

# MASS SPECTROMETRY of NUCLEOSIDES and NUCLEIC ACIDS

---

Edited by  
Joseph H. Banoub  
Patrick A. Limbach

---



CRC Press  
Taylor & Francis Group

# MASS SPECTROMETRY of NUCLEOSIDES and NUCLEIC ACIDS

---

Edited by  
Joseph H. Banoub  
Patrick A. Limbach



**CRC Press**

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the  
Taylor & Francis Group, an **informa** business

CRC Press  
Taylor & Francis Group  
6000 Broken Sound Parkway NW, Suite 300  
Boca Raton, FL 33487-2742

© 2010 by Taylor and Francis Group, LLC  
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed in the United States of America on acid-free paper  
10 9 8 7 6 5 4 3 2 1

International Standard Book Number: 978-1-4200-4402-7 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access [www.copyright.com](http://www.copyright.com) (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

---

#### Library of Congress Cataloging-in-Publication Data

---

Mass spectrometry of nucleosides and nucleic acids / editors, Joseph H. Banoub, Patrick A. Limbach.  
p. cm.

Includes bibliographical references and index.

ISBN 978-1-4200-4402-7 (hardcover : alk. paper)

1. Nucleic acids--Spectra. 2. Nucleosides--Spectra. 3. Mass spectrometry. I. Banoub, Joseph H. II. Limbach, Patrick A. III. Title.

QD433.M37 2010

572.80284--dc22

2009040943

---

Visit the Taylor & Francis Web site at  
<http://www.taylorandfrancis.com>

and the CRC Press Web site at  
<http://www.crcpress.com>

# **MASS SPECTROMETRY of NUCLEOSIDES and NUCLEIC ACIDS**

---

---

# Preface

Nucleic acids and their components play an important role in a variety of fundamental biological processes. The elucidation of the molecular structure of deoxyribonucleic acid (DNA) in 1953 by Watson and Crick led to the so-called *central dogma of molecular biology*. This dogma proclaims that DNA is transcribed to form messenger ribonucleic acid (mRNA) of a complementary sequence that directs its own replication and the sequence of nucleobases in the RNA is then translated to the corresponding sequence of amino acids to form a protein. In addition to mRNA, two other RNA species, transfer RNA (tRNA) and ribosomal RNA (rRNA), are required for this translation process.

The field of genomics has evolved over the last few years, culminating in the sequencing and characterization of the entire human genome. Furthermore, genetic theory contributed to the concept of coding by genes. A gene is defined, biochemically, as a segment of DNA (or, in a few cases, RNA) that encodes the information required to produce a functional biological product—a protein. Current challenges include the unraveling of the relationship between the genome and the proteome. The genome can be used to predict the total potential proteome, including all of the modifications that can be carried out after the initial translation of mRNA to protein by tRNAs and rRNAs.

Consequently, nucleic acid research has had a most profound impact on molecular biology. The study of noncovalent interactions between nucleic acids and peptides or proteins is a highly active research field. Such interactions are involved in many cellular processes, including replication, regulation of gene expression, and DNA packaging. Nearly all the functions of nucleic acids are accomplished by interacting with proteins. Recognition and selectivity are achieved through noncovalent contacts between virtually all biopolymers (proteins, DNA, RNA, polysaccharides, and membrane lipids) that are present in living organisms. The formation and dissociation of these weak interactions are crucial in a vast number of biochemical events. In addition, weak interactions between nucleases and DNA or RNA are central to the recombination, repair, and replication of these molecules in cells.

It is not possible to overestimate the role of metal ions in determining the three-dimensional architecture of nucleic acids, their helices, and higher-order molecular assemblies. Protein–protein, protein–nucleic acid, protein–ligand, and antibody–antigen interactions are other examples that involve noncovalent interactions, which are also of fundamental importance for the pharmaceutical industry in the evaluation of potential drug candidates.

The ability of DNA adducts to induce mutagenesis and carcinogenesis is dependent on their chemical structure, stability, and ability to be recognized by specialized DNA repair proteins. The local DNA sequence context plays a major role in mediating the rate of nucleobase adducts repair and their mispairing potency. Furthermore, the exact location of a DNA adduct within the gene determines whether the resulting mutation affects the structure and function of the gene product.

Among the biophysical techniques employed to study the large biomolecular assemblies involved in these tasks, mass spectrometry presents the greatest potential based on its intrinsic characteristics and accessible information. Therefore, mass spectrometry has had a major role in elucidating the regulatory components and mechanisms involved in the progression from gene to functional protein. Furthermore, analyses of nucleic acids is required to establish their size, purity, and sequence as a prerequisite to their use as molecular probes or therapeutics in biomedical science.

