Report 77-24
Stanford -- KSL

Scientific DataLink

Computer-Assisted Analysis of the High Resolution Mass Spectra of Macrolide Antibiotics. Kent Morrill, Dennis H. Smith, Carl Djerassi, 1977

77-24

COMPUTER-ASSISTED ANALYSIS OF THE HIGH RESOLUTION MASS SPECTRA OF MACROLIDE ANTIBIOTICS[1,2]

Kent Morrill, Dennis H. Smith and Carl Djerassi*

(Contribution from the Department of Chemistry, Stanford University, Stanford, CA 94305)

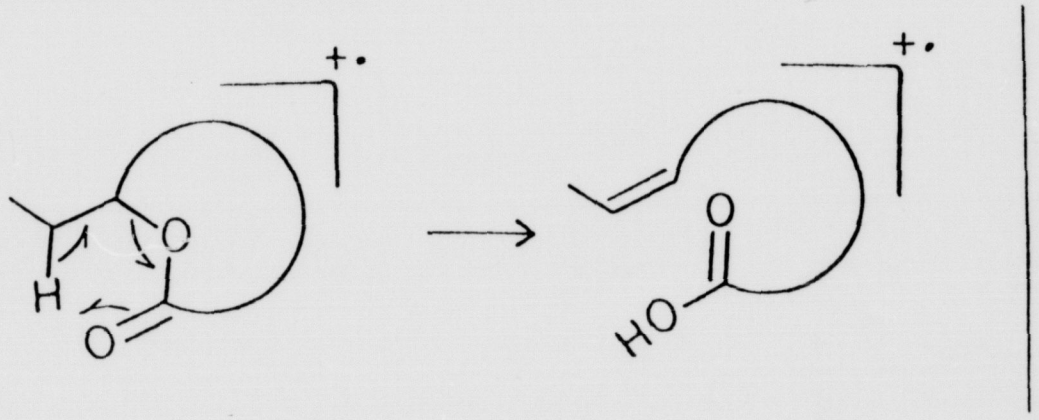
Abstract. The high resolution mass spectra of five 12-membered macrolide aglycones related to methynolide were analyzed with the aid of the Meta-DENDRAL subprogram INTSUM. Metastable defocusing data obtained for several prominent ions in the spectra of two of these compounds support stepwise cleavage processes subsequent to initial ring opening. Their behavior in the mass spectrometer proved to be highly dependent on the substituents present and their location on the macrocyclic ring. A series of empirical mass spectral fragmentation rules, which were derived from these data, proved useful in differentiating between the mass spectra of closely related isomeric structures. These rules were also useful in the analysis of the fragmentation patterns of several 14- and 16-membered macrolide aglycones.

I. INTRODUCTION

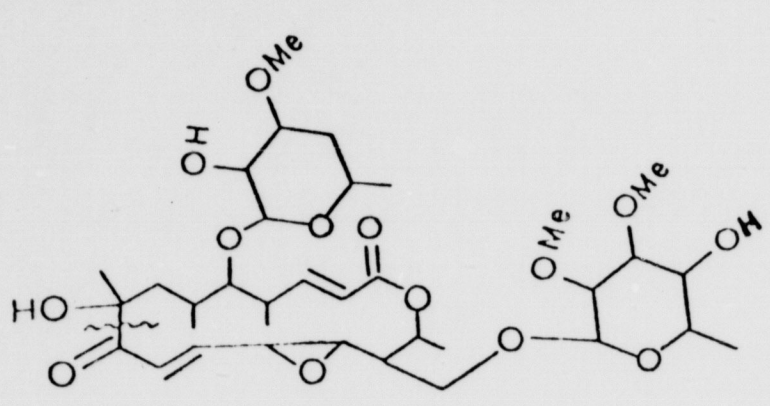
The non-polyene macrolide antibiotics are a class of medically important compounds characterized by a macrocyclic lactone system containing not more than two conjugated double bonds.[3] During the past ten years many researchers have tried to use mass spectrometry to obtain structural information about these compounds.[4-9] However, this use has been hampered by the tendency of the spectra to be dominated by processes involving the attached sugars. Amino sugars have an especially strong effect since virtually all of the peaks above 20% of the base peak contain nitrogen when one is present.[4-9] The fact that many macrolide antibiotics contain more than one sugar further obscures the structural information provided by mass spectrometry. The use of chemical ionization techniques[8] and of N-oxide derivatives[9] of the antibiotics has been helpful, but even with these techniques, mass spectrometry has mostly proven

useful in the determination of molecular weights, the masses of the aglycone and sugar moieties, and in assigning functional groups to the sugar or to the aglycone regions of a molecule.[4-9]

The portion of these antibiotics which is of primary structural interest is the macrocyclic aglycone moiety. However, the conclusions which have been reached about the behavior of the macrolide aglycone in the mass spectrometer are not definitive, partly because of the difficulties outlined above. Mitscher and co-workers concluded from the high resolution mass spectra of a series of compounds related to erythronolide B that structurally important processes for these compounds are initiated by ring opening via a McLafferty rearrangement involving the lactone linkage:[6a,9]



Kinchart and Van Lear observed fragments in the mass spectrum of chalcomycin which were apparently the result of a preferred cleavage α to the keto group:[5]



Achenbach and Karl, however, were not successful in using such relations to locate the positions of various functional groups (epoxy, keto, hydroxyl, etc.) in several aglycones and concluded that they are not general.[7a]

Recently, the utility of the computer program (Meta-DENDRAL) to assist in determining fragmentation rules for known compounds has been demonstrated.[10-12] This program, in particular the INTSUM phase,[8] was applied to the high resolution mass spectra of the macrolide antibiotics. We sought to reduce the effort and time required to correlate the extensive mass spectral data with the many plausible ways of fragmenting the molecules.

