Valérie Gabelica *Editor*

Nucleic Acids in the Gas Phase

Physical Chemistry in Action

Valérie Gabelica

Editor

Nucleic Acids in the Gas Phase

Springer

Editor
Valérie Gabelica
Université de Bordeaux
Institut Européen de Chimie et Biologie
Pessac
France

U869 ARNA Laboratory (Inserm/Univ. Bordeaux)
Institut Européen de Chimie et Biologie
Pessac
France

Preface

Laboratory sciences have bloomed with a variety of techniques to decipher the properties of the molecules of life. Physical chemists and chemical physicists approach the problem by isolating molecules and ruthlessly dissecting them with a variety of tools. One particular way to isolate molecules is to isolate them from their solvent environment. The specific advantages are twofold. First, stripping biomolecules from the solvent makes them amenable to mass spectrometry analysis. Second, isolating the molecules from their solvent enables one to study their *intrinsic* properties. Besides their mass, these properties can be the molecules' structure (from atom connectivity to tridimensional structure), their spectroscopic properties, or their reactivity (for example, with photons, electrons, free radicals, ions, or other molecules) in well-defined energetic conditions. Gas-phase approaches therefore constitute a category of techniques with their own instrumentation, theoretical approaches, and rules for data interpretation.

The objective of this book is to bridge the gap between communities. On the one hand, it aims to give physical chemists a broader view of the potential biological applications of the techniques they develop. On the other hand, the book aims to give chemists, biophysicists, biochemists, molecular biologists, and pharmacologists novel insight into the ways gas-phase techniques can bring unique answers to new types of questions.

Book Outline

This volume is divided into two parts: (I) Methods and (II) Applications.

Part I—Methods—introduces techniques used to investigate the properties of nucleic acids in the absence of solvent and key results: how to transfer nucleic acids from the condensed phase to the gas phase (Chap. 2), the fate of large nucleic acids in the gas phase according to molecular simulations and ion mobility spectrometry experiments (Chap. 3), interaction of nucleic acids in the gas phase with electrons (Chap. 4) or photons (Chap. 5), and gas-phase fragmentation pathways (Chap. 6). Some of these techniques are of course common to the investigation of other

categories of biomolecules (e.g., proteins), but we will highlight here the specificities pertaining to nucleic acid studies.

Part II—Applications—illustrates with four examples the use of gas-phase physicochemical approaches to solve specific questions from fields as diverse as *molecular biology*, with the sequencing of large RNA (Chap. 7), *human health*, with the characterization and quantification of nucleic acid modifications resulting from photodamage (Chap. 8), *pharmacology*, with the screening of drug interactions with nucleic acid targets (Chap. 9), or *structural biology*, with the characterization of RNA folding by mass spectrometry-based approaches (Chap. 10). This list of potential applications is by far not exhaustive, but the key take-home message lies elsewhere. Each of these chapters underlines in its own way the crucial importance of understanding the fundamental physical principles driving nucleic acid structure and reactivity in the gas phase, to conceive tailor-made approaches to solve important problems.

Acknowledgments

Pessac, France   Valérie Gabelica

Reference

1. Schermann J-P (2008) Spectroscopy and modelling of biomolecular building blocks. Elsevier, Amsterdam

Contents

List of Contributors

Annalisa Arcella Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

Joint Research Program in Computational Biology, Institute for Research in Biomedicine and Barcelona Supercomputing Center, Barcelona, Spain

Steen Brøndsted Nielsen Department of Physics and Astronomy, Aarhus University, Aarhus C, Denmark

Kathrin Breuker Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck, Austria

Jennifer S. Brodbelt Department of Chemistry, University of Texas, Austin, TX, USA

Thierry Douki Laboratoire "Lésions des Acides Nucléiques", Université Joseph Fourier – Grenoble 1/CEA/Institut Nanoscience et Cryogénie/SCIB, Grenoble, France

Daniele Fabris The RNA Institute, University at Albany (SUNY), Albany, NY, USA

Valérie Gabelica IECB, ARNA Laboratory, Univ. Bordeaux, Pessac, France

U869, ARNA Laboratory, Inserm, Bordeaux, France

Yang Gao Department of Chemistry, Purdue University, West Lafayette, IN, USA

Gilles Grégoire Laboratoire de Physique des Lasers, Université Paris 13, Sorbonne Paris Cité, Villetaneuse, France

CNRS, UMR 7538, LPL, Villetaneuse, France

Scott A. McLuckey Department of Chemistry, Purdue University, West Lafayette, IN, USA

Modesto Orozco Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

Joint Research Program in Computational Biology, Institute for Research in Biomedicine and Barcelona Supercomputing Center, Barcelona, Spain

Department of Biochemistry and Molecular Biology, University of Barcelona, Barcelona, Spain

Guillem Portella Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

Joint Research Program in Computational Biology, Institute for Research in Biomedicine and Barcelona Supercomputing Center, Barcelona, Spain

Jean-Luc Ravanat Laboratoire "Lésions des Acides Nucléiques", Université Joseph Fourier – Grenoble 1/CEA/Institut Nanoscience et Cryogénie/SCIB, Grenoble, France

Frédéric Rosu UMS 3033 and Inserm US001, IECB, CNRS, Pessac, France

Zhe Xu Department of Chemistry, University of Texas, Austin, TX, USA

Part I

Methods

Introduction: Nucleic Acids Structure, Function, and Why Studying Them In Vacuo

1

Valérie Gabelica

Abstract

This introductory chapter sets the stage for the various methods and application that will be described in the book "Nucleic acids in the gas phase." Using key review articles as references, nucleic acid structures are introduced, with progression from primary structure to the main secondary, tertiary, and quaternary structures of DNA and RNA. Nucleic acid function is also overviewed, from the roles of natural nucleic acids in biology to those of artificial nucleic acids in the biotechnology, biomedical, or nanotechnology fields. Importantly, the question of *why* studying nucleic acids *in the gas phase* is addressed from three different points of view. First, because isolated molecules in vacuo cannot exchange energy with their surroundings, reactivity can be studied in well-defined energetic conditions. Second, isolating molecules from their solvent and environment allow to study their intrinsic properties. Finally, the rapidly expanding field of mass spectrometry, an intrinsically gas-phase analysis method, calls for better understanding of ion structure and reactivity in vacuo.

V. Gabelica (✉)
IECB, ARNA Laboratory, Univ. Bordeaux, 33600 Pessac, France

U869, ARNA Laboratory, Inserm, 33000 Bordeaux, France
e-mail: valerie.gabelica@inserm.fr

Abbreviations

A	Adenine
BIRD	Blackbody infrared radiation-induced dissociation
bp	Base pair
C	Cytosine
DNA	Deoxyribonucleic acid
FDA	Food and Drug Adminstration
G	Guanine
G4-DNA	G-quadruplex DNA
IR	Infrared
LNA	Locked nucleic acid
miRNA	Micro RNA
mRNA	Messenger RNA
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
ncRNA	Noncoding RNA
NMR	Nuclear magnetic resonance
nt	Nucleotide
ODN	Oligodeoxynucleotide
PDB	Protein data bank
PNA	Peptide nucleic acid
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rRNA	Ribosomal RNA
SASA	Solvent-accessible surface area
siRNA	Silencing RNA
T	Thymine
TFO	Triplex forming oligonucleotide
tRNA	Transfer RNA
U	Uracil
UV	Ultraviolet
VEGF	Vascular endothelial growth factor

1.1 Nucleic Acids Structure

This section summarizes the fundamentals and most common nomenclature for nucleic acid structure. The connections between primary, secondary, and tertiary structure are highlighted. Also, the attention of the reader is drawn to the slightly different meanings of "secondary" and "tertiary" structures in the case of nucleic acids compared to the field of proteins.
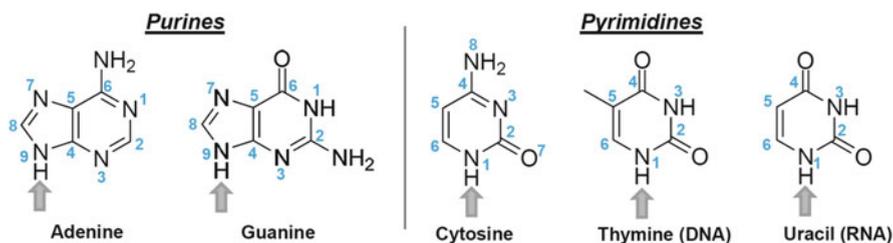
Fig. 1.1 Chemical structure of the five natural nucleobases, in their natural tautomer form (other tautomers are shown in Chaps. 2 and 5). The nomenclature of atom numbering is in *blue*. The position of attachment to the sugar is shown by the *arrow*

1.1.1   Nucleobases, Nucleosides, and Nucleotides

While proteins are expressed with an alphabet of 20 amino acids, nucleic acids language contains 4 main letters: A (adenine), G (guanine), C (cytosine), and either T (thymine) for DNA or U (uracil) for RNA. The chemical structures of the natural nucleobases and the atom numbering nomenclature are given in Fig. 1.1. Note that only the natural tautomer is shown. Tautomers are chemical isomers differing by the hydrogen/proton location. Many nucleic acid scientists ignore their variety, but gas-phase scientists must be aware of them. Tautomer occurrence in the gas phase will be further discussed in Chaps. 2 and 5. In addition to the prevalent natural nucleobases, many other modified bases can be found in nature (for reviews, see [1–3]).

A nucleoside consists of a nucleobase attached to a five-carbon sugar (in green in Fig. 1.2), which is either a ribose for RNA or a deoxyribose for DNA. The difference is a OH group on the 2′ position of the sugar. Nucleotides are the monomers of nucleic acids. They consist of a nucleoside and at least one phosphate group. Figure 1.2 shows the structures of nucleotides with a planar chemical representation, whereas Fig. 1.3 highlights important conformers of nucleotides. Figure 1.3a highlights the two main sugar conformations: C3′-endo and C2′-endo. To visualize them in a tridimensional structure, one has to place the oxygen of the ribose in the back and the 3′ position in the front left. The C3′-endo and C2′-endo are distinguished by the orientation taken by the three bonds highlighted in red in Fig. 1.3a. When the three bonds form a tilted "N," the conformation is C3′-endo. Because of the "N" an ancient but still used nomenclature also calls this conformation "North." When the three red bonds form a tilted "S," the conformation is C2′-endo or "South."

The sugar conformation is linked to base, phosphate, and, in the case of RNA, 2′-OH positions. In the C3′-endo, the base and 2′-OH have axial positions relative to the sugar, whereas the phosphates occupy equatorial positions. In the C2′-endo, the base and 2′-OH are equatorial, and the phosphates are axial. The 2′-OH of riboses (RNA) tend to favor the C3′-endo conformation, whereas deoxyriboses (DNA) are more free to adopt diverse conformations. Another important degree of freedom is
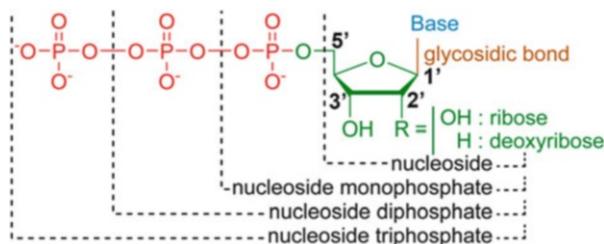
Fig. 1.2 Nucleotide chemical structures. The sugar numbering nomenclature is indicated. Deoxyriboses (constituting DNA) and riboses (constituting RNA) differ by the presence of H or OH on the 2′ carbon
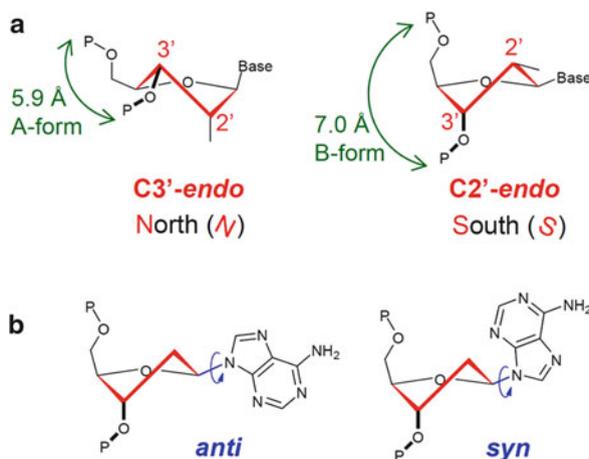
Fig. 1.3 (**a**) Sugar pucker nomenclature. The C2′-endo and C3′-endo conformations are sometimes called "North" or "South" conformations, respectively, because the *red* bonds form the letter "N" or "S" tilted to the *right*. (**b**) Glycosidic bond angle nomenclature (*anti* or *syn*), illustrated with adenosine on the C2′-endo conformation

the rotation of the base around the glycosidic bond (Fig. 1.3b). When the base stays away from the sugar, we have the *anti*-conformation, and when it is rotated closer to the sugar, we have the *syn* conformation. In the C2′-endo conformation, the *anti*-base is energetically favorable, but *syn* conformations can be encountered as well, depending on the environment. In the C3′-endo conformation, because the base is axial, the *syn* is strongly disfavored for steric reasons, and the bases are *anti* [4].

1.1.2 Primary Structure

An oligonucleotide is a base sequence formed by the ligation of the 3′-OH of one nucleotide to the 5′-OH of the next via a phosphate group (see Fig. 1.4a for a four-nucleotide (4-nt) primary sequence). In terms of tridimensional structure, single
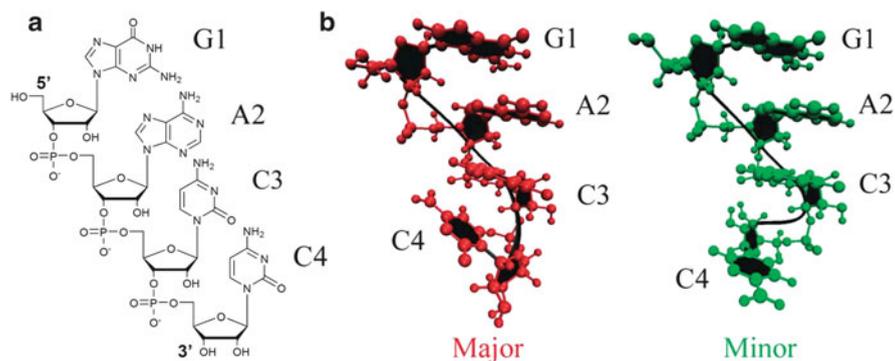
Fig. 1.4 (**a**) Primary structure: plain drawing of the RNA single strand r(GACC). (**b**) NMR and molecular modeling of the tridimensional structure reveal that the unpaired single strand is structured in aqueous solution, with an A-form, evident base stacking in both the major and minor conformer, and occasional inversion of the 3′-terminal cytosine in the minor conformer. (*panel* (**b**) reprinted with permission from [6])

stranded regions are not "unstructured" or "random coils" in solution: bases tend to stack on top of each other and the strand forms a single helix (for a thorough recent discussion of the definition of base stacking, see reference [5]). In particular, RNA is constrained by the preferential C3′-endo/anti-conformation of each nucleotide, and given the spacing between the phosphates and base position, this gives an A-form single helix. DNA single strands can adopt a greater variety of conformers, but stacking between bases remains.

1.1.3 Secondary and Tertiary Structure

The secondary structure of nucleic acids involves hydrogen bonding between bases. The best known hydrogen bonding motif is the Watson–Crick base pairing of guanine with cytosine, and of adenine with thymine or uracil, which leads to double helix structures (Fig. 1.5a). Other base pairs (bp), called noncanonical base pairs, can also occur, either with non-Watson–Crick arrangement of GC or AT/AU base pairs or between different bases (e.g., a GU base pair). The secondary structure is often drawn on a plane, indicating which base is hydrogen-bonded to which one. The tertiary structure is the corresponding tridimensional structure of these motifs (three-dimensional rendering). In the case of double-stranded DNA or RNA, the best known tertiary structures are the right-handed A-helix and B-helix, and the left-handed Z-helix. Because of the C3′-endo conformation favored by the 2′-OH (see Fig. 1.3a), RNA tends to adopt A-helix tertiary structure with all bases in *anti*, whereas DNA tends to adopt a B-helix with all bases in *anti*. An interesting point for the gas-phase chemist is that, while the B-helix is the preferred form in aqueous solution, B-to-A helix transitions are observed upon dehydration [4].
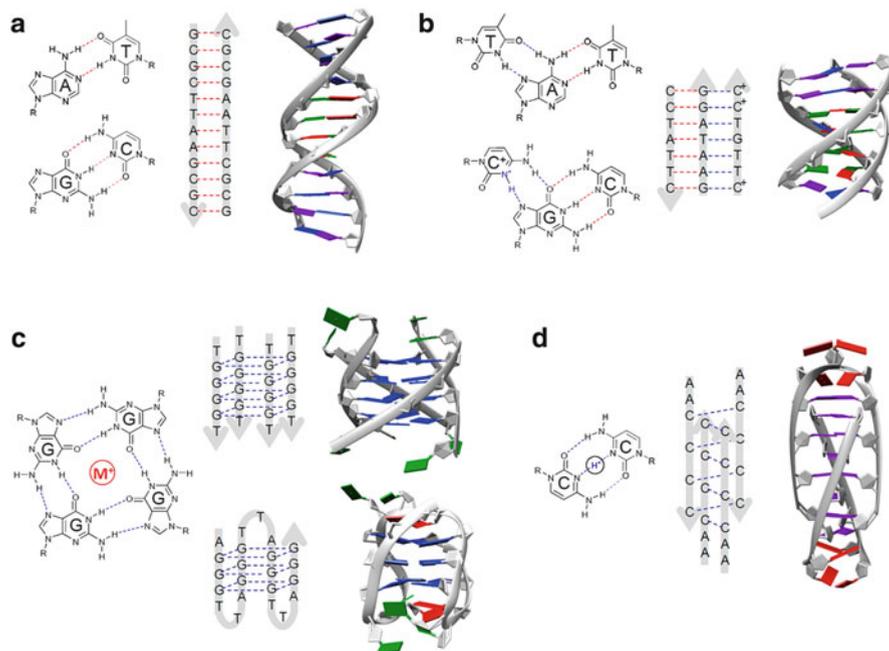
Fig. 1.5 Main hydrogen bonding motifs and secondary structure and tertiary structure of (**a**) a Watson–Crick double helix B-DNA (tridimensional structure from PDB file 1BNA [10]), (**b**) a triple helix DNA (PDB file 149D [11]), (**c**) G-quadruplexes, illustrated with one tetramolecular parallel-stranded quadruple helix (PDB file 2O4F [12]), and one intramolecular antiparallel G-quadruplex structure (PDB file 143D [13]), (**d**) a tetramolecular i-motif structure (PDB file 1YBL [14]). Adenines are in *red*, guanines are in *blue*, cytosines are in *purple*, and cytosines are in *green*

Figure 1.5 also shows several other types of base hydrogen bonding motifs (triplets—Fig. 1.5b; G-quartets that are stabilized by monovalent cations—Fig. 1.5c; and proton-mediated cytosine base pairs—Fig. 1.5d). Examples of secondary and tertiary structures containing these motifs are also shown. Figure 1.5b shows a triple helix DNA, also called "triplex DNA" (for a review, see [7]). Figure 1.5c shows G-quadruplex DNA (G4-DNA), meaning structures that contain at least two stacked guanine quartets (for a detailed review, see [8]): a tetramolecular G-quadruplex—Fig. 1.5c top and an intramolecular G-quadruplex—Fig. 1.5c bottom. Finally, sequences containing adjacent cytosines can form the so-called "i-motif" structure [9], where the cytosines form hemi-protonated base pairs that intercalate alternatively (see Fig. 1.5d). Regarding the nomenclature encountered in the literature, it is important to distinguish the hydrogen bonding motif (a base pair, a triplet, or a quartet) from the molecularity of the assemblies (an intramolecular, a bimolecular, a trimolecular, or a tetramolecular structure). It is therefore particularly important to understand what some authors mean by "duplex" or "dimer" (a double helix or a bimolecular structure?), or "quadruplex"

(a G-quadruplex, i.e., a structure containing G-quartets or a tetramolecular structure?), and to use these terms clearly.

These examples were selected because they will be further discussed in the following chapters, and are the subject of vast amount of literature, but the list is not exhaustive. In particular, X-ray crystallography of RNA structures revealed a great variety of structures with various hydrogen bonding motifs between bases (e.g., the triple helix in Fig. 1.5b contains an unusual A-T/G triple in the middle of the sequence). It is also noteworthy that cations, water molecules, and hydrogen bonds not only between bases but also involving sugars and phosphates sometimes play key roles in the tertiary structures formed. For a key review, see [15].

1.1.4 Quaternary Structure

The quaternary structure refers either to nucleic acid interactions with other molecules, which can be other nucleic acids, metabolites, or proteins. For example, for chromosomic DNA, it refers to the interaction of the double helix with histones and the higher-level organization in the form of chromatin. Examples of quaternary structures involving RNA and proteins are the ribosome (multi-protein and RNA, i.e., a ribonucleoprotein machinery that links amino acids together to form peptides) or the spliceosome (ribonucleoprotein machinery that processes the messenger RNA by cutting out the parts that will not be found in the final peptide or protein). Both ribosomes [16] and spliceosomes [17] have been investigated by mass spectrometry, a gas-phase analysis technique.

1.2 Nucleic Acids Function

1.2.1 Molecules of Life

1.2.1.1 DNA

Nucleic acids are central to life. The central dogma of molecular biology is as follows: DNA is the repository of the genetic information, which will code for the synthesis of RNA (the DNA → RNA step is called "transcription"), then proteins (the RNA → protein step is called "translation"). The genetic information must be transferred from parent cell to daughter cells with high fidelity, and the DNA double-stranded structure with the A–T/G–C base pairing mode ensures this high fidelity upon DNA replication. The cells also host specialized enzymatic machineries whose role is to ensure this high fidelity and to repair damaged sites like strand breaks or mutations. A single damage, if occurring in the wrong place, can cause a disease.

Understanding DNA damage is a field where physical chemistry in general and gas-phase methods in particular have contributed either to analyze the resulting damage or to understand the cause-effect relationships. The damage can have physical or chemical causes: interaction with UV light causing cross-linking

between bases (see Chap. 8), interaction with ionizing radiation causing strand breaks, interaction with reactive oxygen species (ROS) causing oxidation of bases, or interaction with hazardous chemicals (e.g., mustard gases chemical weapons) that form adducts with bases. It is also important to underline that not all DNA modifications are "bad." Some forms of alterations, like the methylation of cytosines, have important regulatory roles in controlling the gene expression. Some chemical agents forming covalent or non-covalent adducts (see Chap. 9) are commonly used in chemotherapies. Duplex DNA binders are general cytotoxic agents that act by perturbing the replication process. Their anticancer properties stem from the fact that cells that divide more frequently, like cancer cells, are more affected by the treatment than somatic cells. Although duplex DNA binders have proven efficacy and are still used in current chemotherapy regimens, their lack of specificity nevertheless causes undesirable side effects.

In recent years, the interest has moved towards peculiar secondary and tertiary structures of DNA that can form in specific sequence locations or at specific moments in the cell cycle when the two strands are transiently separated (e.g., during replication of DNA or during transcription, i.e., when one DNA strand is read and transcribed into RNA). Importantly, the formation of non-B-DNA structures such as G-quadruplexes is linked to specific biological effects [18–20]. This is why structural, thermodynamic, and ligand binding studies on such non-B-DNA structures are currently such a hot topic. There too, gas-phase methods, in particular mass spectrometry, have brought significant contribution.

1.2.1.2 RNA

In the central dogma of molecular biology (DNA → RNA → protein), the RNA that will be translated into protein is called the messenger RNA (mRNA). Other RNAs are involved in the translation step (the RNA → protein step). rRNA (ribosomal RNA) is the name of RNAs that constitute the ribosomes, which are translation machineries. tRNA (transfer RNA, Fig. 1.6) are ~80 bases long RNAs that contain the three-letter nucleic acid codons that will select the amino acid to be included in the peptide chain. A very common secondary structure of RNA is the hairpin structure: parts of the chain are self-complementary and form a local double-stranded structure. The double stranded stems are connected by loops. Importantly, modified bases are very commonly found in naturally occurring RNA, as illustrated in Fig. 1.6 for a typical tRNA. Applications of mass spectrometry to RNA sequencing are highlighted in Chap. 7.

Besides RNA's role of transforming the genomic DNA information into proteins, research in the last decade has revealed that cells express a variety of noncoding RNAs, and novel RNA functions are discovered every day. Importantly, many RNAs exert roles in *regulation* of cell processes, a role that was before thought to be ensured only by proteins. Some of these noncoding RNAs (ncRNA) are long, but some are small regulatory RNAs with sizes entirely compatible with high resolution mass spectrometry analysis, for example, siRNA (silencing RNAs) [21], which are 19 bp–25 bp (base pair) RNA duplexes or miRNA (microRNAs) [22], which are ~22 nt single strands.
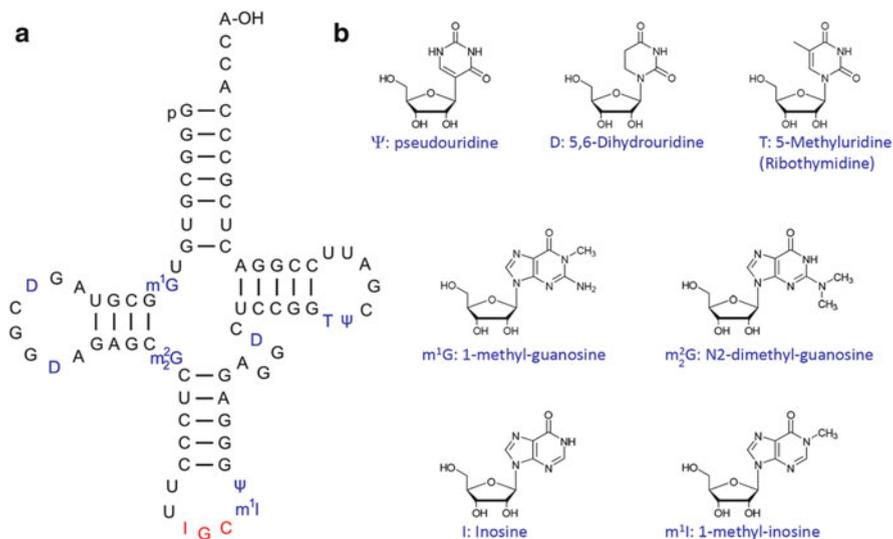
Fig. 1.6 (**a**) tRNA^{Ala} from *S. cerevisiae*. The codon for alanine is shown in *red*. The primary structure contains seven different modifications (bases different from the canonical A,U,G,C, as shown in *Panel* (**b**)), for a total of 10 modified bases out of a total of 76 bases. The structures of the modified bases are shown on the *right*. Panel (**a**) is reproduced from a file licensed under the Wikimedia Creative Commons Attribution-Share Alike 3.0 Unported license

These new regulation mechanisms open broad perspectives both in fundamental research and in pharmacology applications. Instead of targeting the expressed proteins with drugs, one can imagine regulating their production directly from the source, by targeting their regulatory nucleic acids. Importantly, functions are tightly linked to folding into peculiar tridimensional structures or to conformational transitions between structures [23, 24]. The analysis of RNA tridimensional structures with mass spectrometry (MS)-based approaches is covered in Chap. 10.

1.2.2 Artificial Nucleic Acids in Biotechnology

Oligonucleotides are synthesized using the phosphoramidite chemistry, developed in 1981 [25]. The possibility to order oligonucleotides of well-defined length and sequence revolutionized the field of nucleic acid biophysics: the objects studied were much better defined than natural DNA extracted from cellular material. The first biomedical application envisaged for synthetic oligonucleotides was the control of gene expression by the formation of a triple helix with a target duplex DNA sequence [7]. Modified triplex-forming oligonucleotides (TFOs) could be used to bring a drug or cleaving agent close to a gene, targeted thanks to strand specificity, in order to regulate gene expression. Another strategy was the targeting of single-

stranded RNA with a complementary oligodeoxynucleotide (ODN), called the antisense ODN [26].

In the above examples, knowledge of base pairing rules and the principle of complementarity served as rationales for the sequence design. A totally different approach, not based on rational design but on "natural selection," was developed to find *aptamer* nucleic acids: oligonucleotide sequences that are able to bind to selected targets [27, 28]. The variety of targets is infinite. First aptamers were developed against small molecules, proteins, or other nucleic acids, but nowadays aptamer targets range from a single fluorine atom to entire cells. The range of applications of aptamers is also extremely varied, from analytical applications as sensors to medicine, with, for example, an aptamer called Macugen targeting the protein VEGF, approved by the FDA for the treatment of age-related macular degeneration [29].

For many diagnostic or therapeutic applications, the artificial oligonucleotide will be in the presence of biological material containing enzymes called DNAses and RNAses, which are capable to degrade DNA and RNA, respectively. To avoid undesired degradation of the oligonucleotide, chemically modified oligonucleotide backbones have been engineered. Examples are the use of 2′-*O*-methyl-substituted RNA, phosphorothioate, or methylphosphonate groups instead of phosphate groups, peptide nucleic acid (PNA) [30] or locked nucleic acid (LNA) [31] backbones. Also, many applications being based on fluorescence techniques, a wide range of fluorescent base modifications have been developed [32].

The various biological applications outlined above may seem far away from physical chemistry aspects treated in the present book series, but the goal of this brief overview is twofold. First, I wanted to outline the variety of molecules that are available either commercially or from academic labs, and which may become the model objects of future physicochemical studies. Second, most applications actually come down to be based on physical properties such as molecular recognition, reactivity, or electronic properties (see below). Understanding the fundamental principles that would allow to predict these properties from the chemical structure is of prime importance, not only for the fun of science but also to help designing tomorrow's applications.

1.2.3 Nanotechnology Applications

Applications of artificial nucleic acids in biotechnology nowadays go well beyond the introduction of a single sequence in a living system. A whole new area of science, generally called "synthetic biology," aims at bringing artificial functional systems into life, demonstrated by recent examples of introduction of RNA devices in bacteria [33], or the creation of nonnatural nucleic acid-based systems amenable to replication [34]. Here we are at the frontiers between biotechnology and nanotechnology.

Now purely in the world of nanotechnology, nucleic acids can be used for their capacity to form predictable bidimensional and tridimensional structures based on

the principles outlined above. Typical examples are the nanostructures called "DNA origamis," first published by Paul Rothemund [35], that used Watson–Crick base pairing motifs to create 2D shapes at will. For a recent review, see [36]. Besides structural scaffolding applications, specific nucleic acids properties can also be used to perform function and conceive a variety of nanostructures, sensors, or stimuli responsive systems that can even function as logic gates [37–40].

Another example of function is charge transport. Simply put, one of the initial questions on the engineering point of view was whether DNA would be able to conduct electricity. A recent example showed electrical transport through a short G-quadruplex DNA structure [41]. These engineering questions brought about more fundamental physicochemical questions about the nature of charge transport in DNA [42] that inspired some fundamental experimental and theoretical studies in the gas phase. Importantly, it is important to underline that a dehydrated environment is relevant for nanotechnology applications for which water is not necessary or even undesirable.

1.3   Why in the Gas Phase?

1.3.1   Well-Defined Energetics

"In the gas phase" or "in vacuo" means in the absence of bulk solvent. The molecules are isolated. On a statistical mechanics point of view, there is a key difference between condensed phase and gas phase. Molecules in solution are in a canonical ensemble: they can exchange energy with a surrounding heath bath. Isolated molecules in the gas phase are in a microcanonical ensemble: they cannot exchange energy with their environment. The energy of a given molecule is the sum of its translational energy, rotational energy, vibrational energy, electronic energy, and nuclear energy (changes of the latter are not being considered herein). The distribution of internal energy (all degrees of freedom except translation) of a population of molecules (the ensemble) can have different shapes, depending on how the molecules were prepared and brought to a given state. In the presence of a heath bath, the internal energy distribution is a Boltzmann distribution. In the gas phase, this is sometimes the case, for example, in the presence of a gas at sufficiently high pressure that it serves as a heath bath [43], or when the system is equilibrated long enough to establish exchange equilibrium of blackbody radiation with the walls of the instrument [44]. In most other cases, for example when dissociation of the highest-energy molecules depletes the high-energy part of the ensemble, the internal energy distribution differs from a Boltzmann distribution [45].

The key advantage of gas-phase studies is that, when all factors contributing to internal energy build-up or energy dissipation are accounted for, we can study the structure and reactivity of a well-defined system. Collisions, photon absorption (laser irradiation, absorption of blackbody infrared radiation), photon emission (fluorescence, emission if blackbody IR radiation), attachment of electrons of

known kinetic energy, detachment of electrons which kinetic energy can be measured, and fragmentation are phenomena that are all linked by the energy conservation principle. With the knowledge and control of all events affecting the internal energy distribution of molecules or ions in the gas phase, it is possible to study the reactivity of isolated biomolecules with unmatched level of precision. Examples are kinetic energy-resolved collision-induced dissociation studies with guided ion beam tandem mass spectrometers [46] or with the blackbody IR radiation-induced dissociation (BIRD) technique [47]. Both techniques give accurate insight into bond dissociation energetics and, together with theoretical calculations, into fragmentation pathways. They were applied to nucleic acids since their early days [48–50].

1.3.2 Focus on Intrinsic Properties

In the field of chemical physics, the study of chemical processes from the point of view of physics, having a well-defined system, is of prime importance. Solvation is actually not yet a well-understood process, both from its structural, energetics, and dynamics points of view. For nucleic acids, an additional complication comes from the key roles of counter-ions present in solution and their interplay with solvation [51]. In a reductionist approach, interactions can be studied one at a time. This has long motivated the study of biomolecular building blocks in the gas phase, then isolated complexes between these building blocks (e.g., isolated base pairs), isolated molecules that are micro-hydrated in a stepwise manner (e.g., isolated complex with either zero, one, two, three, etc... water molecules attached), isolated biomolecule–metal complexes, and combinations thereof. The reductionist philosophy is that one must first understand the properties of the building blocks in order to lay the basis for understanding more complex molecular architectures, and that understanding the properties of isolated molecules will help understanding their properties in more complex environments. An example is the recurring theme of the emergence of solution-like properties: how many base stacks are required to obtain the same properties of natural DNA? How many water molecules are required to observe the same properties as in bulk solution?

In that vein, removing the solvent in order to understand its role bears analogies with the typical biologists' approach consisting in knocking out a gene in cells or mice and observing the difference between knockout organisms and normal individuals in order to understand the role of that gene, and whether it is essential to life. So, to explain to a biologist why you are studying intrinsic properties of biological molecules *in the gas phase*, you may try and explain that you are doing a "solvent knockout experiment."

Beyond the joke, the idea is worthy of deepening, particularly in the framework of bringing together theory and experiments. Indeed decomposing the effects of the solvent from the effects of purely intermolecular interactions is a very common approach to predict interaction free energies from structures [52, 53]. This decomposition implies a thermodynamic cycle such as the one represented in Fig. 1.7. The

Fig. 1.7 Thermodynamic cycle for folding and binding, including transfer from the solution to the gas phase, illustrated here with a ligand-duplex DNA binding equilibrium. The *single arrow* in the gas phase underlines that most often only the dissociation step can be studied experimentally
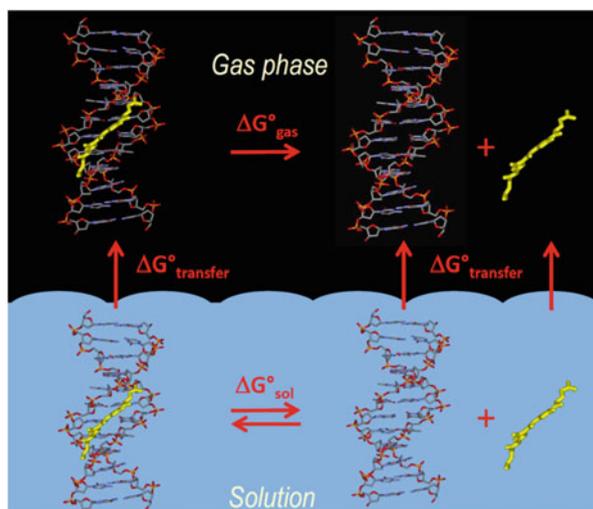
contribution from the differences in free energy of transfer of each component are usually estimated by the difference in solvent-accessible surface area (SASA) in each component, and this computed SASA scales well with the solvophobic contribution to the reaction free energy in solution. Intermolecular interactions, also called intrinsic interactions, can be much more easily computed in the gas phase than in solution, but gas-phase measurements are the only way to verify those computations experimentally. As outlined in the above section, the energetics of unimolecular dissociation (indicated by the single arrow) in the gas phase can in some cases be *measured* accurately. Examples of equilibrium measurements in the gas phase, although more rare, do also exist [54].

1.3.3 The Many Faces of Mass Spectrometry

Finally, another good reason to get interested in what happens to nucleic acids in the gas phase is the booming field of mass spectrometry (MS), and it should come as no surprise that most application chapters in this book deal with mass spectrometry. A mass analyzer separates ions according to their mass-to-charge ratios based on ion movement under electric or magnetic fields. Mass analysis is performed in vacuum so that the ion trajectory is not perturbed by collisions. Mass spectrometers, which can also be coupled to chromatography methods, therefore serve for mass measurement of molecules present in a sample, and with the use of appropriate standards, for quantification. Today's mass spectrometers come with such sophisticated hardware and software that they can often be used as black boxes for mass measurement and quantification, while most of the research and development is focused on the sample preparation and purification aspects. However, when the fundamental principles of what happens to a molecule inside the mass spectrometer are

understood, mass spectrometry can bring much more information than black box machines are essentially conceived for.

Primary structure determination can be inferred from fragmentation experiments, called tandem mass spectrometry (MS/MS) experiments. The different ways of inducing fragmentation, and the fragmentation pathways adopted by different nucleic acids, are covered in Chaps. 6 and 7. Although in the late 1990s mass spectrometry was supplanted by other high-throughput sequencing methods for genome-wide analyses, it remains a crucial technique when the alphabet is not just consisting of A, T/U, G, and C's, for example, for sequencing tRNAs with their numerous modified bases (see Fig. 1.6).

Secondary and tertiary structure analysis by mass spectrometry is the most challenging aspect. Because a mass spectrometer, by definition, only measures masses, tricks have to be imagined to obtain information on the structure of the ions. For example, nucleic acids can be studied by chemical labeling in solution followed by MS analysis of the extent and location of the labels (see Chap. 10 and references therein).

In contrast with in-solution probing, direct probing in the gas phase offers an important advantage when analyzing mixtures: the different constituents of the solution can be first separated in mass. Gas-phase probing methods include ion mobility spectrometry (see Chap. 3), ion spectroscopy (see Chap. 5), gas-phase ion-molecule reactions such as hydrogen/deuterium exchange with deuterated solvents [55, 56], and the analysis of differential fragmentation trends in MS/MS [57]. Moreover, if the gas-phase structure has not changed too much compared to the solution phase structure it is issued from (see Chap. 3 for full discussion), or if the changes occur in a predictable manner, it ultimately becomes possible to use gas-phase probing methods to obtain information on the conformational ensembles present in solution. It is therefore also crucial to cross-check the MS-derived interpretations with independent measurements by traditional solution-phase biophysical methods.

Quaternary structures, where intermolecular non-covalent interactions come into play, are a little more straightforward to study by mass spectrometry, because the mass of the complex is the sum of the masses of the constituents. However, one must also understand and control all the different factors that could lead to undesired disruption of the complexes in the source or in the transfer region before the mass analyzer. All secondary and tertiary structure probing methods also remain important for further characterization of the structures of the complexes, beyond their stoichiometry.

In summary, mass spectrometry plays a role in the characterization of natural and synthetic nucleic acids, in structural biology, and in nucleic acids biophysics. Traditional thermodynamic, spectroscopic, and hydrodynamic techniques all aim at understanding the different conformational or binding states of nucleic acids and how they depend on environmental conditions. Here mass spectrometry has the potential to play a key role, by separating all components present in solution and by characterizing their respective structures. One of the most prominent contributions of mass spectrometry to traditional nucleic acids biophysics is found in the field of

nucleic acid-ligand interactions, where MS is used to determine unambiguously the stoichiometry of the complexes and their relative quantities (see Chap. 9).

1.4   Key Concepts

General Concepts
- Mass spectrometry is a powerful analysis method that operates in vacuum. It is therefore important to understand how biomolecules and their ions behave in vacuo to understand and design mass spectrometry applications.
- Molecules in the gas phase cannot exchange energy with their environment: a population of molecules in vacuo is a microcanonical ensemble.
- Energy can be gained or lost by collisions with a partner, by photon absorption/emission, or by chemical reaction with another biomolecule, ion, or electron.
- Understanding the intrinsic non-covalent interactions in molecules in the absence of solvent is an indirect way to understand the solvent (or solvophobic) effect on folding or binding.

Concepts Specific to Nucleic Acids
- Nucleic acid primary structure is the succession of bases in a sequence.
- Natural DNA contains adenines, thymines, guanines, and cytosines in its canonical form. It is the repository of genome information. DNA is subjected to mutation and damage, which can include modified bases.
- Natural RNA contains uracil instead of thymines, and some functional RNAs can naturally contain many other modified bases. The roles of RNAs as molecules of life are very diverse, and not all well understood yet.
- Nucleic acid secondary structure is the hydrogen bonding pattern between bases, and the tertiary structures are the multiple helices that result from these hydrogen bonding patterns.
- Like for proteins, nucleic acid quaternary structures are supramolecular assemblies between different molecules.
- Nucleic acids derivatives can now be synthesized chemically and have many applications in biotechnology, nanotechnology, and medicine.

References

1. Banoub JH, Newton RP, Esmans E, Ewing DF, Mackenzie G (2005) Recent developments in mass spectrometry for the characterization of nucleosides, nucleotides, oligonucleotides, and nucleic acids. Chem Rev 105:1869–1915
2. Carell T, Brandmayr C, Hienzsch A, Muller M, Pearson D, Reiter V, Thoma I, Thumbs P, Wagner M (2012) Structure and function of noncanonical nucleobases. Angew Chem Int Ed Engl 51:7110–7131
3. Dudley E, Bond L (2013) Mass spectrometry analysis of nucleosides and nucleotides. Mass Spectrom Rev. doi:10.1002/mas.21388

4. Grasby JA, Neidle S, Blackburn GM, Gait MJ, Loakes D, Williams DM, Egli M, Flavell M, Flavell A, Pyle AM et al (2006) In: Blackburn GM, Gait MJ, Loakes D, Williams DM (eds) Nucleic acids in chemistry and biology, 3rd edn. RSC Publishing, pp 13–76

5. Sponer J, Sponer JE, Mladek A, Jurecka P, Banas P, Otyepka M (2013) Nature and magnitude of aromatic base stacking in DNA and RNA: quantum chemistry, molecular mechanics, and experiment. Biopolymers 99:978–988

6. Turner DH (2013) Fundamental interactions in RNA: questions answered and remaining. Biopolymers 99:1097–1104

7. Duca M, Vekhoff P, Oussedik K, Halby L, Arimondo PB (2008) The triple helix: 50 years later, the outcome. Nucleic Acids Res 36:5123–5138

8. Davis JT (2004) G-quartets 40 years later: from 5′-GMP to molecular biology and supramolecular chemistry. Angew Chem Int Ed 43:668–698

9. Gueron M, Leroy JL (2000) The i-motif in nucleic acids. Curr Opin Struct Biol 10:326–331

10. Drew H, Wing R, Takano T, Broka C, Tanaka S, Itakura K, Dickerson RE (1981) Proc Natl Acad Sci USA:2179–2183

11. Radhakrishnan I, Patel DJ (1994) Solution structure of a pyrimidine-purine-pyrimidine DNA triplex containing T*AT, C+*GC and G*TA triples. Structure 2:17–32

12. Creze C, Rinaldi B, Haser R, Bouvet P, Gouet P (2007) Structure of a d(TGGGGT) quadruplex crystallized in the presence of Li+ ions. Acta Crystallogr 63:682–688

13. Wang Y, Patel DJ (1993) Solution structure of the human telomeric repeat d[AG 3 (T 2 AG 3) 3] G-tetraplex. Structure 1:263–282

14. Esmaili N, Leroy JL (2005) i-motif solution structure and dynamics of the d(AACCCC) and d (CCCCAA) tetrahymena telomeric repeats. Nucleic Acids Res 33:213–224

15. Butcher SE, Pyle AM (2011) The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. Acc Chem Res 44:1302–1311

16. Rostom AA, Fucini P, Benjamin DR, Juenemann R, Nierhaus KH, Hartl FU, Dobson CM, Robinson CV (2000) Detection and selective dissociation of intact ribosomes in a mass spectrometer. Proc Natl Acad Sci USA 97:5185–5190

17. Kuhn-Holsken E, Dybkov O, Sander B, Luhrmann R, Urlaub H (2007) Improved identification of enriched peptide RNA cross-links from ribonucleoprotein particles (RNPs) by mass spectrometry. Nucleic Acids Res 35:e95

18. Wells RD (2007) Non-B DNA conformations, mutagenesis and disease. Trends Biochem Sci 32:271–278

19. De S, Michor F (2011) DNA secondary structures and epigenetic determinants of cancer genome evolution. Nat Struct Mol Biol 18:950–955

20. Sarkies P, Reams C, Simpson LJ, Sale JE (2010) Epigenetic instability due to defective replication of structured DNA. Mol Cell 40:703–713

21. Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950–952

22. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233

23. Guil S, Esteller M (2012) Cis-acting noncoding RNAs: friends and foes. Nat Struct Mol Biol 19:1068–1075

24. Dethoff EA, Chugh J, Mustoe AM, Al-Hashimi HM (2012) Functional complexity and regulation through RNA dynamics. Nature 482:322–330

25. Beaucage SL, Caruthers MH (1981) Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Lett 22:1859–1862

26. Francois JC, Lacoste J, Lacroix L, Mergny JL (2000) Design of antisense and triplex-forming oligonucleotides. Methods Enzymol 313:74–95

27. Famulok M, Mayer G, Blind M (2000) Nucleic acid aptamers—from selection in vitro to applications in vivo. Acc Chem Res 33:591–599

28. Ellington AD, Szostak JW (1990) In viro selection of RNA molecules that bind specific ligands. Nature 346:818–822

29. Ng EW, Shima DT, Calias P, Cunningham ET Jr, Guyer DR, Adamis AP (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. Nat Rev Drug Discov 5:123–132
30. Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Norden B, Nielsen PE (1993) Pna hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. Nature 365:566–568
31. Wengel J, Koshkin A, Singh SK, Nielsen P, Meldgaard M, Rajwanshi VK, Kumar R, Skouv J, Nielsen CB, Jacobsen JP et al (1999) LNA (locked nucleic acid). Nucleotides Nucleosides 18:1365–1370
32. Sinkeldam RW, Greco NJ, Tor Y (2010) Fluorescent analogs of biomolecular building blocks: design, properties, and applications. Chem Rev 110:2579–2619
33. Rodrigo G, Landrain TE, Shen S, Jaramillo A (2013) A new frontier in synthetic biology: automated design of small RNA devices in bacteria. Trends Genet 29:529–536
34. Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew SY, McLaughlin SH et al (2012) Synthetic genetic polymers capable of heredity and evolution. Science 336:341–344
35. Rothemund PWK (2006) Folding DNA to create nanoscale shapes and patterns. Nature 440:297–302
36. Sacca B, Niemeyer CM (2012) DNA origami: the art of folding DNA. Angew Chem Int Ed Engl 51:58–66
37. Wilner OI, Willner I (2012) Functionalized DNA nanostructures. Chem Rev 112:2528–2556
38. Krishnan Y, Simmel FC (2011) Nucleic acid based molecular devices. Angew Chem Int Ed Engl 50:3124–3156
39. Seeman NC (2005) From genes to machines: DNA nanomechanical devices. Trends Biochem Sci 30:119–125
40. Tang Y, Ge B, Sen D, Yu HZ (2013) Functional DNA switches: rational design and electro-chemical signaling. Chem Soc Rev 43:518–529
41. Liu SP, Weisbrod SH, Tang Z, Marx A, Scheer E, Erbe A (2010) Direct measurement of electrical transport through G-quadruplex DNA with mechanically controllable break junction electrodes. Angew Chem Int Ed Engl 49:3313–3316
42. Genereux JC, Barton JK (2010) Mechanisms for DNA charge transport. Chem Rev 110:1642–1662
43. McLuckey SA, Goeringer DE (1997) Slow heating in tandem mass spectrometry. J Mass Spectrom 32:461–474
44. Dunbar RC, McMahon TB (1998) Activation of unimolecular reactions by ambient blackbody radiation. Science 279:194–197
45. Vékey K (1996) Internal energy effects in mass spectrometry. J Mass Spectrom 31:445–463
46. Armentrout PB, Ervin KM, Rodgers MT (2008) Statistical rate theory and kinetic energy-resolved ion chemistry: theory and applications. J Phys Chem 112:10071–10085
47. Dunbar RC (2004) BIRD (blackbody infrared radiative dissociation): evolution, principles, and applications. Mass Spectrom Rev 23:127–158
48. Rodgers MT, Armentrout PB (2000) Noncovalent interactions of nucleic acid bases (uracil, thymine, and adenine) with alkali metal ions. Threshold CID and theoretical studies. J Am Chem Soc 122:8548–8558
49. Strittmatter EF, Schnier PD, Klassen JS, Williams ER (1999) Dissociation energies of deoxy-ribose nucleotide dimer anions measured using BIRD. J Am Soc Mass Spectrom 10:1095–1104
50. Klassen JS, Schnier PD, Williams ER (1998) Blackbody infrared radiative dissociation of oligonucleotide anions. J Am Soc Mass Spectrom 9:1117–1124
51. Auffinger P, Hashem Y (2007) Nucleic acid solvation: from outside to insight. Curr Opin Struct Biol 17:325–333
52. Janin J (1997) Angströms and calories. Structure 5:473–479
53. Chaires JB (1997) Energetics of drug-DNA interactions. Biopolymers 44:201–215

54. Dearden DV, Liang Y, Nicoll JB, Kellersberger K (2001) Study of gas-phase molecular recognition using fourier-transform ion cyclotron resonance mass spectrometry. J Mass Spectrom 36:989–997
55. Balbeur D, Dehareng D, De Pauw E (2007) Conformationally driven gas-phase H/D exchange of dinuleotide negative ions. J Am Soc Mass Spectrom 18:1827–1834
56. Vairamani M, Gross ML (2003) G-quadruplex formation of thrombin aptamer detected by electrospray ionization mass spectrometry. J Am Chem Soc 125:42–43
57. Mo J, Hakansson K (2006) Characterization of nucleic acid higher order structure by high-resolution tandem mass spectrometry. Anal Bioanal Chem 386:675–681

Transferring Nucleic Acids to the Gas Phase

2

Gilles Grégoire

Abstract

During the past 30 years, tremendous efforts have been made to implement mass spectrometry and spectroscopy technologies for the characterization of biomolecules in the gas phase. Progresses in the study of gas phase oligonucleotides and DNA have come with the advent of different sources capable to transfer these fragile biomolecules from the condensed phase into the gas phase. These techniques have been largely employed in the spectroscopy and mass spectrometry communities and have stimulated much research with applications of mass spectrometry to structural biology and applications of spectroscopy to detailed understanding of the intra- and intermolecular interactions. The key point in all the experimental techniques is to counterpoise the extremely low volatility of nucleobases, oligonucleotides, and higher-order DNA structures while keeping these thermally fragile molecules intact. The aim of this chapter is to describe the transfer of nucleic acids to the gas phase, including the technical and experimental issues that have been successfully overcome over the last decades.

G. Grégoire (✉)

Laboratoire de Physique des Lasers, Université Paris 13, Sorbonne Paris Cité, 93430 Villetaneuse, France

CNRS, UMR 7538, LPL, 93430 Villetaneuse, France
e-mail: gilles.gregoire@univ-paris13.fr

Abbreviations/Formulae

6-ATT	6-Aza-2-thiothymine
CRM	Charge residue model
DAHC	Diammonium hydrogen citrate
DNA	Deoxyribonucleic acid
ESI	Electrospray ionization
FTICR	Fourier transform ion cyclotron resonance analyzer
FWHM	Full width at half maximum
HPA	3-Hydroxypicolinic acid
HPLC	High performance liquid chromatography
IEM	Ion evaporation model
LILBID	Laser-induced liquid beam(bead) ionization/desorption
MALD(I)	Matrix-assisted laser desorption (ionization)
MS	Mass spectrometry
NB	Nucleobases
NH_4OAc	Ammonium acetate
PA	Picolinic acid
PCA	Pyrazine carboxylic acid
*re*TOF	Reflectron time-of-flight
RNA	Ribonucleic acid

2.1   Introduction

The general objective of studying of biological systems in the gas phase is to gain insights into their behaviors and intrinsic properties free from their native environment. With a bottom-up approach, a first step consists in characterizing separate key components, either individual molecules or building blocks, in order to probe their structures and chemical properties. Increasing progressively the size of the molecular species allows one to investigate the intramolecular interactions that govern the emergence of secondary structures. Subsequently, the intermolecular interactions are explored by probing the non-covalent bonding either with solvent molecules (namely water), cations, or organic ligands in molecular clusters of increasing size and complexity.

Several experimental strategies have been devised to transfer fragile biomolecular species into the gas phase. Molecular beam technology has been mainly applied to the spectroscopic studies of neutral, DNA building blocks, whereas soft ionization methods like electrospray ionization and matrix-assisted laser desorption have been widely used in mass spectrometry to isolate large ionic molecular assemblies. Nowadays, the spectroscopy and mass spectrometry communities get significantly closer and gain strength from each other by sharing their experimental skills of transferring and studying biomolecules in the gas phase.

Among biomolecules, nucleic acids play an important role in a variety of fundamental biological processes, and have thus drawn a particular attention. Research in the field of physical chemistry has been mainly devoted to the study of bare, neutral oligonucleotide building blocks, i.e., the nucleobases (NB) adenine, guanine, cytosine, thymine, and uracil. Optical spectroscopy can provide detailed structural information but requires low internal temperatures in order to reduce the conformational heterogeneity and to get high resolution spectra. Such experimental conditions are commonly achieved by using a supersonic expansion of rare gas atoms seeded with the molecule of interest. Nevertheless, for a long time, those molecules could hardly be studied, in contrast to others organic molecules, for practical reasons. The nucleobases, and particularly guanine, have too low vapor pressures at ambient temperature to be seeded in a molecular beam and they decompose when heated. New experimental strategies developed to tackle this issue will be presented in the first part of this chapter, in particular the coupling of laser desorption with supersonic expansion [1–3] and the use of superfluid helium droplet to pick up single molecules [4].

During the same period and almost independently, the gas-phase studies of large biological molecules have emerged with the advent of soft and sensitive techniques such as matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). The hugely increased mass range accessible by these new ionization techniques has boosted the field of biomolecule mass spectrometry. These techniques have made such an important contribution that they have resulted in the award of the Nobel Prize in Chemistry in 2002 to Fenn [5] for ESI and Tanaka [6] for MALDI. Very soon after the emergence of ESI, oligonucleotides up to 100-mers [7, 8] and DNA oligomers with mass greater than one million Dalton [9] could be detected when coupled to a mass spectrometer, owing to the fact that ESI produces multiply charged species that keeps the m/z ratio in reasonable values. A particular interest of the ESI method is to minimize covalent fragmentation and even to preserve non-covalent complexes. Minimizing the internal energy of the ions produced by ESI sources is a key point for analysis of non-covalent complexes and structural studies [10]. This has promoted the study of gas phase oligonucleotide-drug species as a screening tool in pharmaceutical research in which the ionization state and the stoichiometry are perfectly known and controlled [11]. MALDI, although providing a relatively gentle method of generating large intact biomolecules in the gas phase, suffers from a larger tendency to fragment ions. Special attention has thus been paid to the search of the best matrix in order to avoid extensive fragmentation of oligonucleotides, although methodologies have been employed to get structural information from DNA ion fragmentation. While most of the MALDI experiments are performed from co-crystallized solid matrices, recent developments in laser desorption from liquids, directly under vacuum, have emerged and shown many advantages, in particular in the analysis of large nucleic acids. With all methods, standard oligonucleotide samples exhibit broader distributions of metal ion adducts (Na/K) than what would be obtained with peptides and proteins. Therefore, this chapter will also cover methodologies especially designed for the sample preparation of oligonucleotides.

2.2 Methods

2.2.1 Molecular Beam

To investigate fundamental properties of DNA building blocks, it is desirable to study individual nucleosides in the gas phase without interference from solvent molecules or macromolecular structure. Research in the past two decades has shown considerable progress in the study of biomolecular building blocks in the gas phase, by means of laser spectroscopy on cold, isolated molecular species produced in a molecular beam. The potential energy surface of nucleobases is rich and contains a large number of energy minima corresponding to different structures separated by low energy barriers. It is therefore mandatory to reduce the vibrational energy of the system in order to freeze the molecule in specific conformations. NB can adopt a great number of tautomer forms due to the mobility of their hydrogen atoms which can bind to the nitrogen and oxygen electronegative binding sites, leading to several isomers of the keto and enol forms (Fig. 2.1). For instance, the lowest-energy tautomers of guanine are the keto and enol forms which can each appear in the N7H and N9H form [12, 13]. Enol and keto tautomers can be easily distinguished by infrared spectroscopy due to the specific enol OH stretch around 3,600 cm^{-1}. The groups of de Vries [14–16], Mons [13, 17–19], Kleinermanns [20], and Kim [21, 22] have conducted exhaustive gas phase studies on isolated DNA bases. Nucleosides [23–25], bases pairing [26–30], and hydrates [27, 31–36] have also been investigated. Through IR spectroscopy, structural assignment is achieved by comparison with ab initio quantum chemistry calculations, revealing the different tautomer forms of the nucleobases populated in the supersonic expansion. By means of UV spectroscopy, the electronic spectrum of DNA bases can be recorded and the knowledge of their photophysical properties can be obtained. In any cases, well-resolved spectroscopy can be obtained as long as the molecules are frozen in their lowest vibrational energy levels, which implies efficient cooling of the DNA bases in a molecular beam.

2.2.2 Supersonic Expansion

Vapor pressure of the nucleobases (NB) is rather low at room temperature and thus requires thermal heating. In order to prevent rapid thermal degradation, supersonic expansion of rare gas atoms seeded by the molecules of interest has been widely used for decades.

Figure 2.2a depicts the structure of a jet formed by gas issuing from the exit of a simple converging nozzle into a region of lower pressure. The jet expands axially, accelerating to supersonic velocity as soon as the ratio between the initial backing pressure P_0 and the pressure in the vacuum chamber P_2 is greater than $[(\gamma + 1)/2]^{\gamma/(\gamma-1)}$, where γ is the ratio of heat capacity C_p and C_v of the carrier gas. The generated shockwave expels the residual gas and creates a silent zone where the velocity of carrier gas atoms is greater than the speed of sound (Mach number
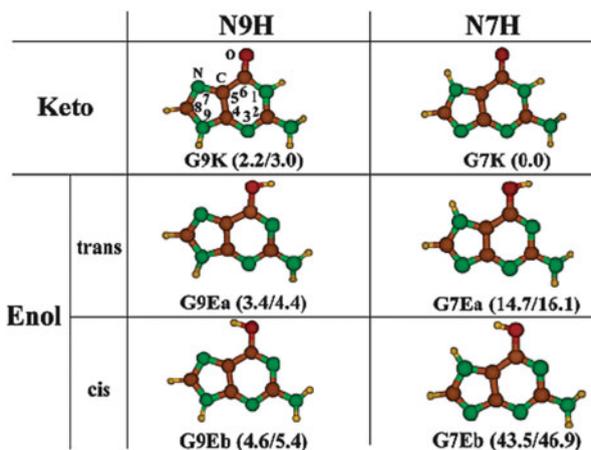
Fig. 2.1 Structures and energies (kJ/mol) of some tautomers of guanine. Each keto (C=O) and enol (O–H) structures can be found in N9H and N7H forms. For the enol structures, the OH group can be either in trans or cis orientation (reproduced with permission from [37]. © (2006) American Chemical Society)
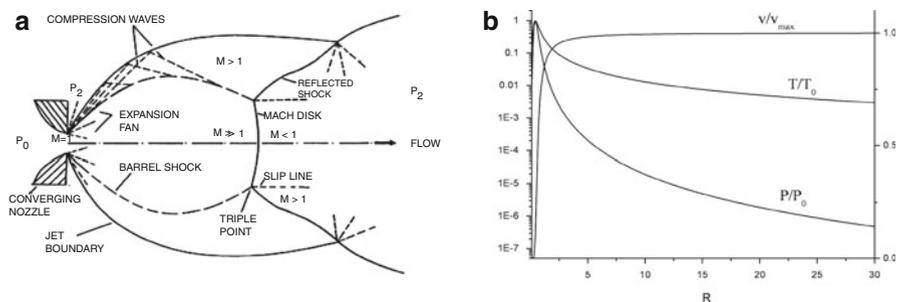
Fig. 2.2 (**a**) Schematic view of a supersonic expansion (reproduced with permission of [38]). (**b**) Evolution of the temperature, pressure, and velocity of the jet as a function of the $R = X_{\mathrm{m}}/d$. See text for details

M = flow velocity/local speed of sound $\gg 1$). The Mach disk defines the position where the molecular beam velocity is equal to the speed of sound ($M = 1$), from where collisions with residual gas atoms heats the seeded molecules of the jet. The Mach distance X_{m} is defined by the position of the Mach disk relative to the source orifice and can be estimated by:

$$X_{\mathrm{m}} = 0.67d\sqrt{\frac{P_0}{P_2}}, \text{where } d \text{ is the diameter of the nozzle.} \quad (2.1)$$

In practice, d is in the order of 100–1,000 μm, P_0 a few bars, and P_2 about $10^{-5}/10^{-7}$ mbar. Therefore, the distance X_{m} is greater than the distance to the mass

spectrometer, which insures that the molecules of interest will be studied in the silent zone, free from any interaction with the residual gas.

The temperature T, pressure P, and the density of the gas n depends on the ratio $R = X_m/d$ and decrease rapidly with the distance to the nozzle, while the hydrodynamic speed of the jet reach its maximum value very close to the exit of the nozzle (Fig. 2.2b).

$$\frac{T(R)}{T_0} = \left(\frac{P(R)}{P_0}\right)^{\frac{\gamma-1}{\gamma}} = \left(\frac{n(R)}{n_0}\right)^{\gamma-1} = \frac{1}{1 + \frac{\gamma-1}{2}M(R)^2} \ \text{with} \ M(R) \approx 3.3R^{\gamma-1} - R^{1-\gamma}. \quad (2.2)$$

The cooling efficiency of the supersonic expansion is essentially determined by the number of inelastic collisions between the rare gas atoms and the heated molecules. The number of collisions per second, N_{coll}, in the supersonic jet depends upon the reservoir density n_0, the nozzle diameter d, the collision cross-section σ with the carrier gas, the local Mach number $M(R)$, and the heat capacity ratio γ [39]:

$$N_{coll} = \sqrt{2}n_0\sigma\bar{v}\frac{1}{\left(1 + \frac{\gamma-1}{2}M(R)^2\right)^{\frac{\gamma+1}{2(\gamma-1)}}}. \quad (2.3)$$

The number of collisions is essentially independent of the nature of the carrier rare gas. However, the cooling rate of the seeded molecules is faster with heavy atoms (Ar, Kr, Xe) than with He or Ne. It is often recommended to work with Ar to produce molecular non-covalent complexes, such as hydrates, while He is sufficient for studying bare, isolated molecules. The main drawback of supersonic expansion is its inability to efficiently cool large molecules due to the rapid decrease of the collision rate along the expansion and the high internal energy in large molecules after vaporization. In practice, species with a molecular weight larger than 500–700 a.m.u. do not show well-resolved laser spectroscopy, which puts an upper limit to the study of large neutral biomolecules.

2.2.3 Thermal Vaporization of Neutral DNA Bases

Among NBs, uracil, thymine, adenine, and cytosine can be heated to obtain sufficiently high vapor pressure before reaching decomposition [40]. The sample in a powder form is heated in a metal oven up to 200 °C, and the beam is expanded through a high temperature pulsed nozzle with rare gas atoms as carrier gas (with or without water vapor) at a backing pressure of few bars. Neutral molecules and clusters streamed in the supersonic expansion are ionized downstream in a second vacuum chamber through resonant two photon ionization (R2PI) scheme and detected in a reflectron time-of-flight (*re*TOF) mass spectrometer. Interestingly, Kim and coworkers [31] have shown that the detection of hydrated adenine and
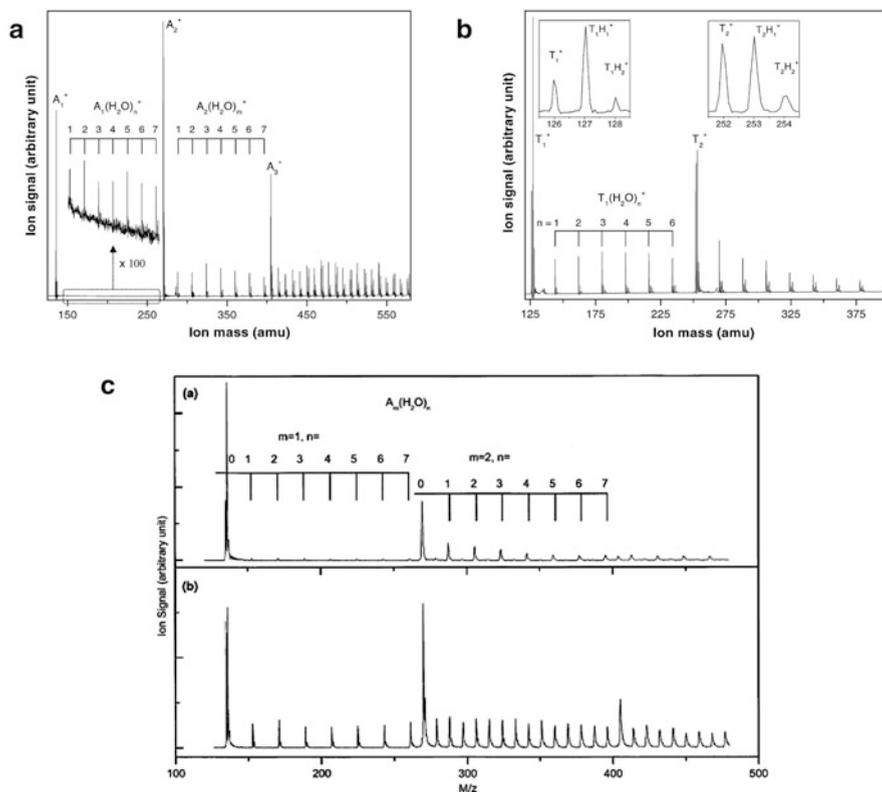
Fig. 2.3 Nanosecond R2PI mass spectrum of hydrated adenine (**a**) and thymine (**b**) (reproduced with permission of [31]). Note that the intensity of the adenine monomer hydrates is anomalously small relative to that of the non-hydrated peak and the thymine case. (**c**) Femtosecond R2PI mass spectrum of hydrated adenine cluster (reproduced with permission of [27])

thymine clusters is dramatically different (Fig. 2.3) when ionized with nanosecond laser. The cluster generation is expected to be essentially similar in both cases, so the large discrepancy in the detection efficiency was related to the excitation/ionization process. They have shown indeed that adenine water clusters were effectively produced and detected when using femtosecond laser pulses (Fig. 2.3) [27].

2.2.4 Laser Desorption of Neutral DNA Bases: The Case of Guanine

Guanine requires more elaborate methods than simple heating. Guanine is difficult to vaporize in the gas phase without significant pyrolysis and requires the use of softer vaporization methods like laser desorption. Pioneered by the group of de Vries [14], and followed by the groups of Mons [41] and Saigusa [42], the coupling

of matrix-assisted laser desorption (MALD) and supersonic expansion have been successfully performed and applied to the vaporization of guanine and guanosine [23].

Sample preparation consists in pressing a mixture of guanine and graphite powders to form a solid disk. Laser desorption is performed with the fundamental or second harmonic of a pulsed Nd:YAG laser light, just in front of the nozzle of a pulsed valve (Fig. 2.4). Typical laser fluences are of the order of 1–5 mJ/cm^2, significantly less than for laser ablation. The neutral molecules streamed in the supersonic expansion are cooled down to low rotational and vibrational temperatures, and this ensures efficient resonant ionization of the seeded molecules. Laser desorption can be considered as relatively mild, given that intact molecules and their complexes can be preserved in the gas phase. However, in some cases, fragmentation can occur, which may depend on the nature of the molecules of interest. Using the same experimental conditions, de Vries et al. have shown that guanosines do not fragment while adenosides do [43]. Recently, Ahmed and coworkers have recorded the Vacuum-Ultraviolet (VUV) photoionization efficiency of guanine using two different methods of vaporization, thermal vaporization, and laser desorption [44]. Interestingly, different tautomers seem to be populated with the two experimental processes. This clearly emphasizes that the source conditions must be carefully controlled and that comparison must be made very carefully.

2.2.5 Pick Up Source with Helium Droplets

Superfluid Helium droplets can be produced in a cryogenic nozzle expansion. These droplets provide a unique matrix for studying molecules at very low temperature of 0.37 K [45]. He droplet beams can be produced by the free expansion of gaseous He through a small nozzle of 5–10 μm diameter at stagnation source temperatures and pressures of 5–20 K and 10–50 bars, respectively [46]. The low stagnation temperature is achieved by attaching the source to a closed-cycle He cryostat. The expanding gas in the jet is cooled by the adiabatic change in state to ambient temperatures and pressures well below the critical point of He, where extensive condensation to small droplets occurs. Mean sizes of the helium droplets are estimated in between 10^4 and 10^5 atoms depending on the initial temperature and backing pressures [4].

The droplets pass downstream a skimmer to enter a pick up cell filled with the species to be embedded (Fig. 2.5). Superfluid helium droplets are known to efficiently pick up molecules due to the large collision cross section of roughly $15 N^{2/3} \text{ Å}^2$, where N is the average number of helium atoms forming the droplet. Therefore, a vapor pressure of only 10^{-5}–10^{-6} mbar is sufficient for picking up a single particle on average, which is particularly well suited for species with low volatility such as DNA bases. Upon attachment to the droplets, the initial kinetic and internal energies of the embedded molecules drop off, the evaporation of each
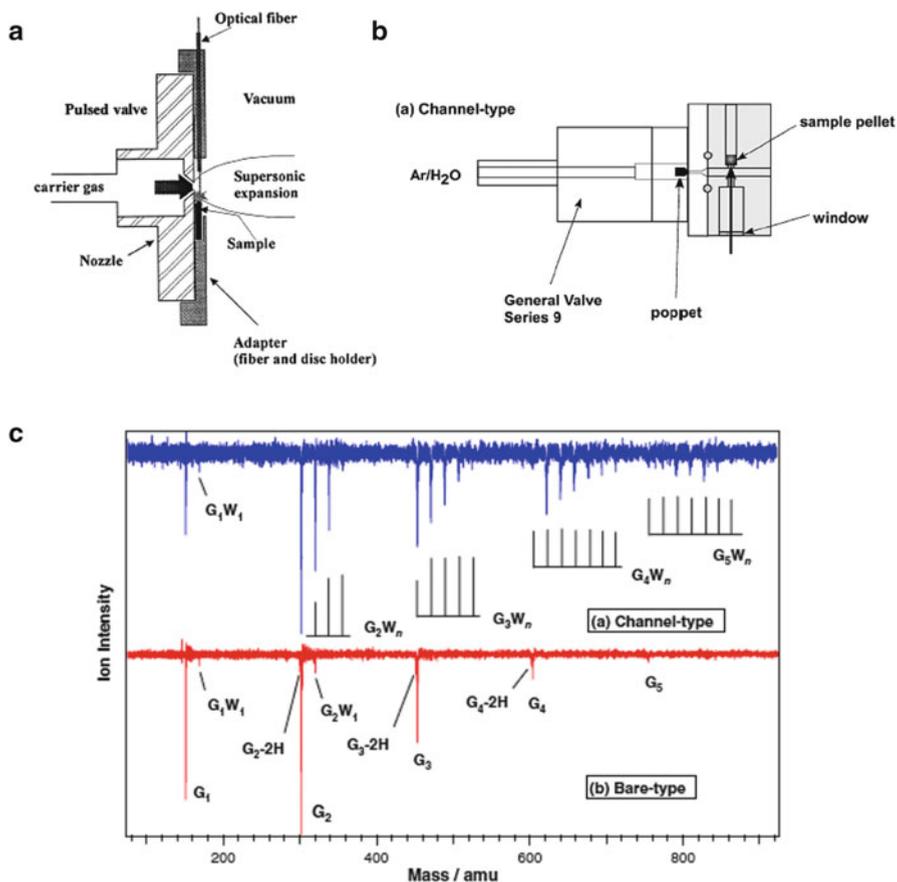
Fig. 2.4 (**a**) Schematic representation of the de Vries and Mons laser desorption source in front of the nozzle of a pulsed valve (reproduced with permission of [41]). (**b**, **c**) The Saigusa source with a channel-type exit enhancing the production of clusters and hydrates (reproduced with permission of [42])
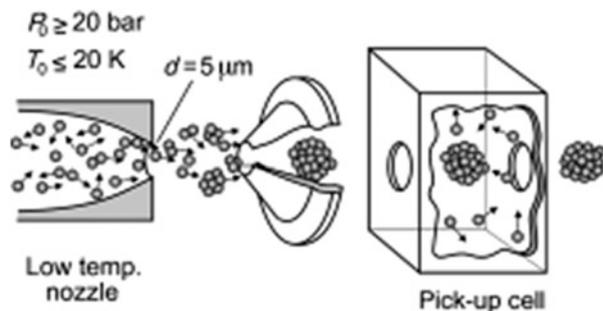
Fig. 2.5 Jet expansion of helium droplets with a pick-up cell downstream. The nozzle is kept at cryogenic temperature with high backing pressure in order to produce superfluid helium droplets (with permission of [45])

helium atom taking off 5 cm^{-1} [4]. Thousands of He atom evaporations are thus required for freezing large molecules.

Several groups have used such experimental setup to study nucleobases and their complexes. Miller et al. [37] have shown that the most stable tautomers of guanine were indeed observed, in contrast to the previous conclusions drawn from the experiment performed using supersonic expansion. Scheier and coworkers [47] have investigated the formation of negative ions of adenine and thymine embedded in He clusters, showing that stable transient anions are stabilized in the helium droplets while being unstable when isolated in the gas phase and that efficient damage of the nucleobases are observed via resonant, low energy electron attachment.

2.2.6 Electrospray Ionization

2.2.6.1 General Mechanism

The principal outcome of the electrospray process is the transfer of analyte species into the gas phase as isolated entities. A solution containing the analyte is introduced in a capillary with a syringe. The application of a high electric field to the capillary accumulates ions attracted by a counter electrode at the surface of the capillary tip. The effect of the high electric field as the solution emerges is to generate a spray of highly charged droplets which travel down the potential and pressure gradients towards the mass spectrometer. During that transition, the droplets shrink by solvent evaporation or by "Coulomb explosion" (droplet fission resulting from the high charge density). The fission process may result in product droplets of roughly equal size, or more often, it may spawn the ejection of a series of much smaller droplets [48]. Ultimately, fully desolvated ions result from complete evaporation of the solvent or by field desorption from the charged droplets [49] (Fig. 2.6).

The electrospray process can be divided into three stages: droplet formation, droplet size reduction, and ion desolvation. Using a negative potential to the tip (as recommended for the study of DNA species), negative ions in solution move towards the counter electrode and will accumulate at the surface, and this establishes a Taylor cone. When the external electric field is high enough (roughly a few keV/cm), the cone is drawn into a filament that produces negatively charged droplets when the applied electrostatic force exceeds the surface tension. The diameter of the droplets formed is influenced by a number of parameters, including the applied potential, the solution flow rate, and solvent properties. Generally, electrospray proceeds at flow rates of 1–100 μL/min. At such flow rates, nebulization of the solution is facilitated by a sheath flow of nebulizer gas and by controlled heating of the interface. Nanoelectrospray is a variation of electrospray in which the spray tip has a much smaller diameter, and the flow rate is drastically reduced to a few nL/min, resulting in a more efficient production of gas phase ions [51, 52]. Charged droplets are only produced by the electric field with no need of an external gas flow [53].
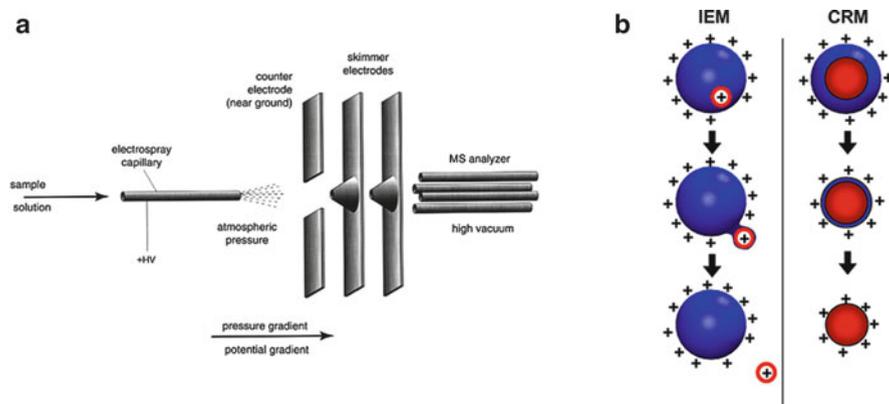
Fig. 2.6 (**a**) Schematic representation of the ESI interfaced with a mass spectrometer (with permission of [49]) (**b**) charged residue (CRM) and ion evaporation (IEM) models introduced to explain the desolvation of the ions (adapted with permission from [50]. © (2013) American Chemical Society)

Two main mechanisms have been proposed for the desolvation of the observed ions. One is the charged residue mechanism (CRM), in which the repulsive effect of the increasing surface charge density completely offsets the droplet surface tension so that the droplet divides, temporarily increasing the overall surface-to-volume ratio (the "Rayleigh limit"). Dole proposed that the fission process continues until the remaining droplets contain only a single charged species [54]. Large molecular ions and folded proteins are thought to be produced by this mechanism [55]. The second proposed mechanism is the ion evaporation mechanism (IEM) [56, 57] in which the increase in surface charge density as a result of solvent evaporation produces a Coulomb repulsion that exceeds the charged species adhesion to the droplet's surface and thus some ions are expelled from the surface [58, 59]. This process also continues as the droplet decreases in surface area due to solvent evaporation. It is believed that atomic ions, small inorganic ions, and denatured proteins are produced by this mechanism [57, 60, 61]. One must say that there is still a debate on the mechanisms by which ions are produced in the gas phase, which has been recently reviewed by Konermann et al. [50, 62].

2.2.6.2 ESI Sample Preparation for DNA
Adduct Removal

The phosphodiester backbone of oligonucleotides in solution is negatively charged ($pK_a \sim 1$) and as a consequence has a high affinity for free cations that are present in solution. Upon ionization, these cations (e.g., Na^+, Mg^{2+}) produce adduct species with polynucleotides and, when present in excess, result in reduced ionization efficiency. Standard oligonucleotide samples typically produce a distribution of molecular ions, varying in the number and type of metal counterions (sodium and/or potassium adducts). Besides decreased signal to noise with an increasing

number of nucleotides in the sample molecules (dispersion of the produced ions among an increasing number of species of different m/z values), such signal distributions require a very high instrumental resolution power if larger species are to be mass analyzed. In addition, mixture analysis becomes difficult as soon as the signal distributions of samples differing by a single nucleotide overlap.

Experimentally, several precautions must be heeded. It is important to work with ultrapure grade solvents that are stored in containers that do not produce salt contamination; in particular, glass containers should be avoided. Capillaries used to transfer the sample to the mass spectrometer are a source of contamination and should be kept clean.

The quality of ESI mass spectra of oligonucleotides can be significantly improved if the metal cations present in the analyte solution are replaced by ammonium ions [8]. Ammonium acetate precipitation is used for desalting and cation exchange. In negative ion mode electrospray, the droplets carry excess negative charges consisting of DNA polyanions and acetate anions. After complete solvent evaporation, further activation of the DNA with its ammonium counterions results in proton transfer reactions from NH_4^+ to PO^-, hence neutralization of phosphates by protons and loss of NH_3. When using about 100 mM ammonium acetate, only a small fraction of phosphates remain negatively charged (on average, 1 deprotonated phosphate group every 5–6 nucleotides). Addition of chelating agents, for instance nitrilotriacetic acid (NTA) and trans-1,2-diamino cydohexane-N,N,N',N'-tetra acetic acid (CDTA), to solutions of nucleic acids results in both an improvement in the signal-to-noise ratio of the mass spectrum as well as reduction of mass spectral peak widths by removal of excess adducted cations [63]. Potier et al. [64] demonstrated that sodium attachment to molecular ions of nucleic acids can be dramatically reduced through the addition of high concentrations of ammonia or organic bases such as trimethylamine (TMA) and triethylamine (TEA), the latter being the most efficient. Organic co-solvents, such as methanol or acetonitrile, may be added to the solution to obtain maximum sensitivity and signal stability. However, the fraction of co-solvent must be kept as low as possible (less than 10–20 % in volume) to prevent any structural rearrangement in solution for analyses aiming at preserving a memory of the native structure in the gas phase.

HPLC-ESI

The on-line sample preparation of nucleic acids by chromatographic separation prior to ESI-MS is very attractive because it not only removes cations from nucleic acid samples but also fractionates nucleic acids in mixtures that are too complex for direct-infusion ESI-MS. Chromatography methods for DNA and RNA have been recently reviewed [65]. HPLC has been successfully applied for the desalting and separation of single-stranded oligodeoxynucleotides. Nevertheless, the applicability of HPLC to efficiently remove cation adducts in large nucleic acid samples is still problematic. Anion-exchange HPLC and ion-pair reversed-phase HPLC (IP-RP-HPLC) are the most popular chromatographic modes for the separation of nucleic acids [66, 67]. Double-stranded DNA fragments ranging in size from 50 to

600 base pairs could be separated and detected by IP-RP-HPLC-ESI-MS [68]. However, accurate mass determination is hardly feasible for DNA strands up to 250-mer, i.e., with a molecular mass of about 160 kDa. Whereas the salt gradients applied to elute the nucleic acids from anion-exchange columns are incompatible with ESI-MS, ion-pair reversed-phase chromatographic phase systems that employ volatile eluent components can be directly coupled to ESI-MS. Huber et al. [69] have shown that systematic variation of the eluent composition, such as concentration of ion-pair reagent, anion in the ion-pair reagent, solution pH, and acetonitrile concentration led to the conclusion that most parameters have opposite effects on chromatographic and mass spectrometric performances. The use of acetonitrile as sheath liquid enabled the rapid and highly efficient separation and detection of phosphorylated and nonphosphorylated oligonucleotides ranging in size from 8 to 40 nucleotides. Nowadays, (nano)-HPLC-ESI method reaches great sensitivity for quantitative analysis of ligand DNA adducts [70–72].

2.2.6.3 ESI-MS for the Study of DNA Non-covalent Complexes
Generalities

ESI typically generates protonated and/or deprotonated molecular ions of the type $(M + nH)^{n+}$ and $(M - nH)^{n-}$, respectively. High charge states are often formed, and the charge state distribution depends on the molecular mass of the analyte, the solvent conditions in the sprayed solution, and ion-source conditions, e.g., skimmer voltage (collisional-induced desolvation and dissociation). Thermal heating, pressure, and acceleration voltage in the nozzle-skimmer region of the ESI are optimized to generate a stable spray and get a complete ion desolvation in order to enhance the ion signal intensities. These parameters also affect the amount of internal energy of the ions [73]. For oligonucleotides, the removal of the last counterions may require harsh ESI condition, increasing the fragmentation. The determination of the internal energy of the ions produced in ESI and the effect on the in-source fragmentation has been recently reviewed by Gabelica and De Pauw [10].

In ESI-MS of oligonucleotides, negatively charged ions yield higher intensity signals than positively charged ions. Many commercial mass spectrometers are equipped with time-of-flight (TOF) or radio frequency-based trapping devices that provide high enough mass resolution. Fourier transform ion cyclotron resonance (FTICR) mass spectrometry represents a unique platform with which very high mass measurement accuracy is achieved. Mass resolution of at least 400,000 (FWHM) and sub-ppm mass measurement error are necessary to resolve U from C in RNA for instance.

The full desolvation (solvent and counterion) of electrosprayed ions is typically completed in less than 1 ms, and the resulting mass spectra are also strongly influenced by the extent of collisional activation and dissociation in the ESI-MS interface. This latter point is particularly important for the removal of the last counterions, which can be preserved for much longer times (seconds) and detected in the mass spectrometer. The loss of non-covalent associations during ESI may be kinetically constrained, allowing non-covalent complexes to remain associated as

residual solvent is removed. This depends on the location and number of charges remaining on each of the associated species (electrostatic interaction), on the effect of the solvent nature in the non-covalent association, on the relative strength and number of attractive interactions between complex constituents compared to solute–solvent interactions, and on the extent of complex activation in the interface. It has been suggested that the absence of water from complexes studied by ESI-MS favors the stability of complexes in which electrostatic interactions play a large role in stabilizing the complex. Interactions that are primarily hydrophobic in nature are more likely to dissociate in the ESI process due to the absence of water. The preservation of nucleic acid higher-order structure from the solution to the gas phase will be discussed in more detail in Chap. 3.

Distinguishing between structurally specific non-covalent interactions arising from solution and nonspecific aggregation in solution (and that might also result from the electrospray process) is feasible with careful selection of experimental conditions. For example, in the case of DNA–ligand complexes, ESI-MS studies should be conducted under the gentlest conditions possible if the relative intensities must correctly reflect the relative abundances of the complexes in solution. Generally, low analyte concentrations (not exceeding 20–50 µM) are mandatory to avoid nonspecific and random aggregation present at higher solution concentrations. Depending on the nature of the intermolecular forces, complexes with weak interactions (as van der Waals, hydrophobic interactions) are more readily dissociated as compared to those having strong electrostatic interactions. It is commonly accepted that "false positives" can be easily avoided by applying proper experimental conditions (in particular with low analyte concentrations). However, "false negatives" has to be carefully checked and viewed with caution since complex stability may be greatly affected in the gas phase and/or if electrospray conditions or instrument tuning in the ion transfer region are too harsh.

The typical solvent conditions used in ESI-MS to achieve maximum sensitivity are not always optimal for maintaining an intact biomolecular complex. Many biomolecular non-covalent interactions are highly dependent on many conditions, including the pH and ionic strength of the solution. Unlike most solution phase spectroscopic techniques, ESI-MS is intolerant to the presence of high salt and nonvolatile buffer concentrations. Therefore, a large majority of the biomolecular non-covalent complexes studied by ESI-MS are done in volatile salts like ammonium acetate (NH_4OAc). Organic solvents such as methanol or acetonitrile may be added to the solution to obtain maximum sensitivity and signal stability, but only with low percentage in order to preserve the native structure. Efforts have been made to obtain ESI-MS data under more physiological conditions. For instance, nanoelectrospray (nanospray) offers great opportunity in that sense since it can work with aqueous solutions at physiological conditions. Besides, the low flow rates (~5 nL/min) obtained with nanospray are also advantageous since the initial droplet size is smaller (<200 nm), desolvation is faster, and less heating is required for desolvation.

Oligonucleotide solutions are commonly heated to 70–90 °C for few minutes and cooled at room temperature for half a day in order to promote duplex or higher

order oligomer formations. Double helical duplex [74, 75] and quadruplex [76] DNA has been first detected in 1993, readily followed by the first observation of DNA–drug non-covalent complexes by Gale and coworkers 1 year later [77]. ESI-MS of non-covalent complexes of small molecules (intercalators, minor groove binders) with ds DNA are obtained by dilution of the complexed DNA just prior to analysis. In any cases, observation of non-covalent complexes by ESI-MS requires relatively gentle ionization/interface conditions, i.e., low capillary voltage, lower cone/orifice voltages, and lower source and/or probe temperatures [78].

DNA Complexes with Metal Ions

Clustering of nucleosides triggered by alkali metals has been studied to reveal magic number clusters, in particular quartets [79–81]. Metal–DNA and metal–NB ion complexes have been investigated for instance in the group of Rodgers [82–84] and Salpin [85–88] in order to get insights into the structure, coordination, and binding energies of these complexes. Several factors must be taken into account when using ESI-MS to study metal–oligonucleotide complexes. First, sensitivity may be reduced because of difficulties in removing excess metal salts and because of the distribution of available ion current among species with different metal isotopes. High resolution may be required for accurate mass measurements, because many metals of interest have complex isotope patterns. The resolution requirements become more stringent while analyzing metal–oligonucleotide complexes because the ions normally have much higher charges compared to the metal complexes alone. Finally, metal complexes may be more labile compared to alkylators, making detection of intact complexes difficult and/or resulting in less informative MS/MS spectra.

DNA Complexes with Small Molecules

Double-stranded (ds) DNA offers a number of potential binding sites and modes of non-covalent interactions for small molecules, namely (a) electrostatic interactions between cationic ligands and the polyanionic backbone of DNA, (b) intercalation of a planar ring system between adjacent base pairs, and (c) hydrogen bonding interactions with the bases usually in the minor groove of DNA. There are a number of examples of drugs or peptides that bind to DNA using combinations of these binding modes [78, 89, 90]. Detailed examples of application of ESI-MS to study nucleic acid–ligand interactions can be found in Chap. 9.

Intercalators typically contain a planar ring system and incorporate a positive charge. These compounds include simple intercalators such as ethidium bromide and daunomycin, threading intercalators such as nogalamycin, and bis-intercalators such as ditercalinium [91, 92]. Compounds (distamycin A, DAPI, beneril, Hoechst 33258, or netropsin) that bind in the minor groove normally incorporate small aromatic ring systems with torsional freedom to allow a twist complementary to that of the minor groove [93, 94]. AT-rich regions of the minor groove also have a greater electrostatic potential, thereby strengthening interactions with these ligands, which also normally contain a positively charged group at one or both ends. Minor groove-binding ligands typically interact with 4–6 bases, whereas intercalators can

bind between every alternate base pair. As a consequence of this requirement, relatively short oligonucleotides (6–8-mers with molecular masses ranging from ~1,800 to 4,500 Da) must be used to observe complexes with low binding stoichiometries. On the other hand, for DNA duplexes, a sufficiently large number of base pairs must be present so that the duplex is sufficiently stable both in solution and in the gas phase. Therefore, typical duplex length for ligand screenings is in the range of 10–16 base pairs.

The groups of Gross [95, 96], Sheil [89, 91], and De Pauw [92–94] have investigated the binding of known minor-groove binders and intercalators to duplex DNA. For instance, de Pauw and coworkers [94] studied the binding of five minor-groove binders to five different duplexed DNAs that contained 12 base pairs with varying AT compositions. Four out of the five minor-groove binders (Hoechst 33258, Hoechst 33342, DAPI and berenil) formed 1:2 DNA–drug complexes in addition to 1:1 complexes. Under the same conditions, Netropsin was observed to form exclusively 1:1 complexes with DNA, indicating that the 1:2 complexes observed with the other drugs are specific and not due to nonspecific aggregation. These investigations also showed that mass spectrometry can be used to assess the preference of drugs for binding to sequences with subtle difference, such as ATAT compared with AATT sequences. Greig and Robinson [97] developed the high-throughput screen DOLCE-MS (detection of oligonucleotide–ligand complexes by electrospray ionization mass spectrometry). In this approach multiple ligands can be screened simultaneously against single-stranded and duplexed oligonucleotides at a rate of 5 min per sample, allowing ~200 samples to be screened in a single day.

DNA and RNA–Protein Complexes

A limitation of ESI-MS lies in its low compatibility with the traditional buffers used to study biomolecular complexes in solution. Hence, in designing experiments, a compromise has to be made between the best ESI-MS solvents and the best solvents for maintenance of the biochemical complex. Volatile salts such as ammonium acetate or ammonium bicarbonate (often 10 mM up to 150 mM) are commonly present in solutions used to acquire ESI mass spectra of oligonucleotides and proteins. The question of salt is also important from the biochemical standpoint. In vivo, DNA binds to proteins under particular conditions of ionic strength, pH, and other dissolved solutes. Changes in these conditions may affect properties of macromolecules such as conformational states of the proteins. The ionic strength experienced by the complex may be of particular significance when electrostatic interactions are important for stability. In addition, some proteins are preferentially stabilized by one counterion over another under identical conditions of ionic strength.

Recent progresses have been made that allow ESI-MS methods to be used for the study of protein–DNA or protein–RNA complexes under relatively mild desolvation conditions with the use of aqueous buffers at physiological pH closed to the native solution conditions [98, 99]. Native mass spectrometry has recently emerged as a new paradigm in ESI-MS [100, 101]. For instance, the group of Robinson has shown that intact ribosome can be transferred into the gas phase and mass-analyzed to reveal its structure [102–105].
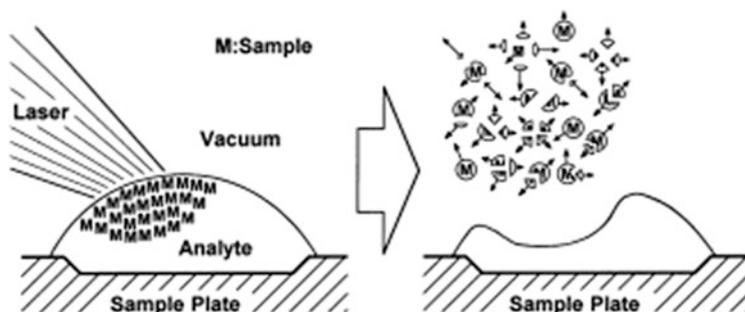
Fig. 2.7 Schematic representation of the MALDI process (with permission of [6])

2.2.7 Matrix-Assisted Laser Desorption Ionization

2.2.7.1 General Mechanism

In MALDI, analytes are embedded in an excess of matrix consisting of small organic molecules or fine metal powder [6, 106] absorbing the beam of UV lasers (to a lesser extent, IR lasers as well). Following absorption of a laser beam by the matrix, a plume is produced that ejects the analytes above the surface in the vacuum (Fig. 2.7). The matrix plays several roles. The first one is isolation to prevent formation of analyte aggregates. In the matrix, analytes are incorporated into the charge state already existing in the solution, in the presence of their counterions to ensure neutrality. Laser energy is absorbed by the UV chromophore of the matrix on a timescale of a few nanoseconds, much shorter than the time required for energy redistribution by thermal diffusion. Depending on the laser fluence, two physical regimes can occur. At low laser fluence, neutral and ionic species are individually transferred into gas phase in a desorption process. When the laser fluence increases, the matrix layer explodes in an ablation regime and ejects molecular matrix clusters as well as individual species [107, 108]. The MALDI mechanism involves a large number of processes occurring on three different timescales. The initial UV excitation ionizes the matrix molecules and ejects free electrons into vacuum. Internal energy conversion induces a fast heating of the surface leading to the formation of a plume, normal to the surface. Chemical reactions between ions and neutrals are thus likely to occur in the expanded plume within a few microseconds as long as there are collisions. The ion-to-neutral ratio in MALDI is rather low, in the order of 10^{-3}–10^{-8} [109]. The temperature in the plume is estimated below 1,000 K [110]. When studying neutral species (MALD), a pellet of carbon is generally used as matrix. While MALDI is frequently performed with acidic matrixes to detect cationic species (peptides and proteins), basic matrixes are used to efficiently abstract protons from the neutral analytes that will be detected as deprotonated species. In the case of oligonucleotides, the analytes are deprotonated in solution ($pK_a \sim 1$) and then form salts as the sample crystallizes. The samples are generally prepared using the "dried-droplet" method; less than 1 mL of the mixed solution of oligomer and matrix solution at equivalent volume ratio are deposited on the stainless steel target and dried in air at room temperature.

2.2.7.2 MALDI-MS for DNA

In the first years after the introduction of MALDI-MS, the generation of intact oligonucleotides longer than 8–10 units caused more difficulties with MALDI than with ESI. Spengler et al. [111] published in 1990 the first mass spectrum of a DNA octamer. However, a considerably lower quality of oligonucleotide mass spectra compared to peptide and protein spectra was commonly observed. Extensive ion fragmentation, far more pronounced in MALDI than in ESI, in addition to the already described problem of adduction of metal cations, was mainly responsible for this limitation. Nowadays, MALDI-MS platforms are widely used for instance in genotyping application [112, 113].

Analyses of non-covalent DNA and RNA complexes often require special buffer conditions and various quantities of salt, which are not always ESI compatible. The high salt tolerance and the matrix assistance for ionization give MALDI-MS some advantage over ESI, in particular with respect to sample preparation [99, 114]. Moreover, because MALDI produces predominantly singly charged ions, spectra are simplified and Coulomb repulsion which affects the stability of multiply charged complexes is reduced [115]. In most cases, instrumentation as well as sample preparation and choice of the matrix has to be very carefully tuned for detection of macromolecular complexes, because MALDI is known to be less soft than ESI [116, 117].

Choice of Matrix

Acceptable mass resolution and sensitivity for DNA oligomers of increased size has been a constant challenge for MALDI-MS. Routine detection of DNA oligomers larger than 200 bases has usually been problematic because adduct formation and fragmentation are still major limitations, and hence, development of new matrixes and MALDI experimental strategies has received much attention in the late 1990s. The 3-hydroxypicolinic acid (HPA) matrix, introduced in 1993, showed significant improvement when compared directly with a set of 47 other compounds which included most matrices used previously [118]. HPA was found to be the best matrix in terms of mass range, signal-to-noise ratio, and ability to analyze mixed-base oligomers below 100 nucleotides. Little et al. has further shown a 50-mer DNA duplex could be even detected with HPA when the samples were prepared at low temperature [119]. Other matrices such as 6-aza-2-thiothymine (6-ATT) and picolinic acid (PA) show very good performance, as evidenced by the detection of oligonucleotides up to 500 bases [120]. 4-nitrophenol was shown to be very effective for the analysis of large DNA oligomers when a cooled sample stage was used to prevent the sublimation of this matrix under vacuum [121]. UV laser desorption from this matrix allowed routine detection of DNA oligomers containing up to about 800 nucleotides at the picomole level. The effectiveness of this matrix was further demonstrated by the observation of a double-stranded DNA oligomer larger than 1,000 base pairs, seen as a denatured single-stranded species, with a molecular ion mass exceeding 300 kDa.

