Anal. Chem. **2004,** *76,* 1746−1753

Prediction of MS/MS Data. 1. A Focus on Pharmaceuticals Containing Carboxylic Acids

Mary L. Bandu, Kathryn R. Watkins, Melinda L. Bretthauer, Christopher A. Moore, and Heather Desaire*

Department of Chemistry, University of Kansas, Lawrence, Kansas 66045

Metabolite identification is a necessary step in developing safe and effective drugs. Metabolite analysis typically involves rapid identification of the chemical composition of the metabolite by automated HPLC−MS methods, followed by the laborious process of identifying the structure of the metabolite. Since MS is typically utilized to identify the metabolite, it is logical to utilize MS/MS to structurally characterize the sample. However, interpretation of MS/MS data may not provide sufficient information, as fragmentation pathways are not well understood or predictable. Therefore, other more time-consuming methods of analysis are often undertaken. If the dissociation rules for low-energy MS/MS experiments were clearly defined for all classes of compounds, more information would be obtained from MS/MS data, and metabolite identification would proceed more rapidly. We are currently developing methods to define these fragmentation rules. By screening ∼100 carboxylic acids at a time and applying knowledge of physical-organic chemistry, predictive rules are under development that describe how compounds dissociate under low-energy collision-induced dissociation conditions. Studies of carboxylic acid dissociation demonstrate that this approach is practical and reliable. Dissociation rules were predicted with a 90% success rate, when tested on acid-containing pharmaceuticals. This predictive power cannot be matched by any commercially available software. This study, and others like it, will be used to develop algorithms that more rapidly identify drug metabolites and degradation products, based on MS/MS data. Such algorithms will benefit drug development for all types of pharmaceuticals.

Since the advent of soft ionization techniques in the 1980s, mass spectrometry (MS) has quickly become the analytical technique of choice for determining structural information about compounds isolated from biological sources.[1] It provides both high sensitivity (detection limits of femtomoles of compound) and high selectivity—with the ability to analyze, quantify, or further study compounds of a specific mass.[1]

Mass spectrometry is heavily relied upon in the pharmaceutical industry to identify drug leads[2] and metabolites of new pharmaceuticals.[3−5] Metabolite identification studies help determine potentially toxic degradation products and metabolic hot spots, points on the drug that quickly decompose in the body.[3] Determining the mass of drug metabolites using mass spectrometry is fairly routine; however, structural characterization of these species (determining where the site of modification occurred) is much more challenging.[6] In fact, this process has been identified as the primary "bottleneck" in determining the chemical structures of metabolites.[3]

While the site of metabolic modification may be determined through manual interpretation of tandem mass spectrometric (MS^n) experiments on some molecules, this process is quite time-consuming. Additionally, some metabolites' structures cannot be determined in this fashion.[6,7] Current approaches to determining the site of modification on these compounds may involve collecting enough sample for analysis by NMR or obtaining partial structural information from MS/MS experiments, followed by synthesis of suspected metabolites and comparison of MS/MS spectra of these standards to the unknown compound.[7] These procedures are labor- and time-intensive; however, they are often necessary because low-energy MS/MS spectra cannot (currently) be predicted, based on a compound's structure.[7]

The studies in this report focus on developing predictive rules for MS/MS spectra of carboxylic acids, because many important pharmaceuticals contain carboxylic acid functional groups. Additionally, several pharmaceuticals become carboxylated as an important metabolic pathway.[8−12] A few examples include Celecoxib, an anti-inflamatory drug,[9] diphenhydramine, (Benadryl),[10]

* To whom correspondence should be addressed. Phone: 785-864-3015. E-mail: hdesaire@ku.edu.

(1) de Hoffman, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications*, 2nd ed.; John Wiley and Sons Ltd.: Chichester, England, 2002.
(2) Kassel, D. B. *Chem. Rev.* **2001**, *101*, 256.
(3) Cox, K. A.; Clarke, N. J.; Rindgen, D.; Korfmacher, W. A. *Am. Pharm. Rev.* **2001**, *4* (1), 45.
(4) Kostiainen, R.; Kotiaho, T.; Kuuranne, T.; Auriola, S. *J. Mass Spectrom.* **2003**, *38* (4), 357.
(5) White, R. E. *Annu. Rev. Pharm. Toxicol.* **2000**, *40*, 133.
(6) Clarke, N. J.; Rindgen, D.; Korfmacher, W. A.; Cox, K. A. *Anal. Chem.* **2001**, *73*, 430A.
(7) Emmer, J.; Vogel, M. *Biomed. Chromatogr.* **2000**, *14*, 373.
(8) Bu, H. Z.; Poglod, M.; Micetich, R. G.; Khan, J. K. *J. Mass Spectrom.* **1999**, *34*, 1185.
(9) Zhang, J. Y.; Wang, Y. F.; Dudkowski, C.; Yang, D. C.; Chang, M.; Yuan, J.; Paulson, S. K.; Breau, A. P. *J. Mass Spectrom.* **2000**, *35*, 1259.
(10) Kumar, S.; Riggs, K. W.; Rurak, D. W. *Drug Metab. Disp.* **1999**, *27*, 463.
(11) Sidelmann, U. G.; Christiansen, E.; Krogh, L.; Cornett, C.; Jornelund, J.; Hansen, S. H *Drug Metab. Disp.* **1997**, *25*, 725.
(12) Davies, I. D.; Allanson, J. P.; Causon, R. C. *J. Chromatogr., B* **1999**, *732*, 173.