Fragmentation rules, whether generated automatically[10] or manually, can be tested for their explanatory power (their ability to predict correct mass spectral peaks) and, more importantly, for analytical purposes, in terms of their discriminatory power (their ability to differentiate between possible candidate structures based on the mass spectrum of a compound). This is done by comparing an actual mass

spectrum and the spectra predicted by the application of the rules to possible structures.

II. RESULTS AND DISCUSSION

A. 12-Membered Macrolides

The unit resolution mass spectra of compounds 1 (3-dehydromethynolide), 2 (methynolide), 3 (neomethynolide), 4 (dihydromethynolide) and 5 (dihydroneomethynolide)[13,14] are shown in Figure 1. The high resolution mass spectral data[15] for these compounds were subjected to analysis using INTSUM - a computer program which generates, within constraints defined by the user, an exhaustive list of all of the mass spectral fragmentations which would give rise to the observed mass spectral peaks.[10] The constraints consist of the maximum number of steps to be allowed in a process, the maximum number of bonds allowed to cleave in a single step, whether more than one (non-hydrogen) bond to the same atom is allowed to undergo fission, and whether double, triple or aromatic bonds are allowed to break. It is also possible to specify the number of hydrogen transfers to and from a fragment as well as neutral losses such as water or carbon monoxide.[12]

For simplicity, only one-step processes allowing a maximum of two bond cleavages were considered. Thus, more complex processes including for example, fragmentations involving loss of side chains in combination with ring cleavage were not considered. Hydrogen transfers of 2, 1, 0, -1 and -2 were considered. Because all of the compounds had several oxygen functionalities including at least one hydroxyl group, loss of up to two water molecules was permitted. Loss of carbon monoxide was not considered because it was noted during initial studies (including such losses) and in every instance where a process involving

of carbon monoxide was proposed by INTSUM, there was an alternative simple one-step explanation which involved the loss of the same atoms as a single fragment.

The initial examination of the INTSUM data for compounds 1-5 revealed that most of the mass spectral peaks had many possible explanations. The redundancy persisted even when only the most simple explanations were considered. The question was often not one of "either/or" since it was possible in some cases to demonstrate that several processes were operative, giving rise to a single mass spectral peak in certain of the compounds but not in others. This redundancy could help explain some of the uncertainty which has been reported regarding the dominant mass spectral fragmentations of the macrolide antibiotics.[5-7] It is highly unlikely that a mass spectrometrist would exhaustively examine a structure for all of the possible explanations for each mass spectral peak. Alternative, plausible explanations would be overlooked.

Because of the large number of possible explanations for most of the mass spectral peaks, we restricted our attention to the important ($\geq$1% Σ_{40}, m/e $\geq$95) mass spectral peaks which could be clearly correlated with the structures of the compounds. The best explanations, when redundancies arise, are taken to be those for which the most consistent evidence exists for the set of five compounds.

Table I shows that macrolides 1-5 differ in three respects through the series. Macrolides 1, 2 and 3 contain an 8-9 carbon-carbon double bond; 1, 2 and 4 have a hydroxyl on carbon 10; 3 and 5 contain a 12-hydroxyl and 3 contains a carbonyl on carbon 3. Classified according to these differences, they form the six overlapping groups listed in Table I. As will be shown, the INTSUM data for these compounds are best understood in light of these classifications.

Table I. Macrolides 1-5 Grouped by Common Characteristics.

Group	Compounds	Common Charact.
1	1, 2, 3, 4, 5	common skeleton
2	1, 2, 3	8-9 double bond
3	1, 2, 4	10-hydroxyl
4	3, 5	12-hydroxyl
5	4, 5	8-9 single bond
6	3	3-carbonyl

The important fragmentations of macrolides 1-5 are summarized in Table II. Each process is represented by one or more pairs of numbers enclosed by parentheses. The numbers correspond to the numbers of the atoms which are separated when the bond between them is cleaved. The number listed first represents the atom on the charge carrying fragment. For example, in process A the bonds between atoms 11 and 14, and 7 and 6 respectively are cleaved with charge retention on the fragment containing atoms 11 and 7 as illustrated schematically in Figure 2.

