or [d ₄ -2H] ³⁻ and/or [(CCGG)+H ₂ O-3H] ³⁻
408.0	0.47	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
386.1	0.59	386.1	[a ₂ -B ₂ H-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.1	0.54	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
346.2	1.84	346.1	w ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	3.04	328.0	[x ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.1	0.60	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻ and/or [(GC)+H ₂ O-2H] ²⁻
307.6	0.17	307.4	[w ₃ -2H] ³⁻ and/or [(GCC)+H ₂ O-3H] ³⁻
306.2	0.15	306.0	d ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
288.1	2.81	288.0	[c ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.2	0.26	275.1	[sps-H] ⁻
	and/or	275.1	[psp+H ₂ O-H] ⁻
266.4	0.07	266.1	y ₁ ⁻ and/or [G _{nt} -p+H ₂ O-H] ⁻
257.1	0.17	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.0	0.05	239.1	[psp-H ₂ O-H] ⁻
230.1	0.27	230.1	[sG-H ₂ O-H] ⁻
195.1	0.67	195.0	[ps+H ₂ O-H] ⁻
177.1	2.85	177.0	[ps-H] ⁻
172.6	1.05	172.5	[w ₁ -H] ²⁻ and/or [G _{nt} +H ₂ O-2H] ²⁻
159.0	0.47	159.0	[ps-H ₂ O-H] ⁻
150.2	4.24	150.0	G ⁻
110.1	5.38	110.0	C ⁻
97.1	2.68	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	14.67	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.39	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-5H]^{5-}$ ion of the oligonucleotide 5'-d(CGGCCG)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
866.0	1.25	866.2	$[a_3-2H]^-$ and/or $[(GGC)-p-H]^-$
715.1	1.18	715.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-CH-2H]^-$ and/or $[(GC)+s-H]^-$
	and/or	715.1	$[(GC)+p+H_2O-H]^-$
635.1	0.75	635.1	w_2^- and/or d_2^- and/or $[(GC)+H_2O-H]^-$
586.0	1.19	586.1	$[y_4-H]^{2-}$ and/or $[b_4-H]^{2-}$ and/or $[(GGCC)-p+H_2O-2H]^{2-}$
555.7	1.18	555.1	y_2^- and/or b_2^- and/or $[(CG)-p+H_2O-H]^-$
536.5	1.74	537.1	$[a_2-2H]^-$ and/or $[z_2-2H]^-$ and/or $[(GC)-p-H]^-$
487.9	0.81	488.0	$[G_{nt}+s+p-H_2O-H]^-$
466.0	0.86	466.0	$[C_{nt}+s+p-H]^-$
426.5	1.55	426.1	$[a_2-CH-2H]^-$ and/or $[z_2-GH-2H]^-$ and/or $[G_{nt}+s-H]^-$
	and/or	426.1	$[G_{nt}+p+H_2O-H]^-$
419.3	0.87	419.3	$[M-C-5H]^{4-}$
416.9	0.94	417.1	$[w_4-2H]^{3-}$ and/or $[d_4-2H]^{3-}$ and/or $[(GGCC)+H_2O-H]^-$
409.3	0.75	409.3	$[M-G-5H]^{4-}$
408.1	1.93	408.1	$[G_{nt}+s-H_2O-H]^-$
	and/or	408.0	$[G_{nt}+p-H]^-$
394.8	0.93	394.8	$[w_5-3H]^{4-}$
386.4	1.86	386.1	$[a_2-B_2H-2H]^-$ and/or $[z_2-GH-2H]^-$ and/or $[C_{nt}+s-H]^-$
	and/or	386.0	$[C_{nt}+p+H_2O-H]^-$
368.1	0.68	368.1	$[C_{nt}+s-H_2O-H]^-$
	and/or	368.0	$[C_{nt}+p-H]^-$
357.5	100	357.5	$[M-5H]^{5-}$
346.2	3.11	346.1	w_1^- and/or $[G_{nt}+H_2O-H]^-$
328.2	7.40	328.0	$[x_1-2H]^-$ and/or $[G_{nt}-H]^-$
317.1	0.93	317.0	$[w_2-H]^{2-}$ and/or $[d_2-H]^{2-}$ and/or $[(GC)+H_2O-2H]^{2-}$
307.5	1.31	307.4	$[w_3-2H]^{3-}$ and/or $[(GCC)+H_2O-3H]^{3-}$
288.1	7.02	288.0	$[c_1-2H]^-$ and/or $[C_{nt}-H]^-$
230.1	1.62	230.1	$[sG-H_2O-H]^-$
195.1	1.31	195.0	$[ps+H_2O-H]^-$
192.5	1.63	192.5	$[a_2-B_2H-3H]^{2-}$ and/or $[z_2-GH-3H]^{2-}$ and/or $[C_{nt}+s-2H]^{2-}$
177.1	11.25	177.0	$[ps-H]^-$
172.6	3.85	172.5	$[w_1-H]^{2-}$ and/or $[G_{nt}+H_2O-2H]^{2-}$
159.0	2.49	159.0	$[ps-H_2O-H]^-$
150.2	11.25	150.0	G^-
110.1	13.67	110.0	C^-
97.1	13.18	97.0	$[a_1-B_1H-2H]^-$ and/or $[s-H]^-$
	and/or	97.0	$[p+H_2O-H]^-$
79.1	68.69	79.0	$[p-H]^-$ (PO_3^-)
42.0	3.79	42.0	CNO^-