Mixed matrixes have shown to yield perhaps even better results, for instance 3-hydroxypicolinic acid (HPA) and pyrazinecarboxylic acid (PCA) [122]. With

respect to reproducibility, resolution, signal-to-noise ratio, and tolerance to metal salts in DNA analysis, this mixed HPA/PCA matrix is superior to HPA alone. The mixed matrix solution of 3-HPA/PCA are prepared by mixing saturated solutions of 3-HPA and PCA, in 50 % 0.5 mol/L DAHC plus 50 % acetonitrile, at volume ratios from 4:1 to 1:4. MALDI-TOF mass spectra with isotopic resolution for DNA segments up to 10.5 kDa in mass have been obtained. Equally impressive is the detection of separate ion peaks for both components of a mixture of two 23-mer DNA segments with a 7 Da difference.

Matrix Additives

Different techniques have been proposed for the suppression of metal-cation adduction, based on the observation that ammonium salts of oligonucleotides yield exclusively molecular ions of the free acids in ESI-MS. Several metal cations, especially Na^+ and K^+, are ubiquitous. Thus, their possible contact with the analyte within the last step of the sample preparation, e.g., from the pipette tips or the sample support surface, must be taken into account. Different techniques have been proposed for the suppression of metal-cation adduction, mainly based on the use of ammonium salts [123, 124]. Ammonium acetate was first used to displace alkali metal ions with NH_4^+ in the MALDI experiment by Currie et al. [125]. Since then, many applications using ammonium salts have appeared with the most successful being the use of diammonium hydrogen citrate (DAHC) by Pieles et al. [126]. Ammonium salts are very useful additives since complexed ammonium ions lose ammonia upon desorption/ionization, leaving protons behind. Li et al. [127] found that NH_4F and $(NH_4)HF_2$ show very good performance in combination with HPA matrices. It was found that, besides the elimination of multiple alkali-ion adduction onto the molecule-ions, the ammonium salt seems to play a significant role in enhancing desorption/ionization of intact oligonucleotide molecules. Asara et al. [128] have shown that spermine, used as a comatrix, results in improved spectra for single-stranded oligonucleotides and that it eliminates the need for a desalting step when such analytes are under study. This latter point may be explained by the fact that spermine can form non-covalent complexes with DNA in solution. Those complexes have indeed been detected both with ESI [129] and MALDI [130] sources.

The use of H^+ ion-exchange resin in situ permitted removal of alkali-metal ions from the oligonucleotide phosphate backbone, and this gave a gain in sensitivity of 1–2 orders of magnitude compared with previous MALDI methods [131]. In other cases, when DNA was not significantly contaminated, NH_4^+ organic salts as matrix additives were used to replace alkali metal ions in the polyphosphate backbone chain. In-line microdialysis has also been used for rapid desalting prior to MALDI-MS [132]. Other approach to sample preparation is based on the extraction of DNA out of solution onto a solid surface with an attached DNA-binding polymer, such as polyethyleneimine or polyvinylpyrrolidone [133]. Binding is strong enough to sustain washing, and thus, desalting and concentration can be performed in a single fast step. After removal of the supernatant solution, the addition of the MALDI matrix material releases the DNA from the surface to

allow the required co-crystallization on the support. The mass spectrometric analysis is then performed directly from this support. A similar, fast, efficient method for desalting for oligonucleotides is the use of solid-phase extraction plates packed with a reverse-phase sorbent that retains the biopolymer analytes while non-retained inorganic ions are washed out with deionized water [134].

IR-MALDI and Liquid Matrices

The use of a liquid matrix such as glycerol is known to give better reproducibility, but since glycerol does not absorb in the UV region, a second component must be included. For example, the binary mix of 4-nitroaniline and glycerol is effective for positive and negative-ion formation for oligonucleotides and other biomolecules [135]. Alternatively, desorption and ionization can be achieved by an IR laser, and Hillenkamp and coworkers [136–138] have shown that IR-MALDI with a glycerol matrix gave excellent results for large double-stranded DNA. Very little fragmentation was observed with an IR laser in the high-mass range compared to UV-MALDI, thus extending the mass limit for IR-MALDI to ca. 500 kDa with a sensitivity in the subpicomole range and a comparable reproducibility.

2.2.7.3 MS and Fragmentation Processes During the MALDI Step

MALDI is a pulsed ion source that is particularly well suited to TOF mass analyzer, but can also be associated with ion trapping devices (FTICR, linear or 3D radio frequency traps, Orbitrap) with longer measurement times than with TOF. For MALDI-MS of nucleic acids, ion fragmentation, and in particular metastable decay, is perhaps the most acute problem. With trap-type mass spectrometers, the measurement time is fairly extended compared to TOF. Typically, ion lifetimes in the millisecond range are needed for detection in FTICR (ion detection after ten to several hundred milliseconds). The time scale of in-source-induced fragmentation monitored by FTICR is significantly longer compared to TOF, where a measurement is typically limited to a few hundred microseconds.

The fragmentation of nucleic acids ions in MALDI-TOF can be classified into four categories; i.e., prompt, fast, early metastable, and metastable fragmentation. Fragment ions are usually called prompt, if the corresponding decay time constant is short or at best comparable to the generation time of the precursor ions (in the sub-microsecond range). Such fragment ions will have flight times corresponding to their authentic m/z value. Fragmentation reactions on a time scale slightly longer than the ion generation time but still short compared to the acceleration time of the precursor ions (fast fragments) will manifest themselves in a linear TOF as tails of the fragment peaks towards higher masses because of their energy deficit. If not too large, this energy deficit will be compensated for by reflectrons, resulting again in a symmetrical peak shape. The performance of MALDI-TOF can also be considerably improved by combination with delayed ion extraction methods employing high accelerating voltages, axial ion extraction, and short (µs) extraction delays [139, 140].

Early metastable fragmentation describes fragmentation reactions with a decay time constant of the same order of magnitude as the acceleration time of the

precursor ions (a few µs). The resulting fragment ions will be dispersed over the *m/z* scale ranging from the *m/z* of the fragment ion to the *m/z* of the precursor ion in both linear TOF and *re*TOF, again depending on their formation time. Early metastable fragmentation is most likely responsible for the limitations of MALDI-MS of larger nucleic acid molecules [141]. Provided that the decay rate is slow compared to the acceleration time, a considerable fraction of the ions will decompose in the field-free-drift region (metastable decays). The resulting fragment ions are frequently designated as "metastable fragment ions" and travel with the speed of the precursor ions. They will arrive at the detector of a linear TOF concurrently with the precursor ions, unless ions undergo another acceleration or deceleration before being detected. Having the same speed but a lower mass than the precursor ions, they can be time-dispersed by a reflectron. Clearly, there are continuous transitions among the four categories of fragmentation and the assignment of ions to one or the other is not always unambiguous.

A high tendency of oligodeoxyribonucleotide ions to fragment, dominated by loss of nucleobases, and strongly increasing with increasing molecular mass, has been addressed in the past as the main problem for UV-MALDI-MS of large oligonucleotides [142, 143]. Recent progresses have been made for instance by Karas and coworkers who have shown that under gentle MALDI conditions, it is indeed possible to detect RNA double strands up to 20 mers [144, 145]. Covalent stabilization with chemical cross-linking is also a suitable alternative to avoid extensive dissociation under MALDI conditions [117].

2.2.8 Laser-Induced Liquid Beam Ionization/Desorption

2.2.8.1 General Mechanism

Neutral and ionized biomolecular systems can be directly infused from their natural medium (water) into the gas phase [146]. Continuous liquid beams can be produced into vacuum by injection of a liquid through a 10–20 µm aperture in diameter at a pressure in between 10 and 100 bars leading to a flow speed of 10–100 m/s [147], much lower than the velocities of rare gas atoms in a supersonic expansion. Although a large amount of water is introduced in vacuum, the background pressure in the vacuum chamber can still be maintained in the 10^{-5}–10^{-6} mbar region by capturing the liquid beam in a liquid nitrogen trap.

Alcohols are easy to use as solvents, but water may present some difficulties due to a rapid freezing [148, 149]. The solubilized biomolecular are ejected from the liquid beam by a pulsed nanosecond infrared laser exciting stretching vibrations of the solvent (e.g., the C–O stretch vibration of methanol at 9.66 µm or the O–H stretch vibration of water in the 3 µm region) [150, 151]. This technique is called LILBID for laser-induced liquid beam ionization/desorption. A first explanation of the observation of ion ejection from the liquid beam has been given by Brutschy [152]. In this model the photon absorption induces a very fast nonequilibrium phase transition beyond the supercritical point, from where the liquid expands explosively. In the decreasing particle density of the rapidly expanding supercritical

phase, the dielectric constant drops to zero, and the no longer shielded ions and counterions start to recombine. Only those ions that are too far away from their counterions escape into vacuum. Preformed ions are detected as the result of an incomplete ion neutralization process. About 10^{-5} of the solvated ions in the liquid escape into vacuum, while the rest being in neutral form. Note that the ion-to-neutral ratio in LILBID is quite low but within the same order of magnitude of the one obtained in laser desorption on solid matrices. By varying the IR laser frequency and power, it is possible to modify its penetration depth in the liquid beam. Fast neutral clusters are also ejected from the beam surface while slow ones are ejected from the bulk region inside the beam [153]. An alternative explanation [154] implies a shock wave produced by the fast energy deposition leading to an explosive thermal volume expansion. The liquid is then dispersed into very tiny droplets, too small to sustain more than a single charge. The charged droplets transfer their charge to the embedded biomolecules. This latter mechanism, "liquid dispersion model," seems to provide a better quantitative prediction of the experimental observations [155]. Such a model predicts a linear behavior of the ion yield as a function of liquid concentration over many orders of magnitude and saturation at the highest concentrations.

The liquid beam experimental setup of Abel's group is depicted in Fig. 2.8. The TOF mass spectrometer operates in the pulsed mode, which enables varying the delay time Δt between the laser pulse and the pulsed extraction/acceleration voltages, to monitor the desorption process at different times. The velocity distribution of the desorbed species can thus be determined. Brutschy and coworkers have shown that non-covalent oligonucleotide duplex can be observed in the gas phase by LILBID [156]. Furthermore, specific non-covalent complex can be dissociated by increasing the energy of the desorbing laser pulse.

2.2.8.2 Liquid Microdroplet Versus Liquid Beam

In order to reduce the large consumption of product needed to detect any ions desorbed from a continuous liquid beam, desorption can be performed on small droplets directly produced in the vacuum [157]. Note that the same acronym (LILBID) stands for this droplet method, in which "bead" replaces "beam". The main experimental challenge of laser desorption from liquid micro-droplet in vacuum is to prevent sample evaporation and freezing of the solution. This requires the generation of droplets at relatively high pressures (not far from 1 atm) and an efficient differential pumping.

In practice, microdroplets (diameter: 50 μm) are produced by a commercial piezo-driven droplet generator and injected into a vacuum chamber, through a small aperture and a differential pumping stage (Fig. 2.9). This droplet generator is located in a first vacuum chamber maintained at 0.1–0.3 Torr to avoid freezing of the aqueous solution that contains biomolecules of interest, at desired pH and salt concentrations. These droplets are ejected from the generator capillary at a speed of a few m/s, and then pass through one or two skimmers to enter a second chamber maintained at low backing pressure ($10^{-5}/10^{-6}$ Torr). There, an IR OPO laser is fired on the droplets so that laser desorption takes place from the droplets directly
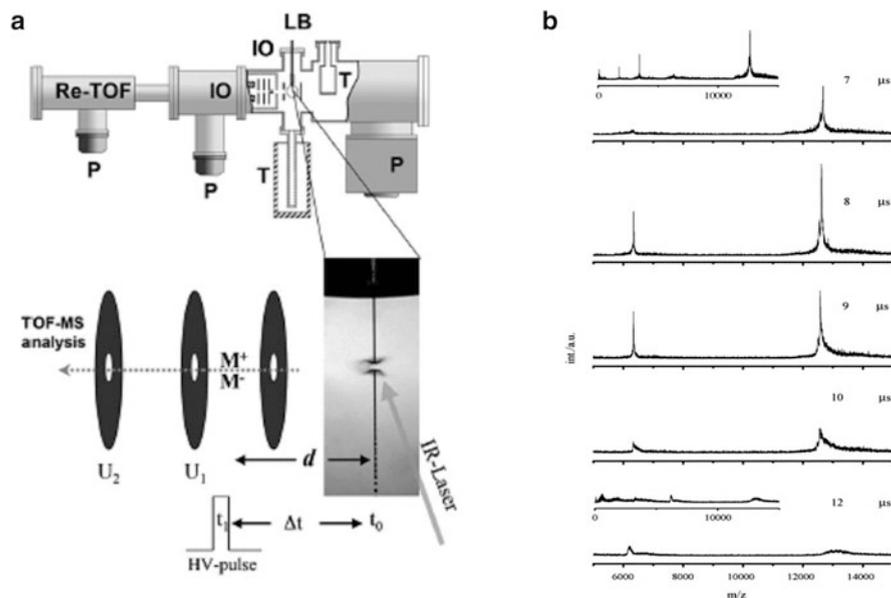
Fig. 2.8 (**a**) Liquid beam setup of the Abel's group, with laser desorption and time-of-flight mass spectrometer. Δt is the delay time between the laser shot and the trigger of the pulsed high voltage of the TOF (with permission of [155]). (**b**) Evolution of the intensity of the desorbed ions (cytochrome c) for different delay times Δt (with permission of [154])
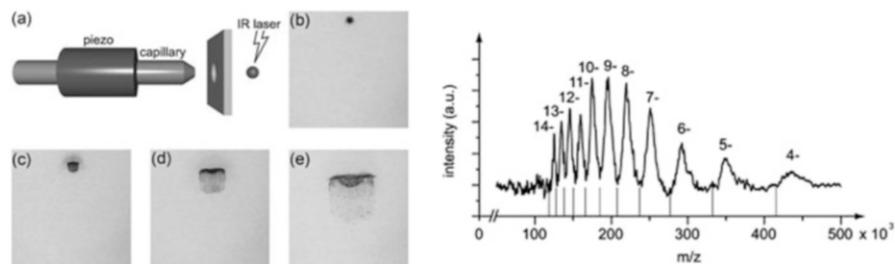
Fig. 2.9 (*Left*) Schematic view of the droplet source of the group of Brutschy: 50 μm droplets are produced in-demand by a piezo-driven generator and are irradiated in vacuum by a IR pulsed laser beam. Pictures of the explosion following laser irradiation at different times. (*Right*) Anion mass spectra of plasmid pUC19 (2,686 bp, 1.66 MDa, 80 nM) averaged over 200 droplets (with permission of [160])

into high vacuum. The repetition rate can be fixed at the laser frequency, 10 or 20 Hz, and the small amount of solution injected into vacuum allows working with reasonable pumping rates. The IR OPO desorption laser can be a simple system (one-stage rotating $LiNbO_3$ crystal pumped by a small YAG laser inside an optical cavity, delivering few mJ/pulse) that is tuned to an absorption band of the water

solution at a wavelength around 3 μm. It can operate in ambient atmosphere (residual water pressure in air does not absorb in the same range) and does not need bandwidth narrowing because liquid water absorption is broad.

Depending on the laser fluence used to evaporate the solute, non-covalently bound complex are preserved in the gas phase and partially or fully dehydrated species are observed [158]. The desorbed ions can then be analyzed by TOF mass spectrometry. However, this desorption source emits ions with a spherical isotropic expansion at velocity of around 1,000 m/s with a broad initial energy dispersion, which reduces the resolving power of the TOF mass spectrometer (about 100 in the kiloDalton range). Another drawback of this method is the careful control of the stability of the droplet (in space and time) in the interaction region with the laser. As reported by Kondow et al. [159], instability in the pointing of the droplet considerably reduces the overall efficiency of the desorption–ionization process down to few percent per laser shot.

2.2.8.3 Comparison Between LILBID and ESI

Such desorption source keeps the major advantage of the electrospray technique for which molecules are initially in solution at room temperature. As for ESI, and may be even more, the soft desorption process allows to study gas-phase non-covalent complexes under conditions that are expected to be closed to the native ones. Moreover, the micrometer size of the droplets allows the characterization of very small amounts of product (down to the picomolar range) that is suitable for analytical studies. A droplet of 50 μm diameter has a volume of roughly 60 pL. Brutschy et al. [160] have shown that they were able to detect ions from rather low solution concentration (μM) leading to an absolute sensitivity in the attomole range. Note that the droplet generator is connected to a reservoir containing about 1 mL of solution, which reduces the ultimate sensitivity of this technique deduced from the volume of a single droplet. It should be stressed that this is a pulsed source with a theoretical duty cycle of one, because the microdroplets can be produced on demand at the same frequency as the laser and the mass spectrometry cycle. This is a main improvement as compared to desorption from liquid beam and ESI, which yield duty cycles in the range of 10^{-4} or less when interfaced with pulsed analyzing lasers and mass spectrometers.

Two of the main drawbacks of the ESI technique are almost suppressed in the LILBID source. With ESI, biomolecular ions are produced at atmospheric pressure, and thus need to be injected in vacuum through capillary and/or ion optics which may induce structural rearrangement and dissociation of non-covalent complexes. Second, ESI is very sensitive to salt contaminations, and the ions are mainly formed in high and broad charge states. One of the major advantages of the LILBID technique is its increased tolerance to pH, salt concentrations, and various solvents used to dissolve the biomolecules with respect to electrospray. For instance, membrane proteins, due to their inherent hydrophobic nature, have to be solubilized by detergents which are not tolerated in conventional ESI source [161]. Laser desorption from micro-droplets directly under vacuum seems to be an ideal

technique to produce in the gas phase biomolecules and their complexes under conditions as close as those encountered in solution.

2.3 Conclusion and Outlook

Because nucleic acids play an important role in a variety of fundamental biological processes, they have triggered many experimental efforts to vaporize them into the gas phase. With the advent of the two novel soft ionization methods, ESI and MALDI, for which Fenn and Tanaka have been awarded by the Nobel Prize in 2002, and the progresses made in the molecular beam community, elegant laser spectroscopy and mass spectrometry studies can be carried out. An advantage of gas-phase based techniques is that a given complex can be analyzed in a variety of different forms, for example, as a function of charge state (anionic, neutral, or cationic), without or with additional water molecules. Nowadays, using soft desorption methods like electrospray ion source and laser desorption from microdroplets, biomolecular systems can be investigated, under carefully chosen conditions in a controlled environment, close to their native forms in solution. Native mass spectrometry has been introduced by the groups of Heck and Robinson as an emerging and promising technology allowing structural investigation of large biomolecular assemblies, like protein complexes, viruses, and ribosomes. Many examples have shown that secondary structures are conserved and that specific binary complexes are preserved during the desorption process. For DNA, this has been clearly evidenced for double strands DNA, triple helical DNA, G-quadruplex DNA, and small molecules targeting specific oligonucleotide sequences in order to control gene expression.

Progress in these applications was accelerated with the improved sensitivity and mass accuracy provided by new sources of vaporization of oligonucleotides and their non-covalent complexes, which is still the focus of active research. However, despite transferring biomolecules in gas phase is practically routine for mass spectrometry with ESI and MALDI, in the case of spectroscopy applications, there is a need for new methods dealing with neutral or cold ionic species. Yamaguchi has developed a cold-spray ionization source [162], which is a variant of ESI operating a low temperature, i.e., from -80 to $10\ ^{\circ}C$, and used for the characterization of triple and quadruple DNA strands and non-covalent complexes [163–165]. Some recent improvements have also led to new methods such as desorption from liquid microdroplets, which could in principle be used also for the study of neutral DNA species, but the LILBID technology is not yet commercially available. Further developments are therefore anticipated. Because the droplet is injected directly into high vacuum, desorption can be performed in front of a cooling supersonic beam, as it is done for neutrals produced in MALD. The small size of the droplets is particularly well suited to perform the desorption very close (1 mm or so) to the exit of a pulsed valve generating a pulsed supersonic cooling beam of rare gas (Ar, Kr, Xe), without disturbing the supersonic expansion. Finally, von Helden and coworkers [166] have very recently succeeded in embedding large

mass-selected ionic peptides (cytochrome c) produced by ESI in liquid helium droplets. This should pave the way to exciting new applications of mass spectrometry and laser spectroscopy on large, mass-selected, cold biomolecules, including nucleic acids.

2.4 Summary of Key Concepts

General Concepts
- Molecular beam and supersonic expansion: technique used to transfer in the gas phase isolated neutral molecular species (bare molecule or cluster) at low internal temperature.
- Thermal vaporization: slow heating method used to get high vapor pressure of the molecule of interest in molecular beam.
- Laser desorption: fast heating method for vaporizing molecules (either neutral or ion) from the solid (or liquid) phase to the gas phase.
- Matrix: used in laser desorption method to absorb the laser light and prevent fragmentation of the analyte.
- Electrospray: soft ionization method used to transfer ionic species from the liquid to the gas phase.
- Mass spectrometry applications: mass spectrometers are used to measure the exact mass of ionic species, to determine the stoichiometry of non-covalent complexes, and to get the primary structure of large molecular systems with sequencing methods.
- Spectroscopy applications: laser spectroscopy is used to probe the optical response of a molecule to the incident light. Could be used for photo-physical and structural studies and sequencing methods.

Concepts Specific to Nucleic Acids
- Tautomer: molecules, such as nucleobases, having labile hydrogen can be found in different structures called tautomers. Most known example is the enol–keto reaction.
- Adduct removal: the phosphodiester backbone of oligonucleotides has a very high binding affinity for free cations present in solution, which should be removed during the preparation of the sample.
- Negative ion mode: since oligonucleotides are deprotonated in native conditions, it is better to work in the negative ion mode of the mass spectrometer.
- DNA fragmentation: some oligonucleotides fragment more easily than peptides and proteins during the vaporization processes.

References

1. Vonweyssenhoff H, Selzle HL, Schlag EW (1985) Laser-desorbed large molecules in a supersonic jet. Z Naturforsch A J Phys Sci 40(7):674–676
2. Cable JR, Tubergen MJ, Levy DH (1987) Laser desorption molecular-beam spectroscopy—the electronic-spectra of tryptophan peptides in the gas-phase. J Am Chem Soc 109 (20):6198–6199
3. Meijer G, Devries MS, Hunziker HE, Wendt HR (1990) Laser desorption jet-cooling of organic-molecules—cooling characteristics and detection sensitivity. Appl Phys B 51 (6):395–403
4. Toennies JP, Vilesov AF (1998) Spectroscopy of atoms and molecules in liquid helium. Annu Rev Phys Chem 49:1–41
5. Fenn JB (2003) Electrospray wings for molecular elephants (Nobel lecture). Angew Chem Int Ed 42(33):3871–3894
6. Tanaka K (2003) The origin of macromolecule ionization by laser irradiation (Nobel lecture). Angew Chem Int Ed 42(33):3860–3870
7. Smith RD, Loo JA, Edmonds CG, Barinaga CJ, Udseth HR (1990) New developments in biochemical mass-spectrometry—electrospray ionization. Anal Chem 62(9):882–899
8. Stults JT, Marsters JC (1991) Improved electrospray ionization of synthetic oligodeoxynucleotides. Rapid Commun Mass Spectrom 5(8):359–363
9. Fuerstenau SD, Benner WH (1995) Molecular weight determination of megadalton DNA electrospray ions using charge detection time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 9(15):1528–1538
10. Gabelica V, De Pauw E (2005) Internal energy and fragmentation of ions produced in electrospray sources. Mass Spectrom Rev 24(4):566–587
11. Hofstadler SA, Sannes-Lowery KA (2006) Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. Nat Rev Drug Discov 5(7):585–595
12. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2001) Guanine tautomerism revealed by UV-UV and IR-UV hole burning spectroscopy. J Chem Phys 115(10):4604–4611
13. Chin W, Mons M, Dimicoli I, Piuzzi F, Tardivel B, Elhanine M (2002) Tautomer contribution's to the near UV spectrum of guanine: towards a refined picture for the spectroscopy of purine molecules. Eur Phys J D 20(3):347–355
14. Nir E, Grace L, Brauer B, de Vries MS (1999) REMPI spectroscopy of jot-cooled guanine. J Am Chem Soc 121(20):4896–4897
15. Nir E, Kleinermanns K, Grace L, de Vries MS (2001) On the photochemistry of purine nucleobases. J Phys Chem A 105(21):5106–5110
16. Nir E, Muller M, Grace LI, de Vries MS (2002) REMPI spectroscopy of cytosine. Chem Phys Lett 355(1–2):59–64
17. Mons M, Dimicoli I, Piuzzi F, Tardivel B, Elhanine M (2002) Tautomerism of the DNA base guanine and its methylated derivatives as studied by gas-phase infrared and ultraviolet spectroscopy. J Phys Chem A 106(20):5088–5094
18. Canuel C, Mons M, Piuzzi F, Tardivel B, Dimicoli I, Elhanine M (2005) Excited states dynamics of DNA and RNA bases: characterization of a stepwise deactivation pathway in the gas phase. J Chem Phys 122(7):074316
19. Mons M, Piuzzi F, Dimicoli I, Gorb L, Leszczynski J (2006) Near-UV resonant two-photon ionization spectroscopy of gas phase guanine: evidence for the observation of three rare tautomers. J Phys Chem A 110(38):10921–10924
20. Plutzer C, Kleinermanns K (2002) Tautomers and electronic states of jet-cooled adenine investigated by double resonance spectroscopy. Phys Chem Chem Phys 4(20):4877–4882
21. Kim NJ, Jeong G, Kim YS, Sung J, Kim SK, Park YD (2000) Resonant two-photon ionization and laser induced fluorescence spectroscopy of jet-cooled adenine. J Chem Phys 113 (22):10051–10055

22. Kang H, Lee KT, Jung B, Ko YJ, Kim SK (2002) Intrinsic lifetimes of the excited state of DNA and RNA bases. J Am Chem Soc 124(44):12958–12959

23. Nir E, Imhof P, Kleinermanns K, de Vries MS (2000) REMPI spectroscopy of laser desorbed guanosines. J Am Chem Soc 122(33):8091–8092

24. Nir E, Hunig I, Kleinermanns K, de Vries MS (2004) Conformers of guanosines and their vibrations in the electronic ground and excited states, as revealed by double-resonance spectroscopy and ab initio calculations. ChemPhysChem 5(1):131–137

25. Abo-Riziq A, Crews BO, Compagnon I, Oomens J, Meijer G, Von Helden G, Kabelac M, Hobza P, de Vries MS (2007) The Mid-IR spectra of 9-ethyl guanine, guanosine, and 2-deoxyguanosine. J Phys Chem A 111(31):7529–7536

26. Nir E, Kleinermanns K, de Vries MS (2000) Pairing of isolated nucleic-acid bases in the absence of the DNA backbone. Nature 408(6815):949–951

27. Kang H, Lee KT, Kim SK (2002) Femtosecond real time dynamics of hydrogen bond dissociation in photoexcited adenine-water clusters. Chem Phys Lett 359(3–4):213–219

28. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2002) Pairing of the nucleobases guanine and cytosine in the gas phase studied by IR-UV double-resonance spectroscopy and ab initio calculations. Phys Chem Chem Phys 4(5):732–739

29. Plutzer C, Hunig I, Kleinermanns K, Nir E, de Vries MS (2003) Pairing of isolated nucleobases: double resonance laser spectroscopy of adenine-thymine. ChemPhysChem 4 (8):838–842

30. Abo-Riziq A, Grace L, Nir E, Kabelac M, Hobza P, de Vries MS (2005) Photochemical selectivity in guanine-cytosine base-pair structures. Proc Natl Acad Sci USA 102(1):20–23

31. Kim NJ, Kim YS, Jeong G, Ahn TK, Kim SK (2002) Hydration of DNA base cations in the gas phase. Int J Mass Spectrom 219(1):11–21

32. Chin W, Mons M, Piuzzi F, Tardivel B, Dimicoli I, Gorb L, Leszczynski J (2004) Gas phase rotamers of the nucleobase 9-methylguanine enol and its monohydrate: optical spectroscopy and quantum mechanical calculations. J Phys Chem A 108(40):8237–8243

33. Abo-Riziq A, Crews B, Grace L, de Vries MS (2005) Microhydration of guanine base pairs. J Am Chem Soc 127(8):2374–2375

34. Crews B, Abo-Riziq A, Grace L, Callahan M, Kabelac M, Hobza P, de Vries MS (2005) IR-UV double resonance spectroscopy of guanine-H2O clusters. Phys Chem Chem Phys 7 (16):3015–3020

35. Canuel C, Elhanine M, Mons M, Piuzzi F, Tardivel B, Dimicoli I (2006) Time-resolved photoelectron and photoion fragmentation spectroscopy study of 9-methyladenine and its hydrates: a contribution to the understanding of the ultrafast radiationless decay of excited DNA bases. Phys Chem Chem Phys 8(34):3978–3987

36. Kim S, Wheeler SE, Schaefer HF (2006) Microsolvation effects on the electron capturing ability of thymine: thymine-water clusters. J Chem Phys 124(20):204310

37. Choi MY, Miller RE (2006) Four tautomers of isolated guanine from infrared laser spectroscopy in helium nanodroplets. J Am Chem Soc 128(22):7320–7328

38. Fenn JB (2000) Mass spectrometric implications of high-pressure ion sources. Int J Mass Spectrom 200(1–3):459–478

39. Schermann JP (2008) Spectroscopy and modelling of biomolecular building blocks. Elsevier, Amsterdam

40. Kim SK, Lee W, Herschbach DR (1996) Cluster beam chemistry: hydration of nucleic acid bases; Ionization potentials of hydrated adenine and thymine. J Phys Chem 100(19):7933–7937

41. Piuzzi F, Dimicoli I, Mons M, Tardivel B, Zhao QC (2000) A simple laser vaporization source for thermally fragile molecules coupled to a supersonic expansion: application to the spectroscopy of tryptophan. Chem Phys Lett 320(3–4):282–288

42. Saigusa H, Tomioka A, Katayama T, Iwase E (2006) A matrix-free laser desorption method for production of nucleobase clusters and their hydrates. Chem Phys Lett 418(1–3):119–125

43. Nir E, de Vries MS (2002) Fragmentation of laser-desorbed 9-substituted adenines. Int J Mass Spectrom 219(1):133–138
44. Zhou J, Kostko O, Nicolas C, Tang XN, Belau L, de Vries MS, Ahmed M (2009) Experimental observation of guanine tautomers with VUV photoionization. J Phys Chem A 113 (17):4829–4832
45. Toennies JP, Vilesov AF (2004) Superfluid helium droplets: a uniquely cold nanomatrix for molecules and molecular complexes. Angew Chem Int Ed 43(20):2622–2648
46. Buchenau H, Knuth EL, Northby J, Toennies JP, Winkler C (1990) Mass-spectra and time-of-flight distributions of helium cluster beams. J Chem Phys 92(11):6875–6889
47. Denifl S, Zappa F, Mahr I, Lecointre J, Probst M, Mark TD, Scheier P (2006) Mass spectrometric investigation of anions formed upon free electron attachment to nucleobase molecules and clusters embedded in superfluid helium droplets. Phys Rev Lett 97(4)
48. Kebarle P, Ho Y (1997) In: Cole RB (ed) Electrospray ionization mass spectrometry. Wiley Interscience, New York, NY
49. Gaskell SJ (1997) Electrospray: principles and practice. J Mass Spectrom 32(7):677–688
50. Konermann L, Ahadi E, Rodriguez AD, Vahidi S (2013) Unraveling the mechanism of electrospray ionization. Anal Chem 85(1):2–9
51. Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379(6564):466–469
52. Wilm MS, Mann M (1994) Electrospray and Taylor-Cone theory, Doles beam of macromolecules at last. Int J Mass Spectrom 136(2–3):167–180
53. Juraschek R, Dulcks T, Karas M (1999) Nanoelectrospray—more than just a minimized-flow electrospray ionization source. J Am Soc Mass Spectrom 10(4):300–308
54. Dole M, Mack LL, Hines RL (1968) Molecular beams of macroions. J Chem Phys 49(5):2240
55. Cole RB (2000) Some tenets pertaining to electrospray ionization mass spectrometry. J Mass Spectrom 35(7):763–772
56. Iribarne JV, Thomson BA (1976) Evaporation of small ions from charged droplets. J Chem Phys 64(6):2287–2294
57. Thomson BA, Iribarne JV (1979) Field-induced ion evaporation from liquid surfaces at atmospheric-pressure. J Chem Phys 71(11):4451–4463
58. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1990) Electrospray ionization-principles and practice. Mass Spectrom Rev 9(1):37–70
59. Fenn JB (1993) Ion formation from charged droplets—roles of geometry, energy, and time. J Am Soc Mass Spectrom 4(7):524–535
60. Kebarle P, Tang L (1993) From ions in solution to ions in the gas-phase—the mechanism of electrospray mass-spectrometry. Anal Chem 65(22):A972–A986
61. Ahadi E, Konermann L (2011) Ejection of solvated ions from electrosprayed methanol/water nanodroplets studied by molecular dynamics simulations. J Am Chem Soc 133(24):9354–9363
62. Konermann L, Rodriguez AD, Liu JJ (2012) On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. Anal Chem 84(15):6798–6804
63. Limbach PA, Crain PF, McCloskey JA (1995) Molecular-mass measurement of intact ribonucleic-acids via electrospray-ionization quadrupole mass-spectrometry. J Am Soc Mass Spectrom 6(1):27–39
64. Potier N, Vandorsselaer A, Cordier Y, Roch O, Bischoff R (1994) Negative electrospray-ionization mass-spectrometry of synthetic and chemically-modified oligonucleotides. Nucleic Acids Res 22(19):3895–3903
65. McGinnis AC, Chen BY, Bartlett MG (2012) Chromatographic methods for the determination of therapeutic oligonucleotides. J Chromatogr B Analyt Technol Biomed Life Sci 883:76–94
66. Huber CG (1998) Micropellicular stationary phases for high-performance liquid chromatography of double-stranded DNA. J Chromatogr A 806(1):3–30

67. Holzl G, Oberacher H, Pitsch S, Stutz A, Huber CG (2005) Analysis of biological and synthetic ribonucleic acids by liquid chromatography-mass spectrometry using monolithic capillary columns. Anal Chem 77(2):673–680

68. Premstaller A, Oberacher H, Huber CG (2000) High-performance liquid chromatography-electrospray ionization mass spectrometry of single- and double-stranded nucleic acids using monolithic capillary columns. Anal Chem 72(18):4386–4393

69. Huber CG, Krajete A (1999) Analysis of nucleic acids by capillary ion-pair reversed-phase HPLC coupled to negative-ion electrospray ionization mass spectrometry. Anal Chem 71(17):3730–3739

70. Sangaraju D, Goggin M, Walker V, Swenberg J, Tretyakova N (2012) NanoHPLC-nanoESI (+)-MS/MS quantitation of bis-N7-guanine DNA-DNA cross-links in tissues of B6C3F1 mice exposed to subppm levels of 1,3-butadiene. Anal Chem 84(3):1732–1739

71. Tretyakova N, Goggin M, Sangaraju D, Janis G (2012) Quantitation of DNA adducts by stable isotope dilution mass spectrometry. Chem Res Toxicol 25(10):2007–2035

72. Owen BC, Haupert LJ, Jarrell TM, Marcum CL, Parsell TH, Abu-Omar MM, Bozell JJ, Black SK, Kenttamaa HI (2012) High-performance liquid chromatography/high-resolution multiple stage tandem mass spectrometry using negative-ion-mode hydroxide-doped electrospray ionization for the characterization of lignin degradation products. Anal Chem 84(14):6000–6007

73. Schmidt A, Bahr U, Karas M (2001) Influence of pressure in the first pumping stage on analyte desolvation and fragmentation in nano-ESI MS. Anal Chem 73(24):6040–6046

74. Ganem B, Li YT, Henion JD (1993) Detection of oligonucleotide duplex forms by ion-spray mass-spectrometry. Tetrahedron Lett 34(9):1445–1448

75. Lightwahl KJ, Springer DL, Winger BE, Edmonds CG, Camp DG, Thrall BD, Smith RD (1993) Observation of a small oligonucleotide duplex by electrospray ionization mass-spectrometry. J Am Chem Soc 115(2):803–804

76. Goodlett DR, Camp DG, Hardin CC, Corregan M, Smith RD (1993) Direct observation of a DNA quadruplex by electrospray ionization mass-spectrometry. Biol Mass Spectrom 22(3):181–183

77. Gale DC, Goodlett DR, Lightwahl KJ, Smith RD (1994) Observation of duplex DNA-drug noncovalent complexes by electrospray-ionization mass-spectrometry. J Am Chem Soc 116(13):6027–6028

78. Beck JL (2011) Developments in electrospray ionization mass spectrometry of non-covalent DNA-ligand complexes. Aust J Chem 64(6):705–717

79. Aggerholm T, Nanita SC, Koch KJ, Cooks RG (2003) Clustering of nucleosides in the presence of alkali metals: biologically relevant quartets of guanosine, deoxyguanosine and uridine observed by ESI-MS/MS. J Mass Spectrom 38(1):87–97

80. Ali OY, Randell NM, Fridgen TD (2012) Primary fragmentation pathways of gas phase M (Uracil-H)(Uracil) plus complexes (M = Zn, Cu, Ni, Co, Fe, Mn, Cd, Pd, Mg, Ca, Sr, Ba, and Pb): loss of uracil versus HNCO. ChemPhysChem 13(6):1507–1513

81. Seo J, Hong ES, Yoon HJ, Shin SK (2012) Specific and nonspecific bindings of alkaline-earth metal ions to guanine-quadruplex thrombin-binding aptamer DNA. Int J Mass Spectrom 330:262–270

82. Yang ZB, Rodgers MT (2005) Influence of methylation on the properties of uracil and its noncovalent interactions with alkali metal ions—threshold collision-induced dissociation and theoretical studies. Int J Mass Spectrom 241(2–3):225–242

83. Ruan CH, Rodgers MT (2009) Modeling metal cation-phosphate interactions in nucleic acids: activated dissociation of Mg+, Al+, Cu+, and Zn + complexes of triethyl phosphate. J Am Chem Soc 131(31):10918–10928

84. Ruan CH, Huang H, Rodgers MT (2007) Modeling metal cation-phosphate interactions in nucleic acids in the gas phase via alkali metal cation-triethyl phosphate complexes. J Phys Chem A 111(51):13521–13527

85. Salpin JY, Guillaumont S, Tortajada J, Lamsabhi AM (2009) Gas-phase interactions between lead(II) ions and thiouracil nucleobases: a combined experimental and theoretical study. J Am Soc Mass Spectrom 20(3):359–369

86. Salpin JY, Guillaumont S, Ortiz D, Tortajada J, Maitre P (2011) Direct evidence for tautomerization of the uracil moiety within the Pb2+/Uridine-5′-monophosphate complex: a combined tandem mass spectrometry and IRMPD study. Inorg Chem 50(16):7769–7778

87. Salpin JY, Gamiette L, Tortajada J, Besson T, Maitre P (2011) Structure of Pb2+/dCMP and Pb2+/CMP complexes as characterized by tandem mass spectrometry and IRMPD spectroscopy. Int J Mass Spectrom 304(2–3):154–164

88. Chiavarino B, Crestoni ME, Fornarini S, Scuderi D, Salpin JY (2013) Interaction of cisplatin with adenine and guanine: a combined IRMPD, MS/MS, and theoretical study. J Am Chem Soc 135(4):1445–1455

89. Beck JL, Colgrave ML, Ralph SF, Sheil MM (2001) Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. Mass Spectrom Rev 20 (2):61–87

90. Alves S, Woods A, Delvolve A, Tabet JC (2008) Influence of salt bridge interactions on the gas-phase stability of DNA/peptide complexes. Int J Mass Spectrom 278(2–3):122–128

91. Kapur A, Beck JL, Sheil MM (1999) Observation of daunomycin and nogalamycin complexes with duplex DNA using electrospray ionisation mass spectrometry. Rapid Commun Mass Spectrom 13(24):2489–2497

92. Rosu F, De Pauw E, Guittat L, Alberti P, Lacroix L, Maillet P, Riou JF, Mergny JL (2003) Selective interaction of ethidium derivatives with quadruplexes: an equilibrium dialysis and electrospray ionization mass spectrometry analysis. Biochemistry 42(35):10361–10371

93. Gabelica V, De Pauw E, Rosu F (1999) Interaction between antitumor drugs and a double-stranded oligonucleotide studied by electrospray ionization mass spectrometry. J Mass Spectrom 34(12):1328–1337

94. Rosu F, Gabelica V, Houssier C, De Pauw E (2012) Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry. Nucleic Acids Res 30(16)

95. Wan KX, Shibue T, Gross ML (2000) Non-covalent complexes between DNA-binding drugs and double-stranded oligodeoxynucleotides: a study by ESI ion-trap mass spectrometry. J Am Chem Soc 122(2):300–307

96. Wan KX, Gross ML, Shibue T (2000) Gas-phase stability of double-stranded oligodeoxynucleotides and their noncovalent complexes with DNA-binding drugs as revealed by collisional activation in an ion trap. J Am Soc Mass Spectrom 11(5):450–457

97. Greig MJ, Robinson JM (2000) Detection of oligonucleotide-ligand complexes by ESI-MS (DOLCE-MS) as a component of high throughput screening. J Biomol Screen 5(6):441–454

98. van den Heuvel RH, Heck AJR (2004) Native protein mass spectrometry: from intact oligomers to functional machineries. Curr Opin Chem Biol 8(5):519–526

99. Schmidt C, Kramer K, Urlaub H (2012) Investigation of protein-RNA interactions by mass spectrometry-techniques and applications. J Proteome 75(12):3478–3494

100. Hernandez H, Robinson CV (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. Nat Protoc 2(3):715–726

101. Heck AJR (2008) Native mass spectrometry: a bridge between interactomics and structural biology. Nat Methods 5(11):927–933

102. Hanson CL, Robinson CV (2004) Protein-nucleic acid interactions and the expanding role of mass spectrometry. J Biol Chem 279(24):24907–24910

103. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA (2008) In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci USA 105(2):512–517

104. Gordiyenko Y, Videler H, Zhou M, McKay AR, Fucini P, Biegel E, Muller V, Robinson CV (2010) Mass spectrometry defines the stoichiometry of ribosomal stalk complexes across the phylogenetic tree. Mol Cell Proteomics 9(8):1774–1783

105. Deroo S, Hyung SJ, Marcoux J, Gordiyenko Y, Koripella RK, Sanyal S, Robinson CV (2012) Mechanism and rates of exchange of L7/L12 between ribosomes and the effects of binding EF-G. ACS Chem Biol 7(6):1120–1127
106. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 Daltons. Anal Chem 60(20):2299–2301
107. Knochenmuss R, Zenobi R (2003) MALDI ionization: the role of in-plume processes. Chem Rev 103(2):441–452
108. Georgiou S, Koubenakis A (2003) Laser-induced material ejection from model molecular solids and liquids: mechanisms, implications, and applications. Chem Rev 103(2):349–393
109. Quist AP, Huthfehre T, Sundqvist BUR (1994) Total yield measurements in matrix-assisted laser-desorption using a quartz-crystal microbalance. Rapid Commun Mass Spectrom 8 (2):149–154
110. Stevenson E, Breuker K, Zenobi R (2000) Internal energies of analyte ions generated from different matrix-assisted laser desorption/ionization matrices. J Mass Spectrom 35(8):1035–1041
111. Spengler B, Pan Y, Cotter RJ, Kan LS (1990) Molecular-weight determination of underivatized oligodeoxyribonucleotides by positive-ion matrix-assisted ultraviolet laser-desorption mass-spectrometry. Rapid Commun Mass Spectrom 4(4):99–102
112. Bradick M, Costa J, Chelo IM (2011) Genotyping with sequenom. In: Methods in molecular biology, vol 772. Springer, pp 193–210
113. Meyer K, Ueland PM (2011) Use of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for multiplex genotyping. In: Advances in clinical chemistry, vol 53. Elsevier Academic, San Diego, CA, pp 1–29
114. Giessing AMB, Kirpekar F (2012) Mass spectrometry in the biology of RNA and its modifications. J Proteome 75(12):3434–3449
115. Terrier P, Tortajada J, Zin G, Buchmann W (2007) Noncovalent complexes between DNA and basic polypeptides or polyamines by MALDI-TOF. J Am Soc Mass Spectrom 18 (11):1977–1989
116. Bolbach G (2005) Matrix-assisted laser desorption/ionization analysis of non-covalent complexes: fundamentals and applications. Curr Pharm Des 11(20):2535–2557
117. Madler S, Erba EB, Zenobi R (2013) MALDI-ToF mass spectrometry for studying noncovalent complexes of biomolecules. In: Applications of Maldi-ToF spectroscopy, vol 331. Springer, Berlin, pp 1–36
118. Wu KJ, Steding A, Becker CH (1993) Matrix-assisted laser desorption time-of-flight mass-spectrometry of oligonucleotides using 3-hydroxypicolinic acid as an ultraviolet-sensitive matrix. Rapid Commun Mass Spectrom 7(2):142–146
119. Little DP, Jacob A, Becker T, Braun A, Darnhofer-Demar B, Jurinke C, van den Boom D, Koster H (1997) Direct detection of synthetic and biologically generated double-stranded DNA by MALDI-TOF MS. Int J Mass Spectrom 169:323–330
120. Tang K, Taranenko NI, Allman SL, Chang LY, Chen CH (1994) Detection of 500-nucleotide DNA by laser-desorption mass-spectrometry. Rapid Commun Mass Spectrom 8(9):727–730
121. Lin H, Hunter JM, Becker CH (1999) Laser desorption of DNA oligomers larger than one kilobase from cooled 4-nitrophenol. Rapid Commun Mass Spectrom 13(23):2335–2340
122. Zhou LH, Deng HM, Deng QY, Zhao SK (2004) A mixed matrix of 3-hydroxypicolinic acid and pyrazinecarboxylic acid for matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of oligodeoxynucleotides. Rapid Commun Mass Spectrom 18(7):787–794
123. Nordhoff E, Ingendoh A, Cramer R, Overberg A, Stahl B, Karas M, Hillenkamp F, Crain PF (1992) Matrix-assisted laser desorption ionization mass-spectrometry of nucleic-acids with wavelengths in the ultraviolet and infrared. Rapid Commun Mass Spectrom 6(12):771–776
124. Nordhoff E, Cramer R, Karas M, Hillenkamp F, Kirpekar F, Kristiansen K, Roepstorff P (1993) Ion stability of nucleic-acids in infrared matrix-assisted laser-desorption ionization mass-spectrometry. Nucleic Acids Res 21(15):3347–3357
125. Currie GJ, Yates JR (1993) Analysis of oligodeoxynucleotides by negative-ion matrix-assisted laser-desorption mass-spectrometry. J Am Soc Mass Spectrom 4(12):955–963

126. Pieles U, Zurcher W, Schar M, Moser HE (1993) Matrix-assisted laser-desorption ionization time-of-flight mass-spectrometry—a powerful tool for the mass and sequence-analysis of natural and modified oligonucleotides. Nucleic Acids Res 21(14):3191–3196

127. Li YCL, Cheng SW, Chan TWD (1998) Evaluation of ammonium salts as co-matrices for matrix-assisted laser desorption/ionization mass spectrometry of oligonucleotides. Rapid Commun Mass Spectrom 12(15):993–998

128. Asara JM, Allison J (1999) Enhanced detection of oligonucleotides in UV MALDI MS using the tetraamine spermine as a matrix additive. Anal Chem 71(14):2866–2870

129. Terrier P, Tortajada J, Buchmann W (2007) A study of noncovalent complexes involving single-stranded DNA and polybasic compounds using nanospray mass spectrometry. J Am Soc Mass Spectrom 18(2):346–358

130. Distler AM, Allison J (2002) Additives for the stabilization of double-stranded DNA in UV-MALDI MS. J Am Soc Mass Spectrom 13(9):1129–1137

131. Langley GJ, Herniman JM, Davies NL, Brown T (1999) Simplified sample preparation for the analysis of oligonucleotides by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 13(17):1717–1723

132. Liu CL, Wu QY, Harms AC, Smith RD (1996) On line microdialysis sample cleanup for electrospray ionization mass-spectrometry of nucleic acid samples. Anal Chem 68(18):3295–3299

133. Smirnov IP, Hall LR, Ross PL, Haff LA (2001) Application of DNA-binding polymers for preparation of DNA for analysis by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 15(16):1427–1432

134. Gilar M, Belenky A, Wang BH (2001) High-throughput biopolymer desalting by solid-phase extraction prior to mass spectrometric analysis. J Chromatogr A 921(1):3–13

135. Williams TL, Fenselau C (1998) p-nitroaniline/glycerol: a binary liquid matrix for matrix-assisted laser desorption/ionization analysis. Eur Mass Spectrom 4(5):379–383

136. Berkenkamp S, Menzel C, Karas M, Hillenkamp F (1997) Performance of infrared matrix-assisted laser desorption/ionization mass spectrometry with lasers emitting in the 3 mu m wavelength range. Rapid Commun Mass Spectrom 11(13):1399–1406

137. Menzel C, Berkenkamp S, Hillenkamp F (1999) Infrared matrix-assisted laser desorption/ionization mass spectrometry with a transversely excited atmospheric pressure carbon dioxide laser at 10.6 mu m wavelength with static and delayed ion extraction. Rapid Commun Mass Spectrom 13(1):26–32

138. Kirpekar F, Berkenkamp S, Hillenkamp F (1999) Detection of double-stranded DNA by IR- and UV-MALDI mass spectrometry. Anal Chem 71(13):2334–2339

139. Colby SM, King TB, Reilly JP (1994) Improving the resolution of matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry by exploiting the correlation between ion position and velocity. Rapid Commun Mass Spectrom 8(11):865–868

140. Brown RS, Lennon JJ (1995) Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser-desorption ionization linear time-of-flight mass-spectrometer. Anal Chem 67(13):1998–2003

141. Wu KJ, Shaler TA, Becker CH (1994) Time-of-flight mass-spectrometry of underivatized single-stranded-DNA oligomers by matrix-assisted laser-desorption. Anal Chem 66(10):1637–1645

142. Schneider K, Chait BT (1993) Matrix-assisted laser-desorption mass-spectrometry of homo-polymer oligodeoxyribonucleotides—influence of base composition on the mass-spectrometric response. Org Mass Spectrom 28(11):1353–1361

143. Li YZ, Tang K, Little DP, Koster H, Hunter RL, McIver RT (1996) High-resolution MALDI Fourier transform mass spectrometry of oligonucleotides. Anal Chem 68(13):2090–2096

144. Bahr U, Ayguen H, Karas M (2008) Detection and relative quantification of siRNA double strands by MALDI mass spectrometry. Anal Chem 80(16):6280–6285

145. Bahr U, Aygun H, Karas M (2009) Sequencing of single and double stranded RNA Oligonucleotides by acid hydrolysis and MALDI mass spectrometry. Anal Chem 81(8):3173–3179

146. Kleinekofort W, Avdiev J, Brutschy B (1996) A new method of laser desorption mass spectrometry for the study of biological macromolecules. Int J Mass Spectrom Ion Process 152(2–3):135–142
147. Wattenberg A, Sobott F, Barth HD, Brutschy B (1999) Laser desorption mass spectrometry on liquid beams. Eur Mass Spectrom 5(2):71–76
148. Mafune F, Takeda Y, Nagata T, Kondow T (1992) Formation and ejection of cluster ions from a liquid beam of aniline ethanol solution by laser photoionization. Chem Phys Lett 199 (6):615–620
149. Kondow T, Mafune F (2000) Structures and dynamics of molecules on liquid beam surfaces. Annu Rev Phys Chem 51:731–761
150. Sobott F, Wattenberg A, Kleinekofort W, Pfenninger A, Brutschy B (1998) Laser desorption mass spectrometry on thin liquid jets. Fresenius J Anal Chem 360(7–8):745–749
151. Charvat A, Stasicki B, Abel B (2006) Product screening of fast reactions in IR-laser-heated liquid water filaments in a vacuum by mass spectrometry. J Phys Chem A 110(9):3297–3306
152. Sobott F, Kleinekofort W, Brutschy B (1997) Cation selectivity of natural and synthetic ionophores probed with laser-induced liquid beam mass spectrometry. Anal Chem 69 (17):3587–3594
153. Horimoto N, Kohno J, Mafune F, Kondow T (2000) Ejection mechanism of molecules and neutral clusters from liquid beam under irradiation of IR laser. Chem Phys Lett 318(6):536–542
154. Abel B, Charvat A, Diederichsen U, Faubel M, Girmannc B, Niemeyer J, Zeeck A (2005) Applications, features, and mechanistic aspects of liquid water beam desorption mass spectrometry. Int J Mass Spectrom 243(2):177–188
155. Charvat A, Abel B (2007) How to make big molecules fly out of liquid water: applications, features and physics of laser assisted liquid phase dispersion mass spectrometry. Phys Chem Chem Phys 9(26):3335–3360
156. Kleinekofort W, Schweitzer M, Engels JW, Brutschy B (1997) Analysis of double-stranded oligonucleotides by laser-induced liquid beam mass spectrometry. Int J Mass Spectrom Ion Process 163(1–2):L1–L4
157. Morgner N, Barth HD, Brutschy B (2006) A new way to detect noncovalently bonded complexes of biomolecules from liquid micro-droplets by laser mass spectrometry. Aust J Chem 59(2):109–114
158. Morgner N, Kleinschroth T, Barth HD, Ludwig B, Brutschy B (2007) A novel approach to analyze membrane proteins by laser mass spectrometry: from protein subunits to the integral complex. J Am Soc Mass Spectrom 18(8):1429–1438
159. Kohno J, Toyama N, Kondow T (2006) Ion formation to the gas phase by laser ablation on a droplet beam. Chem Phys Lett 420(1–3):146–150
160. Hoffmann J, Schmidt TL, Heckel A, Brutschy B (2009) Probing the limits of liquid droplet laser desorption mass spectrometry in the analysis of oligonucleotides and nucleic acids. Rapid Commun Mass Spectrom 23(14):2176–2180
161. Hoffmann J, Sokolova L, Preiss L, Hicks DB, Krulwich TA, Morgner N, Wittig I, Schagger H, Meier T, Brutschy B (2010) ATP synthases: cellular nanomotors characterized by LILBID mass spectrometry. Phys Chem Chem Phys 12(41):13375–13382
162. Yamaguchi K (2003) Cold-spray ionization mass spectrometry: principle and applications. J Mass Spectrom 38(5):473–490
163. Sakamoto S, Yamaguchi K (2003) Low Tm DNA duplexes observed by cold-spray ionization mass spectrometry. Tetrahedron Lett 44(16):3341–3344
164. Sakamoto S, Yamaguchi K (2003) Hyperstranded DNA architectures observed by cold-spray ionization mass spectrometry. Angew Chem Int Ed 42(8):905
165. Kuzuhara T, Sei Y, Yamaguchi K, Suganuma M, Fujiki H (2006) DNA and RNA as new binding targets of green tea catechins. J Biol Chem 281(25):17446–17456
166. Bierau F, Kupser P, Meijer G, von Helden G (2010) Catching proteins in liquid helium droplets. Phys Rev Lett 105(13):133402

Structure of Nucleic Acids in the Gas Phase

3

Annalisa Arcella, Guillem Portella, and Modesto Orozco

Abstract

Evolution has refined nucleic acids to display well-defined three-dimensional structures that are functional under aqueous physiological conditions. While the structure of nucleic acids is well known in solution, it is unclear how nucleic acids react when transferred to a fully anhydrous environment. Simple physical chemistry considerations suggest that a heavily charged poly-anion would adopt fully extended conformations in vacuum, and that multistranded structure would dissociate, to guarantee that charged residues separate as much as possible to reduce Coulomb repulsion. However, and quite counterintuitively, a vast amount of experiments demonstrate that this is not the case and that oligomeric nucleic acids adopt quite compact structures in the gas phase, which in some cases might preserve memories of the original conformation in solution. In this chapter, we review our current understanding of nucleic acid structure in the gas phase.

A. Arcella • G. Portella
Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

Joint Research Program in Computational Biology, Institute for Research in Biomedicine and Barcelona Supercomputing Center, Barcelona, Spain

M. Orozco (✉)
Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain

Joint Research Program in Computational Biology, Institute for Research in Biomedicine and Barcelona Supercomputing Center, Barcelona, Spain

Department of Biochemistry and Molecular Biology, University of Barcelona, Barcelona, Spain
e-mail: modesto.orozco@irbbarcelona.org

Abbreviations

ESI-MS	Electrospray soft ionization mass spectrometry
IMS	Ion mobility spectrometry
IVNT	In vacuum native structure
XFEL	X-ray free electron laser
CCS	Collision cross section
MD	Molecular dynamics
QM	Quantum mechanics
DFT	Density functional theory
CCSD(T)/CBS	Coupled cluster with single, double, and triple excitation/complete basis set

3.1 Introduction

Nucleic acids are flexible and polymorphic molecules [1–4] . Their spatial configuration is the result of numerous effects: (1) intrinsic conformational restrictions related to the nature of the backbone; (2) stabilization effects due to nucleobase stacking (arising for dispersion interaction originated from the overlap of pi-orbitals of nucleobases), which can amount for around 10 kcal/mol per nucleobase dimer in the gas phase [5]; (3) stabilization effects linked to hydrogen-bond base pairing (mostly an electrostatic interaction, which amounts for 12–25 kcal/mol per base pair in the gas phase [6, 7]; (4) desolvation penalty connected to the poorer hydration of stacked or hydrogen-bond paired nucleobases compared to the isolated ones; (5) Coulomb repulsion between phosphate groups; and (6) solvent screening that reduces the intensity of such electrostatic repulsion. The balance between all these interactions is crucial to provide enough stability to nucleic acids, but at the same time to allow them to adopt other conformations required for their biological actions (replication, transcription, and transduction). It is then clear that an unbalance in these interactions will have a major impact in either the structure or the functionality of nucleic acids.

Phosphate–phosphate interactions produce very strong repulsion. For example, simple Coulomb calculations for a canonical B-DNA show that the unscreened repulsion between phosphates (i.e., gas-phase situation) can be as large as 29 kcal/mol (intra-strand) and 51 kcal/mol kcal/mol (inter-strand). These numbers even increased for more compact A-RNA structures 32 kcal/mol (intra-strand) and 59 kcal/mol (inter-strand) and can be even larger in some compact RNA structures. In a long oligonucleotide, phosphate–phosphate repulsion increases nonlinearly with the size of the oligonucleotide (see Fig. 3.1) reaching extreme values even for short chains. Clearly, solvent screening (see also Fig. 3.1) is crucial to reduce this strong repulsion to levels that can be compensated by favorable energy terms.
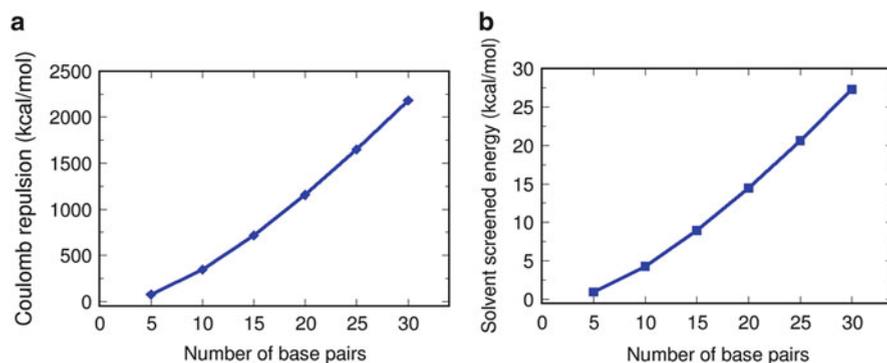
Fig. 3.1 Representation of phosphate–phosphate electrostatic repulsion for B-DNAs of different lengths (from 5 to 30 mer) in the gas phase (*left*) and in physiological conditions (*right*). Coulomb potential was used to determine the electrostatic repulsion in the gas phase, while Mehler–Solmajer potential [8] was used in the case of aqueous simulations

Simple calculations above help to understand the importance of solvent for the structure of nucleic acids. Thus, it is known since the early ages of nucleic acids research that changes in ionic strength or water activity lead to changes in nucleic acid structure or stability [1–4]. For example, addition of co-solvents, such as ethanol, leads to a dramatic conformational change between B and A forms in DNA duplexes [9], high ionic strength is coupled with the B→Z transition in d (G/C) rich DNA duplexes [10], some divalent ions are very important to stabilize triplexes [11], and G-quadruplexes display a dramatic loss of stability as a consequence of the reduction in the monovalent cationic concentration [12, 13]. Even more impressive is that moderate concentrations of some co-solvents of reduced polarity, such as phenol, pyridine, propanol, or 1,4 Dioxane, lead to the complete unfolding of duplex DNA [14], while in some cases can stabilize other forms of DNA [15].

In the light of these considerations, we can be tempted to conclude that the structure of nucleic acids should suffer dramatically when placed out of high dielectric solvents, leading to unfolding when placed in the gas phase. This process is indeed observed in molecular dynamics (MD) simulations of a duplex DNA, which is instantaneously transferred to the gas phase keeping its solution ionic state (i.e., the charge state of residues in physiological DNA; see Fig. 3.2). However, mass spectroscopic studies using mild electrospray vaporization conditions (ESI-MS) [16–30] provide convincing experimental evidence demonstrating that nucleic acids maintain some of the solution-phase structural features when transferred into the gas phase. It is clear, for example, that duplexes, triplexes, and quadruplexes are in the gas phase in the same oligomeric state found in solution. Furthermore, even complexes of different DNA structures with small non-covalent ligands are detected intact in the gas phase (see [16–30] and Chap. 9). This finding suggests that the some structural determinants (for example, groove geometries and hydrogen-bonding pattern) modulating the specific ligand–DNA binding in
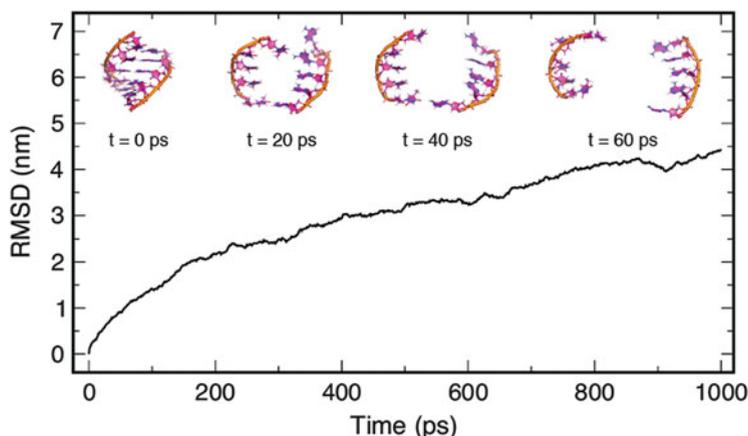
Fig. 3.2 Evolution of the structure of a short DNA duplex subjected to MD simulation in the gas phase. Starting structure is that corresponding to the canonical MD structure of the duplex in aqueous solution. Simulations are done assuming an instantaneous transfer with no water–DNA proton exchange during vaporization, i.e., DNA charge state in the simulation is that for the DNA in solution

aqueous solution are somehow conserved in the gas phase. Even more surprising, ion mobility spectrometry (IMS) experiments suggest that the general shape of nucleic acids (experimentally measured by the collision cross section [31]) in gas phase can be close to the expected values in solution, suggesting that contrary to chemical intuition nucleic acids in the gas phase remain quite compact.

The apparent paradox between experimental observation and theoretical expectations has drawn the attention of the physical chemistry community for the last decades. Several groups have tried to determine the nature of the gas-phase structure of nucleic acids using either experimental or theoretical approaches. The aim of this research effort is to develop a formalism connecting solution and gas-phase structures, as seems to be possible for proteins [32, 33]. It seems now (mid-2013) a purely academic exercise, but we cannot forget that X-Ray Free Electron Laser (XFEL) experiments [34] are expected to provide three-dimensional gas-phase structures of complex nucleic acid motifs, which are difficult to determine by conventional techniques in solution.

Below, we present an overview of our current knowledge on the structure and physicochemical properties of nucleic acids in the gas phase.

3.2   What Gas-Phase Conditions Do We Want to Model?

3.2.1   The Vaporization Process

In order to model gas-phase experiments with theoretical methods we need to understand the process followed to transfer molecules from solution to the gas phase (the vaporization process; see [33, 35], and Chap. 2). In a mild ESI-MS

experiment, the solution containing the nucleic acids is injected in a capillary whose tip is placed in front of the mass spectrometer. A difference in electrostatic potential between the capillary and the entrance of the mass spectrometer is then applied, generating an electric field in which the ions move towards the counter-electrode (contrary to proteins, nucleic acids are typically sprayed in the negative mode). This results in a jet formation and in the emission of charged droplets containing nucleic acids and solvent molecules. As the droplets move to the counter-electrode, water molecules are evaporated, which increases the charge density at the droplet surface up to the Rayleigh limit [33, 36, 37]. This charge density increase induces droplet fission and the emission of a secondary jet or smaller droplets, whose subsequent fission proceeds in a way similar to that of the parent drop. Fully dehydrated nucleic acids are obtained after a few generations of fissions.

The evaporation of solvent from the nucleic acid containing droplets is supposed to occur following the "charge residue mechanism" [33, 38, 39] for oligonucleotide strands. Accordingly, solvent migration is gradual and fast until the last solvent molecule leaves the droplet and the nucleic acids are in a pure gas-phase environment. During this migration and evaporation process part of the original nucleic acid charge moves to the solvent in the form of OH^- ions, i.e., the nucleic acids capture protons from the surrounding medium. The final total charge on the oligonucleotide should correlate then with the solvent accessible surface of the nucleic acid [27, 37–40]. Analysis of nucleic acids' ESI-MS spectra shows clear charge/mass signals that agree well with molecular surfaces similar to those found in solution [40–43]. However, as the charge increases for a given nucleic acid, its shape becomes no longer consistent with the folded state.

Mild vaporization processes imply desolvation of the sample and a significant reduction on the total negative charge on the nucleic acid to a final preferred final charge [16–30]. If the solution sample contains no significant amount of unfolded oligonucleotide and is at a suitable ionic strength (typically 100—150 mM ammonium acetate), and if the vaporization is mild enough, then major signals detected experimentally generally correspond to the preferred charge ± 1. This observation implies that the experimental setup and the structural characteristics of the nucleic acids fix within a narrow range the final charge state of the nucleic acid when it starts migrating in the mass spectrometer or in the ion mobility spectrometer. Ion mobility spectrometry is particularly important because this technique will provide information on gas-phase structure of nucleic acids.

3.2.2 Ion Mobility Spectrometry

In classical ion mobility spectrometry, dehydrated nucleic acids are forced to move in a chamber filled with inert gas at a given density. The drift velocity (v), which determines the time interval a given nucleic acid takes to arrive to the detector, depends linearly on the electric field and the intrinsic mobility of the oligonucleotide:
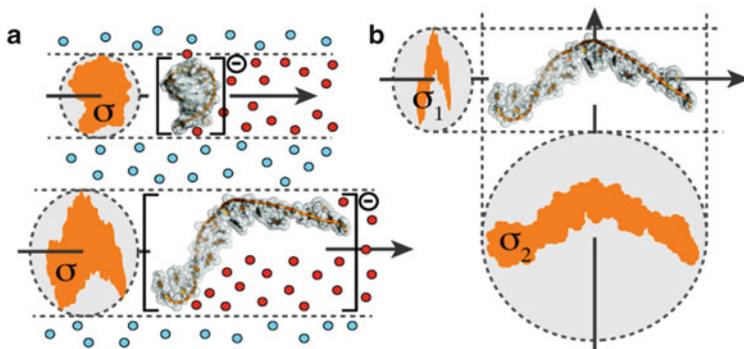
Fig. 3.3 Schematic representation of migration of a nucleic acid in a drift tube during an ESI-MS experiment: (**a**) folded and unfolded conformations display different collision cross sections (σ, represented by the irregular shape enclosed in an ellipsoid). (**b**) The collision cross section depends also on the orientation of the nucleic acid with respect to the drift vector. Other determining factors include the convexity of the nucleic acid surface. Inert gas molecules are displayed as *red spheres* when collisions between the gas molecule and the oligonucleotide are expected (leading to a reduction of nucleic acid mobility) and as *blue spheres* otherwise

$$v = KE, \tag{3.1}$$

with K being the mobility and E the external field. The oligonucleotide mobility depends on its mass and charge, but also its shape and size modulate the kinetic energy lost by collision against the inert gas particles in the drift tube (see Fig. 3.3). In the low field limit, the mobility can be determined by the Mason–Schamp equation [44]:

$$K = \frac{3}{16} N_0 \frac{q}{N^2} \left(\frac{2\pi}{\mu kT} \right)^{\frac{1}{2}} \frac{1}{\Omega}, \tag{3.2}$$

where N is the gas number density, N_0 is the gas number density at standard temperature and pressure, q is the ion charge, μ is the reduced mass ($\mu = mM/(m+M)$, where M is the mass of the gas and m is the mass of the ion), k is the Boltzmann constant, T is the gas temperature, and Ω is a standard property related to the molecular surface that is exposed to collisions with the inert gas in the drift tube [31]).

The exact thermodynamic conditions during the vaporization and the ion migration are not known exactly, hindering the comparison between experimental results and theoretical calculations (for a deep discussion on this topic, we address the reader to [33] and references therein). For example, the effective temperature of ions in the vaporization chamber is difficult to determine, since vaporization is an endothermic process that tends to reduce the effective temperature. All these considerations cannot be ignored when comparing the structural information obtained from simulations with observables derived from ESI-MS experiments,

since the degree of uncertainties does not allow for a straightforward comparison of experiments and simulations.

3.3   What Does Gas-Phase Structure Mean?

3.3.1   Introduction

Gas-phase experiments used to be quite fast, ranging from the extremely rapid femtosecond-scale XFEL experiments to the 1–10 ms timescale of a typical ESI-MS or IMS experiment [33, 45]. Furthermore, the absence of the lubricant effect of water and the strong unscreened interactions typical of the gas phase results in a very rugged conformational landscape for most biomolecules. It is then very unlikely that, except for very small and flexible oligonucleotides, the structure in solution has had time to rearrange itself into the most stable conformation in the gas phase, i.e., the true gas-phase minimum [32, 33] (Fig. 3.4). In other words, structural information derived from gas-phase experiments (assuming mild vaporization conditions) refers always to local minima close to the "solution native state," which will be in most cases a metastable conformation with lifetimes longer than the experimental measuring time. The fast nature of ESI-MS measurements allows us to obtain information of transient structures that are closer to the native structure in solution than the real gas-phase absolute energy minimum. The question is not whether or not the absolute gas-phase minimum is close or far from the solution structure (probably it is very distant), but how far the "solution-like" gas-phase minimum is from the real "solution minimum."

   In the case of proteins, a proteome-scale study [32] suggested that the gas-phase structure detected in ESI-MS experiments is in general very close to the solution conformation for most protein folds. The same study also demonstrates that the "solution-like" gas-phase structure maintains well the main structural characteristics of the fully solvated protein, allowing fold recognition programs to assign correctly the structure to the real solution conformation. It also demonstrates that aqueous MD simulations starting from the "solution-like" gas-phase minimum converged very quickly to the absolute minimum in solution. In other words, protein structure resists well vaporization conditions, and 1–10 ms after vaporization, it still maintains a conformation close to that in aqueous solution.

   Can we extrapolate these results obtained for proteins to multiply charged nucleic acids? Nucleic acids display often non-globular conformations where structural perturbations can be transmitted with great efficiency. We are certainly far from having a complete atlas of the gas-phase structure of nucleic acids, but we have both theoretical and experimental information on the impact of vaporization on some canonical nucleic acid structures (see below).
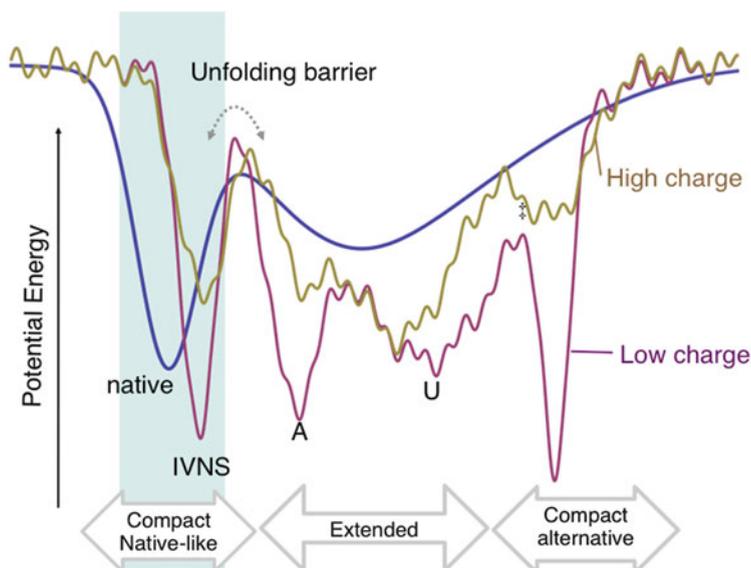
Fig. 3.4 Schematic representation of the conformational landscape of a macromolecule in solution (*blue*) and in the gas phase for (1) low charge density (in *magenta*) and (2) in a high charged density (in *green*). In the case of low charge density (*magenta*) the macromolecule can be trapped in a local minimum, quite compact and not far from the "solution native state," termed *in vacuum* native structure (IVNS). This minimum is very sharp, and accordingly partially (A), totally (U) extended or alternative compact structures are not explored in the millisecond timescale of the experiment. Under high charge conditions (*green*) the Coulomb repulsion is so large that the macromolecule escapes from the "solution native state" minimum and explores a myriad of extended conformations. Adapted with permission from Meyer et al. [33]

3.3.2 The Ionic and Charge State of Nucleic Acids

As noted above, native ESI-MS major signals correspond to a narrow number of charge states (see Fig. 3.5). In fact, at least for canonical duplex structures, the main charge state can be generally approximated from the solvent accessible surface area of the nucleic acid structure in solution [41–43, 46], and higher charge density typically signals unfolded nucleic acids.

Determining the population of charge states, i.e., the different charge substates existing within a given charge state as a consequence of different titration of the oligonucleotide, is much more complicated. Assuming (which might be in some cases a strong and risky assumption) that the reduction in total charge from solution to gas-phase values is achieved by neutralization of the titrable groups, we find that for a specific ionic state the number of different charge states is given by
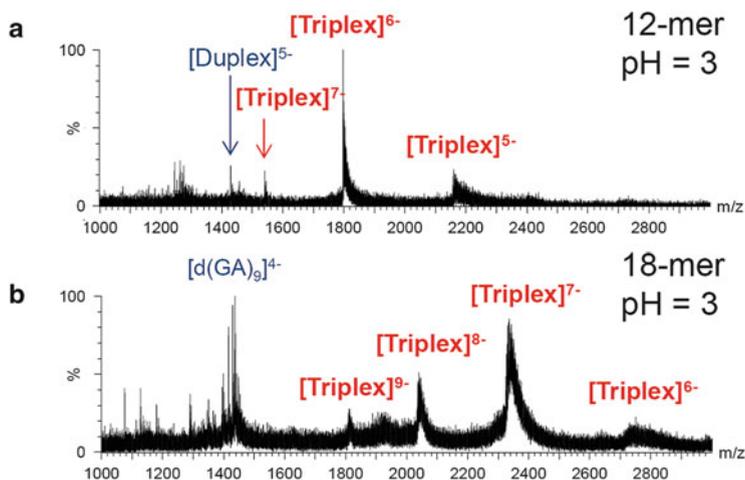
$$N_{\text{tot}} = \frac{m!}{[n!(m-n)!]}, \tag{3.3}$$

Fig. 3.5 Representative ESI-MS spectra obtained for $d(TC)_6$ $d(GA)_6$ (ratio 2:1) and $d(TC)_8$ $d(GA)_8$ (ratio 2:1) in acidic conditions that favor the formation of $d(TC)_6$-$d(GA)_6 \cdot d(TC)_6$ and $d(TC)_8$-$d(GA)_8 \cdot d(TC)_8$ 12-mer and 18-mer triplexes. Note that there is always a dominant signal (-6 for 12-mer and -7 for 18-mer) and dispersion in the ionic state detected is rather small. Data recorded by Gabelica using a Synapt G1 HDMS traveling wave ion mobility mass spectrometer. Figure taken with permission from [41]

where N_{tot} is the number of possible charge states, m is the number of titrable groups and n is the number of protons required to reach the final charge state. Thus, for the 12-mer duplex signal shown in Fig. 3.5 (in blue), assuming that only phosphates are titrable, we have 52,668 different ways to distribute the charge, and even for the single-stranded $d(GA)_9$ oligonucleotide, we can predict up to 2,380 different charge substates for the most probable charge state. These combinatorial figures explode for longer oligonucleotides; for example, for the 18-mer triplex of charge -6 shown in Fig. 3.6, we have in the order 10^{12} different charge substates (even when assuming that the cytosines of the Hoogsteen strand in the triplex are protonated).

From ESI-MS experiments alone it is unclear which distribution of charge substates coexists for a given charge state. We can assume that solvent and buffers present in solution are the main sources of protons for the partial neutralization of phosphates and that such a titration happens during droplet evaporation, mostly at easily accessible phosphates. A reasonable assumption is that the main charge distribution just after the vaporization is the one that minimizes the Coulomb repulsion in the solution structure of the nucleic acid [41–43, 46]. Such charge distribution can be easily determined by using standard optimization routines (e.g., a Monte Carlo procedure). However, it is unclear: (1) what is the distribution of charge along the sequence in cases of (quasi) degeneracy and (2) whether or not the charge distribution evolves with time. With respect to the first point, it is likely that several charge substates coexist for a given charge state, and accordingly, attempts to simulate nucleic acids in the gas phase should account for such diversity [41, 47].

Fig. 3.6 (**a**) Example of interactions between neutral and protonated phosphates sampled during MD simulations of gas-phase DNA. (**b**) Energy profile for the distance dependence of the neutral anionic phosphate contact, as determined from DFT level calculations. Note the very large magnitude of the unscreened (phosphate–phosphate)$^-$ interaction
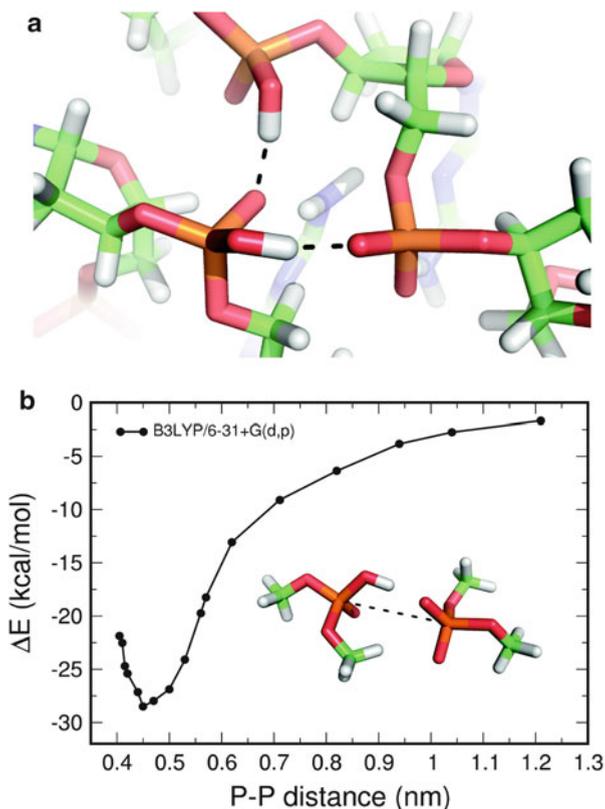
The problem of the time dependency of the distribution of charge substate is very complex, since it is related to the relationship between charge localization and the molecule geometry. In principle, the distribution of charge substates obtained at the end of the desolvation (loss of water and ammonia molecules) might change completely when the nucleic acids reach the detector, and this will likely impact nucleic acid structure.

In a partially anionic DNA fragment, the negatively charged phosphates are likely to be the strongest base accessible to proton capture, while the neutral phosphates are in turn the strongest acids. In fact, in MD simulations of nucleic acids in the gas phase, we have systematically found that the most prevalent Coulomb interaction is that between neutral and anionic phosphates (see and example in Fig. 3.6a), a result that is not surprising considering the strength of these interactions in the gas phase as determined from quantum mechanical calculations (see Fig. 3.6b).

Based on the results shown in Fig. 3.6 and the discussion above, we can predict that most of the variations in charge distribution will occur via proton transfer between neutral and protonated phosphates. As shown in Fig. 3.7 the kinetics of this process (as determined from the kinetic barrier of proton transfer) are extremely dependent on the phosphate–phosphate distance. For short distances, we can reach
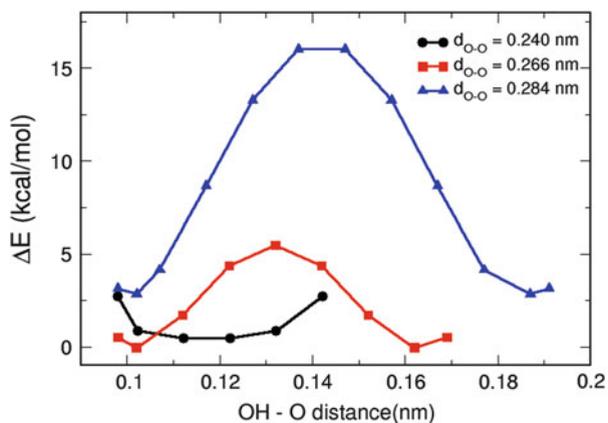
Fig. 3.7 Energy profiles for proton transfer between two phosphate groups at fixed donor–acceptor distances. The lowest energy value of the three profiles was used as reference to rescale the curves. The barrier for proton transfer is sizeable at an interatomic distance of 0.266 nm (around 5 kcal/mol, red curve), corresponding to a separation where the interaction energy of the two fragments reaches its global minima (at 0.45 nm separation between phosphorous atoms; see Fig. 3.6). Larger donor–acceptor separations result in even larger transfer barriers (*blue curve*). Only at very short donor–acceptor distances the proton transfer is barrierless (*black curve*). The calculations were carried out at the DFT level of calculation (B3LYP/6-31 + G(d,p))

the paradigm of the "low barrier hydrogen bond" [48], which would imply ultrafast (diffusion controlled) proton transfer between the two phosphates. However, sizeable energy barriers appear for the most accessible phosphate–phosphate distances. Furthermore, once proton transfer takes place, a pseudo-symmetrical situation is reached, where reverse proton transfer is expected to be often faster than the dissociation of proton donor and acceptors—rearrangement that might yield to major conformational changes in the DNA (see Figs. 3.6 and 3.7). Thus, proton transfer should have reduced structural impact.

3.3.3   The Theoretical Representation of Nucleic Acids in the Gas Phase

High-level quantum mechanical (QM) calculations allow a full characterization of the constituent elements of nucleic acids. State-of-the art Coupled-Cluster (with single, double, and triple excitations) calculation using complete basis set (CCSD (T)/CBS) has characterized the structure and interaction of, for example, nucleobases [5, 7] in the gas phase (mostly hydrogen bond and stacking) with a level of accuracy similar or even better to that reachable by experiments. Extensions to capture higher order excitations and relativistic effects are expected to reduce the error in key experimental observables as the interaction energy to a negligible 1.5 % [49].

By using high-level QM calculations theoreticians have been able to find fundamental differences between the properties of nucleic acid components in

solution and in the gas phase, originated from the tendency of solvents like water to favor electron and nuclear arrangements with high polarity compared to the gas-phase situation [8]. For example, in the gas-phase hydrogen bonds are reinforced, favoring compact arrangements [8, 50, 51]. Less intuitively, but very important, desolvation can produce a change in the tautomeric preferences of nucleobases [50, 52–55], which results in a change in the hydrogen-bonding patterns. For example, cytosine, which adopts a keto-amino form in aqueous solution, moves to an enol-amino form when transferred to the gas phase [53], which will favor pairing with adenine instead of guanosine, highlighting a potential mechanism for spontaneous mutation.

QM calculations are then extremely powerful to define the real nature of nucleic acid constituents in the gas phase. They are accurate, but also very costly, and cannot be applied in the study of real (biologically relevant) nucleic acids, which are expected to adopt a myriad of conformations and that need to be explored from a dynamic rather than static point of view. Thus, most theoretical studies on nucleic acids have been dealt with using classical mechanics, particularly molecular dynamics (MD), a technique that with current computers can be used to study dynamics of medium-sized oligonucleotides in timescales close to those of the experimental ones.

When using classical MD simulations, nucleic acids interactions are reproduced by means of a simple empirical force field that was typically parameterized with a combination of gas phase and solution data. The movements of the system are then determined by integrating corresponding Newton's equations of motion, in either the constant energy or constant temperature ensemble. Even the need to use small time steps (in the order of femtoseconds) for numerical integration of Newton's equations of motion we can now simulate the evolution of nucleic acids in the gas phase for multi-microsecond periods, considering different charge states and substates, obtaining structural information that can be directly compared with experimental one (see below).

We cannot forget, however, the dramatic simplifications assumed in these calculations which hampers their reliability: (1) we assume that force fields calibrated to study nucleic acids in solution are valid also in the gas phase; (2) we explicitly assume that no intramolecular proton transfer occurs during the simulation; and (3) we need to assume an ensemble (constant energy or constant temperature) for our calculations, which might not fit exactly with the real nature of the experiments.

3.4 Gas-Phase Structures of Different DNA Families

3.4.1 Single-Stranded DNA

Single-stranded DNAs in the gas phase have been studied experimentally by several authors [56, 57], but there are not many theoretical works, due probably to the difficulty in determining a reasonable starting conformation for the oligonucleotide

before vaporization. Available data (references above) show a dramatic dependence of the shape of the oligonucleotide on the protonation state, which is detected as sharp variation in CCS. A basic modeling exercise suggests that at low charge states single-stranded oligos adopt a globular like conformation, while at high charge state extended (but not completely extended) conformations are populated. A reasonable selection of the phosphates to be protonated [56] to reach gas-phase total charge seems crucial for reproducing experimental shape observables. Overall, available data suggest that, under mild electrospray conditions, single-stranded conformations in the gas phase are diverse and more compact than expected from the need to reduce phosphate–phosphate repulsions.

3.4.2   The Duplex DNA

Under mild ESI-MS conditions the DNA duplex is maintained in the gas phase [18, 20, 21, 28, 29], with the two strands bound together. MD simulations (see Fig. 3.8) [42, 58] suggest that DNA structure in the gas phase displays a more extended conformation than in aqueous solution, giving structures that resemble the elongated C-form of DNA [3] or mechanically stressed DNAs [59, 60]. Clearly, reduction of Coulomb repulsion in anionic phosphates is one of the main factors behind the gas-phase-driven distortion of DNA duplex.

Several MD simulations of various DNA sequences at different temperatures and neutralized using different schemes provide similar pictures of duplex DNA in the gas phase: the two strands are firmly joined by a multitude of electrostatic and van der Waals contacts (see Fig. 3.8). Despite the evident large distortions, the gas-phase DNA maintains a clear memory of its original conformation in solution, with an overall helical geometry, sugar conformation mostly in the South-East region, and around 84 % of nucleosides in the *anti* region [3]. The gas-phase structure shows an extensive pattern of hydrogen bonding (30–50 % at 448 K and 60–90 % of those found in water at room temperature), but while in Watson–Crick pattern represents nearly 100 % of the detected hydrogen bonds in aqueous solution, it only accounts for a fraction of those hydrogen bonds detected in the gas phase, mostly localized on C · G pairs. Stacking is extremely prevalent in the gas phase and is not limited to that found in canonical duplexes in solution, since many alternative stacked clusters, including T-shape and triads, are detected. Finally, analysis of individual trajectories reveals that the duplex DNA also maintains a good memory of its flexibility in solution, and contrary to our original expectations, the gas-phase DNA is not a random coil, but it is found as a limited ensemble of structures sharing common conformational characteristics and a pattern of flexibility similar to that found in solution.

One of the most visible structural changes in the DNA is the corruption of groove geometry (Fig. 3.8), emerging mostly from the formation of phosphate–phosphate hydrogen bonds and from the change in the pattern of nucleotide interactions. It can be then expected that small ligands binding DNA through the grooves will not resist vaporization and will dissociate from the duplex when the groove geometry is

Fig. 3.8 Schematic representation of Dickerson dodecamer after 1 µs of MD simulation in aqueous solution (*left*) and in the gas phase using simulation conditions as those explained in [42]
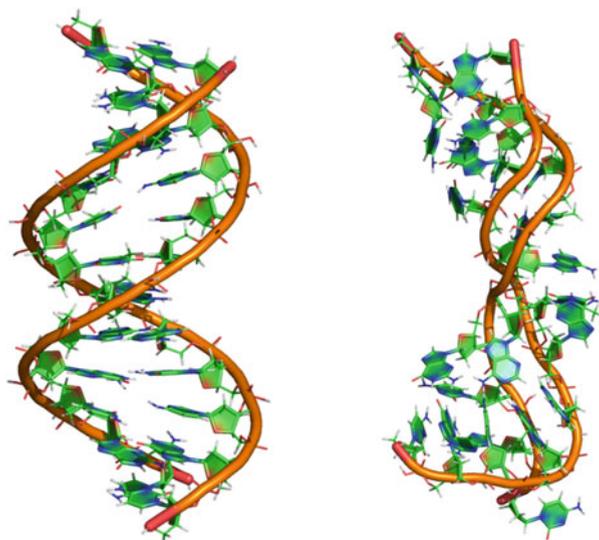
altered. However, as detailed further in Chap. 9, not only groove-binder complexes resist vaporization, but also the relative ratio of bound/unbound ESI-MS signals correlates with the binding free energy of the ligand–DNA complex in solution. Gas-phase MD simulations [43] show that the high flexibility of DNA allows it to wrap fast and tightly around the groove binder, capturing it and hindering its diffusion.

3.4.3 The Triplex DNA

ESI-MS experiments were the first to show that under mild conditions triplexes do not dissociate in the gas phase [23], keeping structural signatures that allow it to make specific non-covalent interactions [22, 25, 27]. More recent accurate CCS measurements on GA triplexes strongly suggest that triplex DNA maintains a compact structure in the gas phase [41]. In the same communication, very extended (sub-millisecond) MD simulations allowed us to characterize the nature of triplex in the gas phase [41]. The excellent agreement between experimental and MD-derived CCS values confirms the quality of the theoretical models, which were quite robust to the exact details of the experimental setup.

The structure of triplex DNA in the gas phase is rigid, very far from a random coil model assumed for an unfolded oligonucleotide. Structural drifts in the 1–30 µs range are generally small, indicating that conformational rearrangements in the DNA are small and slow in the gas phase, confirming our suggestions that ESI-MS experiments are detecting a long-lived metastable conformation not far from the solution structure. Even after 60 µs of MD simulation the solution triple helix motif can be recognized (Fig. 3.9). Interestingly, the overall structure shows a degree of
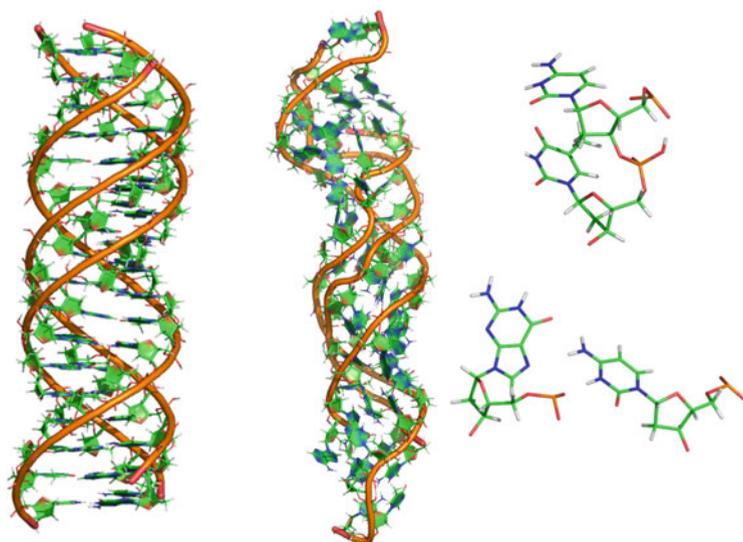
Fig. 3.9 Schematic representation of 18-mer poly(GA) triplex after simulation in aqueous solution (*left*) and gas phase (*middle*) under conditions discussed in [41]; see also text. Structural details at the *right* illustrate some specific interactions detected (*top*: a (phosphate–phosphate)$^-$ contact and *bottom*: a phosphate(−)–cytosine(+) contact)

compactness not too different to that found in aqueous solution (for example, theoretical CCS for aqueous solution: (14.3 ± 0.5) nm^2, theoretical CCS in the gas phase: (13.3 ± 0.4) nm^2, experimental CCS in the gas phase: (13.3 ± 0.4) nm^2; gas-phase values computed or determined from the MD structure considering a −7 charge). Desolvation disrupts many native interactions, but even at the end of our longest trajectories there is a very significant amount of native contacts (for example, around 31 % (Watson–Crick) and 24 % (Hoogsteen) of hydrogen bonds are preserved at the end of 18-mer gas-phase simulations). Many nonnative contacts are formed to stabilize and rigidify the triplex in the gas phase. Thus, for the studied poly(GA) triplex, in addition to the usual (phosphate–phosphate)$^-$ contacts we also detect a significant number of strong ionic contacts between protonated (Hoogsteen) cytosines and anionic phosphates, which surely help to make structure more rigid than those found for duplex DNA in the gas phase.

3.4.4 The G-Quadruplex DNA

Quadruplexes are compact structures mostly formed in poly-guanine tracks, where four guanine stand in a plane forming Watson–Crick and Hoogsteen hydrogen bonds [1–4]. G-tetrads are also stabilized by strong interactions with small monovalent ions, which are placed in the central channel coordinating the carbonyl oxygens that are pointing to the interior [1–4, 12, 61–64]. Experimental evidences

Fig. 3.10 Average structures
of a short (two steps)
G-quadruplex in aqueous
solution (*top*) and in the gas
phase (*bottom*). Ion in the
central channel is shown as a
sphere (*blue* Na^+; *green* K^+).
Gas-phase structures were
considered with a total charge
of −3. Figure adapted from
[46] with author's permission
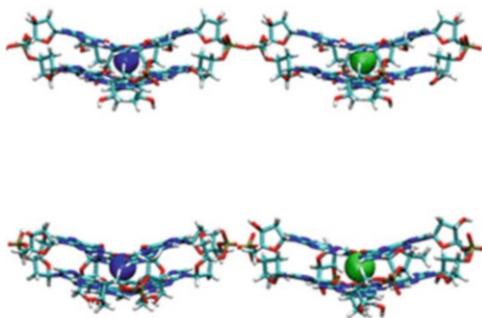
collected by different groups [16, 17, 19, 23, 30] strongly suggest a very significant maintenance of the quadruplex DNA structure in the phase. For example, Vairamani and Groos [30] found already 10 years ago mass spectrometry signals corresponding to the thrombin binding aptamer (a short antiparallel quadruplex) tightly bound to monovalent ions in the gas phase, suggesting a certain preservation of the G-quadruplex structure. De Pauw and coworkers [23] studied different G-DNAs by mass spectroscopy finding signals that agree with those of a stable quadruplex. In the same work [23] the authors also found that non-covalent complexes of quadruplexes with porphyrins resist vaporization, suggesting again significant structural conservation. Bowers and coworkers [19] measured the CCS of G-DNA quadruplexes in the gas phase finding values that agree with a compact structure in the gas phase, consistent with those expected for the solution structure. Using IR spectroscopy, Gabelica et al. [65] found carbonyl stretching signals which could be explained only if these groups remained hydrogen bonded in the gas phase. More recently, gas-phase UV spectral data have supported the maintenance of clear structural signatures for DNA quadruplexes in the gas phase [66] and direct evidences of the role of cations in stabilizing quadruplex structure in the gas phase [67] (see also Chap. 5). Altogether, the experimental evidence clearly suggests that quadruplex DNA can maintain an unusually large number of structural signatures of the structure in solution.

Extensive MD simulations of different G-DNA quadruplexes confirmed that the G-quadruplex core is extremely resistant to desolvation (Fig. 3.10) [46]. In fact, despite a slight overall compaction, gas-phase structures cannot be easily distinguished from solution conformations. Particularly, canonical hydrogen bonds and stacking are fully preserved, and contrary to the situation found for duplex and triplex DNAs the groove geometry of G-quadruplexes is not much altered in the gas phase. ESI-MS signals indicate that cations are tightly bound to G-quadruplex nucleic acids in the gas phase. Interestingly, MD simulations of G-quadruplexes at low charge states suggest (in agreement with experimental data [46, 67] that the removal of ions from the central channel leads to compact structures, which then deviate from the experimental conformation in solution. No MD simulations have been done, to our knowledge, for G-quadruplexes in highly charged state, but it is expected that in this case, unfolded extended conformations will be populated [46, 67].

3.5   Conclusion and Outlook

Nucleic acids are highly charged polymers with well-defined structures in aqueous solution. Most of these well-defined structural motifs are surprisingly retained in the gas phase for significant periods of time (at least in the millisecond range). It is unlikely that the experimentally detected gas-phase structures correspond to the true minima in vacuo; these are probably trapped metastable conformations that are not very far from the solution structure. Experimental data and MD simulation results for different DNA motifs suggest that the degree of structural distortion due to vaporization decreases from duplex to triplex and finally to DNA G-quadruplexes. Ionic contacts involving neutral and charged phosphates are very prevalent and probably crucial to stabilize compact structures in the gas phase. However, there are a myriad of other favorable interactions contributing to nucleic acid stability in the gas phase. Such interactions were already present in solution, but their topology and intensity have drastically changed in vacuum: stacking interactions and noncanonical hydrogen bonds (for example, phosphate–phosphate contact, nucleobase–sugar, etc. . .) are largely increased, while canonical hydrogen bonds become decreased (see for example Figs. 3.8 and 3.9).

All the information presented above is centered in DNA, and in fact there is not much information on the structure of RNA in the gas phase. How much of the data obtained for DNA is transferable for RNA is unclear and probably depends on the nature of the RNA. We can speculate that canonical RNA helices will behave similar to the corresponding DNA ones, while perhaps large RNAs displaying complex 3D arrangement can show a behavior in the gas phase more similar to that of proteins. Clearly, future systematic work is required to answer these important questions.

Experimental and theoretical methods have worked in perfect synergy for nearly one decade, helping us to better understand the nature of nucleic acid structure in the gas phase, its relationship with solution conformation, and the underlying physicochemical reasons that prevent nucleic acid structures to be completely disrupted when transferred to a high vacuum. As ever more powerful technologies become a reality, such as XFEL, new exciting discoveries are likely to be made, which will open novel paths for the young and emerging field of "structural biology in the gas phase."

3.6   Summary of Key Concepts

General Concepts
- Electrospray ionization-mass spectrometry (ESI-MS) experiments are a powerful technique to characterize biomolecules.
- ESI-MS experiments take advantage of the charged nature of biopolymers, such as proteins and nucleic acids, to obtain information on their geometry, shape, and oligomeric state.

- Both computational and experimental studies have shown that protein structure resists vaporization conditions and that even after significant amounts of time (millisecond scale) after complete desolvation they still maintain a conformation close to that found in aqueous solution.
- The stability of solution structures of nucleic acid in the gas phase depends drastically on their oligomeric state.