Only a limited number of processes are observed to occur in all five of the 12 membered macrolides studied (group 1 in Table I). One of them, process A can be envisioned as involving McLafferty rearrangement similar to that described by Mitscher and co-workers.[6]

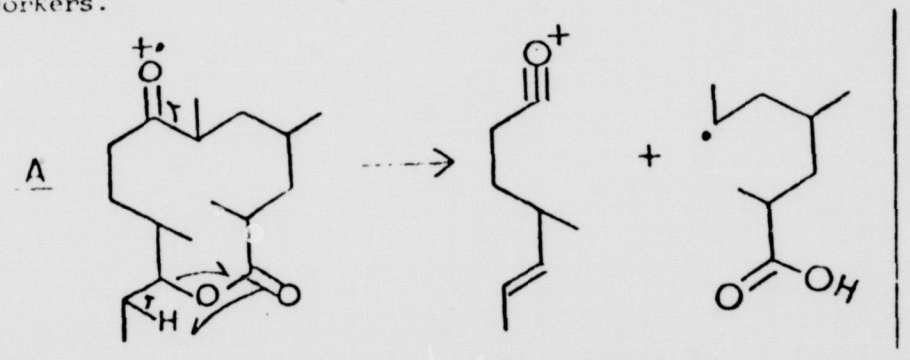
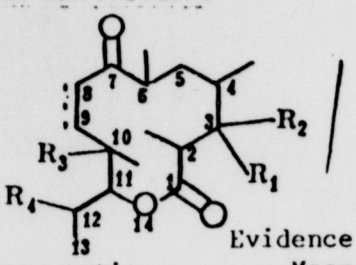


Table II. Fragmentation Processes for which Evidence Exists in the Mass Spectra of Macrolides.[1,5a]



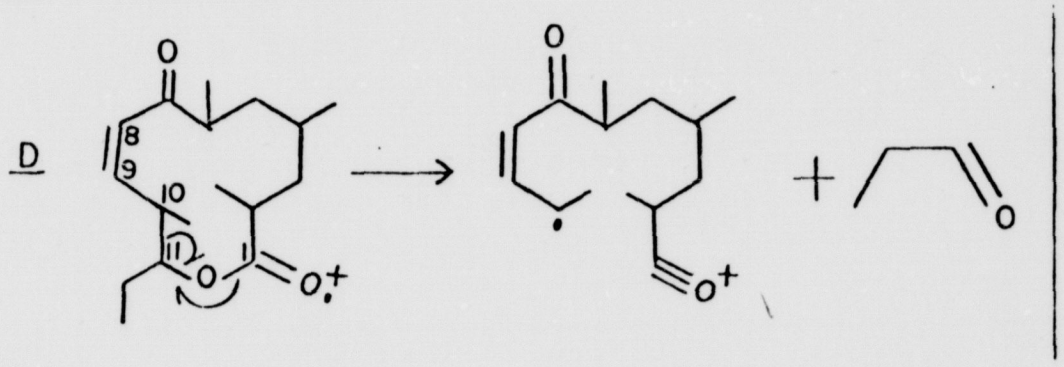
Process	Ion[b]	Representation	Evidence in 12-Membered Macrolides ($\%\Sigma_{40}$)					Neutral Transfers
			1	2	3	4	5	
A	a	(11 14) (7 6)	3.9	1.9	2.5	3.9	1.8	-H
B	b	(1 14) (5 6)	1.1	5.9	1.3	2.0	1.5	-H
			-	2.7	2.0	.8	.7	-H H_2O
C	c	(10 11) (5 4)	3.9	1.9	.9	3.9	1.3	-H
	cc		3.2	3.8	-	.7	.2	$-H_2O$
D	d	(10 11) (1 14)	2.7	3.3	2.9	-	-	-
			.8	1.2	.4	.1	-	$-H_2O$
E	e	(10 11) (3 2)	5.6	2.2	2.0	-	.1	-
F	f	(10 11) (7 6)	3.3[c]	2.6[c]	1.6	.1[c]	.8	-
	ff		1.8[c]	.8[c]	-	2.9[c]	1.0	-H
G	g	(10 11) (6 5)	1.9	5.9	4.9	-	.4	+H
	gg		1.6	2.3	.4	6.4	.2	-
H	h	(3 4) (7 6)	-	-	2.8[c]	-	.4[c]	-H
I	i	(10 11) (4 3)	5.5	-	-	-	-	+H-Me
	ii		1.4	-	-	-	-	-Me

a 1 ^{8,9}, $R_1 = R_2 = =O$, $R_3 = OH$, $R_4 = H$; 2 ^{8,9}, $R_1 = OH$, $R_2 = H$, $R_3 = OH$, $R_4 = H$; 3 ^{8,9}, $R_1 = OH$, $R_2 = H$, $R_3 = H$, $R_4 = OH$; 4 $R_1 = OH$, $R_2 = H$, $R_3 = OH$, $R_4 = H$; 5 $R_1 = OH$, $R_2 = H$, $R_3 = H$, $R_4 = OH$.