The development of soft-ionization techniques for mass spectrometry has, without question, transformed the field of nucleic acid chemistry. Significant progress in the area of accurate mass

determination, sequencing, and the study of noncovalent interactions has been made possible by the use of ionization techniques such as matrix-assisted laser desorption ionization coupled to time-of-flight (MALDI-TOF) mass analyzers and electrospray ionization (ESI) coupled to conventional analyzers and to tandem mass spectrometric (MS/MS) instruments. Mass spectrometry is now one of the definitive methods for the characterization of isolated and synthetic DNA, RNA, and related products. In addition to accurately determining an intact mass, one can obtain information on the molecular structure and biopolymer sequence using a number of different mass spectrometry approaches. These approaches have provided valuable information on structurally specific biomolecular DNA and RNA interactions. They can also enable the direct determination of stoichiometry for post-translational modifications of specific DNA–drug, DNA–protein, RNA–protein, and DNA–RNA noncovalent and covalent associations. Furthermore, these mass spectrometry approaches can be applied to DNA and RNA sequencing analyses, single-nucleotide polymorphisms, genotyping, mutation detection, genetic diagnosis, and probing of viral structures.

The editors acknowledge the special efforts of the contributors to this CRC/Taylor & Francis publication: Diethard K. Bohme (Canada), William Buchmann (France), Ed Dudley (UK), Daniele Fabris (USA), Valérie Gabelica (Belgium), Kristina Håkansson (USA), Filip Lemièr (Belgium), Russell P. Newton (UK), Natalia Y. Tretyakova (USA), Antonio Triolo (Italy), Jean-Jacques Vasseur (France), and Yinsheng Wang (USA). Each contributor, an expert in their particular discipline, has shared their knowledge and understanding of the field of nucleic acid mass spectrometry.

Our aim is to expose you, the reader, to the latest developments in the field of nucleic acid mass spectrometry. The past 15 years have seen dramatic progress in this field. Thus, in addition to contributions that review and summarize these developments and advances, this book also contains chapters describing the next generation of mass spectrometry analyses of nucleic acids and their complexes that have been enabled by such past accomplishments.

This book should be equally accessible to the nucleic acids expert who is a mass spectrometry novice as well as to those with expertise in mass spectrometry but a minimal appreciation of nucleic acids. The exciting developments in mass spectrometry technology have fueled incredible advances in our understanding of nucleic acids and their complexes. We believe these contributions have captured these advances in outstanding fashion, serving to inspire new findings and developments of interest to all in this field.

Finally, no work of this scope is possible without the guidance and support of the publisher. The editors wish to especially acknowledge Ms. Jill J. Jurgensen for her expert help in preparing and producing this book.

**Joseph H. Banoub (Canada)**  
**Patrick A. Limbach (USA)**

---

# Editors

**Joseph H. Banoub** was born in 1947 and obtained his BS degree from the University of Alexandria (Egypt) in 1969. In 1977 he obtained his PhD degree in organic chemistry from the University de Montreal (Quebec) under the supervision of Stephen Hanessian. He spent three years as a research associate with the MDS Health Group at the National Research Council of Canada, Ottawa, synthesizing human neoglycoconjugate vaccines. In 1979 he moved to St. John's, Newfoundland, as a research scientist with the Federal Government of Canada, Department of Fisheries and Oceans. During this period he worked on the structural elucidation of various bacterial antigens and the formation of artificial vaccines for aquatic Gram-negative bacteria. He is currently principal research scientist and head of the special project program of the Department of Fisheries and Oceans and his research focus has evolved into the mass spectrometry and tandem mass spectrometry study of biomolecules (bacterial lipopolysaccharides, nucleic acids, and proteomics). He has been an adjunct professor of biochemistry at Memorial University of Newfoundland since 1984 and an adjunct professor of chemistry since 2006. His present research interests include the uses of tandem mass spectrometry for the structural elucidation of biologically active molecules.

**Patrick A. Limbach** was born in 1966 and obtained his BS degree from Centre College (Kentucky, USA) in 1988. In 1992, he obtained his PhD degree in analytical chemistry from The Ohio State University under the supervision of Alan G. Marshall. He spent two years as a postdoctoral researcher in the laboratory of James A. McCloskey at the University of Utah. In 1995, he accepted his first faculty position in the Department of Chemistry at Louisiana State University in Baton Rouge, Louisiana. In 2001 he moved to his current location in the Department of Chemistry at the University of Cincinnati. He is a professor of chemistry, serves as department head, and is an Ohio Eminent Scholar. His research interests include the development of new mass spectrometry methods for analyzing RNAs, the identification of unknown modified nucleosides, the characterization of RNA-protein complexes, and investigating the role of modified nucleosides in biological systems.