*calculated monoisotopic mass

s = deoxyribose- H_2O ($C_5H_6O_2$), $M_r = 98.0368$ Da

p = PO_3H , $M_r = 79.9663$ Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

$[(B_1...B_n)]$ denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-3H]^{3-}$ ion of the oligonucleotide 5'-d(GCCGGC)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1461.7	0.02	1462.3	y_5^-
1332.9	0.05	1333.2	$[a_5-B_5H-2H]^-$ and/or $[z_5-CH-2H]^-$
1253.1	0.03	1253.2	w_4^- and/or d_4^- and/or $[(CCGG)+H_2O-H]^-$
1235.1	0.01	1235.2	$[x_4-2H]^-$ and/or $[c_4-2H]^-$ and/or $[(CCGG)-H]^-$

1173.0	0.05	1173.2	y4 ⁻ and/or b4 ⁻ and/or [(CCGG)-p+H ₂ O-H] ⁻
1154.8	0.02	1155.2	[z4-2H] ⁻ and/or [a4-2H] ⁻ and/or [(CCGG)-p-H] ⁻
1083.9	0.03	1084.1	[x4-GH-2H] ⁻ and/or [c4-GH-2H] ⁻ and/or [(CCG)+s+p-H] ⁻
1003.9	0.08	1004.2	[a4-B4H-2H] ⁻ and/or [z4-GH-2H] ⁻ and/or [(CCG)+s-H] ⁻
	and/or	1004.1	[(CCG)+p+H ₂ O -H] ⁻
964.1	0.09	964.2	w3 ⁻ and/or [(CGG)+H ₂ O-H] ⁻
945.8	0.02	946.1	[x3-2H] ⁻ and/or [(CGG)-H] ⁻
924.1	0.02	924.1	d3 ⁻ and/or [(CCG)+H ₂ O-H] ⁻
905.9	0.02	906.1	[c3-2H] ⁻ and/or [(CGG)-H] ⁻
884.0	0.16	884.2	y3 ⁻ and/or [(CGG)-p+H ₂ O-H] ⁻
844.2	0.02	844.2	b3 ⁻ and/or [(CCG)-p+H ₂ O-H] ⁻
839.6	0.04	839.7	[M-C ⁻ -3H] ²⁻
819.5	0.17	819.7	[M-G ⁻ -3H] ²⁻
795.0	0.10	795.1	[c3-CH-2H] ⁻ and/or [x3-GH-2H] ⁻ and/or [(CG)+s+p-H] ⁻
770.5	0.68	770.6	[w5-H] ²⁻
755.1	0.06	755.1	[c3-GH-2H] ⁻ and/or [x3-CH-2H] ⁻ and/or [(CC)+s+p-H] ⁻
741.5	0.06	741.6	[a5-B5H-3H] ²⁻
730.5	0.08	730.6	[y5-H] ²⁻
715.1	0.24	715.1	[a3-B3H-2H] ⁻ and/or [z3-GH-2H] ⁻ and/or [(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p+H ₂ O -H] ⁻
697.0	0.04	697.1	[(CG)+p-H] ⁻
	and/or	697.1	[(CG)+s-H ₂ O-H] ⁻