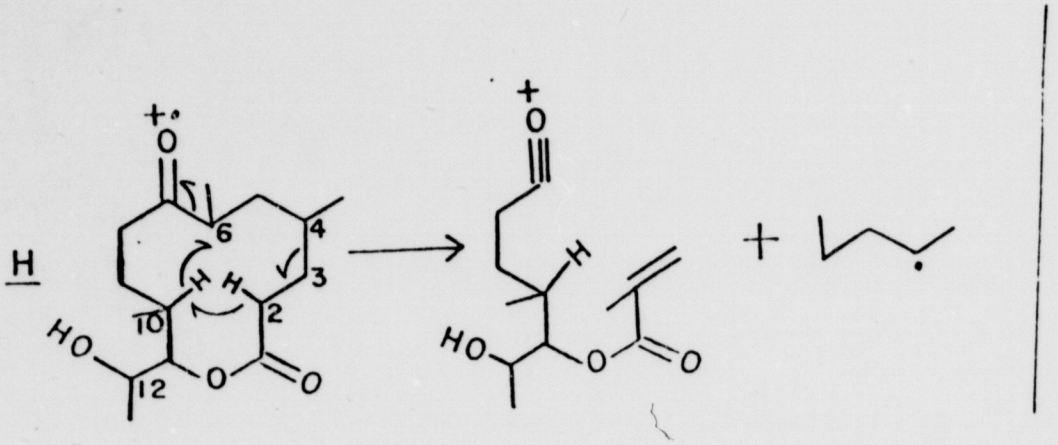
b See Figure 1.

c No other plausible explanations.

The 8-9 carbon carbon double bond in 1, 2 and 3 (group 2 in Table I) has a very strong effect on their mass spectra as demonstrated by the processes which occur most strongly in these compounds. These processes all involve cleavage allylic to the 10-11 bond. Process D is a good example, and a plausible mechanism is illustrated below:



There are several processes which seem to be due to the change in position of the hydroxyl group which is on C-10 in 1, 2 and 4 and C-12 in 3 and 5. Process C is accompanied by loss of water only in 1, 2 and 4. The C-10 OH is implicated because the hydroxyl group on C-10 of these compounds is absent in 3 and 5. Process H occurs in 3 and 5 without alternative explanations. It does not occur at all in 1, 2 and 4. This is intriguing because these two groups of compounds differ in the position of the hydroxyl group which is on C-10 in 1, 2 and 4 and C-12 in 3 and 5. Although C-10 is formally remote from the bonds which break in the process, the fact that process H occurs with loss of hydrogen provides a clue. A mechanism which transfers the hydrogen from C-10 to the neutral fragment would explain these results for 3 and 5 since there is no C-10 hydrogen in 1, 2 and 4. Such a mechanism is illustrated below. The intermediate allylic radical formed on C-10 in 3 would make this mechanism especially favorable and could explain why the resulting mass spectral peak is more intense for 3 and 5.



The only compound which has a carbonyl group at C-3 is 1. The difference which this produces in its mass spectrum can be seen in several cases where processes involving bonds adjacent to C-3 are involved. For example, process E involves a 2-3 bond cleavage and is strongest for 1. There is evidence for the occurrence of process I in 1 if a loss of methyl radical is permitted.[16] There are other processes that are eliminated or which produce peaks of reduced intensity in the mass spectrum of 1. For example, process B with loss of water does not occur for 1, presumably because 1 is the only compound lacking a 3-hydroxyl group.

In conclusion, there are a few processes which can be accounted for by the basic skeleton of 1-5. The most important, easily identified processes are more closely associated with the functional groups present. The strongest fragmentation directing functionalities are the 8-9 carbon-carbon double bond of 1-3, promoting allylic cleavages, together with keto groups on C-7 of all five compounds and C-3 of 1. With so few compounds, however, a detailed hierarchy of effects of various functional groups is difficult to establish.

The correlations which can be made agree generally with the conclusions drawn by Mitscher and co-workers about the McLafferty rearrangement involving the lactone linkage[6] and by Rinehart and Van Lear concerning cleavage α to ketones.[5]

B. Metastable Defocusing Studies

The above analysis of the INTSUM results provides no information on mechanistic details. "Single Step" processes reflect only the bonds ultimately broken to yield a fragment.[10] The question of concertedness vs. stepwise loss of small fragments eventually yielding the observed fragment is important, however, for a deeper understanding of any process which involves widely separated portions of a molecule. Does such a process occur with near simultaneous cleavage of the indicated bonds, or does the ring open by cleavage of one bond and the resulting ion degrade in a stepwise fashion to the observed fragments?

In order to help answer the above questions, the metastable defocusing data[17] were obtained for some of the important mass spectral ions of 1 and 4. These data are presented in Table III and show that most of the daughter ions examined had several parents other than the molecular ion at relatively long times subsequent to ionization. Corresponding conventional metastable transitions were also observed in the low resolution mass spectra of 1 and 4 supporting many of these relationships.

The interrelationships of the daughter and parent ions indicated by the metastable data support the hypothesis that many of the important mass spectral ions are the result of stepwise degradations, sometimes via different pathways. For example, an important pathway to m/e 226 ($C_{13}H_{22}O_3$) indicated by the metastable defocusing data for 4 is via the m/e 296 $(C_{17}H_{28}O_4)M^+-H_2O$ ion through two equivalent branches which contain the same two steps in reverse order. This would

Table III. Metastable Defocusing Data for 3-Dehydromethylnolide (1) and Dihydromethynolide (4).