10.1021/ac0353785 CCC: $27.50
Published on Web 02/07/2004

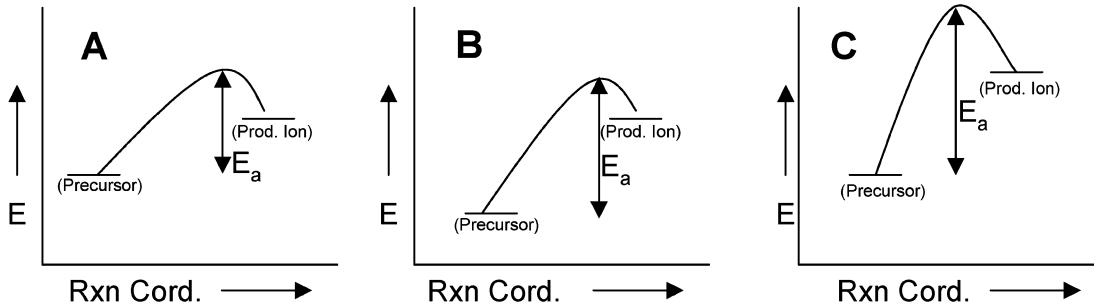


Figure 1. Energy level diagrams for three dissociation reactions. (A) Dissociation with low activation energy. (B) Dissociation with higher energy barrier: Precursor is stabilized relative to the transition state. (C) Dissociation with high energy barrier: Product is high in energy.

and chlorambucil, an anticancer drug.[12] Thus, determining the site of carboxylation in drug metabolites is a significant problem.

To address this and similar problems, some groups have attempted to develop predictive rules for MS/MS data through the use of artificial intelligence.[13] This approach has had phenomenal success in the field of proteomics.[14−17] For nonpeptidic molecules, however, computer algorithms that successfully predict MS/MS spectra do not exist. Some have attempted to develop algorithms for interpreting MS/MS data for pharmaceuticals by evaluating hundreds of collision-induced dissociation (CID) spectra and deriving fragmentation trends via statistical methods.[13] Unfortunately, this approach has its flaws. For example, using this method, Langley and co-workers have determined that all carboxylic acids show losses of 46 Da in their CID spectra.[13] This is not true for many carboxylic acids, and such assumptions may lead to errors in identification of unknown compounds.

The method described herein is used to develop predictive rules for MS/MS data by employing a different strategy. We develop fragmentation rules based on a small set of empirical observations and physical organic principles. Specifically, a set of molecules that undergo a particular dissociation are identified, and then we consider how the dissociation energy barriers for similar molecules will change based on the molecules' properties. To demonstrate how thinking about the energy barrier is useful in predicting fragmentation, consider Figure 1. In Figure 1, there are three energy level diagrams that show potential dissociation reactions. The ion in Figure 1A will dissociate readily, because the activation energy required to reach the transition state is lowest. If the precursor ion is selectively stabilized, as in Figure 1B, more energy for dissociation will be required, because the energy barrier (or activation energy) is higher. Similarly, Figure 1C shows an ion with a high energy barrier, because the product is high in energy. Thus, if the precursor ion is stabilized selectively (Figure 1B), or the product ion is destabilizing (Figure 1C), a dissociation reaction will be stifled. By contrast, raising the energy of the precursor ion and stabilizing the product ion facilitates dissociation.

We incorporate this principle into developing dissociation predictions in the following manner: (1) We identify a set of simple carboxylic acids that undergoes a certain dissociation—in this case, loss of CO_2. (2) We consider how varying the functional groups on those compounds will affect the energy barrier of the dissociation and then test the predictions. (3) We develop rules that describe whether any given carboxylic acid will dissociate. (4) The rules are validated by testing complex, multifunctional molecules (pharmaceuticals).

To summarize, by considering how different functional groups will effect the kinetics of known dissociation processes, one can predict whether a dissociation is (or is not) likely to occur, based on a small set of empirical CID spectra of simple organic molecules. This method is very effective at deriving predictions for CID fragmentation. On an independent test of 20 pharmaceutically active compounds, fragmentation is predicted with a 90% success rate.

EXPERIMENTAL SECTION

The carboxylic acids containing cyclopropyl groups were obtained from Dr. Jack A. Landgrebe at the University of Kansas. All other reagents were obtained from Sigma-Aldrich (St. Louis, MO), LKT Labs (St. Paul, MN), or Lancaster Chemical Co. (Windham, NH). Reagents were used without further purification.

Typically, compounds were dissolved in a minimal amount of HPLC grade methanol. If the compound was not readily soluble in methanol, a more appropriate solvent was chosen. All solutions were diluted with a methanolic solution containing 0.5% concentrated aqueous ammonia. The final concentration of the carboxylic acid was 1.0×10^{-4} M in all cases.

Samples were introduced into the mass spectrometer either by directly infusing the solution via a syringe pump or by automated injection of 25 μL of sample, pumped to the mass spectrometer with a Surveyor MS-Pump HPLC system (Thermo, San Jose, CA.) For automated injections, a mobile phase of 0.5% ammonia in methanol was used. (The method of injection did not affect the presence or absence of ions in the CID spectra.)

All samples were analyzed on an LCQ Advantage, a quadrupole ion trap mass spectrometer (Thermo), and data were acquired in the negative ion mode. Negative ion mode was chosen because comparatively little information is available about how negative ions dissociate. Furthermore, some of the carboxylic acids did not ionize in positive mode, yet all the carboxylic acids tested can be ionized in negative mode. Each sample was analyzed at least twice. Though the method of sample introduction and tuning parameters varied for the replicate injections, these deviations never caused a change in the presence or absence of ions in the

(13) Klagkou, K.; Pullen, F. S.; Harrison, M.; Organ, A.; Firth, A.; Langley, G. J. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1163.
(14) Sadygov, R. G.; Yates, J. R. *Anal. Chem.* **2003**, *75*, 3792.
(15) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976.
(16) Zhang, W.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 2482.
(17) Zhang, N.; Aebersold, R.; Schwilkowski, B. *Proteomics* **2002**, *2*, 1406.