Concepts Specific to Nucleic Acids
- Nucleic acids maintain the memory of their structure in solution when transferred into the gas phase during a mass spectroscopy experiment at mild electrospray vaporization conditions.
- The effect of phosphate groups' protonation during the evaporation process significantly decreases the net charge of the nucleic acids. This reduction in total charge lowers the Coulomb repulsion and it is crucial for the stability of DNA in vacuum.
- Computational studies on nucleic acids in the gas phase show that the structure of most nucleic acids appears in general distorted, but the most important interactions, such as hydrogen bonds and stacking interactions, are partially preserved, and the nucleic acids maintain a clear helicity.
- The order of structural distortion due to vaporization is duplex > triplex > G-quadruplex.

References

1. Blackburn GM, Gait MJ (1990) Nucleic acids in chemistry and biology. IRL Press, Oxford
2. Bloomfield VA, Crothers DM, Tinoco I (2000) Nucleic acids. Structure, properties and functions. University Science, Sausalito, CA
3. Saenger W (1984) Principles of nucleic acid structure. Springer, New York, NY
4. Sinden RR (1994) DNA structure and function. Academic, San Diego, CA
5. Hobza P, Sponer J (2002) Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. J Am Chem Soc 124(39):11802–11808
6. Pérez A, Šponer J, Jurecka P, Hobza P, Luque FJ, Orozco M (2005) Are the hydrogen bonds of RNA (A· U) stronger than those of DNA (A· T)? A quantum mechanics study. Chem Eur J 11 (17):5062–5066
7. Hobza P, Kabelac M, Sponer J, Mejzlik P, Vondrasek J (1997) Performance of empirical potentials (AMBER, CFF95, CVFF, CHARMM, OPLS, POLTEV), semiempirical quantum chemical methods (AM1, MNDO/M, PM3), and ab initio Hartree–Fock method for interaction of DNA bases: comparison with nonempirical beyond Hartree–Fock results. J Comput Chem 18(9):1136–1150
8. Orozco M, Luque FJ (2000) Theoretical methods for the description of the solvent effect in biomolecular systems. Chem Rev 100(11):4187–4226
9. Huey R, Mohr SC (1981) Condensed states of nucleic acids. III. psi(+) and psi(-) conformational transitions of DNA induced by ethanol and salt. Biopolymers 20(12):2533–2552
10. Pohl FM, Jovin TM (1972) Salt-induced co-operative conformational change of a synthetic DNA: equilibrium and kinetic studies with poly (dG-dC). J Mol Biol 67(3):375–396

11. Bernues J et al (1990) DNA-sequence and metal-ion specificity of the formation of *H-DNA. Nucleic Acids Res 18(14):4067–4073
12. Wang Y, Patel DJ (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. Structure 1(4):263–282
13. Williamson JR, Raghuraman MK, Cech TR (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. Cell 59(5):871–880
14. Turner DH (2000) Structure, properties and functions. In: Bloomfield V, Crothers D, Tinoco I (eds) Structure, properties and functions, in nucleic acids. University Science Books, Sausalito, CA, pp 308–310
15. Miyoshi D, Nakamura K, Tateishi-Karimata H, Ohmichi T, Sugimoto N (2009) Hydration of Watson-Crick base pairs and dehydration of Hoogsteen base pairs inducing structural polymorphism under molecular crowding conditions. J Am Chem Soc 131:3522–3531
16. Gabelica V et al (2007) Base-dependent electron photodetachment from negatively charged DNA strands upon 260-nm laser irradiation. J Am Chem Soc 129(15):4706–4713
17. Gabelica V, Rosu F, Witt M, Baykut G, De Pauw E (2005) Fast gas-phase hydrogen/deuterium exchange observed for a DNA G-quadruplex. Rapid Commun Mass Spectrom 19(2):201–208
18. Gale DC, Smith RD (1995) Characterization of noncovalent complexes formed between minor groove binding molecules and duplex DNA by electrospray ionization-mass spectrometry. J Am Soc Mass Spectrom 6(12):1154
19. Gidden J, Baker ES, Ferzoco AL, Bowers MT (2005) Structural motifs of DNA complexes in the gas phase. Int J Mass Spectrom 240:183–193
20. Hofstadler SA, Griffey RH (2001) Analysis of noncovalent complexes of DNA and RNA by mass spectrometry. Chem Rev 101(2):377–390
21. Reyzer ML et al (2001) Evaluation of complexation of metal-mediated DNA-binding drugs to oligonucleotides via electrospray ionization mass spectrometry. Nucleic Acids Res 29(21): E103-3
22. Rosu F, De Pauw E, Gabelica V (2008) Electrospray mass spectrometry to study drug-nucleic acids interactions. Biochimie 90(7):1074–1087
23. Rosu F et al (2002) Triplex and quadruplex DNA structures studied by electrospray mass spectrometry. Rapid Commun Mass Spectrom 16(18):1729–1736
24. Rosu F et al (2010) Tetramolecular G-quadruplex formation pathways studied by electrospray mass spectrometry. Nucleic Acids Res 38(15):5217–5225
25. Rosu F et al (2007) Ligand binding mode to duplex and triplex DNA assessed by combining electrospray tandem mass spectrometry and molecular modeling. J Am Soc Mass Spectrom 18 (6):1052–1062
26. Schnier PD et al (1998) Activation energies for dissociation of double strand oligonucleotide anions: evidence for Watson-Crick base pairing in vacuo. J Am Chem Soc 120(37):9605–9613
27. Wan C et al (2008) Studies of the intermolecular DNA triplexes of C + .GC and T.AT triplets by electrospray ionization Fourier-transform ion cyclotron resonance mass spectrometry. J Mass Spectrom 43(2):164–172
28. Wan KX, Gross ML, Shibue T (2000) Gas-phase stability of double-stranded oligodeoxynucleotides and their noncovalent complexes with DNA-binding drugs as revealed by collisional activation in an ion trap. J Am Soc Mass Spectrom 11(5):450–457
29. Gabelica V, De Pauw E, Rosu F (1999) Interaction between antitumor drugs and a double-stranded oligonucleotide studied by electrospray ionization mass spectrometry. J Mass Spectrom 34(12):1328–1337
30. Vairamani M, Gross ML (2003) G-quadruplex formation of thrombin-binding aptamer detected by electrospray ionization mass spectrometry. J Am Chem Soc 125(1):42–43
31. Wyttenbach T, Bleiholder C, Bowers MT (2013) Factors contributing to the collision cross section of polyatomic ions in the kilodalton to gigadalton range: application to ion mobility measurements. Anal Chem 85:2191–2199
32. Meyer T, de la Cruz X, Orozco M (2009) An atomistic view to the gas phase proteome. Structure 17(1):88–95

33. Meyer T, Gabelica V, Grubmüller H, Orozco M (2012) Proteins in the gas phase. WIRES Comput Mol Sci
34. Neutze R, Huldt G, Hajdu J, van der Spoel D (2004) Potential impact of an X-ray free electron laser on structural biology. Radiat Phys Chem 71:905–916
35. López AT, T, Vilaseca M, Giralt E (2013) New J Chem 37: 1283–1289
36. De la Mora F (2000) J Anal Chim Acta (406):93–104
37. Rayleigh L (1882) Philos Mag: 184–186
38. Fenn JB (1993) Ion formation from charged droplets: roles of geometry, energy, and time. J Am Soc Mass Spectrom 4:524–535
39. Hautreux M, Hue N, Du Fou de Kerdaniel A, Zahir A, Malec V, Laprévote O (2004) Under non-denaturing solvent conditions, the mean charge state of a multiply charged protein ion formed by electrospray is linearly correlated with the macromolecular surface. Int J Mass Spectrom 231:131–137
40. Felitsyn N, Kitova EN, Klassen JS (2002) Thermal dissociation of the protein homodimer ecotin in the gas phase. J Am Soc Mass Spectrom 13(12):1432–1442
41. Arcella A et al (2012) Structure of triplex DNA in the gas phase. J Am Chem Soc 134 (15):6596–6606
42. Rueda M et al (2003) The structure and dynamics of DNA in the gas phase. J Am Chem Soc 125(26):8007–8014
43. Rueda M, Luque FJ, Orozco M (2005) Nature of minor-groove binders-DNA complexes in the gas phase. J Am Chem Soc 127(33):11690–11698
44. McDaniel EW, Mason EA (1988) Transport properties of ions in gases. Wiley, New York, NY
45. Wang B, Valentine S, Plasencia M, Raghuraman S, Zhang X (2010) Artificial neural networks for the prediction of peptide drift time in ion mobility mass spectrometry. BMC Bioinforma 11:182
46. Rueda M, Luque FJ, Orozco M (2006) G-quadruplexes can maintain their structure in the gas phase. J Am Chem Soc 128(11):3608–3619
47. Madsen JA, Brodbelt JS (2010) Asymmetric charge partitioning upon dissociation of DNA duplexes. J Am Soc Mass Spectrom 21:1144–1150
48. Cleland WW, Kreevoy MM (1994) Low-barrier hydrogen bonds and enzymic catalysis. Science 264(5167):1887–1890
49. Rezác J, Hobza P (2013) Describing noncovalent interactions beyond the common approximations: how accurate is then "gold stgandard", CCSD(T), at the Complete basis set limit. J Chem Theory Comput 9:2151–2155
50. Orozco M, Cubero E, Barril X, Colominas C, Luque FJ (1999) Nucleic acid bases in solution. In: Leszczynski J (ed) Computational molecular biology, Theoretical computational chemistry. Elsevier Science, Amsterdam, pp 119–166
51. Colominas C, Teixidó J, Cemelí JM, Luque FJ, Orozco M (1998) J Phys Chem B 102:2269–2276
52. Hernández B, Luque FJ, Orozco M (1996) Tautomerism of xanthine oxidase subtrates hypoxathine and allopurinol. J Org Chem 61:5964–5971
53. Colominas C, Luque FJ, Orozco M (1996) Tautomerism and protonation of guanine and cytosine. Implications in the formation of triplex DNA. J Am Chem Soc 118:6811–6821
54. Blas JR, Luque FJ, Orozco M (2004) Unique tautomeric properties of isoguanine. J Am Chem Soc 126:154–164
55. Orozco M, Canela EI, Mallol J, Lluis C, Franco R (1990) Ab initio study of the protonation and the tautomerism of the 7-aminopyrazolopyrimidine molecule. J Org Chem 55:753–756
56. Hoaglund CS, Liu Y, Ellington AD, Pagel M, Clemmer DE (1997) Gas-phase DNA: oligothymidine ion conformers. J Am Chem Soc 119:9051–9052
57. Gidden J, Bowers MT (2002) Gas-phase conformational and energetic properties of deprotonated dinucleotides. Eur Phys J 20:409–419
58. Gidden J et al (2004) Duplex formation and the onset of helicity in poly d(CG)n oligonucleotides in a solvent-free environment. J Am Chem Soc 126(46):15132–15140

59. Allemand JF et al (1998) Stretched and overwound DNA forms a Pauling-like structure with exposed bases. Proc Natl Acad Sci U S A 95(24):14152–14157
60. Kosikov MK, Gorin AA, Zhurkin VB, Olson WK (1999) DNA stretching and compression: large-scale simulations of double helical structures. J Mol Biol 289:1301
61. Laughlan G et al (1994) The high-resolution crystal structure of a parallel-stranded guanine tetraplex. Science 265(5171):520–524
62. Parkinson GN, Lee MP, Neidle S (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 417(6891):876–880
63. Phillips K et al (1997) The crystal structure of a parallel-stranded guanine tetraplex at 0.95 A resolution. J Mol Biol 273(1):171–182
64. Sen D, Gilbert W (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. Nature 334(6180):364–366
65. Gabelica V et al (2008) Infrared signature of DNA G-quadruplexes in the gas phase. J Am Chem Soc 130(6):1810–1811
66. Rosu F et al (2012) UV spectroscopy of DNA duplex and quadruplex structures in the gas phase. J Phys Chem A 116(22):5383–5391
67. Balthasart F, Plavec J, Gabelica V (2013) Ammonium ion binding to DNA G-quadruplexes: do electrospray mass spectra faithfully reflect the solution-phase species? J Am Soc Mass Spectrom 24(1):1–8

Interactions Between Nucleic Acid Ions and Electrons and Photons

4

Steen Brøndsted Nielsen

Abstract

This chapter deals with nucleic acid ions and their interactions with electrons and photons in the gas phase based on the many different experiments that have been performed relating to this topic within the last 10 years. The fragmentation caused by electron attachment to anions is discussed, and the role of hydration is touched upon. Photoelectron spectroscopy has established the electron binding energies of mononucleotide anions, dinucleotides and larger strands. These are significantly lower than the thresholds for electron-induced electron detachment from anions. Thresholds were measured from electron scattering experiments and product ion masses from mass spectrometry. The site of electron removal is either the base or the phosphate group, and it is likely different for photodetachment and electron detachment. Work has not been limited to anions only, but cations have also been studied. Neutral reionisation of protonated nucleobases has shed light on the lifetime of the neutral intermediate species, which was found to be significantly different to that of the temporary nucleobase anion formed in collisional electron transfer to nucleotide anions. Dissociative recombination experiments involving oligonucleotide monocations have demonstrated that there are certain electron kinetic energies where the cross section for the formation of neutral species is high (resonances), and in closely related electron-capture dissociation experiments on multiply charged cations, the actual fragmentation channels were obtained. Both for oligonucleotide anions and cations, formation of radicals by loss and capture of electrons, respectively, largely governs the dissociation patterns. This is of high relevance for sequencing. Finally, gas-phase absorption spectroscopy has revealed differences in absorption between mononucleotides, single strands, double strands and quadruplexes, which is related to the electronic coupling between two or more bases.

S. Brøndsted Nielsen (✉)
Department of Physics and Astronomy, Aarhus University, 8000 Aarhus C, Denmark
e-mail: sbn@phys.au.dk

Abbreviations

DEA	Dissociative electron attachment
DR	Dissociative recombination
DSB	Double strand break
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EPD	Electron photodetachment dissociation
ETD	Electron transfer dissociation
FTICR	Fourier transform ion cyclotron resonance
HOMO	Highest occupied molecular orbital
IC	Internal conversion
SSB	Single strand break
UV	Ultraviolet
VDE	Vertical detachment energy

4.1   Introduction

Ultraviolet (UV) light or ionising radiation can be hazardous for living organisms, either directly or indirectly. Nucleic acids strongly absorb UV light in the UVC range of the solar spectrum (<290 nm) with a maximum at 260 nm, the excitation being a direct $\pi\pi^*$ transition in the base [1–3]. However, they are surprisingly photostable except for the formation of cyclobutane pyrimidine dimers [4–6] that can lead to cancer if not enzymatically repaired. Photoexcited single bases rapidly convert electronic energy into harmless heat which may have been important for their evolutionary selection as the carriers of genetic information [7–9]. Photodissociation of isolated mononucleotide ions in vacuo occurs on the microsecond timescale [10] and can therefore not compete with vibrational cooling in water (few picoseconds) [7]. The situation is, however, more complicated and less understood for strands due to electronic coupling between two or more bases in the excited state [7–9, 11–13]. Indeed, the quantum dynamics for photoexcited strands is up for much current debate, with the role of charge-transfer states and their lifetimes being particularly unclear. It is an open question whether DNA affords self-protection against UV light.

Fig. 4.1 Quantum yields per incident electron for the induction of double strand breaks (DSBs) (**a**), single strand breaks (SSBs) (**b**) and the loss of DNA supercoiled form (**c**) in DNA solids by low-energy electrons versus incident electron energy. Taken from [16] [Copyright permission: a prize is asked for]
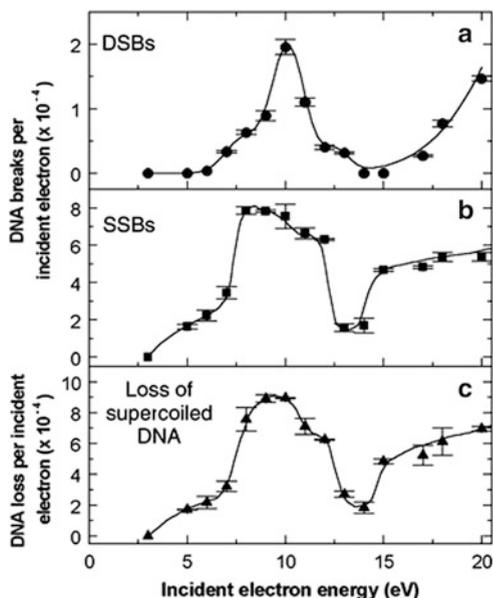
Another cause of radiation damage is the formation of secondary species by ionisation of, for example, water. Thus, ionising radiation (β-, x- or γ-rays) produces free secondary electrons in living cells with energies between ~1 eV and 20 eV [14, 15] that are able to induce genotoxic damage such as single and double DNA strand breaks. This was first demonstrated by Huels, Sanche and co-workers [16] from exposing plasmid DNA on a surface held under ultrahigh vacuum to an electron beam of variable and well-defined energy (Fig. 4.1). The DNA samples were relatively dry and contained only structural water molecules, and the phosphate groups likely had counter ions or protons closely bound to them (overall the samples were neutral). At energies well below the ionisation threshold, it is evident that the electrons induced substantial damage to the DNA caused by rapid decays of transient molecular resonances localised on the DNA's basic components.

In other work, Sanche and co-workers [17] found that even lower-energy electrons produced single strand breaks (Fig. 4.2) with a sharp peak at 0.8 eV and a broader feature centred at 2.2 eV. The shape of the curve is very similar to those obtained from dissociative electron attachment (DEA) cross section measurements on nucleobases in vacuo reported by several researchers [18–24]. Radiation damage of DNA and RNA seems therefore to originate from attachment of low-energy electrons to nucleobases. The formed radical anions induce chemical reactions that lead to strand breaks. In agreement with this, Simons and co-workers [25–29] concluded based on extensive theoretical calculations that the most likely pathway is electron attachment to base π* orbitals followed by sugar–phosphate C–O σ bond scission as this dissociation process occurs with the lowest barrier height (ca. 0.5

Fig. 4.2 Quantum yield of
DNA SSBs and double-strand
breaks as a function of
incident electron energy. The
inset shows the percentage of
SSBs versus time for 0.6 eV
electrons. Taken from
[17]. Copyright (2004) by
The American Physical
Society
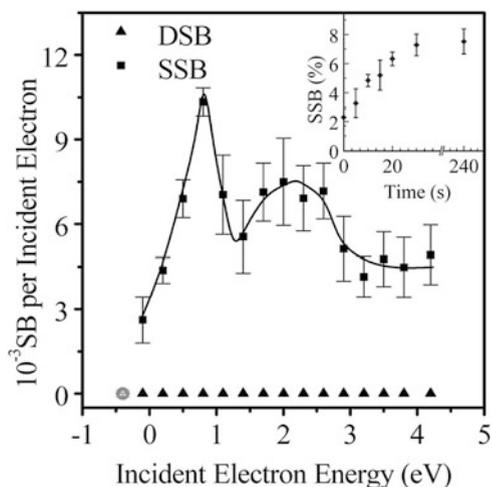
eV) compared to those for others (e.g. sugar-base N–C bond breakage). As the C–O
bond stretches, the electron promptly moves via a through-bond transfer process
from the base π* orbital through the vacant orbitals of the intervening deoxyribose
to the C–O σ* orbital [28]. Interestingly, base π-stacking was found to increase the
barrier thereby lowering the rate of SSBs [26]. Both the theoretical and experimen-
tal data relate to neutral samples in which there are no negative charges from
ionised phosphate groups. Taken together, the studies nicely demonstrated how
the interpretation of results obtained on solid DNA samples was aided by results
from gas-phase experiments and theory on simple model systems.

A full understanding of the interactions between nucleic acids and light or
electrons is nontrivial because of the large complexity of DNA and the
complicating role of the nearby environment (water molecules or counter ions). A
reductionist approach has therefore been taken by many and resulted in studies on
small DNA building blocks such as single bases (as mentioned above),
mononucleotides or smaller oligonucleotides isolated in vacuo. In this chapter
results from such experiments will be presented, and the focus is on charged species
that can easily be made by electrospray ionisation (see Chap. 2) and investigated by
mass spectrometry techniques. While a connection or importance of many of the
results to "real" biology is still to emerge, they are of immediate relevance for the
determination of the sequence of bases in oligonucleotides, are of fundamental
interest to physical chemists and are used to benchmark theoretical calculations.
The hope is that isolated strands in vacuo will serve as good models to bridge the
gap between gas-phase studies of single bases and solution-phase studies on DNA.

This chapter is organised in the following way: First, results on electron attach-
ment to nucleotide ions (both anions and cations) are presented. This section is
followed by a description of electron scattering and electron detachment
experiments on negative ions. Next photoelectron spectroscopy and absorption
spectroscopy on negative ions are discussed. Finally, the status of the results

obtained so far is given. It should be said that the chapter focuses more on electron interactions than on light interactions, which will be covered by Chap. 5 in this volume and in another volume in the same series (by J.M. Weber, J. Marcum and S. Brøndsted Nielsen) dedicated to spectroscopy of the nucleic acids in gas phase. Also fragmentation reactions of nucleic acids in gas phase are only briefly discussed here, as they are presented in Chap. 6.

4.2 Electron Attachment to Nucleotide Ions

Experiments have been done for both anions and cations, the former possess one or more negatively charged phosphate groups and neutral bases while the latter have one or more protonated bases and phosphoric acid groups. The two polarity cases are discussed separately.

4.2.1 Anions

The attachment of free electrons to nucleotide anions in vacuo is prevented by the long-range Coulomb repulsion between the two collision partners. It relies on electron tunnelling through the repulsive Coulomb barrier, which occurs with low probability (Fig. 4.3). To circumvent this obstacle, Hvelplund, Brøndsted Nielsen and co-workers [30, 31] instead did high-energy collisions between monoanions and alkali metal atoms, e.g. sodium and caesium. The latter act as electron donors due to their weakly bound s electrons. An ion accelerated to 50 keV kinetic energies moves with a velocity of about 10^5 m/s and passes by an atom within a few femtoseconds. In this interaction time, there is a finite probability for electron transfer to occur. The energy levels of the anion are Stark shifted as a result of the electric field of the sodium (or caesium) cation facilitating resonant electron transfer at large impact parameters. The collisional electron transfer process is to a good approximation vertical as the nuclei do not have time to move during the brief encounter. Popularly speaking, the so-formed fast-moving dianion leaves the sodium ion behind before the electron realises that it is unfavourable for it to be on the anion. However, now it is trapped as it has to tunnel out through the repulsive Coulomb barrier, which occurs with low probability, just as for the opposite process discussed earlier for the interaction between the anion and a free electron. A similar more well-known situation is alpha particle decay.

In this way, electron transfer to mononucleotide and larger oligonucleotide anions was successfully done [32]. An example illustrating the product ions produced in collisions between AMP anions (adenosine 5′-monophosphate, i.e. deprotonated AMP molecule) and sodium is shown in Fig. 4.4. For comparison the spectrum obtained with neon as the collision gas is also included in the figure. Neon has a similar geometrical cross section to that of sodium, but its high ionisation energy prevents it from participating in electron transfer; hence, peaks of similar magnitude seen in both the Na and Ne spectra are due to collision-

Fig. 4.3 Potential energy between a monoanion and an electron as a function of their separation. At short distances there may be either stable (<0) or unstable (>0) electronic states due to the favourable attraction between the electron and the nuclei. At larger distances, the curve follows Coulomb's law, 14.4 eV Å/R
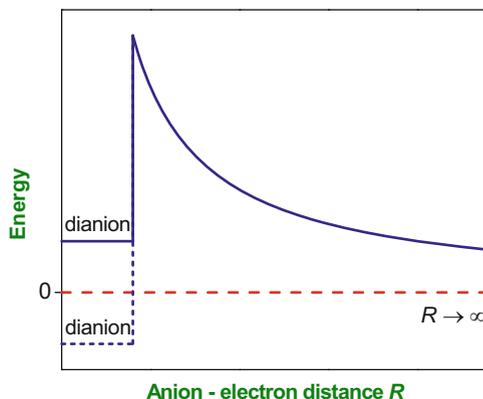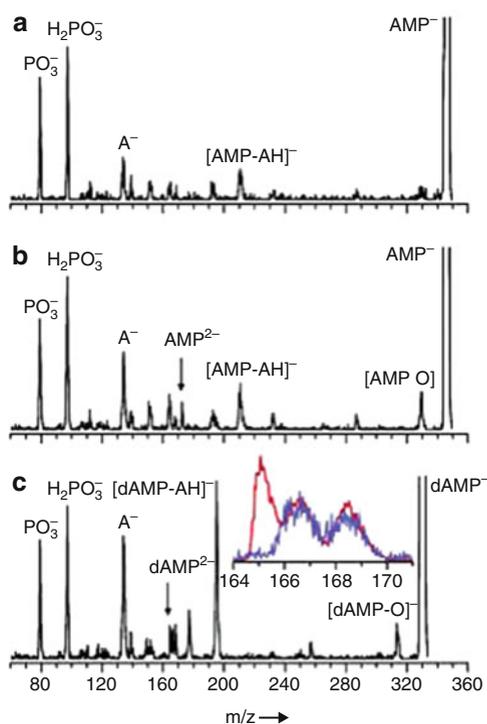

Fig. 4.4 Spectra obtained after high-energy collisions between AMP anions (m/z 346) and neon (**a**) and sodium (**b**) and between dAMP anions (m/z 330) and sodium (**c**). The *inset* in (**c**) shows the region around half the m/z of the dAMP anion for collisions between dAMP anions and neon (*blue curve*) or sodium (*red curve*). Taken from [32]

induced dissociation not involving electron transfer. Evidently, the oxygen loss channel significantly increases in importance from Ne to Na, which is indicative of dissociative electron capture. It is unclear which oxygen is lost, and whether the loss is due to electron attachment to a base π^* orbital or to the phosphate P=O π^* orbital, though the latter is highly unfavourable due to the Coulomb repulsion from the negative charge.

Fig. 4.5 Narrow scan mass spectra in the region of the dianion formed in collisional electron transfer experiments between sodium atoms and AMP anions containing different numbers of deuterium. Zero corresponds to the intact dianion. (a) AMP$^-$. (b) d_1-AMP$^-$. (c) d_2-AMP$^-$. (d) d_3-AMP$^-$. The more deuterium in the initial monoanion, the larger is the peak corresponding to loss of D compared to loss of H. Reprinted with permission from [33]. Copyright 2004, AIP Publishing LLC
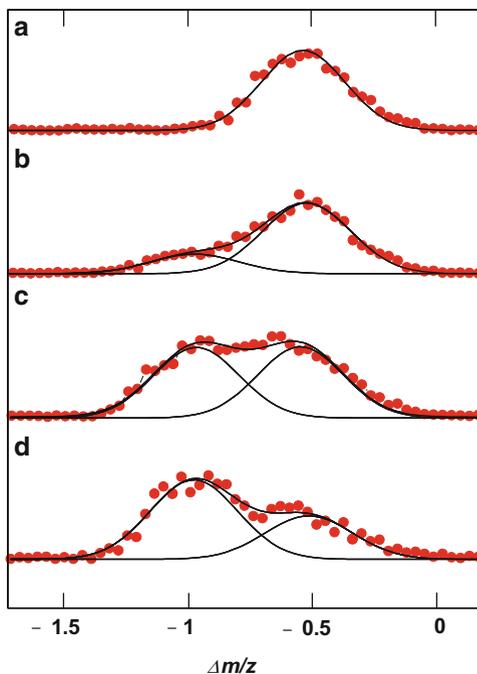
There is a small but discernible sharp peak at about half the m/z of the parent ion, which is only seen in the Na spectrum. A similar result was found for the dAMP mononucleotide (deoxyadenosine 5′-monophosphate) (Fig. 4.4c). These findings indicate that dianions are formed. The extra electron is located in an antibonding covalent orbital on the nucleobase and is predicted to be unbound. Experiments done with higher mass resolution and accuracy show that the dianions are actually dehydrogenated. Hence, electron capture is followed by rapid loss of a hydrogen atom (within the few microsecond travel time to the analyser). Experiments on deuterated ions confirmed this conclusion and revealed that the origin of the hydrogen is a heteroatom (Fig. 4.5) [33], either a nitrogen (NH_2) of the adenine, a sugar oxygen (OH) or a phosphate oxygen (POH). Deuterated ions were formed by dissolving the nucleotide in a solution of D_2O, CD_3OD and CH_3COOD for electrospray ionisation and carefully saturating the region around the electrospray needle with heavy water to prevent back exchange during electrospray. Only hydrogens bound to heteroatoms are exchangeable in such experiments. Loss of hydrogen from the phosphoric acid group seems highly unlikely as doubly deprotonated AMP cannot be made by electrospray ionisation because the Coulomb repulsion between two nearby negatively charged oxygen atoms would be too large. Furthermore, experiments on all eight mononucleotides (AMP, CMP, GMP, UMP, dAMP, dCMP, dGMP, dTMP) revealed strong differences in the cross section for formation of dehydrogenated ions from transient dianions, which indicates that the base plays a key role. Cross sections were determined by varying the pressure of

Fig. 4.6 The cross section for formation of dehydrogenated mononucleotide dianions as a function of the number of NH hydrogens (1, 2, or 3). Uracil and thymine (1), cytosine and adenine (2) and guanine (3). Reprinted with permission from [33]. Copyright 2004, AIP Publishing LLC
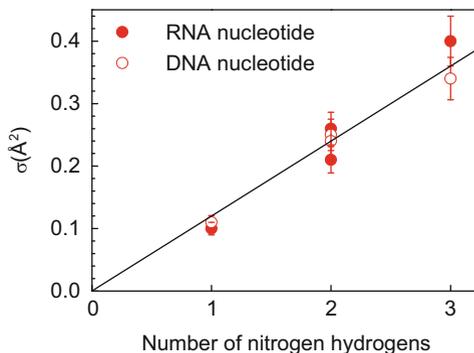
sodium atoms. A clear linear correlation between the cross section and the number of NH hydrogens in the base was found, and an extrapolation to zero NH hydrogens gave a zero cross section (Fig. 4.6), which is in strong support of the hydrogen loss originating solely from the NH base groups. Thus, nucleotides with thymine or uracil bases have the lowest cross sections while those with guanine have the highest, and nucleotides with cytosine or adenine are in between. Also there is no difference between RNA and DNA nucleotides, which implies that substitution of the OH on C2′ with H is unimportant.

These findings are in full agreement with the work done by Desfrançois, Scheier, Farizon, Märk, Illenberger and co-workers [18–24]. They measured the outcome of resonant attachment of low-energy electrons (0–3 eV) to isolated nucleobases as a function of electron energy and found that electron attachment promptly led to hydrogen loss in addition to the formation of a multitude of small fragment anions. Several resonances were identified. From isotope-labelling studies, it was established that NH was the site of dehydrogenation (Fig. 4.7) [24]. The reaction is driven by the large electron affinity of the dehydrogenated radical. Finally, it is worth mentioning that dissociative electron attachment to deoxyribose leads to dehydrogenation but as a minor channel [34].

Next, the role of hydration on nucleotide anions was considered from collisional electron transfer experiments on nanosolvated AMP anions [35]. Spectra for $AMP^-(H_2O)_{10}$ and $AMP^-(H_2O)_{16}$ are shown in Fig. 4.8. Peaks that correspond to a distribution of hydrated $[AMP - H]^{2-}$ dianions are clearly seen. In all cases, electron transfer induces loss of hydrogen and loss of some or all of the water molecules. Careful calibration experiments were carried out to confirm that hydrogen loss had indeed occurred. Interestingly, the cross section for formation of dehydrogenated dianions increased with the number of water molecules m in the initial cluster (Fig. 4.9). Hence, the damage of the nucleotide was smallest for the bare nucleotide ($m = 0$) and more than fifty times larger for $m = 16$. This contrasted with high-energy collision-induced dissociation experiments that showed that the intact AMP anion would survive when surrounded by enough water molecules as evaporation of the water molecules would cool the ion [36]. On the other hand, for electron capture the energy of the dianion state is lowered by hydration thereby

Fig. 4.7 Dissociative
electron attachment to
deuterium-labelled thymine
(C–D) induces selective loss
of hydrogen from NH at low
electron energies. Taken from
[24]. Copyright (2004) by
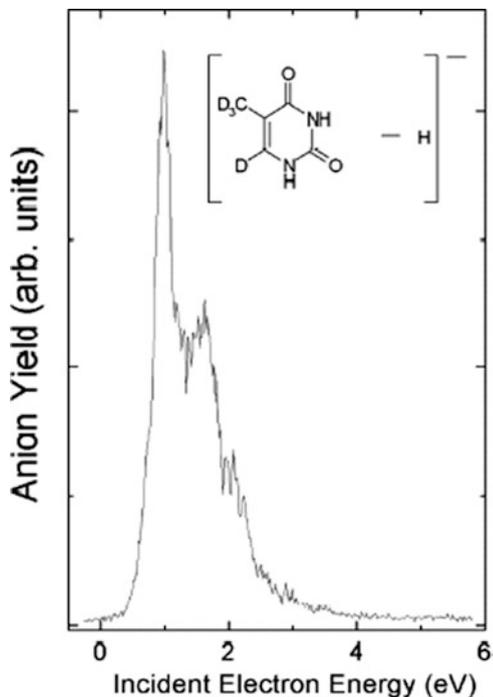The American Physical
Society

Fig. 4.8 Mass spectra
obtained after collisions
between sodium atoms and
either $AMP^-(H_2O)_{10}$ or
$AMP^-(H_2O)_{16}$. The numbers
label the number of water
molecules attached to the
dehydrogenated dianion,
$[AMP - H]^{2-}$. Reprinted with
permission from
[35]. Copyright 2008, AIP
Publishing LLC
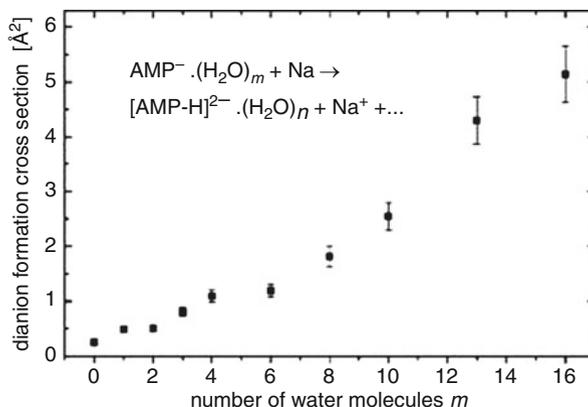
lowering the energy defect of the collisional electron transfer process rendering it more likely.

Finally, it is worth to point out that nucleobase anions can be generated by a supersonic expansion nozzle ion source or through electron transfer collisions between the bases and laser-excited Rydberg atoms as demonstrated by Bowen and co-workers [37] and Desfrançois and co-workers [18], respectively. The extra electron is dipole bound by about 0.1 eV or less and is located in a huge and diffuse orbital at the positive end of the molecule's permanent dipole moment. However, as beautifully shown by Bowen and co-workers [38] through photoelectron spectroscopy of the uracil anion, a dipole bound-to-covalent state transformation occurs after just one water molecule is bound to the anion (Fig. 4.10). Dipole-bound anions exhibit a single, strong, narrow feature at very low electron binding energies in contrast to valence anions that give broad bands due to structural differences between the anion and its corresponding neutral. The water molecule stabilises the covalent anion more than the dipole-bound one as the former has a denser excess electron distribution. With xenon attached both dipole-bound and covalent-like features were seen while in the case of argon and krypton only dipole-bound forms are present (Fig. 4.10).

4.2.2 Cations

Collisional electron transfer experiments have also been performed for protonated nucleobases. Since neutrals are produced, reionisation in a second collision is necessary for mass spectrometric detection. This is done after deflection of all ions from the beam by an electric field, cf., neutral reionisation (NR) mass spectrometry. Figure 4.11 shows $^+$NR$^+$ spectra obtained by Wyer et al. [39] using caesium and molecular oxygen as electron-donating and electron-stripping gases, respectively, for the five protonated nucleobases. While dehydrogenated cations were clearly observed, intact cations were also measured. It implies that some

Fig. 4.10 Photoelectron spectra of (**a**) U^-, (**b**) $U^-(Ar)$, (**c**) $U^-(Kr)$, (**d**) $U^-(Xe)$ and (**e**) $U^-(H_2O)$ recorded using 2.540 eV photons. *U* uracil. Reprinted with permission from [38]. Copyright 1998, AIP Publishing LLC
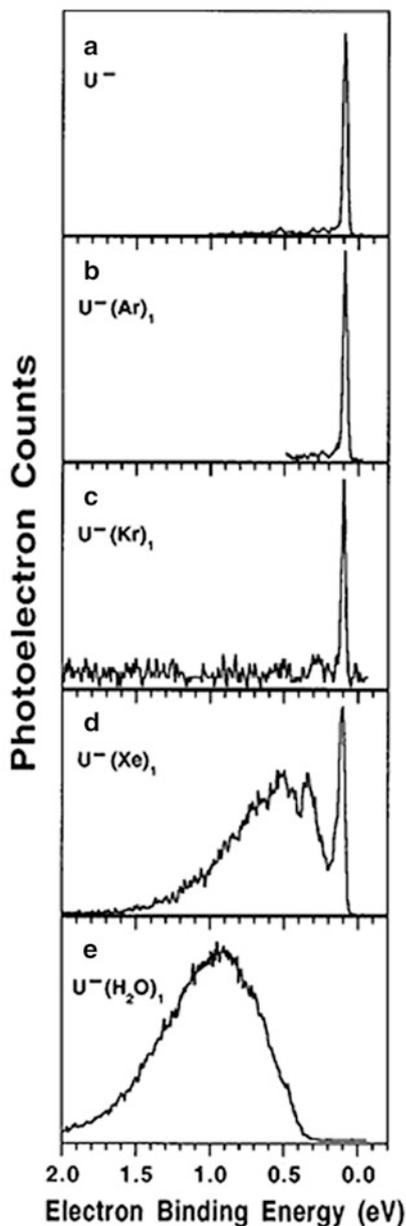

neutralised species survive the 2 µs flight time between the two collision cells. This is in clear contrast to the fact that neutral bases in nucleotides promptly lose hydrogen after electron capture.

Multiply charged oligonucleotide cations have also been subjected to electron capture dissociation (ECD) [40–43] and electron transfer dissociation (ETD) [44]

Fig. 4.11 $^+NR^+$ (i.e. conversion of cations to neutrals and back to cations) spectra resulting from high-energy collisions of MH^+ with Cs and O_2. (**a**) M = G, (**b**) M = A, (**c**) M = T, (**d**) M = U and (**e**) M = C. Taken from [39]
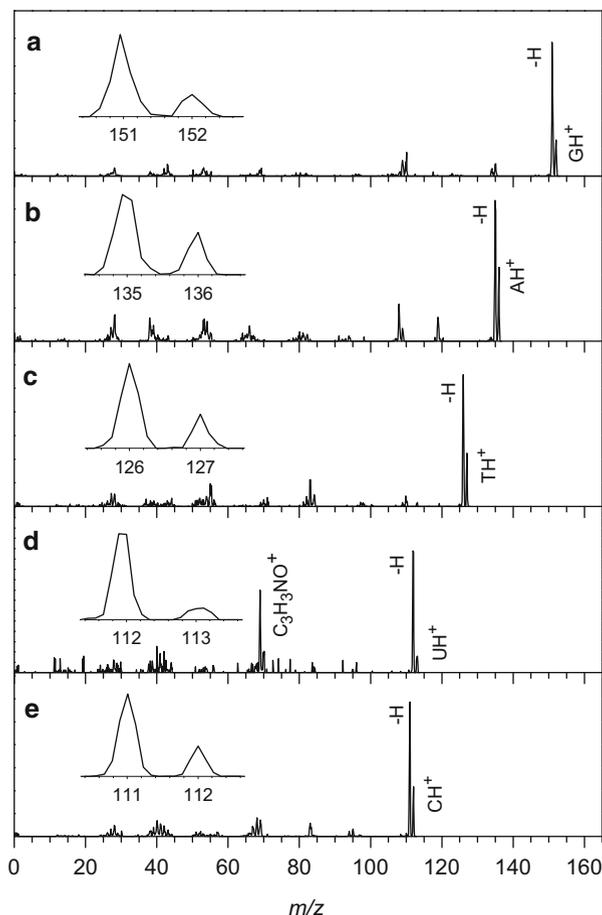
experiments. The former ions recombine with low-energy electrons in the cell of a Fourier-Transform Ion Cyclotron Resonance (FTICR) instrument while the latter ions react with atomic or molecular anions, often fluoranthene anions, in an ion trap. The preferred site of neutralisation is the protonated nucleobase and not the P=O bonds [43, 45]. Schultz and Håkansson [40] showed that ECD followed by infrared heating caused extensive fragmentation of charge-reduced oligonucleotide cations. Of particular importance cleavage of carbon–oxygen bonds at different sides of an inter-residue phosphate group occurred to produce sequence-specific fragment ions. Likewise, Smith and Brodbelt [44] found that ETD followed by low-energy CID provided better sequence coverage than CID alone and a decrease of base loss and internal fragments. Base loss is a low-energy dissociation channel, and it therefore dominates CID. For large oligonucleotides base loss is rather uninformative and undesired, at least when it comes to establishing the oligonucleotide sequence. Similar experiments on oligonucleotide duplexes resulted in specific backbone
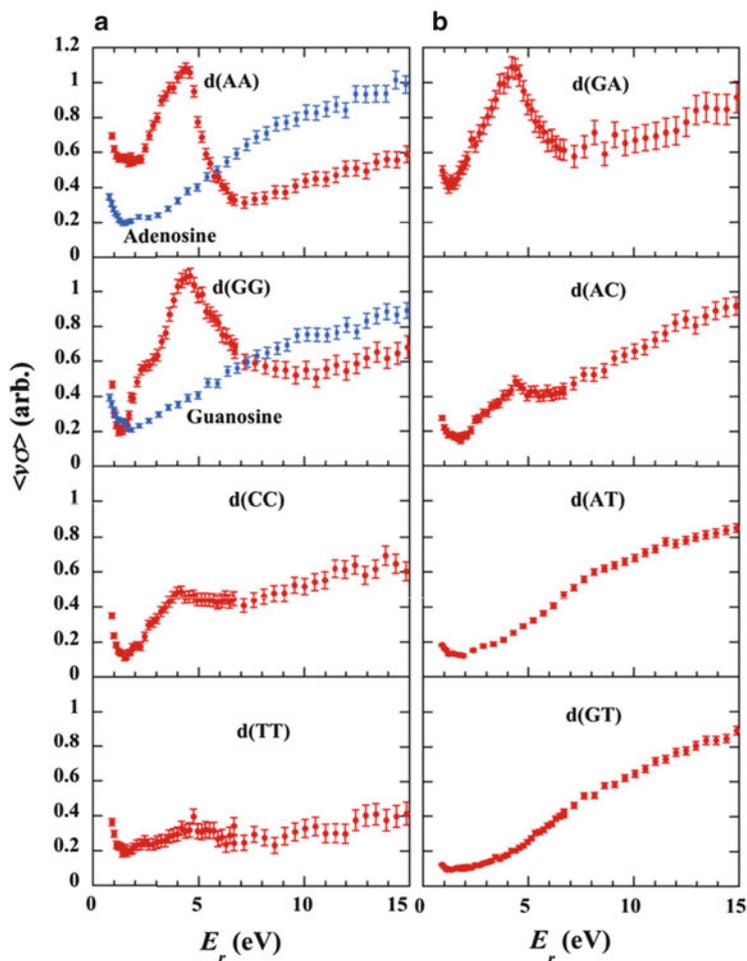
Fig. 4.12 Rates for the formation of neutral species for singly protonated mononucleotides and dinucleotides versus electron energy. Reprinted from [46], Copyright 2006, with permission from Elsevier

cleavages with conservation of weak non-covalent bonds, which is of importance for probing higher-order structures [44].

Finally, Tanabe and co-workers [46] have bombarded mononucleotide and oligonucleotide cations with electrons in an electrostatic ion storage ring (vide infra). Here the rate of formation of neutrals is measured at well-defined electron kinetic energies (Fig. 4.12). The authors found an increased yield in the formation of neutrals at a collision energy of about 4.5 eV in the case of $d(A_2)$, $d(G_2)$, $d(C_2)$, d (GA) and d(AC) DNA dinucleotides and longer strands, and to a lesser extent for d (T_2), whereas no such resonances were observed for mononucleotides and for
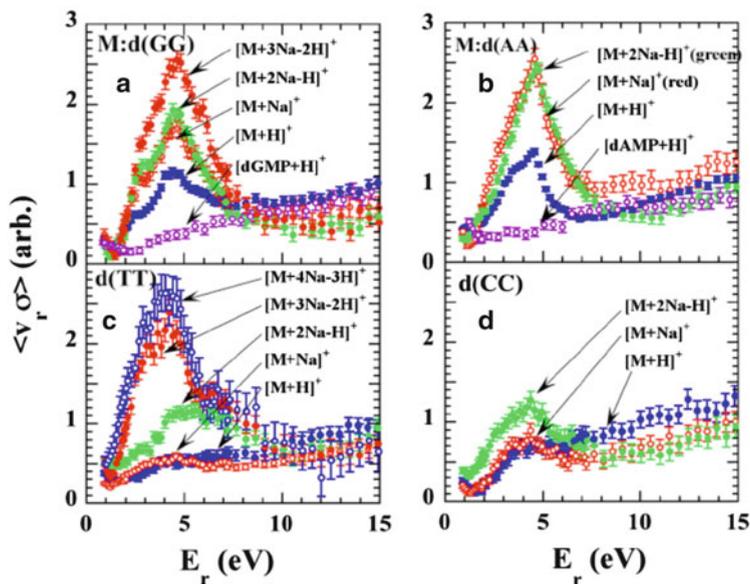
Fig. 4.13 Rates for production of neutral particles in collisions between electrons and [M + n Na − (n–1) H]$^+$ dinucleotide cations where M is d(G$_2$) (**a**), d(A$_2$) (**b**), d(T$_2$) (**c**) and d(C$_2$) (**d**). The rates for protonated dGMP and dAMP are also included. Reprinted from [47], Copyright 2008, with permission from Elsevier

dinucleotides containing one thymine base. These results indicate that the incoming electron excites an electronic state at 4.5 eV that involves two stacked bases, i.e. a delocalised state. The electron thereby loses its kinetic energy and gets captured by the cation. Molecular dissociation likely follows (cf., dissociative recombination, DR). These bumps in the production rates of neutrals become even more pronounced when one or more protons are replaced by sodium ions (Fig. 4.13) [47], which is likely ascribed to conformational differences between protonated and sodiated nucleotides.

4.3   Electron Scattering and Detachment Experiments

Mononucleotide and oligonucleotide anions have been subjected to electron scattering experiments using electrostatic ion storage rings that benefit from easy storage of heavy ions. In such experiments, ion bunches are injected into a ring where they circulate due to bending in electrostatic deflectors. In one side of the ring, the ions interact with an electron target, either in crossed beam (ELISA set-up) [48, 49] or in merged beam (KEK set-up) [50, 51] configurations (see Fig. 4.14). The advantage of using a storage ring for these experiments is that the ions move with high velocities so that the relative energy (E_r) between the two collision partners can be finely controlled and varied. Best resolution is obtained in the
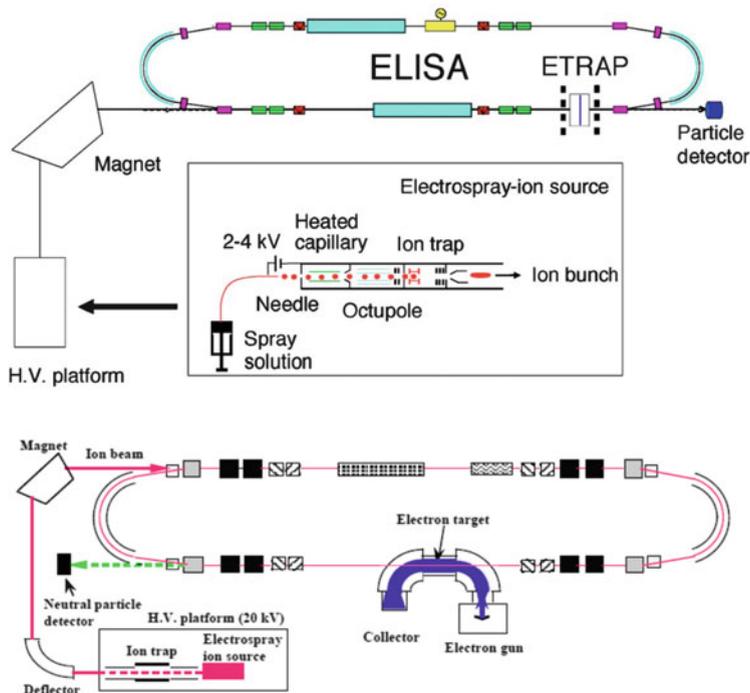
Fig. 4.14 Electrostatic ion storage rings for electron scattering experiments. In the Aarhus set-up (ELISA, *top*) ions cross a beam of electrons that are guided by a magnetic field. Neutral products are detected by a detector positioned at the end of the straight section. In the KEK set-up (*bottom*) the ion and electron beams are instead merged. Taken from [48, 51]. Copyright (2004) by The American Physical Society and © Published under CC BY-NC-SA licence in Journal of Physics: Conference Series by IOP Publishing Ltd

merged beams configuration. At high enough relative energies between the two collision partners, electron detachment may occur to produce neutral fragments, or the ions may become electronically excited also leading to dissociation. The interaction between the incoming electron and the anion is much shorter than vibrational or rotational periods, and the detachment process is therefore vertical. The electron must overcome the Coulomb repulsion to cause detachment, which implies that there is an effective threshold two to three times larger than the vertical detachment energy of the anion. For atomic anions, the relation between the threshold energy and the VDE is $E_{th} = 2^{1/4} \text{VDE}^{3/4}$ (in atomic units) [52].

The yield of neutral species is measured as a function of relative energy to produce curves like the ones shown for AMP^- and UMP^- in Fig. 4.15 [48]. The cross section has a smooth energy dependence and can be described by a simple classical model introduced by Andersen and co-workers [53] for atomic anions, $\sigma = \sigma_0 (1 - E_{th}/E_r)$ where σ_0 is a constant. The model assumes the anion to be spherically symmetric and hence has its limitations for molecular anions. As it is

Fig. 4.15 Cross sections for the production of neutral species in interactions between mononucleotide anions and electrons as a function of the electron energy. *Top*: *AMP* adenosine 5′-monophosphate. *Bottom*: *UMP* uridine 5′-monophosphate. The *curves* are based on a simple classical model that only depends on the vertical detachment energy of the anion and the electron energy. Taken from [48]. Copyright (2004) by The American Physical Society
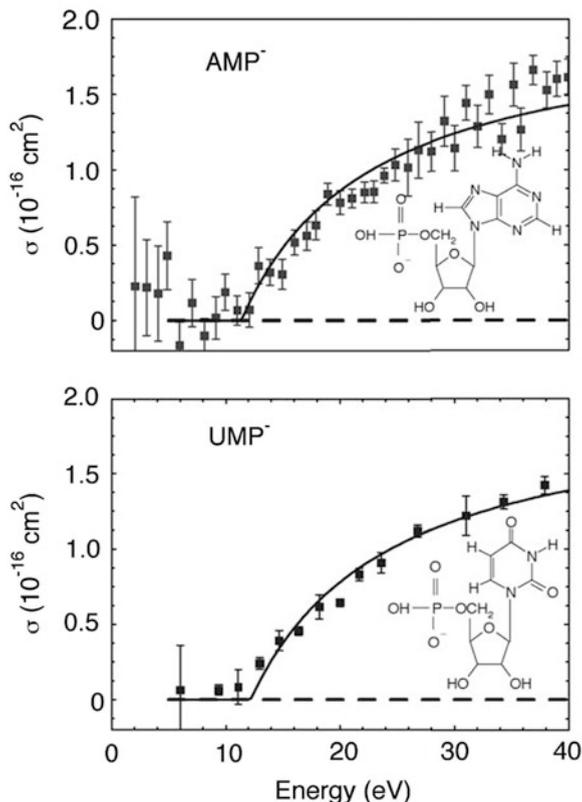
evident from the figure, curves based on this model can be reasonably well fit to the data. The larger polarisability of molecular anions compared to atomic ones implies, however, that the threshold is at higher energies than that predicted from the equation above. In other words, the electrons of the molecular anion avoid the incoming electron by relocating themselves in the electric field set up by the incoming electron, and as a result more energetic incoming electrons are needed to induce detachment [54]. This is seen for the mononucleotides where the measured threshold energy is 12 ± 1 eV while the predicted value based on the equation above is about 10 eV (VDE $= 6.05$ eV for dAMP [55]).

Resonances are sometimes superimposed on the simple detachment curve and are signatures of the population of short-lived dianion states [56]. The electron tunnels through the Coulomb barrier with a higher probability if there happens to be a state at the kinetic energy of the electron. The dianion lives for a very short time (femtoseconds) before it breaks apart, thus higher cross sections for neutral production are seen. Such resonances were, however, not identified for the two mononucleotides.

Fig. 4.16 Production rates of neutral species formed after interactions between oligonucleotide anions and electrons as a function of the electron energy. The charge states of the *x*-mers are given on the figure, where *x* is the number of nucleotides. Sequences are (a) d(TG), (b) d(ATG), (c) d(ATGC), (d) d(ATGCTG), (e), (f) d(ATGCATGC), (g) d(ATGCATGCTG), (h) d(ATGCATGCATGC), (i) d(ATGCATGCATGCAT). Taken from [57]. Copyright (2004) by The American Physical Society
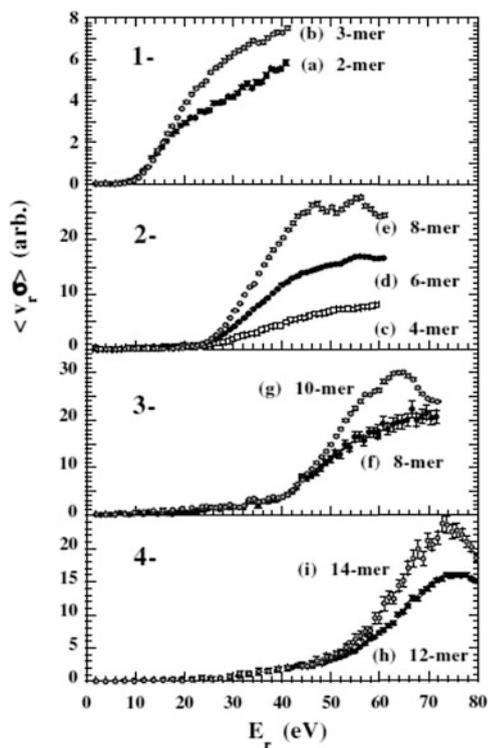
Tanabe and co-workers [57] studied oligonucleotide anions, both singly and multiply charged, up to four negative charges and reported production rates of neutral species as a function of electron energy (Fig. 4.16). They found that the threshold for electron detachment increased in steps of about 10 eV starting from about 10 eV for the singly charged ion. The thresholds seen for the monoanions are similar to those measured for the mononucleotide anions described above. The authors explained their findings based on the collective excitation of all valence electrons of the bases and the sugar-phosphate backbone (plasmon state), and this excitation energy was calculated to be about 10 eV somewhat dependent on the length of the oligomer. The authors noted that the experimental threshold energies approximately agree with the 10 eV excitation energy multiplied by the charge state, which implies that the electron would excite a larger number of plasmon quanta the higher the charge state. Further support for this idea was not given.

It is, however, clear that the larger the charge state of an anion, the more difficult it is for the electron to approach and penetrate the repulsive Coulomb barrier, and higher electron energy is therefore needed to cause excitation or electron detachment. Brøndsted Nielsen, Andersen and co-workers [58] also did electron scattering experiments on doubly charged anions, $Pt(CN)_4^{2-}$ and $Pt(CN)_6^{2-}$, and found that the threshold is simply shifted up by a factor equal to the square root of the charge state (i.e. $2^{1/2}$) compared to that of corresponding monoanions. This does not seem

to be enough by itself to explain the increase in threshold with charge state for the oligonucleotide anions.

The storage-ring experiments did not provide information on the dissociation channels after electron detachment or excitation. Mass spectrometric identification of the product ions formed after electron detachment dissociation (EDD) experiments on multiply charged oligonucleotide anions was done by Håkansson and co-workers [42, 43] using a FTICR instrument. They found that more information on the sequence of nucleobases is obtained from such experiments than from collision-induced dissociation due to cleavages of the phosphodiester backbone with no preceding base loss. Complete sequencing is achieved and has enhanced sensitivity compared to ECD of oligonucleotide cations under similar conditions. There is no apparent nucleobase effect, and as direct detachment occurs, it is therefore likely from the deprotonated phosphate backbone. This contrasts with photodetachment experiments, which are described below. Electron photodetachment followed by collision-induced dissociation (EPD) also provides good sequence coverage [59]. Similar to ECD, EDD and EPD may disrupt covalent bonds and preserve non-covalent ones in the dissociation of duplexes containing two non-covalently bound strands.

Work by McLuckey and co-workers [60, 61] suggests that odd-electron oligonucleotide anions gives less base loss than even-electron species under similar activation conditions. However, the fragmentation behaviour may depend on how the radical anions are produced (ETD, EDD, photodetachment or collisional electron detachment [62]) due to differences in whether the electron is detached from the backbone or from the nucleobase [61].

4.4 Photoelectron Spectroscopy

Wang and co-workers [55] reported the first photoelectron spectra of singly charged DNA nucleotide anions (Fig. 4.17). The spectra of dAMP$^-$, dCMP$^-$ and dTMP$^-$ are similar: They have a spectral onset at about 5.4 eV with broad and continuous spectral features. The spectrum of dGMP$^-$ on the other hand displays a weak and well separated band at much lower binding energies while the spectrum at higher energies is similar to those of dAMP$^-$ and dCMP$^-$. Mononucleotides of adenine, thymine and cytosine have electron binding energies slightly above 5 eV while that of [dGMP–H]$^-$ is 0.7 eV lower. Likewise, di- and trinucleotides containing guanine had lower electron binding energies than other strands. The HOMO is therefore on the guanine base (π orbital) and not on the phosphate group. In contrast, it was concluded that electron detachment takes place from the phosphate group for the other three mononucleotides. Dinucleotides and trinucleotides were found to have higher adiabatic detachment energies than mononucleotides. The work clearly established that guanine has the lowest ionisation energy in oligonucleotides, and therefore that guanine would serve as the oxidation site in DNA damage.

Weber et al. [63] did photoelectron spectroscopy experiments on multiply charged oligonucleotide anions $(dB_5)^{4-}$ (deprotonated four times), where B = A

Fig. 4.17 Photoelectron spectra of deprotonated 2′-deoxy-5′-deoxymononucleotide anions at 157 nm (7.866 eV). (**a**) dAMP⁻, (**b**) dCMP⁻, (**c**) dGMP⁻ and (**d**) dTMP⁻ (*d*). Taken from [55]. Copyright (2004) National Academy of Sciences, USA
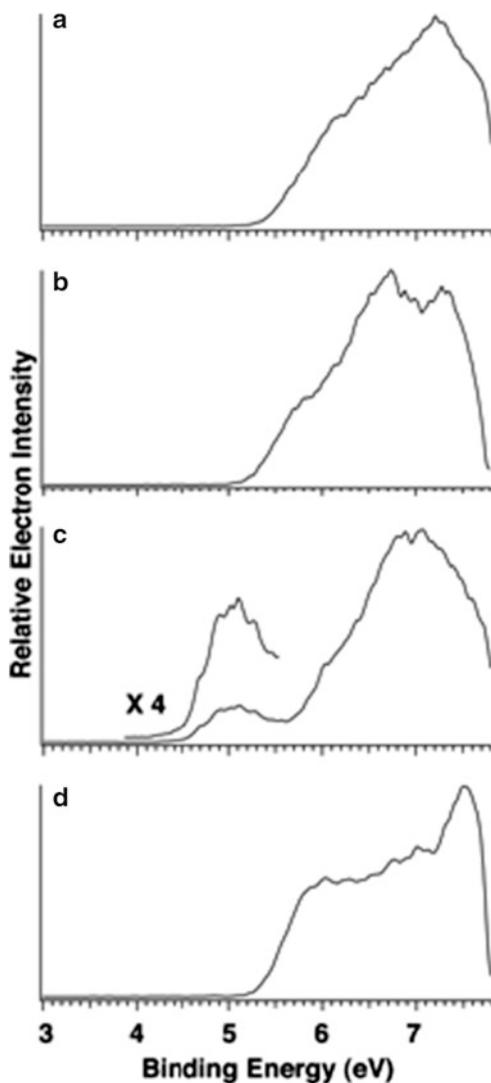
or T, and they observed that the adenine strand was much less stable with respect to electron loss than the thymine one. $(dA_5)^{4-}$ actually has a negative electron binding energy, and its existence is due only to the repulsive Coulomb barrier that prevents immediate electron loss.

In recent exciting work, Vonderach et al. [64, 65] have developed a new instrument that allows for conformer-selective photoelectron spectroscopy by combining an electrospray ionisation source, an ion mobility cell, a quadrupole mass filter and a magnetic bottle time-of-flight photoelectron spectrometer. In the case of highly negatively charged oligonucleotides, their work provided evidence for two

classes of structures: one where the excess electrons are localised at the deprotonated phosphate backbone sites and the other where there is at least one deprotonated base in addition to several deprotonated phosphate groups. Ions belonging to the latter class were shown to have the lowest electron binding energy ascribed to the deprotonated base.

4.5 Gas-Phase Absorption Spectroscopy

Absorption spectroscopy of isolated mononucleotide anions in vacuo has been done by Weber and co-workers [66], and similar experiments on oligonucleotide anions were done by Brøndsted Nielsen [67], Gabelica [59, 68, 69] and their co-workers. As an example the spectra of the dAMP mononucleotide and (dA$_4$) oligonucleotide anions are shown in Fig. 4.18. These were obtained from the yields of photofragment ions in the case of dAMP and photoneutrals in the case of (dA)$_4$ as a function of excitation wavelength. The spectral shapes are quite similar, but the monomer displays an absorption band with maximum at 253 nm while that of the oligomer is slightly blueshifted (by 3 nm) to 250 nm. Such a blueshift has been predicted by theory and is ascribed to the electronic coupling between stacked adenine bases [71]. Thus, the electronic coupling of two bases leads to two new excited state wavefunctions (exciton states) that have energies that are shifted up and down by the same amount relative to the excited state of an isolated monomer. According to theory, the higher exciton state has a larger oscillator strength compared to the lower one, explaining the blueshifted absorption. Excitation to such states may be involved in the DR of oligonucleotide cations as discussed earlier, demonstrating a link between photoexcitation and electron bombardment experiments.

It is also noteworthy that the two gas-phase absorption bands are shifted towards the blue by about 10 nm compared to the recorded spectra of the ions dissolved in water. A similar shift has been observed for the adenine base itself [72], and the shift is therefore not linked to screening of the negative charges in bulk water.

Gabelica and co-workers [59, 66, 67] studied larger multiply charged DNA strands and obtained absorption spectra based on electron photodetachment cross sections, that is, they recorded the yield of the product ion with one less electron as a function of excitation wavelength. Their work indicated that the base is the site of detachment. Guanine has the lowest ionisation energy of the bases, and indeed, guanine strands were found to have the highest electron detachment yields. Theoretical calculations on dinucleotide monophosphates show that the HOMO is located on the base not only for guanine but also for adenine, cytosine and thymine [66]. Hence base excitation triggers electron detachment, clearly demonstrating a link between absorption and electron detachment. Interestingly, the authors also found that internal conversion (IC) could compete with electron detachment, which was evidenced from the formation of fragment ions that are typically formed from vibrationally excited ions. Brøndsted Nielsen and co-workers [67] also demonstrated ultrafast IC even for singly charged strands where the photon energy

Fig. 4.18 Absorption spectra of $(dA)_4$ and dAMP in vacuo and in aqueous solution. *Blue circles*: $(dA)_4$ anions in vacuo [67]. *Grey points*: dAMP anions in vacuo [66]. *Blue curve*: $(dA)_4$ in solution. *Red curve*: dAMP in solution. Each spectrum is normalised to the band maximum at ca. 250 nm (gas phase) and ca. 260 nm (solution phase). Reproduced from [70] with permission from The Royal Society of Chemistry
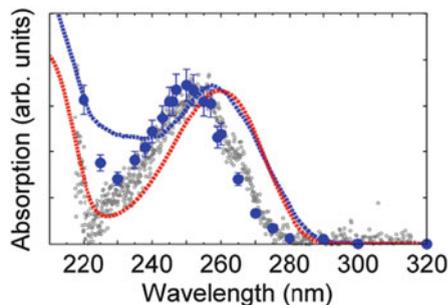
was higher than the electron binding energy; hence, it seems important to establish the competition between electron detachment and IC as a function of excitation energy as the absorption spectrum based on electron detachment would be skewed if this ratio changes with wavelength.

In other work also based on electron photodetachment, Rosu et al. [69] found that duplexes and quadruplexes displayed redshifted absorption bands compared to the corresponding single strands, and such spectroscopy could therefore be useful to obtain information on folding motifs.

4.6   Conclusion and Outlook

The study of DNA/RNA ions, isolated in vacuo, and their interactions with electrons and photons has been a very active research area in the last 10 years. Detailed information at the molecular level has been obtained on, e.g. base dehydrogenation and the origin of the lost hydrogen, electron binding energies, absorption band maxima and the electronic coupling between two or more bases in the excited state. The role of hydration was investigated for the formation of dehydrogenated mononucleotide dianions and for the dipole-bound-to-covalent-state transition in nucleobase anions. The field has taken advantage of the development of new advanced instrumental set-ups such as electrostatic ion storage rings that were used by physicists and chemists in collaborative work. Threshold electron energies needed for electron detachment of singly and multiply charged anions were measured in electron scattering experiments, while dissociative recombination experiments on cations revealed a high formation of neutral species at certain electron energies (resonances). Advances in mass spectrometry and electron-induced dissociation methods have shown that nucleic acid sequencing is best achieved by the formation of radical species that undergo breakages of phosphodiester bonds. These methods also hold great promise for the elucidation of conformational structures as non-covalent bonds can be preserved while covalent

ones are broken. More work in the future on partially hydrated nucleic acid ions would be interesting to bridge the physics of bare ions with that of solvated ones and thereby increase the relevance to the biophysical situation. Further studies on the importance of counter ions would be useful to better mimic the real biological environment. Finally, the link between geometric structure and electronic structure is one that deserves much more exploration.

4.7   Summary of Key Concepts

General Concepts
- The Repulsive Coulomb Barrier prevents direct electron attachment to anions.
- Electron attachment or electron detachment results in radical-driven fragmentation.
- Electronic structure is linked to the folding motif.
- Instrumental set-ups used are mass spectrometers, electrostatic storage rings and ion mobility spectrometers.

Concepts Specific to Nucleic Acids
- Electrons induce damage to nucleotide ions.
- Electron attachment to nucleobases causes dehydrogenation.
- Solvation of bare ions results in spectral shifts in the absorption.
- Microhydration and counter ions play roles on electron attachment and/or detachment.
- Photoelectron spectroscopy provides information on electron binding energies and location of highest occupied molecular orbital.
- There are differences in the absorption between single bases, single strands, duplexes and quadruplexes isolated in vacuo.

References

1. Clark LB (1995) Transition moments of 2'-deoxyadenosine. J Phys Chem 99:4466–4470
2. Matsuoka Y, Nordén B (1982) Linear dichroism studies of nucleic acid bases in stretched poly (vinyl alcohol) film. Molecular orientation and electronic transition moment directions. J Phys Chem 86:1378–1386
3. Petke JD, Maggiora GM, Christoffersen REJ (1990) Am Chem Soc 112:5452–5460
4. Wulff DL (1963) Kinetics of thymine photodimerization in DNA. Biophys J 3:355–362
5. Lamola AA, Eisinger J (1968) On the mechanism of thymine photodimerization. Proc Natl Acad Sci USA 59:46–51
6. Johns HE, Pearson ML, Helleiner CW, Leblanc JC (1964) The ultraviolet photochemistry of thymidylyl-(3′→5′)-thymidine. J Mol Biol 9:503–524
7. Crespo-Hernandez CE, Cohen B, Hare PM, Kohler B (2004) Ultrafast excited-state dynamics in nucleic acids. Chem Rev 104:1977–2019

8. Middleton CT, de La Harpe K, Su C, Law YK, Crespo-Hernández CE, Kohler B (2009) DNA excited-state dynamics: from single bases to the double helix. Annu Rev Phys Chem 60:217–239
9. Kohler B (2010) Nonradiative decay mechanisms in DNA model systems. J Phys Chem Lett 1:2047–2053
10. Brøndsted Nielsen S, Andersen JU, Forster JS, Hvelplund P, Liu B, Pedersen UV, Tomita S (2003) Photodestruction of adenosine 5′-monophosphate (AMP) nucleotide ions *in vacuo*: statistical versus nonstatistical processes. Phys Rev Lett 91:048302
11. Markovitsi D, Gustavsson T, Banyasz A (2010) Absorption of UV radiation by DNA: spatial and temporal features. Mutat Res 704:21–28
12. Markovitsi D (2009) Interaction of UV radiation with DNA helices. Pure Appl Chem 81:1635–1644
13. Markovitsi D, Gustavsson T, Vayá I (2010) Fluorescence of DNA duplexes: from model helices to natural DNA. J Phys Chem Lett 1:3271–3276
14. Cobut V, Frongillo Y, Patau JP, Goulet T, Fraser M-J, Jay-Gerin JP (1998) Monte Carlo simulation of fast electron and proton tracks in liquid water – I. Physical and physicochemical aspects. Radiat Phys Chem 51:229–243
15. LaVerne JA, Pimblott SM (1995) Electron-energy loss distributions in solid, dry DNA. Radiat Res 141:208–215
16. Boudaïffa B, Cloutier P, Hunting D, Huels MA, Sanche L (2000) Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) Electrons. Science 287:1658–1660
17. Martin F, Burrow PD, Cai Z, Hunting D, Sanche L (2004) DNA strand breaks induced by 0-4 eV electrons: the role of shape resonances. Phys Rev Lett 93:068101
18. Desfrancois C, Abdoul-Carime H, Schermann JP (1996) Electron attachment to isolated nucleic acid bases. J Chem Phys 104:7792–7794
19. Gohlke S, Abdoul-Carime H, Illenberger E (2003) Dehydrogenation of adenine induced by slow (<3eV) electrons. Chem Phys Lett 380:595
20. Hanel G, Gstir B, Scheier P, Probst M, Farizon B, Farizon M, Illenberger E, Märk TD (2003) Electron attachment to uracil: effective destruction at subexcitation energies. Phys Rev Lett 90:188104
21. Denifl S, Ptasinska S, Cingel M, Matejcik S, Scheier P, Märk TD (2003) Electron attachment to the DNA bases thymine and cytosine. Chem Phys Lett 377:74–80
22. Denifl S, Ptasinska S, Probst M, Hrusak J, Scheier P, Märk TD (2004) Electron attachment to the gas-phase DNA bases cytosine and thymine. J Phys Chem A 108:6562–6569
23. Ptasinska S, Denifl S, Scheier P, Illenberger E, Märk TD (2005) Bond- and site-selective loss of H atoms from nucleobases by very-low-energy electrons (<3eV). Angew Chem Int Ed 44:6941–6943
24. Abdoul-Carime H, Gohlke S, Illenberger E (2004) Site-specific dissociation of DNA bases by slow electrons at early stages of irradiation. Phys Rev Lett 92:168103
25. Théodore M, Sobczyk M, Simons J (2006) Cleavage of thymine N_3-H bonds by low-energy electrons attached to base π^* orbitals. Chem Phys 329:139–147
26. Anusiewicz I, Berdys J, Sobczyk M, Skurski P, Simons J (2004) Effects of base π-stacking on damage to DNA by low-energy electrons. J Phys Chem A 108:11381–11387
27. Berdys J, Anusiewicz I, Skurski P, Simons J (2004) Damage to model DNA fragments from very low-energy (<1 eV) electrons. J Am Chem Soc 126:6441–6447
28. Berdys J, Skurski P, Simons J (2004) Damage to model DNA fragments by 0.25-1.0 eV electrons attached to thymine π^* orbital. J Phys Chem B 108:5800–5805
29. Barrios R, Skurski P, Simons J (2002) Mechanism for damage to DNA by low-energy electrons. J Phys Chem B 106:7991–7994
30. Nielsen AB, Hvelplund P, Liu B, Brøndsted Nielsen S, Tomita S (2003) Coulomb explosion upon electron attachment to a four-coordinate monoanionic metal complex. J Am Chem Soc 125:9592–9593

31. Liu B, Hvelplund P, Brøndsted Nielsen S, Tomita S (2004) Formation of C_{60}^{2-} dianions in collisions between C_{60}^- and Na atoms. Phys Rev Lett 92:168301
32. Liu B, Tomita S, Rangama J, Hvelplund P, Brøndsted Nielsen S (2003) Electron attachment to "naked" and microsolvated nucleotide anions: detection of long-lived dianions. ChemPhysChem 4:1341–1344
33. Liu B, Hvelplund P, Brøndsted Nielsen S, Tomita S (2004) Hydrogen loss from nucleobase nitrogens upon electron attachment to isolated DNA and RNA nucleotide anions. J Chem Phys 121:4175–4179
34. Ptasińska S, Denifl S, Scheier P, Märk TD (2004) Inelastic electron interaction (attachment/ ionization) with deoxyribose. J Chem Phys 2004(18):8505–8511
35. Liu B, Haag N, Johansson H, Schmidt HT, Cederquist H, Brøndsted Nielsen S, Zettergren H, Hvelplund P, Manil B, Huber BA (2008) Electron capture induced dissociation of nucleotide anions in water nanodroplets. J Chem Phys 128:075102
36. Liu B, Brøndsted Nielsen S, Hvelplund P, Zettergren H, Cederquist H, Manil B, Huber BA (2006) Collision-induced dissociation of hydrated adenosine monophosphate nucleotide ions: protection of the ion in water nanoclusters. Phys Rev Lett 97:133401
37. Hendricks JH, Lyapustina SA, de Clercq HL, Snodgrass JT, Bowen KH (1996) Dipole bound, nucleic acid base anions studied via negative ion photoelectron spectroscopy. J Chem Phys 104:7788–7791
38. Hendricks JH, Lyapustina SA, de Clercq HL, Bowen KH (1998) The dipole bound-to-covalent anion transformation in uracil. J Chem Phys 108:8–11
39. Wyer JA, Cederquist H, Haag N, Huber BA, Hvelplund P, Johansson HAB, Maisonny R, Brøndsted Nielsen S, Rangama J, Rousseau P, Schmidt HT (2009) On the hydrogen loss from protonated nucleobases after electronic excitation or collisional electron capture. Eur J Mass Spectrom 15:681–688
40. Schultz KN, Håkansson K (2004) Rapid electron capture dissociation of mass-selectively accumulated oligodeoxynucleotide dications. Int J Mass Spectrom 234:123–130
41. Håkansson K, Hudgins RR, Marshall AG, O'Hair RAJ (2003) Electron capture dissociation and infrared multiphoton dissociation of oligonucleotide dications. J Am Soc Mass Spectrom 14:23–41
42. Yang J, Håkansson K (2006) Fragmentation of oligonucleotides from gas-phase ion-electron reactions. J Am Soc Mass Spectrom 17:1369–1375
43. Yang J, Mo J, Adamson JT, Håkansson K (2005) Characterization of oligonucleotides by electron detachment dissociation Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 77:1876–1882
44. Smith SI, Brodbelt JS (2009) Electron transfer dissociation of oligonucleotide cations. Int J Mass Spectrom 283:85–93
45. Wah T, Chan D, Choy MF, Chan WYK, Fung YME (2009) A mechanistic study of the electron capture dissociation of oligonucleotides. J Am Soc Mass Spectrom 20:213–226
46. Tanabe T, Starikov EB, Noda K, Saito M (2006) Resonant neutral-particle emission after collisions of electrons with base-stacked oligonucleotide cations in a storage ring. Chem Phys Lett 430:380–385
47. Tanabe T, Starikov EB, Noda K (2008) Resonant neutral-particle emission correlated with base-base interactions in collisions of electrons with protonated and sodiated dinucleotide monocations. Chem Phys Lett 467:154–158
48. Bluhme H, Jensen MJ, Brøndsted Nielsen S, Pedersen UV, Seiersen K, Svendsen A, Andersen LH (2004) Electron scattering on stored mononucleotide anions. Phys Rev A 70:020701
49. Andersen JU, Hvelplund P, Brøndsted Nielsen S, Tomita S, Wahlgreen H, Møller SP, Pedersen UV, Forster JS, Jørgensen TJD (2002) The combination of an electrospray ion source and an electrostatic storage ring for lifetime and spectroscopy experiments on biomolecules. Rev Sci Instrum 73:1284–1287
50. Tanabe T, Noda K, Syresin E (2004) An electrostatic storage ring with a merging electron beam device at KEK. Nucl Instrum Methods Phys Res A 532:105–110

51. Tanabe T (2007) Neutral-particle emission from multiply charged biomolecular ions in collisions with electrons. J Phys Conf Ser 58:17–24

52. Seiersen K, Bak J, Bluhme H, Jensen MJ, Brøndsted Nielsen S, Andersen LH (2003) Electron-impact detachment of O_3^-, NO_3^- and SO_2^- ions. Phys Chem Chem Phys 5:4814–4820

53. Andersen LH, Mathur D, Schmidt HT, Vejby-Christensen L (1995) Electron-impact detachment of D^-: near-threshold behavior and the nonexistence of D^{2-} resonances. Phys Rev Lett 74:892–895

54. El Ghazaly MOA, Svendsen A, Bluhme H, Brøndsted Nielsen S, Andersen LH (2005) Electron scattering on *p*-benzoquinone anions. Chem Phys Lett 405:278–281

55. Yang X, Wang X-B, Vorpagel ER, Wang L-S (2004) Direct experimental observation of the low ionization potentials of guanine in free oligonucleotides by using photoelectron spectroscopy. Proc Natl Acad Sci USA 101:17588–17592

56. Svendsen A, Bluhme H, El Ghazaly MOA, Seiersen K, Brøndsted Nielsen S, Andersen LH (2005) Tuning the continuum ground state energy of NO_2^{2-} by water molecules. Phys Rev Lett 94:223401

57. Tanabe T, Noda K, Saito M, Starikov EB, Tateno M (2004) Regular threshold-energy increase with charge for neutral-particle emission in collisions of electrons with oligonucleotide anions. Phys Rev Lett 93:043201

58. El Ghazaly MOA, Svendsen A, Bluhme H, Nielsen AB, Brøndsted Nielsen S, Andersen LH (2004) Electron scattering on centrosymmetric molecular dianions $Pt(CN)_4^{2-}$ and $Pt(CN)_6^{2-}$. Phys Rev Lett 93:203201

59. Gabelica V, Tabarin T, Antoine R, Rosu F, Compagnon I, Broyer M, De Pauw E, Dugourd P (2006) Electron photodetachment dissociation of DNA polyanions in a quadrupole ion trap mass spectrometer. Anal Chem 78:6564–6572

60. McLuckey SA, Stephenson JL Jr, O'Hair RAJ (1997) Decompositions of odd- and even-electron anions derived from deoxy-polyadenylates. J Am Soc Mass Spectrom 8:148–154

61. Gao Y, McLuckey SA (2013) Electron transfer followed by collision-induced dissociation (NET-CID) for generating sequence information from backbone-modified oligonucleotide anions. Rapid Commun Mass Spectrom 27:249–257

62. Liu B, Hvelplund P, Brøndsted Nielsen S, Tomita S (2003) Electron loss and dissociation in high energy collisions between multiply charged oligonucleotide anions and noble gases. Int J Mass Spectrom 230:19–24

63. Weber JM, Ioffe IN, Berndt KM, Loffler D, Friedrich J, Ehrler OT, Danell AS, Parks JH, Kappes MM (2004) Photoelectron spectroscopy of isolated multiply negatively charged oligonucleotides. J Am Chem Soc 126:8585–8589

64. Vonderach M, Ehrler OT, Weis P, Kappes MM (2011) Combining ion mobility spectrometry, mass spectrometry, and photoelectron spectroscopy in a high-transmission instrument. Anal Chem 83:1108–1115

65. Vonderach M, Ehrler OT, Matheis K, Weis P, Kappes MM (2012) Isomer-selected photoelectron spectroscopy of isolated DNA oligonucleotides: phosphate and nucleobase deprotonation at high negative charge states. J Am Chem Soc 134:7830–7841

66. Marcum JC, Halevi A, Weber JM (2009) Photodamage to isolated mononucleotides-photodissociation spectra and fragment channels. Phys Chem Chem Phys 11:1740–1751

67. Nielsen LM, Pedersen SØ, Kirketerp M-BS, Brøndsted Nielsen S (2012) Absorption by DNA single strands of adenine isolated *in vacuo*: the role of multiple chromophores. J Chem Phys 136:064302

68. Gabelica V, Rosu F, Tabarin T, Kinet C, Antoine R, Broyer M, De Pauw E, Dugourd P (2007) Base-dependent electron photodetachment from negatively charged DNA strands upon 260-nm laser irradiation. J Am Chem Soc 129:4706–4713

69. Rosu F, Gabelica V, De Pauw E, Antoine R, Broyer M, Dugourd P (2012) UV spectroscopy of DNA duplex and quadruplex structures in the gas phase. J Phys Chem A 116:5383–5391

70. Nielsen LM, Hoffmann SV, Brøndsted Nielsen S (2013) Electronic coupling between photo-excited stacked bases in DNA and RNA strands with emphasis on the bright states initially populated. Photochem Photobiol Sci 12(8):1273–1285
71. Lange AW, Herbert JM (2009) Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the (1)pi pi* excitonic bright states. J Am Chem Soc 131:3913–3922
72. Li L, Lubman DM (1987) Ultraviolet visible absorption-spectra of biological molecules in the gas-phase using pulsed laser-induced volatilization enhancement in a diode-array spectropho-tometer. Anal Chem 59:2538–2541

Gas-Phase Spectroscopy of Nucleic Acids

<div style="text-align:right">**5**</div>

Valérie Gabelica and Frédéric Rosu

Abstract

We describe here frequency-resolved gas-phase spectroscopy of nucleic acids. Frequency resolved means that the effect of photons on the nucleic acid molecules is measured as a function of the photon frequency. The present chapter is primarily focused on experimental aspects, and intended as a compass to navigate a rather interdisciplinary field. Indeed, gas-phase spectroscopy usually combines photonics, mass spectrometry (when ions are detected), and theoretical chemistry. Although theory is of prime importance for the interpretation of the results, as it is the comparison between experimental and theoretical energies of the resonance transitions that allow the structural interpretation of the experimental spectra, extended discussion of theory levels will not be provided here, but relevant literature will be indicated along the text. We will cover rotational, vibrational, and electronic spectroscopy from isolated nucleobases to oligonucleotides and nucleic acid higher-order structures.

Keywords

Spectroscopy • Infrared • Ultraviolet • Microwave • Base stacking • Hydrogen bonding • Photoionization • Excited states • Oligonucleotides • Isolated bases • Duplex • Quadruplex • Base pairs • Tautomers

V. Gabelica (✉)
IECB, ARNA Laboratory, Univ. Bordeaux, 33600 Pessac, France

U869, ARNA Laboratory, Inserm, 33000 Bordeaux, France
e-mail: valerie.gabelica@inserm.fr

F. Rosu
UMS 3033 and Inserm US001, IECB, CNRS, 33600 Pessac, France

Abbreviations

CD	Circular dichroism
CID	Collision-induced dissociation
CLIO	Centre Laser Infrarouge d'Orsay
FELIX	Free-Electron Lasers for Infrared eXperiments
FTICRMS	Fourier transform ion cyclotron resonance mass spectrometry
IC	Internal conversion
IP	Ionization potential
IR	Infrared
IRMPD	Infrared multiple photon dissociation
IR-UV	Infrared-ultraviolet double resonance spectroscopy
IVR	Intramolecular vibrational energy redistribution
LA-MB-FTMW	Laser ablation molecular beam Fourier transform microwave spectroscopy
LIF	Laser-induced fluorescence
NMR	Nuclear magnetic resonance
PD	Photodissociation
R1PI	Resonance-enhanced single-photon ionization
R2PI	Resonance-enhanced two-photon ionization
REMPI	Resonance-enhanced multiphoton ionization
S_0	Electronic ground state
S_1	First electronically excited state
UV	Ultraviolet
UVMPD	Ultraviolet multiple photon dissociation
UV-UV	Ultraviolet–ultraviolet double resonance spectroscopy
Vis	Visible
VUV	Vacuum ultraviolet

5.1   Introduction

Molecular spectroscopy is the study of the interaction between electromagnetic radiation (or commonly speaking, "light") and molecules. The electromagnetic spectrum ranges from γ (gamma) rays to radio waves. We will concentrate here on the gas-phase spectroscopy methods that study the interaction between electromagnetic radiation and electronic, vibrational, and rotational states of molecules (Table 5.1). Spectroscopy brings information on the structure of the molecule. Electronic, vibrational, and rotational states are each quantized, and the differences in energy between states depend on the molecular structure, i.e., the atom's positions. Due to the resonance phenomenon, the interaction of a molecule with electromagnetic radiation is stronger when the energy of the radiation matches the energy difference between two states.

Table 5.1 Electromagnetic spectrum and modes of resonant interaction with molecules

	Molecular resonances	Wavelength (λ)	Wavenumber (=1/λ)	Frequency ($v = c/\lambda$)	Photon energy ($E = hv$)
Vacuum ultraviolet (VUV)	Higher electronic states and ionization potential	10–200 nm	10^6– 50,000 cm^{-1}	30–1.5 PHz	124–6.2 eV
Ultraviolet (UV)	Electronic states	10–400 nm	10^6– 25,000 cm^{-1}	30 PHz– 750 THz	124–3 eV
	Lowest-lying nucleic acid base electronic states	240–300 nm	41,667– 33,333 cm^{-1}	1.25–1 PHz	5.17– 4.13 eV
Visible (vis)	Electronic states	390–750 nm	25,641– 13,333 cm^{-1}	770–400 THz	3.2 eV– 1.7 eV
Infrared (IR)	Vibrational states	750 nm– 1 mm	13,333– 10 cm^{-1}	400 THz– 300 GHz	1.7 eV– 1.24 meV
	Highest-energy vibrational states	2.5–40 μm	4,000– 500 cm^{-1}	120–7.5 THz	0.5– 0.062 eV
Microwave	Rotational states	1 mm–1 m	10–0.01 cm^{-1}	300 GHz– 300 MHz	1.24 meV– 1.24 μeV
	Typical FT-microwave spectrophotometers	75 mm– 15 mm	0.13– 0.67 cm^{-1}	20–4 GHz	83 μeV– 16.5 μeV

Spectroscopy is a widely used biophysics method to characterize nucleic acids in solution. Examples of typical IR absorption spectra and UV absorption spectra are shown in Figs. 5.1 and 5.2, respectively. Solution-phase spectroscopy is very easy to carry out with bench-top instruments nowadays, yet its structural interpretation is not always straightforward, for several reasons:

1. Some regions of the spectra may be obscured by the predominant signal of the solvent or buffer. This is illustrated by the H_2O and D_2O bands in the IR spectra of Fig. 5.1. To cover the whole range of structurally interesting bands, actually two spectra in the two solvents must be recorded.

2. The frequency of resonances may shift due to the solvent. In the case of IR spectroscopy, this can for example be due to hydrogen bonding of exposed bases with water: a C=O bond will be weakened when H-bonded to H_2O, and its frequency or wavenumber will diminish. Solvent-induced shift of the electronic transitions frequencies is also a common phenomenon, called solvatochromism. In electronic spectroscopy, because the polarities of the ground and the excited states may differ, the energy difference will be influenced in the presence of a polar solvent like water. Hypsochromism is a shift to higher wavelength, the blue side of the visible spectrum, also called "blue-shift," and in electronic spectroscopy it indicates that the ground state is more polar than the excited state. Bathochromic is a shift to lower wavelength, or "red-shift," and indicates that the electronic excited state is more polar than the ground state. The terms blue-shift or red-shift are widely used, beyond solvatochromism and visible spectroscopy (i.e., these terms are also encountered in UV or IR spectroscopy).
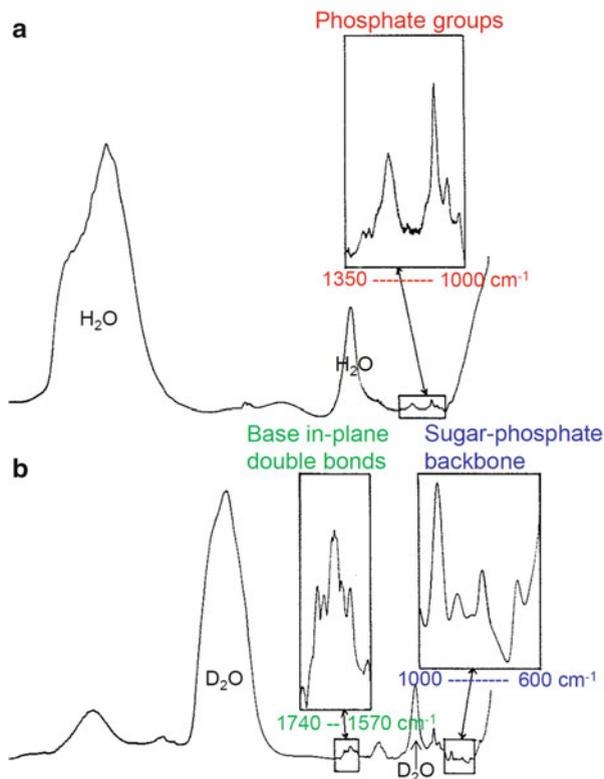
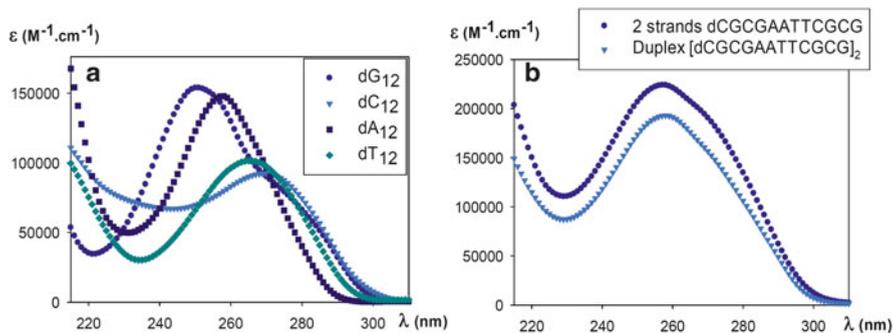**Fig. 5.1** Infrared spectroscopy of nucleic acids in solution (adapted with permission from [1])

Fig. 5.2 Ultraviolet spectroscopy of nucleic acids in solution. *Source*: simulations were carried out on http://biophysics.idtdna.com/, based on parameterization of Owczarzy and coworkers [2]

3. Solutions can contain mixtures of conformers, or mixtures of molecules, or mixtures of molecular aggregation states. This matter of fact may be wanted or unwanted, characterized, or uncharacterized. In the case of mixtures, the

solution spectrum represents a weighted averaged spectrum of all species contributing in the solution.

4. Solution-phase experiments are limited to a temperature range where the solvent phase is liquid.

These four reasons bring us to the specific advantages of gas-phase spectroscopy, which is spectroscopy in the absence of solvent:

1. In the gas phase, solvent does not interfere. For rotational spectroscopy, it is important that the molecules rotate freely to calculate their moments of inertia.
2. Energy levels and the frequencies of the transitions can be computed at much lower cost and to a higher level of accuracy, so it is easier to interpret gas-phase experiments by comparison with theory.
3. The components of a mixture can be sorted in mass if they are in the form of ions and analyzed in a mass spectrometer. Ion sorting can be done before spectroscopy probing of ions produced in a source, after spectroscopy probing (for neutrals produced in the vaporization source, ionization becomes part of the spectroscopy probing scheme, for example, with photoionization), or both.
4. Ions can be cooled down to very low temperatures (close to that of liquid helium, 4 K), which is important for the molecule to be in the ground state not only for electronic but also for vibrational levels. This enables high-resolution vibrational spectroscopy to be carried out, and conformer-specific probing based on sharp energy-resolved transitions.

Gas-phase spectroscopy experiments therefore play a major role in understanding the fundamental principles of spectroscopy and in benchmarking theoretical computations of molecular quantized states. In the field of fundamental spectroscopy of nucleic acids, many questions remain open in particular with regard to the nature of the electronic excited states and the excited states dynamics (what is the fate of the molecule after being promoted to a particular excited state; by which pathway does it relax and how long does it take).

5.2   Gas-Phase Spectroscopy Experiments

Experimental setups can be categorized into absorption spectroscopy and action spectroscopy setups.

- **Absorption spectroscopy** means that light absorption is determined by monitoring the difference of photon flux entering the sample and exiting the sample: the difference was absorbed by the sample due to molecular excitation. Absorbance measurements in UV–Vis spectroscopy, and transmission measurements in IR spectroscopy, are examples of absorption setups in solution.
- **Action spectroscopy** means that light absorption is determined by monitoring a detectable effect of that light on the molecule. Fluorescence spectroscopy can be thought of as a form of action spectroscopy in solution: monitoring the fluorescence intensity at a given wavelength (the action) as a function of the frequency of the excitation would provide a frequency-resolved action spectrum of the molecule. The frequency-resolved fluorescence action spectrum can be similar

to the absorption spectrum, but it might also differ, for example if only a subpopulation of states gives rise to the monitored action.

This important difference between absorption spectroscopy and action spectroscopy must be kept in mind in order to understand gas-phase spectroscopy, as most setups are action spectroscopy setups. One technical reason is that not all molecules can be transferred in the gas phase at sufficient density (vapor pressure) so as to absorb a significant fraction of the photons. But a key advantage of action spectroscopy is that the action can be selected in order to be specific to a molecule or a conformational state of a given molecule.

5.2.1 Absorption Spectroscopy

The first gas-phase spectra of nucleobases were actually the vapor phase absorption spectra published by Tinoco and coworkers in the 1960s [3]. The nucleobase solid samples were heated so as to produce sufficient gas density, and absorbance was measured. Example spectra are shown in Fig. 5.3. Figure 5.3b shows interesting features in the vapor phase spectrum of 9-methylhypoxanthine, an analogue of guanine, compared to the aqueous solution spectrum: the absorption maximum is blue-shifted in the gas phase, and vibrationally resolved features appear between 300 and 260 nm. These vapor spectral data were used for a very long time as benchmarks for calculating electronic excited states of neutral nucleobases. There are several problems however with the absorption vapor spectra: some molecules like guanine show signs of decomposition (e.g., ammonia characteristic bands) when transferred into the gas phase by heating. The exact aggregation state of the bases in the gas phase is uncertain (are all of them isolated monomers), and the nature of the tautomers (isomers differing in proton position) present in the gas phase cannot be assigned. Action spectroscopy schemes were later developed for the probing of neutral bases transferred to the gas phase with softer conditions (e.g., laser ablation and supersonic expansion cooling, see Chap. 2).

5.2.2 Action Spectroscopy

To cope with the low density of neutrals or ions in the gas phase, lasers or synchrotrons must be used in order to have intense and spatially focused light sources. Several actions can follow electronic excitation due to UV light absorption of a chromophore (Fig. 5.4) [4]. As introduced above, light emission (fluorescence) is one of the possible actions. Natural nucleobases, however, are very little fluorescent in solution (the quantum yield is in the order of 10^{-4}) [5], and direct fluorescence has not been detected in the gas phase either. However, some modified nucleobases such as 2-aminopurine are fluorescent in solution, or one can envisage attaching other fluorescent chromophores to oligonucleotides to develop fluorescence detection schemes.

Other possible results of electronic excitation are direct reaction from the electronically excited state (in orange in Fig. 5.4) or internal conversion of the
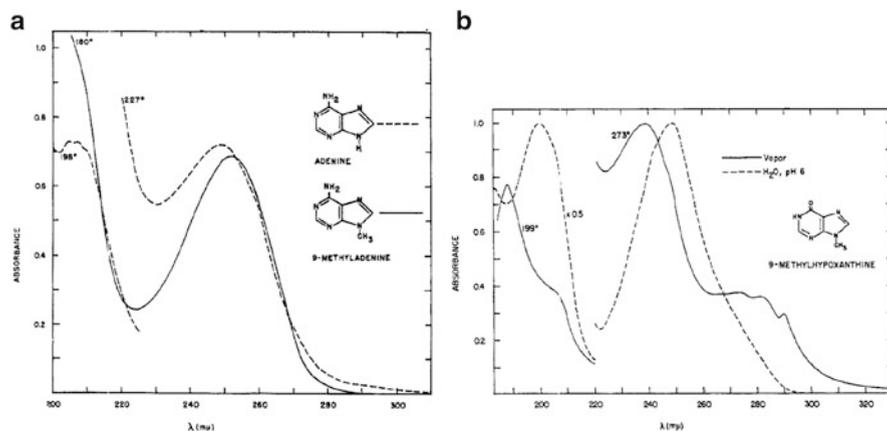
Fig. 5.3 Vapor phase absorption spectroscopy of nucleic bases (reprinted with permission from [3])
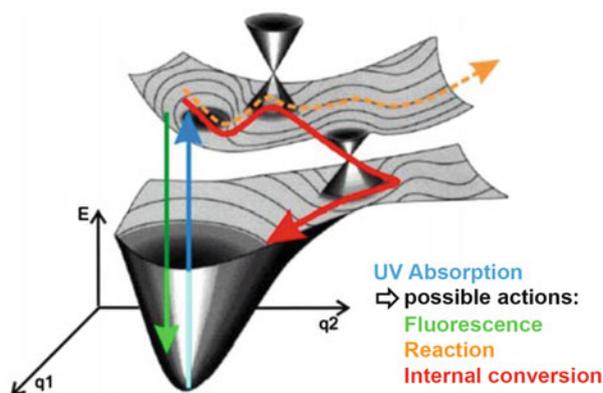

Fig. 5.4 Potential energy surfaces of electronic ground state (S_0) and excited state (S_1) of a neutral or ionic molecule: possible actions following excitation from S_0 to S_1 by UV photon absorption (adapted with permission from [4])

electronically excited state to a vibrationally hot electronically ground state. As shown in red in Fig. 5.4, internal conversion (IC) from one electronic state to another involves a conical intersection, and hence a specific change of geometry. The internal conversion therefore results in the conversion of the energy (in the order of 4–5 eV for nucleobase electronic excitation) in a few selected vibrational normal modes. An important difference between gas phase and solution is that isolated molecules in the gas phase do not have a bath in which this extra energy can be dissipated. The dissipation must therefore be entirely internal, and the process is called IVR (intramolecular vibrational energy redistribution) for the statistical redistribution of the vibrational energy over all vibrational normal modes.
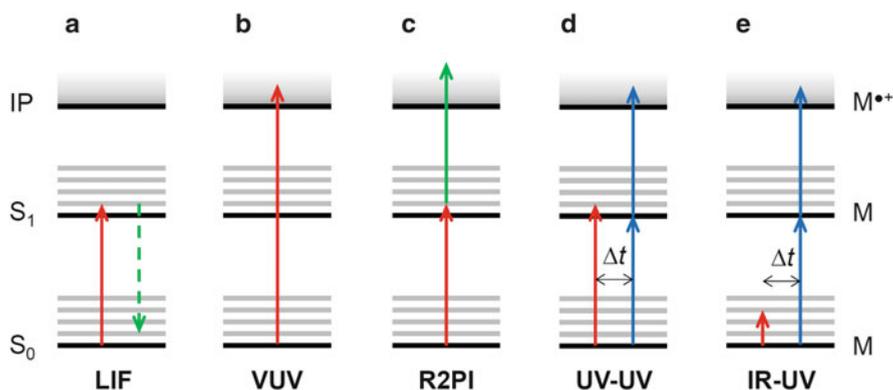
Fig. 5.5 Action spectroscopy of neutrals: typical experimental schemes: (**a**) laser-induced fluorescence, (**b**) direct photoionization with VUV light, (**c**) resonance-enhanced two-photon ionization, (**d**) UV-UV double resonance spectroscopy, (**e**) IR-UV double resonance spectroscopy (adapted with permission from [6])

5.2.2.1 Action Spectroscopy of Neutrals

Action spectroscopy of neutrals (Fig. 5.5) is particular in that, except for laser-induced fluorescence (LIF), where the detected signal is a photon, most action spectroscopy schemes involve the detection of ions, and hence ionization of the molecule is the action. The ionization potential (IP) must be reached. Mass selection of the ion then provides molecular specificity based on the molecular weight of the detected ion. For example, this ensures that the signal of degraded molecules does not contribute to the action spectrum.