---

# Contributors

**Janna Anichina**

York University  
Toronto, ON, Canada

**Joseph H. Banoub**

Fisheries and Oceans Canada  
Science Branch  
Newfoundland, Canada

and

Department of Chemistry  
Memorial University of Newfoundland  
Newfoundland, Canada

**Diethard K. Bohme**

York University  
Toronto, ON, Canada

**William Buchmann**

Analysis and Modeling Laboratory for  
Biology and Environment  
University of Evry-Valley of the Essonne  
Evry, France

**C. Callie Coombs**

Department of Chemistry  
University of Cincinnati  
Cincinnati, Ohio

**Ed Dudley**

Department of Biological Sciences  
Swansea University  
Singleton Park, Swansea,  
United Kingdom

**Daniele Fabris**

Department of Chemistry and  
Biochemistry  
University of Maryland Baltimore  
County  
Baltimore, Maryland

**Valérie Gabelica**

Mass Spectrometry Laboratory  
University of Liège and FNRS  
Liège, Belgium

**Chunang Gu**

Environmental Toxicology Graduate  
Program  
University of California  
Riverside, California

**Nathan A. Hagan**

Department of Chemistry and Biochemistry  
University of Maryland Baltimore County  
Baltimore, Maryland

**Kristina Håkansson**

Department of Chemistry  
University of Michigan  
Ann Arbor, Michigan

**Haizheng Hong**

Environmental Toxicology Graduate  
Program  
University of California  
Riverside, California

**Mahmud Hossain**

Department of Chemistry  
University of Cincinnati  
Cincinnati, Ohio

**Farid Jahouh**

Fisheries and Oceans Canada  
Science Branch  
Newfoundland, Canada

and

Department of Chemistry  
Memorial University of Newfoundland  
Newfoundland, Canada



**Nicolas Joly**

Department of Chemistry  
Rue de l'Université  
Béthune, France

**Filip Lemièr**

Department of Chemistry  
University of Antwerp  
Antwerp, Belgium

**Patrick A. Limbach**

Department of Chemistry  
University of Cincinnati  
Cincinnati, Ohio

**Patrick Martin**

Department of Chemistry  
Rue de l'Université  
Béthune, France

**Judith Miller-Banoub**

Fisheries and Oceans Canada  
Science Branch  
Newfoundland, Canada

and

Department of Chemistry  
Memorial University of Newfoundland  
Newfoundland, Canada

**Russell P. Newton**

Biological Sciences, School of the  
Environment and Society  
Swansea University  
Singleton Park, Swansea, United Kingdom

**Keiichiro Ohara**

Max Mousseron Institute of Biomolecules  
University of Montpellier  
Montpellier, France

**Michael Smietana**

Max Mousseron Institute of  
Biomolecules  
University of Montpellier  
Montpellier, France

**Natalia Y. Tretyakova**

Department of Medicinal  
Chemistry  
University of Minnesota  
Minneapolis, Minnesota

**Antonio Triolo**

Menarini Ricerche Spa  
Florence, Italy

**Kevin B. Turner**

Department of Chemistry and  
Biochemistry  
University of Maryland Baltimore  
County  
Baltimore, Maryland

**Jean-Jacques Vasseur**

Max Mousseron Institute of  
Biomolecules  
University of Montpellier  
Montpellier, France

