Applications of Artificial Intelligence for Chemical Inference. XIV.[1] A General Method for Predicting Molecular Ions in Mass Spectra

R. Geoff Dromey,* Bruce G. Buchanan, Dennis H. Smith, Joshua Lederberg, and Carl Djerassi

Departments of Computer Science, Chemistry, and Genetics, Stanford University, Stanford, California 94305

Received August 2, 1974

A general method for predicting molecular ions is described which is effective whether or not the molecular ion peak is present in a mass spectrum. The method involves generating a set of fragment losses characteristic of a particular spectrum. These losses are then used to predict a ranked set of candidate molecular ions. The method is applicable to both high- and low-resolution spectra and has been successfully applied to a wide range of classes of compounds.

Assignment or identification of the molecular ion plays a key role in structural elucidation from mass spectral data. Knowledge of the molecular ion highly restricts the space of possible molecular structures which might have given rise to a particular spectrum and virtually all spectral interpretations—be they purely conceptual or automated—start with this datum. However, a problem that frequently arises is that a spectral peak corresponding to the molecular ion is absent from the spectrum. To proceed with an automated analysis when the molecular ion is unknown (may be present or absent), as is done with the Heuristic DENDRAL programs,[2,3] it is essential to make inferences as to possible molecular ion candidates before the analysis can proceed.

Several procedures for molecular ion inference have been proposed by McLafferty,[4] Biemann,[5] and Reed.[6] These methods have usually been applied to high-resolution spectra. Molecular ions are selected when a number of criteria are satisfied, *e.g.*, correct parity of mass and nitrogen atom content, correct isotopic distributions, and reasonable composition losses to other peaks in the spectrum. Generation of candidates beyond the highest mass present in the spectrum is done by addition of a set of "good losses" to the highest mass. In this paper we propose an alternative and more general method for molecular ion prediction.

The method we propose parallels the work of Biemann[5] and McLafferty[4] in that candidate molecular ions are accepted or rejected using a set of "bad losses." However, the procedure used here to infer candidate molecular ions does not depend on an arbitrary set of "atoms or groups which can be easily lost from the molecular ion."[5] Rather, a set of "secondary losses" (see below), derived solely from the spectrum being analyzed, is used to infer candidate molecular ions. In addition, we use an intensity index derived from the spectrum to rank our candidate molecular ions rather than the ratio of good losses to bad losses reported previously.[4]

In simplest terms the present method involves a search for all masses (or compositions) X, such that X either appears in the lower half of the spectrum or X is a difference (mass or composition) formed from any pair of ions in the spectrum. Members of this set that do not satisfy certain mass and/or composition constraints are excluded. The range of possible molecular ions then spans every ion which is the sum of an X (from above) and a peak in the upper half of the spectrum. Plausible molecular ions are then filtered from this latter set and ranked using the peak intensities in the upper half of the spectrum and a cumulative intensity index assigned to the X's. Details of these procedures are given below.

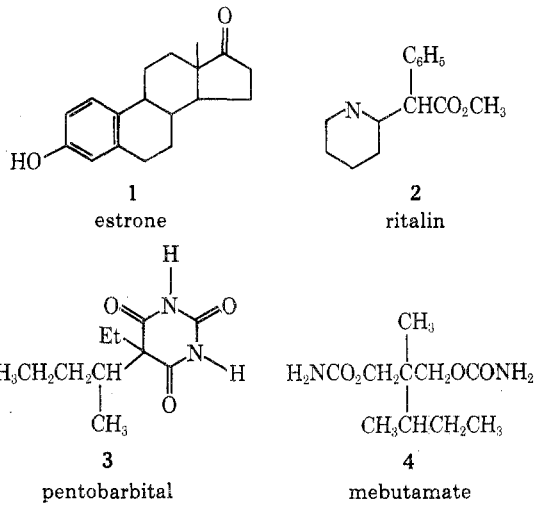
Basis of the Method

Crucial to the proposed method is the following postulate. "*There exists at least one SECONDARY LOSS in a spectrum that will match a PRIMARY LOSS from the molecular ion irrespective of whether the molecular ion is present in the spectrum.*" A PRIMARY LOSS in this context is a composition or mass of a fragment that is lost from the molecular ion. In the spectrum shown in Figure 1 the loss of 55 amu (244 − 189) is an example of a primary loss. The complete set of primary losses for this spectrum is 55, 73, 100, 101, and 111. Additional losses such as 129 amu (244 − 115) are not considered because they exceed an upper loss limit that has been arbitrarily set in the program. Note that primary losses are not known to the program because the molecular ion is unknown.

