# Tandem Mass Spectrometric Strategies for Determining Structure of Biologically Interesting Molecules

MICHAEL L. GROSS

*Departments of Chemistry and Medicine, Washington University, One Brookings Drive, St. Louis, Missouri 63130*

*Received September 17, 1994*

## Introduction

Molecular structure determination has always been one of the fascinations of the chemist. Advanced tools such as nuclear magnetic resonance spectroscopy and X-ray crystallography are favorites because they are general, definitive, and informative. They are not, however, sensitive. As chemists turn more and more to research at the interface of chemistry and biology, they face the requirements of determining structure for picomole and lower quantities of material, a quantity realm where contemporary NMR and X-ray crystallography cannot be applied.

It is no surprise that mass spectrometry can play an important role in the chemist's quest for structure for smaller and smaller quantities of biological materials because the successes of mass spectrometry and GC/MS for standard organic materials are well-known. Structural biological mass spectrometry, however, required a new tack because biologically interesting molecules are not sufficiently volatile and stable to survive the thermal rigors of electron ionization (EI) and GC/MS. New ionization methods for tandem mass spectrometric approaches are a major feature of the new directions needed for determining structure of biologically important molecules.

The principles of tandem mass spectrometry (MS/MS) have been under investigation for over two decades. Sixteen years ago at the University of Nebraska, we commissioned a commercial three-sector instrument, a tandem-sector instrument that was among the first to employ more than two analyzer stages.[1] We were in the fortunate position to exploit the capabilities of tandem mass spectrometry for biomolecule structure determination when fast atom bombardment (FAB) was announced in 1981 by Barber and co-workers.[2] FAB is a routine method that interfaces smoothly with continuous beam instruments such as magnetic sectors and quadrupoles.

Fast atom bombardment, like the more contemporary methods of matrix-assisted laser desorption and electrospray ionization, produces molecular ions in the form $(M + H)^+$, $(M - H)^+$, and $(M + Met)^+$, where Met is a monovalent metal ion (usually $Na^+$). Unfortunately, the desorption process is accompanied by little or no fragmentation. Tandem mass spectrometry not only overcomes the problem of limited fragmentation by submitting the molecule ion to activation, usually via collision with noble gas atoms, but also permits investigation of mixtures and removal of background ions due to bombardment-induced matrix chemistry.[3]

The purpose of this Account is to survey some of the structure-determination developments made at the University of Nebraska laboratory. Most involve the development and understanding of ion chemistry accompanying activation by high-energy collisions (e.g., 8 keV in the laboratory frame) in a tandem mass spectrometer. Unlike the decomposition reactions promoted by EI of organic molecules, those induced by collisions are more readily predicted and understood, and their application to deciphering *primary* structure of biomolecules is usually straightforward. The rather simple connection between ion chemistry and structure has as its basis the fact that the biomolecules are made up of building blocks held together by relatively weak bonds.[4] The fragmentation of peptides at peptide bonds and carbohydrates and conjugates at glycosidic bonds permits sequencing and is consistent with this simple idea of "building blocks".

A current challenge is to utilize this informative ion chemistry for ever-decreasing amounts of sample in ever-increasing complex mixtures. Meeting this challenge requires the development of new sensitive mass spectrometers, usually possessing some form of multichannel advantage, and their interface to separation methods. Some of our research efforts and those of many others have been directed toward the goal of higher sensitivity. The theme of this Account, however, is structure determination by utilizing ion chemistry.

The ion-chemistry subjects that will be reviewed briefly under the rubric of structure determinations are charge-remote fragmentations, metal ion interactions, contrasts in the fragmentation of metal-bearing vs protonated species, and the determination of isomerism in oligonucleotides and carcinogen-modified DNA adducts. Advances in this program have been made with the help of an outstanding group of staff (R. Cerny, F. Crow, D. Giblin, R. Lapp, D. Rempel, and K. Tomer), postdoctoral students (J. Adams, E. Davoli, M. George, and M. Orlando), and graduate students (K. Caldwell, J. Crockett, L. Deterding, R. Grese, P. Hu, N. Jensen, D. MacMillan,

Michael L. Gross was born in St. Cloud, MN, and received the B.A. and Ph.D. degrees in chemistry at St. John's University and the University of Minnesota, respectively. After postdoctoral studies with E. R. Thornton (University of Pennsylvania) and F. W. McLafferty (Purdue University), he joined the faculty of the University of Nebraska—Lincoln in 1968, where he achieved in 1988 the rank of C. Petrus Peterson Distinguished Professor of Chemistry. In 1994, he joined the Departments of Chemistry and Medicine at Washington University, where he intends to apply his long-standing interest in gas-phase ion chemistry and mass spectrometry instrument development to problems in biochemistry and medicine.

(1) Gross, M. L.; Chess, E. K.; Lyon, P A.; Crow, F. W.; Evans, S.; Tudge, H. *Int. J. Mass Spectrom. Ion Phys.* **1982**, *54*, 295–299.
(2) Barber, M.; Bordoli, R. A.; Sedgwick, R. D.; Tyler, A. N. *J. Chem. Soc., Chem. Commun.* **1981**, 325–327.
(3) Caldwell, K.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 72–91.
(4) Tomer, K. B.; Gross, M. L.; Denizer, M. L. *Anal. Chem.* **1986**, *58*, 2527–2534.