**Yinsheng Wang**

Department of Chemistry  
University of California  
Riverside, California

**Jiong Yang**

Department of Chemistry  
University of Michigan  
Ann Arbor, Michigan

---

# Contents

Preface .....	vii
Editors .....	ix
Contributors .....	xi
 <b>Chapter 1</b> Overview of Recent Developments in the Mass Spectrometry of Nucleic Acid and Constituents .....	 1
<i>Joseph H. Banoub, Judith Miller-Banoub, Farid Jahouh, Nicolas Joly, and Patrick Martin</i>	
 <b>Chapter 2</b> Mass Spectrometric Analysis of Deoxyinosines .....	 91
<i>Antonio Triolo</i>	
 <b>Chapter 3</b> Tandem Mass Spectrometry of Nucleic Acids .....	 105
<i>Jiong Yang and Kristina Håkansson</i>	
 <b>Chapter 4</b> Mass Spectrometric Application in the Study of Cyclic Nucleotides in Biochemical Signal Transduction .....	 127
<i>Russell P. Newton</i>	
 <b>Chapter 5</b> Analysis of Urinary Modified Nucleosides by Mass Spectrometry .....	 163
<i>Ed Dudley</i>	
 <b>Chapter 6</b> Mass Spectrometric Determination of DNA Adducts in Human Carcinogenesis .....	 195
<i>Filip Lemière</i>	
 <b>Chapter 7</b> Sequence Distribution of Nucleobase Adducts Studied by Isotope Labeling of DNA–Mass Spectrometry .....	 257
<i>Natalia Y. Tretyakova</i>	
 <b>Chapter 8</b> Electrospray Mass Spectrometry of Noncovalent Complexes between Small Molecule Ligands and Nucleic Acids .....	 283
<i>Valérie Gabelica</i>	

<b>Chapter 9</b>	Electrospray Ionization–Mass Spectrometry for the Investigation of Protein–Nucleic Acid Interactions .....	303
	<i>Daniele Fabris, Kevin B. Turner, and Nathan A. Hagan</i>	
<b>Chapter 10</b>	Characterization of Noncovalent Complexes of Nucleic Acids with Peptides and Proteins by Mass Spectrometry .....	329
	<i>William Buchmann</i>	
<b>Chapter 11</b>	MALDI-TOF Detection of Specific Noncovalent Complexes of Highly Acidic Biomolecules with Pyrenemethylguanidinium .....	371
	<i>Michael Smietana, Keiichiro Ohara, and Jean-Jacques Vasseur</i>	
<b>Chapter 12</b>	Quantitative Identification of Nucleic Acids via Signature Digestion Products Detected Using Mass Spectrometry .....	397
	<i>Mahmud Hossain and Patrick A. Limbach</i>	
<b>Chapter 13</b>	Electrospray Ionization Mass Spectrometry for the Direct Analysis of Modified Nucleosides in Small RNAs .....	423
	<i>C. Callie Coombs and Patrick A. Limbach</i>	
<b>Chapter 14</b>	LC-MS/MS for the Examination of the Cytotoxic and Mutagenic Properties of DNA Lesions <i>In Vitro</i> and <i>In Vivo</i> .....	431
	<i>Chunang Gu, Haizheng Hong, and Yinsheng Wang</i>	
<b>Chapter 15</b>	Influence of Metal Ions on the Structure and Reactivity of Nucleic Acids .....	445
	<i>Janna Anichina and Diethard K. Bohme</i>	
<b>Index</b>	.....	479

---

# 1 Overview of Recent Developments in the Mass Spectrometry of Nucleic Acid and Constituents

*Joseph H. Banoub, Judith Miller-Banoub, Farid Jahouh,  
Nicolas Joly, and Patrick Martin*

## CONTENTS

1.1	Introduction .....	3
1.2	Gas-Phase Techniques .....	4
1.2.1	Electron Impact Ionization-MS for the Detection of DNA Adducts .....	4
1.2.2	Electron-Capture Ionization-MS for the Detection of DNA Adducts .....	5
1.2.3	Photoionization Detection of Nucleic Acid Base Constituents of DNA from Space .....	7
1.3	Desorption Techniques .....	8
1.3.1	MALDI-TOF-MS Technique .....	8
1.3.1.1	MALDI-TOF-MS Characterization of Nucleosides .....	8
1.3.1.2	MALDI-TOF-MS Characterization of Nucleotides .....	9
1.3.1.3	Characterization of Oligonucleotides .....	10
1.3.1.4	MALDI-TOF-MS Characterization of RNA-Protein Complexes .....	24
1.3.1.5	Atmospheric Pressure MALDI-FT-ICR-MS Characterization of RNA-Protein Complexes .....	24
1.3.1.6	MALDI-TOF-MS Characterization of DNA Adducts .....	25
1.3.2	Surface-Enhanced Laser Desorption/Ionization (SELDI)-TOF-MS .....	25
1.3.3	MS/MS Analysis of Oligonucleotides .....	26
1.3.3.1	AP-MALDI-FT-ICR-MS <sup>n</sup> of Normal and Chemically Modified RNA .....	27
1.3.3.2	RNA Fragmentation in MALDI-TOF-TOF-MS/MS Studied by H/D Exchange .....	28
1.3.4	Secondary Ion Mass Spectrometry .....	29
1.3.5	Liquid Secondary Ion Mass Spectrometry .....	30
1.4	Spray Ionization Techniques .....	31
1.4.1	Electrospray Ionization .....	31
1.4.1.1	ESI-MS Characterization of Nucleosides .....	32
1.4.2	ESI-MS Characterization of Nucleotides .....	33
1.4.2.1	Analysis of Nucleotides in Champagne .....	33
1.4.2.2	Determination of Nucleotides in Biomimicking Prebiotic Synthesis .....	33