Compound	Daughter Ion	Parent Ion(s)	Loss Of
3-dehydro-methynolide (1)	268 ($C_{14}H_{20}O_5$)	310 (M^+, $C_{17}H_{26}O_5$)	C_3H_6
	252 ($C_{14}H_{20}O_4$)	310 ($C_{17}H_{26}O_5$)	C_3H_6O
	234 ($C_{14}H_{18}O_3$)	252 ($C_{14}H_{20}O_4$)	H_2O
	196 ($C_{11}H_{16}O_3$)	252 ($C_{14}H_{20}O_4$)	C_3H_4O
	181 ($C_{10}H_{13}O_3$)	196 ($C_{11}H_{16}O_3$)	CH_3
	154 ($C_9H_{14}O_2$)	196 ($C_{11}H_{16}O_3$)	C_2H_2O
	139 ($C_8H_{11}O_2$)	268 (w[a], $C_{14}H_{20}O_5$)	$C_6H_9O_3$
		212 (w, $C_{11}H_{16}O_4$)	$C_3H_5O_2$
		196 (s, $C_{11}H_{16}O_3$)	C_2H_5O
		168 (w, $C_{10}H_{16}O_2$)	C_2H_5
		157 (m, $C_8H_{13}O_3$)	H_2O
		154 (s, $C_9H_{14}O_2$)	CH_3
	112 ($C_6H_8O_2$)	268 (s, $C_{14}H_{20}O_5$)	$C_8H_{12}O_3$
		252 (w, $C_{14}H_{20}O_4$)	$C_8H_{12}O_2$
		234 (w, $C_{14}H_{18}O_3$)	$C_8H_{10}O$
		154 (w, $C_9H_{14}O_2$)	C_3H_6
		130 (w, $C_6H_{10}O_3$)	H_2O
	98 ($C_5H_6O_2$)	252 (s, $C_{14}H_{20}O_4$)	$C_9H_{14}O_2$
		168 (m, $C_{10}H_{16}O_2$)	C_5H_{10}
		140 (w, $C_8H_{12}O_2$)	C_3H_7
		126 (w, $C_7H_{10}O_2$)	C_2H_4

[a] s=50-100% base peak; m=30-50% base peak; w=10-30% base peak in metastable spectrum.[17]

(Table III continued)

Compound	Daughter Ion	Parent Ion(s)	Loss Of
dihydro-methynolide (4)	268 ($C_{16}H_{28}O_3$)	296 (M^+-H_2O, $C_{17}H_{28}O_4$)	CO
	254 ($C_{14}H_{22}O_4$)	312 (w, $C_{17}H_{28}O_5$)	C_3H_6O
		296 (s, $C_{17}H_{28}O_4$)	C_3H_6
	226 ($C_{13}H_{22}O_3$)	296 (m, $C_{17}H_{28}O_4$)	C_4H_6O
		268 (s, $C_{16}H_{28}O_3$)	C_3H_6
		254 (w, $C_{14}H_{22}O_4$)	CO
	141 ($C_8H_{13}O_2$)	254 (s, $C_{14}H_{22}O_4$)	$C_6H_9O_2$
		240 (w)[b]	
		228 (w)[b]	
		214 (s, $C_{11}H_{18}O_4$)	$C_3H_5O_2$
		159 (s, $C_8H_{15}O_3$)	H_2O
		156 (w, $C_9H_{16}O_2$)	CH_3
	128 ($C_7H_{12}O_2$)	296 (w, $C_{17}H_{28}O_4$)	$C_{10}H_{16}O_2$
		210 (s, $C_{13}H_{22}O_2$)	C_6H_{10}
		198 (m)[b]	
	99 ($C_6H_{11}O$)	237 (w)[b]	
		181 (w, $C_{11}H_{17}O_2$)	C_5H_6O
		169 (w, $C_9H_{13}O_3$)	$C_3H_2O_2$
		156 (w, $C_9H_{16}O_2$)	C_3H_5O
		141 (m, $C_8H_{13}O_2$)	C_2H_2O
		127 (s, $C_7H_{11}O_2$)	CO

[b] Not present in high resolution mass spectrum.

occur if the macrolide ring were opened with the cleavage of one bond followed by degradation of the resulting chain from both ends as illustrated in Figure 3.

Thus, the metastable data demonstrate an extensive interrelationship of the various mass spectral ions and the probable presence of branched degradation pathways. However, it is not possible to determine the extent of rearrangements from the data.

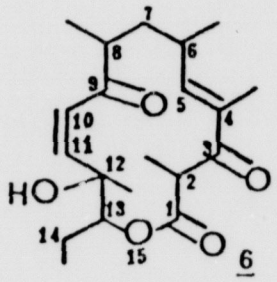
C. 14- and 16-Membered Macrolides

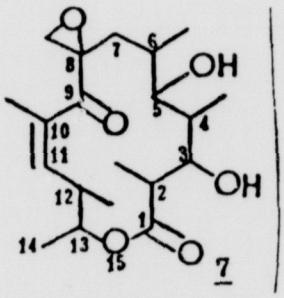
In order to test the generality of the conclusions drawn from the INTSUM data for the 12-membered macrolides, the high resolution mass spectra[15] of several macrolide aglycones with 14- and 16-membered lactones were studied. The unit resolution mass spectra of the 14-membered macrolides kromycin (6) and the aglycone of anhydrooleandomycin (7)[18] and the 16-membered macrolides platenolide I (8) and platenolide II (9)[19] are shown in Figure 4. As was the case for the 12-membered macrolides, the processes group themselves naturally according to the functional groups present in the molecules. The dominant processes in the 14-membered macrolides 6 and 7 and the supporting evidence are listed in Table IV.