675.0	0.06	675.1	[(CC)+p+ H ₂ O -H] ⁻
	and/or	675.1	[a3-GH-2H] ⁻ and/or [(CC)+s-H] ⁻
666.0	0.36	666.1	[a5-B5H-3H] ²⁻
657.0	0.07	657.1	[(CC)+s-H ₂ O -H] ⁻
	and/or	657.1	[(CC)+p-H] ⁻
	and/or	657.1	[(GG)-H] ⁻
635.1	0.41	635.1	w2 ⁻ and/or d2 ⁻ and/or [(CG)+H ₂ O-H] ⁻
626.0	0.43	626.1	[w4-H] ²⁻ and/or [d4-H] ²⁻ and/or [(CCGG)+H ₂ O-2H] ²⁻
617.1	0.31	617.1	[c2-2H] ⁻ and/or [x2-2H] ⁻ and/or [(CG)-H] ⁻
596.4	100	596.4	[M-3H]³⁻
577.5	0.11	577.1	[(CC)-H] ⁻
559.4	0.30	559.4	[M-CH-3H] ³⁻
555.1	0.14	555.1	y2 ⁻ and/or b2 ⁻ and/or [(CG)-p+H ₂ O-H] ⁻
546.1	0.33	546.1	[M-GH-3H] ³⁻
537.1	0.06	537.1	[z2-2H] ⁻ and/or [a2-2H] ⁻ and/or [(CG)-p-H] ⁻
526.8	0.05	526.8	[d5-2H] ³⁻
524.0	0.02	524.1	[w2-CH] ⁻ and/or [d2-CH] ⁻ and/or [G _{nt} +s+p+ H ₂ O-H] ⁻
513.4	0.17	513.4	[w5-2H] ³⁻
506.0	0.19	506.0	[x2-CH-2H] ⁻ and/or [c2-CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
501.6	0.15	501.6	[a4-B4H-3H] ²⁻
488.1	0.04	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.4	0.24	481.4	[w3-H] ²⁻ and/or [(CGG)+H ₂ O-2H] ²⁻
472.7	0.02	472.6	[x3-3H] ²⁻ and/or [(CGG)-2H] ²⁻
466.0	0.14	466.0	[C _{nt} +s+p-H] ⁻
461.5	0.02	461.6	[d3-H] ²⁻ and/or [(CCG)+H ₂ O-2H] ²⁻
452.6	0.01	452.6	[c3-H] ²⁻ and/or [(CCG)-2H] ²⁻
426.1	0.19	426.1	[a2-B2H-2H] ⁻ and/or [z2-CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
417.1	0.01	417.1	[w4-2H] ³⁻ and/or [d4-2H] ³⁻ and/or [(CCGG)+H ₂ O-3H] ³⁻
408.0	0.10	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
385.9	0.06	386.1	[a2-GH-2H] ⁻ and/or [z2-GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
368.0	0.09	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
357.0	0.04	357.1	[a3-B3H-3H] ²⁻