1.4.2.3	Determination of Nucleotides on Pressure-Assisted Capillary Electrophoresis .....	34
1.4.2.4	HPLC-ESI-MS Quantification of Nucleotides .....	34
1.4.3	ESI-MS Characterization of DNA Adducts .....	35
1.4.3.1	Enediyne-DNA Adducts .....	35
1.4.3.2	Etheno-DNA Adducts .....	36
1.4.3.3	Base-Alkylated Adducts .....	36
1.4.3.4	LC-MS Analysis of Oxidized Damaged Bases on Nucleotides .....	36
1.4.4	ESI-MS Genotyping of SNPs .....	37
1.4.5	ESI-MS Characterization of Oligonucleotides .....	39
1.4.5.1	Practical Aspects of Oligonucleotides Analysis .....	39
1.4.5.2	ESI-MS Analysis of Synthetic Oligonucleotides .....	41
1.4.5.3	ESI-MS Oligonucleotides Sequencing .....	41
1.4.5.4	Combined ESI-MS Analyses of DNA Oligonucleotides and DNA Adducts .....	41
1.4.6	ESI-MS Characterization of PCR Products .....	42
1.4.7	ESI-MS Characterization of Covalent or Noncovalent DNA and RNA Complexes .....	42
1.4.7.1	ESI-MS Characterization of Covalent DNA Complexes .....	42
1.4.7.2	ESI-MS Characterization of Noncovalent DNA Complexes .....	44
1.4.7.3	ESI-MS Characterization of Noncovalent Complexes of Peptides and Double-Stranded DNA .....	45
1.4.7.4	ESI-MS Characterization of Complexes of Actinomycin D and Oligonucleotides .....	46
1.4.7.5	ESI-MS Characterization of Complexes of Ribonuclease A and Cytidylic Acid Ligands .....	46
1.4.7.6	ESI-MS Characterization of Complexes of Platinum(II) and Nucleobases .....	46
1.4.7.7	ESI-MS Characterization of DNA Triplexes .....	47
1.4.7.8	ESI-MS Characterization of RNA Noncovalent Interactions .....	47
1.4.7.9	ESI-MS Characterization of RNA Duplexes .....	49
1.4.7.10	Interaction of Cisplatin with DNA .....	50
1.4.8	Charge Detection ESI-MS .....	50
1.4.9	Atmospheric Pressure Photoionization MS .....	50
1.5	Tandem Mass Spectrometry .....	51
1.5.1	FAB-MS/MS .....	51
1.5.1.1	FAB-MS/MS Characterization of Nucleosides .....	51
1.5.1.2	FAB-MS/MS Characterization of Nucleotides .....	52
1.5.2	ESI-MS/MS .....	53
1.5.2.1	MS/MS Characterization of Nucleosides .....	53
1.5.2.2	MS/MS Characterization of Ribonucleosides .....	58
1.5.2.3	MS/MS Characterization of Geometric Deoxynucleoside Adducts .....	58
1.5.2.4	MS/MS Characterization of Nucleotides .....	59
1.5.2.5	CID-MS/MS Sequencing of Small Oligonucleotides .....	60
1.5.2.6	MS/MS Characterization of DNA Complexes .....	62
1.5.2.7	MS/MS Characterization of DNA Adducts .....	63
1.5.3	ESI-FT-ICR-MS <sup>n</sup> Mapping Noncovalent Ligand Binding to Stemloop Domains of the HIV-1 Packaging Signal .....	66
1.6	Novel Mass Analyzers .....	66

1.6.1	Characterization of DNA Noncovalent Complexes Using LTQ-Orbitrap Hybrid Mass Spectrometer .....	67
1.6.2	Ion Mobility Spectrometry .....	67
1.6.2.1	Matrix-Assisted Laser-Desorption Ionization Followed by Ion Mobility Separation and Time-of-Flight Mass Analysis (MALDI-IM-TOF-MS) .....	68
1.7	Summary .....	68
	References .....	68