0001-4842/94/0127-0361$04.50/0 

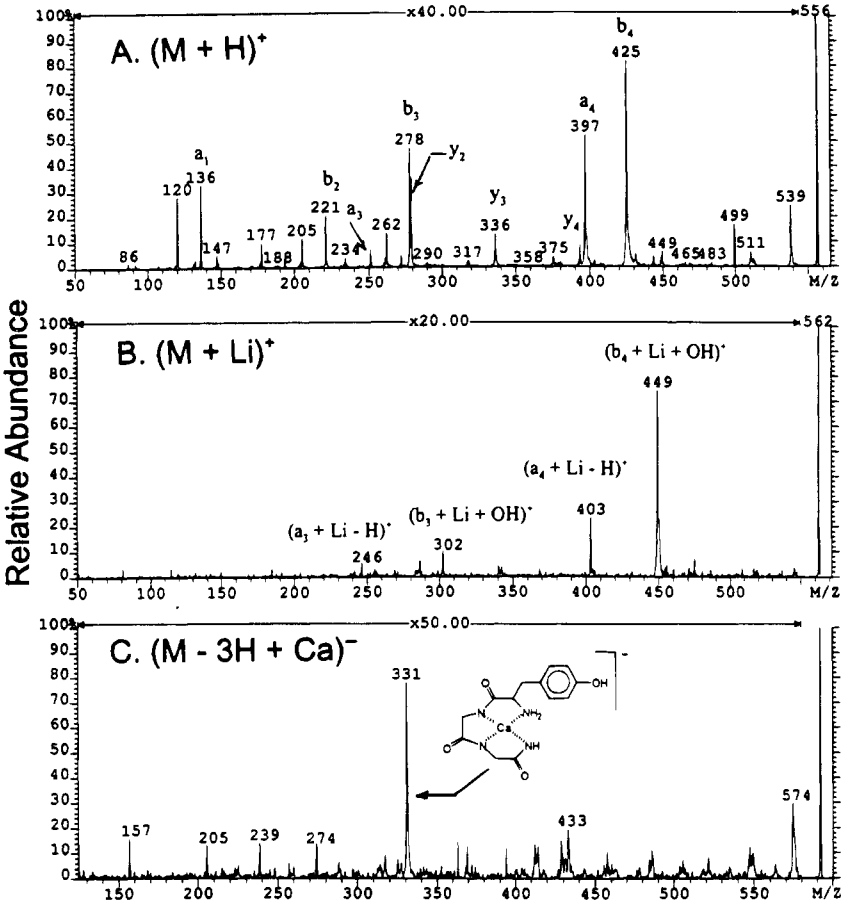



**Figure 1.** Collisionally activated dissociation (CAD) spectra of (A) the $(M + H)^+$, (B) the $(M + Li)^+$, and (C) the $(M - 3H + Ca)^-$ ions of leucine enkephalin (Trp-Gly-Gly-Phe-Leu). The ions were produced by $Cs^+$ FAB and investigated with a VG ZAB-T four-sector tandem mass spectrometer.

D. McCrery, D. Peake, D. Vollmer, and J. Wellemans) under the aegis of an erstwhile National Science Foundation Regional Instrumentation Facility Award to the University of Nebraska.

## Peptides

The major subject of interest to biochemical mass spectrometrists has been the sequencing of linear peptides. This subject is reviewed well in the Account by Biemann and Papayannopoulos in this issue. The ion chemistry associated with the fragmentation of protonated (and even multiply protonated) peptides is reasonably well understood, and the role of mass spectrometry as a complement to the Edman degradation is established. Unlike the Edman procedure, mass spectrometry offers the opportunity of sequencing peptides in mixtures. An elegant example is provided by the efforts of Hunt and co-workers,[5] who are sequencing peptides that are bound to and displayed at cell membranes as major histocompatibility complexes (MHC) that have important ramifications in immunology.

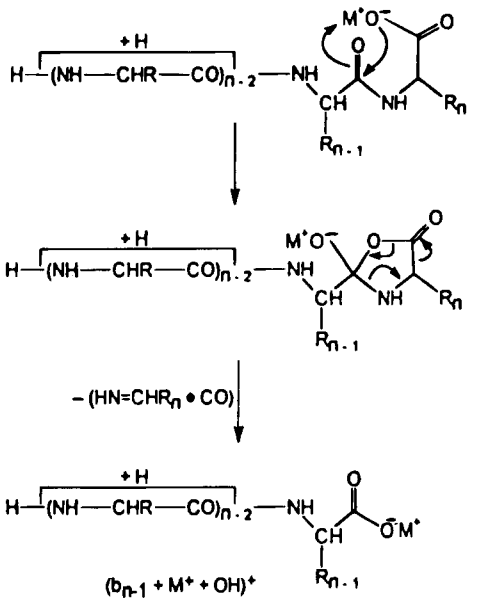
Most of the sequencing has made use of gas-phase $(M + H)^+$ ions. Other approaches may bear fruit. Our first approach was the use of $(M + Met)^+$, where Met is a monovalent metal ion. Unlike protonated peptides, which fragment at nearly every peptide bond (see Figure 1A), peptides cationized with $Li^+$ or $Na^+$ undergo a predominant rearrangement reaction centered at the C-terminus to expel the C-terminal amino acid (minus a water molecule) to produce a truncated peptide containing one less amino acid [labeled as $(b_4 + Li + OH)^+$ in Figure 1B].

The first evidence for a specific interaction of peptides and metal ions was by Russell and co-workers[6] whereas we and others[7-9] showed the correct stoichiometry and proposed mechanisms for the reaction. The mechanism we proposed[9] (Scheme 1) and that of Tang et al.[7] (Scheme 2) differ in details, and extensive studies by Adams and co-workers[10,11] and Leary and co-workers[12] favor the Tang mechanism. Nevertheless, the fragmentation of an alkali-metal-cationized peptide contrasts significantly with that of a protonated peptide (see Figure 1), providing complementary and specific information on the identity of the C-terminal amino acid. It may not be coincidence that the metal-containing enzyme carboxypeptidase catalyzes the removal of the C-terminal amino acid from a peptide or protein.

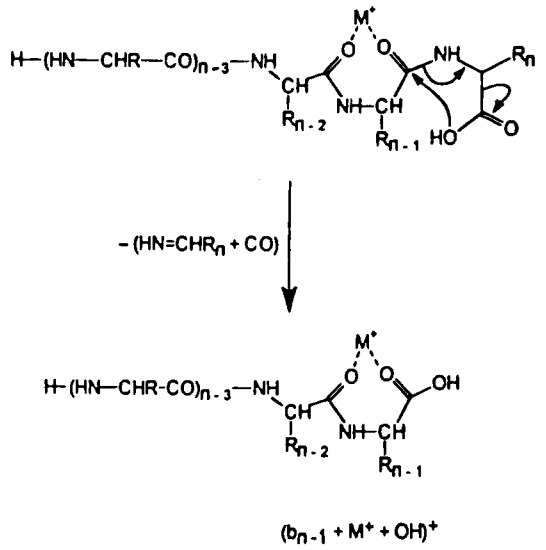
(5) Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A. L.; Appella, E.; Engelhard, V. H. *Science* **1992**, *265*, 1261–1263. Cox, A. L.; Skipper, J.; Chen, Y.; Henderson, R. A.; Darrow, T. L.; Shabanowitz, J.; Engelhard, V. H.; Hunt, D. F.; Slingluff, C. L., Jr. *Science* **1992**, *264*, 716–719 and references cited therein.

(6) Mallis, L. M.; Russell, D. H. *Anal. Chem.* **1986**, *58*, 1076–1080.
(7) Tang, X.; Ens, W.; Standing, K. G.; Westmore, J. B. *Anal. Chem.* **1988**, *60*, 1791–1799.
(8) Renner, D.; Spiteller, G. *Biomed. Mass Spectrom.* **1988**, *15*, 75–77.
(9) Grese, R. P.; Cerny, R. L.; Gross, M. L. *J. Am. Chem. Soc.* **1989**, *111*, 2835–2842.
(10) Teesch, L. M.; Adams, J. *J. Am. Chem. Soc.* **1991**, *113*, 812–820.
(11) Teesch, L. M.; Orlando, R. C.; Adams, J. *J. Am. Chem. Soc.* **1991**, *113*, 3668–3675.
(12) Leary, J. A.; Zhou, Z. R.; Ogden, S. A.; Williams, T. D. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 473–480.