A SECONDARY LOSS is any member of the union of the two sets of fragments defined below. One set consists of losses from all fragment ions other than any possible molecular ion candidate present in the spectrum. The other set is made up of the compositions or masses in the lower half of the spectrum. Referring to the spectrum in Figure 1 the loss of 18 amu (189 − 171) and the peak at mass 88 are examples of secondary losses.

The basic postulate can be restated as follows. For every mass spectrum there is a high probability that any mass or composition lost from the molecular ion will be also observed as a loss from one fragment ion to another or will itself appear as a fragment ion in the lower mass regions of the spectrum. Initial testing of this postulate has indicated that the secondary losses almost invariably contain at least one element of composition or mass which will combine with an element in the spectrum to give the correct molecular composition or mass, independently of the presence or absence of a peak corresponding to the molecular ion.

Chart I
Structures of Some Compounds Whose Molecular Ions Were Predicted



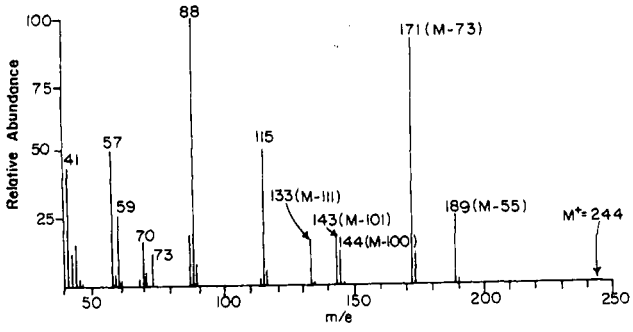


Figure 1. Mass spectrum of dimethylmalonic acid *n*-butyl ester.

An example given in Table I indicates the level of matching of primary losses by secondary losses in the high-resolution mass spectrum of estrone (1) (see Chart I). Each such matching is additional evidence for choice of the correct molecular ion. This is not an isolated example; a high degree of matching exists for most classes of compounds. We point out that some matchings for low-resolution data may be due to numerical accidents (*i.e.*, a difference of 17 mass units may correspond to either a comsitional difference of OH or NH_3). These accidents result in generation of bad candidates that would be rejected in a high-resolution analysis. Although, intuitively, it might seem that this method should generate excessive numbers of candidates, in practice simple evaluation heuristics reject most candidates.

Table I
Matching of Primary Losses by Secondary Losses for Estrone (1)

Primary losses[a] (from the molecular ion)	Other peaks showing identical losses[b]
C_1H_3	214 199 186
C_2H_4	214 199 186
C_2H_5	213
$C_2H_7O_1$	214 213 199
$C_3H_4O_1$	255 242 184
$C_3H_5O_1$	242 233
$C_4H_7O_1$	255 199 186
$C_5H_7O_1$	242
$C_5H_8O_1$	199
$C_5H_{10}O_1$	214 199
$C_4H_{11}O_2$	
$C_6H_8O_1$	255 241 187
$C_6H_9O_1$	242
$C_6H_{11}O_1$	214
$C_5H_{11}O_2$	
$C_5H_{12}O_2$	
$C_5H_{13}O_2$	
$C_7H_{10}O_1$	255

[a] Primary losses are from the molecular ion (270 $C_{18}H_{22}O_2$) of estrone (high-resolution spectrum). Column 2 lists other peaks in the spectrum showing the same losses. [b] Only nominal masses are shown; however, high-resolution masses and elemental compositions derived therefrom were used in the analysis.