## 1.1 INTRODUCTION

Nucleic acids and their components play an important role in a variety of fundamental biological processes.<sup>1</sup> The elucidation of the molecular structure of DNA in 1953 by Watson and Crick led to the so-called central dogma of molecular biology.<sup>2</sup> This dogma proclaims that DNA is transcribed to form a messenger RNA (mRNA) of a complementary sequence that directs its own replication, and the sequence of nucleobases in the RNA is then translated to the corresponding sequence of amino acids to form a protein.<sup>2</sup> In addition to mRNA, two other RNA species, transfer RNA (tRNA) and ribosomal RNA (rRNA), are required for this translation process. Nucleic acid research has consequently had a most profound impact on molecular biology, initiating an avalanche of new disciplines. The field of genomics has evolved over the last few years, culminating in the sequencing and characterization of the entire human genome.<sup>3</sup> Current challenges include the unraveling of the relationship between the genome and the proteome. The genome can be used to predict the total potential proteome, including all of the modifications that can be carried out after the initial translation from mRNA to protein. Mass spectrometry (MS) has had a major role in elucidating the regulatory mechanism involved in the progression from gene to functional protein.<sup>4</sup>

Nucleotides represent the primary level for the storage and transmission of genetic information. They are not only crucial to the organization and execution of protein synthesis but also function as secondary messengers, metabolic regulators, and components of vitamins.<sup>5</sup> They also act as high-energy intermediates driving thermodynamically unfavorable enzyme-catalyzed reactions. Two classes of heterocyclic nitrogenous bases are found in nucleic acids; the two purine derivatives, adenine (A) and guanine (G), and the three pyrimidine derivatives, uracil (U), thymine (T), and cytosine (C). There have been many studies devoted to the separation, isolation, and purification of nucleic acids from biological samples as well as the separation and identification of constituents of hydrolyzed RNA and DNA at the nucleotide, nucleoside, and nucleobase levels.<sup>6</sup> Elucidation of the MS fragmentation patterns for simple nucleosides and nucleotides is a key step in developing an understanding of the MS fragmentation patterns of more complex oligonucleotides and nucleic acids. Seminal reviews by McCloskey and collaborators have established the fundamental importance of MS techniques and applications to nucleic acid research and other related biotechnological fields.<sup>7-9</sup>

The MS analysis of oligonucleotides is of primary importance since establishing the size, purity, and sequence of nucleic acids is a prerequisite to their use as molecular probes in biomedical science. Significant progress in the area of accurate mass determination, sequencing, and study of noncovalent interactions has been made possible by the use of MS ionization techniques such as matrix-assisted laser desorption/ionization coupled to time-of-flight (MALDI-TOF) and electrospray ionization (ESI) coupled to conventional analyzers and to tandem mass spectrometric (MS/MS) instruments.

Nucleic acids and their constituents, due to their polarity and thermal lability, have presented a considerable challenge for analysis by MS. Mixtures of a variety of nucleobases and nucleosides can be analyzed by combined gas chromatography-mass spectrometry (GC-MS), although it is usually

necessary to convert them to more volatile derivatives. The GC-MS analysis of more complex nucleosides, with polar modifications in either the glycone (glycosyl portion) or the nucleobase (aglycone), had not been very successful.<sup>10</sup> However, with the development of newer ionization methods, it has become feasible to perform, without any derivatization, MS analysis on nucleosides, nucleotides, and other large related oligonucleotides (DNA and RNA).

Since the 1960s, it has been known that the mutagenic activity of many substances is linked to the degree with which they covalently bind to DNA. The analysis and measurement of these carcinogen-DNA adducts have been proposed as a means of evaluating the extent of human exposure to specific carcinogens and of monitoring the presence and effects of these specific carcinogens in the environment.<sup>6,11a-c,12</sup> Highly sensitive and structurally informative analytical techniques are thus needed for the general characterization of DNA modifications. MS has the potential to provide structural information and, as such, has played an important role in the structural elucidation of covalently modified nucleic acids.

Many ionization modes, such as electron ionization (EI) and chemical ionization (CI), have been reviewed and employed for the structural elucidation of nucleic acids, their constituents, and their covalently modified nucleic acids.<sup>7-10,12</sup> All of these techniques provide, when coupled with MS/MS, structural information for either single nucleobase-adducts or derivatized single nucleobase adducts. In general, these methods suffer from either a lack of sensitivity necessary for the analysis of trace levels of components derived from small amounts of DNA, or the need to use chemical derivatization techniques to increase volatility and decrease thermal lability. The new technologies developed in the last century, such as fast atom bombardment-mass spectrometry (FAB-MS), MALDI-TOF-MS, ESI-TQ-MS (triple quadrupole, TQ), quadrupole-ion-trap (QIT), and Fourier transform ion cyclotron resonance (FT-ICR), hybrid instruments such as the quadrupole orthogonal time-of-flight (Q-TOF), quadrupole-ion-trap-time-of-flight (QIT-TOF), and quadrupole-ion-trap-Fourier transform-ion cyclotron resonance-tandem mass spectrometry (QIT-FT-ICR-MS/MS) instruments have allowed the analysis of nonderivatized adducts of nucleosides and nucleotides. In addition, capillary electrophoresis (CE) coupled to MS/MS has been used for adduct detection.<sup>6,11a-c,12</sup>