**Scheme 1**




**Scheme 2**




Stronger interactions of metal ions with peptides occur when one species is multiply charged. For example, the interaction of divalent alkaline earth metal ions with negatively charged peptides gives gas-phase (peptide − H + Met)$^+$ species, which have been investigated by Adams and co-workers.[13] An even stronger interaction is of a peptide bearing a charge of 3− with a divalent metal to give a (peptide − 3H + Met)$^-$ species.[14] The structures of these unusual gas-phase complexes involve metal bonding at the C-terminal carboxylate, if the metal is an alkaline earth and the peptide is a tripeptide. For larger peptides, the bonding possibilities become more numerous, yet the CAD spectrum can be simple (see Figure 1C). Divalent transition metals, on the other hand, interact to give more than one structure, some of which involve only metal-to-nitrogen bonds.[15] These bonding properties, which also apply to solution species, are intrinsic, and mass spectrometry can be used to understand these intrinsic interactions.

(13) Teesch, L. M.; Adams, J. *J. Am. Chem. Soc.* **1990**, *112*, 4110–4120.
(14) Hu, P.; Gross, M. L. *J. Am. Chem. Soc.* **1992**, *114*, 9153–9160.
(15) Hu, P.; Gross, M. L. *J. Am. Chem. Soc.* **1993**, *115*, 8821–8828.

Because the interactions are strong, giving tetracoordinated gas-phase ions, the fragmentations for tripeptides and alkaline earth metals involve the N-terminus and side chains.[14] The weakened interaction of transition metal ions at the C-terminal carboxylate is manifested by the losses of $CO_2$ and $H_2$-$CO_2$.[15] Weaker overall interactions apply to complexes of small peptides bearing a charge of 2− with monovalent alkali-metal ions.[16]

The study of metal-cationized peptides in the gas phase is still new, but the variety of complexes is immense especially when one considers that there are a large number of metals and that multiply charged complexes can be produced by electrospray.[17] The dual opportunity of understanding intrinsic interactions as well as opening up new, structurally informative decomposition reactions motivates continued development of this field. Similar structural incentive applies to the investigation of negatively charged peptides, $(M - H)^-$, a subject actively pursued by Bowie and co-workers.[17]

## Cyclic Peptides

Cyclic peptides offer an additional challenge for structural studies because protonation at any amide nitrogen followed by bond cleavage gives a mixture of acylium ions [$H_2N(CHRCONH)_xCHRCO^+$]. Subsequent fragmentation gives a more complex pattern than that of a protonated linear peptide because there is ample structural diversity of the protonated molecule. The first unknown peptide to be sequenced by tandem mass spectrometry was a simple, four amino acid cyclic peptide called HC toxin.[18] A simplifying feature of the ion chemistry is the presence of a proline whose tertiary nitrogen is readily protonated, giving predominantly a single linear acylium ion after ring opening. The effect of proline was verified in a sequel study of larger peptides that also contained one or more prolines.[19]

For cyclic peptides containing amide sites of comparable basicity, the collisionally activated decompositions become more complex because of the multitude of ring-opened structures comprising $(M + H)^+$. Furthermore, fragmentation can take place from the N-terminus to eliminate a substituted imine and CO, in a manner similar to a C-terminal-directed fragmentation.

One approach to solving these problems is to choose dipeptide-type ions (e.g., $H_2NCHRCONHCHRCO^+$).[20] If two different dipeptide fragments can be defined (e.g., AB and CD), then two possible sequences can be proposed [i.e., c(A-B-C-D) and c(D-C-B-A). By extending this reasoning to other dipeptide ions coupled with a knowledge of the molecular weight and the amino acid composition (determined either by MS or by amino acid analysis), the structure of the cyclic peptide can be sometimes reassembled.

(16) Hu, P.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 137–143.
(17) Waugh, R. J.; Bowie, J. H. Gross, M. L.; Vollmer, D. *Int. J. Mass Spectrom. Ion Processes* **1994**, *133*, 165–174 and references contained therein.
(18) Gross, M. L.; McCrery, D. A.; Crow, F. W.; Tomer, K. B.; Pope, M. R.; Ciuffetti, L. M.; Knoche, H. W.; Daly, J. M.; Dunkle, L. D. *Tetrahedron Lett.* **1982**, *23*, 5381–5384.
(19) Tomer, K. B.; Crow, F. W.; Gross, M. L.; Kopple, K. D. *Anal. Chem.* **1984**, *56*, 880–886.
(20) Eckart, K.; Schwarz, H.; Tomer, K. B.; Gross, M. L.; *J. Am. Chem. Soc.* **1985**, *107*, 6765–6769.

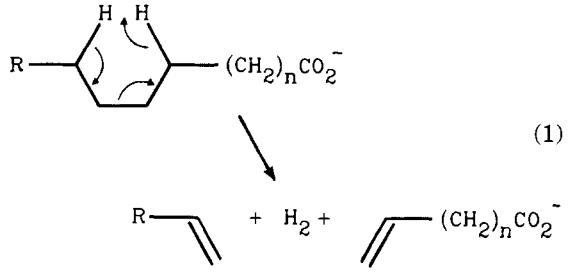
A second approach makes use of conventional mass spectrometric and tandem mass spectrometric data.[21] The following steps are recommended. (1) Obtain a desorption ionization mass spectrum and establish the molecular weight. (2) Determine the exact mass of the molecule ion. (3) Obtain the tandem mass spectrum (both metastable ion and collisionally activated) of the $(M + H)^+$. Compare the mass and tandem mass spectra. (4) On the basis of the comparison, select prominent product ions in the mass spectrum for additional MS/MS study. (5) Determine as necessary the exact masses and the elemental compositions of significant product ions. Data from many of these mass spectrometric steps as well as from NMR have been utilized in recent structural studies of natural-product cyclic peptides.[22-24]

## Fatty Acids/Charge-Remote Fragmentation

Fatty acids, of course, are important components of complex lipids. Their identification has been a major and long-standing challenge to mass spectrometrists. GC/MS has been the major structural tool because the quantities that can be isolated are small and the information from optical spectroscopies (UV, IR, NMR) is limited. But even for EI mass spectrometry, reference materials are usually needed because EI mass spectra of fatty acid isomers are similar and difficult to interpret.

We set out in 1983 to remedy this problem by investigating the high-energy collisional activation mass spectra of fatty acids introduced to the gas phase as $(M - H)^-$ or carboxylate anions. These investigations led to the discovery of a new class of gas-phase processes, which we termed "charge-remote".[25,26]

The charge site of a gas-phase $(M - H)^-$ of a fatty acid would be expected to reside on the carboxylate group, yet most of the fragmentation reactions occur along the hydrocarbon chain to give, for a saturated fatty acid, a series of parallel losses of the elements of alkanes. On the basis of deuterium isotope labeling[26] and other evidence, a mechanism for the apparent alkane elimination was proposed (eq 1).