Working from this postulate concerning primary and secondary losses it is a rather straightforward task to construct a method capable of generating a set of molecular ion candidates that will include the correct molecular ion in almost all instances. It is not necessary to use all secondary losses and the complete spectrum to obtain reliable molecular ion prediction. Details of the procedure presently used will now be outlined.

Method for Molecular Ion Inference

The proposed method for molecular ion prediction falls within the plan–generate–test paradigm common to many artificial intelligence applications, including Heuristic DENDRAL.[2,3] This strategy seems to be an efficient one for solving the present problem.

(A) Planning Phase. The basic function of the planning phase is to establish the appropriate set of secondary losses needed for the molecular ion generation phase. Initially the raw spectrum of mass-intensity pairs is approximately corrected for ^{13}C isotope contributions and then transformed into a list of "mass clusters" as dictated by the natural clustering of the spectral peaks. For our purposes a cluster is a group of peaks in which successive peaks are separated by less than three mass units. Before the clustering procedure is performed the spectrum is passed through a selective thresholding procedure which ensures that eventual clusters do not join because of the presence of ions which are low in abundance compared with the major ions in the local region of the spectrum. From each cluster the three most intense peaks are kept; the rest are discarded. Any peaks that are less than one-third the most intense peak in each cluster are also arbitrarily deleted to give a reduced spectrum. In this transformation relative intensity filtering is achieved without the exclusion of important low intensity peaks that may be present at the high mass end of a spectrum. For the dimethylmalonic acid butyl ester spectrum (Figure 1, molecular ion absent) the reduced low-resolution spectrum of mass–intensity pairs is (41, 44), (57, 48), (59, 23), (70, 15), (73, 12), (88, 100), (115, 46), (133, 14), (143, 18), (144, 16), (171, 95), (189, 26).

Before actually generating the secondary loss set a number of consistency checks are made to determine whether there are any molecular ion candidates present in the spectrum. This is done by checking to see if there are no "bad losses" (see Tables II and III) to neighboring clusters, and if there is a peak in the top cluster of the appropriate parity with respect to the nitrogen content flag (set after the series analysis phase, below). Note that the entries in Table III are selected at the discretion of the user. In esence any molecular ion candidate present in the spectrum must pass the same tests as described below for generated molecular ion candidates. Also, for high-resolution data, any nonsense compositions resulting from numerical accidents are filtered out. In the case of dimethylmalonic acid *n*-butyl ester (Figure 1) the dominant odd mass series 57, 73, 115, 143, 171 (components differ by multiples of 14 or CH_2) together with the greater number of odd than even peaks in the reduced spectrum indicate firstly that there is not an odd

Table II
Bad Compositions Used by the Secondary Loss Generation Phase and the Candidate Filter

Composition	Bad losses
$C_xH_yO_zN_n$	$x < 0$ or $y < 0$ or $z < 0$[a] or $n < 0$[a]
C_x	$x > 0$ $y = 0$ $z = 0$ $n = 0$
H_y	$x = 0$ $y > 2$ $z = 0$ $n = 0$
O_z	$x = 0$ $y = 0$ $z > 0$ $n = 0$
N_n	$x = 0$ $y = 0$ $z = 0$ $n > 0$
H_yO_z	$x = 0$ $z > y$ or $y > 2z$
$C_xH_yO_zN_n$	$x \neq 0$ $y > (2x + 3)$
$C_xH_yO_zN_n$	$x \neq 0$ $z > (x + 1)$
$C_xH_yO_zN_n$	$x \neq 0$ $y \neq 0$ $x > y + 1$

[a] The negative restriction is necessary to exclude losses where the mass difference is valid but the composition difference has negative components.

Table III
Bad and Poor Losses Used by the Secondary Loss Generation Phase and the Candidate Filter

Bad Losses[a]												
4	5	6	7	8	9	10	11	12	13	21	22	23
24	25	26	37	38	50	51	52	53	65	66		
Poor First Losses												
19	39	40	54	62	64	67	68	70	82			
83	84	86	88	89	90	91	92	93	94	95		
96	98	99	103	104	105	107	108	109	110			

[a] Bad losses are taken to be those that can almost always be rejected on chemical grounds. Poor losses are generally of low probability but cannot be excluded on chemical grounds.