346.0	0.11	346.1	d ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
328.1	0.17	328.0	[c ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
317.0	0.09	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻ and/or [(CG)+H ₂ O-2H] ²⁻
310.1	0.01	310.0	[G _{nt} -H ₂ O-H] ⁻
306.1	0.66	306.0	w ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
288.0	0.49	288.0	[x ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.1	0.02	275.1	[sps-H] ⁻
	and/or	275.1	[psp+H ₂ O-H] ⁻
257.0	0.04	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.01	239.1	[psp-H ₂ O-H] ⁻
230.1	0.07	230.1	[sG-H ₂ O-H] ⁻
195.1	0.11	195.0	[ps+H ₂ O-H] ⁻
190.1	0.01	190.1	[sC-H ₂ O-H] ⁻
177.1	0.35	177.0	[ps-H] ⁻
159.0	0.05	159.0	[ps-H ₂ O-H] ⁻
150.1	1.11	150.0	G ⁻
110.0	0.18	110.0	C ⁻
97.0	0.38	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	1.20	79.0	[p-H] ⁻ (PO ₃ ⁻)

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the [M-4H]⁴⁻ ion of the oligonucleotide 5'-d(GCCGGC)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
1173.7	0.03	1173.2	y ₄ ⁻ and/or b ₄ ⁻ and/or [(CCGG)-p+H ₂ O-H] ⁻
1084.3	0.02	1084.1	[x ₄ -GH-2H] ⁻ and/or [c ₄ -GH-2H] ⁻ and/or [(CCG)+s+p-H] ⁻
1004.0	0.04	1004.2	[a ₄ -B ₄ H-2H] ⁻ and/or [z ₄ -GH-2H] ⁻ and/or [(CCG)+s-H] ⁻
	and/or	1004.1	[(CCG)+p+H ₂ O-H] ⁻
963.8	0.03	964.2	w ₃ ⁻ and/or [(CGG)+H ₂ O-H] ⁻
906.1	0.02	906.1	[c ₃ -2H] ⁻ and/or [(CGG)-H] ⁻
884.3	0.05	884.2	y ₃ ⁻ and/or [(CGG)-p+H ₂ O-H] ⁻
794.7	0.06	795.1	[c ₃ -CH-2H] ⁻ and/or [x ₃ -GH-2H] ⁻ and/or [(CG)+s+p-H] ⁻
754.9	0.01	755.1	[c ₃ -GH-2H] ⁻ and/or [x ₃ -CH-2H] ⁻ and/or [(CC)+s+p-H] ⁻
730.6	0.08	730.6	[y ₅ -H] ²⁻
721.6	0.03	721.6	[z ₅ -3H] ²⁻
714.8	0.06	715.1	[a ₃ -B ₃ H-2H] ⁻ and/or [z ₃ -GH-2H] ⁻ and/or [(CG)+s-H] ⁻
	and/or	715.1	[(CG)+p+H ₂ O-H] ⁻
697.2	0.01	697.1	[(CG)+p-H] ⁻
	and/or	697.1	[(CG)+s-H ₂ O-H] ⁻
675.5	0.01	675.1	[(CC)+p+H ₂ O-H] ⁻
	and/or	675.1	[a ₃ -GH-2H] ⁻ and/or [(CC)+s-H] ⁻
665.9	0.27	666.1	[a ₅ -B ₅ H-3H] ²⁻
657.0	0.02	657.1	[(CC)+s-H ₂ O-H] ⁻
	and/or	657.1	[(CC)+p-H] ⁻
	and/or	657.1	[(GG)-H] ⁻
634.8	0.13	635.1	w ₂ ⁻ and/or d ₂ ⁻ and/or [(CG)+H ₂ O-H] ⁻
626.0	0.24	626.1	[w ₄ -H] ²⁻ and/or [d ₄ -H] ²⁻ and/or [(CCGC)+H ₂ O-2H] ²⁻
617.0	0.09	617.1	[c ₂ -2H] ⁻ and/or [x ₂ -2H] ⁻ and/or [(CG)-H] ⁻
586.0	0.14	586.1	[y ₄ -H] ²⁻ and/or [b ₄ -H] ²⁻ and/or [(CCGG)-p+H ₂ O-2H] ²⁻