(1)

If the mechanism is correct, one would expect the pattern of bond cleavages with rearrangement to be interrupted by the presence of multiple bonds. This is indeed the case[27] as is seen in the collisionally activated decomposition (CAD) spectra of two isomeric fatty acids desorbed as lithium-cationized species (Figure 2). The spectra were taken on a new-design four-sector mass spectrometer that is equipped with a large, 150-mm array detector to improve sensitivity.[28] Although the EI mass spectra of simple derivatives of these two fatty acids are similar, the tandem or CAD mass spectra of the $(M - H + 2Li)^+$ ions, as well as those of the carboxylate anions, are dramatically different. One can identify the two fatty acids and locate unambiguously the double bonds *without* reference spectra or standard compounds. Identification of other modified fatty acids, such as branched, multiply unsaturated, ring-containing, and hydroxyl-bearing, can also be achieved[29,30] even when the acids exist in complex mixture.[29] An example is shown in Figure 3, which demonstrates that isomeric, homoconjugated octadecadienoic acids are readily identified on the basis of the tandem mass spectra of the carboxylates.[31]

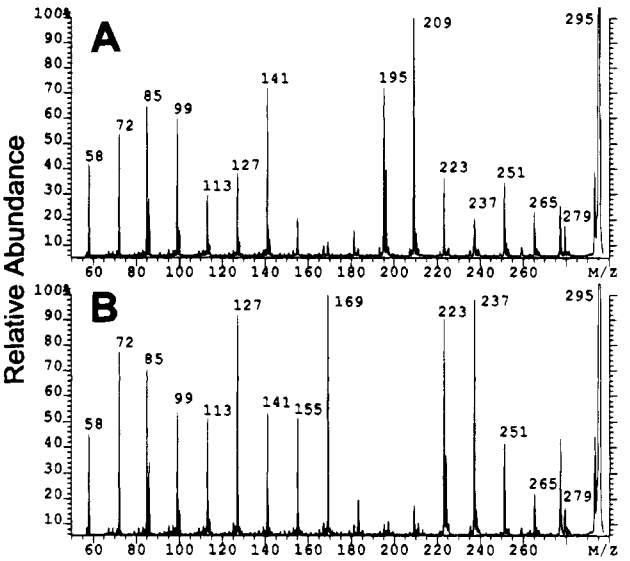



**Figure 2.** CAD spectra of the isomeric oleic (A) and vaccenic (B) fatty acids desorbed as $RCOOLi_2^+$. The spectra were taken with a VG ZAB-T four-sector tandem mass spectrometer and were magnified by 125 and 160 times, respectively, relative to the precursor ions.

Although the spectra in Figure 3 show poor mass resolution because a three-sector rather than a four-sector instrument was used, they do point to two important considerations. First, the tandem mass spectra of the same compound taken on different instruments are nearly identical. Second, the spectra are acquired with considerably more ease on tandem-sector mass spectrometers than on multiple-quadrupole or trapping instruments.

The mechanism for the cleavage of a C–C single bond (eq 1) that we proposed has been criticized by

(21) Cerny, R. L.; Gross, M. L. *Mass Spectrometry of Peptides*; Desiderio, D. M., Ed.; CRC Press: Boca Raton, 1990; pp 289–314.
(22) Helms, G. L.; Moore, R. E.; Niemczura, W. P.; Patterson, G. M. L.; Tomer, K. B.; Gross, M. L. *J. Org. Chem.* **1988**, *53*, 1298–1307.
(23) Pettit, G. R.; Kamano, Y.; Herald, C. L.; Dufresne, C.; Cerny, R. L.; Herald, D. L.; Schmidt, J. M.; Kizu, H. *J. Am. Chem. Soc.* **1989**, *111*, 5015–5017.
(24) Pettit, G. R.; Kamano, Y.; Herald, C. L.; Dufresne, C.; Bates, R. B.; Schmidt, J. M.; Cerny, R. L.; Kizu, H. *J. Org. Chem.* **1990**, *55*, 2989–2990.
(25) Tomer, K. B.; Crow, F. W.; Gross, M. L. *J. Am. Chem. Soc.* **1983**, *105*, 5487–5488.
(26) Jensen, N. J.; Tomer, K. B.; Gross, M. L. *J. Am. Chem. Soc.* **1985**, *107*, 1863–1868.
(27) Jensen, N. J.; Tomer, K. B.; Gross, M. L. *Anal. Chem.* **1985**, *57*, 2018–2021.
(28) Gross, M. L. *Methods in Enzymology*, McCloskey, J. A., Ed.; Academic Press: San Diego, 1990; Vol. 193, pp 131–153.
(29) Tomer, K. B.; Jensen, N. J.; Gross, M. L. *Anal. Chem.* **1986**, *58*, 2429–2433.
(30) Jensen, N. J.; Gross, M. L. *Lipids* **1986**, *21*, 362–365.
(31) Crockett, J. S.; Gross, M. L.; Christie, W. W.; Holman, R. T. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 183–191.

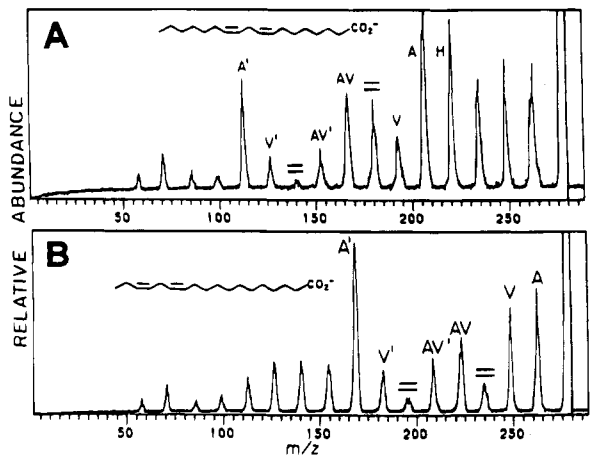



**Figure 3.** CAD spectra of the $(M - H)^-$ anions of two isomeric homoconjugated, 18-carbon fatty acids. The spectra were taken with a Kratos MS-50 three-sector tandem mass spectrometer. The letter designations A, V, =, and AV refer to ions produced by cleavage of the allylic, vinylic, double, and allylic/vinylic bonds, respectively.

others.[32] One key piece of evidence opposing the mechanism is that different types of compounds seem to require different collision energies to promote cleavage of a "remote" C–C bond. If the process is truly "charge-remote", then the nature of the charge-bearing site (e.g., $COO^-$ vs $SO_3^-$) should play little or no role.

Shifts in energy due to differing energy requirements for competitive, charge-driven processes may account for the different collision energies. Moreover, there is a convincing body of evidence that supports the 1,4-$H_2$ elimination mechanism shown in eq 1. First, the mechanism is a thermally allowed process involving a six-membered-ring transition structure. Second, there are thermal analogies. Pyrolysis of fatty acid esters leads to production of alkenes and unsaturated, smaller fatty acid esters.[33] Moreover, the collisionally induced fragmentations of $\beta$-hydroxyalkenoic acids, desorbed as $(M + Li)^+$, $(M - H + 2Li)^+$, or $(M - H)^-$, produce the same products as those of the thermal *O*-hydro-*C*-allyl elimination.[34] Third, collisionally activating the fatty acid as $(M + Li)^+$, $(M - H + 2Li)^+$ [35], or $(M - H + Ba)^+$ ions[36] leads to identical decompositions, demonstrating that the nature of the charge is unimportant. Fatty acid esters containing ester modifications of high proton affinity also undergo charge-remote fragmentations when activated as $(M + H)^+$ ions whereas methyl or ethyl esters undergo only charge-driven chemistry,[37,38] most of which is consistent with the established chemistry of carbocations and oxonium ions.