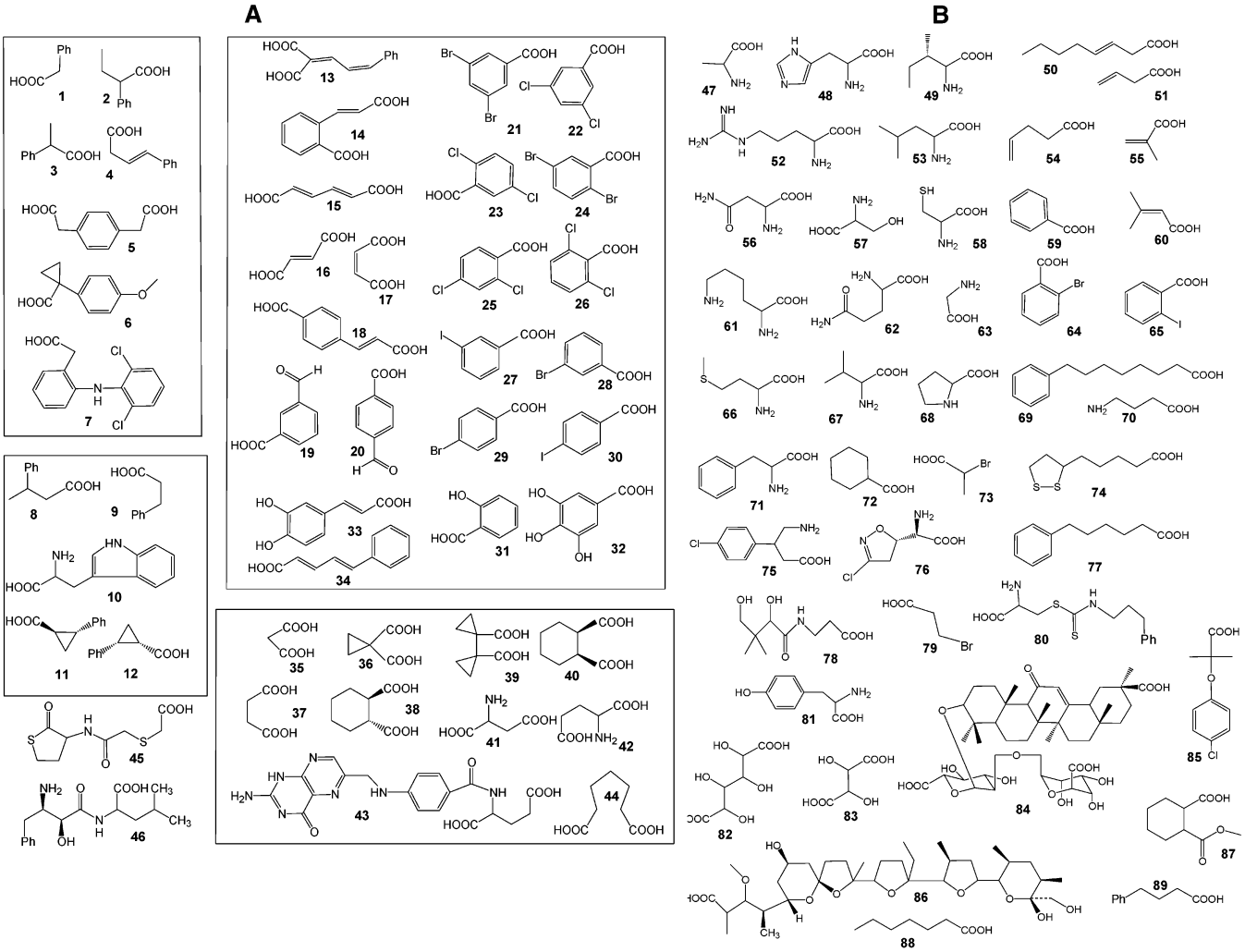


Figure 2. (A) Structures of carboxylic acids that lose CO_2 when subjected to CID in the quadrupole ion trap mass spectrometer. Compounds are grouped according to categories discussed in the text. (B) Acids that do not lose CO_2.

CID data. Variations in ion activation conditions did affect the presence and absence of ions in the CID spectra, so for all the data shown herein, identical ion activation conditions were used. Specifically, the precursor ion isolation width was 5 Da, each precursor ion was activated for 30 ms with 30% normalized collision energy (as defined by the Xcalibur 1.3 software), and a q_z value of 0.25 was used. For very low molecular weight ions, q_z was raised slightly in order to more effectively isolate the precursor ion. This change did not affect the presence or absence of product ions observed in these cases; it simply increased the ion counts. "Presence" of a product ion is specifically defined as an ion consistently larger than 3% relative abundance. This criterion was chosen based on past precedence.[18,19]

RESULTS AND DISCUSSION

The dissociations of a variety of carboxylic acids were investigated on a quadrupole ion trap mass spectrometer. Very simple organic acids were chosen initially, with only one or two different types of functional groups present. These compounds were selected so that effects of the functional groups upon the dissociation of the acid could be readily inferred. Because the structures of these compounds are very simple, their CID spectra did not contain an abundant number of product ions. When more complex molecules (pharmaceuticals) undergo CID, their spectra contain more product ions than the spectra obtained for the standards used here. We demonstrate herein that the "complexity" of the spectra does not effect whether certain dissociation pathways are observed. Rather, the functional groups present may effect the activation energy required for a given dissociation to occur, but those effects can be predicted, by employing basic physical organic principles. The fact that compounds with simple spectra can be effectively used to develop dissociation rules for molecules that have complex spectra is an important principle, and it is demonstrated in this work.

The CID spectra of the carboxylic acid standards commonly contained the loss of CO_2 as an abundant dissociation pathway, and thus, this neutral loss will be dealt with herein. Figure 2A contains compounds that did lose CO_2 upon activation, and Figure 2B contains compounds that did not display a neutral loss of 44 Da (CO_2) during CID. While most of the compounds in Figure 2A produced abundant losses of CO_2 (consistently greater than 10%), eight compounds showed modest losses of carbon dioxide

(18) Desaire, H.; Leary, J. A. *Anal. Chem.* **1999**, *71*, 4142.
(19) Desaire, H.; Leary, J. A. *Int. J. Mass Spectom.* **2001**, *209*, 171.