Table IV
Secondary Losses from Peak Clusters in Dimethylmalonic Acid *n*-Butyl Ester (Figure 1)

Peak cluster	Loss set
189	(18 45 56 74 101)
171	(27 28 56 83 98 101 112 114)
144 and 143	(29 56 71 74 85 103 28 55 70 73 84 86 102)
133	(18 45 60 63 74 76 92)
115	(27 42 45 56 58 74)
88	(15 18 29 31 47)
73 and 70	(14 16 32 29)
59 and 57	(18 16)

number of nitrogens present and also that there is no molecular ion peak.

The secondary losses are generated by exhaustive application of the following procedure to the reduced spectrum. Successive mass and/or composition "inter-cluster" losses are generated for all ions in all clusters of the reduced spectrum. The secondary losses from all clusters in the dimethylmalonic acid *n*-butyl ester spectrum are shown in Table IV. As the losses are calculated impossible composition differences, if one is examining a high-resolution mass spectrum (*e.g.*, a loss of two carbons with a gain of six hydrogens), and bad mass differences (*e.g.*, the loss of 10 mass units generated from masses 143 and 133 in our example) are rejected (see Tables II and III). Other restrictions that have arbitrarily been placed on the losses are that they are limited in magnitude to less than 115 mass units and even if they pass this restriction they cannot be greater than approximately half the largest mass observed in the spectrum. These losses (excluding those from any possible molecular ion candidate) are combined with the list of fragment masses in the lower half of the spectrum. The resulting set is the secondary loss set referred to previously. In some instances this set may contain more than 50 entries. The secondary loss set for the dimethylmalonic acid *n*-butyl ester example (Figure 1) is 14, 15, 16, 18, 27, 28, 29, 31, 32, 41, 42, 45, 46, 47, 55, 56, 57, 58, 59, 60, 63, 70, 71, 73, 74, 76, 83, 84, 85, 86, 87, 88, 92, 101, 102, 103, 112, 114.

At this stage in the procedure two approaches are available, each with certain advantages. If one wishes to avoid using all the secondary losses to generate molecular ion candidates a set of loss series can be formed. The criterion for establishing a loss series is that elements of the series differ in composition multiples of CH_2 or, in the case of low-resolution data, there exist differences that are multiples of 14. The complete set of loss series for dimethylmalonic acid *n*-butyl ester generated from its secondary loss set is given in Table V. To obtain a restricted set of secondary losses for molecular ion prediction one can choose to use

Table V
Loss Series Formed from Secondary Losses for Dimethylmalonic Acid *n*-Butyl Ester (Figure 1)

Loss series no.	Series list
1	(28 42 56 70 84 98 112)
2	(15 29 57 71 85)
3	(58 86 114)
4	(18 32 46 60 74 88 102)
5	(27 41 55 83)
6	(31 45 59 73 87 101)
7	(47 103)

Table VI
Ranked Candidate Molecular Ions for Dimethylmalonic Acid *n*-Butyl Ester (Figure 1)

Candidate	Ranking index
244	100
258	52
246	51
190	37

only the lower members (say the first two or three) of each loss series for molecular ion generation.

If instead of using a restricted set of secondary losses one chooses to use the complete set a likelihood ranking can be given to each of the molecular ion candidates. Details of the method used are given in the discussion of the filtering phase which follows the section on molecular ion generation below.

With the chosen set of secondary losses in hand the program switches to a generative mode.

(B) Molecular Ion Generation. Three main pieces of information are passed from the planning mode to the molecular ion generation mode. They are the chosen set of secondary losses, a list of the compositions or masses in the upper half of the reduced spectrum, and an indication as to whether there is an odd number of nitrogens present in the molecular ion (see above discussion of planning phase).

To generate the molecular ion candidates each member of the secondary loss list is added to each member of the upper half of the reduced spectrum of the appropriate parity. In our example (Figure 1) for dimethylmalonic acid *n*-butyl ester the secondary loss "55" when added to 133, 143, 171, and 189 gives candidates with masses of 198, 226, and 244. Candidates generated that have a molecular weight or composition less than the ion at highest mass (excluding any impurity peaks that have been detected) are immediately excluded. The program copes with impurity peaks at the high mass end in high-resolution spectra by checking the compositions of losses from each of the peaks in the top three clusters against the "bad compositions" listed in Tables II and III. Any impurity peaks that are detected by accurate mass measurements are removed from the spectrum. No attempt is made to detect impurity peaks in low-resolution spectra.