The study of metal-cationized fatty acids also permitted the energy requirements for the 1,4-$H_2$ elimination process to be approximated.[35] On the basis of the competition between release of the metal ion and the charge-remote cleavage of a C–C bond of an $(M + Met)^+$, the activation energy was estimated to be 1.3–1.9 eV, which is in the range of one-third to one-half of a C–C bond energy. It is not unreasonable to conclude that the C–C bond is 33–50% broken in the transition structure.

Classic mechanism proofs in chemistry require that all products be consistent with those predicted by the mechanism. Structure proof of products is not trivial for charge-remote reactions of fatty acids. Because the processes are induced by collision and do not occur in the ion source of the mass spectrometer, the products cannot be simply selected for study. This problem was circumvented by using two stages of collisional activation (an MS/MS/MS experiment) whereby the ionic products were characterized by the second stage of activation, and their structures were found to be in accord with a 1,4-elimination.[34]

A key element of the mechanism proof, however, is the nature of the neutral product(s). The 1,4-elimination should produce $H_2$ and a neutral olefin whereas alternative mechanisms will give an intact alkane or alkyl and hydrogen radicals.[32] At first thought, neutral species produced in mass spectrometric reactions are impossible to characterize because they bear no charge, but the development of neutralization/reionization mass spectrometry (NRMS) by McLafferty and co-workers[39] has changed that. The elegant work of Wesdemiotis[40] with NRMS has shown conclusively that the neutral products of charge-remote fragmentation are principally alkenes and *not* alkanes or alkyl radicals.

Charge-remote fragmentation is not only a useful approach to structure determination but also an important concept in understanding fragmentation processes in general. The concept is applicable to many other classes of compounds in addition to fatty acids, as will be discussed later. Moreover, it offers to the ion chemist a means to use the power of mass spectrometry to investigate thermal reactions that have been, until now, the province of chemists working with neutral species.

## Multiple Bond Location in Neutrals: $Fe^+$ Chemical Ionization

The utilization of charge-remote fragmentation to determine double bonds requires that the analyte be charged and exist as a closed-shell ion. Many structure determinations employ GC/MS, in which the analyte enters the mass spectrometer as a neutral molecule. Here, FAB coupled with tandem mass spectrometry is *not* appropriate.

One approach that does give distinctive spectra and allows for the ready distinction of isomeric olefins and related compounds is $Fe^+$ chemical ionization (CI).[41] The strategy builds on pioneering work done by J. Beauchamp, D. Ridge, and B. Freiser, who first explored the gas-phase reactions of transition metal(I) ions with simple organic molecules (see Account by Freiser is this issue). More proximate to structural applications was a collaborative study with

(32) Wysocki, V. H.; Ross, M. M. *Int. Mass Spectrom. Ion Processes* **1991**, *104*, 179.
(33) Sun, K. K.; Hayes, H. W.; Holman, R. T. *Org. Mass Spectrom.* **1970**, *3*, 1035–1042.
(34) Adams, J.; Gross, M. L. *J. Am. Chem. Soc.* **1989**, *111*, 435–440.
(35) Adams, J.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1986**, *108*, 6915–6921.
(36) Davoli, E.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 320–324.
(37) Deterding, L. J.; Gross, M. L. *Anal. Chim. Acta* **1987**, *200*, 431–445.
(38) Deterding, L. J.; Gross, M. L. *Org. Mass Spectrom.* **1988**, *23*, 169–177.
(39) Danis, P. O.; Wesdemiotis, C.; McLafferty, F. W. *J. Am. Chem. Soc.* **1983**, *105*, 7454–7456.
(40) Cordero, M. M.; Wesdemiotis, C. *Anal. Chem.* **1994**, *66*, 861–866.
(41) Peake, D. A.; Gross, M. L. *Anal. Chem.* **1985**, *57*, 115–120.

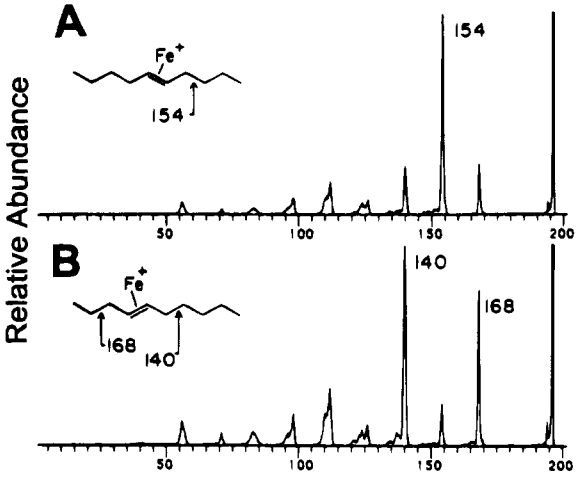



**Figure 4.** CAD spectra of $Fe(C_{10}H_{20})^+$ isomers of 5-decene (A) and 4-decene (B), respectively. The spectra were taken with a Kratos MS-50 three-sector mass spectrometer.

D. Ridge[42] in which we showed that $Fe^+$ oxidatively adds with high selectivity to allylic bonds. This oxidative addition is followed by a $\beta$-H transfer and formation of a bis(olefin) complex. Collisional activation of the bis(olefin) complex causes competitive releases of the two olefins to give a product ion spectrum that can be used to locate the double bond.

An example of data obtained by using this approach is represented by the simple distinction of 5-decene and 4-decene by tandem MS (see Figure 4). Oxidative addition to either allylic bond of 5-decene gives an $Fe^+$/heptene complex of *m/z* 154. Two different oxidative additions are possible for 4-decene to give $Fe^+$/olefin complexes of *m/z* 140 and 168. This approach to structure determination can be used also for alkenyl acetates (often found as insect sex pheromones) and fatty acid esters.[43]

## Complex Lipids

Structure-determination strategies employing charge-remote fragmentations are now being used for a variety of compound classes in addition to fatty acids. Examples include surfactants, steroids,[44] prostaglandins,[45] phospholipids, peptides (see Account by Biemann and Papayannopoulos in this issue), ceramides, glycosphingolipids, carbohydrates, and antibiotics (for recent reviews, see ref 46). In this section we will illustrate a completely instrumental method for the structure determination of a phospholipid. The example is an elaboration of earlier work done in our laboratory[47,48] and at Michigan State University.[49]

The structure elucidation of phospholipids by classical methods is laborious, time-consuming, and material consuming. The phospholipid is broken apart into component fatty acids and polar head group by using enzymatic methods. Many of the components are identified by using derivatization and GC/MS. Given that tandem mass spectrometry also takes apart complex biomolecules into their building blocks, it seems reasonable that an entirely instrumental method using FAB and MS/MS would be efficacious.