Table 1. Neutral Losses Observed for the 89 Model Compounds

compd no. (from Figure 2A)	neutral losses above 3%					compd no. (from Figure 2B)	neutral losses above 3%				
	CO_2 loss	other losses									
1	44					**47**	(no losses)				
2	44	46	93			**48**	17				
3	44	30	32	48	56	**49**	17				
4	44	2				**50**	46				
5	44					**51**	53				
6	44	46	58			**52**	17	42			
7	44	42				**53**	(no losses)				
8	44					**54**	66				
9	44					**55**	(no losses)				
10	44	17	61	87		**56**	18				
11	44					**57**	30	32			
12	44					**58**	(no losses)				
13	44					**59**	(no losses)				
14	44					**60**	(no losses)				
15	44					**61**	(no losses)				
16	44					**62**	18				
17	44					**63**	(no losses)				
18	44					**64**	102	120			
19	44	28				**65**	120				
20	44	28				**66**	48	50			
21	44					**67**	(no losses)				
22	44					**68**	(no losses)				
23	44					**69**	(no losses)				
24	44					**70**	18				
25	44					**71**	17				
26	44					**72**	(no losses)				
27	44	102	120			**73**	72				
28	44	102	120			**74**	34				
29	44	102	120			**75**	17	29	59	73	
30	44	102	120			**76**	53	103			
31	44					**77**	18				
32	44					**78**	72	130			
33	44					**79**	15	72			
34	44					**80**	34	78	177		
35	44					**81**	17	61	87		
36	44					**82**	18	124			
37	44	18				**83**	18	46	62	90	
38	44	18				**84**	18	470			
39	44					**85**	86	128			
40	44	18				**86**	32				
41	44	17				**87**	32				
42	44	18				**88**	(no losses)				
43	44	18	129			**89**	72	104			
44	44	18	62								
45	44	60	92	157							
46	44	177									

(3−10% relative abundance.) Those compounds are as follows: **8**, **27**−**30**, and **44**−**46**. See Table 1 for a comprehensive list of all the neutral losses observed for the compounds.

The fact that certain types of acids lose CO_2 during CID and other acids do not lose CO_2 is important, because this information provides a link between the structure of the molecule and its mass spectrum. Currently, the loss of CO_2 in a CID spectrum provides no information about the structure of the parent ion, other than it likely has a carboxylic acid in it. Yet, knowing what groups promote loss of CO_2 and what groups inhibit it allows researchers to use the CID data to help identify where on the molecule the carboxylic acid may be.

By grouping the compounds in Figure 2A into separate subclasses, based on functional groups in proximity to the carboxylic acid, additional information about the loss of CO_2 becomes apparent. Specifically, it becomes apparent that there are four classes of compounds where the loss of CO_2 is kinetically favored. By considering the effects of varying the functional groups

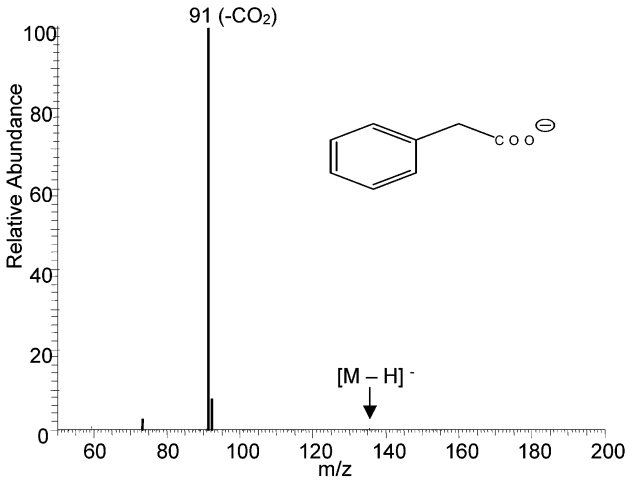
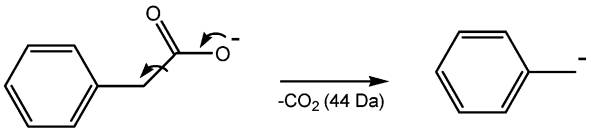


Figure 3. MS/MS of phenylacetic acid (**1**). The predominant product ion, *m*/*z* 91, corresponds to a neutral loss of CO_2.

Scheme 1. Loss of CO_2



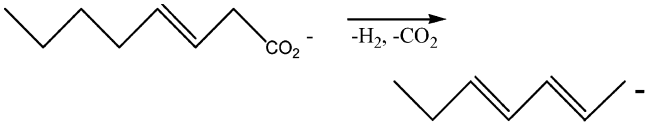
in proximity to the carboxylic acid of these four classes of compounds, general rules about CO_2 dissociation are revealed. The following four groupings describe 96% of the cases where loss of CO_2 was observed, and general rules are provided that explain when any compound loses CO_2 upon CID.

Group 1: Carboxylic Acids One Carbon Removed from a Conjugated System of sp^2 Carbons. All of the compounds tested thus far that have a benzene ring, or other highly conjugated system, at the β position (with respect to the carboxylic acid), show loss of CO_2 in their CID spectra. For an example see Figure 3, which depicts MS/MS data for **1**. Similar compounds that also lost CO_2 upon collisional activation include **2**−**7** in Figure 2A. It seems very likely that this product ion is observed because the product is resonance stabilized (Scheme 1). Resonance stabilization is a fundamental principle of organic chemistry, and these dissociation reactions appear to follow the basic principle that a resonance-stabilized product anion forms easily. In this case, the product of the reaction is stabilized by resonance, while precursor ion is not resonance stabilized by the benzene ring. Stabilizing the product ion results in a low activation energy and a favorable reaction. Numerous accounts of the formation of these resonance-stabilized anions in the gas phase exist in the literature.[20,21]