581.6	0.09	581.6	[(CCG)+s+p+s-H ₂ O-2H] ²⁻
	and/or	581.5	[(CCG)+p+s+p-2H] ²⁻
577.1	0.06	577.1	[a ₄ -3H] ²⁻ and/or [z ₄ -3H] ²⁻ and/or [(CCGG)-p-2H] ²⁻
559.4	0.20	559.4	[M-C ⁻ -4H] ³⁻
555.1	0.13	555.1	y ₂ ⁻ and/or b ₂ ⁻ and/or [(CG)-p+H ₂ O-H] ⁻
546.0	0.29	546.1	[M-G ⁻ -4H] ³⁻
537.0	0.06	537.1	[z ₂ -2H] ⁻ and/or [a ₂ -2H] ⁻ and/or [(CG)-p-H] ⁻
526.7	0.07	526.8	[d ₅ -2H] ³⁻
513.4	0.24	513.4	[w ₅ -2H] ³⁻
506.0	0.15	506.0	[x ₂ -CH-2H] ⁻ and/or [c ₂ -CH-2H] ⁻ and/or [G _{nt} +s+p-H] ⁻
501.6	0.16	501.6	[a ₄ -B ₄ H-3H] ²⁻
487.8	0.06	488.0	[G _{nt} +s+p-H ₂ O-H] ⁻
481.4	0.12	481.4	[w ₃ -H] ²⁻ and/or [(CGG)+H ₂ O-2H] ²⁻
465.8	0.24	466.0	[C _{nt} +s+p-H] ⁻
461.5	0.18	461.6	[d ₃ -H] ²⁻ and/or [(CCG)+H ₂ O-2H] ²⁻
447.1	100	447.1	[M-4H]⁴⁻
432.5	0.07	432.6	[z ₃ -3H] ²⁻ and/or [(CGG)-p-2H] ²⁻
425.9	0.23	426.1	[a ₂ -B ₂ H ₂ H] ⁻ and/or [z ₂ -CH-2H] ⁻ and/or [G _{nt} +s-H] ⁻
	and/or	426.1	[G _{nt} +p+H ₂ O-H] ⁻
419.3	0.08	419.3	[M-CH-4H] ⁴⁻
417.0	0.55	417.1	[w ₄ -2H] ³⁻ and/or [d ₄ -2H] ³⁻ and/or [(CCGG)+H ₂ O-3H] ³⁻
409.3	0.03	409.3	[M-GH-4H] ⁴⁻
407.8	0.10	408.1	[G _{nt} +s-H ₂ O-H] ⁻
	and/or	408.0	[G _{nt} +p-H] ⁻
394.9	0.02	394.8	[d ₅ -3H] ⁴⁻
390.4	0.02	390.3	[c ₅ -5H] ⁴⁻
385.9	0.06	386.1	[a ₂ -GH-2H] ⁻ and/or [z ₂ -GH-2H] ⁻ and/or [C _{nt} +s-H] ⁻
	and/or	386.0	[C _{nt} +p+H ₂ O-H] ⁻
367.9	0.10	368.1	[C _{nt} +s-H ₂ O-H] ⁻
	and/or	368.0	[C _{nt} +p-H] ⁻
345.9	0.11	346.1	d ₁ ⁻ and/or [G _{nt} +H ₂ O-H] ⁻
327.9	0.28	328.0	[c ₁ -2H] ⁻ and/or [G _{nt} -H] ⁻
320.7	0.06	320.7	[w ₃ -2H] ³⁻ and/or [(CGG)+H ₂ O-3H] ³⁻
316.9	0.14	317.0	[w ₂ -H] ²⁻ and/or [d ₂ -H] ²⁻ and/or [(CG)+H ₂ O-2H] ²⁻
306.0	0.69	306.0	w ₁ ⁻ and/or [C _{nt} +H ₂ O-H] ⁻
287.9	0.73	288.0	[x ₁ -2H] ⁻ and/or [C _{nt} -H] ⁻
275.1	0.05	275.1	[sps-H] ⁻
	and/or	275.1	[psp+H ₂ O-H] ⁻
256.9	0.06	257.1	[sps-H ₂ O-H] ⁻
	and/or	257.0	[psp-H] ⁻
239.1	0.01	239.1	[psp-H ₂ O-H] ⁻
230.0	0.06	230.1	[sG-H ₂ O-H] ⁻
195.0	0.14	195.0	[ps+H ₂ O-H] ⁻
190.1	0.01	190.1	[sC-H ₂ O-H] ⁻
177.0	0.46	177.0	[ps-H] ⁻
172.5	0.05	172.5	[d ₁ -H] ²⁻ and/or [G _{nt} +H ₂ O-2H] ²⁻
158.9	0.05	159.0	[ps-H ₂ O-H] ⁻
150.1	1.27	150.0	G ⁻
110.0	0.37	110.0	C ⁻
96.9	0.52	97.0	[a ₁ -B ₁ H-2H] ⁻ and/or [s-H] ⁻
	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	1.59	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.12	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_{nt} denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)