(C) Filtering and Ranking of Molecular Ion Candidates. The filtering or testing part of the method performs three main functions. It uses a list of bad losses (Table III) and compositions (Table II) to reject some of the generated molecular ion candidates. Those candidates rejected in the dimethylmalonic acid *n*-butyl ester example are 198, 200, 220, 202, 212, 236, 226, 248, and 252. For instance the candidate 198 shows a loss of 9 amu (198 − 189) and so is unacceptable. The remaining molecular ion candidates are separated into a probable list and an unlikely list. To perform this latter operation each candidate is tested against two

Table VII
Sample Set of Results Showing Molecular Ion Prediction and Ranking

Compound	Mol formula	Highest mass present	Fragment missing[a]	Mol wt	Ranked at no.[b]
Ritalin (2)	$C_{14}H_{19}NO_2$	172 (M − 61)	$C_2H_5O_2$	233	4
Pentobarbital (3)	$C_{11}H_{18}N_2O_4$	197 (M − 29)	C_2H_5	226	2
Mebutamate (4)	$C_{10}H_{20}N_2O_4$	175 (M − 57)	C_4H_9	232	3
Tridecan-7-one	$C_{13}H_{26}O$	155 (M − 43)	C_3H_7	198	4
Succinic acid methyl ester	$C_6H_{10}O_4$	116 (M − 30)	CH_2O	146	2
Caprylic acid methyl ester	$C_9H_{18}O_2$	129 (M − 29)	C_2H_5	158	3
Glutaric acid methyl ester	$C_7H_{12}O_4$	129 (M − 31)	CH_3O	160	1
Maleic acid butyl ester	$C_{12}H_{20}O_4$	173 (M − 55)	C_4H_7	228	2
N-TFA α-alanine[c] butyl ester	$C_9H_{14}NO_3F_3$	186 (M − 55)	C_4H_7	241	2
N-TFA norleucine butyl ester	$C_{12}H_{20}NO_3F_3$	227 (M − 56)	C_4H_8	283	2
N-TFA valine butyl ester	$C_{11}H_{18}NO_3F_3$	227 (M − 42)	C_3H_6	269	2
N-TFA threonine butyl ester	$C_{12}H_{15}NO_5F_6$	323 (M − 44)	C_3H_8	367	1
N-TFA phenylalanine butyl ester	$C_{15}H_{18}NO_3F_3$	216 (M − 101)	$C_5H_9O_2$	317	4
n-Undecyl alcohol	$C_{11}H_{24}O$	154 (M − 18)	H_2O	172	1
4-Methyloctan-4-ol	$C_9H_{20}O$	129 (M − 15)	CH_3	144	1

[a] These fragment composition losses from the molecular ion are only postulated. Their validity could only be confirmed by high-resolution studies. [b] Note "ranked at number 1" is the program's best choice for a molecular ion candidate. [c] TFA refers to the trifluoroacetyl derivative.

"poor loss" lists. The first set contains a list of poor first losses (Table III) extending up to 115. A first loss is the smallest loss from a candidate molecular ion (*e.g.*, 55 amu for mol wt 244 in Figure 1). The second set of poor losses is a subset of the first set. Losses from each generated ion to every ion in the top three clusters of the spectrum are compared with this second set of losses. If a loss generated from a molecular ion candidate is found to be present in either of these poor loss lists it is given an unlikely ranking; otherwise it is placed on the list of probable molecular ion candidates. The following candidates were given an unlikely ranking for our ester example: 216, 218, 228, 230, 234, 292, and 274. We find that 216 is rejected because it shows a loss of 27 amu (Tables II and III) to the peak at 189, for example.