Thus, the intense interest in the MS analysis of nucleic acids and their components has produced a wealth of published work. The major objectives of this chapter are to put into perspective the progress in current biotechnological research in this area, highlight the most popular ionization methods, and illustrate the diversity of strategies employed in the characterization and sequencing of DNA and RNA oligomers, nucleosides, nucleotides, and adducts of biomedical importance.

## 1.2 GAS-PHASE TECHNIQUES

The analysis of nucleic acids and derived products isolated from complex biological matrixes is synonymous with analysis at very low concentrations. Thus, tremendous efforts have been made to develop and implement MS methodologies for the study of the monomeric constituents of nucleic acids.<sup>4,13-15a-c</sup> Biemann and McColskey elucidated detailed fragmentation pathways for the major ions present in both electron ionization (EI) and chemical ionization (CI) mass spectra of simple purine and pyrimidine nucleosides.<sup>16-18</sup> Derivatization, which is of value in many cases, requires an additional sample manipulation step, and may produce unwanted and unexpected side products.<sup>19-22</sup>

### 1.2.1 ELECTRON IMPACT IONIZATION-MS FOR THE DETECTION OF DNA ADDUCTS

So far the biological significance of cytosine methylation has not been fully understood, but there is growing evidence suggesting a link between the development of human cancer and the perturbation of the methylation patterns due to the lack of DNA cytosine methyltransferase (Mtase) substrate specificity. A novel *in vitro* assay called mass tagging was developed, which permitted quantification of the DNA substrate preferences of cytosine Mtases.<sup>23</sup> This approach involved the labeling of target cytosine residues in the synthetic DNA duplexes with stable isotopes such as <sup>15</sup>N. Methylation

was then measured by GC-EI-MS to determine the formation of 5-methylcytosine (5 mC) and DNA. Substrate selectivity was then determined through MS by the absence or presence of the label in 5 mC. This mass-tagging approach has proven to be a powerful tool for examining the substrate selectivity of cytosine DNA Mtases and could easily be adapted to the study of other kinds of selectivity.

### 1.2.2 ELECTRON-CAPTURE IONIZATION-MS FOR THE DETECTION OF DNA ADDUCTS

Electron capture (EC) is a sensitive ionization technique for MS analysis, providing selectivity toward electrophoric compounds. Advances in instrumentation have led to a more widespread application of this method in biomedical and pharmaceutical analysis.

Fedtke and Swenberg contributed a short review of MS techniques for DNA adducts, with emphasis on electron-capture negative ion chemical ionization-mass spectrometry (EC-MS) detection, especially with regard to their method for *N*<sup>2</sup>,3-ethenoguanine.<sup>24</sup> They included in their work, a summary of earlier and more general reviews on the detection of DNA adducts as well as the normal constituents of DNA. Similarly, Giese et al. emphasized their own work in a short review on the same general subject, but with more attention to method development and instrumental aspects of EC-MS.<sup>25</sup> Detection of DNA adducts by means of different types of ionization techniques has also been reviewed.<sup>26–29</sup>

Electrophores are compounds that can be readily detected by the two closely related gas-phase detection techniques of EC-MS and electron-capture detection (ECD). These two analyses require that the electrophore be both thermally stable and volatile.<sup>30,31</sup> Usually, the compound is delivered to one of these detectors as a peak eluting from a GC column, although a direct insertion probe and a liquid chromatography (LC) belt interface can also be linked to the EC-MS instrument.<sup>30,31</sup>

In ECD, a gas-phase current of electrons is monitored, and the electrophore reduces the current by capturing one of the electrons. Similar EC takes place in the EC-MS, but under partial vacuum conditions, and the ion(s) formed from the analyte are measured instead. Once an electrophore captures an electron, it becomes an anion radical and it can either exist long enough to reach the detector or immediately fragment.