Assignment of product ions observed in the ESI-MS/MS spectrum of the $[M-5H]^{5-}$ ion of the oligonucleotide 5'-d(GCCGGC)-3'.

<i>m/z</i>	<i>Rel. Int. (%)</i>	<i>m/z(calc).*</i>	<i>Assignment</i>
883.5	0.12	884.2	y_3^- and/or $[(CGG)-p+H_2O-H]^-$
794.9	0.09	795.1	$[c_3-CH-2H]^-$ and/or $[x_3-GH-2H]^-$ and/or $[(CG)+s+p-H]^-$
714.5	0.12	715.1	$[a_3-B_3H-2H]^-$ and/or $[z_3-GH-2H]^-$ and/or $[(CG)+s-H]^-$
	and/or	715.1	$[(CG)+p+H_2O-H]^-$
675.5	0.09	675.1	$[(CC)+p+H_2O-H]^-$
	and/or	675.1	$[a_3-GH-2H]^-$ and/or $[(CC)+s-H]^-$
666.0	0.12	666.1	$[a_5-B_5H-3H]^{2-}$
634.7	0.12	635.1	w_2^- and/or d_2^- and/or $[(CG)+H_2O-H]^-$
617.0	0.09	617.1	$[c_2-2H]^-$ and/or $[x_2-2H]^-$ and/or $[(CG)-H]^-$
586.0	0.09	586.1	$[y_4-H]^{2-}$ and/or $[b_4-H]^{2-}$ and/or $[(CCGG)-p+H_2O-2H]^{2-}$
554.9	0.10	555.1	y_2^- and/or b_2^- and/or $[(CG)-p+H_2O-H]^-$
526.7	0.11	526.8	$[d_5-2H]^{3-}$
506.0	0.14	506.0	$[x_2-CH-2H]^-$ and/or $[c_2-CH-2H]^-$ and/or $[G_{nt}+s+p-H]^-$
481.4	0.12	481.4	$[w_3-H]^{2-}$ and/or $[(CCGG)+H_2O-2H]^{2-}$
465.7	0.19	466.0	$[C_{nt}+s+p-H]^-$
443.7	0.14	443.7	$[a_5-B_5H-4H]^{3-}$
441.6	0.21	441.6	$[y_3-H]^{2-}$ and/or $[(CGG)-p+H_2O-2H]^{2-}$
437.0	0.20	437.1	$[(CG)+p+s+p-2H]^{2-}$
	and/or	437.1	$[(CG)+s+p+s-H_2O-2H]^{2-}$
426.0	0.23	426.1	$[a_2-B_2H-2H]^-$ and/or $[z_2-CH-2H]^-$ and/or $[G_{nt}+s-H]^-$
	and/or	426.1	$[G_{nt}+p+H_2O-H]^-$
419.3	0.28	419.3	$[M-C-5H]^{4-}$
417.0	0.43	417.1	$[w_4-2H]^{3-}$ and/or $[d_4-2H]^{3-}$ and/or $[(CCGG)+H_2O-3H]^{3-}$
411.0	0.26	411.0	$[x_4-4H]^{3-}$ and/or $[c_4-4H]^{3-}$ and/or $[(CCGG)-3H]^{3-}$
409.3	0.21	409.3	$[M-G-5H]^{4-}$