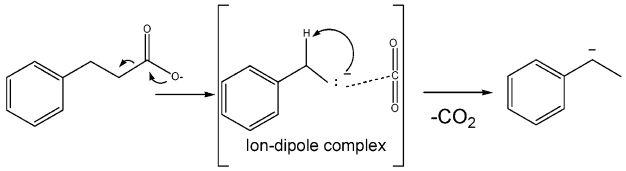
It should be noted that a single C−C double bond does not provide enough stabilization to promote loss of CO_2. For example, **50** and **51** (in Figure 1B) do not dissociate to give loss of 44 Da. Although loss of CO_2 from **51** would produce a resonance-stabilized anion, the loss is not observed. Similarly, **50** also does not exhibit losses of CO_2 upon collisional activation; instead it produces a neutral loss of 46 Da. This is possibly a result of the

(20) Janousek, B. K.; Brauman, J. I. In *Gas-Phase Ion Chemistry*; Bowers, M. T., Ed.; Academic Press: New York, 1979; Vol. 2, pp 53−83.
(21) Bartmess, J. E.; McIver, R. T. In *Gas-Phase Ion Chemistry*; Bowers, M. T., Ed.; Academic Press: New York, 1979; Vol. 2, pp 88−119.

Scheme 2. Proposed Explanation for Loss of 46 Da from Compound **50**



Scheme 3. Proposed Two-Step Mechanism for Loss of CO_2



concerted loss of CO_2 and H_2, forming an anion stabilized by two double bonds (Scheme 2). The loss of H_2 in this case would provide additional resonance stability for the anion. (It is also conceivable that the loss of 46 Da results from the simultaneous loss of CO and H_2O; however, generating a loss of H_2O typically involves transfer of acidic protons to an oxygen, and there are no acidic protons present in **50**.) From these experimental observations, it can be inferred that the presence of one carbon−carbon double bond does not provide enough resonance stabilization to promote loss of CO_2, under the collision conditions chosen.

Group 2: Carboxylic Acids Two Carbons Removed from a Benzene Ring or Other Highly Conjugated System. Not all of the losses of CO_2 (seen in Figure 2A) can be explained based on the presence or absence of a resonance-stabilized product. Compounds **8**−**12** each have an sp^2-hybridized carbon that is two carbons away from the carboxylic acid, and they all produce a loss of CO_2 as well. This dissociation may be explained in a two-step mechanism (Scheme 3). Generation of an ion−dipole complex, through loss of CO_2, is followed by a proton-transfer reaction. Similar proton transfers occurring during CID have been studied extensively elsewhere.[22]

Again, the presence of one carbon−carbon double bond, two carbons removed from the carboxylic acid, does not promote loss of CO_2 under the activation conditions used in this study. For example, **54** in Figure 2B does not lose carbon dioxide during CID. In fact, the presence of a benzene ring, two carbons away from the acid, is likely to be the minimum amount of resonance stabilization required for the rearrangement shown in Scheme 3. This assertion is based on the fact that the product ion representing loss of CO_2 for **8** was only 8% abundant, compared to the parent ion. Additionally, if proton donors are present within hydrogen-bonding distance of the carboxylate ion, the loss of CO_2 may be suppressed or depleted completely. Three examples include **71**, **75**, and **81**, which do not lose CO_2. In these cases, the proton donors can hydrogen bond with the carboxylic acid and selectively stabilize the precursor ion, as in Figure 1B, and the activation energy barrier is raised. This prohibits the dissociation.

Group 3: Carboxylic Acids on sp^2 Carbons. Based on the compounds shown in Figure 2, carboxylic acids located on sp^2-hybridized compounds also may lose carbon dioxide upon collisional activation, provided other conditions are met. Specifically, the compound must contain electron-withdrawing groups or an extended π system. The loss of CO_2 in this case is likely facilitated by the fact that sp^2 carbons can stabilize negative charges, due to the increased s character of the carbon.[23−25] Therefore, the transition state and the product of the reaction would be stabilized, so a low activation energy is required.

All of the carboxylic acids that have substantial electron-withdrawing groups present on double bonds (in conjugation with the acid) lose CO_2 with ease (**13**−**33**). In these cases, the electron-withdrawing groups can lower the activation energy of