Following through our example for dimethylmalonic acid *n*-butyl ester, we find that of the three candidates (198, 226, 244) generated by the secondary loss 55, only 244 is not rejected by the filter. A loss of 9 is shown by 198 and a loss of 37 is shown by 226—both of which have been deemed unacceptable losses.

If the option has been invoked to use all secondary losses (including even losses minus one hydrogen or one mass unit, *e.g.*, if 88 is in the even list 87 is included in the odd list) the probable molecular ion candidates are ranked. Peak intensities are used to obtain this ranking. First, a cumulative intensity measure is found for each secondary loss. This is done by taking the sum of the average intensities for each pair of peaks involved in each secondary loss transition. A reduced weight is assigned to the losses 14, 28, etc. Intensities of peaks in the lower half of the spectrum are also added to any secondary losses with which they correspond. Then by adding in the intensities of peaks in the upper half of the spectrum that generate a given molecular ion candidate it is possible to obtain a likelihood rating for each candidate. In the complete analysis for dimethylmalonic acid *n*-butyl ester (Figure 1) the top four candidates and their associated ratings were as shown in Table VI. The correct molecular ion 244 gets the highest ranking (100). Inspection of the losses from each cluster gives an explanation for this. Losses from 189, 171, 144, 143, 133, and 115 include 55, 56, 73, 74, and 101 a number of times, and since these losses match those from the molecular ion (244) they will make a high contribution to the ranking for the candidate 244. This ranking method is particularly effective when the molecular ion is present, as might be expected because of the high degree of matching of primary losses by secondary losses. It also works well for cases where the molecular ion is absent from the spectrum. Details of ranking performance are given in Tables VII and VIII.

Details of the Computer Program

To cope efficiently with the data manipulations required, the method for molecular ion prediction has been implemented in the list processing language LISP[7] (versions exist in LISP 1.5 and INTERLISP). The symbolic and list processing capabilities are highly desirable when working with the composition lists of high-resolution mass spectra.

The program has been structured to find molecular ions in either a high- or low-resolution mass spectrum. With

Table VIII
A Summary of Ranking Results Is Given for Various Classes of Compounds Tested

Class	Mol ion in top 3	Mol ion in top 5	Total no. of compds
Amines	62	67	68
Alcohols[a]	51	54	57
Ketones	42	44	44
Ethers	33	34	34
Acetals	13	14	14
Amino acid derivatives	9	11	13
Thioethers	11	12	12
Drug compounds	6	8	9
Methyl esters	6	8	8
Butyl esters	4	5	6

[a] Program failed to generate the correct candidate for one alcohol, *tert*-butyl alcohol.

low-resolution spectra the program is given a list of mass–intensity pairs. High-resolution data are accepted as a list of mass–intensity–composition triplets. No other parameterization is necessary. However, changes to the filter set and other parameters can easily be made.

On a normal low-resolution run the program takes between 1 and 4 sec of CPU time on an IBM 360/67 to produce the list of molecular ion candidates. High-resolution determinations which always involve considerable manipulation of composition lists may take from 30 to 60 sec.

Results and Discussion

Referring to the results for dimethylmalonic acid *n*-butyl ester it can be seen that the primary losses (−55), (−73), and (−101) are matched by secondary losses. Closer inspection of the molecular structure for this ester reveals the link between the losses and structural components, *e.g.*, –COOBu has a mass of 101 and –OBu has a mass of 73. The loss of 55 may correspond to the structural unit C_4H_7.

To test the validity and performance of the method for molecular ion prediction, spectra from a wide variety of classes of compounds have been tested, particularly for low-resolution data. The only high-resolution spectra tried have been estrogens, pregnanes, and progesterones most of which show molecular ions.

Among the low-resolution spectra tested (Tables VII and VIII) were derivatized amino acids, barbiturates, and other drug compounds, aliphatic esters, acetals, alcohols, amines, ethers, and ketones, many of which have no molecular ions.

The results tabulated in Tables VII and VIII give some idea of the effectiveness of the program for molecular ion prediction. When the molecular ion is present as for most amines and ketones, and some ethers and alcohols, the program almost always predicts it as the most likely candidate or at least in the top three ranked candidates. It should be emphasized that this high ranking is purely on the basis of the cumulative intensity index and that the same high ranking would have been obtained even if the molecular ions had been removed from the spectra.