Processes J and K are the only fragmentations which are evidenced by strong mass spectral peaks in both compounds. They involve cleavage of the 8-9 bond which is allylic to the 10-11 carbon-carbon double bond and α to the 9-keto group. Process K also involves a possible McLafferty rearrangement such as observed for process A. J and K correspond respectively to processes G and A for the 12-membered macrolides (1-5).

Processes L, M, N and O, which produce important mass spectral peaks in 6 and not in 7, all involve at least one cleavage allylic to the 4-5 double bond in 6. This demonstrates the strong effect of the presence of carbon-carbon

Table IV. Fragmentation Processes for which Evidence Exists in the Mass Spectra of Macrolides 6 and 7.





Process	Ion[a]	Representation[b]	Evidence in 14-Membered Macrolides ($\%\Sigma_{40}$)		Neutral Transfers
			6	7	
J	j	(9 8) (12 13)	1.1	3.8	-
K	k	(9 8) (13 15)	1.0	4.5	-H
L	l	(7 8) (3 2)	1.3	-	-
	ll		4.5	-	-H
M	m	(6 7) (1 15)	1.8	-	-
N	n	(6 7) (3 2)	8.2	-	-
O	o	(8 9) (3 2)	2.5	-	-
P	p	(8 9) (4 3)	.8	3.2	-H
			-	4.5	$-H-H_2O$
Q	q	(5 6) (1 15)	-	3.8	$+H-2H_2O$

[a] See Figure 4.

[b] See footnote c in Table 2.

double bonds on the mass spectral behavior of these compounds. Processes P and Q are processes which predominate in 7 but not in 6. Both involve cleavage of bonds in 7 which correspond to less labile vinylic bonds in 6. Process Q is accompanied by concomitant loss of two water molecules.

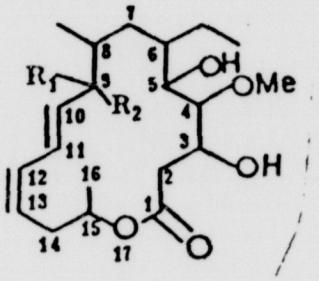
Platenolides I (8) and II (9) are so closely related that it was not possible to distinguish among the possible explanations for many of the strong peaks. In addition, there are many weak but few strong peaks in the high mass regions of the spectra. The processes for which the strongest evidence exists and which produce intense mass spectral peaks are listed in Table V. Processes R and S are the only two for which strong evidence exists in both spectra and each involves at least one cleavage of an allylic bond. The most intense peak in 8 and one of the most intense in 9 are probably due to process R which may involve a McLafferty rearrangement opening the lactone linkage. It corresponds to process A in the 12-membered macrolides and process K in the 14-membered macrolides.

The data are too limited to draw any additional conclusions as to the causes of the differences between the spectra of 8 and 9. The conclusions which can be drawn from these spectra, however, are in line with those already discussed for the 12- and 14-membered macrolides. The substituents play the dominant role in determining the bonds which will be cleaved. Allylic cleavage seems to be particularly important when carbon-carbon double bonds are present.

D. Use of Fragmentation Rules for Prediction and Ranking

Although it was of interest to see if RULEGEN and RULEMOD could form useful fragmentation rules from the INTSUM data for compounds 1-5, this approach

Table V. Fragmentation Processes for which Evidence Exists in the Mass Spectra of Macrolides 8 and 9.[a]



Process	Ion[b]	Representation[b]	Evidence in 16-Membered Macrolides ($\%\Sigma_{40}$)		Neutral Transfers
			8	9	
R	r	(9 8) (15 17)	2.6	2.5	-
	rr		7.6	4.7	-H
S	s	(3 4) (15 14)	1.9	1.5	$-H_2O$

[a] 8 R_1=R_2= =O; 9 R_1=OH, R_2=H.

[b] See Figure 4.

proved unsuccessful. RULEGEN and RULEMOD are designed to generate rules which show generality throughout a series of compounds.[12] We have already concluded that few processes are general to the series. Thus, the number of useful rules which could be generated was limited.

There are, however, rules depending on specific functionalities, some of which were discussed previously. There are two possible approaches which could be used to generate manually a more comprehensive set of rules to predict the mass spectral behavior of 1-9. The first is to use the correlations which are

described in the discussion of the mass spectra of 1-5. This would result, however, in rules finely tuned to the skeleton of the original five compounds. This is not desirable since the original correlation indicates that the location of various functional groups is more important in determining the fragmentation processes of these compounds than the basic skeleton. An approach which overcomes this difficulty defines rules which are combinations of common single bond cleavages such as McLafferty rearrangements, cleavages α to carbonyl and hydroxyl groups, etc. There is still some danger of tuning the results to the data, but it is reduced. In fact, this approach makes it simple to formulate general rules which can be extended to other classes of compounds.