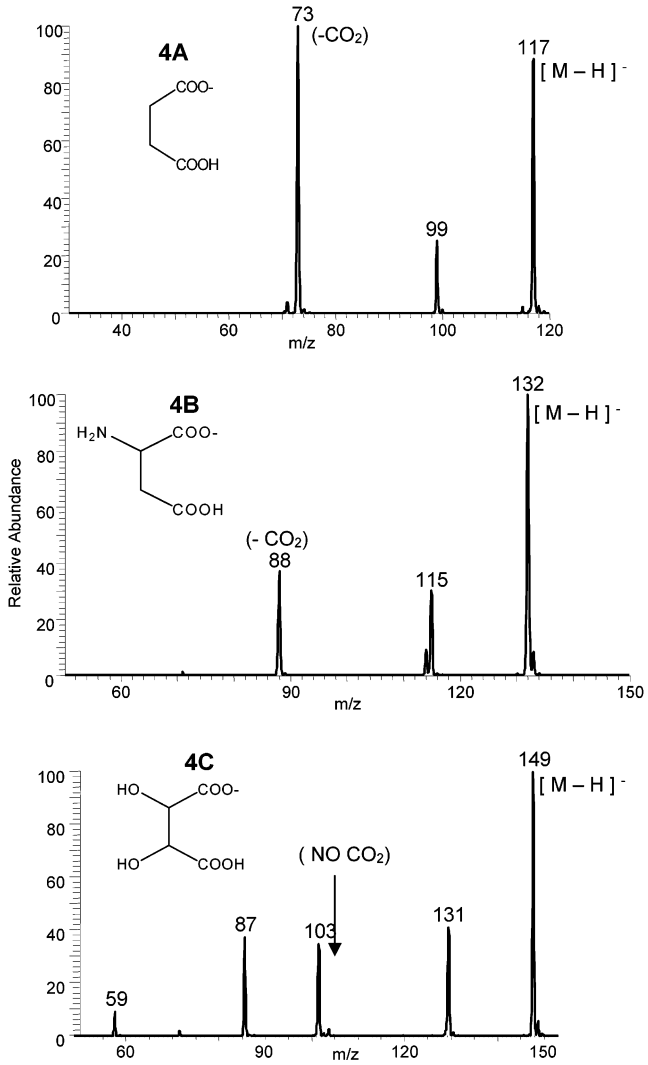


Figure 4. MS/MS of dicarboxylic acids two carbons apart. (A) Succinic acid (**37**) contains no hydroxyl or amino groups; therefore, CO_2 losses are abundant. (B) Arginine (**41**) contains an amino group within hydrogen-bonding distance of the carboxylate. The product ion peak representing the neutral loss of carbon dioxide (*m*/*z* 88) has an intermediate abundance of ∼40%. (C) Tartaric acid (**83**) does not lose CO_2 upon activation. The hydroxyl groups influence the fragmentation.

(22) Dua, S.; Adams, G. W.; Sheldon, J. C.; Bowie, J. H. *J. Chem. Soc., Perkin Trans. 2* **1996**, 1251.

(23) So, S. P.; Wong, M. H.; Luh, T. *J. Org. Chem.* **1985**, *50*, 2632.

(24) Skurski, P.; Simons, J.; Wang, X. B.; Wang, L. S. *J. Am. Chem. Soc.* **2000**, *122*, 4499.

(25) Peerboom, R. A.; de Koning, L. J.; Nibbering, M. M. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 159.

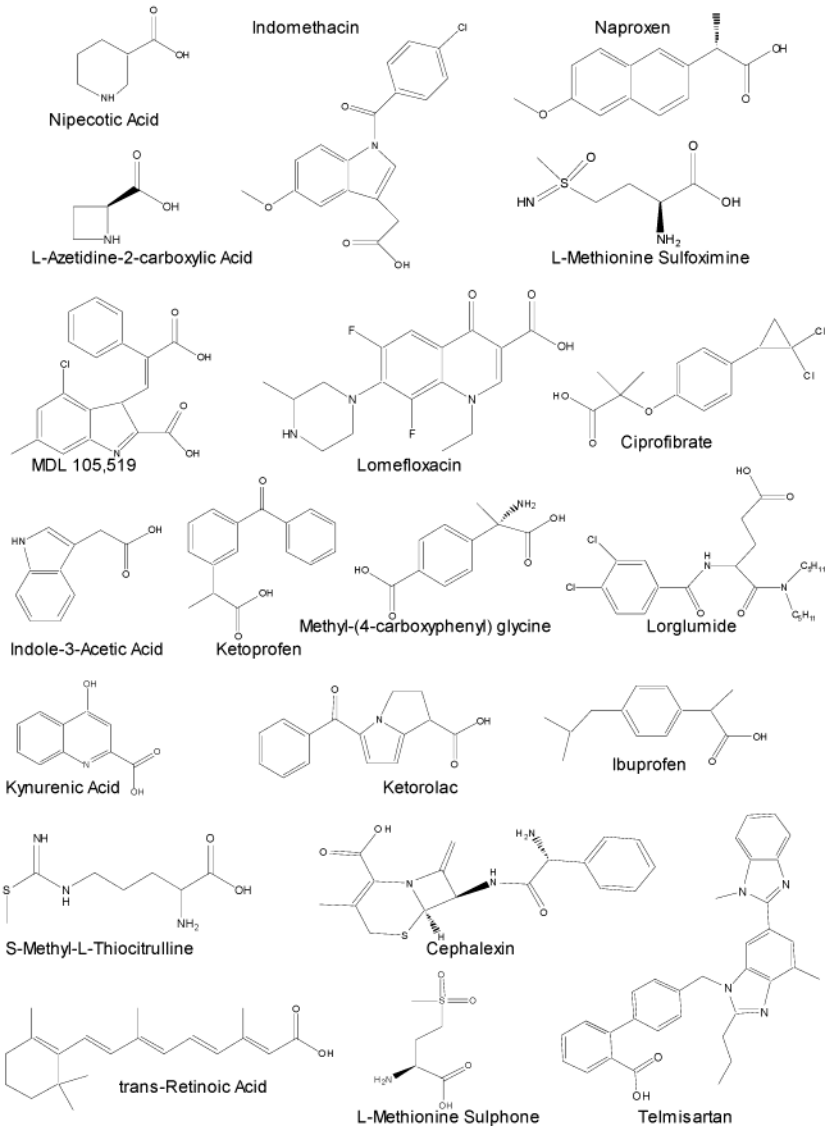


Figure 5. Structures of 20 pharmaceuticals used to test fragmentation predictions.

the reaction by pulling electron density away from the carboxylic acid, stabilizing the resulting transition state and anionic product. The electron-withdrawing groups that stabilize the dissociation the most include carboxylic acids (**13**−**18**), aldehydes (**19**, **20**), or multiple halogens (**21**−**26**). A single halogen present on a benzoic acid afforded modest losses of CO_2 in most cases (**27**−**30**), and hydroxybenzoic acids also lost CO_2 (**31**−**33**).

While electron-withdrawing groups can promote a loss of CO_2, so can an extended π system. For example, **34** contains five double bonds in conjugation but no electron-withdrawing groups. Benzoic acid (**59**), with no electronegative atoms present and three conjugated double bonds, did not readily lose CO_2 upon collisional activation; a small loss of 44 Da was detected below the 3% abundance threshold. Likewise, carboxylic acids on sp^2 carbons with only one double bond (and no electron-withdrawing groups present) did not show loss of CO_2. (See **55** and **60**.) Based on these experimental observations, a rule describing the loss of CO_2 for these types of compounds could be summarized as follows: If a carboxylic acid is present on an sp^2 carbon with a substantial amount of conjugation, CO_2 will be lost upon collisional activation, regardless of the other functional groups present. Yet, if a carboxylic acid is present on a π system containing three or fewer double bonds, electron-withdrawing groups must also be present to facilitate loss of CO_2 during CID, under the collision conditions of this study.