For the derivatized amino acids, aliphatic esters, and barbiturates which do not show molecular ions, performance in prediction was still good. Ranking usually placed the correct candidate within the top five of those predicted. Sometimes when (M − 73) is the highest mass, ranking is poor. This may be due to the decreasing stability of fragments with increasing mass. Alcohols and acetals which quite often show (M − 1) or (M − 15) ions rather than a molecular ion are effectively handled by the program. Results so far indicate that there is a high probability (greater than 95% for 250 spectra of more than 10 classes) that the molecular ion will be in the top five candidates, a workable number for a computer program employed in molecular structure analysis. Of the 250 spectra tested 100 did not show significant molecular ions.

It should be emphasized that the present method for molecular ion prediction has been developed primarily as part of a larger program devoted to the computer interpretation of mass spectra. It is offered as a method for determining molecular ions in routine, conventional electron impact mass spectra, when methods such as chemical ionization and field ionization are not available or applicable.

Limitations of the Program

The method as stated in its most general form can cope with most molecular ion prediction problems. However, by working from a reduced spectrum and by not using all secondary losses an error in prediction may occasionally occur. Also an unusual loss currently on the list of bad losses (*e.g.*, 4 amu) may cause rejection of the correct molecular ion candidate.

The program as presently written may have problems with low-resolution spectra containing more than one chlorine or bromine because of heavy isotope contributions. Impurity peaks at high mass in low-resolution gc–mass spectral runs will also cause problems. Furthermore, if the largest mass peak in the spectrum is less than half the molecular weight (M) or there is a peak at say (M − 101) at highest mass for low M the program will fail. The unusual problems in which the program may fail are likely to be the same problems with which the chemist will have considerable difficulty if given no additional information.

Acknowledgments. One of us (R.G.D.) gratefully acknowledges the support of a CSIRO Postdoctoral Studentship. We are indebted to Dr. Alan Duffield for supplying many of the test spectra. We would also like to thank Dr. N. Sridharan for his comments. This work was supported by a grant (No. RR-612-05A1) from the National Institutes of Health.

References and Notes

(1) Parts XII: L. M. Masinter, N. S. Sridharan, J. Lederberg, and D. H. Smith, *J. Amer. Chem. Soc.*, **96**, 7702 (1974). Part XIII: L. M. Masinter, N. S. Sridharan, R. E. Carhart, and D. H. Smith, *ibid.*, **96**, 7714 (1974). Part XI: R. E. Carhart and C. Djerassi, *J. Chem. Soc., Perkin Trans. 2*, 1753 (1973).
(2) B. G. Buchanan, A. M. Duffield, and A. V. Robertson in "Mass Spectrometry—Techniques and Applications," G. W. A. Milne, Ed., Wiley, New York, N.Y., 1971, p 121.
(3) D. H. Smith, B. G. Buchanan, R. S. Engelmore, A. M. Duffield, A. Yeo, E. A. Feigenbaum, J. Lederberg, and C. Djerassi, *J. Amer. Chem. Soc.*, **94**, 5962 (1972).
(4) R. Venkataraghavan, F. W. McLafferty, and G. E. Van Lear, *Org. Mass Spectrom.*, **2**, 1 (1969).
(5) K. Biemann and W. McMurray, *Tetrahedron Lett.*, 647 (1965).
(6) A. Jardine, R. I. Reed, and M. E. Silva, *Org. Mass Spectrom.*, **7**, 601 (1973).
(7) C. Weissman, "LISP 1.5 Primer," Dickenson, Belmont, Calif., 1967. We will make available to interested persons copies of the program in INTERLISP. However, this will only be worthwhile for those who have access to a computer facility which is committed to maintaining the INTERLISP language. The Stanford University Medical Experimental (SUMEX) computer facility has been established to encourage resource sharing of complex computer programs. For details on how to gain access to this and related programs please write to the authors or to Professor Joshua Lederberg, Director, SUMEX Project, Department of Genetics, Stanford University Medical School, Stanford, Calif. 94305.