For standards, the two detectors are similarly sensitive, but the greater specificity of the EC-MS makes it considerably more powerful for biological samples. Aside from special analytes that are unusually volatile or nonvolatile, it is difficult to reach the low picogram level by GC-ECD when real samples requiring electrophore derivatization are tested.<sup>32,33</sup> Largely, this is because electrophoric contaminants are present or formed in the reagents and solvents during sample preparation.<sup>34</sup>

DNA adducts, due to their lack of volatility and thermal stability, are not good electrophores for detection by GC-EC-MS. Therefore, as part of sample preparation, one or more chemical reactions are necessary to make the DNA adducts available for detection by GC-EC-MS.

It must be kept in mind that the overall electrophore response in the GC-EC-MS is a combination of the recovery of the compound from the GC part into the EC ion source; the inherent EC properties of the compound; other events (e.g., wall effects) in the ion source; and the transmission efficiency (including the lifetime) of the electrophore anion to the detector.<sup>35</sup> A strong electrophore can be detected as a pure standard at the low attomole level on a modern instrument.<sup>36</sup>

Typically, the DNA adduct is converted into an electrophore, which is detected as either a parent anion radical (no dissociative EC) or as an exclusive anion fragment (dissociative EC), which is formed along with the corresponding neutral. Either way, the electrophore that is typically selected forms a single ion (aside from the associated isotope peak) maximizing the sensitivity. Ideally, the electrophore-labeled DNA adduct would give two or three anion fragments in similar amounts, and their ratio would be checked for each sample as an extra guard against interferences. This would be worth the slightly higher detection limit.

Although some standards of derivatized DNA adducts can be detected by GC-ECD (Table 1.1), DNA adducts that are derived from biological samples are usually only detectable by GC-EC-MS.



**TABLE 1.1**  
**DNA Adducts and Monomers Detected by Electrophore-Labeling GC-MS or GC-ECD (Detection is by GC-EC-MS Unless Indicated Otherwise)**

Type of Adduct	Target on DNA	Sample Preparation		Reported Detection Limit		Reference
		Procedure	Yield (%)	Standard	DNA	
4-Aminobiphenyl	C8-G (mostly)	DNA/NaOH/hexane/PFP	—	—	0.32 in 10 <sup>8</sup> (100 g of DNA)	38,39
2-Aminofluorene	C8-G	NH <sub>2</sub> NH <sub>2</sub> /PFBzAl	—	25 amol (S/N = 10)	—	40,41
Benzo[ <i>a</i> ]pyrene	N2-G (mostly)	DNA/HCl/heat/KO <sub>2</sub> /PFBz/Si-SPE	22	14 amol (S/N = 20)	5 in 10 <sup>7</sup> (100 g of DNA)	36,42
	—	DNA/HCl/heat/CH <sub>3</sub> I/C <sub>6</sub> H <sub>6</sub>	—	1 fmol	—	43
Cytosine	—	PFB/Me/GC-ECD	35	7 fg	—	44-46
Etheno	N <sup>2</sup> ,3-G	DNA/HCl/IEC/C18-Si-SPE/PFBz/Si-SPE	—	190 amol (S/N = 10)	60 fmol/mol of G for 1 mg of DNA	24,47
	O <sup>2</sup> -T	PFBz/TLC/Si-FC	56	25 pg (S/N > 100)	—	48
Ethyl	O <sup>4</sup> -T	PFBz/TLC	57	—	—	48,49
Ethyl	O <sup>4</sup> -Tdn	PFBz/TLC/GC-ECD	55	0.44 fmol	—	30,50
Hydroxy	C8-dG	DNA/nucleases/TFA-acid/hydrazine/C18-SPE/ acetyl/hydrolysis/ PFBz/Si-SPE	—	1.8 fmol	4 in 10 <sup>6</sup> for 30 g of DNA	51,52
	C8-dG	Urine/C18(OH)-Si-SPE/acetyl/PFBz/HPLC/ EtOAc	—	1.8 fmol	1.8 pmol in 50 L of urine	51,52
Hydroxy	CH <sub>3</sub> on T	PFBz/TLC/Si-FC/PFBz/TLC	—	200 zmol (S/N = 3)	—	53
4'-Hydroxybutyl	O <sup>6</sup> -G	DNA/HCl/Ab/PFBz/TMS	—	—	0.12-3 mol/mol of G	54
2'-Hydroxyethyl	N7-G	DNA/heat/HCl/HNO <sub>2</sub> /PFBz/Si-SPE	9.7	1.3 amol (S/N = 10)	100 pg in 100 g of DNA	36,55,56

Source: Adapted from Kellersberger, K.A. et al., *J. Am. Soc. Mass Spectrom.* 16, 199, 2005. With permission.