A phosphatidylcholine is desorbed upon FAB as $(M - CH_3)^-$, $[M - HN(CH_3)_3]^-$, and $[M - H_2C{=}CHN(CH_3)_3]^-$; the latter species is likely to be a phosphatidic acid. The three, high-mass ions are characteristic (see Figure 5) and are readily recognized, even in simple mixtures. They can be used to establish molecular weight and to give hints on composition.

Composition is revealed more explicitly by submitting any or all of the high-mass ions to collisional activation (see Figure 6). All decompose to liberate the constituent fatty acids presumably as carboxylate anions of *m/z* 303 and 283 in this case. The former, on the basis of its molecular mass, is likely to be a 20-carbon acid that contains four double bonds, possibly arachidonic acid. The latter *m/z* 283 ion is likely to be a saturated, 18-carbon acid.

The locations of the two fatty acids can be established on the basis of the abundances of the two carboxylate ions. From the two higher mass precursors, the fatty acid located at the 2-position of the glycerol is usually more abundant, whereas that in the 1-position is more readily produced from the lowest mass precursor. Although these abundance relationships are often the case, they are not general. The elimination of the acid in the 2-position, however, is always more facile in the fragmentations of the lowest mass precursor.[49] Thus, we see that collisional activation of the *m/z* 723 ion produces an *m/z* 419 ion that is more abundant than the *m/z* 439 ion, establishing that the carboxylate of *m/z* 303 is located in the 2-position of the complex lipid.

The remaining task is to identify the two fatty acids. Because they are produced upon FAB, the *m/z* 283 and 303 ions can be selected in separate experiments for collisional activation. The former ion is clearly steric acid on the basis of its CAD spectrum (see Figure 7). The fatty acid in the 2-position gives a rather simple spectrum (see Figure 7) that is dominated by a product ion formed by loss of $CO_2$. This charge-driven decarboxylation occurs for polyunsaturated fatty acid carboxylates that can stabilize a carbanion site by interaction with the multiple double bonds. A more definitive spectrum of polyunsaturated fatty acids can be produced by activating the $(M - H + 2Li)^+$ species[50] or the $(M - H + Ba)^+$ species.[36] Because charge-remote chemistry usually has higher energy requirements than charge-driven processes, the latter will dominate if given the opportunity. Charging a polyunsaturated fatty acid with metal ions preempts the charge-driven loss of $CO_2$ and allows the informative charge-remote processes to assert themselves.

A limitation of the methodology discussed in the foregoing is the use of FAB as the ionization methodology. An improvement is to use electrospray.[51] The increase in signal and the reduction in background noise due to matrix ions impart to this ionization mode

(42) Peake, D. A.; Gross, M. L.; Ridge, D. P. *J. Am. Chem. Soc.* **1984**, *106*, 4307–4316.
(43) Peake, D. A.; Huang, S.-K.; Gross, M. L. *Anal. Chem.* **1987**, *59*, 1557–1563.
(44) Tomer, K. B.; Gross, M. L. *Biomed. Environ. Mass Spectrom.* **1988**, *15*, 89–98.
(45) Zirrolli, J. A.; Davoli, E.; Bettazzoli, L.; Gross, M.; Murphy, R. C. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 325–333.
(46) Gross, M. L. *Int. J. Mass Spectrom. Ion Processes* **1992**, *118/119*, 137–165. Adams, J. *Mass Spectrom. Rev.* **1990**, *9*, 141–170.
(47) Jensen, N. J.; Tomer, K. B.; Gross, M. L. *Lipids* **1986**, *21*, 580–588.
(48) Jensen, N. J.; Tomer, K. B.; Gross, M. L. *Lipids* **1987**, *22*, 480–489.
(49) Huang, Z.-H.; Gage, D. A.; Sweeley, C. C. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 71–78.
(50) Adams, J.; Gross, M. L. *Anal. Chem.* **1987**, *59*, 1576–1582.
(51) Han, X.-L.; Gross, R. W. *Proc. Natl. Acad. Sci. U.S.A.*, in press.

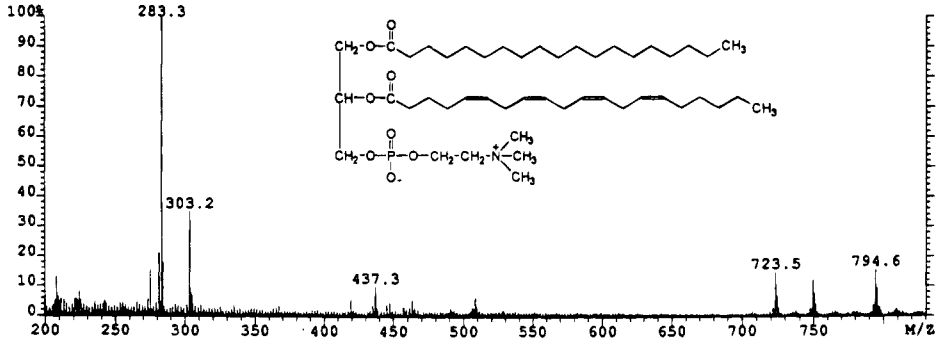



**Figure 5.** $Cs^+$ FAB mass spectrum of the phosphatidylcholine (C18:0/C20:4). Note the three high-mass ions of *m/z* 723, 749, and 794 and the two product carboxylate anions of *m/z* 283 and 303.