Direct Photoionization (VUV)

The ionization potential can be reached by providing a single photon of sufficient energy for the molecule to switch from the electronic ground state directly in the continuum of ionization. The frequency-resolved spectra therefore present a threshold in the ion intensity, and from this threshold the ionization potential (IP) is deduced. For nucleic acid neutral bases, the lowest ionization potential (guanine) is around 8.0 eV. The tunable light sources are synchrotrons, and this range of the ultraviolet spectrum is called VUV (vacuum UV) because air would absorb the photon energy and the entire experimental setup must be under vacuum. Typical direct photoionization VUV data of neutral bases are illustrated in Fig. 5.6, along with the IP values for the four DNA bases, bare and microhydrated (i.e., one or several discrete water molecules attached).

Resonance-Enhanced Multiphoton Ionization (Detection of Ions)

In order to achieve ionization with tabletop lasers, for example, with a frequency-quadrupled Nd-YAG laser at 266 nm (4.66 eV photon energy), the energy of at least two photons is necessary to reach the ionization potential. This is called two-photon ionization or, more generally, multiphoton ionization. There is no strict need that
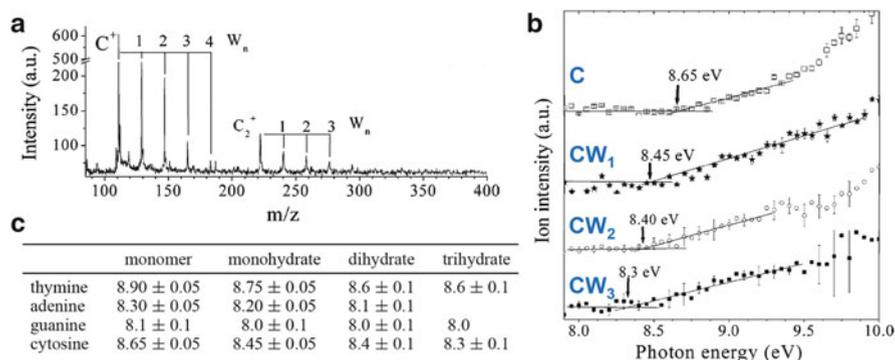
Fig. 5.6 Direct photoionization VUV spectroscopy of neutral bases, bare and hydrated. (**a**) Mass spectrum of ions detected following 10 eV ionization. (**b**) Frequency-resolved spectra, showing that the ionization threshold changes upon hydration. (**c**) Ionization potentials of bare and hydrated DNA bases (adapted with permission from [7])

each absorbed photon reaches an existing quantized state for multiphoton ionization to occur. However, when the first photon absorption is on-resonance with an allowed state, the ionization yield is larger. This process is called resonance-enhanced two-photon ionization (R2PI) or resonance-enhanced multiphoton ionization (REMPI). One can therefore record a frequency-resolved spectrum by scanning a UV laser (in red in Fig. 5.5c) provided that a second photon is simultaneously available for ionization to occur. This multiphoton process requires bright light sources. A disadvantage is that, for converting the action efficiency into a pseudo-absorption efficiency, the number of photons responsible for ionization must be known. This is not always easy to determine, as the number of photons involved in the multiphoton processes can vary from one system to the other or with the laser power.

Double Resonance Spectroscopies, Also Called "Hole-Burning" or "Ion-Dip" Spectroscopies

Double resonance spectroscopy schemes have been developed to obtain single-photon action spectra, either in the UV or in the IR domain. The action monitored is again ionized by R2PI. This probe laser (blue in Fig. 5.5d, e) is parked at a wavelength that resonates with a particular electronically excited state. For cold ions with a lifetime longer than nanoseconds, the R2PI action spectrum can be vibrationally resolved, and proper laser wavelength choice enables the experimentalist to select a single conformer. This feature, coupled to mass selection of the ion, is a unique way to obtain specific spectroscopic data for each component of the mixture.

The probe laser power being stabilized, the ion signal intensity will reflect the amount of molecule or conformer present in the chamber at the electronically ground state (S_0). Therefore if, before the probe laser, another laser (in red in Fig. 5.5d, e) is fired and depletes the electronically ground state due to a transition

to an electronically excited state (Fig. 5.5d) or to a vibrationally excited state (Fig. 5.5e), the ion signal intensity will decrease. Scanning the frequency of the probe UV or IR laser and monitoring the magnitude of the action due to another fixed-frequency laser (e.g., via an R2PI mechanism) is the double resonance experiment.

For those readers familiar with Fourier transform ion cyclotron resonance mass spectrometry (FTICRMS) or with NMR spectroscopy, the term "double resonance" has exactly the same meaning: one frequency depletes the signal by on-resonance depletion or ejection, while the probe signal is taken care of by other resonant frequencies. In the gas-phase spectroscopy jargon, because of the depletion scheme, some authors call it a "hole-burning" experiment, and because an ion signal is depleted, the term "ion-dip" spectroscopy is also used.

5.2.2.2 Action Spectroscopy of Ions

When performing action spectroscopy on ions instead of on neutrals, a main difference is that mass selection can be performed before interrogation of the system by a laser. In contrast, for the spectroscopy of neutrals, detection of an ion of a given mass only means that it was a constituent of the interrogated neutrals, but the detection of a monomer ion can also result from the ionization and then the fragmentation of an aggregate. The second difference is that the energy required for electron ejection changes with the charge state of the interrogated molecule: the ionization energy increases for positive ions compared to neutrals and decreases when the negative charge increases. The adiabatic electron detachment thresholds for monoanions have been determined accurately for singly charged mono, di, and trinucleotides [8]. Multiply charged anions can even become metastable in the gas phase with regard to electron loss, when each phosphate is deprotonated [9, 10]. As a consequence, as illustrated in Fig. 5.7, the possible action spectroscopy schemes may differ from the typical ones used for neutral molecules. For positive ions (Fig. 5.7a), fragmentation is usually observed, whether resulting from excited state reaction pathways or following IC and IVR. For singly charged anions (Fig. 5.7b), fragmentation is usually the dominant pathway, but for some of them (particularly in the case of nucleic acids containing guanines, the base with the lowest ionization potential [8]) electron detachment might also contribute. For multiply charged anions (Fig. 5.7c), electron detachment is often the dominant decay pathway.

Double Resonance Experiments

Double resonance experiments similar to those described above for neutrals are also possible for ions, the only difference being that the action is not caused by R2PI, but can be for example a direct fragmentation from electronically excited states (Fig. 5.7a, orange arrow). High-resolution IR spectroscopy reports on cold peptide and protein cations have provided a brilliant demonstration of conformer-resolved gas-phase vibrational spectroscopy is possible even on large ions [11–13]. Unfortunately, similar experiments could not yet be successfully achieved for nucleic acid ions (Rosu, Gabelica, Zabuga, Stearns, Guidi, Boyarkin and Rizzo, unpublished
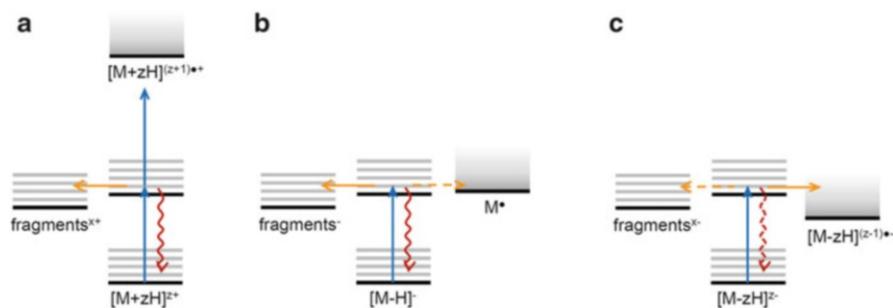
Fig. 5.7   Action spectroscopy of ions: the possibility of electron detachment depends on the ion charge, as the ionization energy decreases as anion charge increases: (**a**) cations, (**b**) monoanion, (**c**) polyanions. In *blue*: UV absorption; in *orange*: reaction from the electronic excited state; in *red*: internal conversion

results). To this date, all electronic spectra were found to be broad even for cold ions. A possible reason might be that the electronic excited states are too short-lived to be energy-resolved, a manifestation of the Heisenberg uncertainty principle. As discussed in the results, this phenomenon was observed for many neutral nucleic acid systems as well.

Electron Detachment from Anions

For anions, a particular action that can be observed even after excitation with a single UV photon is electron detachment. For monoanions, the electron detachment thresholds lie in the same range as the first electronically excited states (see section "From Oligonucleotides to Biologically Relevant Higher-Order Structures" below Sect. 5.3.2.2). For polyanions, electron detachment is strongly favored. Interestingly, in contrast to the direct VUV photoionization spectra shown in Fig. 5.6, the frequency-resolved electron detachment action spectra show resonances at particular frequencies matching expected transitions to base electronically excited states. For polyanions, the process can therefore be coined resonance-enhanced single-photon ionization (R1PI). Photodetachment would be a more proper term for anions than photoionization but the acronym "PD" could be confounded with "photodissociation." Because the process is a single-photon one, normalization of the action with regard to photon flux is linear.

Resonant Multiple Photon Dissociation (UVMPD and IRMPD)

Following electronic excitation, the electronic excited state can convert back to the electronic ground state by internal conversion (in red in Fig. 5.7). This vibrationally hot electronic ground state may be hot enough for the molecule or ion to fragment on the time scale of the experiment. Multiple cycles of resonant UV absorption, IC and IVR, may be necessary for the ion or molecule to fragment to a detectable extent. In such case, the process is called UVMPD (ultraviolet multiple photon dissociation). Here, we specifically use the expression "multiple photon" instead of "multiphoton" to indicate that the photons need not be absorbed in a correlated

manner as in REMPI. UVMPD energy accumulation can be sequential and even take a significant amount of time, as the energy is not dissipated in vacuum except by collisions with residual gas (for that reason, the pressure inside the mass spectrometer can influence the UVMPD efficiency). UVMPD monitored as a function of the UV laser frequency gives an action spectrum.

Similarly, if the primary resonant activation is infrared-induced excitation of a specific vibrational normal mode, several cycles of resonant IR absorption and IVR are often necessary to cause dissociation, and the process is called IRMPD (infrared multiple photon dissociation). The only experimental setup where a single IR photon could be sufficient to cause a detectable action on an ion is when very low-energy bound reporter neutrals (argon atom or a helium cluster from a helium droplet) would be attached to the ion to be probed. Such experimental scheme has been demonstrated for crown ethers [14], small multidentate complexes [15] or proteins [16], and neutral nucleic acid bases [17], but not yet for nucleic acid ions. On the contrary to IR spectra obtained by single-IR photon probing (such as IR-UV double resonance experiments), the relative intensities in the IR spectra obtained by multiple photon processes (IRMPD) are not necessarily well reproduced by oscillatory strength calculations of the theoretical vibrational spectra.

5.3   Results

5.3.1   Neutral Nucleic Acid Bases and Building Blocks

5.3.1.1 Single Bases

The nature of the tautomers [18] observed in the gas phase has been the subject of intense discussion over the past decade, particularly in the case of guanine tautomers (Table 5.2). The most stable tautomer in solution is the keto-N9H. The first IR-UV double resonance experiments on jet-cooled guanine [19] were assigned to four tautomers (keto-N9H, keto-N7H, and the two rotamers of enol-N9H). These four tautomers were also observed in helium nanodroplet experiments [17]. These tautomers correspond to the calculated most stable structures. The controversy started with the reassignment of IR-UV double resonance experiments (REMPI action spectroscopy) in 2006 by Mons and coworkers [20], who revealed the presence of other higher-energy tautomers (enol-N9H-trans, enol-N7H, and two rotamers of the keto-N7H-imino tautomers). More striking was the demonstration that the predicted most stable tautomers were actually undetected in the IR-UV spectroscopy experiments. Mid-IR (400–1,800 cm^{-1}) IR-UV double resonance spectroscopy experiments on 9-Ethyl Guanine, Guanosine, and 2-Deoxyguanosine also confirmed the presence of the enol form for all structures, indicated by the absence of C=O stretching normal mode that should have appeared at 1,800 cm^{-1} for free carbonyl and no lower than 1,680 cm^{-1} for hydrogen-bonded carbonyl [21].

The failure to observe the most stable keto tautomers was later rationalized by the fact that the excited state lifetimes of these tautomers were actually too short

Table 5.2 Guanine tautomers nomenclature, structure, relative energy, and observation in different experimental setups (X indicates the tautomer was observed)

	Keto-N7H	Enol-N9H	Enol-N9H-*trans*	Enol-N9H-*cis*	Enol-N7H-*trans*	Imine 7H-I_d-O	Imine 7H-I_u-O
Structure [24]							
ΔE_{MP2} (cm^{-1}) [24]	0	100	141	347	1,218	2,682	2,723
IR in He [17]	X	X	X	X			
IR-UV REMPI [20]			X		X	X	X
Microwave [24]	X	X	X	X			

(femtosecond time scale) for them to be detected by the R2PI action spectroscopy scheme. In other words, this type of action spectroscopy is blind to conformers that have a lifetime much shorter than the laser pulse width (nanoseconds). This has prompted intense research on excited states dynamics of DNA bases (for reviews, see [4, 22, 23]), and theoretical calculations pointed to a very efficient relaxation by internal conversion (IC). The observation that the natural tautomers of guanine are the most photostable (internal conversion is preventing photodamage, i.e., photo-ionization and reactions from the electronically excited states, see Fig. 5.4) has also led to interesting speculations on the appropriateness of photostable nucleobase tautomers as the molecules are in charge of being the repository of genetic information (see the section below on base pairs).

The guanine tautomer story illustrates one of the limitations of action spectroscopy: not all "actions" are appropriate for probing all molecules or all conformers (tautomers) of each molecule. Absorption spectroscopy is always more likely to reflect the composition of the mixture. Alonso and coworkers have developed microwave spectroscopy as an alternative high-resolution spectroscopy probing method, which is insensitive to excited state dynamics. The assignment of microwave spectra is based on the very accurate and precise measurement of the three principal moments of inertia of the neutral molecule, and on the characterization of the fine structure due to quadrupolar couplings with nitrogen (^{14}N) nuclear spin. Using their laser ablation—molecular beam—Fourier transform microwave (LA-MB-FTMW) setup, they confirmed the presence of the four lowest-energy tautomers of guanine [24] (Table 5.2), five tautomers for cytosine [25] (Fig. 5.8), and that thymine [26] and uracil [27] were present in their diketo forms.

5.3.1.2 Base Pairs

Clusters of molecules can also be produced in supersonic expansion, and ions with a mass corresponding to the sum of two molecules can be detected by REMPI action spectroscopy, indicating the existence of these neutral pairs in the beam. Many such base pairs or dimers have been investigated, both experimentally [28] and
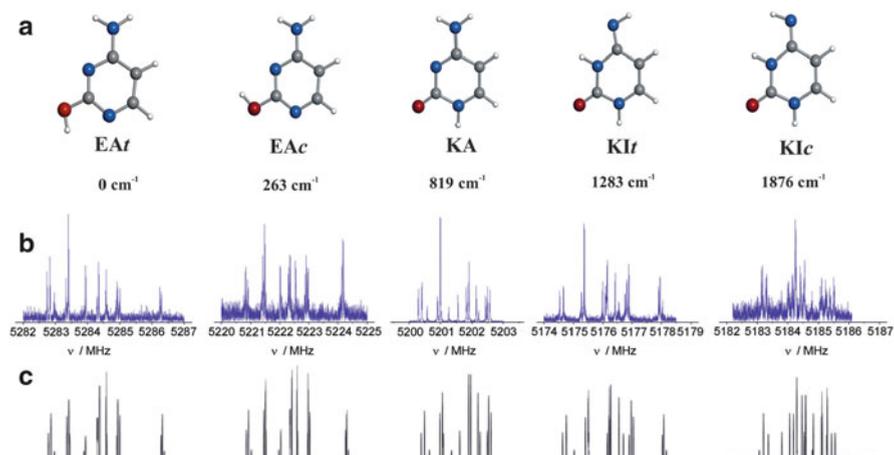
Fig. 5.8 Microwave spectroscopy of neutral cytosine, demonstrating the existence of five different tautomers. (**a**) Tautomer structures: Eat = enol-amino-*trans*, EAc = enol-amino-*cis*, KA = keto-amino = the natural tautomer, KIt = keto-imino-*trans*, KIc = keto-imino-*cis*, (**b**) experimental LA-MB-FTMW spectra, and (**c**) simulated spectra (reprinted with permission from [25])

theoretically [29, 30]. Hydrogen-bonded base pairs are usually observed, as opposed to stacked base pairs (with the exception of methylated adenine dimers, where the methylation blocks the hydrogen bonding sites of the lowest energy H-bonded structures) [31]. The first REMPI investigation concerned a GC base pair [32], and it was concluded that the experimentally observed base pair is not in the Watson–Crick geometry. In the case of GC base pairs, among the 50 possible geometries, the biologically prevalent Watson–Crick geometry is the lowest-energy structure. However, it has a subpicosecond lifetime, preventing the observation of a sharp R2PI spectrum (see Fig. 5.9a, for a methyl-modified base pair that restricts the hydrogen bonding possibilities to the Watson–Crick arrangement), whereas other base pairing arrangements show a sharp R2PI spectrum (see Fig. 5.9b, c). Therefore, as in the case of the natural tautomer of guanine, the natural GC base pair shows remarkably high photostability, and the internal conversion pathways have been the subject of intense debate (e.g., see [33]). In the case of AT base pairs, the Watson–Crick base pair is not the lowest energy structure, as shown both by theory and experiment [34]. The DNA backbone therefore favors the Watson–Crick AT base pair arrangement.
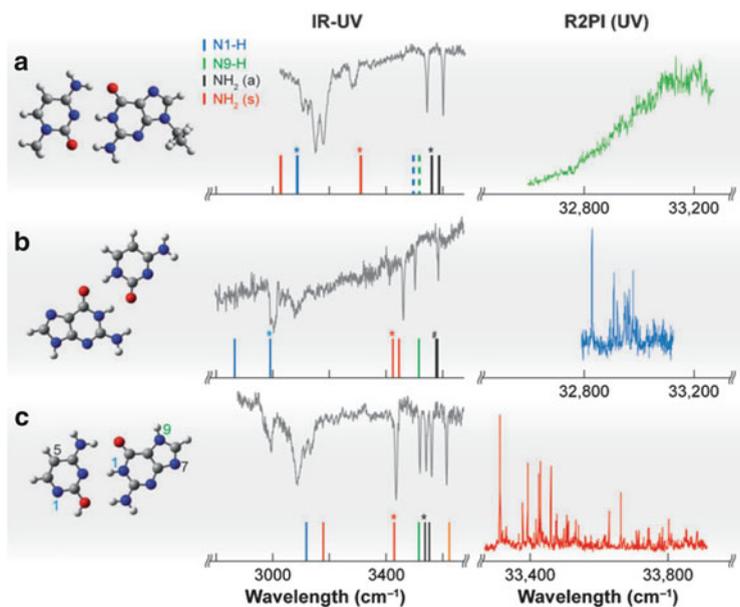
Fig. 5.9 IR-UV double resonance spectroscopy and UV R2PI for isolated guanine–cytosine (GC) base pairs. (**a**) Results for the Watson–Crick structure, (**b**, **c**) the second and third lowest energy structures, respectively. The *second column* shows the experimental IR-UV double resonance experiments together with the calculated IR bands. *Asterisks* represent bands from the guanine moiety. The *third column* shows the UV excitation spectra measured by R2PI (reprinted with permission from [6])

5.3.2   Nucleic Acid Ions

5.3.2.1 Vibrational Spectroscopy of Ions (IRMPD Spectroscopy)

In the case of ions, the usual experimental scheme is IRMPD action spectroscopy, either using tabletop OPA lasers (2,800–3,600 cm^{-1}) or Free Electron Lasers (~400–2,000 cm^{-1}) at facilities such as CLIO (Centre Laser Infrarouge d'Orsay, http://clio.lcp.u-psud.fr, Université Paris Sud, Orsay, France) or Free-Electron Lasers for Infrared eXperiments (FELIX, http://www.ru.nl/felix/, Radboud University, Nijmegen, The Netherlands), which have free electron lasers coupled to FTICRMS (both facilities) or to a quadruple ion trap mass spectrometer (CLIO facility). The ions are prepared from solution using electrospray ionization (see Chap. 2) and transferred to the FTICR trap or to the quadrupole ion trap, where they are stored and mass-selected. IR spectra are obtained by monitoring ion fragmentation as a function of the laser frequency. The fragmentation pathways will be not discussed here (IRMPD fragmentation pathways will be covered in Chap. 6). In the IR action spectra, the band positions do not depend on the monitored pathway. However, the band intensities can vary depending on the absorption efficiency,

especially when consecutive fragmentation pathways are involved. Common practice is to sum all fragments to reconstruct the IRMPD action spectrum.

Numerous IRMPD studies of electrosprayed DNA-related ions have been published to date. Our goal here is not to make a comprehensive review, but to outline how the spectral features and structures change from nucleobases to larger structures. Citations in this chapter were therefore restricted to natural nucleobases and to protonated/deprotonated structures, but we note that a large body of literature exists on cationized nucleic acids and on modified bases. Full list of references can be retrieved from the facilities' respective Web sites.

Single-Base Cations

The first IRMPD spectra of electrosprayed nucleic acids were published by Maitre and coworkers, on the protonated pyrimidine nucleobases (see Fig. 5.10) [35]. The mid-IR range allowed the authors to evaluate the presence of keto conformers, which should give a $C=O$ absorption band around $1{,}800 \text{ cm}^{-1}$. The natural forms of the nucleobases in solution and in the solid are the keto forms. In contrast, the major species observed in the protonated nucleobase ions was the enol form, although a small band around $1{,}800 \text{ cm}^{-1}$ could indicate the presence of a second conformation with lesser abundance. In follow-up studies on uracil [36] and cytosine [37], the authors confirmed that the two tautomers originate from the electrospray process, because they cannot result from interconversion in the gas phase. Among protonated base pairs, particularly interesting on a biological relevance point of view, is the study by Rodgers and coworkers on protonated cytosine dimers, which were found to form the same base pairing motif as the i-motif structure [38]. However, when moving to larger species (nucleotides and oligonucleotides), most studies are carried out on anions issued from the deprotonation of phosphates.

Single-Base Anions

Figure 5.11 shows the IRMPD action spectra of the four natural deprotonated DNA mononucleotides [39] and RNA mononucleotides [40]. The DNA and RNA mononucleotide IRMPD spectra are very similar to each other. The region between 650 and $1{,}400 \text{ cm}^{-1}$ are primarily corresponding to the vibrational signature of phosphate and sugar modes, and the region between 1,400 and $1{,}800 \text{ cm}^{-1}$ contains the base-specific signature. Importantly, in the electrosprayed deprotonated mononucleotides the base tautomers correspond to those naturally present in solution, as indicated by keto bands around $1{,}700 \text{ cm}^{-1}$ for guanine, cytosine, and thymine, but not for adenine, and by NH_2 scissoring bands around 1,600—$1{,}650 \text{ cm}^{-1}$ for adenine, guanine, and cytosine, but not thymine. This contrasts with the laser ablated and jet-cooled neutral deoxyguanosine or with the electrospray protonated purine bases described above. In terms of conformation, the fact that the carbonyl stretching bands appear around $1{,}700 \text{ cm}^{-1}$ and not $1{,}800 \text{ cm}^{-1}$ indicates that carbonyl groups are involved in hydrogen bonding and therefore that the base interacts with the phosphate-sugar backbone even for the mononucleotides. Adenine, cytosine, thymine/uracil favor the *anti*-glycosidic bond angle, whereas guanine prefers the *syn* conformation (for structural nomenclature, see Chap. 1). The
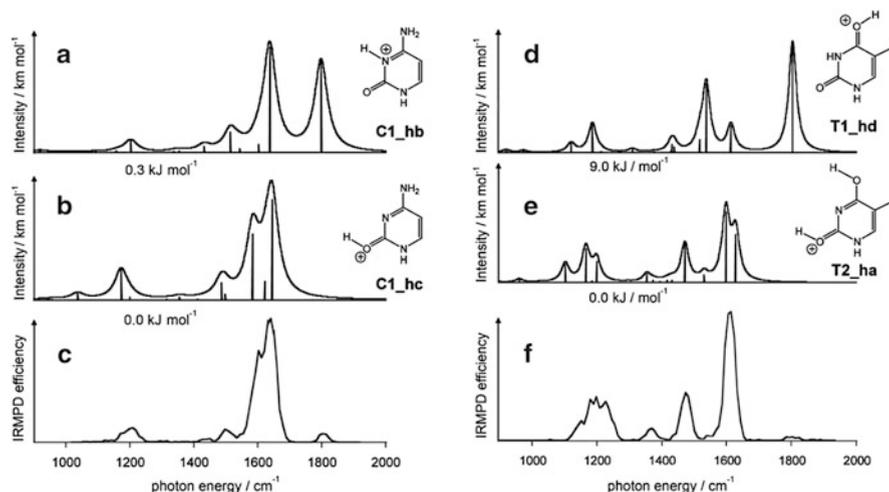
Fig. 5.10 IRMPD spectroscopy of protonated cytosine and thymine, demonstrating the predominance of enol tautomers and a minor keto tautomer. (**a**, **b**) theoretical spectra for cytosine keto and enol tautomer, (**c**) experimental spectrum for cytosine, (**d**, **e**) theoretical spectra for thymine keto and enol tautomer, (**f**) experimental spectrum for thymine (adapted with permission from [35])

ribose adopts a C3′ endo conformation in all cases, and the IRMPD spectra indicate that several distinct conformers are likely to be accessed in the experimental conditions.

From Oligonucleotides to Biologically Relevant Higher-Order Structures

Hydrogen bonding is one of the driving forces for the formation of higher-order structures: base pairs are formed in double helices (Watson–Crick pairing between A and T/U, G and C) or i-motifs (proton-bound C–H$^+$–C base pairs), proton-bound triplets of bases are formed in triplex DNA, and guanine quartets are formed in G-quadruplexes (for more detailed introduction on the different structures, see Chap. 1). Infrared spectroscopy is the ideal method to probe hydrogen bonding in the gas phase. To this end, we investigated the base vibrational signature region (1,500–1,800 cm^{-1}) in more detail for homo-base single strands and for mixed-base higher-order structures.

The homo-base single strands (Fig. 5.12a, b, e, f) show very similar signatures to those of the corresponding DNA mononucleotides in the cases of adenine, thymine, and guanine oligonucleotides. The triply charged 6-mer oligonucleotides are shown here, and similar spectral features were found for 2- to 5-charge states. The only considerable difference between mononucleotides and oligonucleotides is in the case of cytosine, where all bands merge between 1,640 and 1,670 cm^{-1} in the case of the oligonucleotide. Molecular modeling of several possible low-energy conformers of dC$_6^{3-}$ reveals that cytosines are most often involved in hydrogen bonding with other cytosine bases or with the sugar-phosphate backbone (see supporting information of [41]).

Fig. 5.11 Comparison of the measured IRMPD action spectra of [dX5′p-H]⁻ (DNA, *red*) and [X5′p-H]⁻ (RNA, *blue*). Major differences between the analogous spectra are highlighted (reprinted with permission from [40])
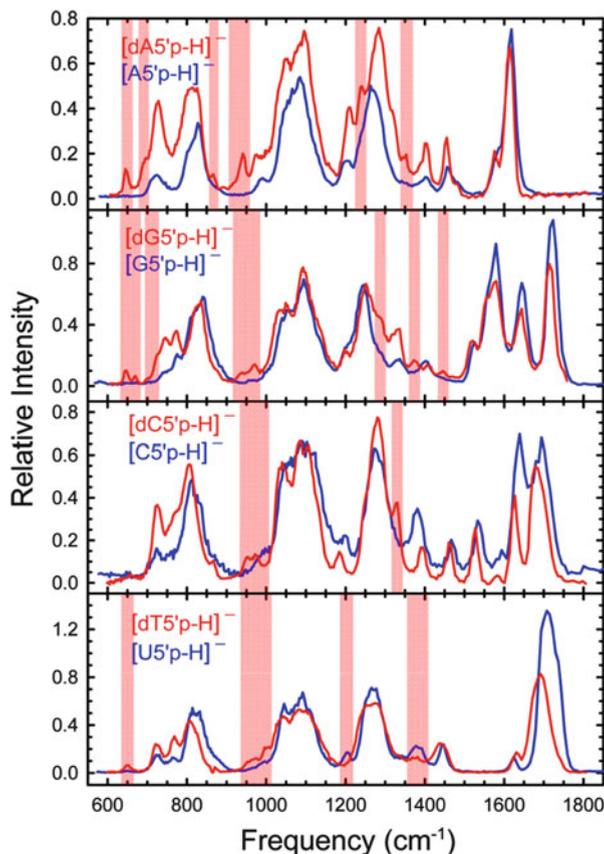
Clearly, the bands corresponding to each type of base are distinct enough in single strands in the gas phase, so as to detect major changes in the hydrogen bonding pattern when forming higher-order structures. For example, when the single strand dC_6 forms a tetrameric i-motif structure in which most cytosines are engaged into proton-bound base pairs, the IRMPD action spectrum changes dramatically, with vibrational absorption bands being red-shifted up to $1,750$ cm^{-1} [41].

Limitations in probing hydrogen bonding in higher-order structures however arise when the studied oligonucleotides involve several different bases simultaneously. In the case of G-quadruplexes stabilized by inner ammonium cations and in which all guanines are involved in quartet formation, significant red-shift of the guanine carbonyl band was detected, from ~1,720 to 1,685 or 1,695 cm^{-1}, a region that is mostly blank for cytosine-free strands [42]. As discussed in detail in Chap. 3, the stability of gas-phase structure ranks in the following order: cation-bound G-quadruplexes > proton-bound triplexes (and i-motifs) > hydrogen-bound base pairs. The preservation of G-quadruplexes and i-motifs in the gas phase being
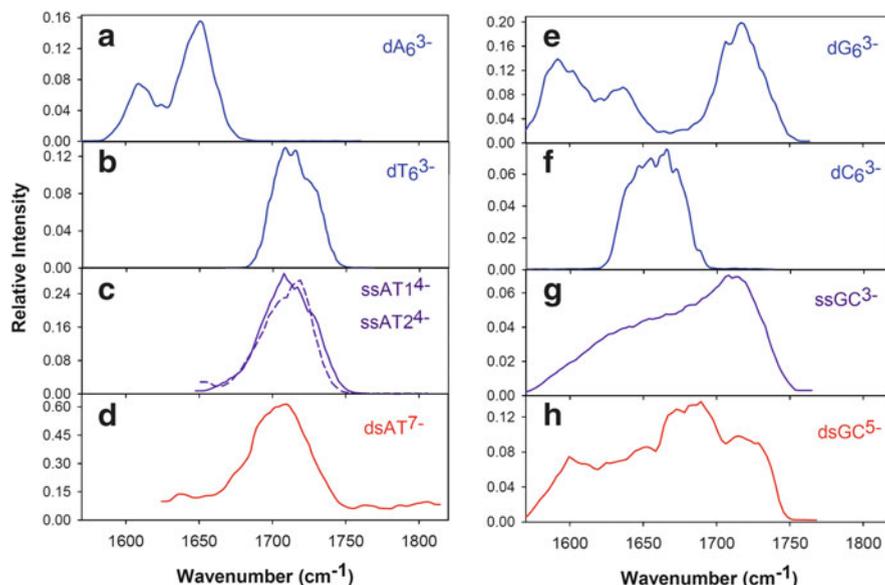
Fig. 5.12 IRMPD spectroscopy of electrosprayed, multiply deprotonated DNA single strands and duplexes (Gabelica and Rosu, previously unpublished results obtained at the CLIO facility): (**a**) adenine oligonucleotide dA_6^{3-}; (**b**) thymine oligonucleotide dT_6^{3-}; (**c**) AT-only single strands $ssAT_1^{4-}$ (sequence: dAAATTATAATATTAAA) and $ssAT_2^{4-}$ (sequence: dTTTAATAT-TATAATTT); (**d**) duplex $dsAT^{7-}$ constituted by $ssAT_1$ and $ssAT_2$; (**e**) guanine oligonucleotide dG_6^{3-}; (**f**) cytosine oligonucleotide dC_6^{3-}; (**g**) GC-only single strand $ssGC^{3-}$ (sequence: dCGCGGGCCCGCG); (**h**) self-complementary duplex $dsGC^{5-}$ (same sequence)

established, our next goal was therefore to probe whether Watson–Crick base pairs can be identified in oligonucleotide dimers by IRMPD action spectroscopy. Reference spectra were those of the constituting single strands (Fig. 5.12c for AT-strands and Fig. 5.12g for GC-strands). The GC-only duplex spectrum (Fig. 5.12h) is markedly different from the single strand, with a specific band around 1,680 cm^{-1} which could be due to the Watson–Crick base pairing. However, the AT-only duplex spectrum (Fig. 5.12d) was not markedly different from those of the constituting single strands. In conclusion, the preservation of hydrogen bonding in gas-phase GC-rich duplexes could be confirmed, but there is much less clear evidence of preservation of hydrogen bonding in gas-phase AT-rich duplexes.

5.3.2.2 Electronic Spectroscopy of Nucleic Acid Ions

Isolated Bases and Mononucleotides

Protonated adenine $AdeH^+$, produced either by electrospray ionization [43, 44] or by ionization of cold neutral adenine dimers [45], was investigated by action spectroscopy in the UV region. The action is neutral loss, resulting in charged fragments that were detected by a mass analyzer. At low laser power, a single photon was sufficient to cause fragmentation [44]. The lowest-energy transitions of

Fig. 5.13 UV spectroscopy
of protonated adenine (*blue
circles*), hydrated adenine
(*black full dots*), and
protonated adenine
monophosphate (*red
diamonds*). The gas-phase
spectra are within
experimental error similar to
the solution phase spectrum
of protonated adenine at low
pH (*green line*).
Figure reprinted with
permission from [44])
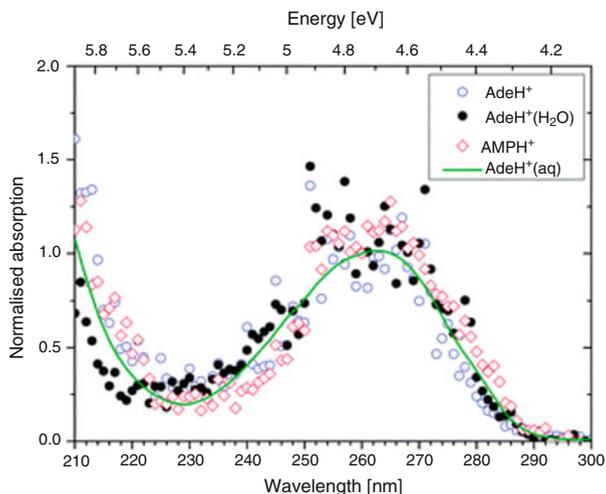
protonated adenine, monohydrated protonated adenine, and protonated adenine
monophosphate produced by electrospray ionization seem identical in the gas
phase as in low-pH solution, indicating that the lowest-energy transitions are
independent on the surroundings (Fig. 5.13). The lowest-energy tautomer produced
by electrospray is the same as the one existing in solution (protonation on adenine's
N1). In contrast, adenine cations formed by dissociation of the dimers have action
spectra shifted to the red, and this was attributed to the formation of a higher-energy
tautomer (protonation on adenine's N3) [45].

At neutral pH in solution, the nucleobases are not protonated and the phosphate
groups are deprotonated. The deprotonated DNA, RNA, and cyclic mono-
nucleotides, ionized by electrospray in the negative mode, were investigated by
Weber and coworkers [46, 47]. The monitored action was fragmentation of the
anions: the nature of the fragments was similar to those observable by collision-
induced dissociation (CID) or IRMPD, so it is possible that electronic excitation is
followed by internal conversion (IC) and IVR to populate a hot electronic ground
state. Nevertheless, as the relative fragment abundance differed from CID and
IRMPD, fragmentation before complete IVR, either from the electronic excited
state or after IC, is not excluded either. The action spectra were however found
mostly independent on the nature of the fragment monitored. The action spectrum
obtained by monitoring the depletion of parent ion with respect to all fragments is
therefore very likely to reflect an absorption spectrum.

Action spectra are shown in Fig. 5.14 for all deprotonated deoxynucleotides,
together with the solution absorption data. For adenine and guanine, the RNA and
cyclic mononucleotides [47] gave virtually similar gas-phase action spectra as those
shown in Fig. 5.14, indicating that the base excited states are not significantly
influenced by the sugar and phosphate backbone. For guanine, no solvatochromic
shift was observed. However, the electronic transitions are clearly red-shifted for
cytosine, and blue-shifted for adenine. For thymine, a second band appears on the
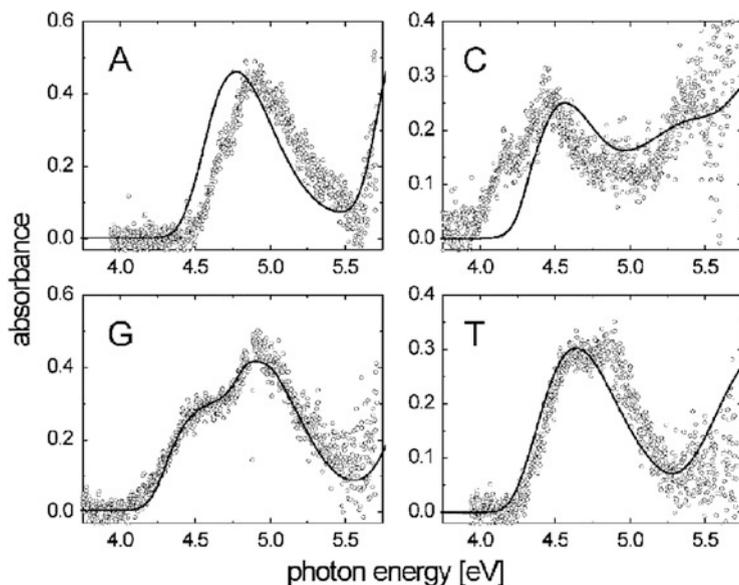
Fig. 5.14 Comparison between parent ion depletion spectra (*open circles*) and aqueous solution absorption spectra (*solid line*) of the mononucleotides under study. Depletion spectra are for parent ions of the form (dBMP-H)⁻ (the base is indicated in each panel). Aqueous absorption spectra are for 70 mM solutions of the disodium salts of the nucleotides. Depletion data (in arbitrary units) have been normalized to the peaks of the aqueous solution absorption spectra in each case for easier comparison (Reprinted with permission from [46])

blue side of the main band in the gas phase. Notably, the action spectra on deprotonated mononucleotide anions match the absorption vapor spectra of neutral bases published 45 years before [3] (see Sect. 5.2.1). This validates action spectral data as reflecting absorption spectra.

From Oligonucleotides to Biologically Relevant Higher-Order Structures
In oligonucleotides and other higher-order structures such as duplexes and G-quadruplexes, bases can interact with hydrogen bonding and stacking. For the electronic properties, stacking is particularly crucial because it can lead to the formation of new types of excited states such as exciplexes (coherent delocalized excitation over several bases) of and charge transfer states (net change of charge distribution from one base to another) [4, 5, 23]. Computational studies of those states remain challenging and would greatly benefit from experimental data on well-defined structures as benchmarks. This is one of the objectives of experimental studies on isolated gas-phase oligonucleotides and higher-order structures.

Excited states dynamics of poly(dA) oligonucleotides have been extensively studied both in solution and in silico. Nielsen et al. have published the first experimental study of dA_n^- ions in the gas phase ($n = 2–4$). The main action was fragmentation several microseconds after laser irradiation, as monitored by

Fig. 5.15 Action spectra of dA_n^- anions recorded from the measurements of the yield of neutrals as a function of excitation wavelength. The yield of neutrals divided by the number of photons for $n = 2$, and the number of photons squared for $n = 3$ or 4, is assumed to represent the absorption. (**a**) $n = 2$; (**b**) $n = 3$, and (**c**) $n = 4$. The action spectrum of $dAMP^-$ is shown on each panel (*gray points*, identical to Fig. 5.14 data) (reprinted with permission from [48])
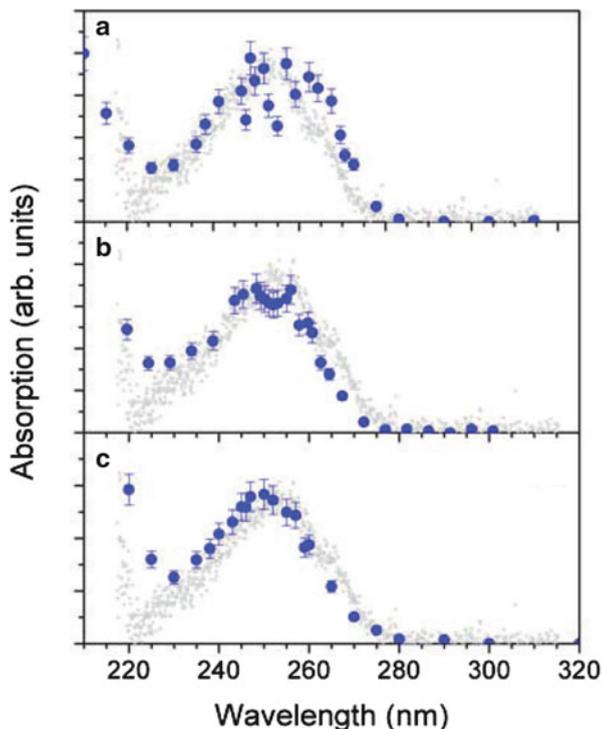
measuring the neutral fragments. dA_2^- ions fragmented with a single photon, whereas dA_3^- and dA_4^- ions required two photons to fragment, via an UVMPD process. The main action is observed at photon energies below the threshold for photodetachment [8].

Remarkably, although the experiments were carried out at room temperature and with no particular conformational selection, they show a clear effect of the number of bases on the action spectra (Fig. 5.15), that is in line with the hypothesis of exciplex formation for strands containing three adenine bases or more. The dA_3^- and dA_4^- action spectra are slightly but significantly blue-shifted compared to that of the isolated base and compared to dA_2^-.

Structures containing more bases, when undergoing electrospray ionization, usually end up as multiply charged anions. In that case, the electron detachment threshold becomes lower than the main DNA excitation bands, so the main action becomes electron detachment and this process requires a single photon [49–51]. Still, resonances are observed (unlike the direct VUV photoionization spectra shown in Fig. 5.6); therefore the electron photodetachment mechanism is believed to be two-step: (1) base excitation followed by (2) electron detachment. Thanks to this efficient action, action spectroscopy can be carried out on large higher-order structures such as duplexes and G-quadruplexes, as shown in Fig. 5.16. Here, the duplex and the G-quadruplex action spectra are compared to those of the single

Fig. 5.16 UV action spectroscopy (action monitored is single-photon electron detachment) of nucleic acid single strands (s.s.) and higher-order structures formed by these strands. (**a**) single strand versus duplex of sequence dCGCGGGCCCGCG. (**b**) Single strand t[dTGGGGT-2H]$^{2-}$ versus tetramolecular G-quadruplex [(dTGGGGT)$_4$•(NH$_4^+$)$_3$-8H]$^{5-}$ of the same sequence (G-quadruplex structure shown in Chap. 1, Fig. 5.5c) (adapted with permission from [51])
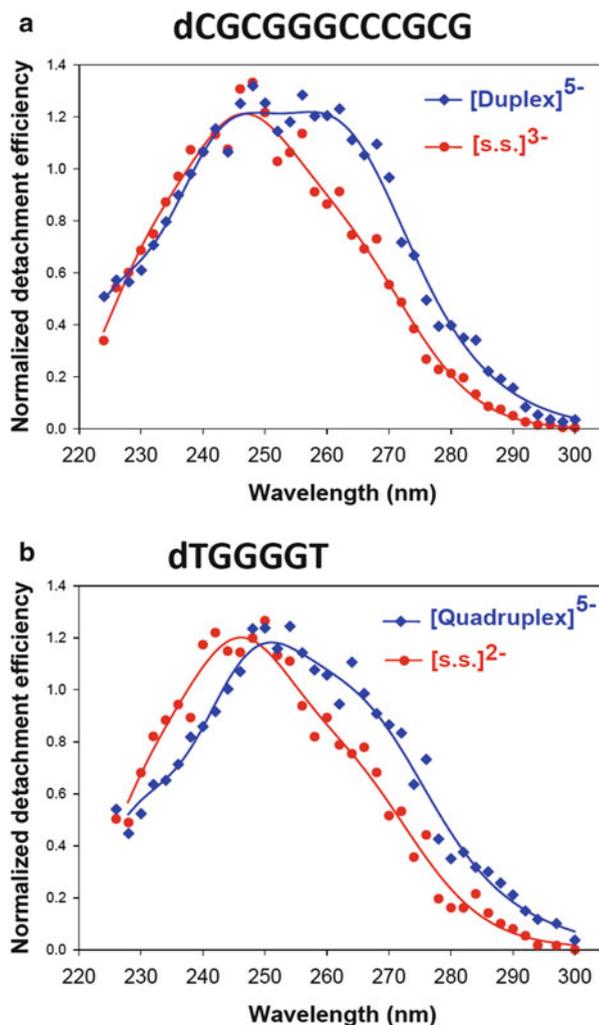
strands with the same sequence. Here the red-shift upon higher-order structure formation is significant. This was attributed to delocalized excited states involving the guanines [51]. Strikingly, the G-quadruplex that has its stacking enforced in the gas phase by the presence of cations between the G-quartets shows the most significant shift, and it is highly likely that the type of excimers in the gas phase are exactly the same as those observed in solution [52].

5.4   Conclusions and Perspectives

In the field of vibrational spectroscopy, harmonic vibrational frequency calculation is now mature enough to interpret IR spectroscopy of nucleobases or mononucleotides, and to predict the most stable conformers, but challenges remain for relative energy and vibrational frequency calculation on large molecules (oligonucleotides) that are conformationally much more flexible. Spectral congestion is also a concern as the number of normal degrees of freedom increases, but each base shows bands at relatively well-defined positions, with significant shifts upon hydrogen bonding, protonation, or cationization. Structural interpretation for larger species should however be based not only on calculations of small model systems but also on carefully selected control experiments, and will nevertheless remain extremely challenging if all types of bases are present in a single sequence. Future cold ion spectroscopy on oligonucleotides could reduce band entanglement, but action spectroscopy on cold multiply charged nucleic acids remains a conceptual and experimental challenge. Alternatively, conformational sorting prior to spectroscopy probing could be carried out using ion mobility spectrometry.

The field of gas-phase electronic spectroscopy was well developed for already 10 years, but mostly with the focus on action spectroscopy of neutrals using double resonance spectroscopy setups. However, the study of the action pathways sparked great interest into the excited states dynamics, with the same questions arising whether the nucleic acids are in solution or in dehydrated environments. Clearly the next challenge now is to probe the effect of higher-order structures, and particularly the effects of different modes of base stacking on electronic excited states energy, nature, and fate. To this aim, gas-phase spectroscopy on mass-charge- and conformation-selected ions could become a method of choice.

The mitigated results of IR and UV spectroscopies in terms of structure characterization in the gas phase are not surprising, given that these methods actually give as poor results for nucleic acid *structure characterization* in solution. Instead, the most widely used solution spectroscopy methods are electronic circular dichroism (CD) spectroscopy for secondary structure characterization and fluorescence spectroscopy for tertiary structure (folding) analysis. Fluorescence ion spectroscopy has been described by several groups [53–55], and although most reports currently concern small molecules, the potential to analyze biomolecules is beyond doubt. For nucleic acids, it is of interest to test modified fluorescent bases, fluorescent ligands, or fluorescent labels that could be used for gas-phase FRET, as suggested in the pioneer work by Danell and Parks [9, 56]. To the best of our knowledge, however, electronic CD ion spectroscopy has never been attempted on biomolecule ions, and using this method to probe base stacking in nucleic acid secondary structures will be one of our highlight projects for the next 5 years.

5.5   **Summary of Key Concepts**

General Concepts
- Gas-phase spectroscopy is usually an "action" spectroscopy: the effect of light on the molecule or ion is monitored as a function of the photon frequency.
- Microwave spectroscopy coupled to highest-level calculations provides the highest resolution structural data but can be applied to small molecules only.
- Vibrational spectroscopy (IR) provides information on structure and is particularly adequate to probe tautomer forms or hydrogen bonding interactions.
- Electronic spectroscopy (UV–Vis) provides information on the electronic environment of the chromophore. Understanding the nature of the action following electronic excitation is however not trivial.
- Different vaporization conditions can lead to different structures, and vibrational cooling is crucial to make high-resolution vibrational spectroscopy experiments possible.
- Conformer selection can be done with lasers (double resonance schemes) or ion mobility spectrometry, whereas mass selection is enabled by mass spectrometry.

Concepts Specific to Nucleic Acids
- Excited state fast relaxation dynamics prevents some conformers or tautomers from being observed by UV double resonance spectroscopy, because they undergo internal conversion faster than the timescale of laser excitation with nanosecond pulses.
- The most stable tautomers in the gas phase are not necessarily the same as those of the natural bases in solution.
- The biologically relevant charge states are usually neutral isolated bases, deprotonated mononucleotides, and multiply charged anions for oligonucleotides. However, ionization energies, and therefore the type of action spectroscopy method to be selected, depend on charge.
- Gas-phase IRMPD spectroscopy provides key information on protonation or metallation sites, and on hydrogen bonding patterns, for small ions. Theory is now able to correctly predict the most stable gas-phase structures, for up to mononucleotides.
- Gas-phase UV spectroscopy is promising for probing the effect of base stacking on excited states nature and excited states dynamics in nucleic acid higher-order structures, but this field is very much in its infancy.

References

1. Taillandier E, Liquier J (1992) Infrared spectroscopy of DNA. Methods Enzymol 211:307–335
2. Tataurov AV, You Y, Owczarzy R (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. Biophys Chem 133:66–70

3. Clark LB, Peschel GG, Tinoco I Jr (1965) Vapor spectra and heats of vaporization of some purine and pyrimidine bases. J Phys Chem 69:3615–3618
4. Kleinermanns K, Nachtigallová D, de Vries MS (2013) Excited state dynamics of DNA bases. Int Rev Phys Chem 32:308–342
5. Vaya I, Gustavsson T, Miannay FA, Douki T, Markovitsi D (2010) Fluorescence of natural DNA: from the femtosecond to the nanosecond time scales. J Am Chem Soc 132:11834–11835
6. de Vries MS, Hobza P (2007) Gas-phase spectroscopy of biomolecular building blocks. Annu Rev Phys Chem 58:585–612
7. Belau L, Wilson KR, Leone SR, Ahmed M (2007) Vacuum-ultraviolet photoionization studies of the microhydration of DNA bases (guanine, cytosine, adenine, and thymine). J Phys Chem A 111:7562–7568
8. Yang X, Wang XB, Vorpagel ER, Wang LS (2004) Direct experimental observation of the low ionization potentials of guanine in free oligonucleotides by using photoelectron spectroscopy. Proc Natl Acad Sci USA 101:17588–17592
9. Danell AS, Parks JH (2003) Fraying and electron autodetachment dynamics of trapped gas phase oligonucleotides. J Am Soc Mass Spectrom 14:1330–1339
10. Weber JM, Ioffe IN, Berndt KM, Loffler D, Friedrich J, Ehrler OT, Danell AS, Parks JH, Kappes MM (2004) Photoelectron spectroscopy of isolated multiply negatively charged oligonucleotides. J Am Chem Soc 126:8585–8589
11. Nagornova NS, Guglielmi M, Doemer M, Tavernelli I, Rothlisberger U, Rizzo TR, Boyarkin OV (2011) Cold-ion spectroscopy reveals the intrinsic structure of a decapeptide. Angew Chem Int Ed Engl 50:5383–5386
12. Nagornova NS, Rizzo TR, Boyarkin OV (2010) Highly resolved spectra of gas-phase Gramicidin S: a benchmark for peptide structure calculations. J Am Chem Soc 132:4040
13. Stearns JA, Seaiby C, Boyarkin OV, Rizzo TR (2009) Spectroscopy and conformational preferences of gas-phase helices. Phys Chem Chem Phys 11:125–132
14. Rodriguez JD, Lisy JM (2009) Infrared spectroscopy of multiply charged metal ions: methanol-solvated divalent manganese 18-crown-6 ether systems. J Phys Chem A 113:6462–6467
15. Garand E, Kamrath MZ, Jordan PA, Wolk AB, Leavitt CM, McCoy AB, Miller SJ, Johnson MA (2012) Determination of noncovalent docking by infrared spectroscopy of cold gas-phase complexes. Science 335:694–698
16. Bierau F, Kupser P, Meijer G, von Helden G (2010) Catching proteins in liquid helium droplets. Phys Rev Lett 105:133402
17. Choi MY, Miller RE (2006) Four tautomers of isolated guanine from infrared laser spectroscopy in helium nanodroplets. J Am Chem Soc 128:7320–7328
18. Shukla MK, Leszczynski J (2013) Tautomerism in nucleic acid bases and base pairs: a brief overview. Comput Mol Sci 3:637–649
19. Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS (2001) Guanine tautomerism revealed by UV–UV and IR–UV hole burning spectroscopy. J Chem Phys 115:4604
20. Mons M, Piuzzi F, Dimicoli I, Gorb L, Leszczynski J (2006) Near-UV resonant two-photon ionization spectroscopy of gas phase guanine: evidence for the observation of three rare tautomers. J Phys Chem A 110:10921–10924
21. Abo-riziq A, Crews BO, Compagnon I, Oomens J, Meijer G, Von Helden G, Kabelac M, Hobza P, de Vries MS (2007) The mid-IR spectra of 9-ethyl guanine, guanosine, and 2-deoxyguanosine. J Phys Chem A 111:7529–7536
22. Crespo-Hernandez CE, Cohen B, Hare PM, Kohler B (2004) Ultrafast excited-state dynamics in nucleic acids. Chem Rev 104:1977–2019
23. Middleton CT, de La Harpe K, Su C, Law YK, Crespo-Hernandez CE, Kohler B (2009) DNA excited-state dynamics: from single bases to the double helix. Annu Rev Phys Chem 60: 217–239
24. Alonso JL, Pena I, Lopez JC, Vaquero V (2009) Rotational spectral signatures of four tautomers of guanine. Angew Chem Int Ed Engl 48:6141–6143

25. Alonso JL, Vaquero V, Pena I, Lopez JC, Mata S, Caminati W (2013) All five forms of cytosine revealed in the gas phase. Angew Chem Int Ed Engl 52:2331–2334
26. Lopez JC, Pena MI, Sanz ME, Alonso JL (2007) Probing thymine with laser ablation molecular beam Fourier transform microwave spectroscopy. J Chem Phys 126:191103
27. Vaquero V, Sanz ME, Lopez JC, Alonso JL (2007) The structure of uracil: a laser ablation rotational study. J Phys Chem A 111:3443–3445
28. Nir E, Plützer C, Kleinermanns K, de Vries M (2002) Properties of isolated DNA bases, base pairs and nucleosides examined by laser spectroscopy. Eur Phys J D 20:317–329
29. Jurecka P, Sponer J, Cerny J, Hobza P (2006) Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. Phys Chem Chem Phys 8:1985–1993
30. Kabelac M, Hobza P (2007) Hydration and stability of nucleic acid bases and base pairs. Phys Chem Chem Phys 9:903–917
31. Kabelac M, Plutzer C, Kleinermanns K, Hobza P (2004) Isomer selective IR experiments and correlated ab initio quantum chemical calculations support planar H-bonded structure of the 7-methyl adenine?adenine and stacked structure of the 9-methyl adenine?adenine base pairs. Phys Chem Chem Phys 6:2781
32. Nir E, Kleinermanns K, de Vries MS (2000) Pairing of isolated nucleic-acid bases in the absence of the DNA backbone. Nature 408:949–951
33. Sobolewski AL, Domcke W, Hattig C (2005) Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: the role of electron-driven proton-transfer processes. Proc Natl Acad Sci USA 102:17903–17906
34. Plutzer C, Hunig I, Kleinermanns K, Nir E, de Vries MS (2003) Pairing of isolated nucleobases: double resonance laser spectroscopy of adenine-thymine. Chemphyschem 4:838–842
35. Salpin JY, Guillaumont S, Tortajada J, MacAleese L, Lemaire J, Maitre P (2007) Infrared spectra of protonated uracil, thymine and cytosine. Chemphyschem 8:2235–2244
36. Bakker JM, Sinha RK, Besson T, Brugnara M, Tosi P, Salpin JY, Maitre P (2008) Tautomerism of uracil probed via infrared spectroscopy of singly hydrated protonated uracil. J Phys Chem A 112(48):12393–12400
37. Bakker JM, Salpin J-Y, Maître P (2009) Tautomerism of cytosine probed by gas phase IR spectroscopy. Int J Mass Spectrom 283:214–221
38. Yang B, Wu RR, Berden G, Oomens J, Rodgers MT (2013) Infrared multiple photon dissociation action spectroscopy of proton-bound dimers of cytosine and modified cytosines: effects of modifications on gas-phase conformations. J Phys Chem B 117(46):14191–201
39. Nei YW, Hallowita N, Steill JD, Oomens J, Rodgers MT (2013) Infrared multiple photon dissociation action spectroscopy of deprotonated DNA mononucleotides: gas-phase conformations and energetics. J Phys Chem A 117:1319–1335
40. Nei YW, Crampton KT, Berden G, Oomens J, Rodgers MT (2013) Infrared multiple photon dissociation action spectroscopy of deprotonated RNA mononucleotides: gas-phase conformations and energetics. J Phys Chem A 117:10634–10649
41. Rosu F, Gabelica V, Joly L, Gregoire G, De Pauw E (2010) Zwitterionic i-motif structures are preserved in DNA negatively charged ions produced by electrospray mass spectrometry. Phys Chem Chem Phys 12:13448–13454
42. Gabelica V, Rosu F, De Pauw E, Lemaire J, Gillet JC, Poully JC, Lecomte F, Gregoire G, Schermann JP, Desfrancois C (2008) Infrared signature of DNA G-quadruplexes in the gas phase. J Am Chem Soc 130:1810–1811
43. Marian C, Nolting D, Weinkauf R (2005) The electronic spectrum of protonated adenine: theory and experiment. Phys Chem Chem Phys 7:3306–3316
44. Pedersen SO, Stochkel K, Byskov CS, Baggesen LM, Nielsen SB (2013) Gas-phase spectroscopy of protonated adenine, adenosine 5'-monophosphate and monohydrated ions. Phys Chem Chem Phys 15:19748–19752

45. Cheong NR, Nam SH, Park HS, Ryu S, Song JK, Park SM, Perot M, Lucas B, Barat M, Fayeton JA et al (2011) Photofragmentation in selected tautomers of protonated adenine. Phys Chem Chem Phys 13:291–295
46. Marcum JC, Halevi A, Weber JM (2009) Photodamage to isolated mononucleotides– photodissociation spectra and fragment channels. Phys Chem Chem Phys 11:1740–1751
47. Marcum JC, Kaufman SH, Weber JM (2011) UV-photodissociation of non-cyclic and cyclic mononucleotides. Int J Mass Spectrom 303:129–136
48. Nielsen LM, Pedersen SO, Kirketerp MB, Nielsen SB (2012) Absorption by DNA single strands of adenine isolated in vacuo: the role of multiple chromophores. J Chem Phys 136: 064302
49. Gabelica V, Rosu F, De Pauw E, Antoine R, Tabarin T, Broyer M, Dugourd P (2007) Electron photodetachment dissociation of DNA anions with covalently or noncovalently bound chromophores. J Am Soc Mass Spectrom 18:1990–2000
50. Gabelica V, Tabarin T, Antoine R, Rosu F, Compagnon I, Broyer M, De Pauw E, Dugourd P (2006) Electron photodetachment dissociation of DNA polyanions in a quadrupole ion trap mass spectrometer. Anal Chem 78:6564–6572
51. Rosu F, Gabelica V, De Pauw E, Antoine R, Broyer M, Dugourd P (2012) UV spectroscopy of DNA duplex and quadruplex structures in the gas phase. J Phys Chem A 116:5383–5391
52. Dao NT, Haselsberger R, Michel-Beyerle ME, Phan AT (2013) Excimer formation by stacking g-quadruplex blocks. Chemphyschem 14:2667–2671
53. Chingin K, Chen H, Gamez G, Zenobi R (2009) Exploring fluorescence and fragmentation of ions produced by electrospray ionization in ultrahigh vacuum. J Am Soc Mass Spectrom 20: 1731–1738
54. Bian QZ, Forbes MW, Talbot FO, Jockusch RA (2010) Gas-phase fluorescence excitation and emission spectroscopy of mass-selected trapped molecular ions. Phys Chem Chem Phys 12: 2590–2598
55. Talbot FO, Rullo A, Yao H, Jockusch RA (2010) Fluorescence resonance energy transfer in gaseous, mass-selected polyproline peptides. J Am Chem Soc 132:16156–16164
56. Danell AS, Parks JH (2003) FRET measurements of trapped oligonucleotide duplexes. Int J Mass Spectrom 229:35–45

Fragmentation Reactions of Nucleic Acid Ions in the Gas Phase

6

Yang Gao and Scott A. McLuckey

Abstract

This chapter summarizes literature describing the gas-phase fragmentation of nucleic acid ions under a variety of reaction conditions. Specifically, the phenomenology of gas-phase dissociation of nucleic acid ions is determined by ion type, charge state, the energy deposition method, and the fragmentation reaction timescale. Various proposed mechanisms are summarized. The chapter is organized by dissociation method. For the most extensively studied collision-induced dissociation (CID), the literature is subcategorized by analyte and ion type. In many cases, no single fragmentation mechanism can account for all the reported products. This suggests that multiple dissociation mechanisms can contribute, depending on ion type, ion charge state, and reaction conditions.

Abbreviations

BIRD	Blackbody infrared radiative dissociation
CID	Collision-induced dissociation
DNA	Deoxyribonucleotide

Y. Gao • S.A. McLuckey (✉)

Department of Chemistry, Purdue University, West Lafayette, IN 47907-2084, USA

e-mail: mcluckey@purdue.edu

DR	Double resonance
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EPD	Electron photodetachment dissociation
ESI	Electrospray ionization
ETD	Electron transfer dissociation
ETcaD	Electron transfer collision-activated dissociation
ET-IRMPD	Electron transfer-infrared multiphoton dissociation
ET-UVPD	Electron transfer-ultraviolet photodissociation
FAB	Fast atom bombardment
FTICR	Fourier transform ion cyclotron resonance
IP	Ionization potential
IR	Infrared
IRMPD	Infrared multiphoton dissociation
LNA	Locked nucleic acid
MALDI	Matrix-assisted laser desorption/ionization
MBO	Mixed-backbone oligonucleotide
MP	Methylphosphonate
MS	Mass spectrometer/mass spectrometry
NETD	Negative electron transfer dissociation
niECD	Negative ion electron capture dissociation
NS	Nozzle-skimmer
PA	Proton affinity
PD	Plasma desorption
PS	Phosphorothioate
PSD	Post-source decay
Q-TOF	Quadrupole/time-of-flight tandem mass spectrometer
rf	Radiofrequency
RNA	Ribonucleotide
RNAi	RNA interference
SID	Surface-induced dissociation
SNP	Single nucleotide polymorphisms
SORI	Sustained off-resonance irradiation
TOF	Time-of-flight
TOF/TOF	Tandem time-of-flight
UV	Ultraviolet
UVPD	Ultraviolet photodissociation

6.1 Introduction

Mass spectrometry (MS) has been used for decades in structure elucidation and characterization via the gas-phase fragmentation of ions derived from analytes of interest. The introduction of "soft" ionization methods, such as electrospray

ionization (ESI) [1] and matrix-assisted laser desorption/ionization (MALDI) [2], allows the formation of intact pseudo-molecular ions ($[M + nH]^{n+}$ or $[M - nH]^{n-}$) as surrogates for the biomolecules. For elucidation of the primary structure of large biopolymers, fragments generated from backbone dissociation have been used to determine the sequence or branching of residues in an oligomer.

Fragmentation behavior of biomolecules such as peptides [3–5], proteins [6–8], oligosaccharides [9, 10], and nucleic acids [11–14] has been extensively studied via tandem MS. Generally, a neutral molecule of interest is introduced into the mass spectrometer in the form of an ion. The ions of interest are then subjected to a tandem MS process that involves at least two stages of mass analysis. A typical tandem MS method is composed of the following steps: (1) mass selection of the ion of interest (parent ion), (2) energy deposition (by collision, photon, electron, etc.) into the isolated parent ion to induce cleavages of the covalent bonds, and (3) mass spectrometric analysis of the product ions. Gas-phase fragmentation of nucleic acids can be generally used in the areas of sequencing, identification of single nucleotide polymorphisms (SNP), identification and location of modified bases and sugar moieties, etc. The dominant dissociation pathways for a particular ion are generally a function of the ion type (e.g., positive ion, negative ion, radical species, etc.), the number of charges, the molecular structure, the energy deposition method, and the fragmentation reaction timescale. Nucleic acids have both acidic (phosphate backbone) and basic (nucleobase) functionalities, and can therefore be protonated (metalated) or deprotonated to form the pseudo-molecular ions. With the evolution of tandem mass spectrometers, biopolymer ions have been subjected to an increasingly wide range of reaction conditions, including a variety of energy deposition methods, reaction time frames, and observation time windows ranging over several orders of magnitude.

In this chapter, we mainly focus on the gas-phase dissociation reactions of nucleic acid ions observed for different ion types generated via ESI or MALDI under a variety of reaction conditions. Generally, these ionization methods are capable of forming either closed shell positively or negatively charged ions with single or, particularly in the case of ESI, multiple charge states. In addition to the variations in ion types and reaction conditions, nucleic acid ions of various sizes and compositions have been investigated. Due to the high dimensionality of variables examined in nucleic acid ion decompositions, a diverse set of observations have been reported in the literature. Hereby, we attempt to give a summary of observations made up to 2013, showing the decomposition phenomena and conditions as well as the proposed mechanisms. In most cases, the way that ions are formed does not affect the dissociation mechanism; thus, this review is organized on the basis of different dissociation methods, analyte identities, and ion types. This chapter begins with a brief description of the nomenclature most frequently encountered in the literature. Then, the discussion is broken down along the lines of the major processes used to probe the ions, the analyte type (e.g., DNA, RNA, modified oligomers, and ion type). Emphasis is placed on current thinking regarding dissociation mechanisms as these ultimately underlie the structural

information that is revealed for a given oligomer type, ion type, subjected to a given structural probe.

6.2   Nomenclature of Oligonucleotide Fragment Ions

In the early 1980s, Grotjahn and coworkers [15–18] referred to the fragment ions generated by dissociation of the phosphodiester bond as "sequence ions." Two major fragment ion series were observed for the backbone cleavages of nucleic acid ions upon FAB-MS. The fragment ions containing the 3′-terminus or 5′-terminus of the precursor ion were designated as 5′-sequence ions or 3′-sequence ions, respectively. The detailed fragment notation scheme is shown in Fig. 6.1a. Viari et al. [19] suggested an alternative nomenclature in the year 1987, which was later extended by Nordhoff et al. [20] in 1995 (see Fig. 6.1b). McLuckey et al. [21] in 1992 proposed a nomenclature scheme that is analogous to the one widely used for peptides and proteins. This nomenclature defines all possible cleavages from the phosphodiester backbone and has been widely adopted by the community. The detailed nomenclature scheme is shown in Fig. 6.1c. The four possible cleavages along the phosphodiester linkage are denoted by the lower case letters a, b, c, and d for fragments that contain the 5′-terminus of the precursor and w, x, y, and z for fragments containing the 3′-terminus of the precursor. The numerical subscripts indicate the number of ribose units from the corresponding termini. Upper case letter B represents a base with the subscript indicating its position in the sequence starting from the 5′-terminus. The identity of the lost base is indicated in parentheses along with the B_n denotation, e.g., $B_2(A)$. The use of $B_n(Base)$, however, has largely been replaced by the use of the base symbol with a subscript indicating location of the nucleotide from which it was lost, e.g., A_2 instead of $B_2(A)$.

6.3   Fragmentation Methods

A variety of methods have been applied to dissociate nucleic acid ions for the purpose of generating sequence information and for localizing modifications. Some are based on depositing energy into the ion via collisions with a gas or surface (ion/neutral interaction) or by some form of photodissociation (ion/photon). Some methods involve either adding or removing an electron to or from an ion derived from ESI to generate a radical ion. In these cases, fragmentation may occur directly as a result of the electronic transition or a subsequent activation step may be used to fragment the radical ion. The energies and timescales associated with these methods can vary widely. It is beyond the scope of this chapter to describe these methods in detail. The reader is directed to a recent overview of dissociation methods in tandem mass spectrometry [22]. Table 6.1 summarizes many of the methods used for fragmenting nucleic acid ions along with the associated timescales and instruments often used to implement them. The remainder of the
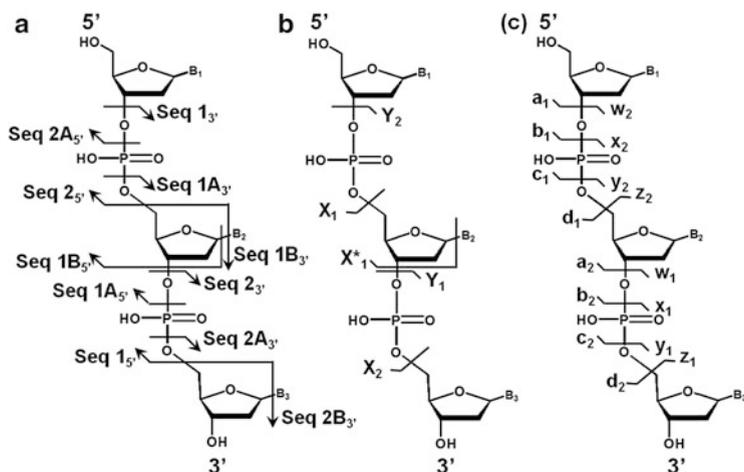
Fig. 6.1 Nucleic acid fragment ion nomenclature proposed (**a**) by Cerny et al. [17], (**b**) by Nordhoff et al. [20], and (**c**) by McLuckey et al. [21]

discussion regarding nucleic acid fragmentation is organized according to ion/neutral, ion/photon, and ion/electron(ion) interactions.

6.3.1   Ion/Neutral Interactions: CID of Nucleic Acids

CID is the most widely used dissociation method in tandem MS [23]. CID can be effected over a wide range of conditions with collision energies ranging from a few electronvolts to thousands of electronvolts, timescales ranging from microseconds to seconds, and pressures ranging from sufficiently low that only single collisions are likely to high enough that hundreds of collisions can occur. The reader is referred a recent review of CID of peptide and proteins for an in-depth discussion of CID applied to large polyatomic ions [24]. For this discussion, it is useful to point out the distinction between "beam-type CID," which applies to transmission instruments (e.g., triple quadrupole instruments), and "ion trap" or "sustained off-resonance irradiation (SORI)" CID, which are implemented in quadrupole ion trap and ion cyclotron resonance instruments, respectively. The latter CID approaches CID employ much longer activation times and sample processes that occur over longer timescales than in beam-type CID. As a result, CID in the ion trapping instruments tends to more strongly favor low-energy processes leading to less sequential fragmentation than does beam-type CID.

6.3.1.1 CID of DNA Anions
Ion Trap CID Phenomenology
The early efforts to elucidate the gas-phase dissociation of nucleic acid ions were focused on deoxyribonucleotides (DNAs) [11, 12, 25, 26]. Various collisional

Table 6.1 Summary of methods used to dissociate nucleic acid ions

Interaction	Experimental timescale	Main result	Common instruments
Ion/neutral			
Beam-type CID	<100 μs	Fragmentation	QQQ, Q-TOF, TOF/TOF, et al.
Ion trap CID	10–100s ms	Fragmentation	3D, linear ion traps
SORI/CID	10–100s ms	Fragmentation	FTICR
Nozzle-skimmer CID	<100 μs	Fragmentation	ESI interface
SID (ion/surface)	ns	Fragmentation	Specialized QQQ, Q-TOF
Ion/photon			
BIRD	100 ms-s	Fragmentation	FTICR
IRMPD	10–100s ms	Fragmentation	FTICR, ion traps
UVPD	ns	Fragmentation	Traps, FTICR, TOF/TOF, et al.
EPD	ns	Electron detachment from anion	Traps, FTICR, TOF/TOF, et al.
Ion/electron			
ECD	10–100s ms	Electron capture by cation/ fragmentation	FTICR
niECD	10–100s ms	Electron capture by anion/ fragmentation	FTICR
EDD	10–100s ms	Electron detachment from anion/ fragmentation	FTICR
Ion/ion			
ETD	10–100s ms	Electron transfer to cation/ fragmentation	Ion traps
NETD	10–100s ms	Electron transfer from anion/ fragmentation	Ion traps

See abbreviations for definitions of the acronyms

activation techniques have been used to investigate the fragmentation pathways of oligonucleotide anions and to explore the potential of MS^n (where $n \geq 2$) for nucleic acid sequencing. The earliest studies of the fragmentation pathways of oligonucleotide anions as a result of ion trap CID were reported by McLuckey and coworkers [21, 27–32]. Beam-type CID of multiply charged oligonucleotide ions was also studied by several groups [33–37]. NS fragmentation was also investigated [38–40]. Several reports are selected to demonstrate the gas-phase fragmentation behavior of the DNA anions under various experimental conditions.

For elucidation of the fragmentation mechanism, CID patterns of mononucleotides, dinucleotides, and trinucleotides have been extensively studied by ESI-MS/MS. The mononucleotides subjected to CID are singly deprotonated, so there is no influence from internal Coulombic repulsion on the unimolecular dissociation process. This provides benchmark information for interpreting the

CID patterns of larger multiply charged nucleic acid ions. Rodgers et al. [41] (SORI/CID) and Habibi-Goudarzi et al. (ion trap CID) [30] systematically examined the low-energy CID behavior of the singly deprotonated ions of the four 2′-deoxyribodinucleoside 3′-monophosphates and their corresponding 5′-monophosphates. They found that CID of the eight 2′-deoxyribonucleoside monophosphates shows loss of the neutral base as the dominant fragmentation channel. Evidence showed that the loss of the nucleobase is greatly facilitated by the presence of 3′-phosphate as opposed to 5′-phosphate. For the 2′-deoxyribonucleoside 3′-monophosphates, formation of PO_3^- is a major channel, whereas the 3′-monophosphate counterpart tends not to fragment through this pathway. Water loss following loss of the neutral base has been observed for all of the 5′-monophosphate species and is far more prominent than their 3′-monophosphate counterparts. The water loss presumably results from elimination of the 3′-hydroxyl group with a 4′-hydrogen, analogous to the (a-B)/w-ion formation for the multiply charged nucleic acid ions [21, 28, 29]. For the two sets of 2′-deoxyribonucleoside monophosphates, the tendency for neutral base loss follows the order of dAp, dTp > dCp > dGp, and pdA, pdT > pdC pdG. According to Rodgers et al. [41], the significant low reactivity of pdG is postulated to result from the strong hydrogen bonding between the 2-NH_2 of guanine and the 5′-phosphate group.

For singly charged dinucleotides, Habibi-Goudarzi and McLuckey [30] have noted that the loss of the 5′-base is significantly preferred over the 3′-base [30, 36]. This phenomenon agrees with the previous postulation that 3′-phosphate facilitates the base loss, given that 3′-base does not have a 3′-phosphate. The 5′-base is lost either as a neutral or as an anion, whereas the 3′-base is preferentially lost as a neutral [30]. Loss of the deprotonated 5′-base follows the trend $A^- > G^- > T^- > C^-$ for a given 3′-base. The identity of the 3′-base also determines the likelihood of observing a 5′-base loss. The tendency for observing the 5′-base anion follows the order of the 3′-base as A > C > T > G [30]. The inhibition to the 5′-base loss with G as the 3′-base is possibly due to the strong interaction between the guanine and the phosphodiester linkage (3′-phosphate for the 5′-base).

Virkic et al. [42] investigated the gas-phase fragmentation pattern of the [M–H]$^-$ ions derived from all 64 trinucleotides and 16 mixed-base tetranucleotides using CID-MS^n in a quadrupole ion trap. The initial fragmentation channel of the singly charged trinucleotides is neutral base loss from the 5′-terminus, along with minor 3′ and internal base loss. Other sequence ions are also present with low or moderate abundances, including c_2^-, x_2^-, y_2^-, and z_2^-. CID of the [M–H–AH]$^-$ ion yields w_2^- as the most abundant fragment peak, which results from the subsequent 3′ C–O bond cleavage of the ribose. As a general rule, loss of a charged base, as opposed to loss of neutral base, increases with the ratio of the charge to the number of phosphate groups in the parent ion. It has also been found that the loss of a charged cytidine is not commonly observed, unless the molecule is highly charged. However, the loss of a neutral cytidine is a facile reaction if Coulombic repulsion does not promote the loss of some other base as a charged species.

McLuckey and Vaidyanathan [32] systematically investigated the effect of precursor ion charge state, oligonucleotide length, and nucleobase composition on the fragmentation behaviors of a series of single-base DNA anions (3- to 8-mers) by ion trap tandem mass spectrometry. Precursor ion charge state has tremendous influences on the major dissociation pathways. Precursor ion charge state has been postulated to determine the threshold energies required for various competitive decomposition pathways. Therefore, changes in precursor ion charge state can lead to dramatic changes in the relative contributions of different fragmentation channels. For a given number of residues, nucleobase composition is an important factor. Poly-T anions show a strong tendency to lose the 5′-thymine, whereas loss of the 5′-guanine is inhibited for poly-G anions. Poly-A and poly-C do not follow any pattern in this regard. Poly-A and poly-C anions show strong evidence that charge location is important for directing fragmentation into particular sequence channels when a single combination of charge site locations is preferred. In contrast to other single-base oligomers, poly-A anions tend to dissociate by consecutive losses of the base, as either an anion or a neutral, competing with the reactions that give rise to $[a-B]^-$- and w-type sequence ions. This tendency is more obvious with increasing size and charge of poly-A oligomers. In the case of oligomers of the same residue number and charge, the ease of fragmentation of single-base DNA follows the order of poly-AT $\approx$ C > G. The ease of adenine loss is possibly relative to the weak N-glycosidic bond strength associated with adenine, whereas enhanced stability of guanine is presumably due to its hydrogen bonding interaction with the 5′-phosphate group.

From the observations made on ion trap CID of a larger set of single-stranded DNA anions under quadrupole ion trap conditions [27–29, 32, 41, 42], several tendencies for gas-phase decomposition of the multiply deprotonated species emerge:

1. The loss of a nucleobase is the initiate step of the fragmentation. Depending on the precursor ion charge state, the base can be lost as an anion or neutral. For precursor ions with low charge states, neutral base loss is predominant. As the charge on the precursor ion increases, loss of charged bases becomes favored and coexists with the neutral base losses. The order of preference for charge base loss is $A^- > T^- > G^-$, C^- [28] or $A^- > G^- > T^- > C^-$ [29]. Neither of these orders is consistent with the order for base loss from the 2′-deoxyribonucleoside 3′-monophosphates (dAp, dTp > dCp > dGp) or from the corresponding 5′-monophosphates (pdA, pdT > pdCpdG) [30]. The different orders of stability reported can be attributed to higher structural complexity of oligonucleotides and the Coulombic repulsion caused by multiple charges.

2. Loss of the 3′-terminal base tends to be disfavored. The propensity for base loss has also been reported to be higher for 3′-phosphorylated mononucleotides than for the corresponding 5′-phosphorylated species [18, 30, 43]. This phenomenon has been attributed to that the base loss is facilitated by the 3′-phosphate. Neither 3′-terminal base nor 5′-phosphorylated mononucleotides have the facilitated 3′-phosphate, and thus such base loss is not favored.

3. The subsequent cleavage of the 3′ C–O bond of the ribose from which the base was lost is a dominant pathway. The dissociation of the 3′ C–O bond generally forms complementary (a-B)- and w-type fragments regardless of the base or its charge. This fragmentation channel outcompetes all the other possible cleavages (b/x-, c/y-, or d/z-type cleavage) along the phosphodiester linkage and therefore greatly simplifies spectral interpretation for DNA-type oligonucleotides.

4. Base losses from first- and second-generation fragments facilitate further backbone cleavages and produce internal fragments. Multiple base losses increase with increasing activation energy and increasing precursor ion charge state. However, loss of a PO_3^- group from 5′-phosphorylated fragment ions, such as w-type products ions, can compete effectively with this base loss channel.

Beam-Type CID Phenomenology

In addition to the observations made by ion quadrupole ion traps, CID behaviors of DNAs have been investigated with beam-type instruments, such as triple quadrupoles. The results differ somewhat from trends observed in quadrupole ion traps. One of the most obvious differences is that no obvious preference for the loss of a specific nucleobase is observed in beam-type CID of DNA anions [33, 36, 44]. Moreover, it has been reported that base loss is not a prerequisite for backbone cleavage [45]. The prominent product ions from beam-type CID are not limited to (a-B)- and w-ions [36, 44]. The differences in dissociation pattern of DNA anions in ion trap versus beam-type instruments likely arise from the differences in activation time frames and precursor ion internal energies. Ion trap collisional activation generally favors slow low-energy processes due to its relatively long time frame, whereas beam-type collisional activation tends to access higher energies that make more dissociation channels competitive. Additionally, ion trap CID uses single-frequency excitation which only affects the precursor ions, whereas, in triple quadrupole instrument, product ions can undergo secondary collision with the bath gas and induce further generations of fragments. It has been reported that beam-type CID generated more extensive sequence information than ion trap CID [44].

Boschenok and Sheil [36] studied the fragmentation patterns of $[M–H]^-$ ions of a series of oligonucleotides on a quadrupole-hexapole-quadrupole mass spectrometer as well as on a magnetic sector-TOF-MS. They found that the dominant fragment in the CID mass spectra of singly deprotonated dinucleotide anions came from the 5′-base anion. For DNA hexamers, beam-type CID of the $[M–2H]^{2-}$ ions generates fragment ions corresponding to loss of bases followed by cleavage of the 3′ C–O bond in the ribose that has lost the base, i.e., (a-B)- and w-fragments. Distinct from ion trap CID results of multiply charged oligonucleotide anions, no strong preference for loss of any particular base is observed in the tandem mass spectra of the hexamers obtained either on a triple quadrupole ($E_{lab} = 40$ eV) or on a magnetic sector-TOF instrument ($E_{lab} = 400$ eV).

Barry et al. [33] studied a series of synthetic oligonucleotides with one modified nucleobase in the sequence of 5′-d(CACGXG)-3′, where X represents ethyldeoxyuridine, bromovinyldeoxyuridine, 5-iododeoxyuridine, or trifluorodeoxythymidine,

in a triple quadrupole mass spectrometer. Triply deprotonated precursor ions were subjected to beam-type CID and an apparent reduction in the total number of product ions in the CID spectra of the oligonucleotides with iodo- and trifluoro-modified bases was observed. However, the CID spectra of the ethyl- and bromovinyl-modified oligomers showed loss of all nonterminal bases with concomitant cleavage of the ribose 3′ C–O bonds. The authors suggested that the multiple collisions induced in a triple quadrupole experiment could allow for multiple competitive pathways to be observed instead of the lowest energy processes as in a quadrupole ion trap. The initial base loss tends to be influenced by the electron affinity of substituents on the nucleobase. Loss of a nucleobase anion becomes more facile when the modification on the base becomes more electron withdrawing. Moreover, the X^- loss in conjunction of the subsequent 3′ C–O cleavage, generating $(a_5\text{-}B_5)$- and w_1-ions, formed the most abundant product ions in the case of the trifluoro- and iodo-modified oligomers such that other backbone cleavage channels were diminished. Another structurally useful fragmentation pathway was also observed by Barry et al. [33]. This fragmentation pathway yields complementary c- and y-ions. It was noted that complete y_n^- series was observed in the CID spectra of the ethyl- and bromovinyl-modified oligomers but disappeared completely in the case of trifluoro- and iodo-modified oligonucleotides. The rationale for this was that fragments containing the base X that is highly electron withdrawing are susceptible to further fragmentation and are therefore less likely to appear in the product spectrum. It is also possible that a y-ion can result from the loss of PO_3^- or HPO_3 from a w-ion. Then, the reduced formation of w-ions could directly impact the observation of y-ions.

CID Phenomenology of Double-Stranded DNA

Besides single-stranded oligodeoxyribonucleotides, DNA duplexes have also been investigated using ESI-MS/MS techniques [46–57]. Ding and Anderegg [55] demonstrated that the relative abundance of DNA duplex ions in ESI mass spectra is qualitatively correlated to the concentration of these duplexes in solution. Williams and coworkers [49] showed that DNA duplexes preserve the Watson–Crick base pairing in the gas phase by comparing the activation energies for dissociation of complementary and noncomplementary DNA duplexes using BIRD in an FTICR mass spectrometer. Gabelica and De Pauw [51] proved the existence of base stacking interactions and the conservation of double helix conformation in the gas phase by evaluating CID MS/MS of various 16-mer DNA duplexes with a hybrid Q-TOF instrument. They also observed that the uneven partition of the negative charges between the two single strands was related to the terminal bases exclusively and suggested that it is correlated with the gas-phase acidities of the [sugar-phosphate-sugar-base] species. Madsen and Brodbelt [57] also investigated the factors that affect the asymmetric charge partitioning observed upon dissociation of DNA duplexes in the gas phase. They showed that the degree of asymmetric charge partitioning increased significantly with duplex size, precursor ion charge state, G/C base composition, and internal energy deposition. They suggested that the conformations of the duplexes and their propensity for changes in

surface area before complete strand separation dictate the degree of asymmetric charge partitioning, in some ways similar to the dissociation mechanism of multimeric protein complexes.

Generally, DNA duplex dissociation follows two competing pathways: noncovalent dissociation of the complex into single strands and covalent fragmentation [48, 49, 52, 58]. Neutral base loss from the duplex competes with separation of the two strands, and both channels are dominant in the product mass spectra. For noncomplementary duplexes, it has been demonstrated that base loss initiated backbone fragmentation [49]. Gabelica and De Pauw [52] studied fragmentation patterns of DNA duplex using quadrupole ion trap and Q-TOF instrument. Their observation showed that neutral base loss from the duplex is favored by slow activation, whereas non-covalent dissociation into single strands is favored under fast activation conditions (Fig. 6.2). CID of the duplex also generated two types of sequence-informative fragments. Type I fragments are single-strand backbone fragments (noted 1, 2, 3,..., in Fig. 6.2), whereas Type II fragments are duplexes with backbone cleavage from one of the two strands (noted a, b, c,..., in Fig. 6.2). The abundance of Type II fragments increases under slow activation conditions. Type I fragments are, however, favored when fast activation methods are used. Moreover, MS [3] experiments suggest that base loss ions from the duplexes are the precursors of Type II fragments. Then, Type I fragments originate from either the complementary parts for Type II fragments or secondary fragmentation of single strands that contain excessive energy. It was noted that the CG pair components significantly alter the CID products of the duplexes. CID of the duplexes with low percentages of CG base pairs yields predominantly single strands and neutral base loss from the duplex, whereas CID of CG-rich duplexes produces significantly abundant backbone fragments. This difference implies that the separation of DNA duplex involves sequential unzipping, which is dependent on the GC pair composition of the duplex.

Fragmentation Mechanisms of DNA Anions

The observations of neutral or charged base loss and the subsequent formation of (a-B)- and w-type fragment ions have been proposed as arising from various mechanisms. The earliest mechanistic study for DNA fragmentation was performed on FAB-MS instrument [18]. Cerny et al. [18] observed that the high-energy CID of all 2′-deoxymononucleotides and 2′-deoxydinucleotides preferentially formed ions with a neutral bass loss from the parent, and this loss was always more facile with the presence of a 3′-phosphate. Moreover, 5′ charged base losses were more abundant than 3′ charged base losses for heterodinucleotides. The mechanism for loss of the charged base was proposed (Scheme 6.1) and was later supported by Phillips and McCloskey [59]. In this mechanism, the formation of a charged base is facilitated by backside nucleophilic attack (S_N2) by the negatively charged phosphate oxygen upon the 1′ carbon. This mechanism explains the favored base loss when a 3′-phosphate is present. However, it could not account for the favored subsequent formation of (a-B)- and w-ions.
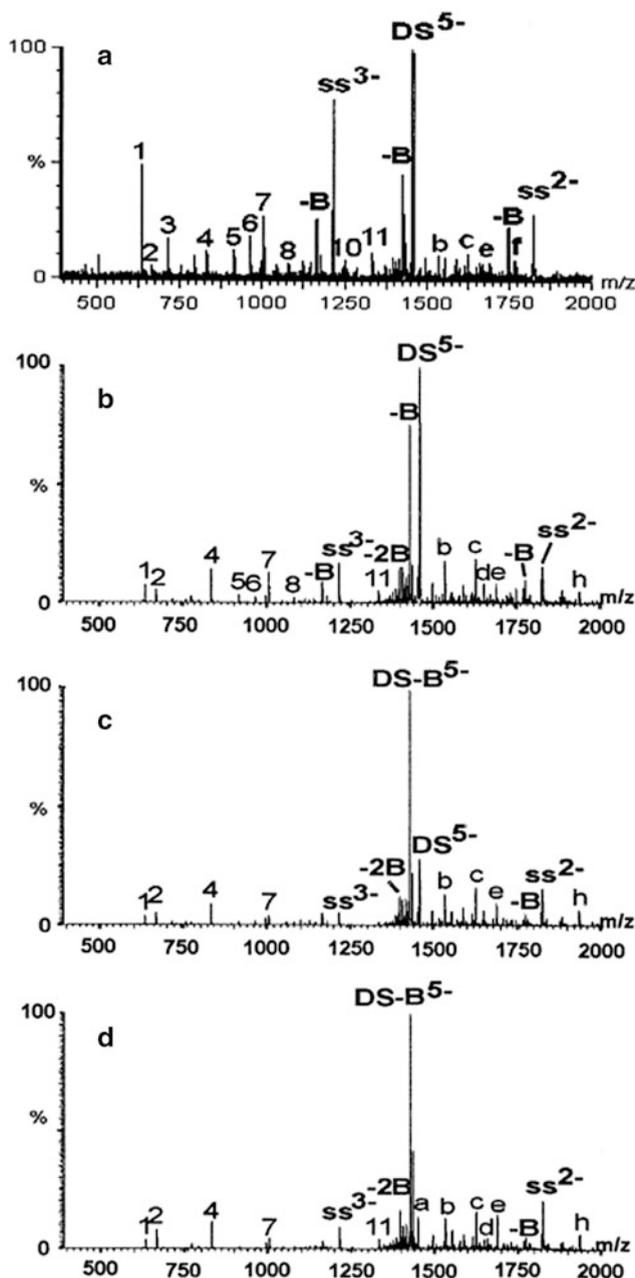
Fig. 6.2 MS/MS spectra of Duplex B [d(CGCGGGCCCGCG)$_2$$^{5-}$] recorded in different collision regimes, fastest to slowest activation from top to bottom. (**a**) Q-TOF, collision energy = 30 eV. (**b**) LCQ (ion trap), activation time = 3 ms, activation amplitude = 24 %. (**c**) LCQ, activation time = 30 ms, activation amplitude = 12 %. (**d**) LCQ, activation time = 300 ms, activation amplitude = 11 %. Reprinted from [52], Copyright 2002 Elsevier, Inc., with kind permission from Springer Science and Business Media
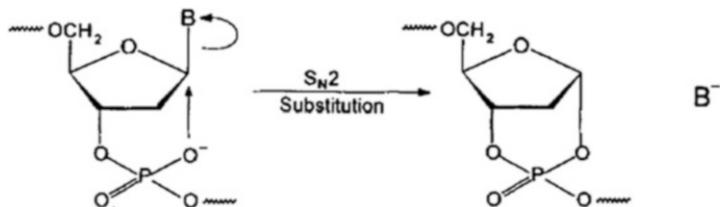
Scheme 6.1 Proposed mechanism whereby 8- is lost by backside nucleophilic attack on C(1′) by negatively charged phosphate. Reprinted with permission from [33]. Copyright 1995 John Wiley & Sons, Ltd

McLuckey and Habibi-Goudarzi [28] proposed a mechanism to account for the formation of the base loss as well as (a-B)/w-fragments (Scheme 6.2). Firstly, a nucleobase is removed by 1,2-elimination on the ribose. For highly charged ions, loss of a nucleobase anion is preferred to relieve electrostatic repulsion. The second step involves another "1,2-elimination reaction" (ribose ring numbering changed due to the abstraction of the nucleobase), which cleaves the 3′ C–O bond and forms a furan ring attached to the 5′-carbon. The second 1,2-elimination step is believed to be independent of the identity of the base initially removed and the precursor ion charge state. The formation of a stable aromatic system (furan) is the major driving force in the formation of the second-generation products.

Wang et al. [60] reported observation of highly abundant w-series ions and an (a_3-B_3)-ion for singly and doubly deprotonated T-rich DNA tetramers from both high-energy and low-energy CID. The authors suggested that the ions of $[M–H]^-$, a_3^-, and $[M–H–B]^-$ are possible precursors for the (a_3-B_3)-ions. However, a_3^- and $[M–H–B]^-$ are either absent or of low abundances in the CID spectra of the tetramers. This phenomenon implied that the (a_3-B_3)-ion is either generated via a short-lived intermediate or a concerted process. Additionally, the formation of the (a_3-B_3)-ion is exclusively preferred over the (a_2-B_2)-ion. The fragmentation of $[M–2H]^{2-}$ of the T-rich 5-mers, d(TGTTT), d(TTGTT), and d(TTTGT), also demonstrated that there is base dependence as well as position selectivity. Loss of neutral G from both a_3- and a_4-ions is favored over loss of T upon low-energy collisional activation. The (a_2-B_2)-ion is not observed unless when G is at the second position from the 5′-terminus. The previously proposed 1,2-elimination mechanisms could not explain the site-specific base loss or the absence of intermediates. A charge-remote mechanism involving transfer of proton from the adjoining phosphate for the formation of (a_n-B_n) was proposed for loss of neutral cytosine from singly charged precursor ions (Scheme 6.3). The first step involves a proton being transferred from the 5′ phosphodiester linkage to the nucleobase and subsequence N-glycosidic bond cleavage. Thus, the preference order of base loss can be explained by the relative gas-phase proton affinities of the nucleobases $(G > A \approx C > T)$. Moreover, the intramolecular proton transfer process requires the existence of a negative charge on the phosphate backbone, a positive charge on an adjacent nucleobase, and an excess negative charge to hold the net −1 charge.

Neutral Base Loss
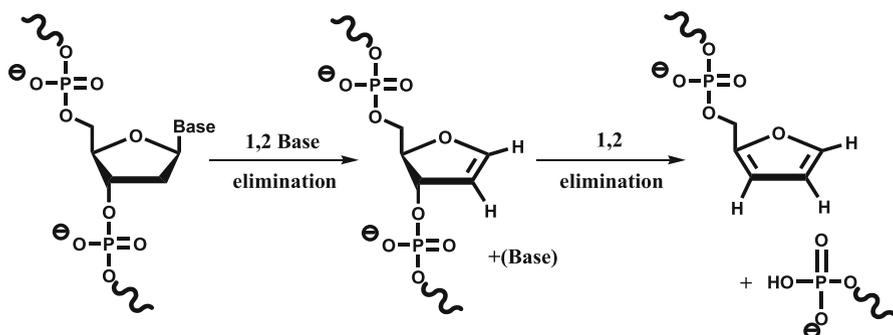
Charged Base Loss
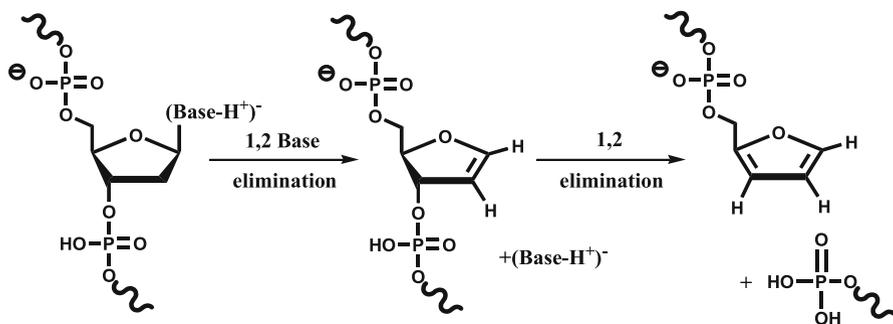
Scheme 6.2 Adapted with permission from [28]. Copyright 1993 American Chemical Society

The nearby acidic hydrogen on the 5′-phosphate is critical, without which the base cannot be protonated in the first step. For the formation of w-ions, the authors suggested that previous mechanisms (e.g., a 1,2-elimination) might still play a role. They found that the w-ions were usually more abundant than (a-B)-ions from doubly charged precursors. However, this mechanism does not explain why loss of the neutral 5′-base from the dinucleotide anion is favored over the loss of 3′-base, nor can it explain fragmentation of "fully" charged precursor ions, in which there is no free acidic 5′-phosphate hydrogen available.