Group 4: Dicarboxylic Acids. One remaining class of compounds consistently lost CO_2 upon activation: compounds containing two carboxylic acids in proximity, each located on sp^3 carbons. See **35**−**44** in Figure 2A. If the carboxylic acids are one carbon apart, loss of CO_2 occurs with ease. (This loss is essentially identical to the mechanism in Scheme 1: Loss of CO_2 produces a resonance-stabilized product ion.) If the carboxylic acids are two carbons apart, and hydroxyl groups are not present, loss of CO_2 also occurs upon collisional activation of the precursor ion. (See **37**−**41**.) The presence of a single amino group suppresses the loss slightly. For example, **37**—with no hydrogen-bonding groups present—produces a neutral loss of CO_2 as the base peak in the spectrum (Figure 4A), while **41**—with one amino group in proximity to the acid—produces the neutral loss of 44 Da to a lesser abundance, 40% (Figure 4B). Compound **83**, which has two hydroxyl groups in proximity to the acids, Figure 4C, does not produce a neutral loss of CO_2 under the same collision

conditions. This trend may be a result of hydrogen-bonding groups stabilizing the carboxylate (the parent ion), increasing the activation energy of the dissociation, and suppressing the loss of CO_2. The energy level diagram would be similar to Figure 1B. This trend was also observed with the acids in group 2. Alternatively, the loss of CO_2 may not be observed, because it occurs in conjunction with loss of H_2O to produce the ion m/z 87. (See Figure 4C.) Mechanistic studies, currently underway, will be useful in determining why **83**, in Figure 4C, does not produce a neutral loss of 44 Da.

Dicarboxylic acids that have three carbons between the acids showed minimal losses of CO_2. We expect that if hydroxyl groups were present, this loss would not occur at all. Compound **82** (in Figure 2B), with carboxylic acids four carbons apart and multiple hydroxyl groups present, did not lose carbon dioxide upon CID. Mechanistic studies rationalizing this dissociation pathway are the subject of a current investigation. These studies will provide insight into the best way to predict the presence or absence of a loss of CO_2 for dicarboxylic acids located on sp^3 carbons.

Of the 89 compounds in this study, almost all the compounds that showed appreciable losses of CO_2 can be described by one of the categories above. There were just two exceptions, **45** and **46**. These compounds showed small losses of CO_2, each at ~8% abundance. Based on the structure of the compounds, it is not immediately obvious why carbon dioxide was lost. It is most likely that the functional groups present in **45** and **46** promote an alternate pathway that results in a kinetically favorable loss of CO_2. While many different carboxylic acids were studied herein, other classes of compounds remain that should be investigated; **45** and **46** provide a good starting point for identifying future classes of compounds to study.

Rules Describing Loss of CO_2. The findings above may be summarized in the following manner. A carboxylic acid-containing compound will lose CO_2, (if subjected to the CID conditions above) if at least one of the following cases is true: (1) The carboxylic acid is one carbon away from a conjugated system of π bonds. At least three double bonds are present (group 1). (2) The carboxylic acid is two carbons away from a conjugated system of at least three double bonds and hydrogen-bonding groups are not interacting with the carboxylate (group 2). (3) The carboxylic acid is on an sp^2-hybridized carbon that is part of a conjugated system of at least four double bonds (group 3). (4) The carboxylic acid is on an sp^2-hybridized carbon and strong electron-withdrawing groups are also part of the conjugated system (group 3). (5) The compound is a dicarboxylic acid, with the two acids in proximity, and acidic hydrogens are not within hydrogen-bonding distance of the acid (group 4).

Test Compounds. To demonstrate the predictability and completeness of the rules described above, 20 different pharmaceuticals that contained a carboxylic acid functional group were purchased. They are depicted in Figure 5. These compounds were analyzed after the fragmentation trends for the original group of compounds were rationalized. Prior to analysis of the 20 pharmaceuticals, predictions were made about whether each compound would produce a neutral loss of carbon dioxide upon CID, based on the rules above.

Specifically, we predicted that indomethacin, ketoprofen, ibuprofen, ketrolac, naproxen, indole-3-acetic acid, and methyl-4-carboxyphenylglycine should lose CO_2, because the resulting product ion would be resonance stabilized. Likewise, kynurenic acid, *trans*-retinoic acid, and telmisartan should lose CO_2 because the acid is on an sp^2 carbon and at least four double bonds are present in the π system. Finally, lomefloxacin and MDL 105,519 both have carboxylic acids on sp^2 carbons, which have electron-withdrawing groups present in their π system, so decarboxylations should be observed for these compounds as well. The remaining compounds do not fit a category of acids (as described above) that lose carbon dioxide; so they should not lose CO_2, according to the prediction method. Table 2 summarizes the predictions as well as the outcomes of the CID experiments, and Table 3 provides a list of all of the neutral losses observed for the pharmaceuticals. As shown, when the above trends were applied, carbon dioxide loss was predicted with a 90% success rate.