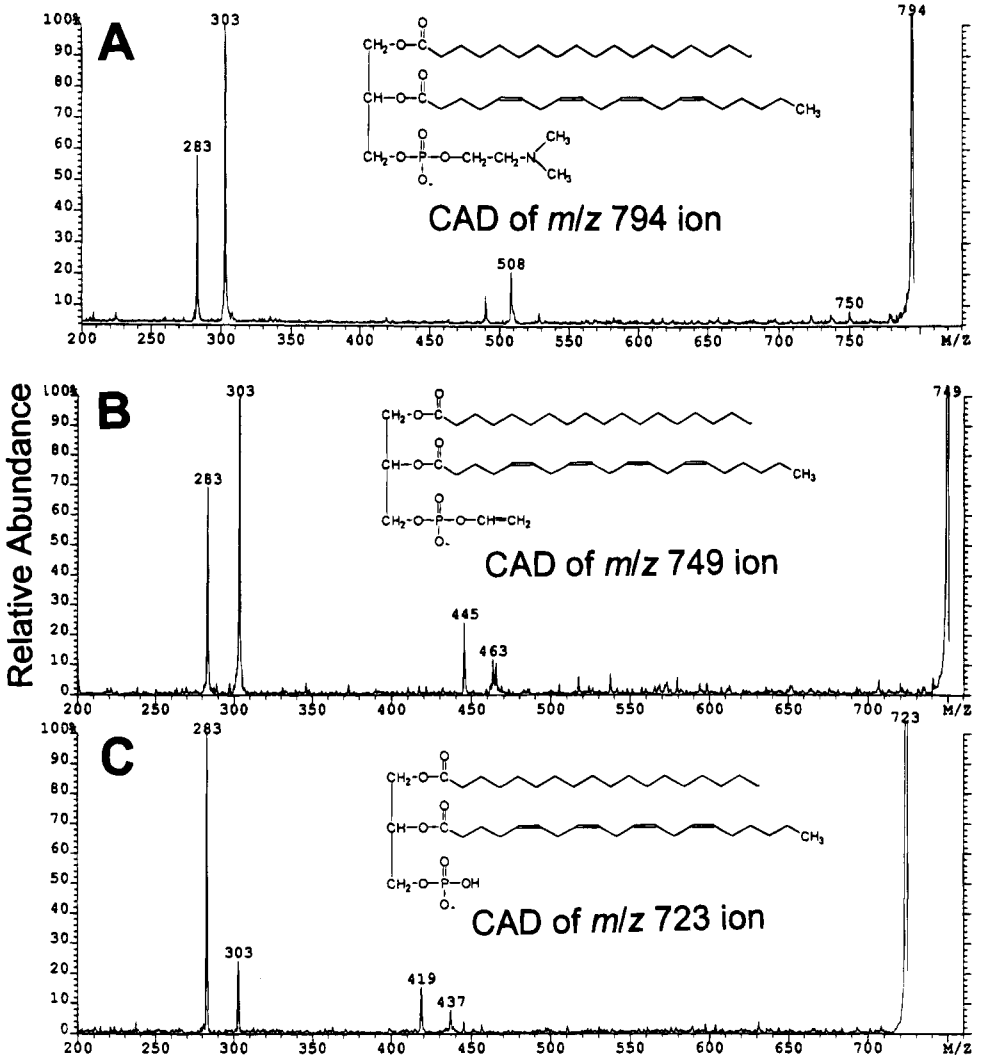



**Figure 6.** CAD mass spectra of the three high-mass ions noted in Figure 5: (A) $(M - 15)^-$, (B) $(M - 60)^-$, and (C) $(M - 86)^-$. All spectra were taken with a VG ZAB-T four-sector tandem mass spectrometer.

an improvement in sensitivity of 2–3 orders of magnitude. The combination of electrospray and charge-remote fragmentations will permit the structure determination of lipids at the picomole and lower levels.

## Oligonucleotides and Glycosides

Mass spectrometry of oligonucleotides is of interest because of the importance of oligonucleotides in the genome sequencing project, and they are now readily introduced into the gas phase by electrospray.[52] Some years ago, we studied a nearly complete set of dinucleotides to explore the opportunity to determine structure by utilizing FAB and tandem MS.[53] The $(M - H)^-$ ions of a dinucleotide decompose upon collisional activation to give sequence ions and to eliminate a base. The base located at the 3-position is preferentially eliminated as BH possibly because the rearranging hydrogen originates from the 2-posi-

(52) McLucky, S. A.; Habibi-Goudarzi, S. *J. Am. Chem. Soc.* **1993**, *115*, 12085–12095. Little, D. P.; Chorush, R. A.; Speir, J. P.; Senko, M. W.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1994**, *116*, 4893–4897.

(53) Cerny, R. L.; Gross, M. L.; Grotjahn, L. *Anal. Biochem.* **1986**, *56*, 424–435.

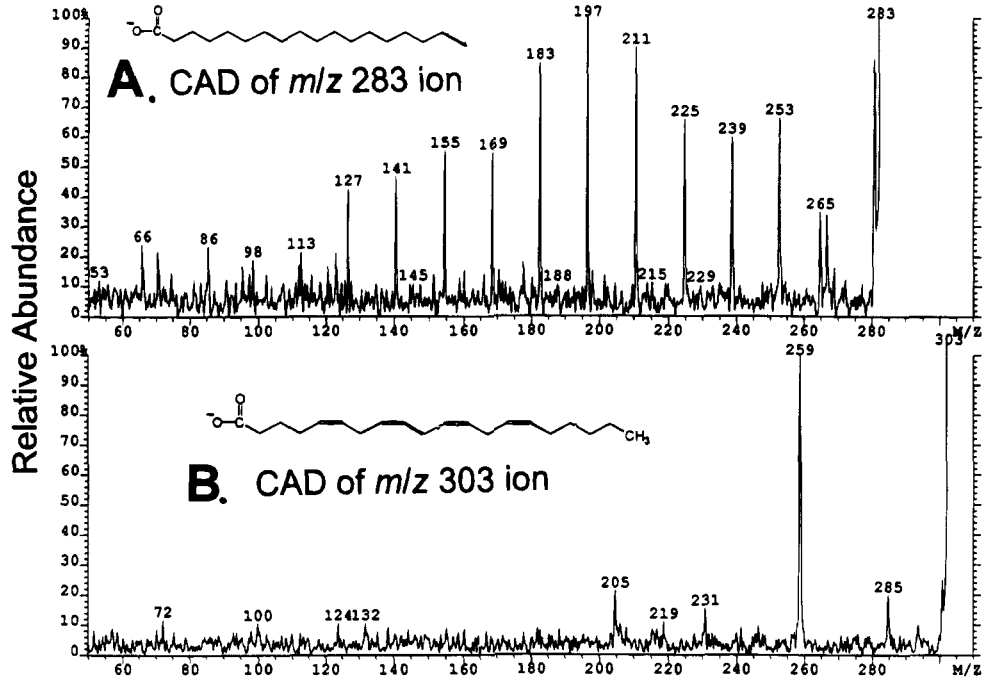



**Figure 7.** CAD spectra of the two carboxylate anions of *m/z* 283 (A) and 303 (B), noted in Figure 5. Both spectra were taken with a VG ZAB-T four-sector tandem mass spectrometer.

tion of the ribose and is activated by the nearby phosphate. In any case, structure proof of simple oligonucleotides is straightforward. More complex oligonucleotides are more difficult to determine, but sequence ions formed by cleavage of glycosidic bonds are seen for oligonucleotides as large as 6-mers.[54]

Another class of compounds in which glycosidic bond cleavages are important are steroid and flavonoid glycosides. Collisional activation in the MS/MS experiment leads to sequential losses of glycoside moieties in a manner that allows their sequence to be established.[55] Some flavonoids occur in nature as homologs, and the homology is also apparent on the basis of CAD spectra.

The method-development research was more recently followed by a structural study of 13 cardenolide glycosides from the roots of *Mandevilla pentlandiana*.[56] FAB and tandem mass spectrometry provided molecular weight and sequence information; and NMR was used later to provide details on stereochemistry and identity of the sugar moieties.

An important approach to structure determination of this class of compounds is to include in the collisional activation study not only the $(M + H)^+$ but also the $(M + Na)^+$ and $(M - H)^-$, if possible. We demonstrated the advantages of this approach with covalently modified nucleosides and nucleotides,[4] but a more recent example is provided by a method-development study directed at macrolide antibiotics.[57] These antibiotics are composed of a macrocycle with one or more sugar groups appended. Because one of the sugar moieties is an amino sugar, the structure of the $(M + H)^+$ ion probably has a proton on the highly basic amino group. The collisionally induced fragmentations are charge-remote glycosidic bond cleavages, giving rise to a simple spectrum that has ions in both the high- and low-mass regions and that provides information on masses of the appended sugar molecules.