Rodgers et al. [41] proposed a mechanism showing that base loss proceeds through a proton-bound dimer intermediate (Scheme 6.4). According to this mechanism, after the nucleobase is lost from 1,2-elimination, it forms a proton-bound dimer with its 3′-phosphate. McLuckey et al. [31] tested this mechanism by investigating MS/MS of a series of adenine containing DNAs with an ESI-quadrupole ion trap. The increasing degree of charged base loss with increasing parent ion charge was attributed to the energy surface involved in the dissociation of the proton-bound intermediate. Increasing charge states favors charged base
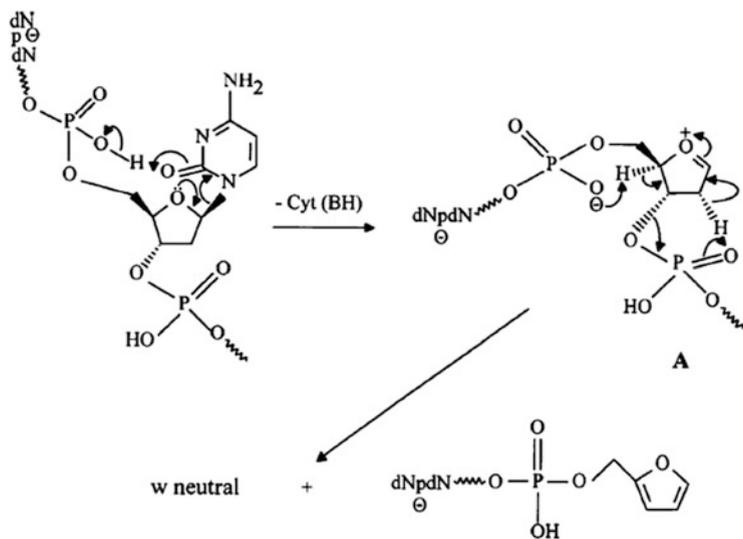
Scheme 6.3 Reprinted from [60]. Copyright 2002 Elsevier, Inc., with kind permission from Springer Science and Business Media

loss because the increasing Coulombic repulsion increases the likelihood of charge separation.

Barry et al. [33] postulated that the base loss reaction is base catalyzed by the phosphate group and undergoes an internal "bimolecular elimination" (E2 reaction). According to their mechanism, the first step involves the negatively charged 3′-phosphate oxygen abstracting the 2′-H via a six-member ring transition state and B⁻ is removed simultaneously. The relief in Coulombic strain in the oligonucleotide from the detachment of B⁻ favors this reaction. In the case when a 3′-phosphate is not present (3′-terminus), other functional groups may act as the base catalyst, i.e., other phosphate groups or a 3′-OH group. However, none of these groups can offer the same low-energy pathway as that provided by the 3′-phosphate. In the second step, the 3′ C–O bond is cleaved to yield a stable furan product, which may be base catalyzed by the 5′-phosphate oxygen. The reduction in Coulombic repulsion by the loss of the w-fragment ion and the formation of a stable substituted furan ring provide the driving force for the reaction.

Most previously proposed mechanisms involve the deprotonated phosphodiester linkage at the cleavage site. In order to investigate the necessity of an adjacent charge in the formation of (a-B)- and w-type fragments, Bartlett et al. [45] investigated the CID of multiply charged polyadenosine oligonucleotides by partially replacing the normal phosphodiester linkages with methylphosphonate (MP) linkages. In all MP-containing oligonucleotides, over 90 % of phosphate groups are observed to be charged, whereas only 60 % are charged in normal oligonucleotides. They postulated that this unexpected effect came from charge

Scheme 6.4 Reprinted from [41], Copyright 1994, with permission from Elsevier
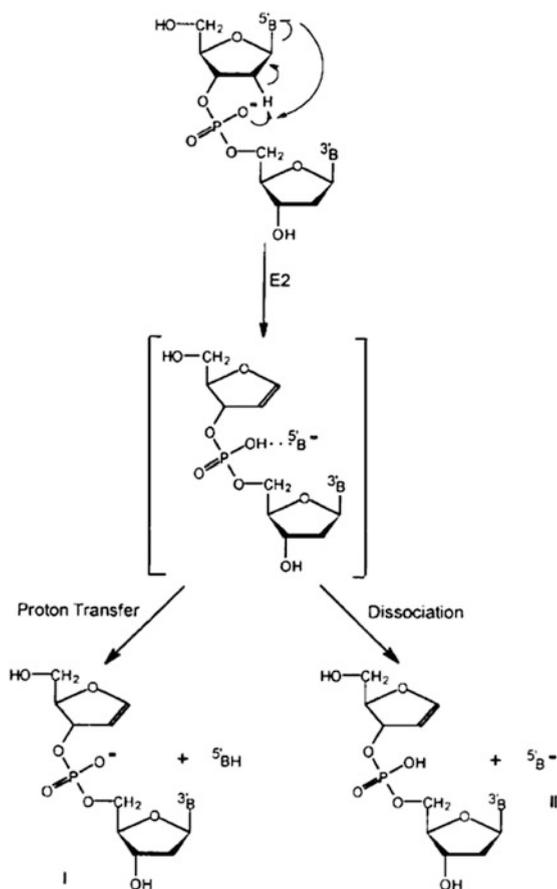
stabilization by interactions of charged sites with uncharged residues. Beam-type CID showed that backbone cleavage was observed at every residue, producing (a-B)- and w-type ions regardless of the phosphate being charged or not. Moreover, the w-type ions were relatively more abundant than those from the normal phosphodiester oligonucleotides. Therefore, two pathways for the backbone cleavages can occur under the multiple collision conditions of the quadrupole collision cell, one involving base loss followed by cleavage of the $3'$ C–O bond (formation of (a-B)/w-fragments) and the other undergoing direct cleavage of the $3'$ C–O bond (formation of a/w-fragments). Two mechanisms were proposed for latter pathway: (1) a proton (e.g., 4′-H) is transferred to a remote charged phosphodiester group (Scheme 6.5a) and (2) a proton is transferred to a MP oxygen at the cleavage site (Scheme 6.5b).

H/D exchange experiments [61] and methylphosphonate substitution experiments [62] supported a proposed formation of a zwitterionic intermediate in dissociation of relatively low charge DNA anions. Wan et al. [61] incorporated
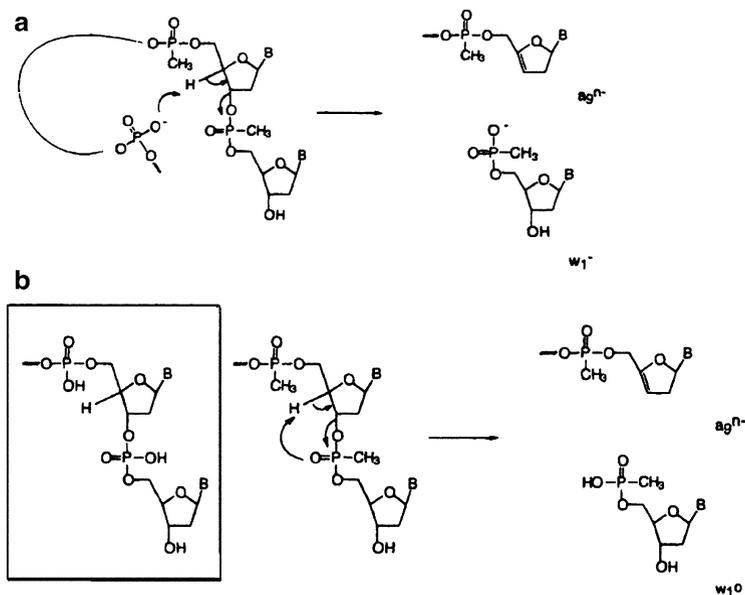
Scheme 6.5 (**a**) Proposed mechanism for transfer of charge from a remote phosphate to methylphosphonate w_1^- ion. Hydrogen transfer from the sugar is depicted as arising from C-4′ for convenience. The second reaction product shown is an intermediate which may further dissociate by base loss. (**b**) Alternative mechanism for chain cleavage at an uncharged site leading to neutral w_1 species. *Inset*: neutral phosphodiester, showing structural analogy to uncharged methlphosphonate. Reprinted with permission from [45]. Copyright 1996 John Wiley & Sons

bases with high proton affinity (PA) at different positions in T-rich tetramers and hexamers for systematic study of the base effect using a quadrupole ion trap. Oligonucleotides were sprayed in D_2O solution to enable H/D exchange. The nucleobases were removed from the single charged anions as deuterated neutral bases (BD). This result is not consistent with the previously proposed 1,2-elimination mechanisms [28, 33], because 1,2-elimination should involve the non-exchangeable 2′-hydrogen and would give rise to BH losses instead of BD losses. The authors postulated that the exchanged deuteron possibly came from a phosphodiester linkage adjacent to the base that was lost. Two mechanisms were proposed for formation of the w_3-ion in the case of d(GTTT): one w_3-ion was 2u less than the completely exchanged w_3 in the MS [3] spectrum of (M–D–GD)⁻ (Scheme 6.6), and the other was 1u less than the completely exchanged species (Scheme 6.7). In Scheme 6.6, the first step involves a deuteron being transferred to the high-PA base guanine. Given that the precursor ion does not have a 5′-phosphate group, the deuteron presumably comes from the nearby 3′-phosphodiester linkage and therefore forms a "zwitterionic intermediate." The negatively charged 3′-phosphodiester oxygen facilitates the elimination of GD via an E2 mechanism. Then, the neutral 3′-phosphate oxygen abstracts the 4′-sugar
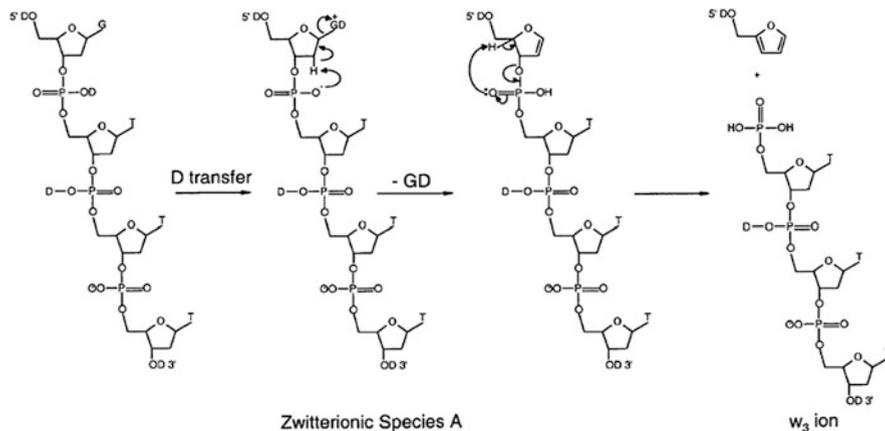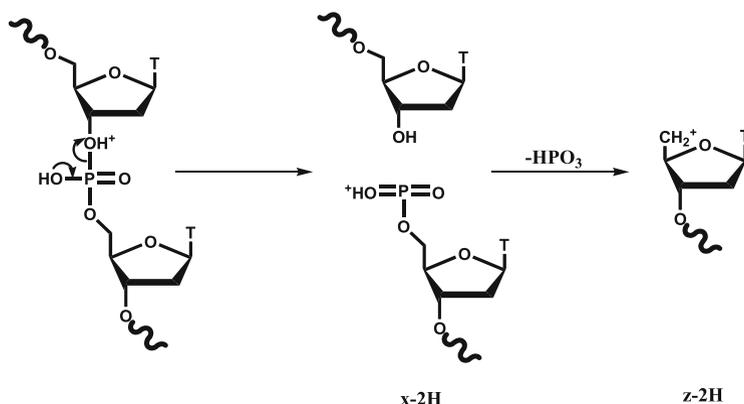
Scheme 6.6 Proposed mechanism for the formation of w_3 that is 2u less than completed exchanged. Reprinted from [61]. Copyright 2001 Elsevier, Inc., with kind permission from Springer Science and Business Media

proton and induces the $3'$ C–O bond cleavage. Another pathway of the w_3-ion involves direct $3'$ C–O bond cleavage as shown in Scheme 6.7.

For the DNA fragmentation, several mechanisms have been reported, each with experimental and/or theoretical support. However, none of them is adequate in explaining all reported phenomenology. For example, the base loss mechanism proposed by Wang et al. [60] was supported well by the CID experiments of T-rich DNA anions. However, it cannot explain either the preferential loss of neutral $5'$-base in dinucleotide anions or the dissociation of "completely" charged precursor ions. Likewise, the zwitterionic proton-bound intermediate mechanism was confirmed by H/D-exchange studies by Wan et al. [61] and can account for the preferential loss of neutral $5'$-base. Nevertheless, it fails to provide an explanation for the T^- loss from highly charged ions. Therefore, it is apparent that several mechanisms can contribute to the gas-phase fragmentation of oligonucleotides depending on the reaction conditions and ion charge state.

6.3.1.2 CID of DNA Cations
Phenomenology
Reports for gas-phase fragmentation of protonated DNAs are relatively few, presumably due to the fact that the negative mode tends to provide higher ion yields, at least under ESI conditions. Most positive ion studies were performed on small oligomers. Boschenok and Sheil [36] investigated the fragmentation pattern of protonated dinucleotides ($[M+H]^+$) using beam-type collisional activation. They observed both protonated base losses and neutral base losses, as well as w-type ions. Vrkic et al. [63] systematically examined CID patterns of all 64 protonated oligonucleotide trimers and 16 mixed-base tetramers by ESI-MS^n in a quadrupole ion trap. They found that the neutral base loss pathway is related to the relative proton

Scheme 6.7 Proposed mechanism for the formation of w_3 that is 1u less than completely exchanged. Reprinted from [61]. Copyright 2001 Elsevier, Inc., with kind permission from Springer Science and Business Media
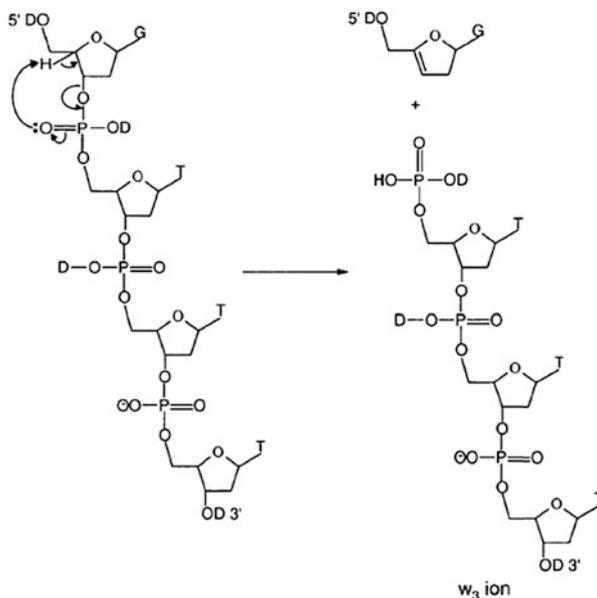
affinity of the nucleobase. Wang et al. [64] studied the CID behavior of positively charge oligonucleotides up to 20-mer level and found that charging levels of oligonucleotides were largely dependent on base composition. Ni et al. [65] confirmed that positive ion CID generates similar product ion types as negative ion CID, namely, (a-B)- and w-type ions. The only exception comes from poly-T, which yields x-2H and z-2H ions exclusively. Weimann et al. [66] examined the high-energy CID behavior of oligonucleotides (6- to 10-mers) on a hybrid EBE-TOF mass spectrometer. Some differences from those reported with triple quadrupoles were observed, presumably due to the significant variations in collision energy and timescale of dissociation. The reader is referred to the studies just mentioned for details of the respective observations.

Fragmentation Mechanisms

Several groups [42, 64, 65] have shown that the CID product spectra of DNA cations qualitatively resemble the CID product spectra of the corresponding anions, showing (a-B)/w-fragment ions. Thus, Wang and coworkers [64] postulated that the CID fragmentation mechanisms of protonated and deprotonated DNAs may share some commonalities. However, DNA cations are believed to be, in most cases, protonated on the nucleobases, and such protonation appears to facilitate the cleavage of the glycosidic bond resulting in base loss and subsequent formation of (a-B)/w-ions. PolydT, being the only exception, disfavors base losses due to the low proton affinity of thymine. CID of polydT cations leads to formation of x-2H and z-2H series ions exclusively, while (a-B)/w-ions dominate the negative ion CID spectrum. This exception has been rationalized in that instead of being protonated

Scheme 6.8 Proposed mechanism for the formation of x-2H and z-2H ions from poly-T oligonucleotides. Adapted with permission from [64]. Copyright 1997 John Wiley & Sons

on the thymines, the polydT are protonated on phosphate backbone. Wang et al. suggested a mechanism with protonation on the bridging oxygen of the phosphodiester linkage attached to the 3′-carbon to form C-OH$^+$-P. The cleavage of the P–O bond leads to charge transfer to the metaphosphoric acid ester group and yields an x-2H ion (Scheme 6.8). Then, loss of the neutral metaphosphoric acid (HPO$_3$) from the x-2H ion forms the corresponding z-2H ion. Furthermore, the absence of the TH$_2^+$ ion in the product ion spectrum of polydT supported the reasoning that thymine is not protonated.

Vrkic et al. [63] proposed two possible pathways for the general base loss from a protonated DNA (Scheme 6.9). The first mechanism is similar to the one that was proposed by Beauchamp and coworkers [67], the base is lost via an E1 reaction, in which a carbocation intermediate is generated followed by the loss of a neutral base to form the oxonium ion (Scheme 6.9a). This pathway does not require the previously proposed 3′-phosphate facilitation in the negative ion CID, allowing for the possibility of the 3′-base loss. The second mechanism occurs via an intramolecular E2 reaction, in which the base is lost simultaneously with the formation of a furan double bond. 5′ or internal base loss can go through either the E1 or E2 mechanism. The authors also proposed a mechanism for the formation of (a-B)/w sequence ions from the [M + H − BH]$^+$ ions (Scheme 6.9a) that involves an ion-molecule complex undergoing either proton transfer or direct dissociation (Scheme 6.10). The instability of the oxonium ion leads to 3′ C–O backbone cleavage, forming an ion-molecule intermediate. This complex can either dissociate directly to yield (a$_1$-B$_1$)$^+$ or undergo intermolecular proton transfer to form a w$_2^+$ ion. Vrkic et al. also postulated a mechanism for the formation of the "(a$_3$-B$_3$)$^+$ ion" that results from water loss from the [M + H − B$_3$H]$^+$ ion of the trimers. The loss of 3′-base via the E1 mechanism produces an intermediate, which can then lose the 3′-OH as a neural water molecule. The key concepts for the ion trap CID fragmentation mechanisms are summarized as follows: (1) the mechanisms are charged
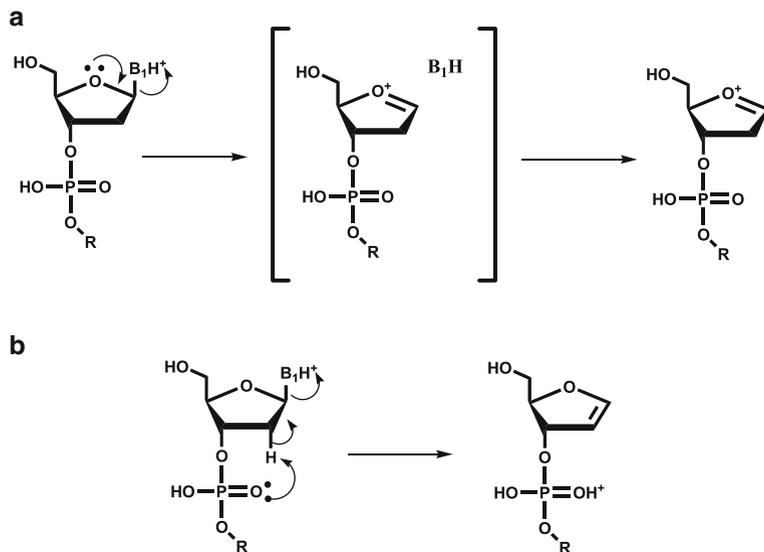
Scheme 6.9 Proposed mechanism for base loss from a protonated oligodeoxyribonucleotide. Adapted from [63], Copyright (2000), with permission from Elsevier

directed; (2) two complementary fragment ions (a-B) and w are competing for the charge via an ion-molecule complex intermediate, and (3) the "mobile proton model" [68, 69], in which a labile proton is able to relocate from one potential cleavage site to another to effect fragmentation, may also be applied in explaining the fragmentation of protonated oligonucleotides.

6.3.1.3 CID of RNA Anions
Phenomenology
Compared to DNA fragmentation, relatively few observations have been made on the dissociation of RNA, most of which were made in the past decade. Schürch et al. [70] examined a series of multiply deprotonated RNA and mixed-sequence RNA/DNA oligonucleotides (5-mer) upon beam-type CID in a hybrid Q-TOF mass spectrometer. They found that the loss of either a charged or a neutral nucleobase is the initial dissociation process, which is similar to the DNA analogs. Loss of neutral nucleobases is preferred for the $[M–2H]^{2-}$ precursor ions, while triply and quadruply charged ions favor the loss of a charged nucleobase. Both neutral and charged base losses predominantly favor the release of the 5′-terminal base, whereas the 3′-base loss is inhibited. Nucleobases with higher proton affinity (C ≫ T > U) are preferentially eliminated. Loss of the base on the 3′-terminus becomes favorable only if the 3′-base has a high proton affinity, e.g., loss of cytosine from UUUUdC. They also found that the 2′-substituent of the ribose is not significantly related to the preference for neutral or charged base loss. The most distinct dissociation pathway of RNA involves backbone fragmentation characterized by the formation of
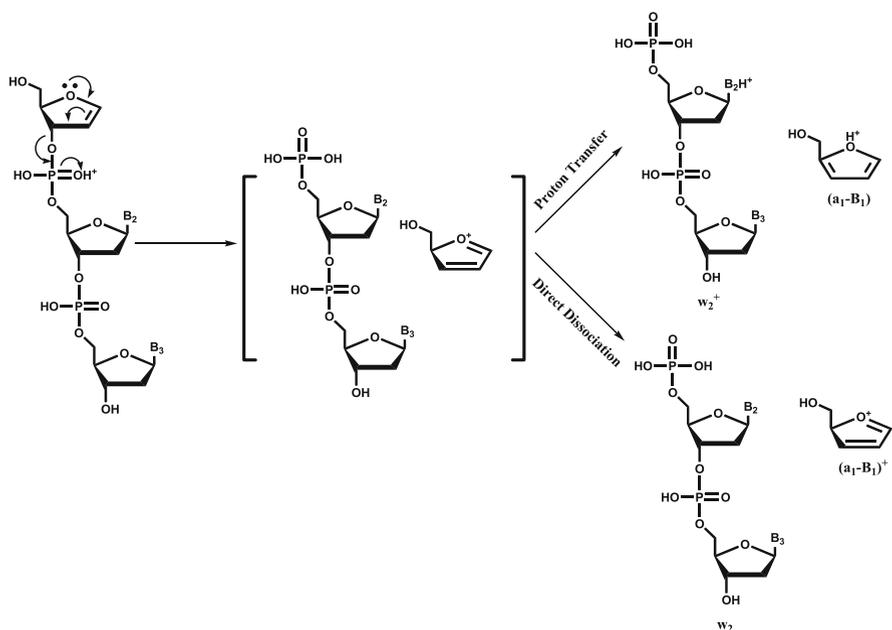
Scheme 6.10 Proposed mechanism for formation of (a-B)/w-fragments from a protonated oligodeoxyribonucleotide. Adapted from [63], Copyright (2000), with permission from Elsevier

abundant c-ions and their complementary y-ions as the major sequence-informative series (Fig. 6.3). Another difference from DNA is that backbone dissociation of RNA is not correlated with base loss. In the case of mixed-sequence RNA/DNA oligonucleotides, CID leads to a combination of DNA-typical (a-B)/w-ions and RNA-typical c/y-ions depending upon the type of nucleotide adjacent to the 5′-side of the cleavage site. The authors suggested that the 2′-hydroxyl group is involved in the fragmentation mechanism because it is the distinguishing structural feature of RNA from DNA.

Later, Tromp and Schürch [71] further investigated low-energy CID behavior of oligoribonucleotides and compared the results with their 2′-modified counterparts using a Q-TOF instrument to investigate their fragmentation pathways. Their findings have established the importance of the 2′-hydroxyl substituent as the structural key element in the fragmentation mechanism of RNA. As opposed to DNA, which preferentially dissociates into (a-B)/w-type ions, the presence of the 2′-hydroxyl group of RNA results in formation of predominant c/y-type ions. They suggested that the electron-withdrawing 2′-substituent induces a stabilizing effect on the N-glycosidic bond by impeding the formation of the carbocation at the 1′-position [72] and therefore makes nucleobase loss less prominent. Moreover, the availability of a proton in the vicinity of the phosphate group is necessary for the RNA-typical backbone cleavage. This was supported by comparing the fragmentation patterns of an RNA tetramer 5′-UAUU-3′ to its 2′-O-methyl and 2′-fluoro
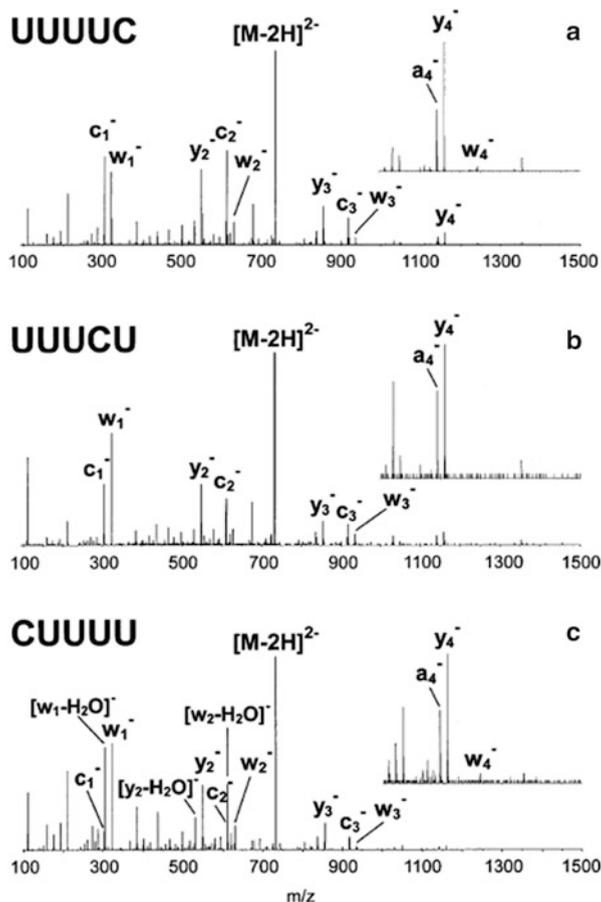
substituted analogs. When the 2′-hydroxyl proton is not available, the phosphodiester linkage is cleaved at all possible bonds, generating all types of fragments, i.e., a/w-, b/x-, c/y-, and d/z-ions. In addition, the experiments on oligoribonucleotides incorporating the dSpacer, a 2′-deoxyribonucleotide without the nucleobase, showed that gas-phase dissociation of the RNA backbone is not strongly influenced by the nucleobases (Fig. 6.4).

Huang et al. [73] studied the dissociation of model RNA anions as a function of charge state and resonance excitation amplitude using ion trap collisional activation. They found that the information from tandem mass spectrometry can be affected by the precursor ion charge state and excitation energy. Generally, formation of c/y-fragments and base losses are two of the preferred dissociation pathways for RNA anions. The c/y-ion signal is predominant in the product ion spectra from low to medium charge states when low excitation amplitudes are applied. However, contributions from cleavages of the other backbone bonds increase at higher excitation amplitudes, including abundant DNA-typical (a-B)/w-ion series at high
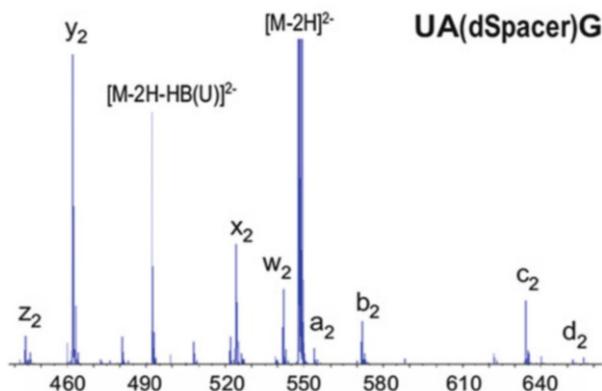
Fig. 6.4 Product ion spectrum of the doubly deprotonated UA(dSpacer)G (m/z 548.11), which incorporates a building block lacking the nucleobase. The spectrum predominantly show the RNA characteristic y_2-ion with m/z 462.12. Other fragments indicate backbone cleavage at alternative positions. Reprinted from [71], Copyright 2005 Elsevier, Inc., with kind permission from Springer Science and Business Media

excitation amplitudes. The dissociation of model dinucleotide systems as a function of activation amplitude showed that the two consecutive cleavages yielding (a-B)/w-ions are almost identical for RNA and DNA systems. The presence of the 2′-OH group has little influence on the energetic requirements for either base loss or the subsequent 3′ C–O bond cleavage when intramolecular proton transfer to a nucleobase does not or cannot occur (as in the case with deprotonated dinucleotides). However, the comparison of protonated RNA and DNA dinucleotides suggested that the 2′-OH group of the RNA tends to impede the base loss channel in positive ion CID. Thus, when intramolecular proton transfer can occur in RNA oligonucleotide anions, most likely at lower charge states, the base loss channel may be affected by the 2′-OH group. Yet, the c/y-fragmentation channel is significantly facilitated by the presence of the 2′-OH group regardless of the protonation state of the nucleobase. The authors suggested that this process is either more facile or comparably facile to base loss. Higher excitation amplitudes were required for production of (a-B)/w-ions than for c/y-ions presumably because the generation of the former ions involves a two-step process which requires more energy. Other backbone cleavages, such as b/x- and d/z-ions, were also observed at relatively low abundances along with the more abundant c/y and (a-B)/w ions at higher excitation amplitudes.

Most early efforts were focused on elucidating the fragmentation mechanism of RNA ions using small model oligoribonucleotides. With small oligomers, ion trap and beam-type activation show little difference in terms of fragment ion types [71, 73]. However, beam-type CID tends to generate complex spectra with fragment ions from two dominant cleavage mechanisms as well as secondary dissociation for large RNA sequences, which inevitably complicates the interpretation of the data. McLuckey and coworkers [74] have investigated the potential of ion trap

CID in generating sequence-informative fragment ions for single-stranded and duplex siRNA anions. They investigated the dissociation behaviors as a function of charge states and resonance excitation amplitude. The siRNA duplexes predominantly dissociated into two complementary single strands through the breaking of noncovalent bonds. The partitioning of the total negative charges into the single-stranded counterpart is found to be sequence dependent. This is consistent with Madsen and Brodbelt's [57] observation for DNA duplexes. By increasing the translational energy in beam-type collisional activation or excitation amplitude in ion trap collision activation, fragment ions from both strands were observed. The single-stranded siRNAs could be fragmented specifically at the 5′ P–O bond and the N-glycosidic bond when precursor ions of relatively low charge states were subjected to ion trap collisional activation at relatively low excitation amplitudes. As the excitation amplitude increased, the other sequence-informative channels, such as 3′ C–O bond cleavages, began to contribute and could therefore be observed (Fig. 6.5).

Fragmentation Mechanisms

Given the fact that the 2′-hydroxyl group of the RNA backbone is the only difference from that of DNA, Schürch et al. [70] suggested that the mechanism for the formation of the major backbone fragmentation products, c/y-ions, should involve the 2′-substituent. The proposed fragmentation mechanism is described in Scheme 6.11a. Backbone fragmentation is initiated by forming an intramolecular cyclic intermediate with 2′-hydroxyl proton bridged to the 3′-phosphate oxygen. The 2′-hydroxyl proton is abstracted from the oxygen and facilitates the cleavage of the 5′ P–O bond, yielding c/y-type fragments. The formation of a 3′-metaphosphoric-acid ester group following subsequent bond rearrangement stabilizes the c-fragment. The complementary y_1-fragment is released as a neutral molecule, and thus will not be detected.

The Schürch mechanism [70] (Scheme 6.11a) was further supported by fragmenting several 2′-modified RNA analogs under similar conditions on a Q-TOF mass spectrometer. Tromp et al. [71] found that by replacing the 2′-hydroxyl group with 2′-O-methyl or 2′-fluoro, the preferential c/y-fragment channel is inhibited and all backbone cleavages are observed. This provides evidence for the necessity of the 2′-hydroxyl group for c/y-ion generation. Moreover, the authors also examined an alternative mechanism similar to the one known for backbone cleavage of RNAs in solution, which could also lead to the same RNA characteristic c/y-type fragment ions (Scheme 6.11b). Although there is no free proton acceptor present in the gas phase, the proton from the 2′-hydroxyl group might still be abstracted by the adjacent 3′-nucleobase via intramolecular proton transfer. A transesterification step follows the deprotonation of the 2′-hydroxyl group, in which the 2′-oxygen undergoes nucleophilic attack by the phosphorus, forming a cyclic intermediate. Then, the cleavage of the phosphodiester backbone is accomplished by bond rearrangement and protonation of the leaving group (y-ion) by the protonated nucleobase, yielding the complementary c/y-fragment ions. Ideally, a 3′-side nucleobase to the given 2′-ribose hydroxyl group is needed as
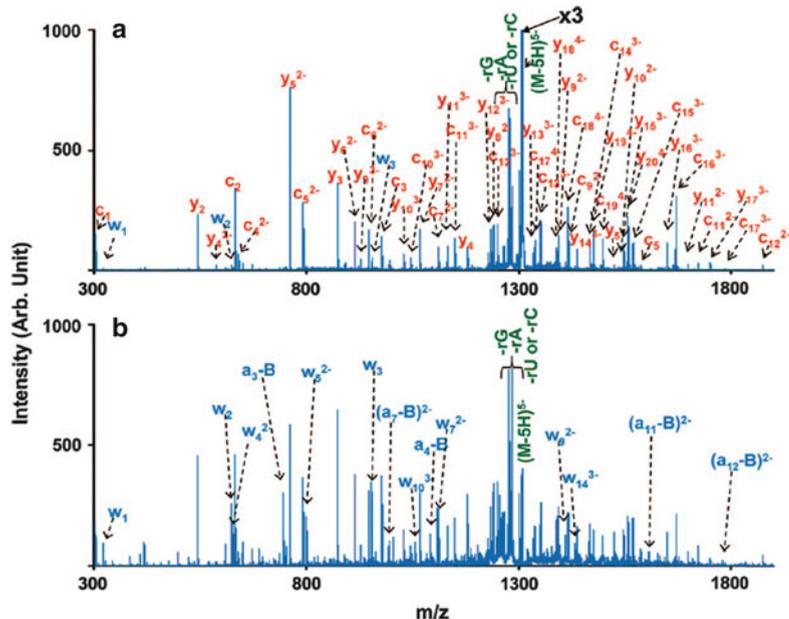
Fig. 6.5 Ion trap CID of the antisense siRNA anions [AS siRNA]$^{5-}$ under relatively low amplitude (**a**) 99.93 kHz, 480 mV, 100 ms and relatively high amplitude (**b**) 99.93 kHz, 800 mV, 100 ms conditions. The c/y-ion series labeled in panel a were not labeled in panel b to avoid figure congestion. Reprinted with permission from [74]. Copyright 2008 American Chemical Society

the proton acceptor for the intramolecular proton transfer. Tetranucleotide UA (dSpacer)G, incorporating a 2′-deoxyribonucleotide lacking the nucleobase, is used to verify such an alternative mechanism. The dSpacer building block enables evaluation of the role of the nucleobase in the RNA fragmentation mechanism, as the dSpacer does not provide the 3′-nucleobase. If the nucleobases were involved in the dissociation of RNA, the CID behavior of UA(dSpacer)G would be different from normal RNA anions. However, 5′ P–O bond cleavage dominates the product spectrum (Fig. 6.4), and this could rule out the possibility for the mechanism that involves the 3′-nucleobase as proton acceptor and support the mechanism in which backbone dissociation is initiated by formation of an intramolecular cyclic intermediate with the 2′-hydroxyl proton bridged to the 5′-phosphate oxygen.

6.3.1.4 CID of RNA Cations

Phenomenology

Kirpekar and Krogh [75] have studied the fragmentation behavior of short, singly protonated oligoribonucleotides generated by Q-TOF instrument to locate posttranscriptional modifications of RNA molecules. Complex MS/MS spectra were generally observed, showing nearly all types of cleavages along the phosphodiester
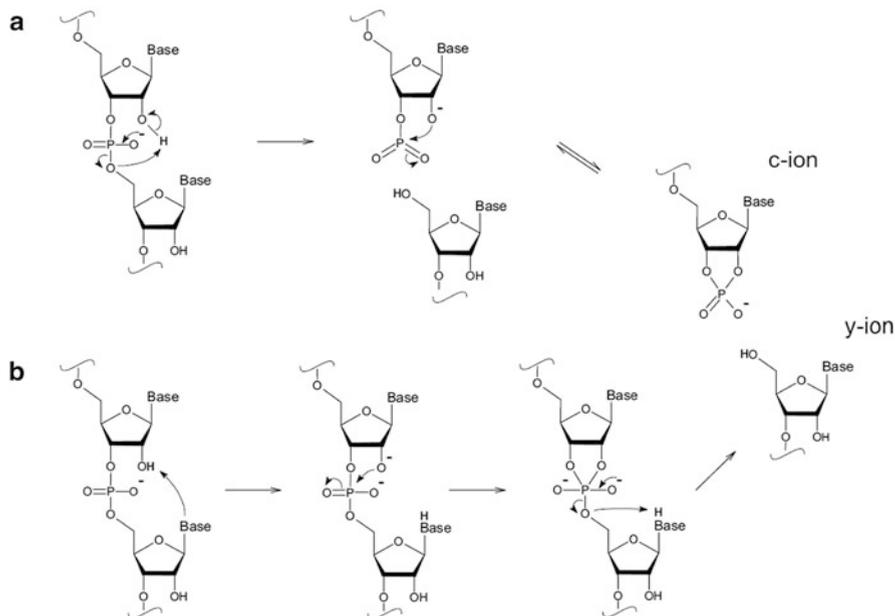
Scheme 6.11   Dissociation mechanisms of RNA. (**a**) Previously proposed mechanism describing the dissociation of RNA. (**b**) Alternative dissociation mechanism similar to the one known for backbone cleavage of oligoribonucleotides in solution. Both mechanisms include the 2′-hydroxyl proton and lead to the RNA-typical fragment ions. In contrast to the mechanism (**a**), the alternative mechanism (**b**) does involve the nucleobase. Reprinted from [71], Copyright 2005 Elsevier, Inc., with kind permission from Springer Science and Business Media

backbone and of the N-glycosidic bonds (and combinations of these) at different relative abundances. The most prevalent backbone fragment ions were found to be c- and y-ions, whereas the least prevalent ions were b/x-ions. Loss of neutral cytosine and guanine required low-energy deposition and therefore were preferred, whereas neutral loss of adenine was less prevalent. Loss of uracil, as either a neutral or cation, was never observed. By varying the collision energy, ions arising from loss of neutral GH and CH and that of y-sequence ions are typically the first fragments to appear. This fragmentation behavior is distinct from the that of singly protonated DNAs, where loss of a nucleobase is the key initiating step followed by predominant cleavage of the 3′ C–O bond, resulting in (a-B)/w-fragments. As suggested by Tang et al. [72], the 2′-hydroxyl stabilizes the N-glycosidic bond in RNA, making it less susceptible to cleavage upon protonation of the nucleobase. As a result, the energies required to generate MH^+-GH, MH^+-CH in RNA cations are elevated to the similar level as the c/y-formation channel.

For elucidation of the gas-phase unimolecular dissociation behavior of RNA and DNA cations, Andersen et al. [76] performed hydrogen/deuterium exchange on a series of RNA and DNA tetranucleotides and studied their fragmentation patterns on a MALDI TOF-TOF instrument. They observed the loss of a neutral nucleobase

as a fully deuterated species for all RNA tetramers, which is consistent with the DNA result reported by Gross et al. [77]. They have also found that all y-ions and all c_2- and c_3-ions of the RNA tetramers analyzed are observed as fully deuterated species. Therefore, no carbon-bound hydrogens are transferred to mobile sites in the cleavage reaction. Moreover, nucleobase loss or other hydrogen mobilizing reactions generally do not precede cleavage of the 5′ P–O bond. Initial loss of a nucleobase would have induced peak splitting of the c/y-ion fragments due to intramolecular H/D exchange before 5′ P–O cleavage. In addition to the significant difference in fragmentation pathways between RNA and DNA (c/y-fragment vs. (a-B)/w-fragments), water loss from the precursor ion is only observed for RNA ions, implying that this water is lost from or by the interaction of 2′-OH. Water loss was not observed in the fragmentation of the 2′-O-methylated UAUC tetramer, further indicating the involvement of 2′-OH in the reaction. Tandem mass spectrometry of the ribonucleosides shows no evidence of water loss and suggests that the presence of 2′-OH alone is not sufficient to induce the loss of water. Other reactive groups present in RNA, most likely phosphate, are required in the reaction as well.

Fragmentation Mechanisms

Andersen et al. [76] have proposed a common RNA-specific mechanism to explain the water loss as well as 5′ P–O cleavage (Scheme 6.12). This mechanism involves an intramolecular nucleophilic attack on the phosphorus atom by the adjoining 2′-OH to yield a phosphorane transition state. The phosphorane transition state is well documented in the mechanism of RNase [78]. In the gas phase, the deuteron abstracted from the 2′-OD presumably delocalizes between the oxygens of the phosphorane. Two reaction (1 and 2) pathways could occur depending on which oxygen acts as the leaving group: (1) The delocalized deuteron leaves with the –OD group leading to neutral loss of D_2O and formation of an internal 2′-3′-cyclic phosphate (reaction 1 in Scheme 6.13) and (2) 5′ P–O cleavage and formation of a y-fragment and a c-fragment with a 3′-terminal 2′-3′ cyclic phosphate (reaction 2). The mechanism of Scheme 6.13 agrees well with the detection of all deuterated c- and y-ions from the deuterated precursor and also accounts for the neutral loss of water observed only for the RNA precursors. Furthermore, this mechanism explains why the water loss from deuterated RNA is a loss of D_2O and not HDO, which would otherwise be expected from a simple dehydration reaction.
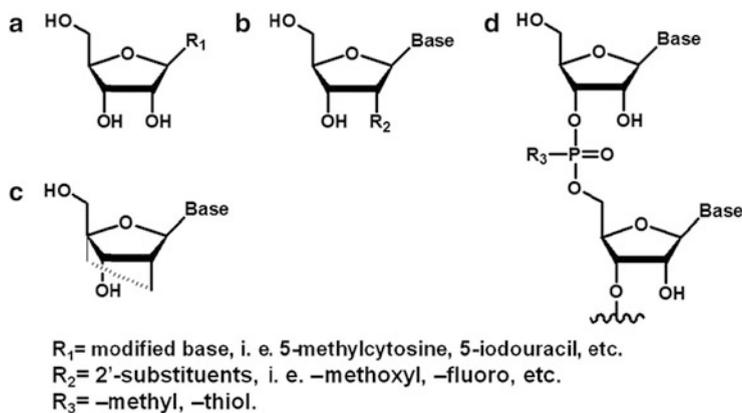
6.3.1.5 CID of Chemically Modified Nucleic Acids

Phenomenology

Chemically modified oligonucleotides have been extensively studied in biomedical and pharmaceutical research due to their potential antisense or RNA interference (RNAi) applications [79–82]. Most antisense and RNAi compounds are heavily chemically modified oligonucleotides, such as analogs with one or more unnatural nucleobases, substituted ribose (2′-substitution or 2′-4′ bridging), or modified phosphodiester linkages [83–87] (Scheme 6.13). Traditional enzymatic digestion and chemical degradation are limited for characterizing the modified
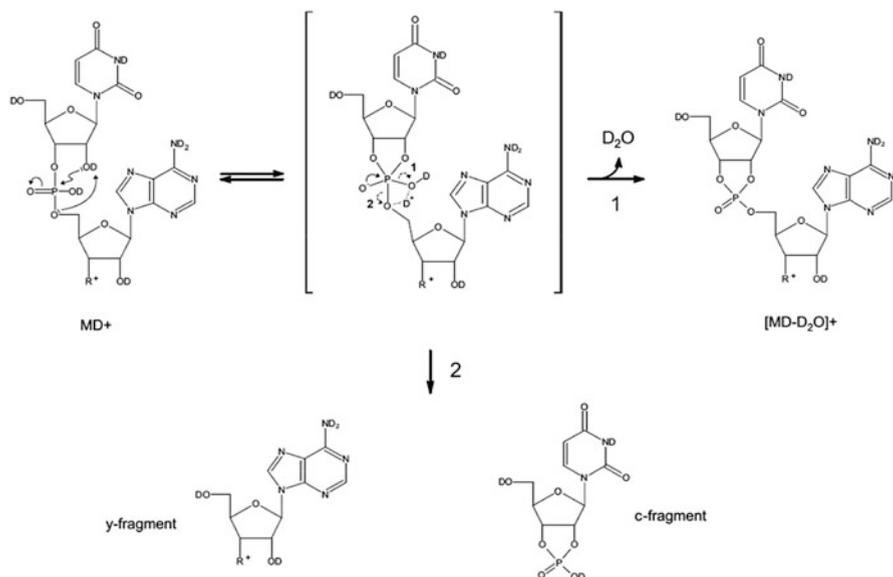
Scheme 6.12 Proposed fragmentation scheme for c- and y-ion formation from a deuterated RNA precursor. Reprinted from [76], Copyright 2005 Elsevier, Inc., with kind permission from Springer Science and Business Media

R_1= modified base, i. e. 5-methylcytosine, 5-iodouracil, etc.
R_2= 2'-substituents, i. e. –methoxyl, –fluoro, etc.
R_3= –methyl, –thiol.

Scheme 6.13 Chemical modifications of oligonucleotides. (**a**) Modified nucleobase; (**b**) 2'-substitution; (**c**) locked nucleotide (2'-4' bridging); (**d**) modified phosphodiester linkage

oligonucleotides due to their enhanced stability in solution. Direct sequencing of such oligonucleotides by tandem mass spectrometry is attractive because of potential advantages in speed and sensitivity.

CID patterns of chemically modified DNAs and RNAs have been investigated in various contexts [29, 33, 88–93]. Early studies have been focused on modified

nucleobases. McLuckey and Habibi-Goudarzi [29] examined two isomeric oligodeoxyribonucleotide hexamer anions with the ESI-MS^n approach. They found that the preference for a particular base loss is important in locating the modified bases. A report by Baker et al. [94] showed the CID of a multiply protonated permethylphosphonate antisense oligonucleotide. Recently, efforts have been directed to elucidating the fragmentation patterns of nucleic acids with modified sugar or modified phosphate backbones, due to their potential pharmaceutical and biological applications [95–97]. Tromp and Schürch [71] investigated oligonucleotide anions with different 2′-substitutions, including 2′-O-methyl, 2′-fluoro, and 2′-hydroxyl in a Q-TOF mass spectrometer. These modifications led to production of all possible types of backbone fragments in relatively similar abundance. Monn and Schürch [88] revisited the unimolecular fragmentation mechanisms of nucleic acids by tandem mass spectrometry of unmodified and methylphosphonate-modified oligonucleotides. Gas-phase dissociation of locked nucleic acids (LNA) and LNA-DNA chimeras has been studied by Huang et al. [89] using ion trap CID. Recently, Gao and McLuckey have reported the unimolecular fragmentation behavior of oligonucleotide mix-mers with 2′-O-methyl and 2′-fluoro modification [92] as well as the phosphorothioate (PS) modification [93]. Smith and Brodbelt [91] investigated the positive ion dissociation pattern of a partially 2′-O-methyl-modified single-strand 12-mer and an entirely phosphorothioate-modified 12-mer with various energy deposition methods.

The striking differences between the fragmentation pathways of DNA and RNA reflect the importance of 2′-substituents on the ribose. In order to further evaluate the influence of different 2′-substituents, Tromp and Schürch [71] analyzed several tetraribonucleotides incorporated with a 2′-O-methyl-ribose or a 2′-fluoro-ribose. Figure 6.6 shows the product ion spectra of [M–2H]$^{2-}$ of U(2′-OMeA)UU and U(2′-FluoroA)UG under beam-type activation conditions. Although the first and third phosphodiester linkages fragment into the RNA-typical c- and y-ions, no preferred dissociation pathway has been observed for the second phosphodiester linkage and all four possible decomposition channels appear in the product spectra with similar abundance. This experiment also shows that the electronegativity of the 2′-substituent helps to stabilize the N-glycosidic bond. The effect is noted by the decreased formation of the w_2-ion (m/z 668.10) in Fig. 6.6b. These findings were later confirmed by Nyakas et al. [90] and Gao and McLuckey [92]. Oligonucleotide mix-mers that do not have the same 2′-substituents are observed to have "cleavage gaps" due to significant differences in the energetic requirements associated with cleavages of the adjacent phosphodiester linkage. For instance, the presence of DNA moieties in the sequence, along with residues with 2′-OMe and 2′-fluoro substituents, can result in a situation where all of the backbone decomposition occurs via the low-energy channels associated with DNA, whereas little or no structurally diagnostic cleavages take place on the 3′-sides of the 2′-OMe and 2′fluoro substituted residues. The gas-phase stability of the 3′-side phosphodiester bond follows the order: 2′-fluoro > 2′-OMe > 2′-OH > 2′-H according to a systematic study of a series of chemically modified 21- and 23-mers [98].

Fig. 6.6 (**a**) Product ion spectrum of the doubly deprotonated modified tetraribonucleotide U (2′-OMeA)UU with m/z 598.60. The spectrum gives no evidence of a preferred dissociation pathway. (**b**) Product ion spectrum of doubly deprotonated U (2′-FluoroA)UG with m/z 612.11. The a_2 fragment ion (m/z 556.10) is overlapped by the very abundant [M–2H–HB(U)]$^{2-}$ ion. Reprinted from [71], Copyright 2005 Elsevier, Inc., with kind permission from Springer Science and Business Media
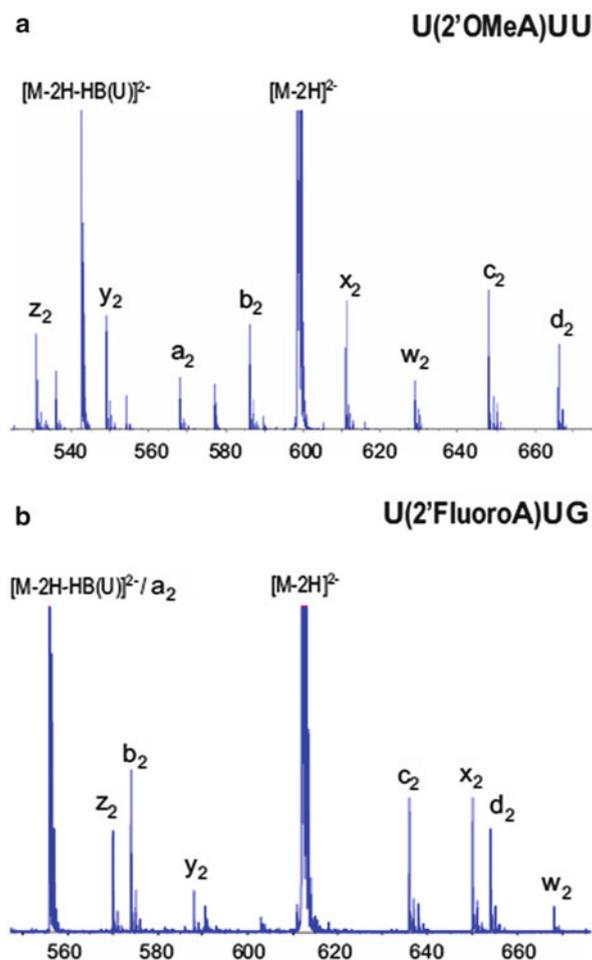
Huang et al. [89] studied the dissociation behavior of (locked nucleic acid) LNA anions under ion trap collision activation conditions. They examined model LNA 5- and 8-mers containing all four LNA monomers in the sequence and noted the cleavage of all backbone bonds, generating a/w-, b/x-, c/y-, and d/z-type fragments. No significant preference for fragment channels was observed and complete sequence coverage was achieved. Similar to the results from completely 2′-modified oligonucleotides, (a-B)-ions were not observed in high abundance, unlike their DNA counterparts [71, 92]. MS/MS of the [M–2H–GH]$^{2-}$ ion of the LNA 8-mer (5′-ATC^{m}GATC^{m}G-3′), where the "m" superscript represents the 5-methylation on cytosine, indicated that subsequent cleavage to yield the (a$_4$-G)- and w$_4$-ions, while observed, was not the dominant peaks, as expected based on the DNA fragmentation mechanism [14]. Instead, consecutive losses of two water molecules were the predominant dissociation channels. Low abundance fragment

peaks from the subsequent loss of neutral bases (-C^mH > -($C^mH + 18$) > -AH > -GH) were also observed. This result supports the conclusion that the modification associated with LNA inhibits the subsequent cleavage of the 3′ C–O bond cleavage after the base loss in DNA anions. CID of the LNA-DNA chimeras generated similar "gapped" fragmentation pattern as the 2′-modified mix-mers [92]. The 3′-backbones of the LNA subunits are much more stable than those of the DNA subunits at lower charge states and therefore remain intact upon collision activation when DNA subunits are present in the same sequence. However, with higher precursor ion charge states and excitation amplitudes, more sequence-informative fragments can be generated by cleavages at the LNA residues.

Oligomers with modifications/alterations of the phosphodiester linkages have also been investigated. Monn and Schürch [88] investigated the partially backbone-modified pentanucleotides TxGxGGG and TxGxGxGG and the fully modified TxGxGxGxG, where x denotes the position of a methylphosphonate modification. In this model, the location of the charge is a key for the fragmentation pathways. With a natural phosphodiester backbone, the negative charge is likely present on a deprotonated phosphate group, whereas for the methylphosphonate-modified backbone deprotonation is highly unlikely. They noticed that the methylphosphonate-modified oligonucleotides exhibit reduced stability compared to their natural counterparts. No (a-B)-ions were observed in the product spectrum of TxGxGxGxG (Fig. 6.7), while a complete w-fragment ion series was observed with similar peak abundances in the spectra of natural and modified oligonucleotides. d/z-ion series were also observed with significant abundances. The result suggests that the formation of (a-B)-type fragments is dependent on the deprotonated phosphate group. They further supported this by performing CID on two partially modified oligomers, TxGxGGG and TxGxGxGG. The two sequences differ only by the backbone linkage between the third and fourth nucleosides. The absence of the (a_4-B_4)-ion in the product ion spectrum of TxGxGxGG supports the DNA-cleavage mechanism proposed by Wang et al. [60], whereas the (a_4-B_4)-ion with m/z 1055.2 appears in the product spectrum of TxGxGGG (Fig. 6.8). Gao and McLuckey [93] examined the ion trap CID of mixed-backbone oligonucleotides (MBOs) in which the phosphodiester backbones are partially replaced with phosphorothioate. Ion trap CID of doubly deprotonated PS 6-mer (5′-dA*dCdC*dGdA*dG-3′) and PS 2′-OMe 6-mer (5′-mA*mCmC*mGmA*mG-3′) resulted in extensive backbone cleavages on the PS linkages. For CID of the PS 6-mer, complete series of (a-B)/w-ions were observed, consistent with the DNA fragmentation mechanism [14], whereas in the product spectrum of PS 2′-OMe 6-mer, cleavages from only the PS bonds were observed.

Modified oligonucleotides have also been investigated in the positive polarity. Smith and Brodbelt [91] examined the positive ion dissociation patterns of a partially 2′-O-methyl-modified single-strand 12-mer and an entirely phosphorothioate-modified 12-mer with various energy deposition methods. CID of the triply protonated 2′-OMe-modified single-strand 12-mer generated a result similar to that of the negative ion CID, the 3′-side of the 2′-OMe sugar moieties being resistant to cleavage upon CID. However, the positive ion CID fragmentation
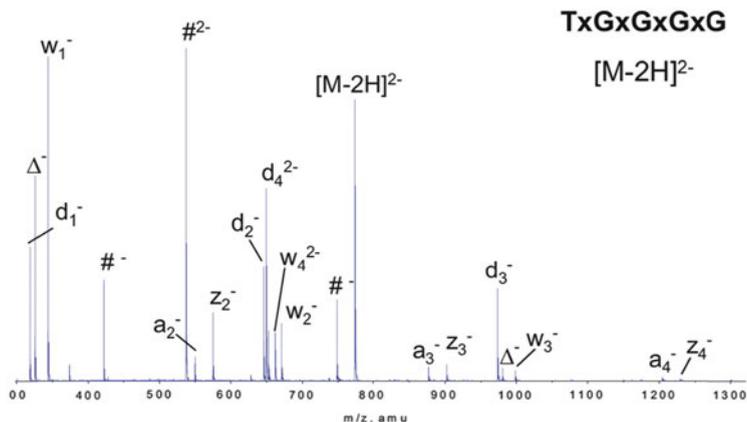
Fig. 6.7 Product ion spectrum of the fully methylphosphonate-substituted pentanucleotide TxGxGxGxG; a-B- and M–B-ions are not detected. Besides the w- and a-ions, abundant fragments of the d/z-series are observed. Further peaks indicate the presence of the internal fragment ions, generated by repetitive backbone cleavage, e.g., [w-d]- (#) and [z-d]/[a-w]-ions (Δ). Reprinted from [88], Copyright 2007 Elsevier, Inc., with kind permission from Springer Science and Business Media

pattern of the phosphorothioate-modified single-strand 12-mer (PSss12) showed completely different fragmentation behavior from the negative polarity. For PSss12 5′-TAGCTAGTCsGAC-3′, where s represents the phosphothioate, CID of the triply charge cation generated complete (a-B)/w-ion pairs for almost all phosphodiester linkages, except for the thymine due to its low gas-phase basicity [64]. However, fragment ions resulting from cleavage of the phosphorothioate linkage were observed in considerably low abundance. This difference is presumably due to the different locations that hold the charges in the two opposite polarities. Phosphorothioate has low gas-phase acidity so that it is more likely to be deprotonated in negative ESI, and therefore the PS linkages are more prone to fragmentation due to the proposed charge-directed mechanism [14]. In contrast, oligonucleotide cations are believed to be protonated on the nucleobase so that the PS modification does not facilitate the dissociation. The authors attributed the lack of cleavage of the phosphorus–oxygen bond on the 5′ side of the sulfur atom to the resulting change in basicity of the phosphorothioate site that suppresses the proton migration necessary for the adjacent C–O bond cleavage.

Fragmentation Mechanisms

The fragmentation mechanisms of 2′-modified nucleic acids have rarely been discussed. Most papers have compared the fragmentation patterns of the 2′-modified oligonucleotides and the corresponding DNA and RNA. It has been noted that the 2′-modification inhibits the facile base loss pathway and the subsequent 3′ C–O bond cleavage pathway because of the absence of a 2′-hydrogen.

Fig. 6.8 Product ion spectrum of TxGxGGG. Abundant peaks of w_3^{2-} and w_4^{2-} indicate that charge is preferably located on unmodified phosphate groups. # indicates the [w-d]-ions; * indicates M–B-ions. Reprinted from [88], Copyright 2007 Elsevier, Inc., with kind permission from Springer Science and Business Media
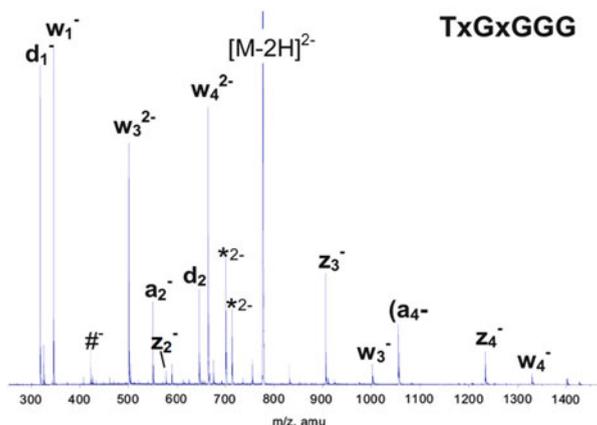
Similarly, the 5′ P–O bond cleavage is no longer favored for 2′-modified oligomers because the RNA fragmentation mechanism requires a 2′-hydroxyl group.

Recently, Schürch and coworkers [90] have suggested a mechanism for the formation of a/w-, b/x-, c/y-, and d/z-type fragments for the gas-phase decomposition of 2′-modified nucleic acid anions. Since the oligomer is completely 2′-modified, the authors noted that the presence of uninterrupted series of c- and y-ions is remarkable. According to the generally accepted mechanism for dissociation of unmodified RNA [70, 71], the intermediate for 5′ P–O bond cleavage pathway requires a bridging oxygen from the 2′-position to form a cyclic structure. However, such bridging oxygen is not available in the 2′-O-methyl substituent. Thus, they proposed an alternative dissociation pathway accessible for 2′-O-methyl oligoribonucleotides.

The formation of c/y-fragments is explained by the free electron pair of the phosphate oxygen rearranging randomly to either the 3′- or the 5′-side of the oligonucleotide upon CID (Scheme 6.14). The subsequent cleavage of either of the P–O bonds gives rise to the complementary c/y- or b/x-fragments. It is noted that the c- and x-ions contain a metaphosphoric acid ester group. Pathway A in this mechanism explains the exceptional abundance of b- and x-ions. The author suggested that the formation of (a-B)/w-type fragments can be explained by the known DNA-typical mechanism. The mechanism is further supported by MS [3] experiments of the b- and x-ions. Loss of a metaphosphoric acid ester group (HPO_3) from c_3^{2-} was observed for 2′-OMe CCGGUU, whereas the unmodified counterpart generated abundant neutral base loss species. Furthermore, the generation of the $_yG_c$- as well as $_xG_c$-fragments gives evidence for dissociation of the P–O bonds as suggested in Scheme 6.14 (the internal fragment ions are assigned in the way that the subscripted letters indicate the type of backbone cleavage and the capital letters represent the nucleobase(s) present).
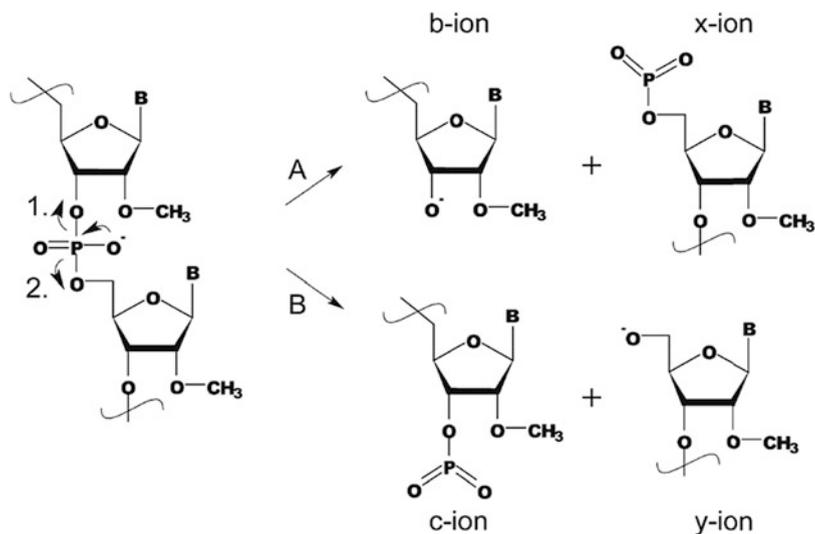
Scheme 6.14 Proposed gas-phase dissociation mechanism of 2′-*O*-methyl oligoribonucleotides resulting in either b/x- or c/y-ion pairs. Reprinted from [88], Copyright 2011 American Society for Mass Spectrometry, with kind permission from Springer Science and Business Media

6.3.2 Ion/Photon Interactions: Photodissociation of Nucleic Acids

6.3.2.1 Infrared Multiphoton Dissociation

Infrared multiphoton dissociation (IRMPD) [99, 100] involves multiple photon absorption over a relatively long time frame to induce fragmentation via vibrational excitation. Since low-energy CID and IRMPD both give rise predominantly to vibrational excitation, they are likely to lead to similar fragmentation mechanisms. The typical IRMPD experiment is implemented with a continuous wave CO_2 laser (10.6 μm) in an ICR cell [101]. The precursor ions are isolated and slowly heated to their dissociation threshold. The irradiation time required for fragmentation to occur is dependent upon the ion structure. IRMPD is well suited to the FTICR mass analyzer because it eliminates the need to introduce a collision gas for CID so that there is no pump-down delay, degradation of the high vacuum, or subsequent deterioration in analyzer performance. However, IRMPD can also be implemented in quadrupole ion traps [102, 103]. A variation of the IRMPD experiment is BIRD, which relies on blackbody radiation as the energy source, which has been far less used in the characterization of oligonucleotide ions.

IRMPD has been widely used in the identification and characterization of nucleic acid ions. McLafferty and coworkers [39, 40, 99] applied IRMPD to various DNA and RNA ions ranging in size from 8-mer to 108-mer. Complete sequence information of a 50-mer DNA was achieved [39]. They determined that oligonucleotides are more susceptible to infrared irradiation than peptides or proteins, which was later proved to be due to the strong absorption at 10.6 μm

(the typical wavelength used) by backbone phosphate groups [104, 105]. Later, Hofstadler et al. [106] demonstrated the use of a peptide as a stable internal mass standard to obtain accurate mass measurements of IRMPD-generated phosphorothioate oligonucleotide fragment ions by manipulating the duration and power of the IRMPD event. Sannes-Lowery and Hofstadler [107] examined the fragmentation behavior of modified oligonucleotides (2′-O-methoxylethyl ribose, phosphorothioate) using IRMPD in an external ion reservoir of an ESI-FTICR mass spectrometer. Keller and Brodbelt [108] have compared CID and IRMPD of oligonucleotide cations and anions in a quadrupole ion trap. Yang and Håkansson [109] demonstrated the use of double resonance experiments to interrogate the IRMPD fragmentation mechanism of DNAs. Their results indicated that IRMPD of the relatively low charge state ions examined proceeds through the zwitterionic intermediate mechanism proposed by Gross and coworkers [62]. Parr and Brodbelt [110] demonstrated that IRMPD can significantly promote the fragmentation of thymine-rich oligodeoxyribonucleotides. Gardner et al. [111] performed IRMPD on protonated and deprotonated single-stranded small interfering RNA as well as duplexes and compared the IRMPD result with the result from CID in a quadrupole ion trap. Generally, IRMPD of various oligonucleotides (DNA, RNA, modified oligonucleotides) shows significant resemblance to the respective CID results.

Sannes-Lowery and Hofstadler [107] explored the capability of IRMPD in sequencing and characterizing modified and unmodified oligonucleotides. They examined ten different sequences with various modifications (2′-deoxyribose, 2′-O-methoxylethyl ribose, phosphorothioate, and 5-methylated cytosine). The experiment was carried out in an external ion reservoir, in which fragment ion yield and sequence coverage can be significantly improved because ions are exposed to a range of laser irradiation times, and metastable ions are stabilized by the high gas pressure [103]. With IRMPD, they found that abundant (a-B)/w-fragment ions are generally observed for DNAs, regardless of whether the backbone comprises phosphodiester or phosphorothioate linkages. For 2′-modified oligonucleotides, b/x- and c/y-fragment ions provide significant sequence information. If there is a mixture of 2′-substituents, fragmentation preferentially takes place at 3′-backbone of the 2′-deoxy groups (DNA residues). Despite the 2′-substituent, backbone cleavage on the 3′-side of thymine is observed at either very low abundance or not at all. The authors attributed the effect of the 2′-substituents to the interaction between the 3′-phosphate and 2′-position. Both the hydroxyl group and the methoxyethyl group are electron withdrawing, which can stabilize the N-glycosidic bond and make the intramolecular proton transfer to the nucleobase unfavorable. In addition, the steric hindrance of the 2′-substituents may prohibit the attack of the 3′-phosphate at the 2′-position.

Keller and Brodbelt [108] examined IRMPD of deprotonated and protonated oligonucleotides (5- to 40-mers) using a quadrupole ion trap coupled with an IR laser. They found that IRMPD and CID produce comparable sequence-informative fragments. However, there are two major differences in outcome between the two techniques: (1) the predominant base loss channel in CID spectra is generally observed in low abundance in IRMPD spectra and (2) phosphate and nucleobase

ions can be observed directly in IRMPD experiments. The first variation comes from the different extents to which product ions undergo further excitation. Ion trap CID uses single-frequency dipolar excitation which targets a very narrow m/z range, and therefore, the generated products are not further activated and survive as first-generation fragments (bass loss ions). In the IRMPD experiment, further dissociation of first-generation product ions readily occurs in the IR laser beam since the fragment population is continuously activated during the irradiation period. The second difference arises from the fact that a higher low-mass cutoff is required in ion trap CID because the parent ion is accelerated. This is less of an issue for the IRMPD experiment since there is no need for parent ion acceleration, which allows for the use of a lower low-mass cutoff.

In order to gain further understanding of the IRMPD fragmentation mechanism, Yang and Håkanssen [109] applied Fourier transform ion cyclotron double resonance (FT-IC DR) experiments on a series of model DNAs. The FTICR DR experiment serves as an approach to provide information on intermediates in dissociation pathways, which involves selective excitation at particular cyclotron frequencies. If intermediate ions with corresponding m/z values exist in the ICR cell, they will be ejected and the corresponding product ions from this intermediate will disappear from the MS/MS spectra as long as the lifetime of the intermediate is long enough to be ejected by the excitation waveform [112]. Figure 6.9 shows results of DR experiments on the doubly deprotonated oligonucleotide ion, 5′-d (GCATAC)-3′. In the IRMPD spectrum without DR, complementary $(a_5\text{-A})/w_1$, $(a_3\text{-A})/w_3$, and $(a_2\text{-C})/w_4$ product ion pairs are observed as well as a w_5^{2-}-ion. When $[M-A]^{2-}$ is ejected during the IRMPD process, 838 % drop in abundance is observed for $(a_5\text{-A})/w_1$ and $(a_3\text{-A})/w_3$ (Fig. 6.9b). Figure 6.9c shows that $(a_2\text{-C})/w_4$ ion pair has 858 % drop due to the ejection of $[M-C]^{2-}$. However, the signal of w_5^{2-}-ion only drops by 40 % with ejection of $[M-G]^{2-}$. This result agrees with the mechanism described by Gross and coworkers [62], who suggested that protons from an adjacent 5′-phosphate constitute the principal source for proton transfer to the nucleobases during collisional activation. Since the guanine is located on the 5′-terminus of the sequence, there is no available 5′-phosphate. Therefore, the w_5^{2-}-ion is generated, at least in part, via a different mechanism and is less affected by the DR experiment.

Gardner et al. [111] explored the capability of IRMPD for sequencing siRNA by investigating the fragmentation of both siRNA anions and cations on a linear ion trap. They observed that single-strand siRNA anions dissociate through cleavage of the 5′ P–O bonds and give rise to c/y-type fragments. The base loss channel can be significantly reduced by increasing irradiation time to induce secondary fragmentation. With longer irradiation times, internal fragment ions corresponding to consecutive cleavage 5′ P–O bonds also become abundant as well as (a-B)/w-ions. IRMPD of siRNA cations mainly yields complementary c/y-fragments, whereas other fragmentation channels are minimal. Comparing to ion trap CID, the primary difference of IRMPD is the total number of product ions observed. With longer irradiation times, larger product ions further dissociate into small fragment ions; however, full sequence coverage can still be achieved. Short
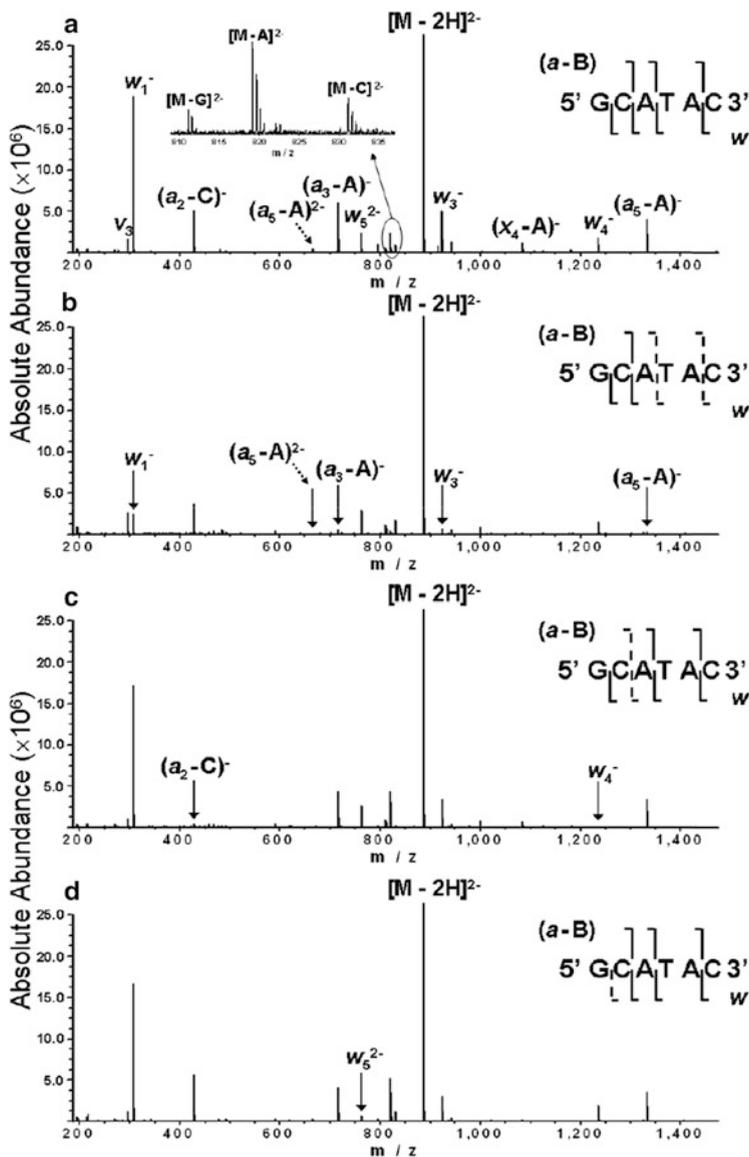
Fig. 6.9 IRMPD spectra of d(GCATAC) (**a**) without DR and (**b**) with DR ejection at the ICR frequencies of $[M–A]^{2-}$, (**c**) $[M–C]^{2-}$, and (**d**) $[M–G]^{2-}$ (*solid arrows* indicate product ions that were significantly decreased in abundance). Reprinted with permission from [109]. Copyright (2009) IM Publications LLP

irradiation times typically give rise to nearly complete series of c/y-type fragments. This suggests that IRMPD analysis of siRNA cations may be a potential tool for de novo sequencing because of its single fragmentation pattern. Additionally, the

charged nucleobase fragments, $[G+H]^+$, $[A+H]^+$, and $[C+H]^+$, which cannot be detected by conventional CID in ion traps due to the alleviation of low-mass cutoff, are observed in the IRMPD-produced spectra.

Thymine-containing oligonucleotides are notoriously troublesome in sequencing due to the high gas-phase stability of the $3'$-side backbone of the thymidine [14]. CID of DNAs containing thymine and non-thymine bases results primarily in cleavage on the $3'$-side of the non-thymine nucleobases. Parr and Brodbelt [110] have demonstrated that IRMPD leads to cleavages between all the nucleobases and provides complete sequence information for thymine-containing DNAs (Fig. 6.10), generating (a-B)/w-sequence ions. In addition, energy-variable CID and time-variable IRMPD data were collected for one mixed-base oligonucleotide and one thymine-rich sequence to compare the trends in the product ions as a function of energy deposition. They observed that the distribution of the fragment ions remained unchanged with increase in collision energy, whereas fragmentation pattern of IRMPD changed as a function of irradiation time. With low-irradiation times, the same products are observed as CID. As the irradiation time increases, more (a-B)/w-ions from other backbone cleavages are observed. At very high-irradiation times, c-ions, internal fragment ions, and individual base ions are observed, presumably due to the further decomposition of (a-B)- and w-ions. The authors suggested that IRMPD overcomes the limitation of CID via extensive secondary dissociation of the primary (a-B)- and w-ions to generate a more informative fragmentation pattern.

6.3.2.2 UV Photodissociation

The limited number of studies that describe the irradiation of nucleic acid ions with UV light (see below) has been restricted to multiply charged anions, and in these cases, photodetachment of an electron is the major process. For this reason, UVPD is included with the ion/electron(ion) techniques discussed below because those techniques also generate radical ions.

6.3.3 Ion/Electron(Ion) Interactions Techniques

Based on experience with peptides and protein ions, the fragmentation pathways accessed by radical ions can allow for extensive [113], de novo sequencing [114] and characterization of posttranslational modifications [115]. Therefore, fragmentation of oligonucleotide radical ions is of potential interest for the generation of sequence information complementary to that from conventional CID. ECD, for example, has been shown to fragment oligonucleotide ions in a different pattern, giving rise to d radical cations, a/z-, c/x-, and (c/x-base)-ions (a- and z-, c- and x-ions have the same mass in palindromic sequences) [116, 117]. However, generation of protonated oligonucleotides is not always as efficient as the generation of

Fig. 6.10 Comparison of product ions observed with CID and IRMPD for seven different oligodeoxyribonucleotides. Reprinted with permission from [110]. Copyright 2003 John Wiley & Sons, Ltd
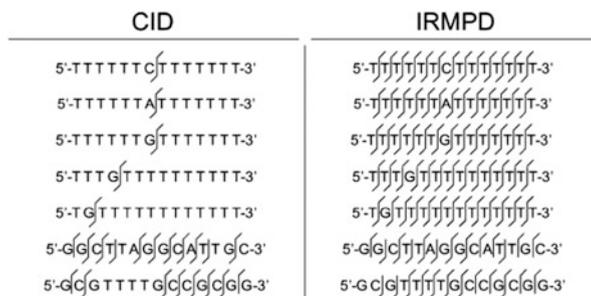
the corresponding anions via ESI. Therefore, several techniques are used to generate radical anions, including electron detachment dissociation (EDD) [118], negative electron transfer dissociation (NETD) [119], negative ion ECD (niECD) [120], as well as UVPD and EPD. Electrons can be removed from a multiply charged anion via electron detachment, electron transfer to a cation, or photodetachment. From the limited data published to date, it is not possible to draw firm conclusions regarding any differences in initial sites of radical formation or how such differences might lead to differences in structural information generated by a subsequent activation step. In any case, the radical anions formed via these different electron removal methods appear to fragment similarly on a qualitative basis when subjected to CID. For this reason, we discuss photodetachment as an illustrative case.

UVPD at 193 nm [91, 121] and electron photodetachment dissociation [122, 123] at 260 nm have been applied to multiply charged oligonucleotide anions. McLafferty and coworkers [124] demonstrated for the first time that 193 nm photon irradiation can fragment multiply charged dT_{30} via electron photodetachment and formation of d/w-and a/z-ions (which were initially interpreted as w- and a-ions since they have the same masses as d-and z-ions in the sequence dT_{30}). Gabelica et al. [122, 123] examined the electron photodetachment of single-stranded oligonucleotide anions and duplexes at 250–285 nm. They found that laser irradiation at 260 nm results in minimal fragmentation of the DNA anion, generating predominant charge-reduced radical anions. CID of the charge-reduced radical anions arising from the irradiation produced complementary a/w- and d/z-fragment ions. Smith and Brodbelt [121] explored UVPD of a series of modified oligonucleotides (6- to 20-mer) at 193 nm and observed all types of backbone fragments and internal fragments. EPD also resulted in abundant sequence ions in terms of a/(a-B)/w- and d/z-products. Generally, UVPD and EPD are almost identical in fragmentation pathways. The only difference between the two is that EPD uses lower irradiation photon energy and requires further collision activation to fragment the radical species generated via the electron photodetachment process.
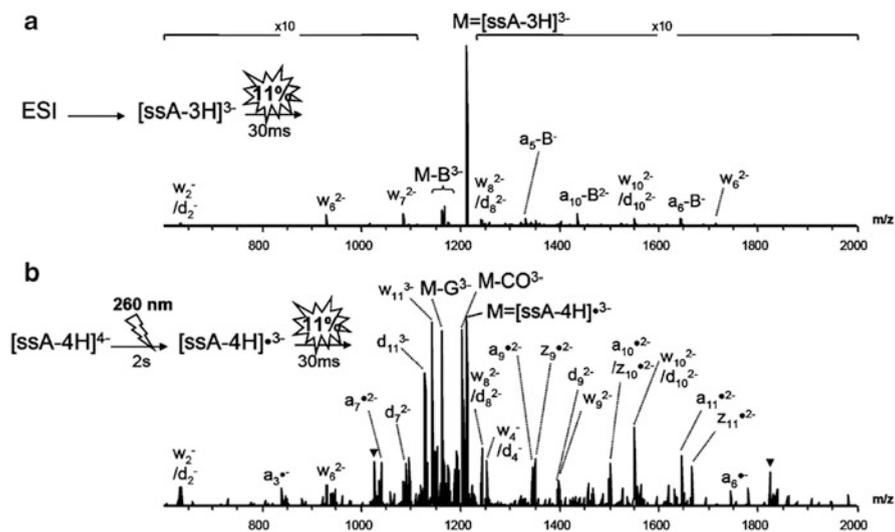
Fig. 6.11 Comparison of CID and EPD of $[ssE]^{4-}$. (**a**) CID during 30 ms at 12 % activation amplitude on $[ssE-4H]^{4-}$ produced by electrospray. (**b**) EPD: CID during 30 ms at 12 % activation amplitude on $[ssE-5H]^{4-\cdot}$ produced by electron photodetachment of $[ssE-5H]^{5-}$ under 2 s irradiation at 260 nm. The parent ion is noted M. Reprinted with permission from [122]. Copyright 2006 American Chemical Society

To understand the photodetachment mechanism, Gabelica and coworkers [122] have investigated the electron detachment efficiency as a function of irradiation wavelength and found that maximum effect is observed between 260 nm and 275 nm. The 260 nm irradiation mainly gives rise to electron detachment with very low abundance of backbone fragments. They also suggested that electron detachment efficiency of DNA duplexes is correlated with their G/C composition, with higher electron detachment efficiency of G/C-rich duplexes. This phenomenon is attributed to the low ionization potential (IP) of guanine [125]. With further CID on the charge-reduced radical anions, nearly complete a/w- and d/z-ion series were observed (Fig. 6.11). In contrast to fragment ions produced by CID of normal DNA anions, a- and z-ions are suggested to be radicals, which is consistent with the fragment ions observed in the EDD experiment [118]. A mechanism was proposed for the fragmentation pathway observed in EPD (Scheme 6.15). The guanine base absorbs the UV photon energy due to its low oxidation potential and leads to electron detachment from the base. Then the radical delocalizes to the phosphodiester backbone and facilitates the cleavage of the 3′ C–O and 5′ C–O bonds.
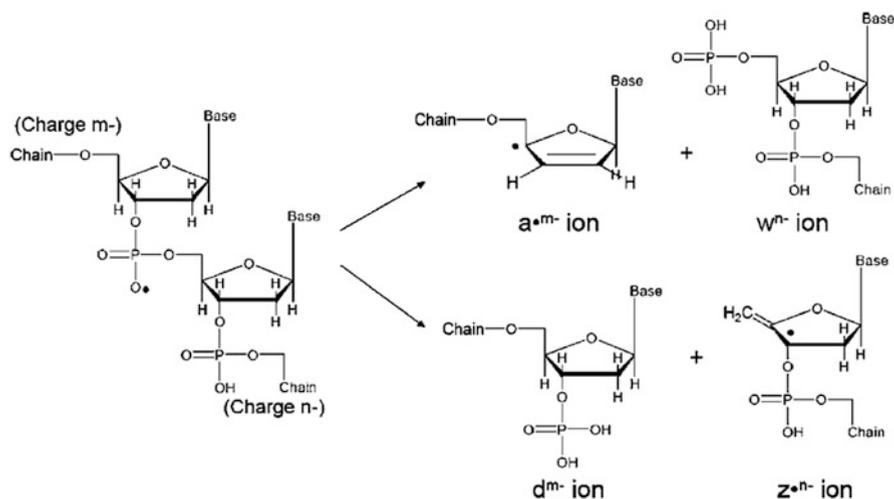
Scheme 6.15 Proposed mechanism for the EPD fragmentation pathways. Reprinted with permission from [122]. Copyright 2006 American Chemical Society

Smith and Brodbelt [121] explored the fragmentation patterns of deprotonated unmodified and modified oligodeoxyribonucleotides (6- to 20-mer) using UVPD and EPD with a higher energy photon (193 nm). Electron detachment efficiency for this wavelength was also shown to be nucleobase related, with the trend to be $dA_6 > dG_6 > dC_6 > dT_6$. This trend is different from that observed by EPD at 260 nm ($dG_6 > dA_6 > dC_6 > dT_6$). The authors suggested that the difference is not only caused by ionization potentials of the individual bases but also by their photoabsorption efficiencies (molar extinction coefficients). Moreover, UVPD at 193 nm produces efficient charge-reduced species as well as an impressive variety of backbone ions. EPD was also performed on oligonucleotides with modifications, including 5-methylated cytosine and phosphorothioate. For the oligonucleotide containing 5-methylated cytosine, EPD and CID provide similar levels of structural information, both being able to pinpoint the location of modified cytosine. For the phosphorothioate-modified oligodeoxyribonucleotide, EPD provides better sequence information than CID. Figure 6.12 shows the comparison between CID and EPD of the phosphorothioate containing oligomer PSss12 (5′-TAGCTAGTCsGAC-3′). For CID, the w_4- and a_{10}-AH ions are modified, indicating that the modification locates within the region where the two fragment ions overlap, i.e., the cytidine nucleoside, the two adjacent phosphate groups, and the ribose at the tenth residue. For EPD, the presence of modified z_4- and d_9-ions suggests a more precise location for the phosphorothioate: cytidine nucleoside and the 3′ adjacent phosphate group. Due to the absence of the modified a_9-ion, the phosphorothioate is pinpointed on the phosphate backbone between C_9 and G_{10}.
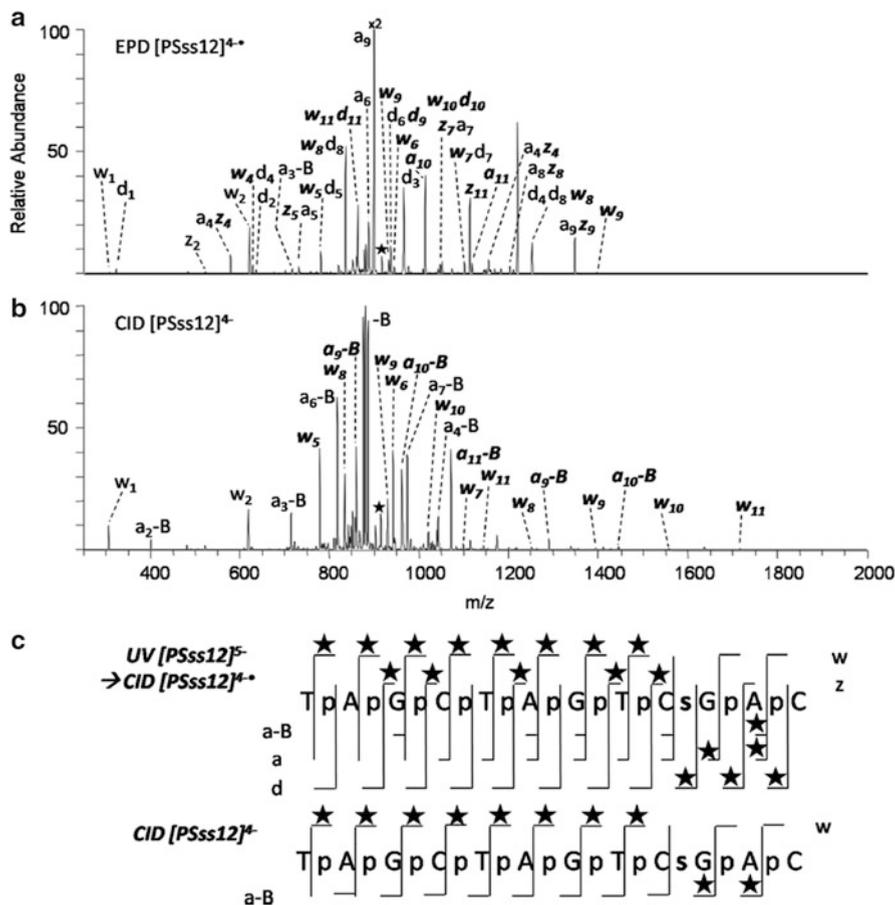
Fig. 6.12 MS/MS spectra of PSss12 by (**a**) EPD of $4^{-\bullet}$ (one UV pulse, 5 mJ; then 8 % normalized collision energy, 30 ms) and (**b**) CID 4^- (8 % normalized collision energy, 30 ms). Precursor ions are noted with a star. *Bolded* and *italicized* product ions retain the modification. Not all of the EPD product ions are labeled due to the complexity of the spectrum. A summary of the product ions formed from EPD and CID for the phosphorothioate single-strand 12-mer (**c**), in which the *slash marks* represent cleavages at that location. A *star* above the *slash mark* represents a product containing the phosphorothioate modification. The phosphorothioate modification is shown as an s in the sequence. Reprinted with from [121]. Copyright 2010 American Chemical Society

6.4 Conclusion and Outlook

With the development of mass spectrometry instrumentation and ionization methods, the fragmentation of gaseous nucleic acid ions of both polarities has been investigated with a variety of energy deposition methods and over a wide range of reaction time frames. Observations made by many laboratories over

several decades have been broadly self-consistent. However, some inconsistencies in the literature have been noted, which may arise from differences in instrument platform, different discrimination effects, and variations in conditions. For instance, different base loss trends (orders) have been reported for what nominally appear to be similar conditions. Moreover, distinct correlations between electron photodetachment efficiency and nucleobase identity have been reported.

As a tool for the sequencing of oligonucleotides and characterization of modifications, tandem MS has become particularly powerful for small oligomers. Its utility is derived from the ability to generate a variety of ion types, access a wide variety of dissociation conditions, and yield high mass accuracy product analysis [126, 127]. Mass spectrometry is not competitive with conventional sequencing techniques for high-throughput sequencing applications. However, it is uniquely suited to the characterization of modified oligomers [128]. These include modified bases, for which there are many in RNA, and modified backbones, which are being explored for therapeutic applications, for example. In terms of the length of oligomers that can be sequenced (i.e., the "read length"), tandem MS is largely limited to oligomers with a common backbone smaller than a few tens of residues. In a few exceptional cases, high sequence coverage for species greater than about 30 nucleotides has been reported using CID [129]. This performance, however, is hardly routine. Current technology is able to support "bottom-up" scenarios that employ RNases, for example, to generate oligomers small enough for characterization by CID [130]. The relatively recent introduction of techniques to generate radical ions, however, has enabled the sequencing of mixed-backbone oligomers of the size of typical siRNAs [98] and relatively high sequence coverage for natural RNA molecules as large as tRNA [131, 132]. These constitute "top-down" applications. Several factors contribute to limiting the size of oligonucleotides amenable to top-down structural characterization. Among these is the fragmentation chemistry. The introduction of hybrid techniques (e.g., EPD followed by CID) that combine the generation of a new ion type with another structural probe has opened up new possibilities for increasing sequence coverage in oligonucleotide analysis. Such hybrid approaches appear to be promising for improving the state of the art in the primary structural characterization of oligonucleotides by tandem mass spectrometry and will likely see further exploration/development in the coming years.

6.5   Summary of Key Concepts

- **DNA**: Even-electron DNA ion fragmentation is dominated by base loss followed by cleavage of the 3′ C–O bond of the ribose from which the base was lost to yield complementary (a-B)- and w-ions. Which bases are preferentially lost depends upon base identity, ion polarity, and charge state. DNA ions are also highly dependent upon ion activation conditions. Overall, the loss of adenine tends to be favored, whereas loss of thymine tends to be inhibited. Complete sequencing of small oligomers can be challenging for a given charge state,

polarity, and set of conditions due to the sensitivity of the mechanism to base identity and the general tendency for DNA ions to show extensive sequential fragmentation reactions at high enough energies to allow for a wider array of base loss channels to compete.

- **RNA**: A channel that leads to complementary c/y-ions becomes competitive with the presence of the 2′-OH group of RNA. At low activation energies, long times, and relatively low precursor ion charge it is possible to favor c/y-cleavage, whereas at higher energies and higher charge, base loss followed by (a-B)/w-ion formation becomes competitive. The c/y-cleavage is relatively insensitive to base identity, which facilitates the sequencing of oligomers with several tens of residues.

- **Modified backbone**: Sugars without a labile hydrogen at the 2′-position show little preference for either base loss or c/y-ion formation. Rather, cleavages all along the phosphodiester linkage tend to be observed when the modification is present for each residue.

- **Mixed modifications**: Oligomers with mixed modifications along the backbone can be challenging to fully sequence, particularly if DNA residues are present. Cleavages at residues with sugars without a labile hydrogen on the 2′-carbon tend to be less competitive than those from DNA residues. This complicates sequencing via common CID approaches. However, the 3′-phosphodiester backbones of 2′-modified residues and RNA residues have comparable level of dissociation threshold.

- **Radical ions**: A variety of techniques can remove an electron from oligonucleotide anions, including electron detachment with electron irradiation, UV photon irradiation, or ion/ion electron transfer. This introduces a radical site into the pseudo-molecular ion and gives rise to fragmentation reactions that can give information complementary to that derived from fragmentation of even-electron ions upon CID. Too few observations are currently available to draw general conclusions regarding the influence of the form of electron removal on subsequent fragmentation or on the generality of mechanisms currently proposed for radical anion cleavage. However, it has already been demonstrated that the ability to generate radical anions is desirable when conventional dissociation methods applied to even-electron ions are inadequate. Examples have already been provided for the utility of radical ion dissociation in dealing with oligomers with mixed-backbone modifications and for moderate size RNA ions.