A set of rules of the type discussed above was generated in the following manner. Single bond cleavages of various types (McLafferty rearrangements, cleavage α to carbonyl and hydroxyl groups, etc.) corresponding in part to our observations of the mass spectra of 1-9, were combined in all possible ways to generate candidate rules which would result in the fragmentation of a compound into two separate pieces. These rules were tested against the data for macrolides 1-5. Each rule was applied to each individual structure. Each time a rule predicted a mass spectral peak which was found in the actual spectrum of the compound, it received a score of 1; if it predicted a nonexistent peak, -1. Only rules with at least twice as much positive evidence as negative evidence (i.e. $|P/N|>2$ where P equals the sum of the positive and N equals the sum of the negative evidence, respectvely) for all five compounds were retained. This manner of scoring was used rather than a simple summation because it permitted the retention of good rules which were not necessarily applicable in every compound - i.e. a

type of rule with low generality but relatively high discriminatory power. The intensity of a peak correctly predicted by a rule was not taken into consideration. The rules were not allowed to have a range of hydrogen transfers, but were given the specific transfers indicated by the mechanism involved. For example, a rule involving a McLafferty rearrangement was considered to have a loss or a gain of one hydrogen depending on the fragment retaining the charge. Possible loss of water was also considered for each rule.

When the candidate rules were scored, 37 rules survived.[15] However, it was not possible to determine among these "best" rules which types of single bond cleavages were preferred since most of the rules involved the cleavage of two types of bonds.

The rules generated as described above were used to predict mass spectra for compounds 1-5 and the predicted spectra were scored against the actual spectra for each of the five compounds using the scoring function $S = \sum_i m_i I_i$ where m is the mass and I is the intensity for the mass spectral peak of a predicted and observed ion (i). In each case, the correct structure received the highest score.

The successful identification of the compounds using the rules which had been generated from their own spectra, of course, is only significant in that it demonstrated that the rules do have discriminatory power. It remained to be seen if the rules could successfully be used to identify a compound from among a large number of possible structures other than those from which they were obtained. Therefore, the computer program CONGEN[20] was used to generate closely related isomeric structures for macrolides 1-5. In order to limit the number of structures generated, CONGEN was constrained to keep the same carbon macrolactone

Table VI. Ranking of 12-Membered Macrolide Structures.

Structure	Rank/Original 5 Structures	Rank/Total Number of CONGEN Structures
1	1/5	2/60
2	1/5	1/105
3	1/5	2/105
4	1/5	5/105
5	1/5	5/105

skeleton for all structures; only the three remaining oxygens were allowed to change position. Furthermore, these three oxygens were only allowed to exist as alcohols, ketones and aldehydes. When the resulting structures were ranked using the experimental mass spectra as described above, the correct structures were ranked as shown in Table VI.

The structures ranked higher than the correct structures were closely related.[15] For example, the structure ranked higher than 1 is shown in Figure 5. The ranking is most successful for the macrolides containing an olefinic double bond (1-3). This is in agreement with the INTSUM data which indicated that carbon-carbon double bonds strongly influence the mass spectral behavior of these compounds.

Since the rules used to score the 12-membered macrolide structures were not dependent on the skeleton, they are in principle applicable to other classes of macrolides. CONGEN was therefore used to generate isomeric structures for macrolides 6-9. These structures were generated under similar constraints as those used for the 12-membered macrolides (1-5). Only the oxygen functionalities not involved in the lactone linkage were allowed to change position on the

macrolide skeletons. In addition, it was necessary to further restrict the number of isomers generated by CONGEN for kromycin (6) by forcing the exocyclic epoxide to remain on carbon 8, and for the 16-membered macrolides (8 and 9) by forbidding the generation of structures containing aldehyde or primary alcohol groups. The predicted mass spectra were used to rank these structures in the same manner described for the 12-membered macrolide structures. The results of these rankings are listed in Table VII.

Table VII. Ranking of 14 and 16-Membered Macrolide Structures.

Structure	Rank/Total Number of CONGEN Structures
6	5/70
7	8/140
8	3/196
9	11/84

Only platenolide II (9) is not ranked among the top 10% of the candidate structures. Again, the structures ranked highest are closely related.[15] For example, the structures ranked higher than 6 are illustrated in Figure 6.

The results of the structure rankings in Tables VI and VII show that it is possible to use empirical evidence to generate useful mass spectral fragmentation rules. The correct structure may not be ranked first in all cases, but as the data in Tables VI and VII indicate, they can focus on a set of most plausible structures which contains the correct one. The success of this approach can probably be improved in at least three ways. Greater flexibility in the possible hydrogen transfer and neutral loss constraints in the rule definitions could

account for mass spectral peaks such as those resulting from the loss of water from ions predicted by a rule. This occurrence was common in the mass spectra of compounds 1-9. The scoring function used for the above rankings is a very simple one which does not consider negative evidence (predicted peaks which are not observed) and can probably be improved. Another source of improvement could result from the inclusion of intensity information in a rule. Often, the application of a rule to an incorrect structure generates a predicted ion with the same composition as an actual peak in the spectrum while the spectrum of the correct structure does not display the ion. Such an ion causes an incorrect structure for which the rule applies to receive a higher ranking relative to the correct structure. These ions are often of different intensities than those which lead to the formulation of the rule in the first place, and intensity information can be used to filter them out.

III. CONCLUSION

The Meta-DENDRAL subprogram INTSUM was used to assist in determining the important mass spectral fragmentations of several classes of macrolide antibiotic aglycones. The nature of these fragmentations proved to be highly dependent on the types of functional groups present in the molecules, especially carbon-carbon double bonds. Metastable defocusing data for two of the compounds studied indicated that many of the fragmentation processes were multi-step rather than concerted.