By way of contrast, the $(M + K)^+$ ions have structures in which $K^+$ is bound to the macrocycle, and these ions give a completely different set of products that are formed by processes that are remote to a different charge site. The simple expedient of studying two readily produced gas-phase molecule ions provides two different "looks" at the molecule and is recommended for the structure determination or verification of biomolecules, when appropriate.

## Carcinogen-Modified DNA

To understand the mechanisms of carcinogenesis requires that the nature of modified DNA be determined. The DNA is usually isolated and degraded into nucleosides, which can be determined by a mass spectrometric procedure. GC/MS is recommended because it has high sensitivity, but its use requires derivatization of the modified nucleosides or nucleotides. A more direct approach, and one that takes advantage of the structurally distinctive fragmentation induced by collisional activation, is FAB and tandem mass spectrometry.

For polycyclic aromatic hydrocarbon (PAH) carcinogenesis, the usual assumption is that carcinogenic PAHs are activated by oxidation to a diol epoxide, which then reacts via the epoxide with nucleophilic sites on DNA bases. A simpler oxidation is the removal of one electron to give a PAH radical cation. This radical cation is expected to react in a different manner to form modifications that are characteristic of this new mechanism, which was proposed originally by Cavalieri and Rogan.[58]

(54) Cerny, R. L.; Tomer, K. B.; Looker, J. H.; Gross, M. L.; Grotjahn, L. *Anal. Biochem.* **1987**, *165*, 175–182.
(55) Crow, F. W.; Tomer, K. B.; Looker, J. H.; Gross, M. L. *Anal. Biochem.* **1986**, *155*, 286–307.
(56) Cabrera, G.; Deluca, M. E.; Seldes, A. M.; Gros, E. G.; Oberti, J. C.; Cocket, J.; Gross, M. L. *Phytochemistry* **1993**, *32*, 1253–1259.
(57) Cerny, R. L.; MacMillan, D. K.; Gross, M. L.; Mallams, A. K.; Pramanik, B. N. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 151–158.

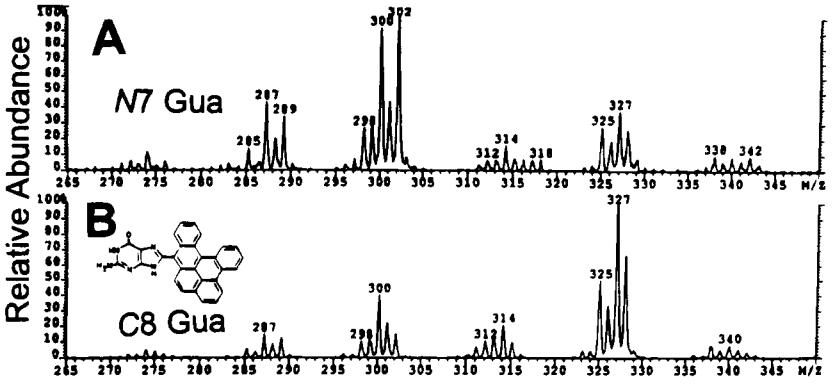



**Figure 8.** Partial CAD spectra of the two isomeric dibenzo[*a*,*l*]pyrene/guanine adducts: (A) DBP-10-N7Gua and (B) DBP-10-C8Gua. Both spectra were taken with a VG ZAB-T four-sector tandem mass spectrometer.

FAB combined with tandem mass spectrometry provided the first convincing evidence that this mechanism is operable in in vivo studies. Benzo[*a*]pyrene was activated to the radical cation by electrochemical oxidation or horseradish peroxidase[59] as well as by cytochrome P-450[60] in separate experiments and allowed to react with DNA in vivo. In all cases, a modified base that is characteristic of the radial cation mechanism was isolated and identified by using tandem MS.

Mass spectrometric studies of not only benzo[*a*]pyrene but also 6-methylbenzo[*a*]pyrene,[61] 6-fluorobenzo[*a*]pyrene,[61] and the highly carcinogenic dibenzo[*a*,*l*]pyrene[62] show that subtle differences in isomerism of modified bases and nucleosides can be detected by using tandem mass spectrometry. One typical example, that of guanine modified by dibenzo[*a*,*l*]pyrene, shows that the abundance ratio of two fragment ions is important (see Figure 8). The product ion of *m/z* 302 is the radical cation of dibenzo[*a*,*l*]pyrene (DBP), and it is produced abundantly when the DBP is substituted at a nitrogen atom. Conversely, when substitution occurs at carbon, the abundant product is of *m/z* 327, which is the radical cation of the cyano-substituted DBP. This latter ion is formed readily by a cycloreversion of the imidazole ring when carbon is substituted; when the PAH is bonded to nitrogen, the less-stable isonitrile forms with greater difficulty.

## Conclusions

As for many classes of biomolecules, the utilization of the informative ion chemistry in meaningful experiments requires that detection limits be lowered to the femtomole range and that coupling of tandem MS with separation be accomplished. These appear likely in the near future given the high sensitivity and fast scanning of the new generation of tandem mass spectrometers. Although the mass spectrometric experiments usually reveal little about higher order structure of biomolecules, they do provide considerable information on group connectivity (primary structure) at a sample-size level that cannot be touched by current NMR and X-ray methods.

The inability of tandem mass spectrometry to provide a complete picture of biomolecule structure can be overcome in collaborative studies in which bioorganic chemists prepare sufficient reference materials so that rules of ion chemistry can be convincingly established before unknowns are tackled. This is the approach we are taking in PAH carcinogenesis research.

*We gratefully acknowledge the National Institutes of Health (Grants 1PO1CA49210-05 and 2P41RR00954-18) for support in preparing this Account.*

(58) Cavalieri, E. L.; Rogan, E. G. In *Free Radicals in Biology*; Pryor, W. A., Ed.; Academic Press: New York, 1984; Vol. VI, pp 323–369.
(59) Rogan, E. G.; Cavalieri, E. L.; Tibbels, S. R.; Cremonesi, P.; Warner, C. D.; Nagel, D. L.; Tomer, K. B.; Cerny, R. L.; Gross, M. L. *J. Am. Chem. Soc.* **1988**, *110*, 4023–4029.
(60) Cavalieri, E. L.; Rogan, E. G.; Devanesan, P. D.; Cremonesi, P.; Cerney, R. L.; Gross, M. L.; Bodell, W. J. *Biochemistry* **1990**, *29*, 4820–4827.
(61) Ramakrishna, N. V. S.; Li, K.-M.; Rogan, E. G.; Cavalieri, E. L.; George, M.; Cerny, R. L.; Gross, M. L. *Chem. Res. Toxicol.* **1993**, *6*, 837–845.
(62) Ramakrishna, N. V. S.; Padmavathi, N. S.; Cavalieri, E. L.; Rogan, E. G.; Cerny, R. L.; Gross, M. L. *Chem. Res. Toxicol.* **1993**, *6*, 554–560.