Table 2. Fragmentation Predictions for 20 Pharmaceuticals

compound	prediction for loss of CO_2	actual loss of CO_2
indomethacin	yes	yes
ketoprofen	yes	yes
ibuprofen	**yes**	**no**
ketorolac	yes	yes
lorglumide	no	no
naproxen	yes	yes
(±)-α-methyl-(4-carboxy-phenyl)glycine	yes	yes
kynurenic acid	yes	yes
S-methyl-L-thiocitrulline	no	no
L-methionine sulfoximine	no	no
indole-3-acetic acid	yes	yes
cephalexin	**no**	**yes**
ciprofibrate	no	no
lomefloxacin	yes	yes
nipecotic acid	no	no
L-methionine sulfone	no	no
L-azetidine-2-carboxylic acid	no	no
MDL 105,519	yes	yes
telmisartan	yes	yes
trans-retinoic acid	yes	yes

Table 3. Neutral Losses Observed for Pharmaceuticals

pharmaceuticals	neutral losses above 3%			
	CO_2 loss	other losses		
indomethacin	44			
ketoprofen	44	56		
ibuprofen		46		
ketorolac	44			
lorglumide		157		
naproxen	44			
(±)-α-methyl-(4-carboxy-phenyl)glycine	44	46		
kynurenic acid	44			
S-methyl-L-thiocitrulline		45	48	
L-methionine sulfoximine		101	117	
indole-3-acetic acid	44			
cephalexin	44	34	78	113
ciprofibrate		86	202	
lomefloxacin	44	26	59	72
nipecotic acid		no losses		
L-methionine sulfone		101	244	
L-azetidine-2-carboxylic acid		28		
MDL 105,519	44	88		
telmisartan	44			
trans-retinoic acid	44			

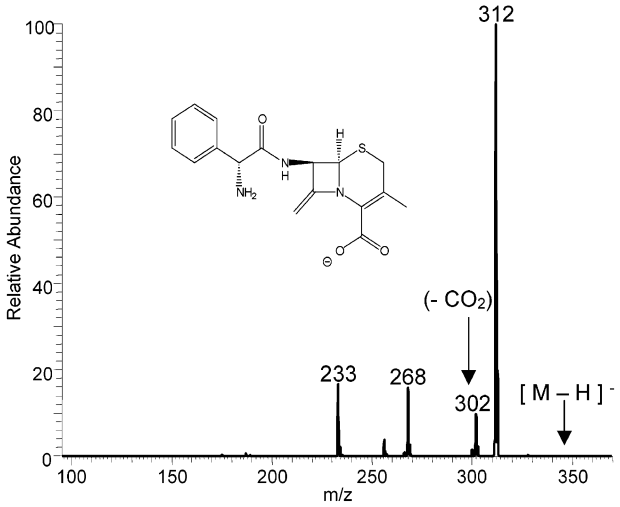


Figure 6. MS/MS of cephalexin. Loss of carbon dioxide was incorrectly predicted due to a lack of information on the effects of the amide. Carbon dioxide loss (*m/z* 302) has a low abundance of ~9%.

Although 18 acids fragmented as predicted, two acids, ibuprofen and cephalexin, did not follow expected trends. Ibuprofen was predicted to lose carbon dioxide and produce a resonance-stabilized anion, but the spectrum indicates a loss of 46 Da from the parent ion, not 44 Da. While this loss was not predicted, it is similar to the behavior of **50**, which also loses 46 Da (Scheme 2). It is possible that the loss of 46 Da corresponds to a loss of CO_2 and H_2 and that the loss of H_2 creates an additional double bond conjugated to the benzene ring, increasing stability of the product ion. (The 46 Da loss is also consistent with loss of CO and H_2O.)

Loss of CO_2 for cephalexin was predicted to be unfavorable. This prediction was based on the fact that the carboxylic acid was located on a double bond that was not part of a highly conjugated system. As the spectrum in Figure 6 indicates, a peak representing loss of CO_2 is present (~9% abundance). In this case, it is probable that the electron-withdrawing properties of the allylic amine, conjugated to the double bond, were underestimated.

In general, the results of this study are very promising, because the rules describing loss of CO_2 were highly predictive. Of the 89 "standards" used to develop these rules, 87 compounds' dissociations were consistent with the rules developed. (Only 2% of the 89 compounds in the study underwent an unexplained loss of CO_2.) Most of these 89 "standards" contained very simple structures. When the rules developed for these simple compounds were used to predict fragmentation for more complex compounds, the 20 pharmaceuticals in Figure 5, the rules were still very reliable: They had a 90% success rate.

By comparing the predictions made by the method described herein to the predictions of commercially available software designed to identify compounds from their MS/MS data, the value of this method becomes apparent. When the structures of the carboxylic acids in this study are input into Mass Frontier 3.0, a well-known MS/MS data interpretation program, the software program predicts that all of the compounds lose the carboxylic acid functional group upon collisional activation. Therefore, Mass Frontier could not be used to determine the location of a carboxylic acid in a metabolite, because the program does not differentiate among carboxylic acids in different chemical environments. Additionally, the program may misguide users into thinking that if a loss of the carboxylic acid is not apparent in the CID spectrum, the compound is not a carboxylic acid. The method described herein does not suffer these limitations. By determining rules that govern when a loss of CO_2 will appear in the CID spectrum, more information about the structure of unknown compounds is obtained.

CONCLUSION

This study represents the first step in developing a rational method of predicting the occurrence of product ions in MS/MS data for nonpeptidic molecules. A single neutral loss, loss of CO_2, was focused on herein. By determining all the possible scenarios in which the neutral loss of CO_2 may be detected in the product ion spectrum of a molecule, structural information about that molecule becomes apparent. Specifically, loss of CO_2 (44 Da) in a product ion spectrum indicates more than just the presence of a carboxylic acid in a molecule. It indicates that a carboxylic acid is present and certain other functional groups, which promote the loss of CO_2, are also in proximity.

By developing basic fragmentation rules that apply to low-energy CID spectra, MS/MS data (acquired under specified conditions) will be predictable, based on a compound's structure. This predicting power will be a great step forward in the field of pharmaceutical research, because it would introduce new technology that can more rapidly identify structures of unknown compounds, including drug metabolites and degradation products.

ACKNOWLEDGMENT

The authors thank Dr. Jack Landgrebe for donating some of the compounds used in the study. H.D. thanks Dr. Michael D. Leavell for helpful discussions and the University of Kansas for financial support.

Received for review November 20, 2003. Accepted January 12, 2004.

AC0353785