References

1. Fenn J, Mann M, Meng C, Wong S, Whitehouse C (1989) Electrospray ionization for mass-spectrometry of large biomolecules. Science 246:64–71
2. Hillenkamp F, Karas M (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal Chem 63:A1193–A1202
3. Tabb D, Smith L, Breci L, Wysocki V, Lin D, Yates J (2003) Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides. Anal Chem 75:1155–1163

4. Breci L, Tabb D, Yates J, Wysocki V (2003) Cleavage N-terminal to proline: analysis of a database of peptide tandem mass spectra. Anal Chem 75:1963–1972

5. Huang Y, Triscari JM, Pasa-Tolic L, Anderson GA, Lipton MS, Smith RD, Wysocki VH (2004) Dissociation behavior of doubly-charged tryptic peptides: correlation of gas-phase cleavage abundance with Ramachandran plots. J Am Chem Soc 126:3034–3035

6. McLafferty FW (2001) Tandem mass spectrometric analysis of complex biological mixtures. Int J Mass Spectrom 212:81–87

7. Stephenson JL, McLuckey SA, Reid GE, Wells JM, Bundy JL (2002) Ion/ion chemistry as a top-down approach for protein analysis. Curr Opin Biotechnol 13:57–64

8. Reid GE, McLuckey SA (2002) 'Top down' protein characterization via tandem mass spectrometry. J Mass Spectrom 37:663–675

9. Harvey DJ (1996) Matrix-assisted laser desorption ionization mass spectrometry of oligosaccharides and glycoconjugates. J Chromatogr A 720:429–446

10. Zaia J (2004) Mass spectrometry of oligosaccharides. Mass Spectrom Rev 23:161–227

11. Nordhoff E, Kirpekar F, Roepstorff P (1996) Mass spectrometry of nucleic acids. Mass Spectrom Rev 15:67–138

12. Murray KK (1996) DNA sequencing by mass spectrometry. J Mass Spectrom 31:1203–1215

13. Muddiman DC, Smith RD (1998) Sequencing and characterization of larger oligonucleotides by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry. Rev Anal Chem 12:1–68

14. Wu J, McLuckey SA (2004) Gas-phase fragmentation of oligonucleotide ions. Int J Mass Spectrom 237:197–241

15. Grotjahn L, Frank R, Blocker H (1982) Ultrafast sequencing of oligodeoxyribonucleotides by FAB-MASS spectrometry. Nucleic Acids Res 10:4671–4678

16. Grotjahn L, Blocker H, Frank R (1985) Mass spectroscopic sequence-analysis of oligonucleotides. Biomed Mass Spectrom 12:514–524

17. Cerny RL, Tomer KB, Gross ML, Grotjahn L (1987) Fast-atom bombardment combined with mass spectrometry for determining structures of small oligonucleotides. Anal Biochem 165:175–182

18. Cerny RL, Gross ML, Grotjahn L (1986) Fast-atom bombardment combined with tandem mass spectrometry for the study of dinucleotides. Anal Biochem 156:424–435

19. Viari A, Ballini JP, Vigny P, Shire D, Dousset P (1987) CF-252-Plasma desorption mass-spectrometry of oligonucleotides. 3. Sequence-analysis of unprotected tri-deoxyribonucleoside diphosphates by CF-252-plasma desorption mass-spectrometry. Biomed Environ Mass Spectrom 14:83–90

20. Nordhoff E, Karas M, Cramer R, Hahner S, Hillenkamp F, Kirpekar F, Lezius A, Muth J, Meier C, Engels JW (1995) Direct mass spectrometric sequencing of low-picomole amounts of oligodeoxynucleotides with up to 21 bases by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom 30:99–112

21. McLuckey SA, Van Berkel GJ, Glish GL (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. J Am Soc Mass Spectrom 3:60–70

22. McLuckey SA, Mentinova M (2011) Ion/neutral, ion/electron, ion/photon, and ion/ion interactions in tandem mass spectrometry: do we need them all? Are they enough? J Am Soc Mass Spectrom 22:3–12

23. McLuckey SA (1992) Principles of collisional activation in analytical mass spectrometry. J Am Soc Mass Spectrom 3:599–614

24. Wells JM, McLuckey SA (2005) Collision-induced dissociation (cid) of peptides and proteins. Methods Enzymol 402:148–185

25. Limbach PA (1996) Indirect mass spectrometric methods for characterizing and sequencing oligonucleotides. Mass Spectrom Rev 15:297–336

26. Burlingame AL, Boyd RK, Gaskell SJ (1998) Mass spectrometry. Anal Chem 70:647R–716R

27. McLuckey SA, Stephenson JL, O'Hair RAJ (1997) Decompositions of odd- and even-electron anions derived from deoxy-polyadenylates. J Am Soc Mass Spectrom 8:148–154

28. McLuckey SA, Habibi-Goudarzi S (1993) Decompositions of multiply charged oligonucleotide anions. J Am Chem Soc 115:12085–12095
29. McLuckey SA, Habibi-Goudarzi S (1994) Ion trap tandem mass spectrometry applied to small multiply charged oligonucleotides with a modified base. J Am Soc Mass Spectrom 5:740–747
30. Habibi-Goudarzi S, McLuckey SA (1995) Ion trap collisional activation of the deprotonated deoxymononucleoside and deoxydinucleoside monophosphates. J Am Soc Mass Spectrom 6:102–113
31. McLuckey SA, Vaidyanathan G, Habibi-Goudarzi S (1995) Charged vs. neutral nucleobase loss from multiply charged oligonucleotide anions. J Mass Spectrom 30:1222–1229
32. McLuckey SA, Vaidyanathan G (1997) Charge state effects in the decompositions of single-nucleobase oligonucleotide polyanions. J Mass Spectrom Ion Process 162:1–16
33. Barry JP, Vouros P, Van Schepdael A, Law SJ (1995) Mass and sequence verification of modified oligonucleotides using electrospray tandem mass spectrometry. J Mass Spectrom 30:993–1006
34. Wolter MA, Engels JW (1995) Nanoelectrospray mass spectrometry mass spectrometry for the analysis of modified oligoribonucleotides. Eur Mass Spectrom 1:583–590
35. Gentil E, Banoub J (1996) Characterization and differentiation of isomeric self-complementary DNA oligomers by electrospray tandem mass spectrometry. J Mass Spectrom 31:83–94
36. Boschenok J, Sheil MM (1996) Electrospray tandem mass spectrometry of nucleotides. Rapid Commun Mass Spectrom 10:144–149
37. Ho YH, Kebarle P (1997) Studies of the dissociation mechanisms of deprotonated mononucleotides by energy resolved collision-induced dissociation. Int J Mass Spectrom 165:433–455
38. Little DP, Chorush RA, Speir JP, Senko MW, Kelleher NL, McLafferty FW (1994) Rapid sequencing of oligonucleotides by high-resolution mass spectrometry. J Am Chem Soc 116:4893–4897
39. Little DP, McLafferty FW (1995) Sequencing 50-mer DNAs using electrospray tandem mass spectrometry and complementary fragmentation methods. J Am Chem Soc 117:6783–6784
40. Little DP, Aaserud DJ, Valaskovic GA, McLafferty FW (1996) Sequence information from 42-108-mer DNAs (complete for a 50-mer) by tandem mass spectrometry. J Am Chem Soc 118:9352–9359
41. Rodgers MT, Campbell S, Marzluff EM, Beauchamp JL (1994) Low-energy collision-induced dissociation of deprotonated dinucleotides: determination of the energetically favored dissociation pathways and the relative acidities of the nucleic acid bases. Int J Mass Spectrom Ion Process 137:121–149
42. Vrkic AK, O'Hair RAJ, Foote S (2000) Fragmentation reactions of all 64 deprotonated trinucleotide and 16 mixed base tetranucleotide anions via tandem mass spectrometry in an ion trap. Aust J Chem 53:307–319
43. Sidona G, Uccella N, Weclawek K (1982) Structure determination of isomeric oligodeoxynucleotide salts by fast-atom bombardment mass-spectrometry. J Chem Res Synop 1982:184–185
44. Crain PF, Gregson JM, McCloskey JA, Nelson CC, Peltier JM, Phillips DR, Pomerantz SC (1996) Reddy, D.M. In: Burlingame AL, Carr SA (eds) Mass spectrometry in the biological sciences. Humana Press, Totowa, NJ, p 497
45. Bartlett MG, McCloskey JA, Manalili S, Griffey RH (1996) The effect of backbone charge on the collision-induced dissociation of oligonucleotides. J Mass Spectrom 31:1277–1283
46. Doktycz MJ, Habibi-Goudarzi S, McLuckey SA (1994) Accumulation and storage of ionized duplex DNA molecules in a quadrupole ion trap. Anal Chem 66:3416–3422
47. Aaserud DJ, Kelleher NL, Little DP, McLafferty FW (1996) Accurate base composition of double-strand DNA by mass spectrometry. J Am Soc Mass Spectrom 7:1266–1269

48. McLafferty FW, Aaserud DJ, Guan ZQ, Little DP, Kelleher NL (1997) Double stranded DNA sequencing by tandem mass spectrometry. Int J Mass Spectrom 165:457–466

49. Schnier PD, Klassen JS, Strittmatter EE, Williams ER (1998) Activation energies for dissociation of double strand oligonucleotide anions: evidence for Watson-Crick base pairing in vacuo. J Am Chem Soc 120:9605–9613

50. Gabelica V, De Pauw E (2001) Comparison between solution-phase stability and gas-phase kinetic stability of oligodeoxynucleotide duplexes. J Mass Spectrom 36:397–402

51. Gabelica V, De Pauw E (2002) Collision-induced dissociation of 16-mer DNA duplexes with various sequences: evidence for conservation of the double helix conformation in the gas phase. Int J Mass Spectrom 219:151–159

52. Gabelica V, De Pauw E (2002) Comparison of the collision-induced dissociation of duplex DNA at different collision regimes: evidence for a multistep dissociation mechanism. J Am Soc Mass Spectrom 13:91–98

53. Ganem B, Li YT, Henion JD (1993) Detection of oligonucleotide duplex forms by ion-spray mass-spectrometry. Tetrahedron Lett 34:1445–1448

54. Bayer E, Bauer T, Schmeer K, Bleicher K, Maier M, Gaus HJ (1994) Analysis of double-stranded oligonucleotides by electrospray mass-spectrometry. Anal Chem 66:3858–3863

55. Ding JM, Anderegg RJ (1995) Specific and nonspecific dimer formation in the electrospray-ionization mass-spectrometry of oligonucleotides. J Am Soc Mass Spectrom 6:159–164

56. Greig MJ, Gaus HJ, Griffey RH (1996) Negative ionization micro electrospray mass spectrometry of oligonucleotides and their complexes. Rapid Commun Mass Spectrom 10:47–50

57. Madsen JA, Brodbelt JS (2010) Asymmetric charge partitioning upon dissociation of DNA duplexes. J Am Soc Mass Spectrom 21:1144–1150

58. Wan KX, Gross ML, Shibue T (2000) Gas-phase stability of double-stranded oligodeoxynucleotides and their noncovalent complexes with DNA-binding drugs as revealed by collisional activation in an ion trap. J Am Soc Mass Spectrom 11:450–457

59. Phillips DR, McCloskey JA (1993) A comprehensive study of the low-energy collision-induced dissociation of dinucleoside monophosphates. Int J Mass Spectrom Ion Process 128:61–82

60. Wang Z, Wan KX, Ramanathan R, Taylor JS, Gross ML (1998) Structure and fragmentation mechanisms of isomeric T-rich oligodeoxynucleotides: a comparison of four tandem mass spectrometric methods. J Am Soc Mass Spectrom 9:683–691

61. Wan KX, Gross J, Hillenkamp F, Gross ML (2001) Fragmentation mechanisms of oligodeoxynucleotides studied by H/D exchange and electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 12:193–205

62. Wan KX, Gross ML (2001) Fragmentation mechanisms of oligodeoxynucleotides: effects of replacing phosphates with methylphosphonates and thymines with other bases in T-rich sequences. J Am Soc Mass Spectrom 12:580–589

63. Vrkic AK, O'Hair RAJ, Foote S, Reid GE (2000) Fragmentation reactions of all 64 protonated trimer oligodeoxynucleotides and 16 mixed base tetramer oligodeoxynucleotides via tandem mass spectrometry in an ion trap. Int J Mass Spectrom 194:145–164

64. Wang PP, Bartlett MG, Martin LB (1997) Electrospray collision-induced dissociation mass spectra of positively charged oligonucleotides. Rapid Commun Mass Spectrom 11:846–856

65. Ni JS, Mathews MAA, McCloskey JA (1997) Collision-induced dissociation of polyprotonated oligonucleotides produced by electrospray ionization. Rapid Commun Mass Spectrom 11:535–540

66. Weimann A, Iannitti-Tito P, Sheil MM (2000) Characterisation of product ions in high-energy tandem mass spectra of protonated oligonucleotides formed by electrospray ionization. Int J Mass Spectrom 194:269–288

67. Rodgers MT, Campbell S, Marzluff EM, Beauchamp JL (1995) Site-specific protonation directs low-energy dissociation pathways of dinucleotides in the gas phase. Int J Mass Spectrom Ion Process 148:1–23

68. Dongre AR, Jones JL, Somogyi A, Wysocki VH (1996) Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: evidence for the mobile proton model. J Am Chem Soc 118:8365–8374
69. Harrison AG, Yalcin T (1997) Proton mobility in protonated amino acids and peptides. Int J Mass Spectrom Ion Process 165:339–347
70. Schürch S, Bernal-Mendez E, Leumann CJ (2002) Electrospray tandem mass spectrometry of mixed-sequence RNA/DNA oligonucleotides. J Am Soc Mass Spectrom 13:936–945
71. Tromp JM, Schürch S (2005) Gas-phase dissociation of oligoribonucleotides and their analogues studied by electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 16:1262–1268
72. Tang W, Zhu L, Smith LM (1997) Controlling DNA fragmentation in MALDI-MS by chemical modification. Anal Chem 69:302–312
73. Huang T-Y, Kharlamova A, Liu J, McLuckey SA (2008) Ion trap collision-induced dissociation of multiply deprotonated RNA: c/y-ions versus (a-B)/w-ions. J Am Soc Mass Spectrom 19:1832–1840
74. Huang T-Y, Liu J, Liang X, Hodges BDM, McLuckey SA (2008) Collision-induced dissociation of intact duplex and single-stranded siRNA anions. Anal Chem 80:8501–8508
75. Kirpekar F, Krogh TN (2001) RNA fragmentation studied in a matric-assisted laser desorption/ionization tandem quadrupole/orthogonal time-of-flight mass spectrometer. Rapid Commun Mass Spectrom 15:8–14
76. Andersen TE, Kirpekar F, Haselmann KF (2006) RNA fragmentation in MALDI mass spectrometry studied by H/D-exchange: mechanisms of general applicability to nucleic acids. J Am Soc Mass Spectrom 17:1353–1368
77. Gross J, Leisner A, Hillenkamp F, Hahner S, Karas M, Schafer J, Lutzenkirchen F, Hordhoff E (1998) Investigations of the metastable decay of DNA under ultraviolet matrix-assisted laser desorption/ionization conditions with post-source decay analysis and hydrogen/deuterium exchange. J Am Soc Mass Spectrom 9:866–878
78. Horton HR, Moran LA, Ochs RS, Rawn JD, Scrimgeour KG (2002) Principles of biochemistry, 3rd edn. Prentice-Hall, Upper Saddle River, NJ, pp 605–607
79. Goodchild J (2004) Oligonucleotide therapeutics: 25 years agrowing. Curr Opin Mol Ther 6:120–128
80. Opalinska JB, Gewirtz AM (2002) Nucleic acid therapeutics: basic principles and recent applications. Nat Rev Drug Discov 1:503–514
81. Raal FJ, Santos RD, Blom DJ, Marais AD, Charng MJ, Cromwell WC, Lachmann RH, Gaudet D, Tan JL, Chasan-Taber S, Tribble DL, Flaim JD, Crooke ST (2010) Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. Lancet 375:998–1006
82. Brody EN, Gold L (2000) Aptamers as therapeutic and diagnostic agents. J Biotechnol 74:5–13
83. Chiu Y-L, Rana TM (2009) siRNA function in RNAi: a chemical modification analysis. RNA 9:1034–1048
84. Butora G, Kenski DM, Cooper AJ, Fu WL, Qi N, Li JJ, Flanagan WM, Davies IW (2005) Potent and persistent in vivo anti-HBV activity of chemically modified siRNAs. Nat Biotechnol 23:1002–1007
85. Manoharan M (2004) RNA interference and chemically modified small interfering RNAs. Curr Opin Chem Biol 8:570–579
86. Mouritzen P, Nielsen AT, Pfundheller HM, Choleva Y, Kongsbak L, Moller S (2003) Single nucleotide polymorphism genotyping using locked nucleic acid (LNA). Expert Rev Mol Diagn 3:27–38
87. Crinelli R, Bianchi M, Gentilini L, Magnani M (2002) Design and characterization of decoy oligonucleotides containing locked nucleic acids. Nucleic Acids Res 30(11):2435–2443

88. Monn STM, Schürch S (2007) New aspects of the fragmentation mechanisms of unmodified and methylphosphonate-modified oligonucleotides. J Am Soc Mass Spectrom 18:984–990

89. Huang T-Y, Kharlamova A, McLuckey SA (2010) Ion trap collision-induced dissociation of locked nucleic acids. J Am Soc Mass Spectrom 21:144–153

90. Nyakas A, Stucki SR, Schürch S (2011) Tandem mass spectrometry of modified and platinated oligoribonucleotdes. J Am Soc Mass Spectrom 22:875–887

91. Smith SI, Brodbelt JS (2011) Hybrid activation methods for elucidating nucleic acid modifications. Anal Chem 83:303–310

92. Gao Y, McLuckey SA (2012) Collision-induced dissociation of oligonucleotide anions fully modified at the 2'-position of the ribose: 2'-F/-H and 2'-F/-H/-OMe mix-mers. J Mass Spectrom 47:364–369

93. Gao Y, McLuckey SA (2013) Electron transfer followed by collision-induced dissociation (NET-CID) for generating sequence information from backbone-modified oligonucleotide anions. Rapid Commun Mass Spectrom 27:249–257

94. Baker TR, Keough T, Dobson RLM, Riley TA, Hasselfield JA, Hesselberth PE (1993) Antisense DNA oligonucleotides.1. The use of ionspray tandem mass-spectrometry for the sequence verification of methylphosphonate oligodeoxyribonucleotides. Rapid Commun Mass Spectrom 7:190–194

95. Altmann K, Dean NM, Fabbro D, Freier SM, Geiger T, Häner R, Hüsken D, Martin P, Monia BP, Müller M, Natt F, Nicklin P, Phillips J, Pieles U, Sasmor H, Moser HE (1996) Second generation of antisense oligonucleotides: from nuclease resistance to biological efficacy in animals. Chimia 50:168–176

96. Eckstein F (2000) Phosphorothioate oligonucleotides: what is their origin and what is unique about them? Antisense Nucleic Acids Drug Dev 10:117–121

97. Kurreck J (2003) Antisense technologies: improvement through novel chemical modifications. Eur J Biochem 270:1628–1644

98. Gao Y, Yang J, Cancilla MT, Meng F, McLuckey SA (2013) Top-down interrogation of chemically modified oligonucleotides by negative electron transfer and collision induced dissociation. Anal Chem 85:4713–4720

99. Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW (1994) Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. Anal Chem 66:2809–2815

100. Woodin RL, Bomse DS, Beauchamp JL (1978) Multiphoton dissociation of molecules with low power continuous wave infrared laser radiation. J Am Chem Soc 100:3248–3250

101. Håkansson K, Cooper HJ, Hudgins RR, Nilsson CL (2003) High resolution tandem mass spectrometry for structural biochemistry. Curr Org Chem 7:1503–1525

102. Stephenson JL Jr, Booth MM, Shalosky JA, Eyler JR, Yost RA (1994) Infrared multiphoe-photon dissociation in the quadrupole ion trap via a multipass ootical arrangement. J Am Soc Mass Spectrom 5:886–893

103. Hofstadler SA, Sannes-Lowery KA, Griffey RH (1999) Infrared multiphoton dissociation in an external ion reservoir. Anal Chem 71:2067–2070

104. Crowe MC, Brodbelt JS (2005) Differentiation of phosphorylated and unphosphorylated peptides by high-performance liquid chromatography-electrospray ionization-infrared multi-photon dissociation in a quadrupole ion trap. Anal Chem 77:5726–5734

105. Flora JW, Muddiman DC (2001) Selective, sensitive, and rapid phosphopeptide identification in enzymatic digests using ESI-FTICR-MS with infrared multiphoton dissociation. Anal Chem 73:3305–3311

106. Hofstadler SA, Griffey RH, Pasa-Tolic L, Smith RD (1998) The use of a stable internal mass standard for accurate mass measurements of oligonucleotide fragment ions using electrospray ionization fourier transform ion cyclotron resonance mass spectrometry with infrared multi-photon dissociation. Rapid Commun Mass Spectrom 12:1400–1404

107. Sannes-Lowery KA, Hofstadler SA (2003) Sequence confirmation of modified oligonucleotides using IRMPD in the external ion reservoir of an electrospray ionization fourier transform ion cyclotron mass spectrometer. J Am Soc Mass Spectrom 14:825–833

108. Keller KM, Brodbelt JS (2004) Collisionally activated dissociation and infrared multiphoton dissociation of oligonucleotides in a quadrupole ion trap. Anal Biochem 326:200–210

109. Yang J, Håkansson K (2009) Characterization of oligodeoxynucleotide fragmentation pathways in infrared multiphoton dissociation and electron detachment dissociation by Fourier transform ion cyclotron double resonance. Eur J Mass Spectrom 15:293–304

110. Parr C, Brodbelt JS (2010) Increased sequence coverage of thymine-rich oligodeoxynucleotides by infrared multiphoton dissociation compared to collision-induced dissociation. J Mass Spectrom 45:1098–1103

111. Gardner MW, Li N, Ellington AD, Brodbelt JS (2010) Infrared multiphoton dissociaiton of small-interfering RNA anions and cations. J Am Soc Mass Spectrom 21:580–591

112. Comisarow MB, Grassi V, Parisod G (1978) Fourier transform ion cyclotron double resonance. Chem Phys Lett 57:413–416

113. Cooper HJ, Hakansson K, Marshall AG (2005) The role of electron capture dissociation in biomolecular analysis. Mass Spectrom Rev 24:201–222

114. Horn DM, Ge Y, McLafferty FW (2000) Activated ion electron capture dissociation for mass spectral sequencing of larger (42 kDa) proteins. Anal Chem 72:4778–4784

115. Kjeldsen F, Haselmann KF, Budnik BA, Sorensen ES, Zubarev RA (2003) Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. Anal Chem 75:2355–2361

116. Håkansson K, Hudgins RR, Marshall AG, O'Hair RAJ (2003) Electron capture dissociation and infrared multiphoton dissociation of oligodeoxynucleotide dications. J Am Soc Mass Spectrom 14:23–41

117. Schultz KN, Hakansson K (2004) Rapid electron capture dissociation of mass-selectively accumulated oligodeoxynucleotide dications. Int J Mass Spectrom 234:123–130

118. Yang J, Mo JJ, Adamson JT, Håkansson K (2005) Characterization of oligodeoxynucleotides by electron detachment dissociation Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 77:1876–1882

119. Huang T-Y, McLuckey SA (2011) Gas-phase ion/ion reactions of rubrene cations and multiply charged DNA and RNA anions. Int J Mass Spectrom 304:140–147

120. Yoo HJ, Wang N, Zhuang SY, Song HT, Håkansson K (2011) Negative-ion electron capture dissociation: radical-driven fragmentation of charge-increased gaseous peptide anions. J Am Chem Soc 133:16790–16793

121. Smith SI, Brodbelt JS (2010) Characterization of oligodeoxynucleotides and modifications by 193 nm photodissociation and electron photodissociation dissociation. Anal Chem 82:7218–7226

122. Gabelica V, TabarIn T, Antolne R, Rosu F, Compagnon I, Broyer M, De Pauw E, Dugourd P (2006) Electron photodetachment dissociation of DNA polyanions in a quadrupole ion trap mass spectrometer. Anal Chem 78:6564–6572

123. Gabelica V, Rosu F, Tabarin T, Kinet C, Antoine R, Broyer M, De Pauw E, Dugourd P (2007) Base-dependent electron photodetachment from negatively charged DNA strands upon 260-nm laser irradiation. J Am Chem Soc 129:4706–4713

124. Guan Z, Kelleher NL, O'Connor PB, Aaserud DJ, Little DP, McLafferty FW (1996) 193 nm Photodissociation of larger multiply charge biomolecules. Int J Mass Spectrom Ion Processes 157/158:357–364

125. Wetmore SD, Boyd RJ, Eriksson LA (2000) Electron affinities and ionization potentials of nucleotide bases. Chem Phys Lett 322:129–135

126. Oberacher H, Pitterl F (2009) On the use of ESI-QqTOF-MS/MS for the comparative sequencing of nucleic acids. Biopolymers 91:401–409

127. Flosadóttir HD, Gíslason K, Sigurdsson ST, Ingólfsson O (2012) Mass spectro-metric study on sodium ion induced central nucleotide deletion in the gas phase. J Am Soc Mass Spectrom 23:690–698
128. Giessing AMB, Kirpekar F (2012) Mass spectrometry in the biology of RNA and its modifications. J Proteomics 75:3434–3449
129. Taucher M, Rieder U, Breuker K (2010) Minimizing base loss and internal fragmentation in collisionally activated dissociation of multiply deprotonated RNA. J Am Soc Mass Spectrom 21:278–285
130. Meng Z, Limbach PA (2005) Shotgun sequencing small oligonucleotides by nozzle-skimmer dissociation and electrospray ionization mass spectrometry. Eur J Mass Spectrom 11:221–229
131. Taucher M, Breuker K (2010) Top-down mass spectrometry for sequencing of larger (up to 61 nt) RNA by CAD and EDD. J Am Soc Mass Spectrom 21:918–929
132. Taucher M, Breuker K (2012) Characterization of modified RNA by top-down mass spectrometry. Angew Chem 124:11451–11454

Part II

Applications

Characterization of Ribonucleic Acids and Their Modifications by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

7

Kathrin Breuker

Abstract

Recent advances in electrospray ionization and the extension of radical ion based dissociation techniques to ribonucleic acids (RNA) were key factors for developing top-down mass spectrometry as a powerful method for the detailed characterization of posttranscriptional and synthetic modifications of RNA. This new approach identifies and localizes all mass-altering modifications without the need for labeling reactions, and can be used for characterization of RNA of unknown sequence.

Keywords

RNA • Posttranscriptional modifications • Synthetic modifications • Collisionally activated dissociation • Electron detachment dissociation • Top-down mass spectrometry • Fourier transform ion cyclotron resonance • Mass spectrometry

Abbreviations/Formulae

BIRD	Blackbody infrared multiphoton dissociation
CAD	Collisionally activated dissociation (synonymous with CID)
CID	Collision-induced dissociation (synonymous with CAD)
cDNA	Complementary desoxyribonucleic acid
DNA	Desoxyribonucleic acid
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EPD	Electron photodetachment dissociation
ESI	Electrospray ionization

K. Breuker (✉)

Institute of Organic Chemistry and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, CCB, Innrain 80/82, 6020 Innsbruck, Austria
e-mail: kathrin.breuker@uibk.ac.at

FTICR   Fourier transform ion cyclotron resonance
IRMPD   Infrared multiphoton dissociation
MALDI   Matrix-assisted laser desorption/ionization
MS   Mass spectrometry
nETD   Negative electron transfer dissociation
nt   Nucleotide(s)
PTM   Posttranscriptional modification
RNA   Ribonucleic acid
siRNA   Small interfering RNA
tRNA   Transfer RNA

7.1   Introduction

In living organisms, the genetic information encoded in desoxyribonucleic acids (DNA) is transcribed only in part into ribonucleic acids (RNA) by RNA polymerases, and only a fraction of the RNA is translated into proteins at the ribosome. This protein fraction is stunningly small in homo sapiens, <2 % of the entire genome, even though more than 50 % of the human DNA is transcribed, or predicted to be transcribed, into RNA [1]. While these percentages differ for different organisms [1], the considerable scale of transcription into non-protein coding RNA raises the general question if it serves any purpose. This important research subject can be addressed by studying the DNA ("genome"), RNA ("transcriptome"), and proteins ("proteome") found in a given organism. DNA sequencing generally involves amplification of the DNA under study by polymerase chain reaction or molecular cloning [2], which is also used in RNA sequencing after biosynthesis of its complementary DNA (cDNA) by reverse transcriptase. These methods have in common that possible chemical alterations of the DNA or RNA under study are lost during reverse transcriptase and amplification reactions [3]. However, in recent years, it was shown that noncoding RNA, which can play an important role in gene expression, frequently carries posttranscriptional modifications such as base methylations [4] that can be critical to function. Moreover, synthetic RNA for pharmaceutical applications is usually highly modified for reasons of stability.

Helm and coworkers have discussed different approaches for the detection of RNA modifications that are based on either physicochemical properties, differential enzymatic turnover, or differential chemical reactivity [3]. They concluded that methods for identification based on physicochemical properties, among them mass spectrometry, offer the most direct and unambiguous evidence for modifications, but frequently at the expense of sequence information. Thus, a major challenge in characterizing native or synthetic RNA is the development of new methodology that can identify, localize, and possibly quantify (bio)chemical modifications of RNA. This contribution is not meant as a comprehensive review of the field of MS of RNA, but instead highlights recent advances in the characterization of modified

ribonucleic acids by top-down mass spectrometry, including important experimental details.

The truly groundbreaking methods electrospray ionization (ESI) [5] and matrix-assisted laser desorption/ionization (MALDI) [6] in the late 1980s made possible the generation of intact gaseous ions of larger biomolecules for examination by mass spectrometry (MS). This initiated the era of mass spectrometry based "proteomics" research, which, judging from the number of publications in the field, continues to grow at impressive speed (Fig. 7.1). Although the growth in the number of publications found in a search on mass spectrometry and RNA is similarly steep, the corresponding number of publications is more than an order of magnitude smaller. Moreover, not all publications found in the latter search actually deal with MS of RNA. This data shows that, at present, mass spectrometry plays a far more important role in protein analysis than in RNA characterization. However, because high-throughput RNA sequencing methods fail to characterize posttranscriptional or synthetic modifications, mass spectrometry is likely to gain far more importance in the analysis of modified RNA in the near future. For this, the top-down approach holds great promise, although complementary data can be provided by bottom-up mass spectrometry [7–12].

This chapter is organized according to the order in which the different techniques involved in top-down MS of RNA play a role in a typical experiment.

First, mass spectrometry of RNA requires gaseous ions of RNA, for which both ESI and MALDI are well suited. However, ESI generally produces more highly charged ions than MALDI and is thus compatible with the majority of mass spectrometers whose mass-to-charge (m/z) range is limited, including Fourier transform ion cyclotron resonance (FTICR) instruments, so only ESI of RNA is considered here. Second, some unique features of Fourier transform ion cyclotron resonance instrumentation relevant to top-down MS of RNA are discussed. Third, the top-down approach is described, and its advantages over bottom-up MS highlighted. Fourth and fifth, collisionally activated dissociation (CAD) and electron detachment dissociation (EDD) of RNA are reviewed in the context of top-down MS of RNA. Finally and sixth, data analysis and fragment ion assignments in top-down MS of RNA are discussed.

7.2 RNA Characterization: Methods and Discussion

7.2.1 Electrospray Ionization of RNA

Depending on the polarity of the electrostatic potential applied to the ESI emitter, electrospray ionization of aqueous solutions generally yields $(M + nH)^{n+}$ or $(M - nH)^{n-}$ ions of the analyte under study. In ESI of RNA, however, alkali ion adduction can give rise to complex mixtures of ion species in which the $(M + nH)^{n+}$ or $(M - nH)^{n-}$ ions can be of low abundance or even completely absent [13]. ESI spectra of such mixtures are characterized by low signal-to-noise (S/N) ratio and difficult to interpret. Even after extensive removal of alkali ions from the

Fig. 7.1 Number of publications found in a literature search ("Web of Science," Thomson Reuters) on "mass spectrometry" and "peptide" or "protein" (*squares*, *left axis*), and "mass spectrometry" and "RNA" or "ribonucleic acid" (*triangles*, *right axis*) versus publication year
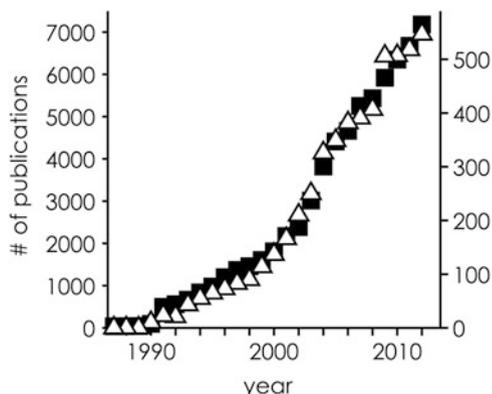
electrospray solution ("desalting"), alkali ion adduction can be severe unless suitable solution additives are used. Such additives include ammonium salts [14, 15], carboxylic acids [16], and organic bases [13, 17], all of which also affect the net charge of the $(M+nH)^{n+}$ or $(M-nH)^{n-}$ ions from ESI. As will be shown in the sections on RNA dissociation by CAD and EDD, it is very important that ion net charge can be manipulated efficiently as it is critical to ion dissociation.

In our laboratory, we use the same procedure for desalting of solutions of proteins and RNA. Briefly, 500 µl of the aqueous protein or RNA solution (1–100 µM) is concentrated to 100 µl using centrifugal concentrators (Vivaspin 500, Sartorius, polyethersulfone membrane), and 400 µl of 100 mM ammonium acetate in H_2O (18 MΩ cm) added. The process is repeated ~5 times, followed by at least five cycles of concentration and dilution with H_2O (18 MΩ cm). For transfer RNA (tRNA), which has a highly stable solution fold, we found that efficient desalting required a slightly modified protocol using instead denaturing solutions, 50 mM ammonium citrate in 1:1 H_2O/CH_3OH and 1:1 H_2O/CH_3OH, in the first and last steps of desalting, respectively. Furthermore, it is important to use high-purity solvents and ammonium salts that contain only minimal amounts of alkali ions. For ESI of RNA, we found that 1–2 µM solutions in 1:1 H_2O/CH_3OH with suitable additives generally give $(M+nH)^{n+}$ or $(M-nH)^{n-}$ ions of sufficiently high abundance for MS dissociation experiments.

The mechanism of RNA hydrolysis in solution is both acid and base catalyzed, and therefore strongly affected by solution pH [18, 19], showing minimal hydrolysis rates at pH values around 5–6 [19]. For acidic and basic ESI additives, not only their efficiency in promoting the formation of $(M+nH)^{n+}$ or $(M-nH)^{n-}$ ions of the desired charge values but also their effect on solution pH and thus rate of RNA hydrolysis needs to be considered. For example, we found that the addition of 0.05 % by volume of acetic acid to a 1 µM solution in 1:1 H_2O/CH_3OH at pH 2.5 caused complete hydrolysis of 34 nt RNA within 5 h, and hydrolysis of larger RNA (57 and 61 nt) was too fast to measure CAD spectra [20]. RNA hydrolysis at pH 9.5 (1 % vol. triethylamine) and 11.5 (25 mM piperidine and 25 mM imidazole) was far slower, with only ~50 % of 34 nt RNA being hydrolyzed after 5 h [20]. In our

laboratory, we therefore prefer the use of basic over acidic additives for generation of $(M - nH)^{n-}$ ions of RNA by ESI. For efficient manipulation of $(M - nH)^{n-}$ ion charge (Fig. 7.2), organic bases such as quinuclidine, 7-methyl-1,5,7-triazabicyclo [4.4.0]dec-5-ene (MTBD), piperidine, and imidazole can be employed as additives [13, 21]. In a systematic study, we found that the average charge of $(M - nH)^{n-}$ ions of RNA and proteins generally increased with increasing gas phase basicity of the additive used, suggesting that gas phase reactions play an important role in determining the charge of $(M - nH)^{n-}$ ions in ESI [13]. However, as illustrated in Fig. 7.2, when piperidine is used together with other basic additives in a 1:1 mixture, it has only little effect on RNA $(M - nH)^{n-}$ ion charge but can significantly reduce alkali ion adduction [13, 21]. For example, efficient suppression of alkali ion adducts in ESI of tRNA was achieved by addition of piperidine/quinuclidine (10 mM each) or piperidine/imidazole (100 mM each) for generation of highly (~0.5 charges/nucleotide) or lowly (~0.2 charges/nucleotide) charged $(M - nH)^{n-}$ ions, respectively [21].

Another important aspect to consider in MS of RNA, which is often overlooked by the less experienced operator, is the cleanliness of the ESI interface of the mass spectrometer. The droplets and ions from ESI can readily pick up (alkali) ions and molecules from surfaces in the interface region that were deposited in previous experiments, which can seriously degrade spectral quality. Moreover, compounds that enter the vacuum system, such as solvent molecules, ESI additives, or volatile analytes, can react with $(M - nH)^{n-}$ ions of RNA by proton transfer, which can reduce $(M - nH)^{n-}$ ion charge. For these reasons it is important that surfaces in the ESI interface region are cleaned frequently, for example, by washing with a 49:49:2 $H_2O/CH_3OH/CH_3COOH$ solution and that time is allowed between experiments for pumping away compounds that could react with $(M - nH)^{n-}$ ions of RNA in undesired proton transfer reactions.

7.2.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Fourier transform ion cyclotron resonance (FTICR) instruments [22, 23] provide unique capabilities such as high mass resolving power and mass accuracy, along with ion isolation and trapping capabilities. Moreover, modern FTICR mass spectrometers are usually equipped with quadrupole and hexapole devices in the higher to medium pressure regions that interface the ESI ion source and the FTICR cell regions at atmospheric and low ($<10^{-9}$ mbar) pressure, respectively. The multipoles are used for ion transfer and accumulation ("linear ion traps") [24, 25] and can optionally serve as ion isolation devices and for vibrational ion activation by energetic collisions with inert gas. Finally, ions trapped in the FTICR cell can be exposed to laser irradiation or react with electrons of varying energy, typically between 0 and 100 eV. Thus, the design of modern FTICR instruments allows for both unambiguous identification of ions according to their m/z values and the design of sophisticated experiment sequences. For example, to study the effect of

Fig. 7.2 Negative ion mass spectra of 22 nt RNA (GGACG AUACG CGUGA AGCGU CC) electrosprayed from 1 µM solutions in 1:1 H$_2$O/CH$_3$OH, using the additives indicated, 7-methyl-1,5,7-triazabicyclo [4.4.0]dec-5-ene (MTBD, 0.1 % Vol.), piperidine (20 mM), and a 1:1 mixture of imidazole and piperidine (25 mM each); adopted from [13]
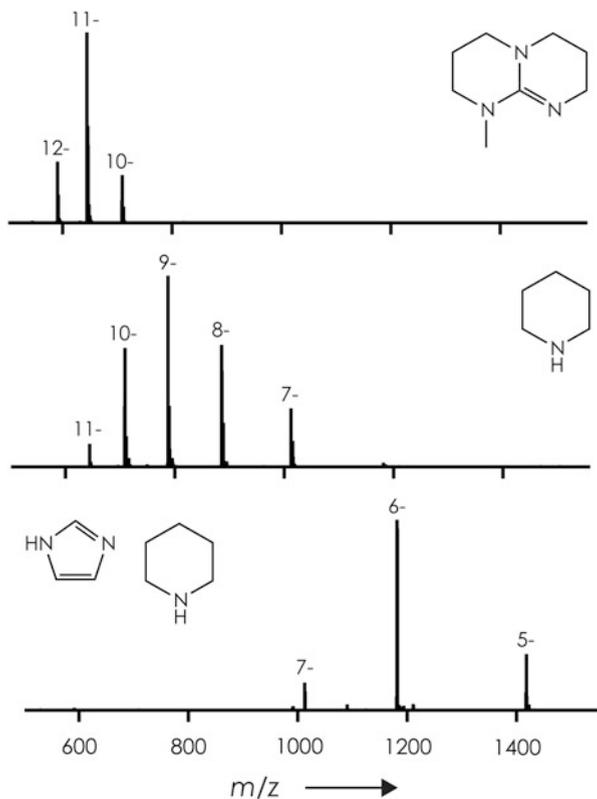
vibrational energy on dissociation of $(M - nH)^{n-}$ ions of RNA in electron detachment dissociation (EDD), we have recently used a complex experimental sequence involving $(M - nH)^{n-}$ ion isolation, their collisional activation, electron detachment, isolation of the $(M - nH)^{(n-1)-\bullet}$ product ions, and their dissociation by infrared laser radiation [21].

Fragment ion assignment and determination of $(M+nH)^{n+}$ or $(M - nH)^{n-}$ ion mass values in ESI mass spectra of larger biomolecules requires both high mass resolving power and mass accuracy. Although FTICR mass spectrometers inherently provide high mass spectrometric performance, this can be limited by "space charge" effects caused by the ions trapped in the FTICR cell, and ion interactions with their image charge on the detection electrodes, all of which affect the observed ion cyclotron frequencies [26–32]. In other words, in the FTICR cell, each ion directly or indirectly affects the cyclotron frequency of another ion, and this effect increases with increasing number of ions and charge density. Because the number of ions trapped in the FTICR cell is different in every different experiment, external calibration of *m/z* spectra is always afflicted with effects from varying numbers and, in most cases, different *m/z* values and proportions of the trapped ions. As a consequence, external calibration can introduce systematic errors in mass

accuracy on the order of typically 10 ppm. Internal calibration of the frequency spectra, on the other hand, largely eliminates the problems caused by ion–ion and ion–image charge interactions and routinely gives mass accuracies to within 1 ppm on our 7 Tesla FTICR instrument, even for larger (>20 kDa) proteins and RNA.

Mass spectra can be calibrated internally when they contain a sufficient number of signals of ions of known mass, the calibrant ions, but this is usually not the case for fragment spectra for which a selected "precursor ion" of the compound under study is isolated prior to dissociation. However, when a sufficient number of precursor ions that correspond to the same compound but differ in net charge are selected for simultaneous dissociation, and provided that dissociation is not complete, internal calibration can use compound ion signals. This is possible even if the compound is not known, as its mass can be determined to within 1 ppm in a separate spectrum with internal calibration [21].

7.2.3 Top-Down Mass Spectrometry

The advantages and limitations of the two general approaches for analysis of biomolecules by mass spectrometry, "top-down" and "bottom-up" MS, have been described and discussed for proteins previously [33–42]. Briefly, in bottom-up MS of proteins, the protein of interest is digested into small peptides prior to analysis by mass spectrometry, whereas top-down MS involves at least two separate mass measurements: one of the intact protein and another one of its fragment ions from unimolecular dissociation. In our laboratory, top-down mass spectrometry generally comprises four separate steps, as illustrated in Fig. 7.3. First, an ESI mass spectrum of the intact protein or RNA is recorded to reveal sample heterogeneity and relative abundances of the species observed (Fig. 7.3a). Second, a compound, usually a water-soluble polymer such as polyethylene glycol, is added to the spray solution for internal calibration and determination of all mass values to within 1 ppm (Fig. 7.3b). Third, a mass spectrum of fragment ions from unimolecular ion dissociation by radical ion chemistry is recorded (Fig. 7.3c); for precursor ions with positive and negative net charge, we utilize electron capture dissociation (ECD) [43] and electron detachment dissociation (EDD) [44], respectively. Fourth, a mass spectrum of fragment ions from slow heating [45] of even-electron $(M + nH)^{n+}$ or $(M - nH)^{n-}$ ions, typically using CAD, is recorded (Fig. 7.3d). For increased sequence coverage, steps three and four can be repeated with precursor ions of different net charge. This top-down approach is rather time-consuming and therefore incompatible with online coupling of liquid separation devices to the mass spectrometer, but fully justified by the extent of information that can be obtained.

Bottom-up MS of proteins is routinely used in many laboratories and is well suited for protein *identification*. Detailed *characterization* of proteins, on the other hand, includes determination of the complete protein sequence and the identification and localization of all modifications. However, full sequence information and any correlations that may exist between individual modifications are lost during the digestion step into small peptides in bottom-up MS. Thus, for protein
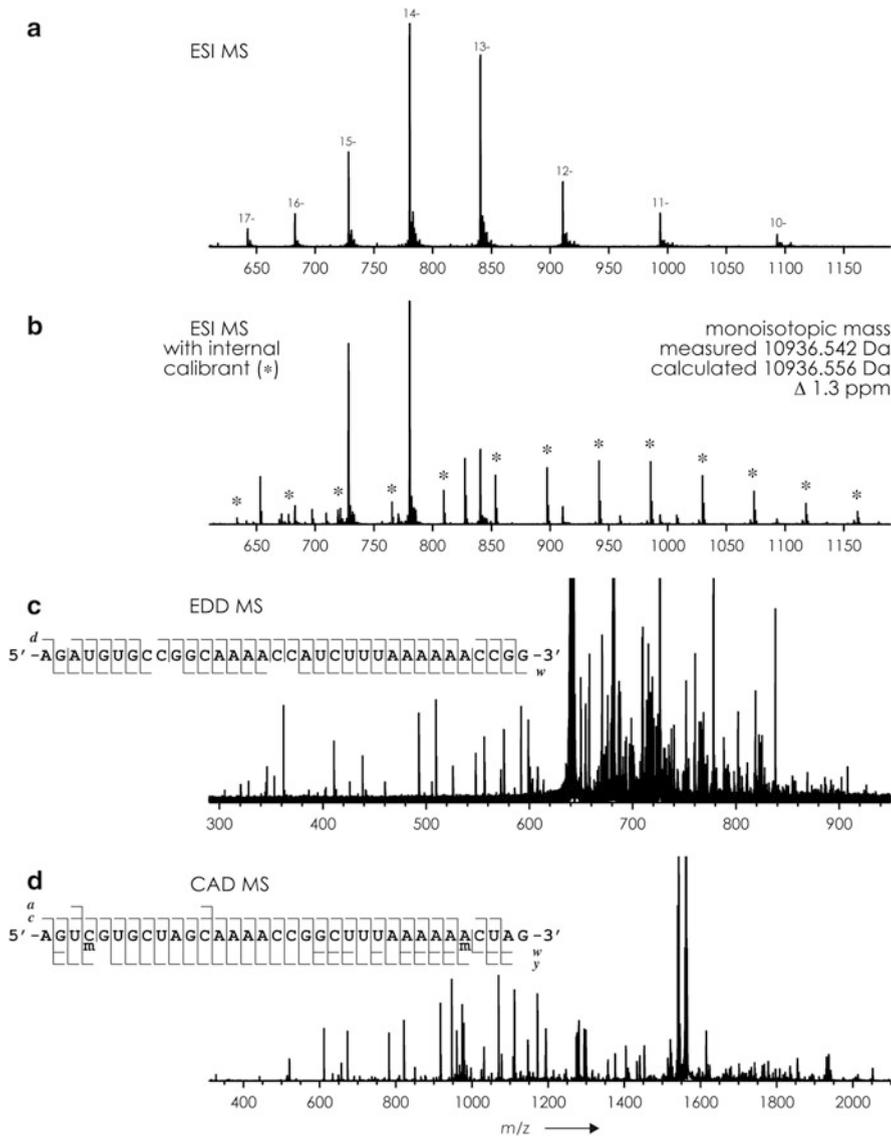
Fig. 7.3 Illustration of the measurements for top-down mass spectrometry, shown here for three different 34 nt RNA sequences for variety; ESI MS spectra of 5′-AGUCmG UGCUA GCAAA ACCGGm CUUUA AAAAA CUAG-3′ (methylation at the ribose C2′-positions of C4 and G20) without (**a**) and with (**b**) internal calibrant (polyethylene 1000), EDD MS spectrum of 5′AGAUG UGCCG GCAAA ACCAU CUUUA AAAAA CCGG-3′, CAD MS spectrum of 5′-AGUCmG UGCUA GCAAA ACCGG CUUUA AAAAAm CUAG-3′ (methylation at the ribose C2′-positions of C4 and A30); fragment ion maps in (**c**, **d**) illustrate sequence coverage

characterization, the top-down approach is far better suited, as it immediately reveals sample heterogeneity, the types of modifications present, and the relative abundances of the different protein forms. Moreover, each protein form can be isolated and dissociated in the mass spectrometer, which can provide full sequence information and disclose the sites of modifications. Top-down MS also has the potential for detailed analysis of "combinatorial codes" in posttranslationally modified proteins [42].

The above considerations apply equally to top-down MS of RNA and its posttranscriptional or synthetic modifications, although as a result of the different chemical nature of proteins and RNA, the dissociation techniques used can be different. Accordingly, top-down MS of RNA in our laboratory (Fig. 7.3) involves recording ESI mass spectra of the intact RNA with and without internal calibrant, separate experiments using two techniques suitable for dissociation of RNA, collisionally activated dissociation and electron detachment dissociation, and thorough data analysis.

As can be seen from the small number of publications on the topic, top-down mass spectrometry of RNA is only beginning to be developed: In 2004, Fabris and coworkers used top-down MS for characterization of RNA and DNA, consisting of up to 22 nucleotides, that was modified by "structural probes" to obtain information about base pairing and steric protection in nucleic acid structures [46]. In 2008, McLuckey and coworkers demonstrated top-down MS of 21 nt small interfering RNA (siRNA) [47]. In 2010, our group first demonstrated full sequence coverage in top-down MS of RNA consisting of 34 nucleotides [16] and in the same year extended this limit to 61 nt RNA [20]. Top-down MS of tRNA was reported by McLuckey and coworkers also in 2010 [48] and by our group in 2012 [21].

7.2.4 Collisionally Activated Dissociation of RNA

Multiple, low-energy collisions of RNA $(M - nH)^{n-}$ ions with inert background gas (collisionally activated dissociation, CAD) can incrementally increase their vibrational energy to the point at which unimolecular ion dissociation is observed. Thresholds for ion fragmentation can just as well be reached by slow ion heating using infrared radiation in IRMPD (infrared multiphoton dissociation) or BIRD (blackbody infrared multiphoton dissociation) experiments [49, 50], but CAD is by far the most commonly used technique for slow ion heating.

The net negative charge of $(M - nH)^{n-}$ ions critically affects the types of fragment ions observed in CAD of RNA. McLuckey and coworkers have shown that more lowly charged RNA ions primarily dissociate into *c* and *y* ions (Scheme 7.1), whereas CAD of more highly charged ions causes mostly fragmentation into *a* and *w* ions, with *a* ions exhibiting a high tendency for base loss [52]. Although the type of backbone fragmentation (*c/y* versus *a/w* or *a*-base/*w*) does not generally limit RNA sequencing by MS, base loss from *a* ions can seriously complicate spectral interpretation as the extent of base loss can differ substantially for different *a* ions [53]. In other words, in the same spectrum from

Scheme 7.1 Nomenclature for fragment ions from unimolecular dissociation of ribonucleic acids [51]
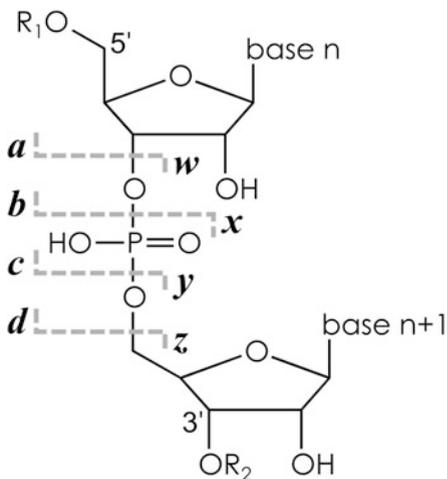
CAD of a given RNA, one may find only *a* ions or only *a*-base ions from some backbone cleavage sites and varying fractions of *a* and *a*-base ions from other sites. If base loss from *a* ions were instead uniform, one could use *a*-base and *w* ions in a similar way that we use *d* and *w* ions from EDD for assignment of fragment ion identity as described in the section on data analysis and fragment ion assignments.

Backbone cleavage into *c* and *y* ions is sufficiently unselective with respect to the residues framing the cleavage site to allow for 100 %, 98 %, and 89 % sequence coverage in a single spectrum from CAD of 34, 61, and 76 nt RNA, respectively, provided that secondary fragmentation and base loss are kept at a minimum [16, 20, 21]. This can be achieved by collisional cooling of primary fragment ions and selection of precursor ions of relatively low negative net charge (~0.2/nt); the latter at the same time disfavors formation of *a*, *a*-base, and *w* ions [16, 20, 52]. CAD spectra recorded under these conditions are far easier to interpret than those from CAD of more highly charged ions, because they show predominantly *c* and *y* ions [16, 20, 21, 47].

7.2.5 Electron Detachment Dissociation of RNA

Electron detachment dissociation (EDD) was introduced by Zubarev and coworkers in 2001 for dissociation of $(M - nH)^{n-}$ ions of peptides [44], and the method has since been extended to acidic proteins [54], carbohydrates [55–58], DNA [59–61], and RNA [20, 21, 62]. EDD provides a means of generating radical $(M - nH)^{(n-1)-\bullet}$ from even-electron $(M - nH)^{n-}$ ions by electron *detachment*, similar to ECD converting $(M+nH)^{n+}$ into radical $(M+nH)^{(n-1)+\bullet}$ ions by electron *attachment*. However, the chemistry involved in unimolecular dissociation of $(M - nH)^{(n-1)-\bullet}$ and $(M+nH)^{(n-1)+\bullet}$ ions is generally different because not only the location of radical sites, but also the types of radicals, hydrogen-deficient versus hydrogen

abundant, are generally different [20, 43, 54, 62–64]. In EDD, $(M - nH)^{n-}$ ions are exposed to fast (>10 eV) electrons that, depending on experimental factors such as electron energy and density, cause detachment of one or more electrons from the $(M - nH)^{n-}$ ions. Because radiofrequency fields and higher background pressures in mass spectrometers other than FTICR instruments severely limit electron mean free paths, EDD is restricted to FTICR MS. However, the recently developed methods negative electron transfer dissociation (nETD) [65–68] and electron photodetachment dissociation (EPD) [69–73] can also produce $(M - nH)^{(n-1)-\bullet}$ from $(M - nH)^{n-}$ ions, but further studies are needed to evaluate possible differences between fragmentation mechanisms in EDD, nETD, and EPD, especially for RNA. The first study on EDD of $(M - nH)^{n-}$ ions of RNA, which consisted of six nucleotides, was performed in the Håkansson group [62]. In 2010, our group has extended EDD to RNA consisting of up to 61 nt and suggested a mechanism for dissociation of $(M - nH)^{(n-1)-\bullet}$ ions of RNA in EDD [20]. In 2012, we have further extended EDD to highly modified tRNA consisting of 76 nt [21].

EDD of $(M - nH)^{n-}$ ions of RNA produces noncomplementary, even-electron **d** and **w** fragments (Scheme 7.1), but no **a**, **b**, **c**, **x**, **y**, or **z** fragments [20]. This finding is astonishing because the radical species $(M - nH)^{(n-1)-\bullet}$ from electron detachment dissociates into even-electron **d** and **w** fragments, and because it requires breaking of both the C3′–O3′ and the P–O5′ bond of the same phosphodiester moiety. Our proposed mechanism for backbone cleavage in EDD is accordingly complex [20].

A detailed analysis of fragment ion abundances from EDD of a 22 nt RNA revealed a strong correlation between site-specific extent of fragmentation and nucleobase ionization energy, suggesting that electrons are detached from the nucleobases in EDD of RNA [20]. Considering the strong evidence for electron detachment from negatively charged sites in EDD of $(M - nH)^{n-}$ ions of proteins [54], this finding was surprising, as one could, by analogy, expect electron detachment from the negatively charged phosphate moieties in EDD of RNA. After all, the anion $H_2PO_4^-$ has a far lower electron detachment energy (4.57 ± 0.01 eV) (http://webbook.nist.gov) than the neutral nucleobases adenine (8.3 ± 0.1 eV), cytosine (9.0 ± 0.1 eV), guanine (8.0 ± 0.2 eV), and uracil (9.4 ± 0.1 eV) [74]. However, calculations suggest that in nucleic acids, the negative charge of the phosphodiester can reduce nucleobase ionization energies by several eV and that phosphate electron detachment energies can be increased by hydrogen bonding to sugar or base moieties, such that electron detachment from the nucleobases becomes energetically favorable [75, 76].

Vibrational activation of $(M - 8H)^{8-}$, $(M - 9H)^{9-}$, and $(M - 10H)^{10-}$ ions of 22 nt RNA, either before or after electron detachment, showed no appreciable effect on sequence coverage, suggesting that higher order gas phase structure is not a limiting factor in EDD of more highly charged ($0.36-0.46$ charges/nt) 22 nt RNA. Likewise, vibrational activation of $(M - nH)^{n-}$ ions of tRNA (76 nt) with n $= 27-35$, also corresponding to $0.36-0.46$ charges/nt, prior to EDD, did not significantly increase sequence coverage (81 % compared to 80 % without activation)

[21]. However, for 22 nt RNA, sequence coverage did increase from 90 % for $(M - 8H)^{8-}$ and $(M - 9H)^{9-}$ ions to 100 % for $(M - 10H)^{10-}$ ions, and increased net charge gave more random backbone cleavage, probably by leveling differences between nucleobase ionization energies [21].

7.2.6 Data Analysis and Fragment Ion Assignments

FTICR mass spectra of products from unimolecular dissociation of RNA by CAD and EDD are highly complex. Each fragment ion gives rise to a number of isotopic peaks arranged as "isotopic clusters," whose isotopes are spaced by approximately the reverse of the ion charge on the m/z axis. Isotopic resolution as provided by FTICR instrument thus allows for determination of ion charge, and therefore ion mass, in an m/z spectrum. Most algorithms for identification of isotopic clusters and assignment of monoisotopic mass values in complex spectra were originally developed for proteins [77–82] but can also be used for RNA, provided that the averaged elemental residue composition [83], which is naturally different for proteins and RNA, can be defined by the user. For establishing fragment mass ladders [37], differences between monoisotopic fragment mass values are analyzed. However, analysis of fragments from dissociation of an unknown sequence with unknown modifications is tedious or even impossible because monoisotopic mass values cannot simply be assigned as fragments carrying either the 5′ terminus (*c* and *d* ions) or the 3′ terminus (*w* and *y* ions). In other words, the "sequence direction" of the mass values in a CAD or an EDD spectrum is generally not known. However, when data from CAD and EDD are analyzed together, fragment ion identity (*c*, *d*, *w*, or *y*) can be assigned with high confidence as outlined below.

The dissociation chemistry observed in EDD of $(M - nH)^{n-}$ ions of RNA is unique in that it consistently yields noncomplementary *d* and *w* ions, whose mass values differ from those of the complementary *c* and *y* ions from CAD by 18.01 and 79.97 Da, respectively. So not only are the mass values of *c* and *d* ions, and *y* and *w* ions, different, on top of that, the mass differences between *c* and *d* ions, and *y* and *w* ions, are different (Scheme 7.2)! For this reason, a comparison of mass values from only one EDD and only one CAD spectrum can give four fragment ion mass ladders with known sequence direction, as previously demonstrated for highly modified, 76 nt tRNA [21]. Importantly, neither CAD nor EDD showed fragment ions consistent with loss of any modifications [21].

This new top-down MS approach can reveal the presence, type, and location of all mass-altering posttranscriptional or synthetic modifications of RNA and even identify so far unknown types of modifications. RNA residues with the same elemental composition and mass, for example, pseudouridine and uridine, cannot be distinguished from each other with this approach. Such "mass-silent" modifications can, however, be detected by complementary experimental strategies based on mass spectrometry [12, 84–87] or chemical reactivity [3].
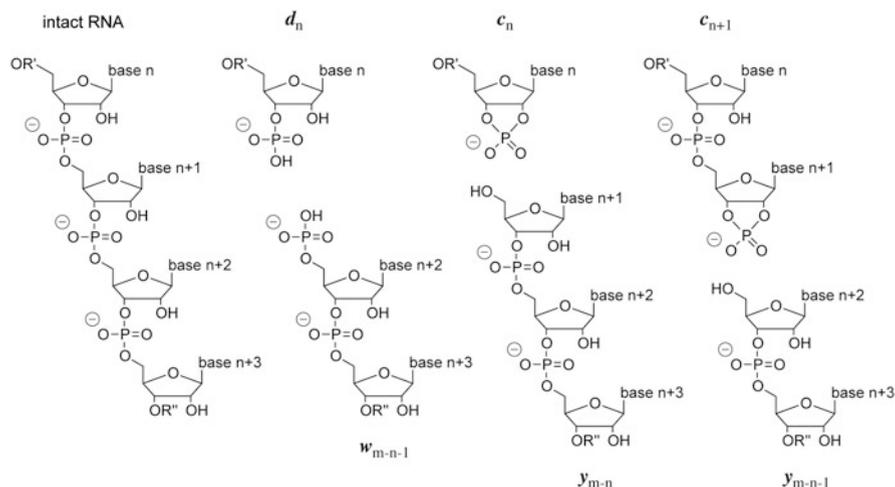
Scheme 7.2 RNA consisting of m nucleotides (*left*) dissociated by EDD (*d* and *w* ions) and CAD (*c* and *y* ions), illustrating the mass differences of 18.01 Da (H_2O) between d_n and c_n ions and 79.97 Da (HPO_3) between w_{m-n-1} and y_{m-n-1} ions [21]. Dissociation into *d* and *w* ions in EDD involves loss of a neutral, radical nucleotide unit carrying base $n + 1$ [20]

7.3   Conclusion and Outlook

Recent developments in top-down mass spectrometry of RNA now allow for detailed characterization of posttranscriptionally and synthetically modified ribonucleic acids. These include efficient manipulation of $(M - nH)^{n-}$ ion charge by use of organic bases as additives to the ESI solution and the extension of CAD and EDD to larger (up to 76 nucleotides) and highly modified RNA. Future challenges for top-down MS of RNA include the development of new methodology for the relative quantification of modifications and improved algorithms for automated MS data analysis.

7.4   Summary of Key Concepts

- Efficient manipulation of $(M - nH)^{n-}$ ion charge and suppression of alkali ion adducts in electrospray ionization of ribonucleic acids can now provide highly (~0.5 charges/nt) and lowly (~0.2 charges/nt) charged precursor ions for top-down MS.
- Extensive RNA sequence coverage can be obtained from electron detachment dissociation, EDD, and collisionally activated dissociation, CAD, using Fourier transform ion cyclotron resonance instruments.
- The unique chemistry involved in unimolecular dissociation of ribonucleic acid anions by EDD and CAD produces unique fragment ions from RNA backbone

cleavage that can be aligned as four separate fragment mass ladders with known sequence direction.

- Posttranscriptional modifications were fully preserved in EDD and CAD of highly modified transfer RNA, highlighting the potential of top-down MS for detailed characterization of modified ribonucleic acids.

References

1. Hüttenhofer A, Schattner P, Polacek N (2005) Non-coding RNAs: hope or hype? Trends Genet 21:289–297
2. Shendure J, Ji HL (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145
3. Kellner S, Burhenne J, Helm M (2010) Detection of RNA modifications. RNA Biol 7:237–247
4. Le DD, Fujimori DG (2012) Protein and nucleic acid methylating enzymes: mechanisms and regulation. Curr Opin Chem Biol 16:507–515
5. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71
6. Karas M, Bachmann D, Hillenkamp F (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. Anal Chem 57: 2935–2939
7. Bahr U, Aygun H, Karas M (2009) Sequencing of single and double stranded RNA oligonucleotides by acid hydrolysis and MALDI mass spectrometry. Anal Chem 81: 3173–3179
8. Berhane BT, Limbach PA (2003) Stable isotope labeling for matrix-assisted laser desorption/ionization mass spectrometry and post-source decay analysis of ribonucleic acids. J Mass Spectrom 38:872–878
9. Meng ZJ, Limbach PA (2004) RNase mapping of intact nucleic acids by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICRMS) and O-18 labeling. Int J Mass Spectrom 234:37–44
10. Meng ZJ, Limbach PA (2005) Shotgun sequencing small oligonucleotides by nozzle-skimmer dissociation and electrospray ionization mass spectrometry. Eur J Mass Spectrom 11:221–229
11. Hossain M, Limbach PA (2007) Mass spectrometry-based detection of transfer RNAs by their signature endonuclease digestion products. RNA 13:295–303
12. Taucher M, Ganisl B, Breuker K (2011) Identification, localization, and relative quantitation of pseudouridine in RNA by tandem mass spectrometry of hydrolysis products. Int J Mass Spectrom 304:91–97
13. Ganisl B, Taucher M, Riml C, Breuker K (2011) Charge as you like! Efficient manipulation of negative ion net charge in electrospray ionization of proteins and nucleic acids. Eur J Mass Spectrom 17:333–343
14. Greig MJ, Gaus HJ, Griffey RH (1996) Negative ionization micro electrospray mass spectrometry of oligonucleotides and their complexes. Rapid Commun Mass Spectrom 10:47–50
15. Griffey RH, Sasmor H, Greig MJ (1997) Oligonucleotide charge states in negative ionization electrospray mass spectrometry are a function of solution ammonium ion concentration. J Am Soc Mass Spectrom 8:155–160
16. Taucher M, Rieder U, Breuker K (2010) Minimizing base loss and internal fragmentation in collisionally activated dissociation of multiply deprotonated RNA. J Am Soc Mass Spectrom 21:278–285

17. Greig M, Griffey RH (1995) Utility of organic bases for improved electrospray mass spectrometry of oligonucleotides. Rapid Commun Mass Spectrom 9:97–102
18. Perreault DM, Anslyn EV (1997) Unifying the current data on the mechanism of cleavage - transesterification of RNA. Angew Chem Int Ed 36:432–450
19. Oivanen M, Kuusela S, Lonnberg H (1998) Kinetics and mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by Brønsted acids and bases. Chem Rev 98:961–990
20. Taucher M, Breuker K (2010) Top-down mass spectrometry for sequencing of larger (up to 61 nt) RNA by CAD and EDD. J Am Soc Mass Spectrom 21:918–929
21. Taucher M, Breuker K (2012) Characterization of modified RNA by top-down mass spectrometry. Angew Chem Int Ed 51:11289–11292
22. Comisarow MB, Marshall AG (1974) Fourier transform ion cyclotron resonance spectroscopy. Chem Phys Lett 25:282–283
23. Marshall AG, Hendrickson CL, Jackson GS (1998) Fourier transform ion cyclotron resonance mass spectrometry: a primer. Mass Spectrom Rev 17:1–35
24. Senko MW, Hendrickson CL, Emmett MR, Shi SDH, Marshall AG (1997) External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 8:970–976
25. Schwartz JC, Senko MW, Syka JEP (2002) A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom 13:659–669
26. Taylor PK, Amster IJ (2003) Space charge effects on mass accuracy for multiply charged ions in ESI-FTICR. Int J Mass Spectrom 222:351–361
27. Masselon C, Tolmachev AV, Anderson GA, Harkewicz R, Smith RD (2002) Mass measurement errors caused by "local" frequency perturbations in FTICR mass spectrometry. J Am Soc Mass Spectrom 13:99–106
28. Zhang LK, Rempel D, Pramanik BN, Gross ML (2005) Accurate mass measurements by Fourier transform mass spectrometry. Mass Spectrom Rev 24:286–309
29. Leach FE, Kharchenko A, Heeren RMA, Nikolaev E, Amster IJ (2010) Comparison of particle-in-cell simulations with experimentally observed frequency shifts between ions of the same mass-to-charge in Fourier transform ion cyclotron resonance mass spectrometry. J Am Soc Mass Spectrom 21:203–208
30. Jeffries JB, Barlow SE, Dunn GH (1983) Theory of space charge shift of ion cyclotron resonance frequencies. Int J Mass Spectrom Ion Process 54:169–187
31. Gorshkov MV, Marshall AG, Nikolaev EN (1993) Analysis and elimination of systematic errors originating from Coulomb mutual interaction and image charge in Fourier transform ion cyclotron resonance precise mass difference measurements. J Am Soc Mass Spectrom 4: 855–868
32. Hendrickson CL, Beu SC, Laude DA (1993) Two-dimensional Coulomb-induced frequency modulation in Fourier transform ion cyclotron resonance: a mechanism for line broadening at high mass and for large ion populations. J Am Soc Mass Spectrom 4:909–916
33. Kelleher NL (2004) Top-down proteomics. Anal Chem 76:196A–203A
34. Fridriksson EK, Beavil A, Holowka D, Gould HJ, Baird B, McLafferty FW (2000) Heterogeneous glycosylation of immunoglobulin E constructs characterized by top-down high-resolution 2D mass spectrometry. Biochemistry 39:3369–3376
35. Bogdanov B, Smith RD (2005) Proteomics by FT-ICR mass spectrometry: top-down and bottom-up. Mass Spectrom Rev 24:168–200
36. Loo RRO, Hayes R, Yang YN, Hung F, Ramachandran P, Kim N, Gunsalus R, Loo JA (2005) Top-down, bottom-up, and side-to-side proteomics with virtual 2D gels. Int J Mass Spectrom 240:317–325
37. Chait BT (2006) Bottom-up or top-down? Science 314:65–66
38. Breuker K, Jin M, Han XM, Jiang HH, McLafferty FW (2008) Top-down identification and characterization of biomolecules by mass spectrometry. J Am Soc Mass Spectrom 19: 1045–1053

39. Ma MM, Chen RB, Ge Y, He H, Marshall AG, Li LJ (2009) Combining bottom-up and top-down mass spectrometric strategies for *de novo* sequencing of the crustacean hyperglycemic hormone from cancer borealis. Anal Chem 81:240–247

40. Wu S, Lourette NM, Tolic N, Zhao R, Robinson EW, Tolmachev AV, Smith RD, Pasa-Tolic L (2009) An integrated top-down and bottom-up strategy for broadly characterizing protein isoforms and modifications. J Proteome Res 8:1347–1357

41. Cui WD, Rohrs HW, Gross ML (2011) Top-down mass spectrometry: recent developments, applications and perspectives. Analyst 136:3854–3864

42. Lanucara F, Eyers CE (2013) Top-down mass spectrometry for the analysis of combinatorial post-translational modifications. Mass Spectrom Rev 32:27–42

43. Zubarev RA, Kelleher NL, McLafferty FW (1998) Electron capture dissociation of multiply charged protein cations. A nonergodic process. J Am Chem Soc 120:3265–3266

44. Budnik BA, Haselmann KF, Zubarev RA (2001) Electron detachment dissociation of peptide di-anions: an electron-hole recombination phenomenon. Chem Phys Lett 342:299–302

45. McLuckey SA, Goeringer DE (1997) Slow heating methods in tandem mass spectrometry. J Mass Spectrom 32:461–474

46. Kellersberger KA, Yu E, Kruppa GH, Young MM, Fabris D (2004) Top-down characterization of nucleic acids modified by structural probes using high-resolution tandem mass spectrometry and automated data interpretation. Anal Chem 76:2438–2445

47. Huang TY, Liu J, Liang XR, Hodges BDM, McLuckey SA (2008) Collision-induced dissociation of intact duplex and single-stranded siRNA anions. Anal Chem 80:8501–8508

48. Huang TY, Liu JA, McLuckey SA (2010) Top-down tandem mass spectrometry of tRNA Via ion trap collision-induced dissociation. J Am Soc Mass Spectrom 21:890–898

49. Hofstadler SA, Sannes-Lowery KA, Hannis JC (2005) Analysis of nucleic acids by FTICR MS. Mass Spectrom Rev 24:265–285

50. Gardner MW, Li N, Ellington AD, Brodbelt JS (2010) Infrared multiphoton dissociation of small-interfering RNA anions and cations. J Am Soc Mass Spectrom 21:580–591

51. McLuckey SA, Van Berkel GJ, Glish GL (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. J Am Soc Mass Spectrom 3:60–70

52. Huang TY, Kharlamova A, Liu J, McLuckey SA (2008) Ion trap collision-induced dissociation of multiply deprotonated RNA: *c*/*y*-ions versus (*a*-B)/*w*-ions. J Am Soc Mass Spectrom 19:1832–1840

53. Tromp JM, Schürch S (2006) Electrospray ionization tandem mass spectrometry of biphenyl-modified oligo(deoxy)ribonucleotides. Rapid Commun Mass Spectrom 20:2348–2354

54. Ganisl B, Valovka T, Hartl M, Taucher M, Bister K, Breuker K (2011) Electron detachment dissociation for top-down mass spectrometry of acidic proteins. Chem Eur J 17:4460–4469

55. McFarland MA, Marshall AG, Hendrickson CL, Nilsson CL, Fredman P, Mansson JE (2005) Structural characterization of the GM1 ganglioside by infrared multiphoton dissociation/electron capture dissociation, and electron detachment dissociation electrospray ionization FT-ICR MS/MS. J Am Soc Mass Spectrom 16:752–762

56. Adamson JT, Håkansson K (2007) Electron detachment dissociation of neutral and sialylated oligosaccharides. J Am Soc Mass Spectrom 18:2162–2172

57. Wolff JJ, Chi LL, Linhardt RJ, Amster IJ (2007) Distinguishing glucuronic from iduronic acid in glycosaminoglycan tetrasaccharides by using electron detachment dissociation. Anal Chem 79:2015–2022

58. Wolff JJ, Laremore TN, Busch AM, Linhardt RJ, Amster IJ (2008) Electron detachment dissociation of dermatan sulfate oligosaccharides. J Am Soc Mass Spectrom 19:294–304

59. Yang J, Mo JJ, Adamson JT, Håkansson K (2005) Characterization of oligodeoxynucleotides by electron detachment dissociation Fourier transform ion cyclotron resonance mass spectrometry. Anal Chem 77:1876–1882

60. Mo JJ, Håkansson K (2006) Characterization of nucleic acid higher order structure by high-resolution tandem mass spectrometry. Anal Bioanal Chem 386:675–681

61. Kinet C, Gabelica V, Balbeur D, De Pauw E (2009) Electron detachment dissociation (EDD) pathways in oligonucleotides. Int J Mass Spectrom 283:206–213

62. Yang J, Håkansson K (2006) Fragmentation of oligoribonucleotides from gas phase ion-electron reactions. J Am Soc Mass Spectrom 17:1369–1375

63. Moore B, Sun QY, Hsu JC, Lee AH, Yoo GC, Ly T, Julian RR (2012) Dissociation chemistry of hydrogen-deficient radical peptide anions. J Am Soc Mass Spectrom 23:460–468

64. Moore BN, Ly T, Julian RR (2011) Radical conversion and migration in electron capture dissociation. J Am Chem Soc 133:6997–7006

65. Huzarska M, Ugalde I, Kaplan DA, Hartmer R, Easterling ML, Polfer NC (2010) Negative electron transfer dissociation of deprotonated phosphopeptide anions: choice of radical cation reagent and competition between electron and proton transfer. Anal Chem 82:2873–2878

66. Wolff JJ, Leach FE, Laremore TN, Kaplan DA, Easterling ML, Linhardt RJ, Amster IJ (2010) Negative electron transfer dissociation of glycosaminoglycans. Anal Chem 82:3460–3466

67. Huang TY, McLuckey SA (2011) Gas-phase ion/ion reactions of rubrene cations and multiply charged DNA and RNA anions. Int J Mass Spectrom 304:140–147

68. Coon JJ, Shabanowitz J, Hunt DF, Syka JEP (2005) Electron transfer dissociation of peptide anions. J Am Soc Mass Spectrom 16:880–882

69. Gabelica V, Tabarin T, Antoine R, Rosu F, Compagnon I, Broyer M, De Pauw E, Dugourd P (2006) Electron photodetachment dissociation of DNA polyanions in a quadrupole ion trap mass spectrometer. Anal Chem 78:6564–6572

70. Antoine R, Joly L, Tabarin T, Broyer M, Dugourd P, Lemoine J (2007) Photo-induced formation of radical anion peptides. Electron photodetachment dissociation experiments. Rapid Commun Mass Spectrom 21:265–268

71. Antoine R, Joly L, Allouche AR, Broyer M, Lemoine J, Dugourd P (2009) Electron photodetachment of trapped doubly deprotonated angiotensin peptides. UV spectroscopy and radical recombination. Eur Phys J D 51:117–124

72. Larraillet V, Antoine R, Dugourd P, Lemoine J (2009) Activated-electron photodetachment dissociation for the structural characterization of protein polyanions. Anal Chem 81:8410–8416

73. Larraillet V, Vorobyev A, Brunet C, Lemoine J, Tsybin YO, Antoine R, Dugourd P (2010) Comparative dissociation of peptide polyanions by electron impact and photo-induced electron detachment. J Am Soc Mass Spectrom 21:670–680

74. Verkin BI, Sukhodub LF, Yanson IK (1976) Potentials of ionization of nitrogen bases of nucleic acids. Dokl Akad Nauk SSSR 228:1452–1455

75. Zakjevskii VV, King SJ, Dolgounitcheva O, Zakrzewski VG, Ortiz JV (2006) Base and phosphate electron detachment energies of deoxyribonucleotide anions. J Am Chem Soc 128:13350–13351

76. Rubio M, Roca-Sanjuan D, Serrano-Andres L, Merchan M (2009) Determination of the electron detachment energies of 2′-deoxyguanosine 5′-monophosphate anion: influence of the conformation. J Phys Chem B 113:2451–2457

77. Horn DM, Zubarev RA, McLafferty FW (2000) Automated de novo sequencing of proteins by tandem high-resolution mass spectrometry. Proc Natl Acad Sci USA 97:10313–10317

78. Horn DM, Zubarev RA, McLafferty FW (2000) Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J Am Soc Mass Spectrom 11:320–332

79. Kaur P, O'Connor PB (2006) Algorithms for automatic interpretation of high resolution mass spectra. J Am Soc Mass Spectrom 17:459–468

80. Chen L, Sze SK, Yang H (2006) Automated intensity descent algorithm for interpretation of complex high-resolution mass spectra. Anal Chem 78:5006–5018

81. Park K, Yoon JY, Lee S, Paek E, Park H, Jung HJ, Lee SW (2008) Isotopic peak intensity ratio based algorithm for determination of isotopic clusters and monoisotopic masses of polypeptides from high-resolution mass spectrometric data. Anal Chem 80:7294–7303

82. Strohalm M, Kavan D, Novak P, Volny M, Havlicek V (2010) A cross-platform software environment for precise analysis of mass spectrometric data. Anal Chem 82:4648–4651
83. Senko MW, Beu SC, McLafferty FW (1995) Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. J Am Soc Mass Spectrom 6:229–233
84. Bruckl T, Globisch D, Wagner M, Müller M, Carell T (2009) Parallel isotope-based quantification of modified tRNA nucleosides. Angew Chem Int Ed 48:7932–7934
85. Durairaj A, Limbach PA (2008) Matrix-assisted laser desorption/ionization mass spectrometry screening for pseudouridine in mixtures of small RNAs by chemical derivatization, RNase digestion and signature products. Rapid Commun Mass Spectrom 22:3727–3734
86. Durairaj A, Limbach PA (2008) Mass spectrometry of the fifth nucleoside: a review of the identification of pseudouridine in nucleic acids. Anal Chim Acta 623:117–125
87. Durairaj A, Limbach PA (2008) Improving CMC-derivatization of pseudouridine in RNA for mass spectrometric detection. Anal Chim Acta 612:173–181

Quantification of DNA Damage Using Mass Spectrometry Techniques

8

Thierry Douki and Jean-Luc Ravanat

Abstract

The role of DNA in cells relies on its chemical structure. Unfortunately, a number of physical and chemical agents may damage DNA, leading to modification of genetic information or to cell death. Detecting DNA damage is thus a major issue in numerous studies. Analytical methods have therefore been developed in order to quantify modified DNA bases. In particular, mass spectrometry is used as a very specific and sensitive detector when combined to gas or liquid chromatography. This review will first briefly present the major classes of DNA damage and then focus on two mass spectrometry-based approaches for the quantification of modified bases following DNA hydrolysis into monomers.

Abbreviations/Formulae

64PPS	6-4 Pyrimidine dimer photoproducts (35)
B(a)P	Benzo[a]pyrene
BPDE	Benzo(a)pyrene diol epoxide
CPDs	Cyclobutane pyrimidine dimers (34)
DEW	Pyrimidine dimer Dewar valence isomer (36)
EIMS	Electron impact ionization mass spectrometry
GC	Gas chromatography
HPLC	High-performance liquid chromatography

T. Douki (✉) • J.-L. Ravanat
Laboratoire "Lésions des Acides Nucléiques", Université Joseph Fourier – Grenoble 1/CEA/
Institut Nanoscience et Cryogénie/SCIB, UMR-E3, Grenoble, France
e-mail: thierry.douki@cea.fr

LMDS   Locally multiply damaged sites
MMS   Methyl methanesulfonate
MS/MS   Tandem mass spectrometry
MS   Mass spectrometry
NCIMS   Negative chemical ionization mass spectrometry

8.1   Introduction

In cells, the key role played by DNA in the synthesis of specific proteins and in the transmission of genetic information from one cell generation to the other relies on its chemical structure. Indeed, the complementarity between adenine and thymine on the one hand and guanine and cytosine on the other hand makes possible the synthesis of messenger RNA and the replication of DNA. Formation of A:T and G:C base pairs by hydrogen bonding is also at the origin of the double-stranded structure of DNA. It thus appears that any modification of the chemical structure of DNA, of either the bases or the phosphodiester backbone, can drastically impair its normal functions and threaten the cell or the organism. DNA polymerases replicate the double helix before cell division, When encountering a modified base, DNA polymerases may incorporate in the strand under synthesis a base different from the original one. This misincorporation in the synthesized sequence is a point mutation that alters the function of the coded protein, if the considered portion of DNA is located in a gene. Another possible consequence of the presence of modified bases and of breaks in the sugar-phosphate backbone may be the blocking of the synthesis leading to either cell death or chromosome breakage. These processes leading to the degradation of the genetic information are potentially at the origin of tumors and cancer. A precise understanding of the nature and the amount of DNA damage produced by physical agents (ionizing radiation, ultraviolet light) and chemical compounds (pollutants, industrial products) is thus of outmost importance in toxicology. In this field, mass spectrometry is increasingly used. It has been a major tool for many years in mechanistic studies on model systems aimed at precisely determining the chemical structure of the reaction products between DNA and genotoxics. More recently, analytical techniques combining chromatographic separation and mass spectrometry detection became very valuable approaches due to their sensitivity and specificity. The present chapter, after a survey of the main genotoxic pathways, will focus on the analytical application of mass spectrometry to the field of DNA damage.

8.2   The Main Types of DNA Damage

DNA plays a key role in maintaining cell integrity. Therefore, modification of DNA structure could have severe biological consequences, including cell death or, even worse, mutagenicity. Therefore, since 1950 efforts have been made to identify the

main types of DNA damage that could be generated in cells. In the following chapter we will describe the lesions that could be produced under conditions of oxidative stress, in the presence of alkylating agents and following exposure to UV light. The main DNA decomposition products that have been identified, mostly based on extensive studies of free nucleosides, will be presented together with their mechanisms of formation under physiological conditions. In other contexts for instance in the gas phase, the chemistry of DNA bases may be drastically different but such processes are not covered in this chapter.

8.2.1 Oxidation Products

As much as about 60 different oxidation-induced decomposition products of DNA constituents have been identified so far [1, 2]. The decomposition reactions of the main DNA constituents mediated by hydroxyl radical (HO·), one-electron oxidation, and singlet oxygen ($^1O_2$) will be described. The decomposition reactions of DNA bases following one electron oxidation and HO·-mediated oxidation are presented in parallel, because the final decomposition products are common to these two decomposition pathways. Reaction of singlet oxygen with DNA constituent is simpler since its reactivity is limited to guanine.

8.2.1.1 DNA Oxidation Mediated by Singlet Oxygen

Singlet molecular oxygen ($^1O_2$) is an excited state of molecular oxygen which exhibits a strong reactivity towards double bonds and electron-rich molecules. Regarding DNA it has been clearly demonstrated that its reactivity is limited to the guanine base (*1*) [3]. Decomposition of this purine base (Fig. 8.1) following reaction with $^1O_2$ involves first the formation of an endoperoxide *2* between C4 and C8 generated through a Diels–Alder addition. In double-stranded DNA, conversion of the unstable endoperoxide gives rise exclusively to 8-oxo-7,8-dihydro-2′-deoxyguanosine (***8-oxodGuo, 5***).

8.2.1.2 HO· and One Electron-Induced Oxidation of DNA Bases

Similar decomposition products are obtained following either one-electron oxidation or HO·-mediated decomposition of DNA bases. This is explained by the fact that common transiently produced radicals may be generated by the two above-mentioned decomposition mechanisms. Indeed, following one electron-mediated oxidation of DNA bases, base radical cations are generated. These radical cations could undergo two distinct decomposition reactions, either through hydration or through deprotonation. Hydration gives rise to radical intermediates that are similar to those generated by addition of HO· onto the base moiety, and deprotonation produces neutral radical intermediates that are similar to those produced by HO·-mediated H-abstraction reactions. The detailed mechanisms are presented for the four different DNA bases.
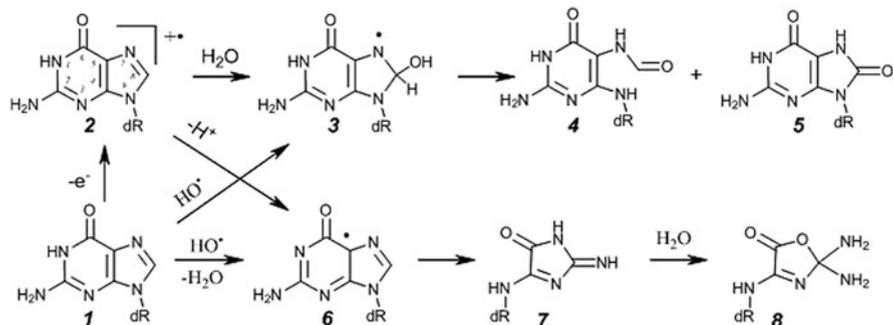
Fig. 8.1 HO˙ and one-electron oxidation mediated decomposition of guanine

Purine Bases

Hydration of the guanine radical cation *2* (Fig. 8.1) produced by one-electron oxidation of guanine *1* occurs at the C8 position producing radical *3* that is also generated through addition of HO˙ at C8. Radical *3* further decomposes to produce 8-oxodGuo (*5*) under oxidizing conditions or 2,6-diamino-4-hydroxy-5-formamido-pyrimidine (*4*) under reducing conditions. The neutral radical *6* produced through either deprotonation of guanine radical cation *2* or dehydration of HO˙ adduct is at the origin through a cascade of events of 2-amino-5-[(2-deoxy-ß-D-*erythro*-pentofuranosyl)amino]-4*H*-imidazol-4-one *7*. The latter nucleoside undergoes spontaneous hydrolysis to produce 2,2-diamino-4-[(2-deoxy-ß- D-*erythro*-pentofuranosyl) amino]-5-(2*H*)-oxazolone (*8*). It should also be noted that guanine is a critical target for one-electron oxidation of DNA because among the DNA constituents this base has the lowest ionization potential and because efficient charge transfer could occur in double-stranded (ds) DNA. Thus 8-oxodGuo *5* is the major lesion in dsDNA following exposure to UV or visible light is the presence of specific photosensitizers. After absorption of a photon, this class of molecules reaches an excited state oxidizing enough to abstract one electron from DNA. Ionization of DNA by high-intensity UV laser pulses triggers the same chemistry. A similar decomposition pathway is observed for adenine, except that HO˙ addition occurs only on C8 generating formamidopryrimidine and 8-oxo derivatives as reported for guanine. Another difference with guanine is that deprotonation of radical cation occurs on the exocyclic amino group generating deaminated hypoxanthine.

Pyrimidine Bases

As reported for purine bases, the decomposition of pyrimidine radical cation (Fig. 8.2) could occur by two competitive decomposition pathways: hydration or deprotonation. Deptrotonation of thymine radical cation (*10*) generated upon one-electron oxidation of the thymine base (*9*) occurs on the methyl group generating the 5-methylyl-2′-deoxyuridine radical *11* that, following reaction with molecular oxygen, produces 5-hydroperoxymethyl-2′-deoxyuridine *12*. That
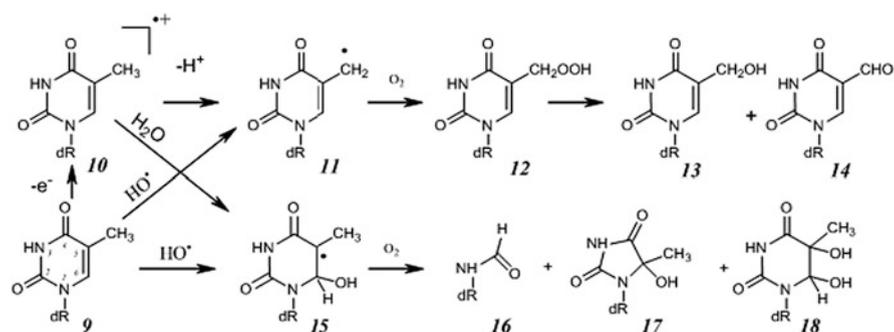
Fig. 8.2 HO˙ and one-electron oxidation mediated decomposition of thymine

unstable hydroperoxide further decomposes to produce 5-hydroxymethyl-2′-deoxyuridine (*13*) and 5-formyl-2′-deoxyuridine (*14*). Hydration of the thymine radical cation (*10*) occurs mostly onto C5–C6 double bound, producing mainly the 5-yle radical *15*. Following reaction with molecular oxygen, the generated unstable hydroperoxide radical further decomposes to produce mostly thymidine glycols *18* and to a minor extent formyl amine *16* and hydantoin *17*. These thymidine decomposition products (*16*, *17*, *18*) were also found to be major radiation-induced decomposition products of aqueous aerated solution of thymidine. Under these conditions their formation is explained by addition of HO˙ to the C5–C6 double bound to produce radical *15*, which is at the origin of products *16*, *17*, and *18*. In addition, 5-hydroxymethyl-2′-deoxyuridine (*13*) and 5-formyl-2′-deoxyuridine (*14*) are also produced under these conditions, albeit with a relatively lower yield. Their presence is explained by the formation of intermediate *11* by HO˙-mediated H-abstraction, which mostly occurs onto the methyl group of thymine. For cytosine, deprotonation of radical cation occurs (as reported for adenine) on the exocyclic amino group, generating uracil. HO˙ addition takes place onto the C5–C6 double bound, as reported for thymine, and generates formylamine, non-methylated hydantoin, and cytosine glycols. The latter oxidation products are unstable and could decompose through either dehydration or deamination.

8.2.1.3 Complex Lesions

During the last two decade several works have reported the formation of so-called complex DNA lesions that involves at least modification of two DNA constituents [4]. A common feature of these complex DNA lesions is that they are generated by a single oxidation event. Thus, these modifications are different to the LMDS, locally multiple damage sites, which are specifically produced by ionizing radiation through multi-ionization processes produced in the track of the particle. It has been shown that complex DNA lesions arose from two distinct mechanisms. The first one involves the reaction of an initially produced radical that is able to react with surrounding DNA bases creating cross-links or tandem lesions consisting of two adjacent DNA modifications. The second process is mediated by the initial

formation of a deoxyribose decomposition product containing a reactive aldehyde moiety that is thus able to react with a surrounding nucleophilic base. DNA–protein cross-links have also unfortunately been poorly studied to date.

Tandem DNA Lesions

Intra-strand cross-links between a purine and a vicinal pyrimidine are generated in double-stranded DNA mostly in the absence of oxygen. Formation of such tandem lesions (Fig. 8.3) in a purine–pyrimidine sequence context (*19*) involves first the formation of a pyrimidine radical located onto the methyl group of thymine (*20*). In the absence of oxygen, this radical can bind to the C8 position of a vicinal purine base, generating an intra-strand cross-link *21* [5]. Interestingly, it has been demonstrated that such a reaction is facilitated when the purine base is located at the 5′-position to the pyrimidine base. Formation of these intra-strand cross-links is significantly reduced in the presence of oxygen, since pyrimidine carbon-centered radicals are known to react efficiently with O_2. Interestingly, it has recently been demonstrated that the peroxyl radicals thus generated could be also at the origin of formation of tandem DNA lesions [6]. Indeed, pyrimidine peroxyl radicals *22* are able to add to the C8 position of and adjacent purine (guanine or adenine) base [7], creating an unstable endoperoxide *23* [8] that, following decomposition, generates an 8-oxodGuo containing tandem DNA lesion (*24*). The fact that, when involved in tandem lesions, 8-oxodGuo is less efficiently repaired by DNA glycosylases highlights the biological importance of such tandem DNA lesions.

Complex Lesions Arising from Sugar Decomposition Products

HO˙ is also able to react with the sugar, resulting in single strand breaks. However, recent results have reported that decomposition of the ribose could generate reactive aldehydic derivatives that could further react with nucleophilic DNA bases [9]. An example is provided by the formation of cytosine adducts (Fig. 8.4) generated through a cascade of reactions, the first one involving H-abstraction at the C4-position of the deoxyribose moiety in dsDNA [10]. Decomposition of the C4′ radical *25* produces an abasic site *26* that following a ß-elimination produces an SSB together with a reactive conjugated aldehyde *27*. Nucleophilic addition of a vicinal (mostly on the complementary strand) cytosine base *28* is at the origin of the identified cytosine adducts *29* [11] that may exist in a ring close form (*30*) depending on their conformation. Other similar mechanisms involving the initial formation of reactive aldehydes have been also reported.

8.2.2   Alkylated Bases

A large number of DNA adducts, usually referred to as alkylated bases, have been identified (Fig. 8.5). Their formation is explained by the nucleophilicity of the heteroatoms of the base moieties of DNA. For instance, methyl methanesulfonate (MMS) is able to methylate DNA bases mostly on N7 of 2′-deoxyguanosine *31* (that may decompose through depurination) and N3 of 2′-deoxyadenosine. Bulky
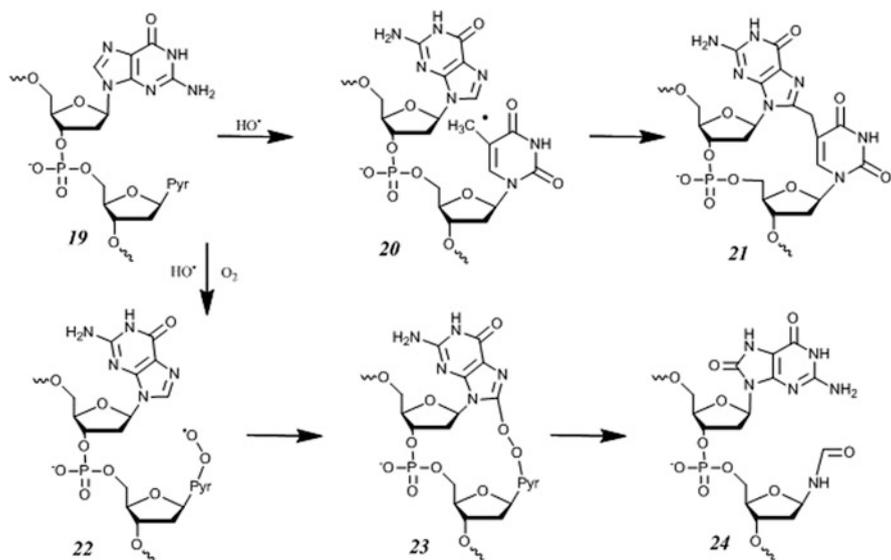
Fig. 8.3 Mechanisms of formation of tandem DNA lesions
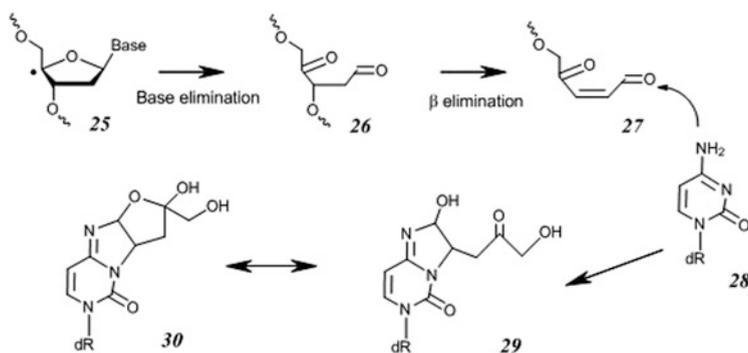
Fig. 8.4 Formation of a complex cytosine modification following C4′ H-abstraction

adducts could be generated with for example by benzo[a]pyrene (B(a)P) a well-known genotoxic polycyclic aromatic hydrocarbon [12]. In this case, formation of the DNA adducts requires metabolization of B(a)P into a diol epoxide derivative BPDE (Benzo Pyrene Diol Epoxide) that following nucleophilic addition of DNA base generates mostly guanine adducts **32**. More complex DNA lesions involving two adjacent DNA bases could be also generated by cis-platin derivatives [13]. In that case the N7 amino group of two adjacent guanine bases reacts with the drug and generates mostly intra-strand cross-links **33**. Formation of adducts between two adjacent A and G bases or two G located onto the two complementary strand is also possible.
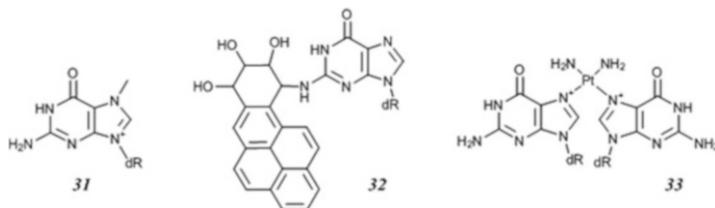
Fig. 8.5 Examples of alkylated DNA bases

8.2.3 UV-Induced Photoproducts

It is well documented that DNA is the main critical cellular target for the carcino-
genic effects UV radiation. Discovered in the late 1950s, bipyrimidine photo-
products, and in particular cyclobutane pyrimidine dimers, were amongst the first
identified classes of DNA damage. Nowadays, a large amount of information on the
basic photochemistry of DNA [14], either in solution or in cells, has been
accumulated for such damage [15]. Following excitation with UVC (200–
280 nm) and to a lower extent UVB (280–320 nm), formation of two types of
bipyrimidine photoproducts has been reported (Fig. 8.6). The first class involves
formation of cyclobutane pyrimidine dimers (CPDs, *34*) that are generated through
a [2 + 2] cycloaddition reaction between the C5–C6 double bonds of two adjacent
pyrimidine bases. Thus, according to the sequence (TT, TC, CT, CC), four possible
combinations of the pyrimidine bases could lead to the formation of CPDs. In
dsDNA, the two bases in the cyclobutane dimer are in parallel orientation and
located on the same side with respect to the cyclobutane ring, generating so-called a
cis,syn cyclobutane dimer. In addition to CPDs, 6-4 photoproducts *35* (64PPs)
could be also generated. These dimers are produced by a [2 + 2] cycloaddition
between the C5–C6 double bond of the 5′-end base and the C4 carbonyl group of a
3′-end thymine or cytosine when this one is in an exocyclic imine tautomeric form.
Thus, as reported for CPDs, four possible combinations of products are possible.
The generated intermediates, either oxetane for thymine or azetidine for cytosine,
decompose into the final 64PPs which exhibit a 3′-end pyrimidone ring and where
the C4 substituent of the 3′-end base has migrated to the C5 position of the 5′-end
one. It should also be mentioned that the 64PPs may be converted into their Dewar
valence isomer *36* following UVA excitation, for example, when simulated sunlight
is used.

8.3 Methods for the Quantification of DNA Bases

The frequency of modified bases is low in DNA. In cells it is below one modification
per millions of normal bases and sometimes even one or two orders of magnitude
lower. The frequency of damage is much higher when experiments are performed
in isolated DNA, but real biologically relevant data are only obtained at
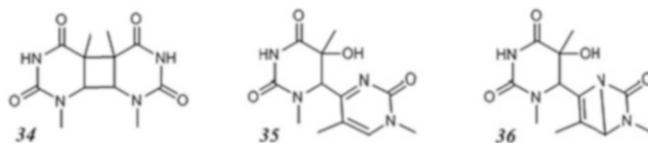
Fig. 8.6 Structure of CPDs (*34*) 64PPs (*35*) and its dewar isomer DEW (*36*) thymidine dimers

limited degradation yield in order to prevent secondary reactions to occur. Because of this low modification yield, one cannot envision to analyze DNA as a whole polymer and detect changes in molecular weight. For quantification of damage, DNA is hydrolyzed and modified bases are therefore detected as abnormal monomers. The challenge encountered for the quantification of modified bases in DNA is thus first a necessity of sensitive detection, because trace amounts have to be detected. In addition, the targeted compounds are mixed with the matrix of normal bases present in much larger amounts. High specificity is thus a second requisite. Mass spectrometry is well suited for both aspects. This technique possesses an intrinsic good sensitivity and can specifically target specific molecules on the basis of their mass spectrometry features. Two main approaches have been used in the two last decades for this purpose: gas chromatography combined with mass spectrometry and high-performance liquid chromatography associated with tandem mass spectrometry.

8.3.1 DNA Hydrolysis

As explained above, the first step in the analysis of modified DNA bases is the hydrolysis of the biopolymer into a mixture of monomers (Fig. 8.7). Of course, DNA has first to be extracted when cellular samples are analyzed. Several treatments may be applied to the subsequent release of the modified bases. Historically, the first used procedure was acidic treatment at high temperature, for example, 96 % formic acid at 90 °C for 30 min. Under these conditions, the phosphodiester backbone is cleaved, but more importantly, the N-glycosidic bond between the 2-deoxyribose moiety and the base is hydrolyzed. A limitation of this approach is that several modified bases are unstable under the strong hydrolysis conditions. Yet, hot acidic hydrolysis is suitable for numerous lesions including UV-induced cyclobutane pyrimidine dimers and some oxidized purines and pyrimidines. In contrast, limitations were documented for formamidopyrimidine derivatives of purines (*4*) and (6-4) photoproducts (*35*).