407.8	0.19	408.1	$[G_{nt}+s-H_2O-H]^-$
	and/or	408.0	$[G_{nt}+p-H]^-$
394.8	0.28	394.8	$[d_5-3H]^{4-}$
390.4	0.15	390.3	$[c_5-5H]^{4-}$
367.8	0.23	368.1	$[C_{nt}+s-H_2O-H]^-$
	and/or	368.0	$[C_{nt}+p-H]^-$
357.5	100	357.5	$[M-5H]^{5-}$
345.9	0.24	346.1	d_1^- and/or $[G_{nt}+H_2O-H]^-$
327.9	0.29	328.0	$[c_1-2H]^-$ and/or $[G_{nt}-H]^-$
320.8	0.53	320.7	$[w_3-2H]^{3-}$ and/or $[(CCGG)+H_2O-3H]^{3-}$
316.9	0.15	317.0	$[w_2-H]^{2-}$ and/or $[d_2-H]^{2-}$ and/or $[(CG)+H_2O-2H]^{2-}$
312.6	0.23	312.6	$[w_4-3H]^{4-}$ and/or $[d_4-3H]^{4-}$ and/or $[(CCGG)+H_2O-4H]^{4-}$
306.0	0.67	306.0	w_1^- and/or $[C_{nt}+H_2O-H]^-$
288.0	1.25	288.0	$[x_1-2H]^-$ and/or $[C_{nt}-H]^-$
275.2	0.11	275.1	$[sps-H]^-$
	and/or	275.1	$[psp+H_2O-H]^-$
256.9	0.10	257.1	$[sps-H_2O-H]^-$
	and/or	257.0	$[psp-H]^-$
230.0	0.09	230.1	$[sG-H_2O-H]^-$
212.5	0.13	212.5	$[a_2-B_2H-3H]^{2-}$ and/or $[G_{nt}+s-2H]^{2-}$
	and/or	212.5	$[G_{nt}+p+H_2O-2H]^{2-}$
211.0	0.10	211.0	$[w_2-2H]^{3-}$ and/or $[d_2-2H]^{3-}$ and/or $[(CG)+H_2O-3H]^{3-}$
195.0	0.29	195.0	$[ps+H_2O-H]^-$
177.0	1.02	177.0	$[ps-H]^-$
172.5	0.09	172.5	$[d_1-H]^{2-}$ and/or $[G_{nt}+H_2O-2H]^{2-}$
158.9	0.09	159.0	$[ps-H_2O-H]^-$
150.1	1.90	150.0	G^-
110.1	0.83	110.0	C^-
97.0	0.97	97.0	$[a_1-B_1H-2H]^-$ and/or $[s-H]^-$

	and/or	97.0	[p+H ₂ O-H] ⁻
79.0	2.76	79.0	[p-H] ⁻ (PO ₃ ⁻)
42.0	0.49	42.0	CNO ⁻

*calculated monoisotopic mass

s = deoxyribose-H₂O (C₅H₆O₂), M_r = 98.0368 Da

p = PO₃H, M_r = 79.9663 Da

B_n denotes a mononucleotide which may be either psB or sBp

[(B₁...B_n)] denotes a polynucleotide which may be either (psB₁...+psB_n) or (sB₁p...+sB_np)