A widely developed milder alternative to acidic hydrolysis relies on enzymatic treatments. By combining specific enzymes, lesions are released as nucleosides, namely, the base linked to the 2-deoxyribose unit. First, endonucleases cleave the long pieces of genomic DNA into shorter fragments. Then, the DNA strands are incubated with phosphodiesterases. These enzymes cleave the phosphodiester bonds between the phosphate groups and the sugar units, thereby sequentially
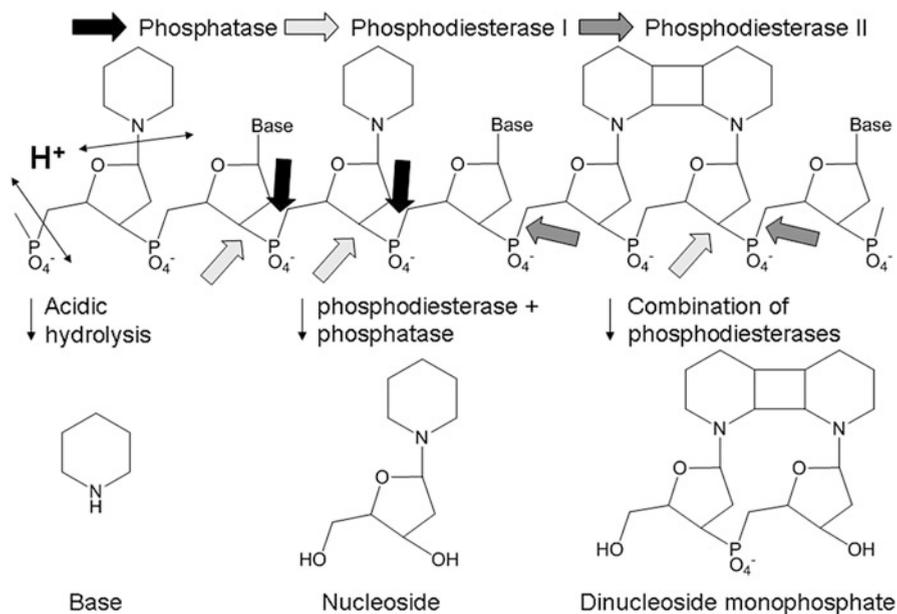
Fig. 8.7 Different procedures for the hydrolysis of DNA. Monomeric base damages are released as base and nucleoside following acidic and enzymatic hydrolysis, respectively. Dimeric lesions are enzymatically hydrolyzed into dinucleoside monophosphates

releasing monomeric nucleotides. Phosphatases are added to remove pending phosphate groups and yield a mixture of nucleosides. A noticeable exception to this scheme is related to DNA lesions involving formation of covalent bonds between two adjacent bases. In that case, the result of the hydrolysis is a dinucleoside monophosphate because the phosphodiester link between the modified nucleotides is not a substrate for the enzymes. As an alternative to the complete hydrolysis of the DNA polymer, a few groups have proposed to use repair enzymes in order to release the modified bases. Measurements can also be made in biological fluids such as urine, with the idea that the detected modified bases arise from DNA repair in tissues.

8.3.2 GC-MS

Association of chromatographic separation with specific mass spectrometry detection is an interesting approach for sensitive analysis and in particular for traces of DNA damage in the overwhelming matrix of normal bases. Chromatography involves a mobile phase containing the analytes and mass spectrometers are operated under vacuum. Therefore, the two are not easily compatible because the mobile phase may be too difficult to pump out from the system. This is less a

problem for gas chromatography where the mobile phase is a gas and does not introduce too much material in the mass spectrometer. Consequently, GC-MS has been one of the first efficient coupling systems and was applied to the quantification of modified bases in DNA. Although nucleosides can be analyzed, this approach is mostly performed following acidic hydrolysis of DNA into bases. However, bases are not volatile and have to be modified before analysis in a so-called derivation step. Derivation aims at blocking the heteroatoms of DNA bases and making them volatile. An additional requirement for accurate GS-MS analyses is linked to the fact that the derivation step is not quantitative and that the injection in the gas phase is not highly reproducible. To avoid possible lack of reproducibility, a common strategy is the use of isotopically labeled internal standard. In analytical chemistry, internal standards are compounds similar to the targeted analyte which are added to the samples. The output of the analysis is the ratio between the responses to the analyte and the internal standard, thus limiting the variations. When mass spectrometry is used as detection, the internal standard can be the targeted analyte itself if some of the atoms are replaced by their heavier isotopes (^{2}H, ^{13}C, ^{15}N).

A widely used protocol for derivation involves silylation, which converts bases into their trimethylsilyl or *tert*-butyl dimethylsilyl derivatives (Fig. 8.8). This reaction is performed with chlorinated silylation agents at high temperature in organic solvents. The samples are then analyzed by GC coupled to a quadrupolar mass spectrometer and using electron impact ionization (EIMS). In the vast majority of the systems, the energy of the electrons is set at 70 eV and positive ions are monitored. Because a specific detection is required to detect traces of modified bases in the matrix, only the ions produced by the analytes of interest are monitored. This mode is referred to as ion monitoring. The monitored ions are either molecular ions ($M^{\bullet+}$) or more frequently some of their fragments. This approach has been widely used in the 1990s for the quantification of oxidized bases [16, 17]. Data were gathered on both isolated and cellular DNA. Unfortunately, concern was raised on the measurement of 8-oxo-7,8-dihydroguanine **5** in cellular samples [18, 19]. This oxidized base was also quantified by HPLC combined with electrochemical detection and large discrepancies were observed between the results obtained by the two methods with a general trend to a higher level determined by GC-EIMS. After intensive discussion and inter-laboratory comparisons, it was shown that, in GCEIMS analysis, the derivation step induced the oxidation of the samples [20–22]. The yield of this reaction is low but sufficient to interfere with the trace amounts of oxidized bases to be detected. Several strategies such as degassing, adding antioxidants, or removing normal bases by pre-purification were proposed. They were either not efficient at totally preventing the spurious oxidation or made the analyses much more time-consuming. GC-EIMS analysis of oxidized bases is now rarely used. GC-EIMS analysis of hydrolyzed DNA following silylation has also been proposed for thymine cyclobutane dimer **34**, for which no artifact is expected [23].

In contrast with the case described above, the derivation step required for GC-MS can be an opportunity for the introduction of chemical functions allowing specific detection with increased sensitivity. This is in particular the case for GC associated
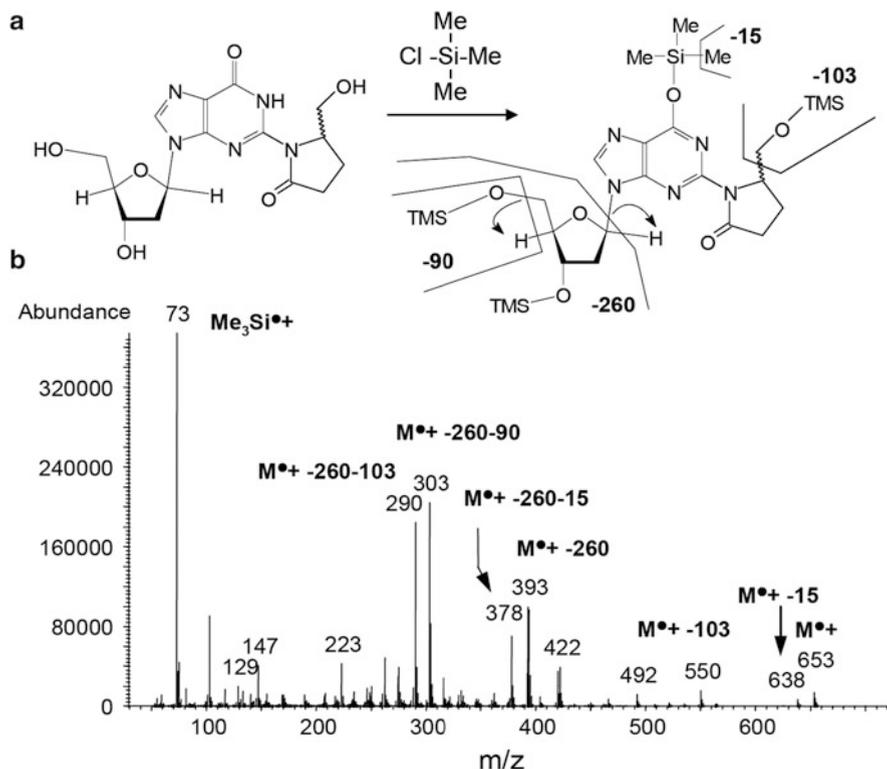
Fig. 8.8 GS-MS characterization of the trimethylsilylated derivative of the adduct between 2′-deoxyguanosine and 4,5-dioxovaleric acid. (**a**) Silylation reaction using trimethylsilyl chloride and (**b**) GC-MS mass spectrum recoded after ionization by electron impact at 70 eV [24]

with a mass spectrometer operating with negative chemically induced ionization (NCIMS). In this ionization technique, the source of the mass spectrometer contains as a reactive gas like methane, which is subjected to the 70 eV electron beam. The resulting ions then gently ionize the analytes eluting from the capillary column of the GC with limited fragmentation. When derivation is performed with penta-fluorobenzyl bromide, the group added to the bases is highly electro-attractive and easily ionizable into negative ions. The resulting detection is thus very sensitive, at least one order of magnitude larger than in the positive mode after silylation. This approach has been used by several groups to quantify modified bases, and in particular DNA adducts, with high sensitivity. First examples are adducts resulting from exposure to chemicals such as ethylene oxide [25–27]. This approach also made it possible to quantify etheno bases identified as DNA lesions induced upon exposure to vinyl chloride and possibly associated with endogenous lipid peroxidation [28–31]. GC-NCIMS has also been applied to the detection of the adduct between guanine and malonaldehyde, a by-product of the oxidation of membrane lipids [32, 33]. Application to oxidized bases has been also reported [34]. More

recently, the GCNCIMS assay was applied to the detection of oxidation products of 2-deoxyribose [35].

8.3.3 HPLC-MS/MS

Coupling mass spectrometry to liquid chromatography has been a challenge for years because it requires making compatible spectrometers operated under vacuum with chromatographic systems using hundreds of microliters of mobile phase per minute. A breakthrough in that respect was the development of electrospray ionization which makes it possible to eliminate the solvent, transfer the analytes in the gas phase and ionize them at atmospheric pressure. Ionization and desolvation take place simultaneously. The outlet of the HPLC column is transformed into a spray placed in an electric field (4–5 kV) which induces a cascade of fragmentation of the droplets. Ultimately, the analytes are released in the gas phase as either protonated $[M+H]^+$ or deprotonated $[M-H]^-$ ions depending on the sign of the electric field. The ions are then transferred into the mass spectrometer operated under high vacuum. For the detection of traces of modified DNA bases, most reports use tandem mass spectrometry with triple quadrupole mass spectrometer. To achieve optimal sensitivity, analyses are performed in the so-called ion monitoring mode. The first quadrupole isolates the ions exhibiting the m/z ration corresponding to that of the pseudomolecular ion of the analyte. These ions are then directed into a second quadrupole containing traces of an inert gas. Collision between the ions and the molecules of gas leads to fragmentation of the analyte: this is collision-induced dissociation (CID, see Chap. 6). The third quadrupole is used to collect fragments known to be specific of the targeted analyte. Altogether, this type of detection offers two levels of specificity and exhibits a very low noise. Detection limits for this type of analyses can be in the femtomolar range. An additional parameter for the identification of analyte is the chromatographic retention time. HPLC-MS/MS is also a powerful for characterization tool since CID occurs at specific positions of the molecule which are dependent on the chemical structure of the analytes, including short pieces of DNA (see Chap. 6).

HPLC-MS/MS is nowadays the most widely used technique for the quantification of modified bases in DNA [36, 37]. It does not always exhibit the same sensitivity as the older 32P-post-labeling approach but is much more specific and quantitative. HPLC-MS/MS has been applied by numerous groups to the quantification of adducts and oxidized bases. Isotopic dilution is sometimes used. However, this is much less mandatory than in GC-MS when analyses are performed directly after hydrolysis without further treatment of the sample. Isotopic dilution is in contrast absolutely required when pre-purification steps are added. In most cases, DNA lesions are quantified as modified nucleosides using positive electrospray ionization in order to take advantage of weakness of the N-glycosidic bond. Some modified bases are also analyzed as free bases. For instance, this strategy is used for products of alkylation of guanine at the N7 position (*31*) by 1,2-dihaloethanes [38],

ethylene oxide [39], 4-hydroxyestradiol [40, 41], 1,3-dutadienne [42], or sulfur mustard [43, 44]. These adducts are known to easily depurinate; namely, the bond between the base and the 2-deoxyribose moiety is spontaneously hydrolyzed. Another example is the formamidopyrimidine derivatives of purines (**4**) which are enzymatically released as four isomers of the sugar units and thus more easily quantified as bases [45].

The number of applications of HPLC-MS/MS to DNA adducts quantified as nucleosides is too long for an exhaustive presentation and new works are continuously published. The studied lesions include among other damage induced by chemicals used in the industry [38, 46, 47], pollutants [12, 48–54], or antitumor drugs [13, 55, 57]. Adducts produced by endogenously generated alkylating agents such as products of lipid peroxidation can also be quantified in this way [56, 57]. Another wide field of application includes lesions arising from oxidative and radical reaction. One of the first oxidized base quantified by HPLC-MS/MS was 8-oxo-7,8-dihydro-2′-deoxyguanosine **5** (Fig. 8.9) [58]. The assay was then extended to the measurement of series of oxidized nucleosides in enzymatically hydrolyzed DNA [45]. More recently, the assay was applied to cyclonucleosides [59]. Efficiency of the enzymes used for hydrolysis decreases due to the presence of these lesions, and the protocol had thus to be adjusted. It should be stressed that inconsistencies of unknown origin on the in vivo level of cyclonucleosides can be observed between different groups [59, 60]. In addition, HPLC-MS/MS has been used to quantify modified bases or nucleosides in urine [61–63]. These DNA derivatives are proposed to be products of DNA repair and indirectly reflect damage to the genome. In these techniques, a key step is the pre-purification aimed at removing salts, proteins, and the numerous compounds found in urine.

HPLC-MS/MS also permits the quantification of modified bases in forms bearing a phosphate group originating from the DNA backbone. These derivatives are easily detected as negative ions with usually reduced background of detection. Such approach may be applied to nucleotides but is mainly interesting for the quantification of DNA damage involving the formation of covalent links between adjacent bases. The resulting dimeric lesions are enzymatically released as dinucleoside monophosphates. This technique was very successful for the quantification of thymine–purine cross-links **21** in DNA exposed to ionizing radiation [5]. Another interesting application is the quantification of UV-induced photoproducts **34**, **35**, and **36** [64] which was used for the determination of the yield and distribution of DNA damage in isolated DNA, cell culture, and skin (Fig. 8.10) [15, 65–66]. The assay was also applied to spores and bacteria [67, 68]. Monitoring dinucleoside monophosphates has also been applied to tandem lesions where two adjacent bases are oxidized by a single radical event [69].

All the examples discussed above used HPLC-MS/MS as a very specific and sensitive detector, on the basis that the targeted modified DNA bases had been identified in preliminary experiments. Isolation of standards molecules is a requisite to set up the assay. Recently, HPLC-MS/MS was applied in slightly different ways to gather data on DNA damage on novel unidentified lesions. A nice example was the identification of a complex lesion induced upon exposure to ionizing radiation
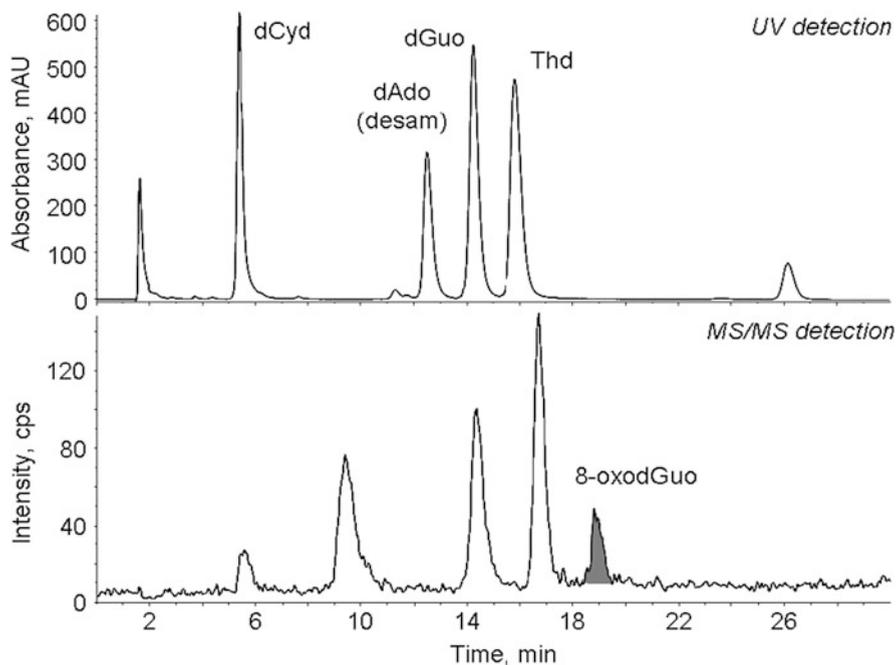
Fig. 8.9 Detection of 8-oxodGuo in the DNA of mouse brain. The amount of DNA was 15 µg and the level of damage was 1 8-oxodGuo per million normal bases. The *Upper trace* is the UV monitoring of the elution of normal nucleosides. 2′-Deoxyadenosin is deaminated as the result of an adenine deaminase contamination in one of the enzymes. The *lower trace* is the MS/MS detection of 8-oxodGuo (fragmentation *m/z* 284 → 168)

consisting in a cross-link between an oxidized deoxyribose and a cytosine **30** [10, 73]. This lesion, which could not have been found in irradiated solution of monomers, was detected on the basis of analyses performed in the so-called neutral loss mode. Under these conditions, the first and third quadrupoles scan mass-to-charge ratio with a constant difference. In order to identify novel modified nucleosides, this difference was set at 116, the part of 2-deoxyribose commonly lost by nucleosides upon CID. This loss of 116 mass units was also the basis of an approach referred to as adductomic [70]. It consists in analyzing hydrolyzed DNA with a series of fragmentation reactions of the type $[M+H]^{+} \rightarrow [M+H–116]^{+}$. Under these conditions, detection of modified nucleosides is performed without required knowledge of the structure. A similar approach was developed using an ion trap mass spectrometer [71].
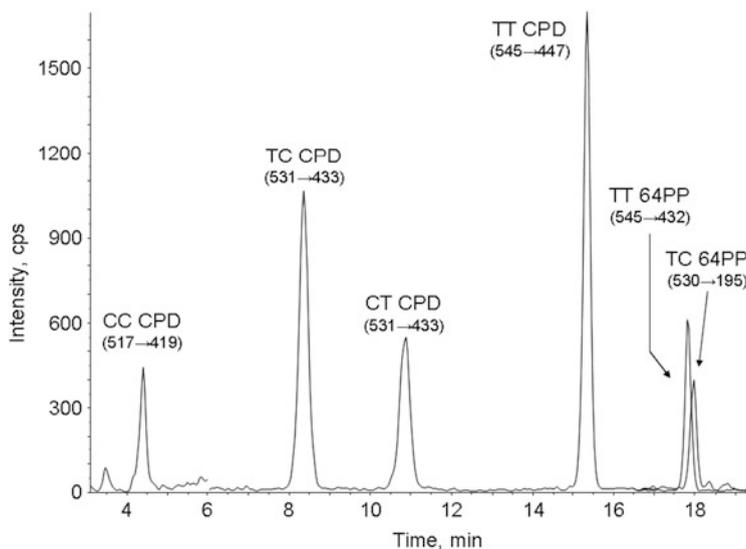
Fig. 8.10 HPLC-MS/MS detection of the main pyrimidine dimeric photoproducts in the DNA of a biopsy of human skin explant exposed to 1 MED, the dose leading to a sunburn in fair skin. TT, TC CT, and CC CPDs are thymine–thymine *34*, thymine–cytosine, cytosine–thymine, and cytosine–cytosine cyclobutane dimer. TT and TC 64PPs are (6-4) photoproducts formed at thymine–thymine (*35*) and thymine–cytosine sequences. Each photoproduct is detected using a specific fragmentation reaction shown in *brackets*

8.4 Conclusion and Outlook

Damage to DNA is a major issue in health and suspicion of carcinogenicity for various substances is discussed almost daily. Identification of DNA lesions is thus a very important issue. Through its application to the quantification of modified bases, mass spectrometry combined with chromatography is becoming a major tool in this field. Gas chromatography has been a first technique used for this purpose, but liquid chromatography combined to electrospray mass spectrometry is becoming the most widely applied approach. Recent woks have shown that direct identification of novel lesion in DNA is possible. The power of the HPLC-MS/MS technique allows us to study a very wide range of toxic chemicals. In that respect, identification of new lesions in model systems remains a key step before biological application. For all these assays, the ultimate goal is obviously application to studies in human samples. In the general population, this is often hampered by sensitivity issues. However, the constant development of more sensitive and more resolutive mass spectrometers (linear ion traps in triple quadrupole systems, orbitrap, quadrupople/Time-of-Flight detector) will increase in the future the predominant place of HPLC-mass spectrometry in the field of DNA damage. One may add that the strategy developed for DNA is also relevant for other biomolecules such as RNA and proteins although only few applications have been reported.

8.5   **Summary of Key Concepts**

General Concepts
- Association of mass spectrometry with chromatographic techniques provides sensitive and specific analytical tools.
- Gas chromatography is easily combined to mass spectrometry detection. The derivation of DNA bases aimed at making them volatile may be an opportunity to increase the sensitivity of the detection by adding adequate substituents.
- High-performance liquid chromatography can be associated with mass spectrometry electrospray ionization. Because it does not require derivation, it is increasingly used in the field of DNA damage.
- All these analytical developments require purified standards of the targeted compounds. This often represents the most time-consuming step of the setup of an assay.

Concepts Specific to Nucleic Acids
- DNA is a reactive molecule which can be modified by endogenous and exogenous agents.
- Modification of a base or induction of a break in DNA may lead to a mutation and ultimately cancer.
- A very large variety of modified DNA bases is produced by genotoxic agents.
- The frequency of these damages is very low in cells.
- Mass spectrometry associated with gas or liquid chromatography is a powerful approach for the detection of these lesions in DNA.

References

1. Ravanat J-L, Cadet J, Douki T (2012) Oxidatively generated DNA lesions as potential biomarkers of in vivo oxidative stress. Curr Mol Med 12:655–674
2. Cadet J, Douki T, Ravanat J-L (2010) Oxidatively generated base damage to cellular DNA. Free Radic Biol Med 49:9–21
3. Ravanat J-L, Martinez GR, Medeiros MGH, Di Mascio P, Cadet J (2006) Singlet oxygen oxidation of 2′-deoxyguanosine. Formation and mechanistic insights. Tetrahedron 62:10709–10715
4. Cadet J, Ravanat J-L, Taverna-Porro M, Menoni H, Angelov D (2012) Oxidatively generated complex DNA damage: tandem and clustered lesions. Cancer Lett 327:5–15
5. Bellon S, Ravanat J-L, Gasparutto D, Cadet J (2002) Cross-linked thymine-purine base tandem lesions: synthesis, characterization, and measurement in gamma-Irradiated isolated DNA. Chem Res Toxicol 15:598–606
6. Bergeron F, Auvré F, Radicella JP, Ravanat J-L (2010) HO˙ radicals induce an unexpected high proportion of tandem base lesions refractory to repair by DNA glycosylases. Proc Natl Acad Sci USA 107:5528–5533
7. Dupont C, Patel C, Ravanat JL, Dumont E (2013) Addressing the competitive formation of tandem DNA lesions by a nucleobase peroxyl radical: a DFT-D screening. Org Biomol Chem 11(18):3038–3045

8. Douki T, Riviere J, Cadet J (2002) DNA tandem lesions containing 8-oxo-7,8-dihydroguanine and formamido residues arise from intramolecular addition of thymine peroxyl radical to guanine. Chem Res Toxicol 15:445–454

9. Dedon PC (2008) The chemical toxicology of 2-deoxyribose oxidation in DNA. Chem Res Toxicol 21:206–219

10. Regulus P, Duroux B, Bayle P-A, Favier A, Cadet J, Ravanat J-L (2007) Oxidation of the sugar moiety of DNA by ionizing radiation or bleomycin could induce the formation of a cluster DNA lesion. Proc Natl Acad Sci USA 104:14032–14037

11. Hong IS, Carter KN, Sato K, Greenberg MM (2007) Characterization and mechanism of formation of tandem lesions in DNA by a nucleobase peroxyl radical. J Am Chem Soc 129: 4089–4098

12. Tarantini A, Maitre A, Lefebvre E, Marques M, Marie C, Ravanat J-L, Douki T (2009) Relative contribution of DNA strand breaks and DNA adducts to the genotoxicity of benzo [a]pyrene as a pure compound and in complex mixtures. Mutat Res 671:67–75

13. Baskerville-Abraham IM, Boysen G, Troutman JM, Mutlu E, Collins L, Dekrafft KE, Lin W, King C, Chaney SG, Swenberg JA (2009) Development of an ultraperformance liquid chromatography/mass spectrometry method to quantify cisplatin 1,2 intrastrand guanine-guanine adducts. Chem Res Toxicol 22:905–912

14. Cadet J, Mouret S, Ravanat JL, Douki T (2012) Photoinduced damage to cellular DNA: Direct and photosensitized reactions. Photochem Photobiol 88:1048–1065

15. Mouret S, Baudouin C, Charveron M, Favier A, Cadet J, Douki T (2006) Cyclobutane pyrimidine dimers are predominant DNA lesions in whole human skin exposed to UVA radiation. Proc Natl Acad Sci USA 103:13765–13770

16. Dizdaroglu M (1994) Chemical determination of oxidative DNA-damage by gas-chromatography mass-spectrometry. Oxygen Radic Biol Syst D 234:3–16

17. Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H (2002) Free radical-induced damage to DNA: mechanisms and measurement. Free Radic Biol Med 32:1102–1115

18. Collins AR, Brown J, Bogdanov M, Cadet J, Cooke M, Douki T, Dunster C, Eakins J, Epe B, Evans M, Farmer P, Gedik CM, Halliwell B, Herbert K, Hofer T, Hutchinson R, Jenner A, Jones GDD, Kasai H, Kelly F, Lloret A, Loft S, Lunec J, McEwan M, Moller L, Olinski R, Podmore I, Poulsen H, Ravanat JL, Rees JF, Reetz F, Shertzer H, Spiegelhalder B, Turesky R, Tyrrell R, Vina J, Vinicombe D, Weimann A, de Wergifosse B, Wood SG (2000) Comparison of different methods of measuring 8-oxoguanine as a marker of oxidative DNA damage. Free Radic Res 32:333–341

19. Gedik CM, Collins A (2004) Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study. FASEB J 18:82

20. Cadet J, D'Ham C, Douki T, Pouget J-P, Ravanat J-L, Sauvaigo S (1998) Facts and artifacts in the measurement of oxidative base damage to DNA. Free Radic Res 29:541–550

21. Douki T, Delatour T, Bianchini F, Cadet J (1996) Observation and prevention of an artefactual formation of oxidized DNA bases and nucleosides in the GC-EIMS method. Carcinogenesis 17:347–352

22. Ravanat J-L, Turesky RJ, Gremaud E, Trudel LJ, Stadler RH (1995) Determination of 8oxoguanine in DNA by gas chromatography-mass spectrometry and HPLC-electrochemical detection: overestimation of the background level of the oxidized base by the gas chromatography-mass spectrometry assay. Chem Res Toxicol 8:1039–1045

23. Podmore ID, Cooke MS, Herbert KE, Lunec J (1996) Quantitative determination of cyclobutane thymine dimers in DNA by stable isotope-dilution mass spectrometry. Photochem Photobiol 64:310–315

24. Douki T, Onuki J, Medeiros MHG, Bechara EJH, Cadet J, Di Mascio P (1998) DNA alkylation by 4,5-dioxovaleric acid, the final oxidation product of 5-aminolevulinic acid. Chem Res Toxicol 11:150–157

25. Giese RW (1997) Detection of DNA adducts by electron capture mass spectrometry. Chem Res Toxicol 10:255–270

26. Ranasinghe A, Scheller N, Wu KY, Upton PB, Swenberg JA (1998) Application of gas chromatography electron capture negative chemical ionization high-resolution mass spectrometry for analysis of DNA and protein adducts. Chem Res Toxicol 11:520–526

27. Wu KY, Scheller N, Ranasinghe A, Yen TY, Sangaiah R, Giese R, Swenberg JA (1999) A gas chromatography/electron capture/negative chemical ionization high-resolution mass spectrometry method for analysis of endogenous and exogenous N7-(2-hydroxyethyl)guanine in rodents and its potential for human biological monitoring. Chem Res Toxicol 12:722–729

28. Ham AJL, Ranasinghe A, Morinello EJ, Nakamura J, Upton PB, Johnson F, Swenberg JA (1999) Immunoaffinity/gas chromatography/high-resolution mass spectrometry method for the detection of N-2,3-ethenoguanine. Chem Res Toxicol 12:1240–1246

29. Morinello EJ, Ham AJL, Ranasinghe A, Nakamura J, Upton PB, Swenberg JA (2002) Molecular dosimetry and repair of N(2),3-ethenoguanine in rats exposed to vinyl chloride. Cancer Res 62:5189–5195

30. Chen HJ, Lin TC, Hong CL, Chiang LC (2001) Analysis of 3, N(4)-ethenocytosine in DNA and in human urine by isotope dilution gas chromatography/negative ion chemical ionization/mass spectrometry. Chem Res Toxicol 14:1612–1619

31. Nair U, Bartsch H, Nair J (2007) Lipid peroxidation-induced DNA damage in cancer-prone inflammatory diseases: a review of published adduct types and levels in humans. Free Radic Biol Med 43:1109–1120

32. Chandhary AK, Nokubo M, Marnett LJ, Blair IA (1994) Analysis of the malondialdehyde-2′-deoxyguanosine adduct in rat-liver DNA by gas-chromatography electron-capture negative chemical-ionization mass-spectrometry. Biol Mass Spectrom 23:457–464

33. Rouzer CA, Chaudhary AK, Nokubo M, Ferguson DM, Reddy GR, Blair IA, Marnett LJ (1997) Analysis of the malondialdehyde-2′-deoxyguanosine adduct pyrimidopurinone in human leukocyte DNA by gas chromatography electron capture negative chemical ionization mass spectrometry. Chem Res Toxicol 10:181–188

34. Teixeira AJR, Ferreira MR, Vandijk WJ, Vandewerken G, Dejong A (1995) Analysis of 8-hydroxy-2′-deoxyguanosine in rat urine and liver DNA by stable-isotope dilution gas-chromatography mass-spectrometry. Anal Biochem 226:307–319

35. Chan W, Chen BZ, Wang LR, Taghizadeh K, Demott MS, Dedon PC (2010) Quantification of the 2-deoxyribonolactone and nucleoside 5′-aldehyde products of 2-deoxyribose oxidation in DNA and cells by isotope-dilution gas chromatography mass spectrometry: differential effects of gamma-radiation and Fe2+-EDTA. J Am Chem Soc 132:6145–6153

36. Singh R, Farmer PB (2006) Liquid chromatography-electrospray ionization-mass spectrometry: the future of DNA adduct detection. Carcinogenesis 27:178–196

37. Tretyakova N, Goggin M, Sangaraju D, Janis G (2012) Quantitation of DNA adducts by stable isotope dilution mass spectrometry. Chem Res Toxicol 25:2007–2035

38. Huang H, Jemal A, David C, Barker SA, Swenson DH, Means JC (1998) Analysis of DNA adduct, S-[2-(N-7-guanyl)ethyl]glutathione, by liquid chromatography mass spectrometry and liquid chromatography tandem mass spectrometry. Anal Biochem 265:139–150

39. Liao PC, Li CM, Hung CW, Chen SH (2001) Quantitative detection of N-7-(2-hydroxyethyl) guanine adducts in DNA using high-performance liquid chromatography/electrospray ionization tandem mass spectrometry. J Mass Spectrom 36:336–343

40. Saeed M, Rogan E, Fernandez SV, Sheriff F, Russo J, Cavalieri E (2007) Formation of depurinating N3Adenine and N7Guanine adducts by MCF-10F cells cultured in the presence of 4-hydroxyestradiol. Int J Cancer 120:1821–1824

41. Bransfield LA, Rennie A, Visvanathan K, Odwin SA, Kensler TW, Yager JD, Friesen MD, Groopman JD (2008) Formation of two novel estrogen guanine adducts and HPLC/MS detection of 4-hydroxyestradiol-N-7-guanine in human urine. Chem Res Toxicol 21:1622–1630

42. Goggin M, Anderson C, Park S, Swenberg J, Walker V, Tretyakova N (2008) Quantitative high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry analysis of the adenine-guanine cross-links of 1,2,3,4-diepoxybutane in tissues of butadiene-exposed B6C3F1 mice. Chem Res Toxicol 21:1163–1170

43. Batal M, Boudry I, Clery-Barraud C, Mouret S, Douki T (2013) Relative yields of monomeric and dimeric adducts induced by sulphur mustard in isolated and cellular DNA as determined by HPLC/tandem mass spectrometry. Toxicol Environ Chem 95:260–276

44. Wei YX, Yue LJ, Liu Q, Chen J, Xie JW (2011) A sensitive high performance liquid chromatography-positive electrospray tandem mass spectrometry method for N-7-[2-[(2-hydroxyethyl)thio]-ethyl]guanine determination. J Chromatogr B Anal Technol Biomed Life Sci 879:1707–1712

45. Frelon S, Douki T, Ravanat J-L, Tornabene C, Cadet J (2000) High performance liquid chromatography - tandem mass spectrometry measurement of radiation-induced base damage to isolated and cellular DNA. Chem Res Toxicol 13:1002–1010

46. Tompkins EM, Jones DJL, Lamb JH, Marsden DA, Farmer PB, Brown K (2008) Simultaneous detection of five different 2-hydroxyethyl-DNA adducts formed by ethylene oxide exposure, using a high-performance liquid chromatography/electrospray ionisation tandem mass spectrometry assay. Rapid Commun Mass Spectrom 22:19–28

47. Yang YN, Nikolic D, Swanson SM, van Breemen RB (2002) Quantitative determination of N-7-methyldeoxyguanosine and O-6-methyldeoxyguanosine in DNA by LC-UV-MS-MS. Anal Chem 74:5376–5382

48. Harsch A, Sayer JM, Jerina DM, Vouros F (2000) HPLC-MS/MS identification of positionally isomeric benzo[c]phenanthrene diol epoxide adducts in duplex DNA. Chem Res Toxicol 13:1342–1348

49. Marie C, Maitre A, Douki T, Gateau M, Tarantini A, Guiraud P, Favier A, Ravanat JL (2008) Influence of the metabolic properties of human cells on the kinetic of formation of the major benzo[a]pyrene DNA adducts. J Appl Toxicol 28:579–590

50. Randall KL, Argoti D, Paonessa JD, Ding Y, Oaks Z, Zhang YS, Vouros P (2010) An improved liquid chromatography-tandem mass spectrometry method for the quantification of 4-aminobiphenyl DNA adducts in urinary bladder cells and tissues. J Chromatogr A 1217:4135–4143

51. Zhang SM, Chen KM, Aliaga C, Sun YW, Lin JM, Sharma AK, Amin S, El-Bayoumy K (2011) Identification and quantification of DNA adducts in the oral tissues of mice treated with the environmental carcinogen dibenzo[a, l]pyrene by HPLC-MS/MS. Chem Res Toxicol 24:1297–1303

52. Beland FA, Churchwell MI, Von Tungeln LS, Chen SJ, Fu PP, Culp SJ, Schoket B, Gyorffy E, Minarovits J, Poirier MC, Bowman ED, Weston A, Doerge DR (2005) High-performance liquid chromatography electrospray ionization tandem mass spectrometry for the detection and quantitation of benzo[a]pyrene-DNA adducts. Chem Res Toxicol 18:1306–1315

53. Chen C, Ma XC, Malfatti MA, Krausz KW, Kimura S, Felton JS, Idle JR, Gonzalez FJ (2007) A comprehensive investigation of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) metabolism in the mouse using a multivariate data analysis approach. Chem Res Toxicol 20:531–542

54. da Costa GG, Singh R, Arlt VM, Mirza A, Richards M, Takamura-Enya T, Schmeiser HH, Farmer PB, Phillips DH (2009) Quantification of 3-nitrobenzanthrone-DNA adducts using online column-switching HPLC-electrospray tandem mass spectrometry. Chem Res Toxicol 22:1860–1868

55. Jones CR, Sabbioni G (2003) Identification of DNA adducts using HPLC/MS/MS following in vitro and in vivo experiments with arylamines and nitroarenes. Chem Res Toxicol 16:1251–1263

56. Beland FA, Churchwell MI, Doerge DR, Parkin DR, Malejka-Giganti D, Hewer A, Phillips DH, Carmichael PL, da Costa GG, Marques MM (2004) Electrospray ionization-tandem mass spectrometry and P-32-postlabeling analyses of tamoxifen-DNA adducts in humans. J Natl Cancer Inst 96:1099–1104

57. da Costa GG, Marques MM, Beland FA, Freeman JP, Churchwell MI, Doerge DR (2003) Quantification of tamoxifen DNA adducts using on-line sample preparation and HPLC-electrospray ionization tandem mass spectrometry. Chem Res Toxicol 16:357–366

58. Douki T, Odin F, Caillat S, Favier A, Cadet J (2004) Predominance of the 1, N-2-propano 2′-deoxyguanosine adduct among 4-hydroxy-2-nonenal-induced DNA lesions. Free Radic Biol Med 37:62–70

59. Loureiro APM, Marques SA, Garcia CCM, Di Mascio P, Medeiros MHG (2002) Development of an on-line liquid chromatography-electrospray tandem mass spectrometry assay to quantitatively determine 1, N-2-etheno-2′-deoxyguanosine in DNA. Chem Res Toxicol 15: 1302–1308

60. Ravanat JL, Duretz B, Guiller A, Douki T, Cadet J (1998) Isotope dilution high-performance liquid chromatography - electrospray tandem mass spectrometry assay for the measurement of 8-oxo-7,8-dihydro-2′-deoxyguanosine in biological samples. J Chromatogr B 715:349–356

61. Belmadoui N, Boussicault F, Guerra M, Ravanat JL, Chatgilialoglu C, Cadet J (2010) Radiation-induced formation of purine 5′,8-cyclonucleosides in isolated and cellular DNA: high stereospecificity and modulating effect of oxygen. Org Biomol Chem 8:3211–3219

62. Wang J, Yuan B, Guerrero C, Bahde R, Gupta S, Wang Y (2011) Quantification of oxidative DNA lesions in tissues of Long-Evans Cinnamon rats by capillary high-performance liquid chromatography-tandem mass spectrometry coupled with stable isotope-dilution method. Anal Chem 83:2201–2209

63. Harri M, Kasai H, Mori T, Tornaeus J, Savela K, Peltonen K (2007) Analysis of 8-hydroxy-2′-deoxyguanosine in urine using high-performance liquid chromatography-electrospray tandem mass spectrometry. J Chromatogr B 853:242–246

64. Malayappan B, Garrett TJ, Segal M, Leeuwenburgh C (2007) Urinary analysis of 8-oxoguanine, 8-oxoguanosine, fapy-guanine and 8-oxo-2′-deoxyguano sine by high-performance liquid chromatography-electro spray tandem mass spectrometry as a measure of oxidative stress. J Chromatogr A 1167:54–62

65. Weimann A, Belling D, Poulsen HE (2001) Measurement of 8-oxo-2′-deoxyguanosine and 8-oxo-2′-deoxyadenosine in DNA and human urine by high performance liquid chromatography-electrospray tandem mass spectrometry. Free Radic Biol Med 30:757–764

66. Douki T (2013) The variety of UV-induced pyrimidine dimeric photoproducts in DNA as shown by chromatographic quantification methods. Photochem Photobiol Sci 12(8): 1286–1302

67. Douki T, Cadet J (2001) Individual determination of the yield of the main-UV induced dimeric pyrimidine photoproducts in DNA suggests a high mutagenicity of CC photolesions. Biochemistry 40:2495–2501

68. Douki T, Court M, Cadet J (2000) Electrospray-mass spectrometry characterization and measurement of far-UV induced thymine photoproducts. J Photochem Photobiol B Biol 54: 145–154

69. Douki T, Court M, Sauvaigo S, Odin F, Cadet J (2000) Formation of the main UV-induced thymine dimeric lesions within isolated and cellular DNA as measured by high performance liquid chromatography-tandem mass spectrometry. J Biol Chem 275:11678–11685

70. Douki T, Setlow B, Setlow P (2005) Effects of the binding of α/β -type small, acid-soluble spore proteins on the photochemistry of DNA in spores of Bacillus subtilis and in vitro. Photochem Photobiol 81:163–169

71. Matallana-Surget S, Meador JA, Joux F, Douki T (2008) Effect of the GC content of DNA on the distribution of UVB-induced bipyrimidine photoproducts. Photochem Photobiol Sci 7: 794–801

72. Bourdat A-G, Douki T, Frelon S, Gasparutto D, Cadet J (2000) Tandem base lesions are generated by hydroxyl radicals within isolated DNA in aerated aqueous solution. J Am Chem Soc 122:4549–4556

73. Regulus P, Spessotto S, Gateau M, Cadet J, Favier A, Ravanat JL (2004) Detection of new radiation-induced DNA lesions by liquid chromatography coupled to tandem mass spectrometry. Rapid Commun Mass Spectrom 18:2223–2228

74. Singh R, Teichert F, Seidel A, Roach J, Cordell R, Cheng MK, Frank H, Steward WP, Manson MM, Farmer PB (2010) Development of a targeted adductomic method for the determination

of polycyclic aromatic hydrocarbon DNA adducts using online column-switching liquid chromatography/tandem mass spectrometry. Rapid Commun Mass Spectrom 24:2329–2340

75. Bessette EE, Goodenough AK, Langouet S, Yasa I, Kozekov ID, Spivack SD, Turesky RJ (2009) Screening for DNA adducts by data-dependent constant neutral loss-triple stage mass spectrometry with a linear quadrupole ion trap mass spectrometer. Anal Chem 81:809–819

Ligand Binding to Nucleic Acids

9

Jennifer S. Brodbelt and Zhe Xu

Abstract

Among the preeminent compounds that bind to DNA are numerous anticancer and antibacterial therapeutics. The development of new chemotherapeutics has accelerated the need for sensitive and versatile analytical techniques that are capable of characterizing DNA/ligand interactions including determination of binding stoichiometries, selectivities, and affinities. Electrospray ionization mass spectrometry (ESI-MS) has emerged as a useful technique for the analysis of complexes formed between DNA and small molecules due to its low sample consumption and fast analysis time. This chapter describes the exploration, optimization, and validation of ESI-MS methods for characterizing DNA–ligand interactions.

Keywords

DNA–ligand complexes • DNA-interactive agent • Intercalation • Minor groove binding • Quadruplex • Electrospray ionization • Collision-induced dissociation

Abbreviations

ESI   Electrospray ionization
CID   Collision-induced dissociation

J.S. Brodbelt (✉) • Z. Xu
Department of Chemistry, University of Texas, Austin, TX 78712, USA
e-mail: jbrodbelt@cm.utexas.edu

9.1   Introduction

The interaction of small molecules with DNA is the basis for many clinically important drugs and biophysical probes [1–12]. Continued interest in DNA-interactive agents focuses on developing a better understanding of the nature of their interactions with DNA and the biological consequences of DNA binding. These studies are ultimately directed towards the design of compounds with higher levels of selectivity for particular DNA sequences or DNA structures and increased utility in biology and medicine. In the context of understanding both fundamental aspects of DNA–ligand interactions and applications related to ligand screening and drug discovery, electrospray ionization (ESI) has proven to be one of the most versatile analytical methods owing to its ability to transfer both covalent and noncovalent complexes from a native solution environment to the gas phase for high sensitivity molecular analysis by mass spectrometry [13–19]. ESI is particularly well suited for analysis of DNA complexes for two reasons. First, ESI generates multi-charged ions, meaning that even large DNA–ligand complexes with high molecular weights may be observed in a much lower m/z range, thus making the analysis of DNA–ligand complexes feasible even on benchtop triple quadrupole and ion trap platforms. Second, ESI can be tuned to operate very efficiently in the negative polarity mode. This is beneficial for analysis of nucleic acids which have a phosphodiester backbone that is readily deprotonated to create negative charged ions. ESI-MS allows the determination of nucleic acid:ligand stoichiometries, evaluation of binding affinities, and investigation of sequence selectivities. Moreover, ESI-MS provides insight that is complementary to that obtained from other conventional methods, such as NMR and gel-based assays. The dual goals of unraveling the mode of action of many current drugs and developing new compounds with specific pharmacological activities motivate many of the applications of ESI-MS in the DNA/drug arena.

9.2   Overview of DNA-Interactive Ligands

An impressive array of DNA-interactive ligands have been isolated or synthesized in the past two decades, with some demonstrating sufficient pharmacological activities (generally cytotoxicities) to make them successful as drugs [1–12]. Although the specific mechanisms of action are not fully elucidated, some DNA-interactive compounds induce conformational changes to or cleavage of DNA, whereas others form covalent bonds to DNA in a way that interferes with DNA transcription. The three most prevalent categories of DNA-interactive ligands include those that bind via intercalation, those that bind in the minor groove, and those that form covalent bonds to nucleobases, the latter leading to mono-adducts or crosslinks (Fig. 9.1).

Some examples of the most notable DNA-interactive ligands are illustrated in Fig. 9.2. Intercalators typically contain one or more planar aromatic groups that insert between base pairs of DNA, often in G/C-rich stretches, thus causing

Fig. 9.1 Types of DNA–
ligand complexes
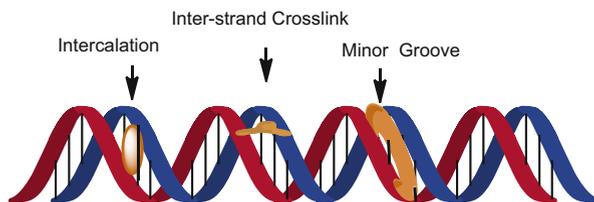
elongation and unwinding of DNA to accommodate the ligand [2, 3]. Some of the
more well-known intercalators include daunomycin, actinomycin-D, and echino-
mycin (Fig. 9.2). Some intercalators, notably daunomycin and actinomycin-D,
possess side chains that associate with the minor groove of the duplex after
intercalation of the planar chromophore. The ligands that bind to the curved
minor groove of DNA are frequently crescent-shaped molecules whose interactions
with DNA are mediated by hydrogen bond formation and van der Waals
interactions. Minor groove binding agents are more likely to interact with A/T-
rich stretches of DNA due to greater opportunities for hydrogen bonding
interactions. Characteristic minor groove ligands include distamycin, netropsin,
and Hoechst 33342.

Another important subclass of DNA-binding drugs exert their activities in a
metal-dependent fashion, based on coordination or recognition of a specific metal
ion [3]. The discovery of cisplatin ushered in a new era of DNA-interactive
anticancer agents based upon coordination chemistry [8]. Displacement of each
of the chloride moieties of cisplatin by water allows the platinum to coordinate a
basic site in DNA, favoring the cross-linking of adjacent guanine residues in one
strand of DNA. Other metal-mediated DNA ligands include the bleomycins, the
antitumor antibiotic, streptonigrin, which requires Zn^{2+} or Cu^{2+}, the antibacterial
fluoroquinolones, the aureolic acid antibiotics, mithramycin and chromomycin,
which bind in the presence of Mg^{2+}, the antitumor quinobenzoxazines, various
phenanthroline derivatives, and other platinum antitumor drugs (Fig. 9.2).

Ligands that selectively bind to unusual higher order DNA structures, such as
G-quadruplexes, have been found to have promising biological effects, as
exemplified by the inhibition of the cancer cell-associated enzyme telomerase by
perylene diimides [6, 7]**.** G-quadruplex DNA arises from the stacking of guanine
quartets commonly found at the overhangs of telomeric DNA at the termini of
linear chromosomes [7]. Ligands that stabilize the G-quartet motifs inhibit telo-
merase, an enzyme that regulates the lengths of telomeres and is implicated in
cell immortality [6, 7]. Telomeres play a critical role in protecting DNA from
degradation and fusion, thus ensuring the complete replication of DNA sequences
[7]. The telomere length shortens with each stage of DNA replication, ultimately
preventing replication of certain regions of coding DNA and leading to cell death
[7]. The length of the telomeres is modulated by telomerase, a reverse transcriptase
enzyme. Because telomerase is overexpressed in most tumor cells, the possibility of
regulating telomerase activity via the stabilization of G-quadruplex structures
has spurred interest in the development of quadruplex-interactive agents

a

Intercalator compounds

Daunomycin

Ellipticine

Actinomycin D

Echinomycin

Nogalamycin

b
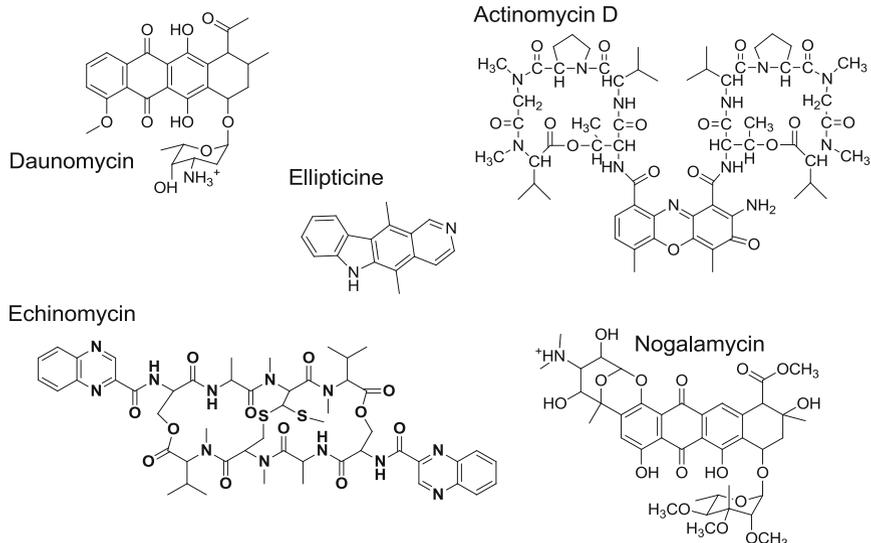
Minor groove binding compounds

4',6-Diamidino-2-phenylindole (DAPI)

Netropsin

Distamycin

Hoechst 3342

c

Crosslinking agents

Cisplatin

Psoralen

RH$_1$
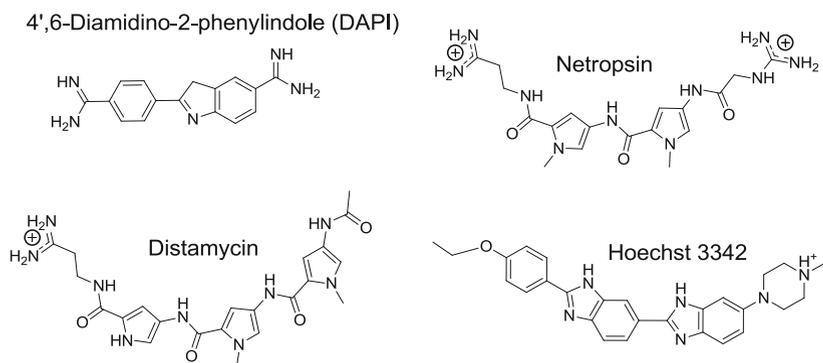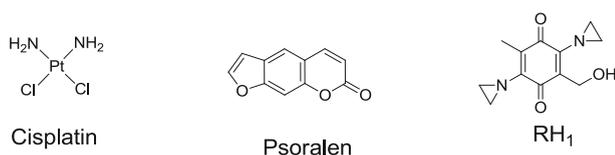
Fig. 9.2 Representative structures of DNA-interactive ligands: (**a**) intercalators, (**b**) minor groove binders, (**c**) cross-linking agents

[7]. Quadruplex-binding ligands include some porphyrin compounds, carbo-cyanines, perylene diimides, and ethidium bromide [7].

Some ligands form covalent bonds with reactive sites of DNA, creating stable adducts or crosslinks [8–10]. The general class of alkylating agents includes electrophilic compounds with high affinities for nucleophilic centers in nucleic acids. Numerous potential reaction sites for alkylation sites have been identified among the four nucleobases, with the N7 position of guanine and the N3 position of adenine being the most reactive. For example, mitomycin C preferentially alkylates guanine residues in the sequence 5′-CpG-3′, and duocarmycin and its analogs bind at the N^3 position of adenine [10]. Crosslinking agents represent a particularly important class of DNA-interactive agents because the formation of cross-links prevents DNA strand separation, leading to interruption of replication and transcription and inducing cell death. Such ligands include the nitrosoureas, mitomycin C, cisplatin and its derivatives, both nitrogen and sulfur mustards, and photoactivated psoralens.

9.3   Analytical Methods for Examining DNA–Ligand Interactions

An enormous variety of methods are currently used to characterize drug/DNA complexes, some providing thermodynamic information and others used for structural analysis [20–27]. The most popular method for measuring binding constants and assessing stoichiometries is isothermal titration calorimetry [20, 21]. Also used are spectrophotometric methods [23] involving variation of the drug concentration or mole fraction relative to the DNA to create Job plots for binding stoichiometries or Scatchard plots for binding constants. Thermal melting curves, a third option, give a cruder estimation of binding constants. Structural information is typically obtained by linear and circular dichroism methods, NMR, and X-ray crystallography [24–26]. The dichroism methods, based on polarized light spectroscopy, give information about the orientation of the bound drug and DNA, whereas the NMR spectroscopic methods can be used to determine the mode of binding and changes in DNA conformation upon binding. The X-ray crystallographic methods give three-dimensional structures of the drug–DNA complexes.

The most common methods for determination of drug binding sites involve footprinting techniques [26, 27] in which drug/DNA complexes are cleaved by enzymatic or chemical reactions prior to electrophoretic analysis. The bound drug blocks or enhances the cleavage of the DNA around the binding site, thus creating changes in the resulting gel-based fragment map relative to maps created by the DNA alone. For the footprinting methods, the DNA is often labeled by ^{32}P or fluorescent tags at the 3′ or 5′ ends to allow convenient detection and quantification after polyacrylamide gel electrophoresis. DNase I, a glycoprotein endonuclease that causes single strand nicks in the phosphodiester backbone, is the most commonly used enzymatic footprinting reagent [27]. Hydroxyl radical footprinting is appropriate for DNA ligands that do not induce major structural perturbations upon

binding to DNA and whose binding occludes access to the deoxyribose hydrogens, such as DNA minor groove ligands. Other reagents that are employed for probing drug–DNA ligand complexes include those that react at specific DNA bases. For example, diethylpyrocarbonate promotes cleavage at N7 position of adenine (and to a lesser extent guanine) via carboethoxylation of the purine ring, leading to ring opening between N7 and C8 [26, 27]. Dimethylsulfate methylates the N7 positions of guanines with little sequence specificity, rendering the 5-membered ring susceptible to cleavage and leading to strand breakage [27]. These reagents are useful for footprinting ligands that bind in the DNA major groove or which alter the conformation of the DNA when bound.

9.4   Electrospray Ionization Mass Spectrometry for Analyzing DNA–Ligand Complexes

The traditional methods of measuring selectivities, binding affinities, and structures of drug/DNA complexes summarized above have a rich history but also have some shortcomings in terms of sensitivity and sample consumption, in addition to limited compatibility with targeted screening methods. The design, development, and understanding of new DNA-interactive agents has accelerated the need for sensitive, versatile analytical methods that can be used to determine the nature and stoichiometry of drug/DNA binding, including evaluation of the selectivity of drugs for certain sequences or conformations of DNA, and the strengths of the binding interactions, along with techniques suitable for library-based screening. Electrospray ionization mass spectrometry (ESI-MS) has proven to be one of the most versatile methods for examining many of the issues related to drug/DNA (and drug/RNA) binding [13–19], as evidenced by numerous studies in the past two decades which have reported the application of ESI-MS for evaluating binding of drugs to DNA and RNA [28–75]. Several reports have reviewed the application of ESI-MS for studying the noncovalent binding of drugs to DNA [13–18] and RNA [13, 19] as well as drugs that bind to DNA via covalent interaction [20] and have summarized many of the milestones in this field. To date, ESI-MS has been used both for drugs that have been well studied by conventional methods as well as new emerging drug candidates for which quantities are limited and for which ESI-MS offers an efficient and sensitive screening option.

In the earliest studies by Gale et al., complexes formed between minor groove binders including distamycin, pentamidine, and Hoechst 33258 and a self-complementary 12-mer duplex were transferred to the gas phase by ESI [28, 29]. At low distamycin/DNA molar ratios, 1:1 complexes were detected, whereas 2:1 complexes were observed at higher molar ratios. The binding stoichiometries of the DNA complexes determined by ESI-MS were consistent with previous NMR reports. These seminal studies also established appropriate experimental conditions for the analysis of drug/DNA complexes by ESI-MS, including using low ESI voltages and heated capillary temperatures, as well as replacing common alkali metal buffers with the more volatile and ESI-friendly ammonium acetate buffer.
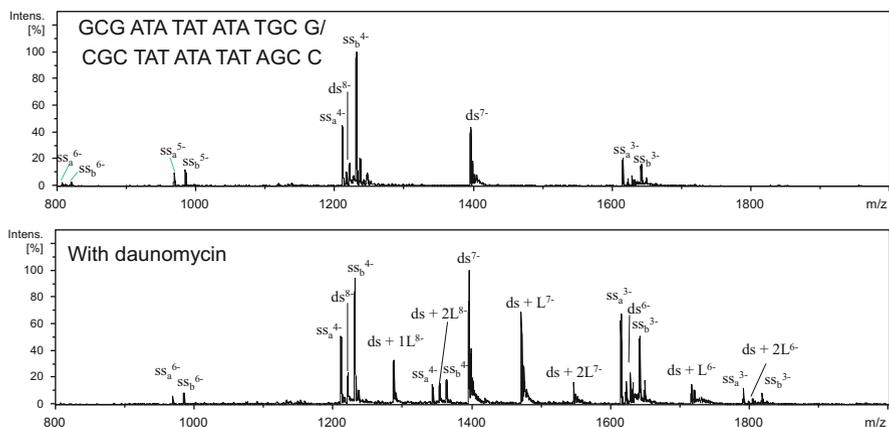
Fig. 9.3 ESI-mass spectra of AT-rich 16-bp duplex prior to and after incubation with daunomycin (L) in solution. *ds* duplex, *ss* single strand. The charge state is shown as a superscript. The DNA solution was prepared at 50 µM in 50 µL of 50 mM ammonium acetate. Daunomycin (structure shown in Fig. 9.2) was added at 100 µM, and the solution was allowed to equilibrate for 30 min. Electrospray ionization mass spectra were collected on a ThermoFisher LTQ linear ion trap mass spectrometer. Each solution was infused at 3 µL/min at a net DNA concentration of 10 µM. The ESI source was operated in the negative ion mode with an electrospray voltage of 3.5 kV and a heated capillary temperature of 90° C

Ammonium acetate is amenable to ESI-MS because of the greater lability of the ammonium ion over Na^+ and K^+, thus reducing the extent of adduction of counterions to the DNA ions. In another key study, Gabelica et al. [31] observed that 1:2 duplex:distamycin A complexes were the dominant products detected upon ESI-MS of an equimolar solution of a dodecamer oligonucleotide and distamycin A. However, exclusively 1:1 complexes were observed with netropsin and berenil and the same dodecamer. The fact that different stoichiometries were observed supported the hypothesis that the ESI process was capable of retaining specific information about complexation in solution.

Representative examples of ESI mass spectra for DNA duplex/ligand complexes are shown in Figs. 9.3 and 9.4 for solutions consisting of a duplex, either AT-rich d (GCGATATATATATGCG) or GC-rich d(CGCGCGCGCTACGCGC) with intercalator daunomycin. The ESI mass spectra display some ligand-free duplex ions and single strands, labeled as "ds" and "ss," as well as 1:1 and 1:2 duplex/ligand complexes labeled as ds + 1L and ds + 2L in each spectrum to designate the number of ligands bound to the duplex. These products are observed in the 8-, 7-, and 6- charge states, as denoted by the superscript values on each of the ion labels. Most significantly, the detection of these types of DNA/ligand complexes reflects the ability to transport desolvated noncovalent complexes into the gas phase for mass spectrometric analysis. The abundances of the duplex/ligand complexes relative to the ligand-free duplex are notably different in Figs. 9.3 and 9.4, with far more abundant duplex/ligand complexes observed for daunomycin with the
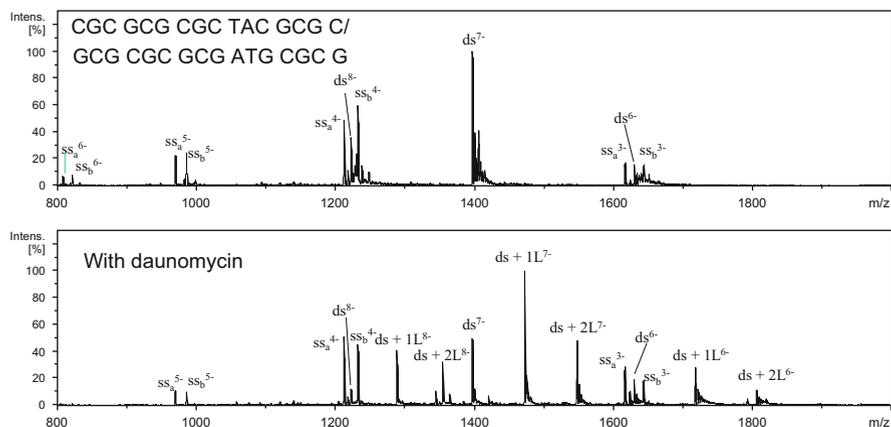
Fig. 9.4 ESI-mass spectra of GC-rich 16-bp duplex prior to and after incubation with daunomycin (L) in solution. *ds* duplex, *ss* single strand. The *charge state* is shown as a *superscript*. The DNA solution was prepared at 50 μM in 50 μL of 50 mM ammonium acetate. Daunomycin (the structure is shown in Fig. 9.2) was added at 100 μM, and the solution was allowed to equilibrate for 30 min. Electrospray ionization mass spectra were collected on a ThermoFisher LTQ linear ion trap mass spectrometer. Each solution was infused at 3 μL/min at a net DNA concentration of 10 μM. The ESI source was operated in the negative ion mode with an electrospray voltage of 3.5 kV and a heated capillary temperature of 90° C

GC-rich duplex in Fig. 9.4 and significantly more abundant complexes than observed for the AT-rich duplex in Fig. 9.3. Based on the simple visual inspection of the spectra, the affinity for GC-rich or AT-rich duplexes is readily evaluated and consistent with the known preferences of intercalating and minor groove binding ligands in solution. Whether assessing sequence selectivities of different ligands or comparing relative binding affinities of different ligands for individual duplexes, ESI-MS provides fast feedback with minimal sample consumption, thus making it an appealing screening tool. ESI-MS studies are particularly useful when undertaken in parallel with other conventional solution-based analytical methods or with cytotoxicity assays to provide basic correlations of relative binding affinities with biological activities.

9.5   Survey of ESI-MS Binding Studies

Since the capability of ESI for transferring noncovalent complexes to the gas phase was reported, a number of research groups have exploited the use of ESI-MS for evaluating numerous types of DNA/ligand complexes. Over the past decade, ESI-MS has been used to study complexes involving acridines, cationic porphyrins, cryptolepine alkaloids, protoberbine alkaloids, benzopyridoindole and benzo-pyridoquinoxaline compounds, and acridine- and naphthalene-derived macrocyclic
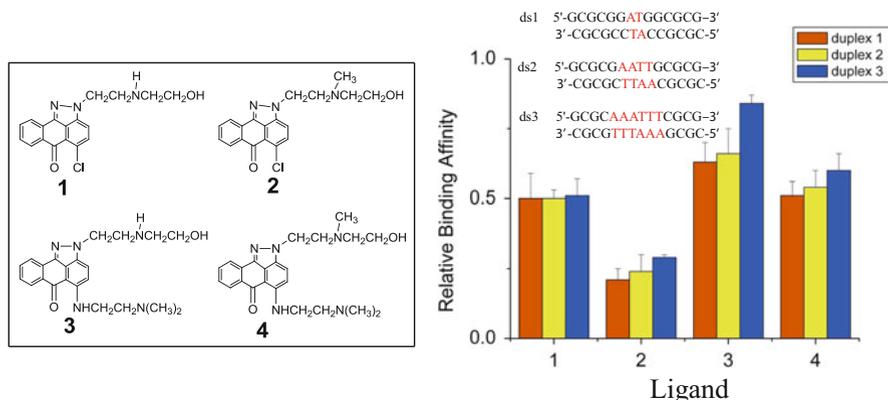
Fig. 9.5 Four anthrapyrazole ligands and their relative binding affinities for three duplexes of variable AT content. For more details, see [39]

bis-intercalators, among others [32–69]. Several representative examples are summarized in the following sections.

The ability of ESI-MS to monitor subtle differences in sequence selectivity and binding affinity is illustrated in Fig. 9.5 for a series of four anthrapyrazole ligands [44]. These intercalator compounds are abbreviated analogs of anthracyclines, anticancer intercalating agents. Shown in Fig. 9.5 in bar graph format are the relative binding affinities across the series of anthrapyrazoles for three duplexes of varying AT/GC content. The relative binding affinity is actually a measure of the fraction of DNA bound to a ligand relative to the free DNA duplex, and it is calculated by summing all the peaks areas of DNA–ligand complexes observed in the ESI mass spectra and dividing by the sum of the areas of all DNA–ligand complexes and free DNA. A higher value for the fraction of bound DNA reflects a higher ligand binding affinity. Because ESI has the potential to produce ions of different charge states, the spectra include DNA ions and DNA–ligand complexes in more than one charge state, all of which contribute to the determination of relative binding affinities. As shown in Fig. 9.5, anthrapyrazole **2** exhibits the lowest DNA binding affinity as well as relatively low sequence selectivity. Anthrapyrazole **3** shows the greatest DNA binding affinity with a slight enhancement for interaction with the most AT-rich duplex. The simple trends in Fig. 9.5 convey the impact of substituting a methyl group for a hydrogen atom on the amine side chain and replacement of the chlorine atom by a second amine side chain.

Some of the most systematic studies of DNA–ligand complexation have been reported by DePauw, Gabelica, and Rosu in which both positive and negative ESI as well as collision-induced dissociation were used to characterize complexes containing a range of minor groove binding agents (Hoechst 33258 and 33342, netropsin, DAPI) and intercalators (daunomcyin, doxorubicin, actinomycin D, ethidium, cryptolepine, neocryptolepine, m-amsacrine, proflavine, ellipticine, mitoxantrone) [31, 37, 47]. These studies provide excellent benchmark data for

Table 9.1 Binding constants of phenothiazine ligand RP12274 for three duplexes [40]

	$K_1 \, [\text{M}^{-1}]$	$K_2 \, [\text{M}^{-1}]$
d(CGTAAATTTACG)$_2$ DK33	$(2.8 \pm 0.7) \times 10^4$	$(5.5 \pm 0.9) \times 10^6$
d(CGCGAATTCGCG)$_2$ DK66	$(2.1 \pm 0.4) \times 10^4$	$(3.0 \pm 0.7) \times 10^6$
d(CGCGGGCCGCG)$_2$ DK100	$(1.6 \pm 0.3) \times 10^5$	$(1.7 \pm 0.8) \times 10^5$

assessing the stoichiometries of DNA:drug complexes and evaluating the fragmentation patterns of the resulting complexes.

Establishing correlations between ion abundances of DNA/ligand complexes and binding properties of the ligands in solution motivated the same group to derive equilibrium association constants (or equilibrium dissociation constants K_d) for duplex/ligand complexes based on ESI-MS [47]. The association constants were based on the ratio of the abundance of a particular duplex/ligand complex to the product of the abundances of the ligand-free duplex and the free ligand in the ESI mass spectra. It was assumed that the ionization efficiencies for the various ions (duplex, duplex–ligand complexes) were similar for this simple ratiometric strategy, and the association constants obtained by this method showed good agreement with those determined by conventional solution titration-based methods. Normalization of the differences in ionization efficiencies of the free and ligand-bound duplexes (by using an internal standard) and using a comprehensive mass balance analysis which necessitated that all complexes produced in solution must be detected in the ESI mass spectra further improved the accuracy of the ESI-MS method [49, 50]. More detailed insight into the factors that influence the determination of solution binding constants by ESI-MS and approaches for making more accurate quantitative measurements is presented in [49, 50].

Examples of the binding constants derived from ESI mass spectra are summarized in Table 9.1 for a minor groove binding phenothiazine ligand (RP12274; see Fig. 9.6), originally developed as a DNA photosensitizer [40]. The relevant ESI mass spectra are shown in the upper half of Fig. 9.6, and the distributions of 1:1 and 2:1 phenothiazine:duplex complexes are shown in the lower half. The ESI mass spectra are easily interpreted—the free duplex (5-) and both 1:1 and 2:1 ligand: duplex (5-) complexes have unambiguous m/z values that make them readily assigned. The ESI-MS results revealed the sequence-dependent nature of the binding of phenothiazine to duplex DNA with a preference for A/T-rich duplexes as evidenced by the increase in binding constant with the increasing AT content of the duplex (Table 9.1). The binding constants for the 2:1 complexes are two orders of magnitude greater than for the 1:1 complexes for the two AT-rich duplexes, thus confirming the cooperativity of DNA:phenothiazine interactions [40].

An alternative method for the estimation of relative binding affinities was based on collision-induced dissociation of DNA–ligand complexes in which the abundances of product ions containing one ligand were compared to those containing a
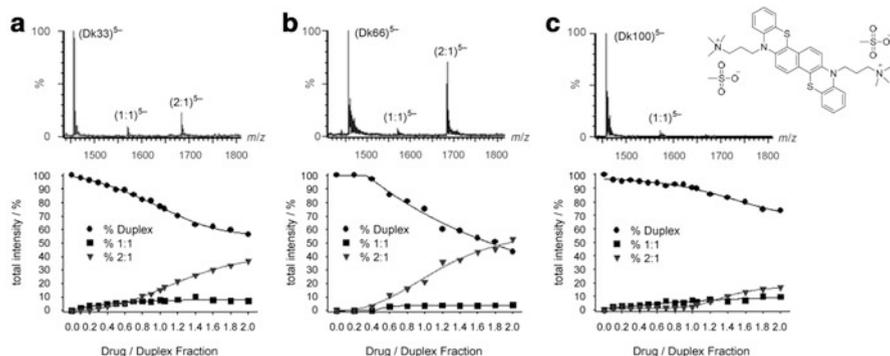
Fig. 9.6 ESI-MS spectra and companion kinetic plots for solutions containing 8 μM RP12274 (structure shown in *top right corner*) and 5 μM DNA for three duplexes of varying AT content. (**a**) DK33 = (CGTAAATTTACG)$_2$ duplex, (**b**) DK66 = (CGCGAATTCGCG)$_2$ duplex, and (**c**) DK100 = (CGCGGGCCCGCG)$_2$. Graphs of the relative abundance of the free duplex (*filled circle*), the 1:1 (*filled square*), and 2:1 (*filled inverted triangle*) complexes versus the drug molar fraction added to a 5 μM duplex solution are shown in the *lower panels*. Reprinted with permission from [40]

different duplex. A particular benefit of this method was that it did not require accurate information about the concentration of the ligand in solution. This method was demonstrated using a *N*-methylpyrrole polyamide ligand and was a facile way to order the relative affinities of a ligand as a function of the type of duplex [48].

ESI-MS has been used to evaluate a number of other types of DNA-interactive ligands in order to correlate structural features that mediate DNA binding. For example, the complexation of ten flavonoid aglycones and ten flavonoid glycosides with five DNA duplexes of varying AT/GC content was reported [41]. The 4′ OH group of the flavonoid aglycones was found to be critical for interaction with DNA, and in addition those flavonoids with sugars attached to the A or B ring exhibited higher binding affinities to DNA. In another study, ESI-MS was used to evaluate the binding of cosmomycin, an anthracycline analog with two trisaccharide chains, to DNA. Based on the fact that cosmomycin exhibited higher relative binding affinities to DNA than the analogs daunorubicin or doxorubicin, it was proposed that the intercalation of the anthracycline-like moiety was supplemented by binding interactions of the sugar chains in the minor groove, thus resulting in complexes with greater stability and with a greater footprint [43].

Another compelling application entailed using ESI-MS to identify poly-imidazole lexitropsin derivatives (structures given in Fig. 9.7) that recognized and bound to T:G mismatches in DNA [45]. The bar graphs in Fig. 9.8 illustrate the relative DNA binding affinities of three lexitropsin analogs. One of the polyimidazole ligands exhibited a significantly higher affinity for a DNA sequence that had a T:G mismatched base pair. In particular, ligand NMS-057 showed over a
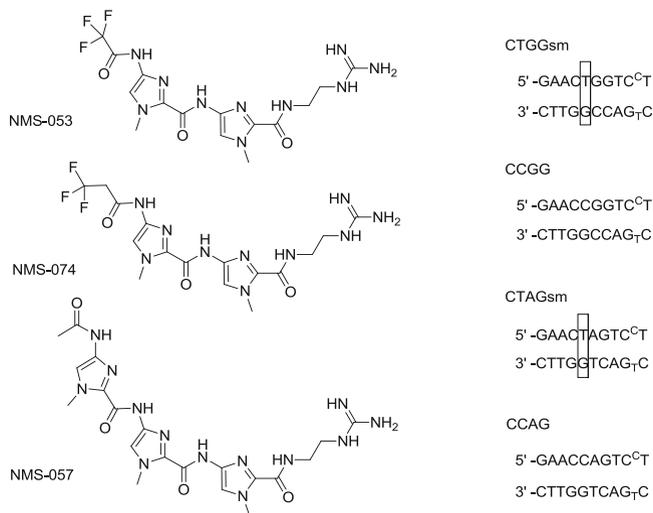
Fig. 9.7 Structures of three polyimidazole lexitropsin analogs and four DNA hairpins
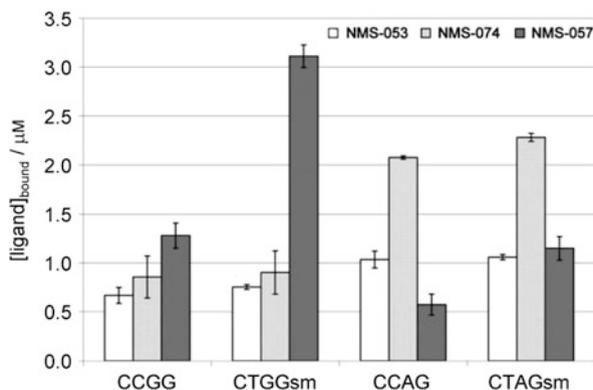
Fig. 9.8 Amount of bound ligand determined for three lexitropsin analogs (NMS-053, NMS-074, and NMS-057, structures in Fig. 9.7) at a concentration of 20 μM and for four different DNA sequences. Reprinted with permission from [45]

twofold higher affinity for mismatched CTGGsm than matched CCGG based on the abundances of 1:1 and 1:2 complexes relative to the abundances of ligand-free duplexes in the ESI mass spectra.
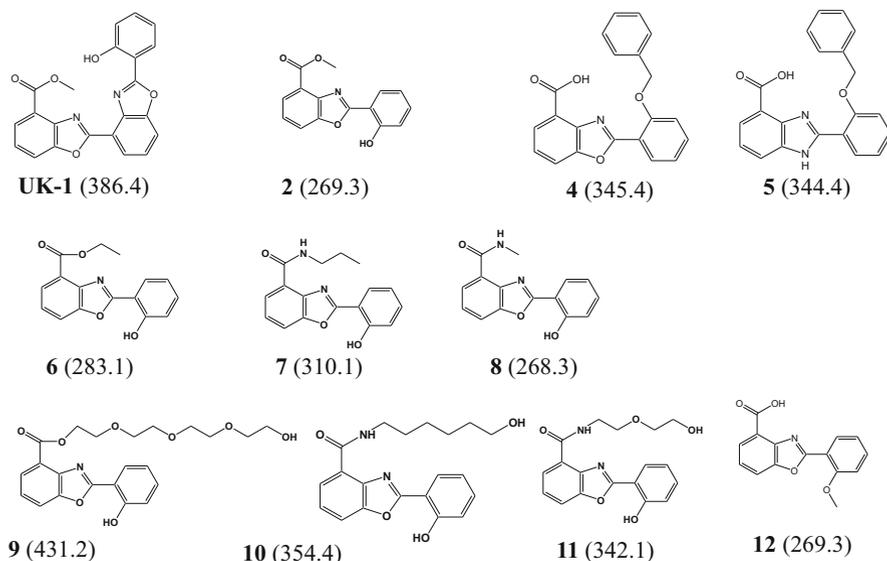
Fig. 9.9 Structures of the benzoxazole and benzimidazole ligands. Molecular weights are given (in Da) in the *parentheses*

9.6   Metal-Mediated Binding

The influence of metal ions on the binding of ligands to DNA has also been investigated by ESI-MS. One of the original studies confirmed the metal-mediated binding of mithramycin A and chromomycin A_3 to duplexes occurred selectively in the presence of transition metals Co^{2+} and Ni^{2+} [55]. In another study, the impact of the identity of the metal on the DNA complexation of a series of benzoxazoles possessing different amide- and ester-linked side chains was evaluated by ESI-MS (Fig. 9.9) [55, 56]. The benzoxazole family has generated substantial interest stemming from the original anticancer activity found for a bis(benzoxaole) meta-bolite isolated from *Streptomyces*, along with the additional finding that the DNA/benzoxazole interactions were metal mediated. A number of synthetic analogs have been produced and evaluated for cytotoxicity using conventional bioassays as well as spectrophotometric methods and ESI-MS (Figs. 9.9 and 9.10). As revealed by ESI-MS, some of the ligands only formed DNA complexes in the presence of metal cations, most notably for Cu^{2+} and to a lesser degree for Ni^{2+} and Zn^{2+}. An example is shown in Fig. 9.10 for one benzoxazole incubated with a self-complementary duplex, 5′-GCGAATTCGC-3′, and two different metal salts [55]. The compound shown in Fig. 9.10 exhibited formation of DNA–drug–metal complexes in the presence of Ni^{2+}, but no complexation in the presence of Mg^{2+}. The ligands exhibiting the most dramatic metal-enhanced DNA binding also demonstrated the greatest cytotoxic activity against A549 lung cancer and MCF7
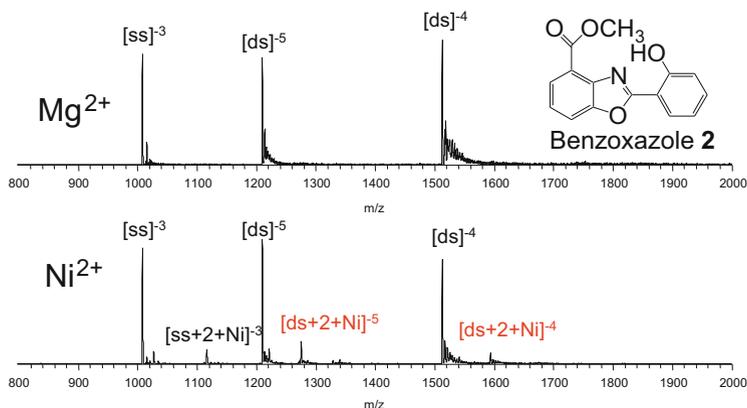
Fig. 9.10 ESI-mass spectra obtained for incubates of duplex 5′-GCGAATTCGC-3′ with benzoxazole 2. For more details, see [55, 56]

breast cancer cell lines. A follow-up study that encompassed a larger set of benzoazole ligands with both amide- and ester-linked side chains demonstrated that the side chains played a significant role in the DNA interactions, and some of the ligands with the longest side chains no longer displaying metal-mediated binding [56].

By monitoring the abundances of complexes via ESI-MS, the interactions of ruthenium and nickel complexes, $[Ru(phen)_2L]^{2+}$ and $[Ni(phen)_2L]^{2+}$ (where phen = 1,10-phenanthroline and L = dipyrido[3,2-*a*:2′,3′-c]phenazine, dipyrido [3,2-*a*:2′,3′-c] (6,7,8,9-tetrahydro)phenazine or 1,10-phenanthroline), with DNA duplexes were used to estimate their orders of relative binding affinities. The binding affinities correlated with the size of the ligand, and the complexes containing ruthenium led to consistently higher binding affinities than the complexes containing nickel [57, 58]. Metal-mediated binding has also been investigated for other DNA-interactive ligands, including ruthenium metallointercalators [52], porphyrins [54], and telomestatin [59].

9.7 Ligand Binding to Quadruplexes

ESI-MS has also been used to evaluate ligand binding to G-quadruplex DNA. The higher order DNA structure known as "quadruplex DNA" is assembled from three or more tetrads comprised of guanine bases from one to four DNA strands. The first report of quadruplex/ligand complexes by ESI-MS assessed the unique selectivity, binding stoichiometry, and mode of binding of a perylene diimide [61]. More recently, Gornall et al. evaluated the interactions of berberine alkaloids with various quadruplexes relative to duplexes [66, 68]. The quadruplex/berberine complexes
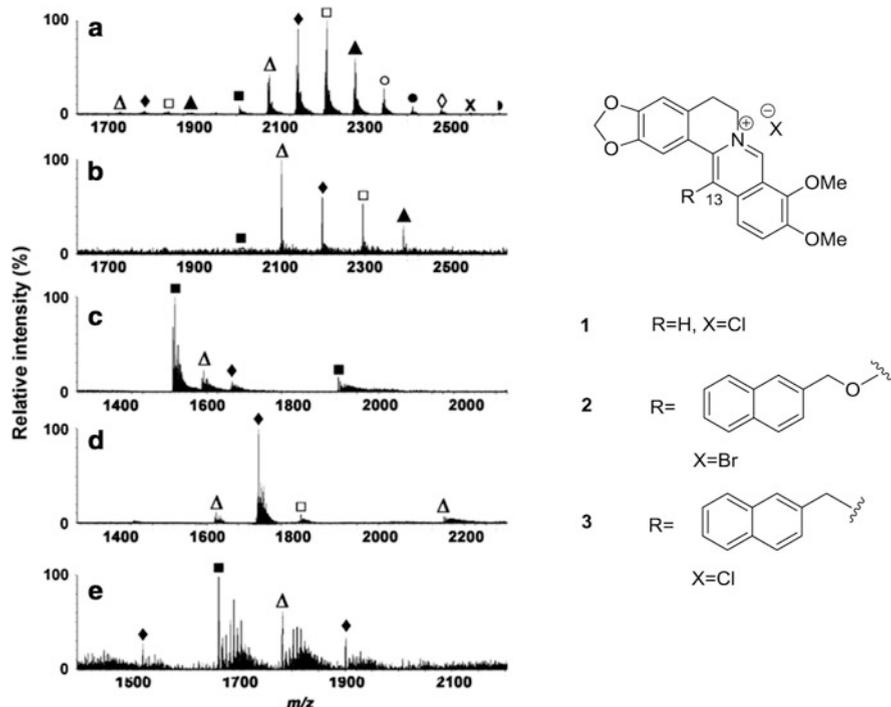
Fig. 9.11 Negative ion ESI mass spectra of 1:9 molar quadruplex–ligand mixtures. (**a**) Q4/1. (**b**) Q4/7. (**c**) Q2/1. (**d**) Q2/6. (**e**) Q1/7. The most abundant ion in each spectrum was $[Q4 + 4NH_4^+ + 3(1) + 9H]^{5-}$, m/z 2211.8; $[Q4 + 4NH_4^+ + 1(7) + 9H]^{5-}$ m/z 2105.3; $[Q2 + 2NH_4^+ + 7H]^{5-}$, m/z 1521.9; $[Q2 + 3NH_4^+ + 2(6) \, 8H]^{5-}$, m/z 1721.5; and $[Q1 - 4H]^{4-}$, m/z 1662.3, respectively. Ions are labeled as follows: *filled square* qDNA alone; *open triangle* qDNA + 1 ligand; *filled diamond* DNA + 2 ligands; *open square* qDNA + 3 ligands; *filled triangle* qDNA + 4 ligands; *open circle* DNA + 5 ligands; *filled circle* DNA + 6 ligands; *open diamond* DNA + 7 ligands; X qDNA + 8 ligands. The quadruplex sequences are as follows: Q1 is $G_3(T_2AG_3)_3$, Q2 is $(G_4T_4G_4)_2$, and Q4 is $(T_2G_5T)_4$. Reprinted from [68]

were readily identified based on the number of associated ammonium ions (one, two, or three), number of ligands (zero, one, or two), and the charge state of the complexes. Examples of the ESI-mass spectra of the quadruplex/ligand complexes are shown in Fig. 9.11 for 1:9 molar mixtures of three quadruplexes and three ligands, as follows: Q4 with ligands **1** or **7**, Q2 with ligands **1** or **6**, and Q1 with ligand **7**, in which Q1 is G3(T2AG3)3, Q2 is (G4T4G4)2, and Q4 is (T2G5T)4. The binding patterns varied significantly depending on the quadruplex sequence and the specific structure of the ligand. Interestingly, none of the 13-substituted berberine ligands showed any binding to duplex DNA. Ligand **7** displayed a notable preference for Q4, whereas ligand **6** exhibited little selectivity among the three quadruplexes. Berberine showed similar high affinities for Q1 and Q4 and low affinity for Q2. Although the specific structural features and chemical properties that influence the striking selectivities of the berberine compounds remain elusive,

the ability of ESI-MS to evaluate binding selectivities and provide feedback with respect to the impact of substituents on binding selectivities was confirmed.

Other groups have used ESI-MS to evaluate the quadruplex binding affinities of ligands based on the antitumor drug ditercalinium [62], ethidium derivatives [63], perylene diimides [65], phenanthroline complexes [69], platinum-based phenanthroimidazole complexes [71], perylene and coronene compounds [72], tris-triazole structures [73], an array of natural products including jatrorrhizine hydrochloride and tetrandrine [74], and tridentate N-donor terpyridine palladium (II) complexes [75].

9.8   Ligands That Form Covalent Bonds to DNA

The effects of many chemotherapeutic agents arise from the formation of covalent bonds to DNA, causing damage which can terminate replication or transcription steps during processing of the genetic code or activating cellular repair processes that can cleave the DNA [76, 77]. There is an array of electrophilic compounds that form covalent bonds to DNA, primarily via alkylation of specific sites of nucleobases [78–83]. These include mitomycin [78], acridine ligands [79], enediynes [80], nitrogen mustards [81], tamoxifen [82], and platinum derivatives [83, 84]. The most reactive sites of nucleobases are typically the N7 position of guanine and the N3 position of adenine. Those compounds with two reactive groups may react with two nucleophilic sites in DNA, thus leading to the formation of cross-links. Interstrand crosslinks are particularly deadly to cells because they prevent DNA strand separation during replication and transcription. This general class of DNA alkylating agents includes nitrogen mustards, cisplatin, and psoralens.

The observation of a DNA–ligand complex in an ESI mass spectrum does not reveal whether it is a covalent or noncovalent complex because the measurement of mass alone only conveys stoichiometric information. A simple initial test to determine whether the observed complexes are covalent or noncovalent is to heat the solutions and/or decrease the ionic strength to denature the dsDNA. Noncovalent complexes should disassemble (in conjunction with denaturation of the DNA), whereas the strands should remain bound for the covalent complexes. MS/MS methods may be used to assist in unraveling the nature of the binding mode. Often noncovalent complexes will disassemble upon activation in the gas phase or dissociate by rather simple pathways, whereas complexes held together through covalent bonds may follow more elaborate fragmentation routes, such as loss of a single base plus the attached ligand. The first ESI-MS report of the identification of a covalent crosslink between a duplex containing six base pairs and cisplatin occurred in 1995 [85], and the application of ESI-MS to characterization of covalent DNA adducts has flourished over the past 15 years. A recent review focused on the tandem mass spectrometric strategies developed for characterization of covalent adducts formed between DNA and chemotherapeutic ligands, and thus only a couple of examples are described in this chapter [20].

Fig. 9.12 Structures of acridine mustards and proposed pathway for interaction of mustard with guanine. For more details, see [44]

A recent study reported the use of ESI-MS to assess the binding mode (covalent versus noncovalent) of a series of sulfur-containing acridine mustards (Fig. 9.12) which have the potential to act as strong alkylating agents [44]. Ligands possessing a 2-chloroethyl sulfide moiety alkylate DNA very effectively (Fig. 9.12), and those equipped with an acridine moiety were designed to intercalate DNA and thus enhance the ability of the ligands to interact with their target (DNA) and facilitate subsequent alkylation. Ultimately these types of bioreductive antitumor agents may be effective for targeting tumors with elevated levels of sulfur oxide reducing enzymes. As reported by Smith et al., all of the acridines formed abundant complexes with duplex DNA. AS4, an acridine sulfide mustard, formed complexes with duplex DNA that were consistent with alkylation of the DNA as supported by tandem mass spectrometry. MS/MS of the duplex/AS4 complexes (via infrared multiphoton dissociation) in the 6- charge state resulted in strand separation, with the AS4 ligand bound to either strand (Fig. 9.13). Most importantly, two products were detected that corresponded to each of the constituent single strands with retention of both the AS4 moiety in addition to a single guanine base, i.e., [ss + AS2 + GH] that must originate from the other strand. This product confirmed that the AS4 ligand covalently bound to guanine; then the guanine from one strand remained bound to the complementary strand via the AS4 crosslink. The complexes containing the other acridine ligands dissociated via disassembly of the complex and ejection of the ligand, indicative of noncovalent interactions [44].

Fig. 9.13 IRMPD mass spectrum of $[ds + AS4]^{6-}$ with an irradiation time of 1.0 ms (structure of AS4 is in Fig. 9.12). The precursor ion is denoted with an *asterisk*. Reprinted with permission from [44]

ESI-MS/MS was also used to evaluate the interactions of pyrrolobenzodiazepine (PBD) dimers, which are synthetic sequence-selective DNA minor groove cross-linking agents, with oligonucleotides of varying length and sequence [86]. It was demonstrated that the type of adduct formed between PBD dimers and DNA depended on the distance between the two reacting guanines. Interstrand cross-linking was favored with shorter DNA sequences, whereas intrastrand cross-linking was preferred for the longer sequences along with monoalkylated adducts as minor products. For even longer sequences, PBD could not span the distance between the two guanines when the bases were located on either the same or opposite strands, and so in this case monoalkylated adducts were observed.

Another study reported the use of ESI-MS/MS to evaluate the relative binding affinities of a series of psoralens (Fig. 9.14) to DNA duplexes and to determine the formation of crosslinks, monoadducts, and noncovalent complexes (47). Psoralens are a class of furocoumarin ligands that have proven effective for the treatment of an array of skin diseases. Psoralens are good intercalators and in the presence of UV light may form cycloaddition monoadducts with pyrimidine bases, typically thymine. The cycloaddition reaction occurs between the 5,6 double bond of a thymine base and the 3,4 (pyrone) or 4′,5′ (furan) double bond of the psoralen, as illustrated in Fig. 9.14. The monoadducts may convert to crosslinked DNA structures upon a second photoabsorption event.

Incubation of P, 5-MOP, 8-MOP, or TMP with each duplex in the absence of UV irradiation led to no detectable products by ESI-MS [46]. However, after UV irradiation for 30 min at 365 nm, DNA/psoralen complexes were readily detected by ESI-MS, as illustrated in Fig. 9.15 for 8-MOP. Each psoralen analog had the ability to intercalate, form monoadducts, or cross-link DNA, and all of these products had identical masses. However, these three distinctive types of products were differentiated by MS/MS experiments, as discussed later. The duplex/8-MOP complexes were produced with solely 1:1 binding stoichiometries, and there were no complexes containing the single strands, thus indicating the specificity of the psoralen interactions with DNA. In contrast, incubation of AMP with each duplex resulted in formation of abundant duplex/AMP complexes with stoichiometries
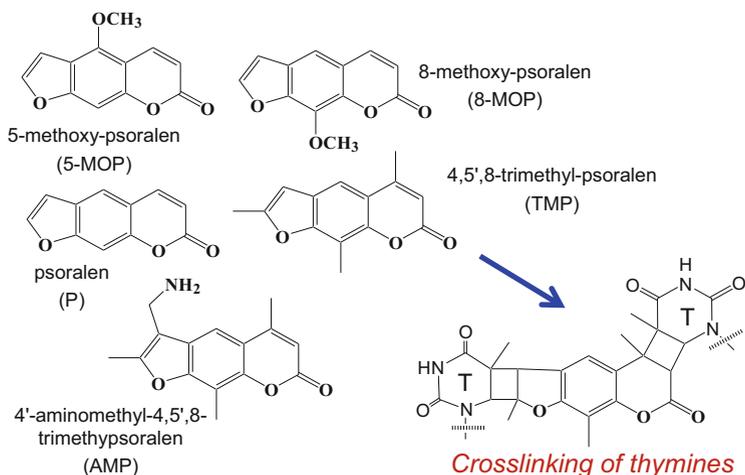
Fig. 9.14 Structures of psoralen ligands and proposed crosslink to thymines
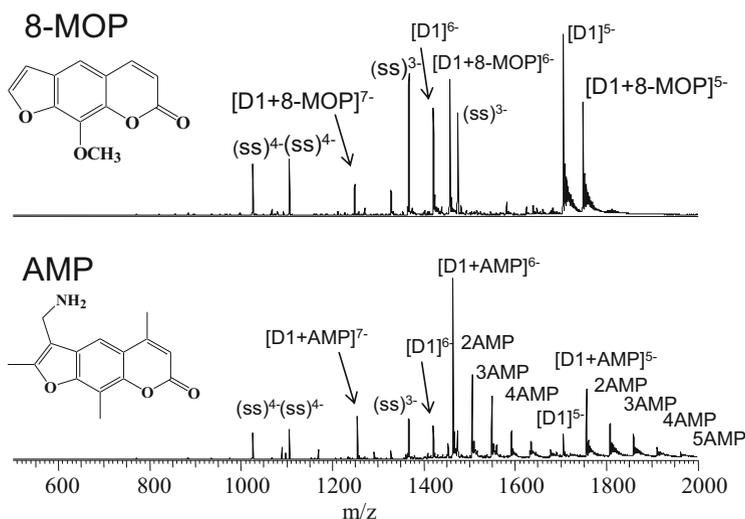
Fig. 9.15 ESI mass spectra of solutions containing the 5′-GCGGGGTAGGGGCG/ 5′-CGCCCCTACCCCGC duplex incubated with 8-MOP or AMP. Reprinted with permission from [46]

ranging from 1:1 to 1:5, even in the absence of UV irradiation. The ESI mass spectrum in Fig. 9.15 confirmed the large degree of nonspecific noncovalent binding to DNA and possible self-aggregation of AMP.

The percentages of bound DNA, which reflected the binding affinities and sequence selectivities, measured by ESI-MS are summarized in Fig. 9.16.
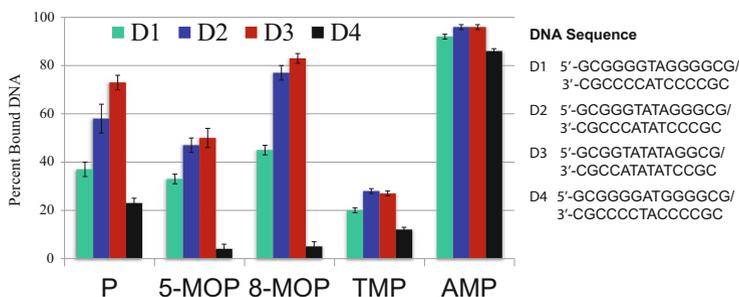
Fig. 9.16 Percent bound DNA for each psoralen based on ESI-MS peak areas. The percentages of bound DNA were determined after adding 1 μL of saturated psoralen to the DNA solution and subjecting the incubate to 30 min of UVA irradiation. The bound DNA complexes include the intercalation products, mono-adducts, and crosslinks. However, intercalation products are only observed for AMP. Reprinted with permission from [46]

In general, the percentages of bound DNA indicated that P, 5-MOP, 8-MOP, and to a lesser extent TMP were selective for specific DNA sequences, with a preference for the AT-rich duplex (D3) and low affinity for the duplex with the reversed orientation 5′-AT binding site [46]. The percentages of bound DNA obtained for AMP were the highest of all the psoralens, yet with no selectivity among the four duplexes. For AMP, the duplex:ligand stoichiometries of the complexes ranged from 1:1 up to 1:6, indicative of nonspecific binding.

To evaluate the nature of the interactions (noncovalent intercalation or covalent adduction and crosslink formation), MS/MS via infrared multiphoton dissociation was used to characterize the resulting duplex/psoralen structures [46]. The abundances of unique types of fragment ions that were characteristic for each type of product were tabulated to determine the distribution of crosslinks [based on (ds + psoralen - B)$^{6-}$, and (ss + 8-MOP + w) and (ss + 8-MOP + a–B) fragment ions], monoadducts [based on summation of (ss + psoralen)$^{3-}$ and (ss + psoralen - B)$^{3-}$ fragment ions], and noncovalent complexes [summation of (ss)$^{3-}$, base loss, and (ds)$^{6-}$ ions]. Note that single strands of DNA typically dissociate via base loss and subsequent formation of a–B and w ions in which B represents a nucleobase. The resulting product distributions are shown in Table 9.2. The overall trend for cross-link formation was 8-MOP > P ~ 5-MOP > TMP ≫ AMP. The distribution of cross-links to monoadducts showed remarkable agreement with the product distributions obtained by a conventional HPLC-UV method in which the various products were separated and the chromatographic peaks were integrated [46]. This study confirmed that ESI-MS/MS afforded a successful means to determine product distributions and confirm types of product structures for DNA-interactive ligands that have several possible binding modes.

Arguably one of the most well-known anticancer drugs is cisplatin, *cis*-[PtCl$_2$(NH$_3$)$_2$], a compound that readily crosslinks DNA via binding at N7 atoms of guanine [84, 87]. DNA–cisplatin adducts have been characterized using collision-induced dissociation by several groups [88–91]. A recent study evaluated

Table 9.2 Percent distribution of various products upon incubation of duplex 5′-GCGGGGTAGGGGCG/5′-CGCCCCTACCCCGC with each psoralen

Psoralen ligand	Cross-links	Mono-adducts	Non-covalent complexes	Cross-links/mono-adducts
P	53	22	25	2.4
5-MOP	56	24	20	2.3
8-MOP	60	14	26	4.3
TMP	46	32	22	1.4
AMP	28	35	37	0.8

Adapted from [46]
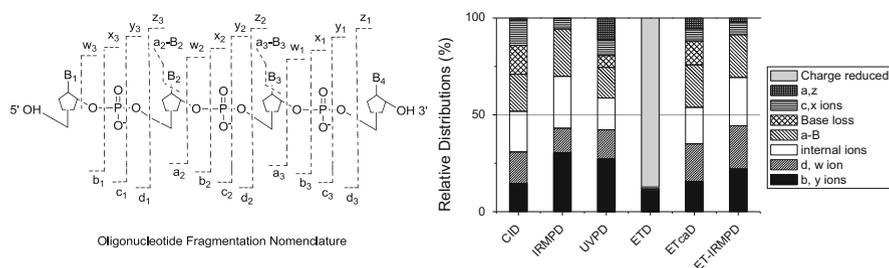
Fig. 9.17 Relative distributions of all ion types upon CID, IRMPD, UVPD, ETD, ETcaD, ETIRMPD for 3+ (or 3$^{+\bullet}$) charge state of platinated CCC CCC GGC CCC CC. The relative portion of each category of ion is obtained by dividing the summed abundance for all fragments in a particular category by the total abundance of all fragment ions. Internal fragments result from a double backbone cleavage. They have a phosphate at their 5′ end and a furan at the 3′ terminal. Reprinted with permission from [90]

the use of an array of MS/MS methods, including CID, infrared multiphoton dissociation (IRMPD) (10.6 μm), electron transfer dissociation (ETD), negative ETD (NETD), or ultraviolet photodissocation (UVPD) (193 nm) and hybrid processes termed electron transfer collisional activated dissociation (ETcaD), electron transfer-infrared multiphoton dissociation (ET-IRMPD), or electron transfer-ultraviolet photodissociation (ET-UVPD), to optimize the structural characterization of DNA–cisplatin adducts, particularly to identify the sites of cisplatin binding [91]. The hybrid MS/MS methods entailed the use of electron transfer reactions to produce odd electron ions, followed by activation of the resulting charge-reduced precursors in order to generate a more diverse array of fragment ions. The types and distributions of fragment ions are summarized in Fig. 9.17. Two laser-based methods, IRMIPD and UVPD, yielded the most informative array of fragment ions, in terms of providing full coverage of the oligonucleotide sequence as well as platinated product ions that were useful for pinpointing the adduction site [91].

9.9   Ligand Binding to RNA

The use of electrospray ionization mass spectrometry for screening ligand binding to RNA targets has also had a rich history [92–97], particularly for aminoglycoside antibiotics such as ribostamycin and paromomycin [94], 2-deoxystreptamine [95], bacterial-based natural product libraries [96], as well HIV-1 Tat peptide binding to transactivation-responsive RNA [93]. Even low binding affinity (in the mM range) RNA complexes have been successfully transferred to the gas phase and analyzed with respect to their stoichiometries and binding affinities [95]. The investigation of ligand binding to RNA has been described in more detail in several reviews [13, 19].

9.10   Conclusion and Outlook

ESI-MS has enabled the characterization of a wide array of DNA–ligand complexes, including those involving duplex, quadruplex, and other alternative DNA structures. Ligands ranging from intercalators to minor groove binding agents to those that stack on the end of higher order DNA structures and ligands that covalently bind and form crosslinks have been evaluated as DNA-interactive agents. The ability to transfer DNA complexes to the gas phase by the electrospray process offers remarkable insight into species assembled by either noncovalent interactions or covalent bonds. In fact, ESI-MS is now recognized as one of the most versatile methods for investigating DNA–ligand interactions, and its application has extended far beyond systematic validation studies to those that rely on ESI-MS as a routine strategy to: confirm the stoichiometries of complexes, estimate DNA binding affinities and sequence selectivities, assess binding modes, and screen series of new ligands. The minimal sample consumption and high sensitivity of ESI-MS make it a first choice when studies involve high-throughput screening or precious drug candidates isolated or synthesized in limited quantities. The capability for studying even more elaborate macromolecular architectures and assemblies containing DNA with other biological molecules has already been achieved [13, 14] and promises to make ESI-MS/MS an even more powerful strategy in the upcoming years.

9.11   Summary of Key Concepts

- ESI can be used to transfer noncovalent DNA–ligand complexes into the gas phase.
- Ion abundances in ESI mass spectra can be used to determine binding constants of DNA–ligand complexes and sequence selectivities of DNA-interactive ligands.

- Complexes arising from intercalation, binding in the minor groove, interactions with quadruplexes, and both monoadducts and cross-links can be detected and characterized in the gas phase.
- Elucidation of ligand binding sites for covalent DNA complexes may be facilitated by use of more than one MS/MS method, including the use of hybrid activation methods to yield more diverse series of fragment ions that may pinpoint structural details.

References

1. Goodman LS, Hardman JG, Limbird LE, Gilman AG (2001) Goodman & Gilman's the pharmacological basis of therapeutics, 10th edn. New York, McGraw-Hill
2. Brana MF, Cacho M, Gradillas A, de Pascual-Teresa B, Ramos A (2001) Intercalators as anticancer drugs. Curr Pharm Des 7:1745–1780
3. Liu HK, Sadler PJ (2011) Metal complexes as DNA intercalators. Acc Chem Res 44(5): 349–359
4. Bischoff G, Hoffmann S (2002) DNA-binding drugs used in medicinal therapy. Curr Med Chem 9:321–348
5. Baraldi PG, Bovere A, Fruttarolo F, Preti D, Tabrizi M, Pavania M, Romagnoli R (2004) DNA minor groove binders as potential antitumor and antimicrobial agents. Med Res Rev 24:475–528
6. Hartley JA, Hochhauser D (2012) Small molecule drugs-optimizing DNA damaging agent-based therapeutics. Curr Opin Pharmacol 12(4):3980402
7. Huang Z-S, Tan J-H, Ou T-M, Li D, Gu L-Q (2011) G-quadruplex DNA and its ligands in anticancer therapy. In: Zhang L-H, Xi Z, Chattopadhyaya J (eds) Medicinal chemistry of nucleic acids. Wiley, Hoboken, NJ, pp 206–257
8. Reedijk J, Lohman PHM (1985) Cisplatin: synthesis, antitumour activity and mechanism of action. Pharm World Sci 7:173–180
9. Rajski SR, Williams RM (1998) DNA cross-linking agents as antitumor drugs. Chem Rev 98: 2723–2795
10. Borowy-Borowski H, Lipman R, Tomasz M (1990) Recognition between mitomycin C and specific DNA sequences for cross-link formation. Biochemistry 29:2999–3006
11. Zhang L-H, Xi Z, Chattopadhyaya J (2011) Medicinal chemistry of nucleic acids. Wiley, Hoboken, NJ
12. Singh Y, Palombo M, Sinko PJ (2008) Recent trends in targeted anticancer prodrug and conjugate design. Curr Med Chem 15:1802
13. Hofstadler SA, Griffey RH (2001) Analysis of noncovalent complexes of DNA and RNA by mass spectrometry. Chem Rev 101:377–390
14. Beck JL, Colgrave ML, Ralph SF, Sheil MM (2001) Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. Mass Spectrom Rev 20: 61–87
15. Rosu F, De Pauw E, Gabelica V (2008) Electrospray mass spectrometry to study drug-nucleic acid interactions. Biochimie 90:1074–1087
16. Beck JL (2011) Developments in electrospray ionization mass spectrometry of noncovalent DNA-ligand complexes. Aust J Chem 64:705–717

17. Gabelica V (2010) Electrospray mass spectrometry of noncovalent complexes between small molecule ligands and nucleic acids. In: Banoub JH, Limbach PA (eds) Mass spectrometry of nucleosides and nucleic acids, vol 8. CRC, Hoboken, NJ, pp 283–302
18. Brodbelt JS (2009) Evaluation of DNA/ligand interactions by electrospray ionization mass spectrometry. Annu Rev Anal Chem 3:67–87
19. Hofstadler SA, Griffey RH (2000) Mass spectrometry as a drug discovery platform against RNA targets. Curr Opin Drug Discov Devel 3(4):423–431
20. Silvestri C, Brodbelt JS (2012) Tandem mass spectrometry for characterization of covalent adducts of DNA with anticancer therapeutics. Mass Spectrom Rev 32:247–266
21. Haq I, Jenkins TC, Chowdhry BZ, Ren J, Chaires JB (2000) Parsing free energies of Drug-DNA interactions. Methods Enzymol 323:373–405
22. Haq I, Ladbury J (2000) Drug-DNA recognition: energetics and implications for design. J Mol Recognit 13:188–197
23. Eriksson M, Norden B (2001) Linear and circular dichroism of drug-nucleic acids complexes. Methods Enzymol 340:68–98
24. Lane AN (2001) Nuclear magnetic resonance studies of drug-DNA coplexes in solution. Methods Enzymol 340:252–281
25. Peek ME, Williams LD (2001) X-Ray crystallography of drug-DNA complexes. Methods Enzymol 340:282–289
26. Lilley DMJ (1992) Probes of DNA structure. Methods Enzymol 212:133–139
27. Fox KR, Waring MJ (2001) High-resolution footprinting studies of drug-DNA complexes using chemical and enzymatic probes. Methods Enzymol 340:412–430
28. Gale DC, Smith RD (1995) Characterization of noncovalent complexes formed between minor groove binding molecules and duplex DNA by electrospray ionization-mass spectrometry. J Am Soc Mass Spectrom 6:1154–1164
29. Gale DC, Goodlett DR, Light-Wahl KJ, Smith RD (1994) Observation of duplex DNA-drug noncovalent complexes by electrospray ionization mass spectrometry. J Am Chem Soc 116:6027–6028
30. Kapur A, Beck JL, Sheil MM (1999) Observation of daunomycin and nogalamycin complexes with duplex DNA using electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom 13:2489–2497
31. Gabelica V, De Pauw E, Rosu F (1999) Interaction between antitumor drugs and a double-stranded oligonucleotide studied by electrospray ionization mass spectrometry. J Mass Spectrom 34:1328–1337
32. Wan KX, Shibue T, Gross ML (2000) Noncovalent complexes between DNA-binding drugs and double-stranded oligodeoxynucleotides: a study by ESI ion-trap mass spectrometry. J Am Chem Soc 122:300–307
33. Colgrave ML, Beck JL, Sheil MM, Searle MS (2002) C electrospray ionisation mass spectrometric detection of weak noncovalent interactions in nogalamycin-DNA complexes. Chem Commun: 556–557
34. Gupta R, Beck JL, Ralph SF, Sheil MM, Aldrich-Wright JR (2004) Comparison of the binding stoichiometries of positively charged DNA-binding drugs using positive and negative ion electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 15:1382–1391
35. Chen WH, Chan CL, Cai Z, Luo GA, Jiang ZH (2004) Study on noncovalent complexes of cytotoxic protoberbine alkaloids with double-stranded DNA by using electrospray ionization mass spectrometry. Bioorg Med Lett 14:4955–4959
36. Chen WH, Quin Y, Cai Z, Chan CL, Luo GA, Jiang ZH (2005) Spectrometric studies of cytotoxic protoberberine alkaloids binding to double-stranded DNA. Bioorg Med Chem 13:1859–1866
37. Rosu F, Pirotte S, DePauw E, Gabelica V (2006) Positive and negative ion mode ESI-MS and MS/MS for studying drug-DNA complexes. Int J Mass Spectrom 253(3):156–171

38. Mazzitelli CL, Chu Y, Reczek JJ, Iverson BL, Brodbelt JS (2007) Screening of threading Bis-intercalators binding to duplex DNA by electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 18:311–321

39. Smith SI, Guziec LJ, Guziec FS, Hasinoff BB, Brodbelt JS (2007) Evaluation of relative DNA binding affinities of anthrapyrazoles by electrospray ionization mass spectrometry. J Mass Spectrom 42:681–688

40. Rosu F, Gabelica V, DePauw E, Mailliet P, Mergny J-L (2008) Cooperative 2:1 binding of a bisphenothiazine to duplex DNA. ChemBioChem 9(6):849–852

41. Wang Z, Cui M, Song F, Lu L, Liu Z, Liu S (2008) Evaluation of flavonoids binding to DNA duplexes by electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 19:914–922

42. Wang Z, Guo X, Liu Z, Cui M, Song F, Liu S (2008) Studies on alkaloids binding to GC-rich human survivin promoter DNA using positive and negative ion electrospray ionization mass spectrometry. J Mass Spectrom 43:327–335

43. Kelso C, Tillott V, Rojas JD, Furlan RLA, Padilla G, Beck JL (2008) Mass spectrometric investigation of the DNA-binding properties of an anthracycline with two trisaccharide chains. Arch Biochem Biophys 477:348–355

44. Smith S, Guziec FS, Guziec L, Brodbelt JS (2009) Interactions of sulfur-containing acridine ligands with DNA by ESI-MS. Analyst 134:2058–2066

45. Sirtori FR, Aldini G, Colombo M, Colombo N, Malyszko J, Vistoli G, D'Alessio R (2012) Molecular recognition of T:G mismatched base pairs in DNA as studied by electrospray ionization mass spectrometry. ChemMedChem 7:1112–1122

46. Smith SI, Brodbelt JS (2010) Rapid characterization of cross-links, mono-adducts, and noncovalent binding of psoralens to deoxyoligonucleotides by LC—UV/ESI-MS and IRMPD mass spectrometry. Analyst 135(5):943–952

47. Rosu F, Gabelica V, Houssier C, De Pauw E (2002) Determination of affinity, stoichiometry and sequence selectivity of minor groove binder complexes with double-stranded oligodeoxynucleotides by electrospray ionization mass spectrometry. Nucleic Acids Res 30:e82/1–e82/9

48. Buchmann W, Boutorine A, Halby L, Tortajada J, De Pauw E (2009) A new method for the determination of the relative affinity of a ligand against various DNA sequences by electrospray ionization mass spectrometry. Application to a polyamide minor groove binder. J Mass Spectrom 44:1171–1181

49. Gabelica V, Galic N, Rosu F, Houssier C, De Pauw E (2003) Influence of response factors on determining equilibrium association constants of noncovalent complexes by electrospray ionization mass spectrometry. J Mass Spectrom 38:491–501

50. Gabelica V, Rosu F, De Pauw E (2009) A simple method to determine electrospray response factors of noncovalent complexes. Anal Chem 81:6708–6715

51. Reyzer ML, Brodbelt JS, Kerwin SM, Kumar D (2001) Evaluation of complexation of metal-mediated DNA-binding drugs to oligonucleotides via electrospray ionization mass spectrometry. Nucleic Acids Res 29:e103/1–e103/12

52. Beck JL, Gupta R, Urathamakul T, Williamson NL, Sheil MM, Aldrich-Wright JR, Ralph SF (2003) Probing DNA selectivity of ruthenium metallointercalators using ESI mass spectrometry. Chem Commun: 626–627

53. Urathamakul T, Beck JL, Sheil MM, Aldrich-Wright JR, Ralph SF (2004) Comparison of mass spectrometry and other techniques for probing interactions between metal complexes and DNA. Dalton Trans 17:2683–2690

54. Ramos CIV, Barros CM, Fernandes AM, Santana-Marques MG, Ferrer Correia AJ et al (2005) Interactions of cationic porphyrins with double-stranded oligodeoxynucleotides: a study by electrospray ionisation mass spectrometry. J Mass Spectrom 40:1439–1447

55. Oehlers L, Mazzitelli CL, Brodbelt JS, Rodriguez M, Kerwin S (2004) Evaluation of complexes of DNA duplexes and novel benzoxazoles or benzimidazoles by electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 15:1593–1603

56. Mazzitelli CL, Rodriguez M, Kerwin SM, Brodbelt JS (2008) Evaluation of metal-mediated DNA binding of benzoxazole ligands by electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 19:209–218

57. Urathamakul T, Waller DJ, Beck JL, Aldrich-Wright JR, Ralph SF (2008) Comparison of mass spectrometry and other techniques for probing interactions between metal complexes and DNA. Inorg Chem 47(15):6621–6632

58. Talib J, Harman DG, Dillon CT, Aldrich-Wright J, Beck JL, Ralph SF (2009) Does the metal influence noncovalent binding of complexes to DNA? Dalton Trans 3:504–513

59. Rosu F, Gabelica V, Smargiasso N, Mazzucchelli G, Shin-Ya G, De Pauw W (2010) Cation involvement in telomestatin binding to G-quadruplex DNA. J Nucleic Acids 2010:121259

60. Schilter D, Urathamakul T, Beck JL, Harding MM, Rendina LM (2010) ESI-MS and thermal melting studies of nanoscale platinum (II) metallomacrocycles with DNA. Dalton Trans 39(46):11263–11271

61. David WM, Brodbelt JS, Kerwin SM, Thomas PW (2002) Investigation of quadruplex oligonucleotide-drug interactions by electrospray ionization mass spectrometry. Anal Chem 74:2029–2033

62. Carrasco C, Rosu F, Gabelica F, Houssier C, De Pauw E, Garbay-Jaureguidberry C et al (2002) Tight binding of the antitumor drug ditercalinium to quadruplex DNA. Chembiochem 3: 1235–1241

63. Rosu F, De Pauw E, Guittat L, Alberti P, Lacroix L, Mailliet P, Riou JF, Mergny JL (2003) Selective interaction of ethidium derivatives with quadruplexes: an equilibrium dialysis and electrospray ionization mass spectrometry analysis. Biochemistry 42:10361–10371

64. Pothukuchy A, Mazzitelli C, Salazar M, Brodbelt JS, Kerwin SM (2005) Duplex and quadruplex DNA binding and photocleavage by trioxatriangulenium ion. Biochemistry 44: 2163–2172

65. Mazzitelli CL, Brodbelt JS, Kern JT, Rodriguez M, Kerwin SM (2006) Evaluation of binding of perylene diimide and benzannulated perylene diimide ligands to DNA by electrospray ionization mass spectrometry. J Am Soc Mass Spectrom 17:593–604

66. Gornall KC, Samosorn S, Talib J, Bremner JB, Beck JL (2007) Selectivity of an indolyl berberine derivative for tetrameric G-quadruplex DNA. Rapid Commun Mass Spectrom 21:1759–1766

67. Mazzitelli CL, Wang J, Smith JI, Brodbelt JS (2007) Gas-phase stability of G-quadruplex DNA determined by electrospray ionization tandem mass spectrometry and molecular dynamics simulations. J Am Soc Mass Spectrom 18:1760–1773

68. Gornall KC, Samosorn S, Talib J, Samorsorn S, Tanwirat B, Suksamrarn A, Bremner JB, Kelso MJ, Beck JL (2010) A mass spectrometric investigation of novel quadruplex DNA-selective berberine derivatives. Chem Commun 46:6602–6604

69. Talib J, Green C, Davis KJ, Urathamakul T, Beck JL, Aldrich-Wright JR, Ralph SF (2008) A Comparison of the binding of metal complexes to duplex and quadruplex DNA. Dalton Trans: 1018–1026

70. Collie G, Reszka AP, Haider SM, Gabelica V, Parkinson GN, Neidle S (2009) Selectivity in small molecule binding to human telomeric RNA and DNA quadruplexes. Chem Commun: 7482–7484

71. Pierce SE, Kieltyka R, Sleiman HF, Brodbelt JS (2009) Evaluation of binding selectivities and affinities of platinum-based quadruplex interactive complexes by electrospray ionization mass spectrometry. Biopolymers 91:233–243

72. Casagrande V, Antonello A, Armandodorianao B, Ortaggi G, Franceschin M (2009) Study of binding affinity and selectivity of perylene and coronene derivatives towards duplex and quadruplex DNA by ESI-MS. J Mass Spectrom 44:530–540

73. Lombardo CM, Martinez IS, Haider S, Gabelica V, De Pauw E, Moses JE, Neidle S (2010) Structure-based design of selective high-affinity telomeric quadruplex-binding ligands. Chem Commun 46(48):9116–9118

74. Liu Y, Zheng B, Xu X, Yan G (2010) Probing the binding affinity of small molecule natural products to the G-quadruplex in C-myc oncogene by electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom 24:3072–3075
75. Largy E, Hamon F, Rosu F, Gabelica V, De Pauw E, Guedin A, Mergny J-L, Teulade-Fichou M-P (2011) Tridentate N-donor palladium (II) complexes as efficiency coordinating quadruplex DNA binders. Chem Eur J 17:13274–13283
76. Waring MJ (1986) Overview of the interaction between chemotherapeutic agents and DNA. Drugs Exp Clin Res 12:441–453
77. Chaney SG, Sancar A (1996) DNA repair: enzymatic mechanisms and relevance to drug response. J Natl Cancer Inst 88:1346–1360
78. Tomasz M, Palom Y (1997) The mitomycin bioreductive antitumor agents: cross-linking and alkylation of DNA as the molecular basis of their activity. Pharmacol Ther 76:73–87
79. Lerman LS (1961) Structural considerations in the interaction of DNA and acridines. J Mol Biol 3:18–30
80. Nicolaou KC, Dai W-M (1991) Chemistry and biology of the enediyne anticancer antibiotics. Angew Chem Int Ed Engl 30:1387–1416
81. Colvin M, Brundrett RB, Kan M-NN, Jardine I, Fenselau C (1976) Alkylating properties of phosphoramide mustard. Cancer Res 36:1121–1126
82. Phillips DH, Carmichael PL, Hewer A, Cole KJ, Poon GK (1994) α-Hydroxytamoxifen, a metabolite of tamoxifen with exceptionally high DNA-binding activity in rat hepatocytes. Cancer Res 54:5518–5522
83. Sherman SE, Lippard SJ (1987) Structural aspects of platinum anticancer drug interactions with DNA. Chem Rev 87:1153–1181
84. Reedijk J (1987) The mechanism of action of platinum anti-tumor drugs. Pure Appl Chem 59: 181–192
85. Wickham G, Iannitti P, Boschenok J, Sheil MM (1995) Electrospray ionization mass spectrometry of covalent ligand-oligonucleotide adducts: evidence for specific duplex ion formation. Rapid Commun Mass Spectrom S197:203
86. Rahman KM, James CH, Thurston DE (2011) Effect of base sequence on the DNA cross-linking properties of pyrrolobenzodiazepine (PBD) dimers. Nucleic Acids Res 39:5800–5812
87. Lippert B (1999) Cisplatin: chemistry and biochemistry of a leading anticancer drug. VHCA &Wiley-VCH, Zurich
88. Iannitti-Tito P, Weimann A, Wickham G, Sheil MM (2000) Structural analysis of drug-DNA adducts by tandem mass spectrometry. Analyst 125(4):627
89. Egger AE, Hartinger CG, Ben Hamidane H, Tsybin YO, Keppler BK, Dyson PJ (2008) High resolution mass spectrometry for studying the interactions of cisplatin with oligonucleotides. Inorg Chem 47(22):10626
90. Nyakas A, Eymann M, Schuerch S (2009) The influence of cisplatin on the gas-phase dissociation of oligonucleotides studied by electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 20(5):792
91. Xu Z, Shaw JB, Brodbelt JS (2012) Comparison of MS/MS methods for characterization of DNA/cis-platin adducts. J Am Soc Mass Spectrom 24:265–272
92. Hofstadler SA, Sannes-Lowery KA, Crooke ST, Ecker DJ, Sasmor H, Manalili S, Griffey RH (1999) Multiplexed screening of neutral mass-tagged rna targets against ligand libraries with electrospray ionization FTICR MS: a paradigm for high-throughput affinity screening. Anal Chem 71:3436–3440
93. Sannes-Lowery KA, Hu P, Mack DP, Mei HY, Loo JA (1997) HIV-1 Tat peptide binding to TAR RNA by electrospray ionization mass spectrometry. Anal Chem 69:5130–5135
94. Griffey RH, Hofstadler SA, Sannes-Lowery KA, Ecker DJ, Crooke ST (1999) Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. Proc Natl Acad Sci USA 96:10129–10133

95. Griffey RH, Sannes-Lowery KA, Drader JJ, Mohan V, Swayze EE, Hofstadler SA (2000) Characterization of low-affinity complexes between RNA and small molecules using electrospray ionization mass spectrometry. J Am Chem Soc 122:9933–9938
96. Cummins LL, Chen S, Blyn LB, Sannes-Lowery KA, Drader JJ, Griffey RH, Hofstadler SA (2003) Multitarget affinity/specificity screening of natural products: finding and characterizing high-affinity ligands from complex mixtures by using high performance mass spectrometry. J Nat Prod 66:1186–1190
97. Sannes-Lowery KA, Cummins LL, Chen S, Drader JJ, Hofstadler SA (2004) High throughput drug discovery with ESI-FTICR. Int J Mass Spectrom 238:197–206

MS-Based Approaches for Nucleic Acid Structural Determination

<div align="right">

10

</div>

Daniele Fabris

Abstract

The vast majority of the human genome consists of sequences that do not code for proteins. Understanding their function must rely on approaches that are capable of providing more than mere sequence information. Alone or in combination with other techniques, mass spectrometry (MS) can provide an excellent platform for investigating non-coding nucleic acids (NA) at many different levels. This chapter reviews MS-based approaches developed to pursue the structural elucidation of species that are not readily amenable to the classic high-resolution techniques. Indeed, MS has recently found increasing applications as a detection platform for chemical probes used to interrogate NA structure in solution. These developments have been riding on the concomitant advances of computational approaches, which are rapidly closing the resolution gap with NMR and crystallography by taking full advantage of the sparse constraints afforded by alternative techniques. Further, the hierarchic nature of NA structure, which is characterized by the three-dimensional organization of discrete structural elements, lends itself well to the investigation by techniques that are capable of revealing the position of structure-defining interactions. For this reason, novel strategies are being developed to study secondary, tertiary, and quaternary interactions in the gas phase, which may retain memory of the solution architecture. The popularization of ion mobility spectrometry (IMS) has opened new avenues for investigating the overall topology of non-coding elements, which promise to contribute significantly to the elucidation of progressively larger NA systems.

D. Fabris (✉)
The RNA Institute, University at Albany (SUNY), Life Sciences Research Building room 1109, 1200 Washington Avenue, Albany, NY 12222, USA
e-mail: fabris@albany.edu

Abbreviations/Formulae

A	Adenine nucleotide
BDG	4,4′-Bihenyl-diglyoxal
BKT	Bikethoxal
C	Cytosine nucleotide
CASP	Critical assessment of protein structure prediction
CCS	Collisional cross section
CID	Collision-induced dissociation
CMCT	1-Cyclohexyl-3-(2-morpholinoethyl)-carbodiimide metho-p-toluene sulfonate
CPT	Cisplatin
DMS	Dimethylsulfate
DNA	Deoxyribonucleic acid
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
FIV	Feline immunodeficiency virus
G	Guanine nucleotide
GUI	Graphic user interface
HDX	Hydrogen–deuterium exchange
HIV-1	Immunodeficiency virus type 1
IMS	Ion mobility spectrometry
IRMPD	Infrared multiphoton dissociation
KT	Kethoxal, β-ethoxy-α-ketobutyraldehyde
MALDI	Matrix-assisted laser desorption ionization
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS3D	Mass spectrometry three-dimensional
NA	Nucleic acids
NB	Neomycin B
NC	HIV-1 nucleocapsid protein
NM	Nitrogen mustard
NMIA	*N*-methylisatoic anhydride

NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
PDG	1,4-Phenyl-diglyoxal
PPT	Polypurine tract
RMSD	Root mean square deviation
RNA	Ribonucleic acid
SAXS	Small-angle X-ray scattering
SHAMS	2′-Hydroxyl acylation by mass spectrometry
SHAPE	2′-Hydroxyl acylation by primer extension
SL1 though 4	HIV-1 stemloop 1 through 4
T	Thymine nucleotide
U	Uracil nucleotide
UV	Ultraviolet
WC	Watson–Crick
Ψ-RNA	HIV-1 packaging signal

10.1   Introduction

The significance of investigating the 3D structure of nuclei acids stems from the observation that only ~2 % of the three billion base pairs present in the human genome code for actual proteins. Ribozymes and riboswitches are proof that the function of nucleic acids is not defined exclusively by the genetic information encoded in their primary structure, but may hinge also on their ability to interact with other cellular components, which is determined by their higher order structure. The RNA cleaving activity of ribozymes rests upon the proper folding of the catalytic core and its placement onto a susceptible region [1, 2]. Allosteric regulation of mRNA expression is a direct function of the ability of riboswitches to form alternative structures in response to the binding of specific ligands (Fig. 10.1) [3, 4]. For this reason, the investigation of structure and dynamics assumes a central role in the elucidation of the biological function of non-coding sequences that represent the vast majority of the human genome [5, 6].

The higher order structure of nucleic acids (NA) can be described as the hierarchic arrangement of constitutive structural units [7, 8]. Fundamental elements of secondary structure are characterized by helical regions that, in the case of stemloop domains, originate from the annealing of contiguous sequences of the same strand [9]. Tertiary structure requires annealing of distal stretches of the same strand, or other types of long-range contacts between discrete domains, which combine to shape the topology assumed by the entire molecule. Quaternary assemblies involve annealing of separate strands or binding of proteins and other ligands, which lead to the formation of functional complexes [10]. The necessary stabilization is typically contributed by a combination of base pairing, stacking, metal binding, and hydration effects [7, 8]. The prominence of double-stranded structures, however, makes base pairing the most important factor in determining

Fig. 10.1 General riboswitch mechanism. (**a**) In the absence of ligand (L) the AUG starting codon may be masked by secondary structure folded by the 5′ non-coding region of mRNA. (**b**) Ligand binding mediates a rearrangement of the secondary structure, which exposes the starting codon and initiates gene translation
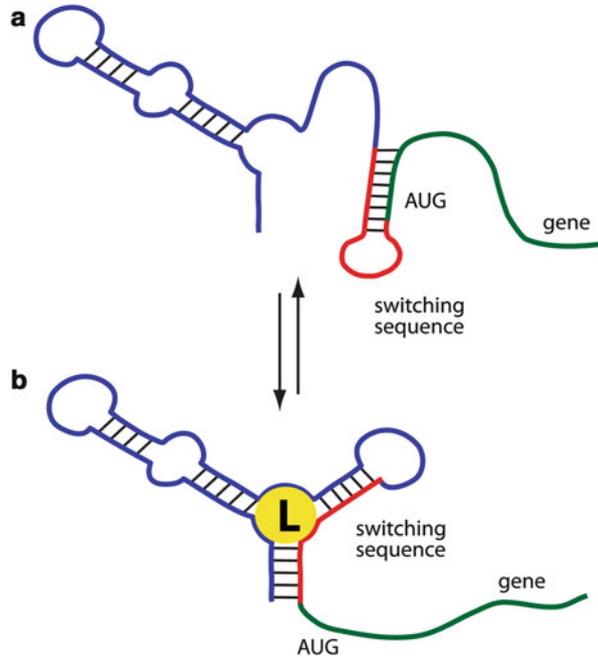
the position of individual elements in the overall architecture. Governed by the Watson–Crick (WC) rules, duplex formation is specifically driven by the presence and location of complementary sequences, which result in constraining the fold into well-defined configurations. Further, nucleotides can engage also in non-canonical pairing and different modular motifs that contribute to the great diversity of RNA structure [11, 12]. These considerations explain why the correct identification of interactions represents an important step towards the structural determination of NA species.

The classic approaches for NA structure determination employ crystallography [13, 14] and nuclear magnetic resonance (NMR) [15, 16], which are capable of providing the 3D structure of macromolecules with atomic resolution. Recent advances in instrumentation and sample preparation methodologies have spurred a very rapid increase of the number of RNA structures deposited in the Protein Data Bank [17]. In particular, the increased availability of synchrotron light and the development of methods for heavy-atom substitution, which facilitate the phasing of diffraction maps, have enabled the solution of crystal structures well in excess of 100 nucleotides (nt) [18, 19]. In a similar fashion, high-field magnets and improved strategies for targeted [13]C- and ^{15}N-labelling have pushed the size limit of NMR to exceed ~100 nt [20], while larger samples have become accessible by combining this technique with small-angle X-ray scattering (SAXS) experiments [21]. The insufficient chemical-shift dispersion associated with the presence of only four fundamental units and the faster relaxation rates typical of large macromolecules

are intrinsic factors that limit the applicability of NMR to these types of biopolymers. Further, technical advances notwithstanding, both crystallography and NMR still pose stringent sample requirements. Amount, purity, solubility, and flexibility present hurdles that may prevent the elucidation of the desired structure. For this reason, the interest in alternative approaches has seen an intense surge, fuelled by the need to obtain valid structural information for intractable samples, or by the desire to complement high-resolution data and extend their reach. At the same time, the development of alternative approaches has been stimulated also by recent advances of computational methods for structure prediction and refinement, which have demonstrated the ability to compensate for the typically lower resolution afforded by them. A close examination of available computational tools can provide a greater understanding of the type of information demanded from the new experimental approaches and a more accurate evaluation of their potential for structure elucidation.

10.2 Computational Approaches for NA Structure Determination

If the central dogma of biology is that DNA contains all the information necessary to orchestrate the full range of cellular functions and activities, then its corollary is that a biopolymer's sequence must determine its fold. Based on this axiom, also known as the biophysics' central dogma, a wide range of computational approaches has been developed to infer NA higher order structure from primary structure information. Since the pioneering work performed in 1973 by Tinoco and coworkers [22], predictive models for RNA secondary structure have been proposed on the basis of thermodynamic calculations, statistical analysis, and knowledge of fundamental properties of NA molecules. The mFold algorithm [23, 24], the Vienna RNA package [25, 26], and the RNA structure suite [27, 28] are examples of tools based on free-energy calculations, which can predict with reasonable accuracy secondary structures of moderate size. The mFold algorithm looks for putative base pairs one at a time and evaluates the stability of all possible folds that contain them. Overall free energy is calculated from the contributions of each base pair, single-base stacking at the end of helices, and mismatched pairs in loops and bulges. The program then allows for the graphic analysis of suboptimal structures with the lowest free energies, which represent the most likely secondary structures folded by the construct of interest (Fig. 10.2). The mFold suite contains also additional tools that utilize thermodynamic calculations to assess the stability of duplexes produced by mutual hybridization of individual strands.

The Vienna RNA package utilizes dynamic programming to combine free-energy minimization with partition function calculations and suboptimal folding predictions. The equilibrium partition method involves calculating in recursive fashion the probability that any possible base pair may occur within a given construct [29]. The pair binding probabilities are then visualized as a full ensemble of possible alternative structures, which can reveal the most probable fold and
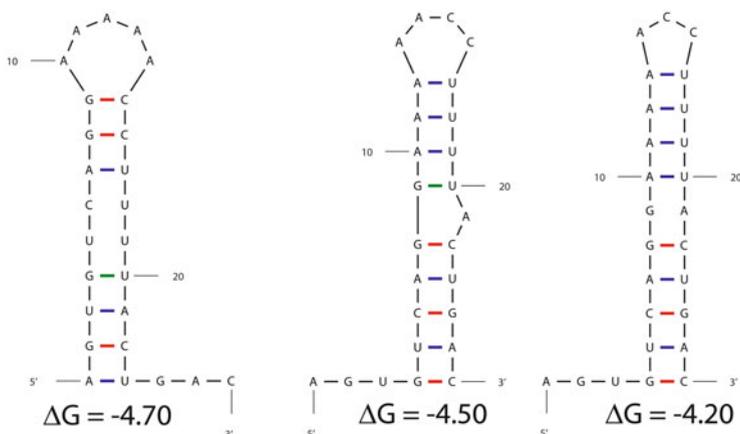
Fig. 10.2 Alternative secondary structures calculated by mFold (http://mfold.rna.albany.edu/) for the following sequence: AGU GUC AGG AAA AAC CUU UUA CUG AC. Calculations were performed for a fixed folding temperature of 37 °C and an environment containing 1 M NaCl

highlight intermediates and regions of structural instability. In contrast, the suboptimal folding algorithm generates all secondary structures between the minimum free energy and a target upper limit [26]. The calculations approximate statistical quantities that account for equilibrium partitioning and structural diversity. The possible states tend to cluster close to the free-energy minimum and, taken together, provide an excellent representation of what the ground state looks like. In a similar fashion, the RNA structure package [27, 28] contains algorithms that perform thermodynamic and partition function calculations, as well as programs for evaluating bimolecular folding/hybridization and supporting the stochastic sampling of putative structures.

Comparative sequence analysis is effectively employed to identify complementary regions involved in possible secondary structures [30, 31]. Sequence alignment algorithms have been developed to reveal highly conserved couples of complementary nucleotides and identify co-variations that preserve putative base pairs (e.g. by replacing a GC-pair with an AU, or vice versa). For example, thousands of sequences accumulated over the years were analysed to obtain complete secondary structures of 16S and 23S rRNA [31]. When the high-resolution structures of 30S and 50S ribosomal subunits became available, they showed that nearly all the conserved and co-variating pairs included in the comparative models were indeed present in the crystal structures [31]. When co-variations are not readily ascribable to secondary structure, they are typically explained as reflecting long-range tertiary interactions. Based on this assumption, a full-fledged 3D model was derived for the conserved core of group I introns by aligning a total of 87 available sequences and applying stereochemical modelling techniques [32]. In many cases, however, non-canonical base pairs complicate the co-variation interplay and lead to blurred predictions. Furthermore, comparative methods require the examination of a large

number of sequences, which makes availability a major limiting factor for these types of approaches.

Alternative strategies have expanded the near-neighbour concept to improve prediction accuracy and overcome the limitations of thermodynamics and statistical calculations in forecasting long-range interactions. The MC-Fold/MC-Sym pipeline [33, 34] utilizes minimum cyclic motifs—groups of three to five nucleotides present in direct spatial contiguity—to build 3D models through an iterative process that considers all possible overlapping cycles in the construct. The cyclic motifs capture the full range of spatial relationships between contiguous nucleotides, including not only base-pairing and stacking information, but also non-canonical pairs and interactions involving phosphoribose functional groups [35]. The 3D structures of representative cyclic motifs were initially extracted from high-resolution structures of large RNA samples. These fundamental elements were then incorporated in the algorithm as modular blocks to be utilized in model building operations [33, 34]. In analogous fashion, knowledge of RNA interactions and architecture has led to recognize numerous recurring motifs that are responsible for shaping RNA's higher-order structure [11, 12]. The Assemble suite [36, 37] employs a library of tertiary modules and long-range interactions, which guide the organization of the underlying secondary structure into full-fledged 3D models. In this case, models are built in interactive fashion by using an integrated graphic user interface (GUI). Simulated annealing and energy minimization can be subsequently performed to refine initial models and produce the final structures.

The development of advanced computational approaches has greatly increased the ability to predict RNA structure to very high levels of accuracy (Fig. 10.3) [37]. On the wings of these developments, an initiative has been recently launched to evaluate de novo prediction capabilities, which is analogous to the CASP competition for proteins [38]. Called RNA-Puzzles, the initiative consists of a blind experiment in 3D structure prediction meant to uncover strengths and weaknesses of the various approaches [39]. The results of the first exercise completed in 2012 served as a reminder of the intrinsic challenges of translating all the factors that govern folding into a comprehensive prediction algorithm. These challenges are expected to increase with the size and complexity of the RNA under consideration. Riboswitches and similar systems, which are capable of adopting different stable structures as a function of ligands, metals, and other environmental conditions, will continue to represent a challenge to computational approaches, regardless of the recent advances. For this reason, it is not surprising that the more refined algorithms offer the ability of incorporating experimental constraints into the model building process. The community is keenly aware of the fact that experimental methods will continue to play critical roles by providing not only validation to predicted structures, but also valuable constraints for model building and refinement.

Fig. 10.3 When available, high-resolution structures are used to assess model accuracy. The root mean square deviation (RMSD) is calculated by minimizing the sum of the distances between corresponding atoms. In this case, the *red* and *green traces* correspond to the model and crystal structure, respectively, of the ribonuclease P specificity domain of *B. subtilis*. Except for the P9 domain, the traces show remarkable overlap (Reproduced with permission from [37])
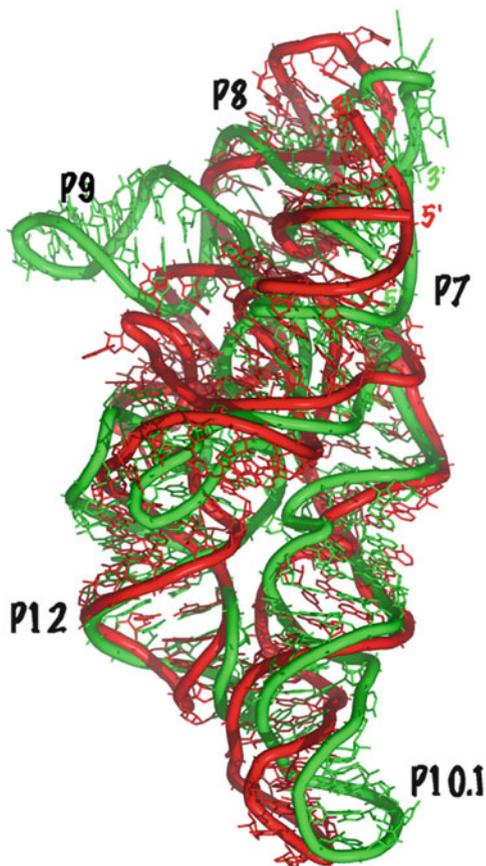
10.3 Structural Probing

Among the possible experimental alternatives, approaches based on chemical probing have arguably benefitted the most from the ability of new algorithms to maximize the dividends of the attainable information and narrow the resolution gap with the classic high-resolution techniques. Excellent versatility, relative ease of use, and absence of major instrumentation requirements are additional characteristics that are sustaining the recent progress of alternative technologies for structure elucidation.

Chemical probing approaches can be classified according to the type of constraint they can provide. The majority of reagents employed in classic footprinting procedures are alkylating agents that modify exposed nucleophilic groups [40–42]. The absence of reactivity is interpreted as steric protection induced by the nucleotide's structural context (Fig. 10.4a). Unfortunately, the mono-functional

nature of such probes precludes the recognition of the protection source. In the case
of cross-linking reagents, the probing reaction establishes a covalent bridge
between nucleotides that are placed within mutual striking distance and, thus,
provides direct information on their spatial relationship (Fig. 10.4b) [43–45]. For
these reasons, the results of footprinting experiments must be interpreted through
the lenses of the Watson–Crick rules to infer the correct composition of base pairs,
whereas those of cross-linking experiments enable the unambiguous pairwise
identification of nucleotides situated in close proximity or direct mutual contact.
The interpretation of the latter is made more straightforward by the elimination of
any inference process.

Mono-functional probes can be distinguished on the basis of their ability to
attack blindly all types of nucleotides or to engage in nucleotide-specific chemistry.
Dimethylsulfate (DMS), β-ethoxy-α-ketobutyraldehyde (kethoxal, KT), and
1-cyclohexyl-3-(2-morpholinoethyl)-carbodiimide   metho-p-toluene   sulfonate
(CMCT) are examples of nucleotide-specific reagents that target, respectively,
N1-adenine and N3-cytidine, N1- and N2-guanine, and N3-uridine (Fig. 10.5)
[46, 47]. Taken together, these reagents reveal the pairing-edge status of all four
nucleotides and are consequently capable of capturing the full range of information
on possible base-pairing interactions. Further, the reaction of DMS with
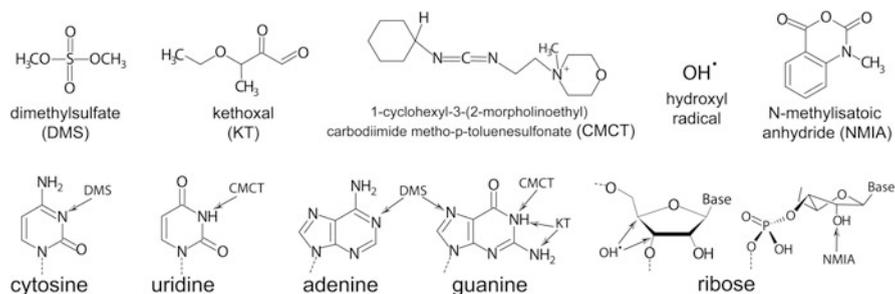
Fig. 10.5 Structures and targets of typical mono-functional probes

N7-guanine is inhibited by typical stacking interactions established in helical motifs (Fig. 10.4a), which indicates the participation of this specific type of nucleotide in RNA secondary structure. Based on the same principles, a variety of reagents have been investigated for their ability to target the different functional groups present on NA molecules. As discussed below, however, the need for probe-specific chemistry to achieve detection is frequently perceived as an obstacle to their broader application in structural biology.

Hydroxyl radicals represent an example of non-discriminating probes that can attack any type of exposed nucleotide (Fig. 10.5) [48, 49]. Produced in situ by either the Fenton reaction between iron(II) EDTA and hydrogen peroxide [50] or synchrotron radiolysis of water [51], hydroxyl radicals induce cleavage of the RNA backbone by abstracting a proton from either the C4′ or C5′ positions of ribose. The attack is possible only when these functional groups are exposed to the solvent, which results in the identification of segments located on the surface of the structure of interest. The ability to perform radiolysis within very short times and to accurately quantify the extent of cleavage has offered the opportunity to study the dynamics of RNA folding [51]. Another reagent employed to gauge local backbone flexibility is *N*-methylisatoic anhydride (NMIA), which modifies the 2′-hydroxyl of ribose units that are not constrained in the C3′-endo conformation (Fig. 10.5) [52, 53]. Given that this puckering conformation is typical of base-paired nucleotides, NMIA can effectively distinguish unreactive double-stranded regions from reactive single-stranded ones. Probe reactivity may be also inhibited in single-stranded regions with flexibility restricted by the fold context, or accessibility limited by buried placement and ligand binding. Therefore, data interpretation requires a great deal of attention to avoid ambiguous assignments.

In a similar fashion, cross-linking strategies are also differentiated according to their base-specific activity or lack thereof. The nitrogenous bases are per se prone to form cross links upon exposure to UV radiation, which is the basis of its mutagenic properties [54]. Without significant base discrimination, this process induces the conjoining of any couple of nucleotides that are within short distance and assume at least transiently the correct geometry [43, 55]. For example, 13 cross-linking sites were observed in *E. coli* 16S rRNA after irradiating 30S ribosomal subunits [56].
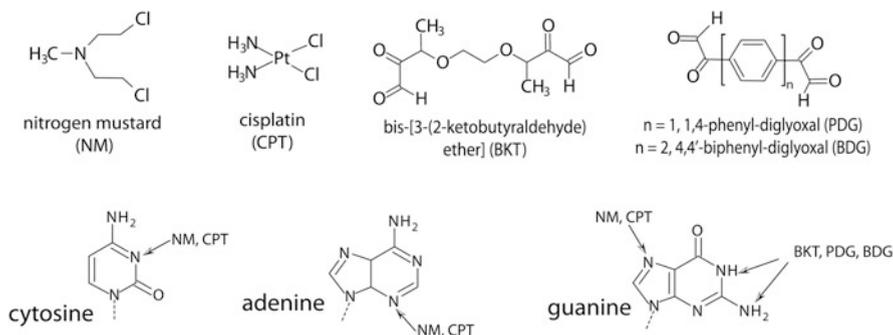
Fig. 10.6 Structures and targets of bifunctional probes

At least eight of them were recognized as involving nucleotides within discrete secondary structures or adjacent junction regions, while the others were attributed to long-distance interactions that constrained the domains architecture. When a more targeted approach is desired, site-specific bridging can be achieved by incorporating modified nucleotides in the construct's sequence, which contain photosensitive labels [57]. While the types of functional groups may be quite different, a common denominator consists of the ability to obtain localized cross-linking by utilizing longer wavelengths that are less likely to induce indiscriminate damage by hydroxyl-radical reaction.

With or without photoaffinity labels, UV cross-linking produces zero-length conjugates between structures that must lay very close or establish direct contact to fulfil the stringent proximity requirements posed by photochemical reactions. The utilization of bifunctional alkylating agents, in which two reactive groups are separated by a spacer, enables the identification of structures that are not necessarily in mutual contact, but are situated within well-defined distances. Examples of bifunctional probes include nitrogen mustard (NM) [58], bikethoxal (BKT) [59], and cisplatin (CPT) [60] (Fig. 10.6), which can bridge functional groups placed, respectively, 9.5 ± 1.5, 5.4 ± 1.0, and 3.2 Å apart [45]. The conformational flexibility of spacers that frequently exhibit numerous degrees of freedom is reflected in the uncertainty associated with the average cross-linking distance. The more narrowly defined distance afforded by cisplatin is consistent with a bridging structure that involves direct metal coordination. Although spacer flexibility can be readily dealt with during simulated annealing and energy minimization, more rigid spacers have been evaluated for their narrower uncertainty ranges. The concept is illustrated, for example, by the nested bikethoxal analogues 1,4-phenyl-diglyoxal and 4,4′-biphenyl-diglyoxal. These probes contain two identical 1,2-dicarbonyl functions spaced by one or two phenyl groups, which are capable of cross-linking guanines placed 6.14 ± 0.64 and 10.44 ± 0.80 Å apart (Fig. 10.6) [61]. A direct comparison between 1,4-phenyl-diglyoxal and bikethoxal shows that substituting an aliphatic with an aromatic spacer reduced significantly the uncertainty range. Further, the modular design employed for these probes demonstrated the ability of

nested analogues to act as "molecular rulers" and cover progressively greater distances between susceptible targets.

Probe selection must account not only for the type of constraint sought for a certain task, but also for practical considerations dictated by the detection platform. In general, mono-functional probes rely on the analysis of oligonucleotide ladders that are produced by either strand cleavage or polymerization inhibition at the site of modification. The products are typically analysed by electrophoresis to determine their size and sequence, which reveal the position of the modification. Nucleotide-specific chemistry is frequently necessary to achieve strand cleavage at the modification site. In the case of DMS footprinting, hydrazinolysis is employed to achieve base loss and backbone hydrolysis at methylated C and T, whereas alkaline treatment is used for G, and mild acid followed by alkaline hydrolysis for A [40]. Cleaved products may be analysed immediately or after amplification by polymerase chain reaction (PCR) with ad hoc primers. PCR-like strategies are used also in the absence of strand-cleavage procedures, when the probe modification is bulky enough to inhibit primer extension by the specific polymerase or reverse transcriptase employed in the reaction. This case is exemplified by NMIA that is the probe used in selective 2′-hydroxyl acylation by primer extension (SHAPE) [52, 53]. The readout afforded by electrophoretic methods can achieve single-nucleotide resolution for lengths of up to 1,000 nts. The fact that the same modified oligonucleotide can serve as a template for multiple copy cycles results in amplification of the corresponding signal. The utilization of appropriately labelled primers may enable also single-molecule detection, which is at the basis of recently developed high-throughput sequencing techniques. The broad availability of software packages for data interpretation and quantification has minimized major bottlenecks in the structural probing workflow. The recent advances of computational approaches have greatly increased the appeal of structural probing, which combine a relatively straightforward design with the ability to pursue ex vivo and in cell applications.

10.4   MS Detection for Structural Probing

Mass spectrometry (MS) has the potential for accessing all possible levels of NA higher order structure. The advent of soft ionization techniques, such as electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), has facilitated the analysis of progressively larger NA samples, ranging from fundamental units, such as nucleosides and nucleotides [62, 63], to entire plasmids and ribosomal assemblies [64, 65]. Both solution and gas-phase approaches have been developed to elucidate the primary structure of these biopolymers. The former involve the mass determination of oligonucleotide ladders from strand cleavage or polymerization termination [66, 67], in analogy with methods based on electrophoretic detection. The latter, instead, employ tandem mass spectrometry (MS/MS) to obtain backbone fragmentation of mass-selected oligonucleotides, which generates characteristic ion series consistent with their

sequence [68, 69]. A variety of gas-phase activation methods have been effectively employed to perform sequencing of both unmodified and modified oligonucleotides, as discussed in depth in Chap. 6. Larger NA targets are typically addressed by performing controlled nuclease digestion to obtain mixtures of hydrolytic products, which can be subsequently mass mapped and sequenced [70–72]. Alternative to these types of bottom-up approaches, top-down strategies have been explored by completing the direct MS/MS analysis of large NA samples with no hydrolytic procedures [73, 74]. These capabilities garnered a great deal of attention at the inception of the Human Genome Project, when a variety of analytical platforms were carefully evaluated for taking on the challenge [75]. The choice fell on methods based on strand amplification and electrophoretic analysis by virtue of their higher throughput and lower cost per sequenced base [76].

MS methods can be effectively applied to characterize oligonucleotide ladders obtained by structural probing, in ways that mimic the role of electrophoresis in sequence analysis. Indeed, MALDI-MS was used to evaluate ladders that were produced by DMS methylation/cleavage for sequencing purposes, but could have well represented the products of probing reactions [77]. A valuable characteristic of MS approaches, however, which sets them apart from those based on electrophoresis, is the intrinsic ability to identify covalent NA modifications with no need for strand cleavage at modified sites. Thanks to this capability, MS has contributed to the elucidation of over one hundred modified ribonucleotides catalogued in the RNA Modifications Database [78, 79], in addition to many deoxyribonucleotide modifications produced by endogenous processes and environmental exposure, which are implicated in DNA damage [80, 81]. Typical primer-extension methods are severely hampered by the absence of nucleotides that are complementary to the target modification and may be therefore capable of sustaining strand amplification. Even when the modifier is capable of halting polymerization, the size/sequence information afforded by electrophoretic methods cannot provide the identity of the modification, which must be pursued by other means. In contrast, no such limitations exist for MS approaches, which can identify virtually any type of modifier from its characteristic mass shift and, in many cases, isotopic signature.

In structural biology applications, this capability translates into the ability of detecting probe modifications regardless of chemical and spectroscopic properties. No nucleotide-specific chemistry is required to induce strand cleavage at the modified site. Instead, the utilization of a broad range of endonucleases is usually preferred when bottom-up strategies call for hydrolytic procedures to obtain smaller oligonucleotides for mass mapping/sequencing [70–72]. Product characterization is also possible in top-down fashion by activating the gas-phase dissociation of mass-selected precursors [82]. This strategy has been combined with NMIA to achieve what has been called 2′-hydroxyl acylation by MS (SHAMS) [83]. In the SHAMS investigation of HIV-1 polypurine tract (PPT), a DNA–RNA hybrid involved in the reverse transcription of viral genome, the number of susceptible nucleotides was initially monitored by direct infusion ESI-MS to highlight regions of local flexibility. The position of modifications introduced by NMIA reaction was subsequently obtained by activating the fragmentation of the selected adducts. Analogous to

SHAPE, this technique represents a valuable alternative for the investigation of strands that are too short for supporting the stable annealing of ad hoc primers [83].

The combination of MS detection with structural probing has led to the development of so-called MS3D strategies for structure determination [84]. The widespread utilization of this analytical platform has spurred a major expansion of the probe repertoire by facilitating the utilization of additional bifunctional cross linkers. In fact, the characterization of typical cross-linked products requires the simultaneous sequencing of two individual oligonucleotides that are connected by a covalent bridge. This task represents a challenge to electrophoretic platforms that may be faced with reading overlapped oligonucleotide ladders. On the contrary, MS techniques offer the opportunity to combine mass mapping with gas-phase sequencing, which together enable the complete characterization of cross-linked moieties and correct identification of bridged nucleotides. Furthermore, the fact that the MS platform is readily applicable to oligonucleotide–peptide conjugates has opened the door for the utilization of bifunctional reagents to probe ribonucleoproteins and identify the points of contact between bound components [85, 86].

In analogy with footprinting experiments, the products of bifunctional cross-linking can be effectively characterized by either bottom-up or top-down approaches. For example, the former was used to characterize the products obtained by probing HIV-1 stemloop-1 (SL1) with cisplatin and NM [45]. Mass mapping of digestion mixtures obtained by treatment with either RNase A or T1 provided an excellent measure of the ability of these reagents to bridge nucleotides within reciprocal range. In particular, NM was found to cross-link Gs and As in the single-stranded loop and bulge regions (Fig. 10.7), which were approximately 9.5 ± 1.5 Å from one another. Consistent with its shorter span of 3.2 Å, cisplatin produced cross links between adjacent Gs in both single- and double-stranded regions. The MS determinations showed also that an initial alkylation step was not always followed by a bridging reaction, if a second susceptible group was not present within range. This type of situation results in mono-functional adducts with a "dangling" arm that may be hydrolyzed by the solvent. In the case of NM, incomplete cross links exhibit either 2-chloroethyl- or 2-hydroxyethyl- moieties with characteristic mass shifts. On the one hand, the formation of incomplete conjugates tends to increase the complexity of digestion mixtures to be mass mapped. On the other hand, however, dangling products afford the same type of constraint afforded by mono-functional agents, which relates to the level of solvent accessibility exhibited by the modified nucleotide. These types of constraints, which are accessible only to the MS platform, contribute indirectly to the high information content associated with bifunctional probes.

The ability to establish stable bonds with susceptible nucleotides represents a common denominator between mono-functional and cross-linking probes. The irreversible nature of these covalent modifications is crucial to enable their characterization by either MS or electrophoretic methods. Conversely, the need to withstand the detection process represents a major obstacle to the utilization of reversible modifications as possible structural probes. The ability of ESI to preserve non-covalent interactions during MS analysis has prompted us to explore the

Fig. 10.7 Cross-linking maps of HIV-1 SL1 after probing with (**a**) cisplatin and (**b**) nitrogen mustard. Products obtained by RNase A hydrolysis are labelled by their first and last nucleotides separated by a *colon*. Cyclic phosphate is indicated by the "*greater than*" symbol [44]
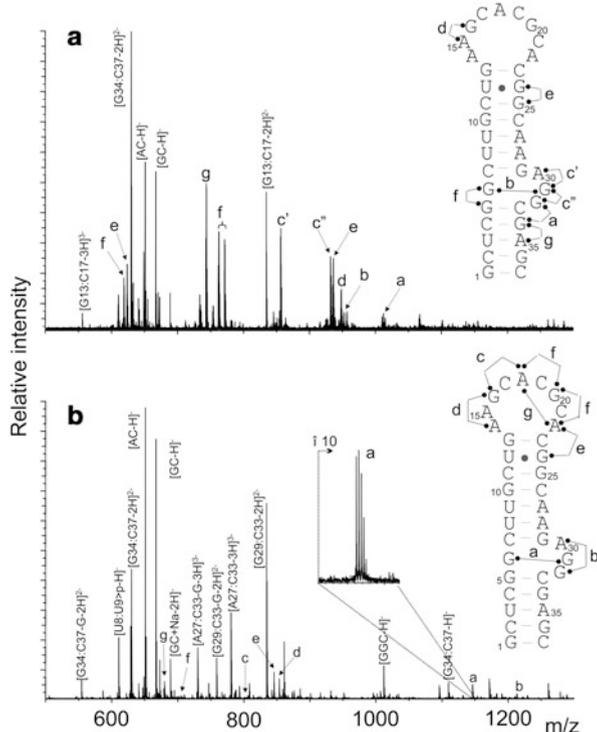
utilization of typical NA ligands to detect the presence of specific structural elements [87–89]. Indeed, the gentle character of this ionization technique allows for reversible non-covalent complexes to be observed intact in the gas phase [90, 91]. We took advantage of this favourable characteristic to evaluate the binding of general intercalators, minor-groove binders, mixed-mode intercalator/groove binders, and multifunctional polycationic aminoglycosides (Fig. 10.8). Owing to their well-characterized tropism for specific NA features, these ligands can report respectively on base stacking patterns, groove depth and geometry, and the presence of electronegative pockets associated with bulging nucleobases, backbone kinks, and helix anomalies. Upon mixing the ligands with target substrates, the stoichiometry and relative affinity afforded by MS analysis indicated the presence/state of the corresponding structural element. In some instances, it was possible to employ MS/MS to identify nucleotides making direct contact with the ligand, which enabled us to map the location of the corresponding motif [87–89]. This type of experiment, which could be defined as gas-phase footprinting, was made possible by the ability of stronger interactions to survive intact the activation process (Fig. 10.9) [92]. The information afforded by non-covalent probes provided valuable insights into anomalous pairing interactions within the HIV-1 and Ty3 PPT complexes [87, 88] and the structural rearrangements that mediate the dimerization and isomerization processes of HIV-1 SL1 domain [89].
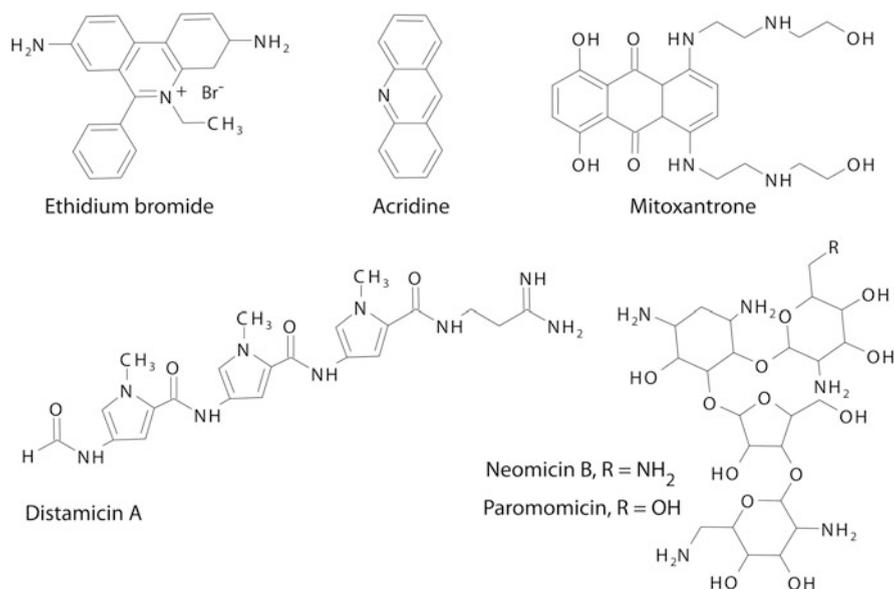
Fig. 10.8 Archetypical NA ligands explored as possible non-covalent structural probes. Ethidium bromide and acridine are intercalators. Distamycin A is a minor-groove binder. Mitoxantrone is a mixed-mode intercalator/groove binder. Neomycin B and paromomycin are polycationic aminoglycosides

10.5   MS-Based Structure Elucidation

Experimental design is expected to have significant consequences on the ability to obtain the spatial constraints necessary to generate meaningful 3D structures. In the case of structural probing, the actual reaction is typically completed before any analysis is carried out to identify the location of modified nucleotides. The fact that the sought-after information is irreversibly captured in a unique modification pattern dispenses with the need to preserve the native fold during subsequent analysis. This observation affords a great deal of flexibility in terms of solution composition and environmental conditions (i.e. temperature, pH, ionic strength, type of buffer, presence of metals, cofactors, etc.), which must ensure structure integrity rather than direct compatibility with the selected analytical technique. Decoupling the probing reaction ensures that, regardless the platform, the conditions can be adjusted to faithfully capture the structure of the sample in solution. At the same time, however, different platforms may afford unique capabilities that should be considered when designing and executing probing experiments. For example, the mass determination of entire probing mixtures affords information that may not be directly accessible to other detection techniques, such as the number of hits sustained by the probed material. This type of information offers the possibility of monitoring the effects of environmental
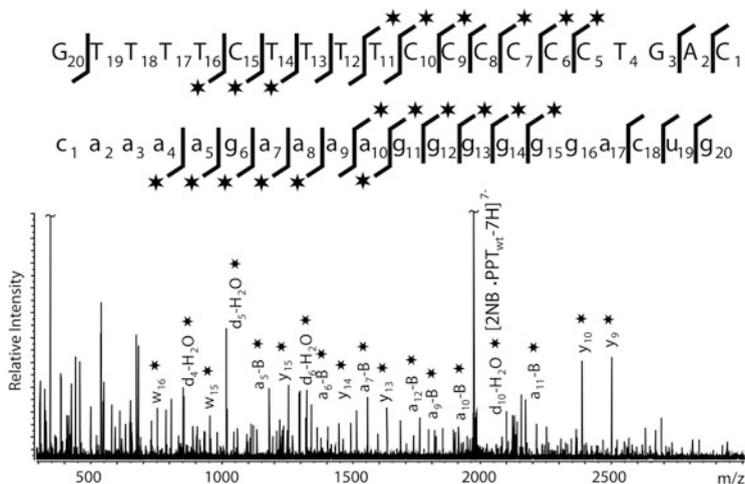
Fig. 10.9 MS/MS spectrum of the 1:2 complex between HIV-1 PPT and neomycin B. For clarity, only some of the characteristic ion series were labelled on the spectrum. –B indicates loss of nucleobase at the cleavage site. The *asterisk* identifies products containing non-covalently bound neomycin B. All product ions observed for the DNA and RNA strands of PPT were reported on the respective sequences to highlight nucleotides that were protected from fragmenting by the presence of bound ligand [91]. Two distinct gaps are evident in the ion series observed for the DNA and RNA components of the construct, which revealed the presence of two separate binding sites [86]

parameters on the stability of the system under consideration, which facilitates the optimization of probing conditions and the investigation of structure dynamics. This possibility is illustrated by the footprinting investigation of individual stemloop domains of the HIV-1 packaging signal (Ψ-RNA) [47, 82]. When stemloop-2 (SL2) was treated at room temperature with the U/G-specific probe CMCT, up to three stable adducts were observed as discrete products spaced by 251.2 Da (Fig. 10.10a). This outcome was consistent with the presence of one U and three G nucleotides in the single-stranded loop of the stemloop structure. The number of hits increased to seven when the reaction was repeated at 85 °C under otherwise identical conditions, suggesting that melting of the double-stranded stem had induced the exposure of additional susceptible targets (Fig. 10.10b). This explanation was corroborated by the position of modified nucleotides, which was obtained through top-down experiments involving direct gas-phase activation of mass-selected adducts [82]. The results clearly indicated that the low-temperature adducts were localized exclusively in the single-stranded loop, whereas the high-temperature ones were distributed over the entire sequence, as expected from the presence of an unstructured species in solution.

These types of experiments can be employed not only to identify the optimal temperature for probing a native fold, but also to investigate the dynamics of the system of interest. In the case of protein samples, structure dynamics can be effectively evaluated by performing hydrogen–deuterium exchange (HDX) to
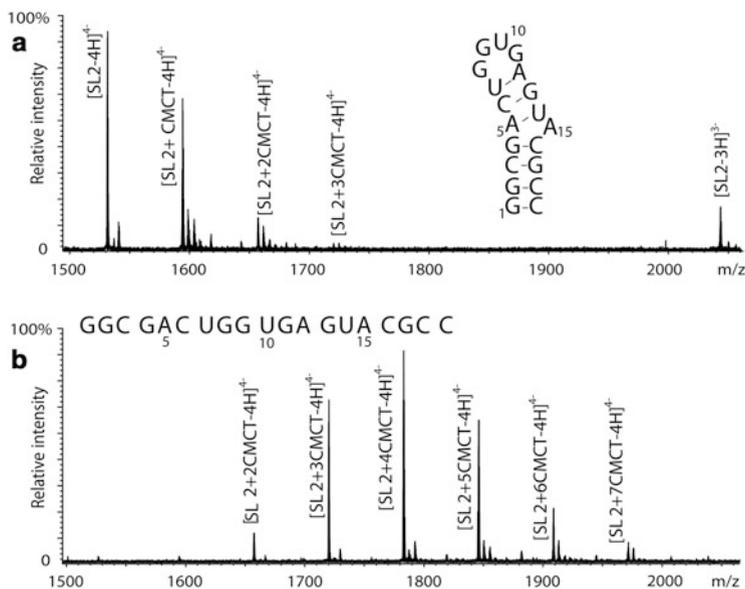
Fig. 10.10 Effects of heat denaturation on base protection ESI-MS spectrum of HIV-1 SL2 treated with CMCT at (**a**) 25 and (**b**) 85 °C under identical conditions. Both reactions were allowed to proceed until no further increase in the number of hits was detected (Adapted with permission from [81])

identify regions that may be transiently exposed to the solvent [93, 94]. Unfortunately, the fast rate of proton exchange characteristic of NA molecules in solution hampers their investigation by this technique [95, 96]. Alternatively, footprinting agents may be utilized to induce irreversible alkylation and capture the transient exposure of NA structures without the exchange problems associated with HDX. In the investigation of full-length packaging signal (Ψ-RNA) of HIV-1 genome, the mono-functional probes DMS, KT, and CMCT were applied in concert to verify the formation of discrete stemloop domains under various refolding protocols and in the presence of different magnesium(II) concentrations [97]. The footprinting maps clearly pointed out base-paired nucleotides at the ends of helical structures or near internal bulges, which were prone to prominent dynamics substantiated by transient dissociation or "breathing."

Monitoring the number of hits as a function of reagent concentration represents an excellent strategy for diagnosing possible artefacts. A common source of ambiguities is encountered when footprinting alkylation of exposed nucleotides induces conformational changes that expose previously protected sites. The outcome may be an inflated number of modifications that do not reflect the native fold but provide a distorted picture of the structure of interest. In a similar fashion, the observed number of cross links may artificially increase when functional groups that spend only little time within striking distance are inadvertently bridged, thus facilitating the cross-linking of additional sites on the now trapped conformation.

These types of situations are generally limited by employing low probe-to-substrate ratios that ensure the statistical introduction of only one modification per substrate molecule (i.e. single-hit probing). MS determinations offer the opportunity to work with progressively higher ratios, following titration schemes in which any sudden increase of the number of hits is attributed to unwanted distortion or kinetic traps. In this way, any constraints obtained at ratios that led to questionable increases can be discarded as possible artefacts of probe-induced conformational changes. In contrast, those obtained at lower ratios can be safely utilized in modelling operations, regardless of whether they were produced under single-hit conditions. These types of strategies allow one to enhance the yield of probing reactions, which facilitates product characterization, without compromising the validity of the information attained.

Experimental constraints obtained from structural probes can support different modelling strategies. In the absence of additional sources of information, the data afforded by MS-based approaches were utilized to generate full-fledged models of retroviral pseudoknots by using the constraint satisfaction algorithm MC-Sym [98]. The program started from the sequence of interest and general NA information, such as chemical structure, bond distances, torsion and dihedral angles, etc., which is hard-coded in the cyclic motif modules. Actual experimental data, such as base-pair composition and distance constraints from structural probes, were then combined to generate all-atom models. As the program progressed through the sequence by computing all possible cyclic motifs, the experimental constraints were used to more narrowly define the spatial position of nucleotides and strands. The resulting models were subjected to several rounds of energy minimization to eliminate any steric clashes and backbone strains. In the case of the ribosomal frameshifting pseudoknot of mouse mammary tumour virus (Fig. 10.11), the final coordinates were compared atom by atom with those of the corresponding high-resolution NMR structure. The comparison afforded an average 3 Å root mean square deviation (rmsd), which provided a measure of the excellent quality of the final model. In the same study, the absence of a high-resolution structure for the feline immunodeficiency virus pseudoknot did not allow for direct comparisons, but made this sample the first unknown RNA structure ever solved by MS-based approaches [98].

A different strategy was implemented to solve the structure of HIV-1 Ψ-RNA, which could count on the availability of high-resolution coordinates for separate sections of its full-length 118 nt sequence [97]. As taught us by the discovery of riboswitches, the larger an RNA molecule, the greater is the probability that different pairing arrangements may lead to alternative folds with comparable stabilities. For this reason, mono-functional probes were employed to verify the consensus secondary structure that had been proposed on the basis of phylogenetic and mutation analysis. The fact that these preliminary experiments confirmed the formation of all the expected subdomains enabled the utilization of their high-resolution coordinates obtained separately by different groups. These building blocks were combined with connecting linkers that, in the absence of high-resolution structures, were generated by MC-Sym from sequence information.
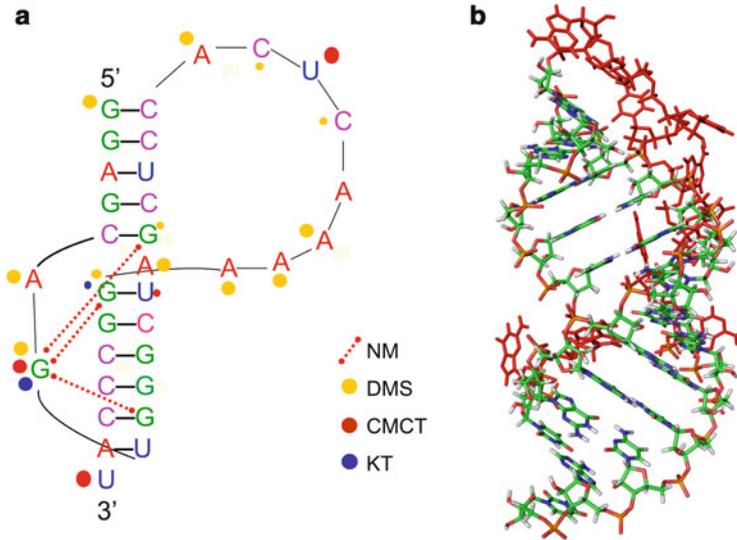
Fig. 10.11 (**a**) Two-dimensional map of mouse mammary tumour virus frameshifting pseudoknot, showing the probed nucleotides. The size of the dots is proportional to the extent of modifications. (**b**) All-atom 3D structure obtained by MC-Sym. Single-stranded loops are marked in *red* [97]

Mimicking the hierarchic nature of RNA folding, the geometric relationships between blocks and the overall fold architecture were established by using probing information, which was accomplished by subjecting the initial model to simulated annealing restrained by cross-link distances. After rounds of energy minimization, the final structure retained the atomic resolution of the initial building blocks and matched the experimental constraints afforded by both footprinting and bifunctional reagents (Fig. 10.12). Also in this case, the absence of NMR/crystal structures for Ψ-RNA made it impossible to gauge the quality of the model by completing a direct comparison. Instead, a close examination of the model itself provided the elements necessary to assess its validity. Indeed, the modelling operations had accounted for the numerous cross links between the SL1 and SL4 domains by translating them into a long-range tertiary interaction, which was deemed responsible for stabilizing the rather compact topology assumed by Ψ-RNA. These cross links were eliminated when the loop of SL4 was mutated by substituting a critical nucleotide. The fact that the model had enabled valid predictions of structural determinants offered validation to the probing and modelling approaches employed in the project [97].

The increasing power of structure prediction algorithms could enable the potential implementation of yet different modelling strategies, in which experimental constraints could guide the search for the best possible match within libraries of theoretical models. Strategies can be envisioned in which any of the algorithms discussed earlier—based on sequence alignment, thermodynamics, statistical, and
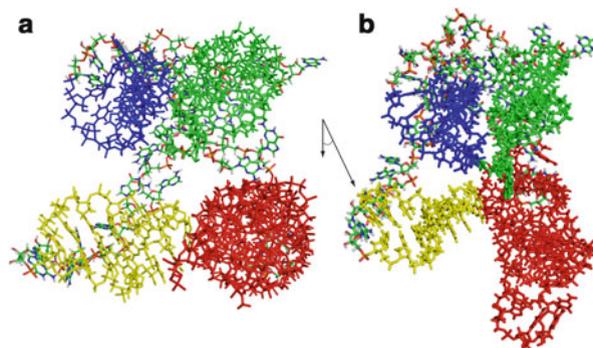
Fig. 10.12 (**a**) *Top* and (**b**) *rotated view* of all-atom structure of HIV-1 packaging signal (Ψ-RNA). The SL1 domain is marked in *red*, SL2 in *green*, SL3 in *blue*, and SL4 in *yellow*. The *top view* shows that the stemloops exhibit nearly parallel axes, with the exception of SL4 that points towards the SL1 stem. The *rotated view* shows a long-range interaction between SL4 and SL1 [96]

knowledge-based principles—is employed to generate large libraries of putative models, or decoys, from the sequence of interest. Constraints from structural probing are then applied as filters to reject decoys that are incompatible with the experimental observations. A scoring function is then calculated to evaluate the degree of fitness between hits and selection conditions, which could be further improved through simulated annealing restrained by probe distances. In order for these strategies to be effective, the libraries would have to cover the widest possible conformational space to increase the probability of matching the greatest possible number of experimental constraints. It is expected that the combinatorial nature of typical filtering processes, in which independent conditions have multiplying rather than summing effects, would result in rather stringent selections with only a limited number of constraints, thus maximizing the impact of probing experiments. However, it is also expected that further model refinement will still remain a data-intensive operation resting on the broad application of structural probes.

10.6   MS Techniques as Direct Source of Structural Information

As a versatile readout for probing reactions, MS techniques play a rather indirect role in NA structure elucidation. A more direct and perhaps larger role can be prognosticated on the basis of the ability to observe base pairing and other weak interactions in the gas phase. Historically, short duplexes of 8–20 base pairs (bp) were among the first examples of non-covalent complexes detected intact by ESI-MS, which highlighted the soft character of this ionization technique [99, 100]. This valuable feature allowed us to investigate the mechanism of dimerization and isomerization of the SL1 stemloop domain, which establishes a critical quaternary interaction between like copies of HIV-1 genomic RNA [101]. In this system,

annealing of palindromic loop sequences induces the formation of a 6-bp duplex to obtain a loop–loop kissing complex (Fig. 10.13). Intervention of the viral nucleo-capsid (NC) protein leads to an intermolecular rearrangement of the double-stranded stems to form a 28-bp extended duplex dimer. The ability to observe both types of complexes without dissociation was demonstrated by using conformer-specific mutants that could not undergo structure interconversion. This experiment enabled the application of collisional activation to discriminate the conformers according to the stability imparted by the different number of inter-molecular pairs (i.e. 6 bp for the kissing complex versus 28 bp for the duplex). More recently, we have demonstrated that two intermolecular pairs are sufficient to stabilize kissing complexes not only in the gas phase, but also in solution, when the paired nucleotides are flanked by unpaired purines (Fig. 10.14) [102]. These conclusions were reached by using ESI analysis to monitor the extent of dimeri-zation exhibited by mutant constructs in which the flanking bases were systemati-cally varied. Mechanical pulling (i.e. optical tweezers) experiments and computational analysis helped demonstrate that the stacked flanking purines were essential to stabilize the two base-pair duplex.

Under selected conditions, base-pairing interactions have been shown to with-stand different types of activation processes employed in MS/MS experiments. For example, it was demonstrated that collisional activation of double-stranded precursors could produce either duplex dissociation with detection of individual strands, or partial opening of intermolecular pairs with fragmentation of the unzipped regions [103]. The partitioning between pathways was influenced by the kinetics of the activation event, in ways that suggested a concerted unzipping process involving multiple transitions states. In addition to determining the mecha-nism of strand dissociation, the stability of pairing interactions in the gas phase has been credited as the possible cause for the limited sequence information attainable from structured samples upon activation by electron-based techniques, such as electron capture dissociation (ECD) and electron detachment dissociation (EDD) [104, 105], as discussed in detail in Chaps. 6 and 7. Undesirable for sequencing purposes, this effect was minimized by introducing a pre-activation step to achieve base-pair dissociation before any radical ion chemistry initiated [105]. Following this strategy, top-down experiments that combined collisional pre-activation with EDD enabled gas-phase sequencing of structured samples up to 76-nt long with nearly complete sequence coverage [106].

These observations prompted us to explore the possibility of combining different activation techniques to discriminate between single- and double-stranded regions of structured ions. We noted that, under selected experimental conditions, stemloop constructs displayed similar fragmentation patterns whether they were produced by radical dissociation processes induced by ECD [104] and EDD [105] or closed-shell ion fragmentation activated by collision-induced dissociation (CID) [107] and infrared multiphoton dissociation (IRMPD) [108]. In particular, the observed fragments could be attributed almost exclusively to the single-stranded loops of precursor ions [109]. At the same time, the extent of fragmentation was increased by factors that are known to decrease the stability of double-stranded stems,

Fig. 10.13 Dimerization of HIV-1 SL1 in a loop–loop kissing complex, followed by NC-mediated isomerization into an extended duplex structure. In identical MS/MS experiments, the former dissociated into individual stemloops, whereas the latter remained unaffected [100]

Fig. 10.14 Mass spectrum of stemloop complexes. The H3 "light" and H4 "heavy" stemloops were mutated to eliminate stem complementarity and prevent their annealing into intermolecular duplexes. Therefore, while H3·H3 and H4·H4 could potentially adopt both kissing and duplex conformations, the H3·H4 dimer can only exist as a kissing complex (Adapted with permission from [101])

Fig. 10.15 IRMPD product ion spectra obtained from FIV pseudoknot. *Panels* (**a**) and (**b**) show different regions of the same spectrum with increased magnification. The *letter p* indicates a phosphate group; *multiplication sign*, a base-loss product. In the *inset*, *arrows* point towards the end contained by a fragment. *Absence of arrow* indicates internal fragments that do not contain construct ends (Adapted with permission from [108])

including greater A-T/U content, shorter length, and higher charge state. The enhancing effects were replicated by performing gentle collisional activation immediately after IRMPD to induce the dissociation of any weak interaction that had survived the irradiation event. The results clearly indicated that IRMPD had indeed achieved widespread cleavage of the precursor backbone, but failed to dissociate underlying base-pairing interactions that, in the absence of post-activation, prevented the detection of backbone fragments as stand-alone species. These types of experiments showed that IRMPD could be finely tuned to correctly distinguish the single-stranded loops of stemloop and pseudoknot structures from their intact double-stranded stems (Fig. 10.15) [109]. Considering the importance of base pairing in defining NA higher order structure, these fragmentation strategies

could become a valuable tool for structure elucidation. Indeed, a greater under-standing of the rules that govern the gas-phase dissociation of recurring elements of tertiary structure could lead to the broader utilization of these techniques to obtain experimental constraints for modelling operations.

The recent commercialization of ion mobility spectrometry (IMS) mass spectrometers has opened new avenues for obtaining direct information on NA structure and dynamics, as described in detail in Chap. 3. This technique can assess the size/shape of ions from the amount of time it takes them to cross a low-pressure region of the instrument, which may be subjected to a continuous [110, 111] or oscillating [112, 113] electric field. The relationship between an ion's collisional cross section (CCS) and its arrival time is such that these types of determinations can provide valuable information on the overall fold of macromolecular structures [114]. IMS determinations have demonstrated that B-helices typical of DNA are conserved in the gas phase, under appropriate conditions [115, 116]. Similar experiments were also employed to investigate elements defined by helical regions, such as stemloop hairpins, pseudoknots, and cruciform [117, 118], as well as structures that do not rely on base-pairing interactions, such as G-quadruplexes [119, 120] and ligand complexes [117, 118]. These studies employed well-characterized systems to help establish firm correlations between experimental observations and specific structural features. The possibility of extracting actual constraints for modelling operations will hinge on the ability to recognize such features within the complex data provided by progressively larger samples of unknown structure.

10.7 Conclusion and Outlook

Perhaps more than any other type of biopolymer, NA exhibit specific features that favour the application of MS approaches for structural determination. The Watson–Crick rules identify pairwise relationships between nucleotides with a level of confidence that cannot be matched by any known rule describing amino acid interactions. Further, base pairing is more meaningful and defining for NA higher order structure than any single type of interaction might be for protein structure. Therefore, it could be argued that the information afforded by footprinting experiments may go a little farther for NA than protein substrates. From a pragmatic point of view, it certainly helps that convenient reagents are available for each of the fundamental nucleobases, whereas not all amino acid side chains can be probed in practical fashion. This fact ensures that NA footprinting experiments can provide comprehensive information on the base-pairing state of all nucleotides in the target structure.

MS technologies are uniquely poised to take structural probing to full fruition by exceeding the capabilities of typical readout methods based on strand amplification and electrophoresis. On the one hand, the ability of this platform to support multiplexed probe application presents the opportunity of maximizing the infor-mation attainable by individual experiments [121]. On the other hand, its versatility

will be expected to reawaken the interest in cross-linking probes as powerful tools for obtaining direct distance constraints. In addition, the advances of computational approaches for the effective utilization of sparse experimental constraints are already contributing in significant ways to the rapid diffusion of structural probing methodologies, including those that rely on MS detection.

The stability of base-pairing interactions under selected conditions has opened the door for the investigation of NA higher order structure in the gas phase. Ion activation and mobility approaches have clearly demonstrated the possibility of studying determinant features of structural elements that rely on pairing interactions, such as duplexes, stemloops, cruciforms, and pseudoknots. The evidence supports excellent correlation between solution and gas-phase observations, which stems from the fact that hydrogen bonding and base stacking thrive in solvent-free environment. This apparent memory effect offers a strong rationale for employing direct MS analysis for structure elucidation purposes. It remains to be seen whether these types of analyses may be capable of moving beyond the investigation of isolated structural elements to provide the level of information necessary to elucidate complex biological systems. The modular nature of NA structure and the demonstrated ability to observe quaternary interactions, such as those present in kissing complexes, are encouraging facts that may make this proposition possible.

In this direction, the acquisition of valid constraints for modelling operations could greatly benefit from the development of robust gas-phase strategies for the unambiguous characterization of base pairs. Ideal strategies would enable one to map base pairs onto target NA strands in the same way that post-translational modifications are typically mapped onto proteins. The information contained in such maps would represent an excellent basis for generating full-fledged models by a variety of structure-prediction algorithms. At the same time, the field could greatly benefit from a better understanding of the correlation between the mobility profiles observed by IMS determinations and the underlying structural determinants. In particular, learning how to recognize unique signatures corresponding to specific structural elements would enable one to tease out the contributions of individual modules to the overall substrate topology, thus providing important clues about their organization. Based on this knowledge, IMS determinations could be employed also to validate the outcome of modelling operations by comparing theoretical and observed values of CCS or to guide the selection of the model with the best possible fit among a large library of plausible decoys. Realizing the potential of these approaches will help establish mass spectrometry at the forefront of the alternative techniques employed for the elucidation of structure and dynamics of nucleic acid samples.

10.8   Summary of Key Concepts

- The fact that less than 2 % of the human genome codes for proteins highlights the need for more than sequence information to understand their function.
- Directly or indirectly, the MS platform can sustain the investigation of all levels of NA structure—from primary to quaternary—as a possible complement to classic high-resolution techniques.
- Owing to their excellent versatility, MS technologies are rapidly revolutionizing the application of structural probes to the elucidation of NA structures in solution.
- Concomitant advances of computational methods have enabled a more effective utilization of the sparse distance constraints afforded by structural probes, closing the resolution gap typical of probing approaches.
- MS techniques offer the opportunity to investigate tertiary and quaternary structure by virtue of their ability to detect base pairing and other weak interactions in the gas phase.
- IMS promises to open new avenues for observing the effects of such interactions on the global architecture of progressively larger NA structures.

References

1. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) Cell 31: 147–157
2. Gesteland RF, Cech TR, Atkins JF (2006) The RNA world: the nature of modern RNA suggests a prebiotic RNA world. CSHL Press, New York, NY
3. Winkler W, Nahvi A, Breaker RR (2002) Nature 419:952–956
4. Nudler E, Mironov AS (2004) Trends Biochem Sci 29:11–17
5. Mattick JS, Makunin IV (2006) Hum Mol Genet 15 Spec No 1:R17–R29
6. Pheasant M, Mattick JS (2007) Genome Res 17:1245–1253
7. Brion P, Westhof E (1997) Annu Rev Biophys Biomol Struct 26:113–137
8. Draper DE (2013) Biopolymers 99(12):1105–1113
9. Bloomfied VA, Crothers DM, Tinoco I Jr (2000) Nucleic acids—structures, properties, and functions, 1st edn. University Science, Sausalito, CA
10. Draper DE (1999) J Mol Biol 293:255–270
11. Leontis NB, Westhof E (2001) RNA 7:499–512
12. Leontis NB, Lescoute A, Westhof E (2006) Curr Opin Struct Biol 16:279–287
13. Lietzke SE, Barnes CL, Kundrot CE (1995) Curr Opin Struct Biol 5:645–649
14. Mooers BHM (2009) Methods 47:168–176
15. Varani G, Tinoco I Jr (1991) Q Rev Biophys 24:479–532
16. Fürtig B, Richter C, Wöhnert J, Schwalbe H (2003) Chembiochem 4:936–962
17. Rose PW, Bi C, Bluhm WF, Christie CH, Dimitropoulos D, Dutta S, Green RK, Goodsell DS, Prlic A, Quesada M, Quinn GB, Ramos AG, Westbrook JD, Young J, Zardecki C, Berman HM, Bourne PE (2013) Nucleic Acids Res 41:D475–D482
18. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA (1996) Science 273:1678–1685

19. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) Science 289:905–920
20. D'Souza V, Dey A, Habib D, Summers MFJ (2004) Mol Biol 337:427–442
21. Wang Y-X, Zuo X, Wang J, Yu P, Butcher SE (2010) Methods 52:180–191
22. Tinoco I Jr, Borer PN, Dengler B, Levin MD, Uhlenbeck OC, Crothers DM, Bralla J (1973) Nat New Biol 246:40–41
23. Zuker M (1989) Science 244:48–52
24. Zuker M (2003) Nucleic Acids Res 31:3406–3415
25. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P (1994) Monatshefte Für Chem Chem Mon 125:167–188
26. Wuchty S, Fontana W, Hofacker IL, Schuster P (1999) Biopolymers 49:145–165
27. Mathews DH, Andre TC, Turner DH, Zuker M (1997) In: Leontis NB, Santa Lucia J Jr (eds) Molecular modeling of nucleic acids. American Chemical Society, San Francisco, CA
28. Bellaousov S, Reuter JS, Seetin MG, Mathews DH (2013) Nucleic Acids Res 41: W471–W474
29. McCaskill JS (1990) Biopolymers 29:1105–1119
30. Woese CR, Magrum LJ, Gupta R, Siegel RB, Stahl DA, Kop J, Crawford N, Brosius R, Gutell R, Hogan JJ, Noller HF (1980) Nucleic Acids Res 8:2275–2294
31. Gutell RR, Lee JC, Cannone JJ (2002) Curr Opin Struct Biol 12:301–310
32. Michel F, Westhof E (1990) J Mol Biol 216:585–610
33. Cedergren R, Major F (1998) RNA structure and function. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
34. Parisien M, Major F (2008) Nature 452:51–55
35. Lemieux S, Major F (2006) Nucleic Acids Res 34:2340–2346
36. Jossinet F, Ludwig TE, Westhof E (2010) Bioinformatics 26:2057–2059
37. Westhof E, Masquida B, Jossinet F (2011) Cold Spring Harb Perspect Biol 3:309–320
38. Marti-Renom MA, Madhusudhan MS, Fiser A, Rost B, Sali A (2002) Structure 10:435–440
39. Cruz JA, Blanchet M-F, Boniecki M, Bujnicki JM, Chen S-J, Cao S, Das R, Ding F, Dokholyan NV, Flores SC, Huang L, Lavender CA, Lisi V, Major F, Mikolajczak K, Patel DJ, Philips A, Puton T, Santalucia J, Sijenyi F, Hermann T, Rother K, Rother M, Serganov A, Skorupski M, Soltysinski T, Sripakdeevong P, Tuszynska I, Weeks KM, Waldsich C, Wildauer M, Leontis NB, Westhof E (2012) RNA 18:610–625
40. Peattie DA, Gilbert W (1980) Proc Natl Acad Sci U S A 77:4679–4682
41. Ehresmann C, Baudin F, Mougel M, Romby P, Ebel J-P, Ehresmann B (1987) Nucleic Acids Res 12:9109–9128
42. Stern S, Moazed D, Noller HF (1988) Methods Enzymol 164:481–489
43. Zwieb C, Ross A, Rinke J, Meinke M, Brimacombe R (1978) Nucleic Acids Res 5: 2705–2720
44. Brimacombe R, Stiege W, Kyriatsoulis A, Maly P (1988) Methods Enzymol 164:287–309
45. Zhang Q, Yu ET, Kellersberger KA, Crosland E, Fabris D (2006) J Am Soc Mass Spectrom 17:1570–1581
46. Krol A, Carbon P (1989) Methods Enzymol 180:212–227
47. Yu E, Fabris D (2003) J Mol Biol 330:211–223
48. Tullius TD, Dombroski BA (1985) Science 230:679–681
49. Tullius TD, Greenbaum JA (2005) Curr Opin Chem Biol 9:127–134
50. Pogozelski WK, McNeese TJ, Tullius TDJ (1995) Am Chem Soc 117:6428–6433
51. Ralston CY, Sclavi B, Sullivan M, Deras ML, Woodson SA, Chance MR, Brenowitz M (2000) Methods Enzymol 317:353–368
52. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KMJ (2005) Am Chem Soc 127:4223–4231
53. Wilkinson KA, Merino EJ, Weeks KM (2006) Nat Protoc 1:1610–1616
54. Cadet J, Sage E, Douki T (2005) Mutat Res 571:3–17
55. Zwieb C, Brimacombe R (1980) Nucleic Acids Res 8:2397–2411
56. Wilms C, Noah JW, Zhong D, Wollenzien P (1997) RNA 3:602–612
57. Harris ME, Christian EL (2009) Methods Enzymol 468:127–146

58. Lawley PD, Brookes P (1967) J Mol Biol 25:143–160
59. Brewer LA, Goelz S, Noller HF (1983) Biochemistry (Mosc) 22:4303–4309
60. Lippard SJ (1982) Science 218:1075–1082
61. Zhang Q, Crosland E, Fabris D (2008) Anal Chim Acta 627:117–128
62. Biemann K, McCloskey JA (1962) J Am Chem Soc 84:2005–2007
63. Quinn R, Basanta-Sanchez M, Rose RE, Fabris D (2013) J Mass Spectrom 48(6):703–712
64. Cheng X, Camp DG, Wu Q, Bakhtiar R, Springer DL, Morris BJ, Bruce JE, Anderson GA, Edmonds CG, Smith RD (1996) Nucleic Acids Res 24:2183–2189
65. Benjamin DR, Robinson CV, Hendrick JP, Hartl FU, Dobson CM (1998) Proc Natl Acad Sci U S A 95:7391–7395
66. Keough T, Baker TR, Dobson RL, Lacey MP, Riley TA, Hasselfield JA, Hesselberth PE (1993) Rapid Commun Mass Spectrom 7:195–200
67. Limbach PA (1996) Mass Spectrom Rev 15:297–336
68. McLuckey SA, Habibi-Goudarzi S (1993) J Am Chem Soc 115:12085–12095
69. Nordhoff E, Kirpekar F, Roepstorff P (1996) Mass Spectrom Rev 15:67–138
70. Kowalak JA, Pomerantz SC, Crain PF, McCloskey JA (1993) Nucleic Acids Res 21:4577–4585
71. Kirpekar F, Douthwaite S, Roepstorff P (2000) RNA 6:296–306
72. Hossain M, Limbach PA (2007) RNA 13:295–303
73. Little DP, Aaserud DJ, Valaskovic GA, McLafferty FW (1996) J Am Chem Soc 118:9352–9359
74. Taucher M, Breuker KJ (2010) Am Soc Mass Spectrom 21:918–929
75. Henry C (1997) Anal Chem 69:243A–246A
76. Venter JC, Smith HO, Hood L (1996) Nature 381:364–366
77. Wang BH, Biemann K (1993) Proceedings of the 41st ASMS conference on mass spectrometry and allied topics, San Francisco, CA
78. Limbach PA, Crain PF, McCloskey JA (1994) Nucleic Acids Res 22:2183–2196
79. Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FAP, Fabris D, Agris PF (2011) Nucleic Acids Res 39:D195–D201
80. Stone MP, Huang H, Brown KL, Shanmugam G (2011) Chem Biodivers 8:1571–1615
81. Swenberg JA, Lu K, Moeller BC, Gao L, Upton PB, Nakamura J, Starr TB (2011) Toxicol Sci 120:S130–S145
82. Kellersberger KA, Yu E, Kruppa GH, Young MM, Fabris D (2004) Anal Chem 76:2438–2445
83. Turner KB, Yi-Brunozzi HY, Brinson RG, Marino JP, Fabris D, Le Grice SF (2009) RNA 15:1605–1613
84. Fabris D, Yu ET (2010) J Mass Spectrom 45:841–860
85. Jensen ON, Kulkarni S, Aldrich JV, Barofsky DF (1996) Nucleic Acids Res 24:3866–3872
86. Steen H, Petersen J, Mann M, Jensen ON (2001) Protein Sci 10:1989–2001
87. Turner KB, Brinson RG, Yi-Brunozzi HY, Rausch JW, Miller JT, Le Grice SFJ, Marino JP, Fabris D (2008) Nucleic Acids Res 36:2799–2810
88. Brinson RG, Turner KB, Yi-Brunozzi HY, Le Grice SF, Fabris D, Marino JP (2009) Biochemistry (Mosc) 48:6988–6997
89. Turner KB, Kohlway AS, Hagan NA, Fabris D (2009) Biopolymers 91:283–296
90. Ganem B, Li YT, Henion JD (1991) J Am Chem Soc 113:6294–6296
91. Loo JA (1997) Mass Spectrom Rev 16:1–23
92. Turner KB, Hagan NA, Kohlway A, Fabris D (2006) J Am Soc Mass Spectrom 17:1401–1411
93. Hvidt A, Linderstrøm-Lang K (1954) Biochim Biophys Acta 14:574–575
94. Englander SWJ (2006) Am Soc Mass Spectrom 17:1481–1489
95. Johnston PD, Redfield AG (1977) Nucleic Acids Res 4:3599–3615
96. Englander SW, Sosnick TR, Englander JJ, Mayne L (1996) Curr Opin Struct Biol 6:18–23
97. Yu ET, Hawkins AE, Eaton J, Fabris D (2008) Proc Natl Acad Sci U S A 105:12248–12253
98. Yu ET, Zhang Q, Fabris D (2005) J Mol Biol 345:69–80

99. Light-Wahl KJ, Springer DL, Winger BE, Edmonds CG, Camp DG, Thrall BD, Smith RD (1993) J Am Chem Soc 115:803–804
100. Ganem B, Li Y-T, Henion JD (1993) Tetrahedron Lett 34:1445–1448
101. Turner KB, Hagan NA, Fabris D (2007) J Mol Biol 369:812–828
102. Stephenson W, Asare-Okai PN, Chen AA, Keller S, Santiago R, Tenenbaum SA, Garcia AE, Fabris D, Li PTXJ (2013) Am Chem Soc 135:5602–5611
103. Gabelica V, De Pauw E (2002) J Am Soc Mass Spectrom 13:91–98
104. Yang J, Mo J, Adamson JT, Hakansson K (2005) Anal Chem 77:1876–1882
105. Mo J, Hakansson K (2006) Anal Bioanal Chem 386:675–681
106. Taucher M, Breuker K (2012) Angew Chem Int Ed 51:11289–11292
107. McLuckey SA (1992) J Am Soc Mass Spectrom 3:599–614
108. Woodin RL, Bomse DS, Beauchamp JL (1978) J Am Chem Soc 100:3248–3250
109. Fabris D, Kellersberger KA, Wilhide JA (2012) Int J Mass Spectrom 312:155–162
110. Revercomb HE, Mason EA (1975) Anal Chem 47:970–983
111. Von Helden G, Hsu M-T, Kemper PR, Bowers MTJ (1991) Chem Phys 95:3835–3837
112. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH (2004) Rapid Commun Mass Spectrom 18:2401–2414
113. Shvartsburg AA, Smith RD (2008) Anal Chem 80:9689–9699
114. Clemmer DE, Jarrold MF (1997) J Mass Spectrom 32:577–592
115. Gidden J, Ferzoco A, Baker ES, Bowers MTJ (2004) Am Chem Soc 126:15132–15140
116. Baker ES, Bowers MTJ (2007) Am Soc Mass Spectrom 18:1188–1195
117. Gidden J, Baker ES, Ferzoco A, Bowers MT (2005) Int J Mass Spectrom 240:183–193
118. Baker ES, Dupuis NF, Bowers MTJ (2009) Phys Chem 113:1722–1727
119. Gabelica V, Baker ES, Teulade-Fichou MP, De Pauw E, Bowers MTJ (2007) Am Chem Soc 129:895–904
120. Porter KC, Beck JL (2011) Int J Mass Spectrom 304:195–203
121. Yu ET, Fabris D (2004) Anal Biochem 344:356–366

Index

Note: for some index terms with numerous references, those in **bold** indicate where the term is most thoroughly introduced.