Because the behavior of these compounds was so highly dependent on the substituents present, there were few fragmentation processes for which evidence could be found throughout a series of compounds. It was possible, however, to manually generate a series of useful rules based on combinations of simple, commonly accepted one-bond cleavages. Spectral predictions based on these rules

proved useful in ranking structural candidates based on the mass spectrum alone. This approach may eventually prove useful in the automation of structure elucidations of compounds from their mass spectra.

IV. EXPERIMENTAL

The high resolution mass spectra were measured by Ms. A. Wegmann using a Varian-MAT 711 mass spectrometer at 70 ev with the direct insertion probe. Low resolution mass spectra and metastable defocusing measurements were recorded by Mr. R. G. Ross on an AEI MS-9 mass spectrometer. Samples of compounds 1-5 were available from previous investigations in this laboratory.[13,14] Samples of 6 and 7 were generously furnished by Professor Haruo Ogura[18] and samples of 8 and 9 by Dr. Makoto Suzuki.[19] The computer programs are written in Interlisp and run on the DEC 10 SUMMEX AIM computer facility at Stanford University.

REFERENCES

1. Applications of Artificial Intelligence for Chemical Inference part XXVI. For part XXV see T. M. Mitchell and G. M. Schwenzer, Organic Magnetic Resonance, submitted for publication.

2. We wish to thank the National Institutes of Health, Grant No. 2 R24 RR00612.

3. R. Morin and M. Corman, "Encyclopedia of Chemical Technology", A. Standen Ed., John Wiley & Sons, Inc., New York, 1967, Vol. 12, p. 632.

4. K. L. Rinehart, Jr. and G. E. Van Lear, "Biochemical Applications of Mass Spectrometer", G. R. Waller Ed., John Wiley & Sons, Inc., New York, 1972, p. 449.

5. K. L. Rinehart, Jr. and G. E. Van Lear, Abstr. papers, 15th Ann. Conf. Mass Spectrom. Allied Topics, ASTM E-14, Denver, Colo., May 16-19, 1967, Paper 120.

6. a. R. L. Folz, L. A. Mitscher and M. I. Levenberg, Abstr. papers, 17th Ann. Conf. Mass Spectrom. Allied Topics, ASTM E-14, Dallas, Tex., May 18-23, 1969, Paper 119; b. I. Maezawa, A. Kinumki and M. Suzuki, J. Antibiotics, 27, 84 (1974).

7. a. H. Achenbach and W. Karl, Chem. Ber., 108, 772 (1975); b. A. Kinumaki, I. Takamori, Y. Sugawara and M. Suzuki, J. Antibiotics, 27, 107 (1974); c. A. Kinumaki, I. Takamori, Y. Sugawara, Y. Seki, M. Suzuki and T. Okuda, J. Antibiotics, 27, 117 (1970).

8. L. A. Mitscher, R. L. Folz and M. I. Levenberg, Org. Mass Spectrom., 5, 1229 (1971).

9. L. A. Mitscher and H. D. H. Showalter, J.C.S. Chem. Comm., 796, (1972).

10. D. H. Smith, B. G. Buchanan, W. C. White, E. A. Feigenbaum, J. Lederberg and C. Djerassi, Tetrahedron, 29, 3117 (1973).

11. a. S. Hammerum and C. Djerassi, Tetrahedron, 31, 2391 (1975); b. L. L. Dunham, C. A. Henrick, D. H. Smith and C. Djerassi, Org. Mass Spectrom., 11, 1120 (1976).

12. B. G. Buchanan, D. H. Smith, W. C. White, R. J. Gritter, E. A. Feigenbaum, J. Lederberg and C. Djerassi, J. Amer. Chem. Soc., 98, 6168 (1976).

13. C. Djerassi and J. A. Zderic, J. Amer. Chem. Soc., 78, 6390 (1956).

14. C. Djerassi and O. Halpern, Tetrahedron, 3, 255 (1958).

15. K. A. Morrill, Ph.D. Dissertation, Stanford University, May 1977.

16. Loss of methyl was not considered in the original INTSUM analysis, but it was considered in a later analysis and provided additional plausible explanation in only a few instances.

17. a. M. Barber and R. M. Elliott, ASTM Committee E-14, 12th Annual Conference on Mass Spectrometry and Allied Topics, Montreal, June 1964; b. K. R. Jennings, J. Chem. Phys., 43, 4176 (1965).

18. a. H. Ogura, K. Furuhata, H. Kuwano and N. Harada, J. Amer. Chem. Soc., 97, 1930 (1975); b. Aglycone of Anhydro-oleandomycin, Haruo Ogura, private communication. See also ref. 1.

19. T. Furumai, Y. Seki, K. Takeda, A. Kinumaki and M. Suzuki, J. Antibiotics, 26, 708 (1973).

20. R. E. Carhart, D. H. Smith, H. Brown and C. Djerassi, J. Amer. Chem. Soc., 97, 5755 (1975).

LEGENDS TO FIGURES

Figure 1. Unit resolution mass spectra of 12-membered macrolides 1-5.

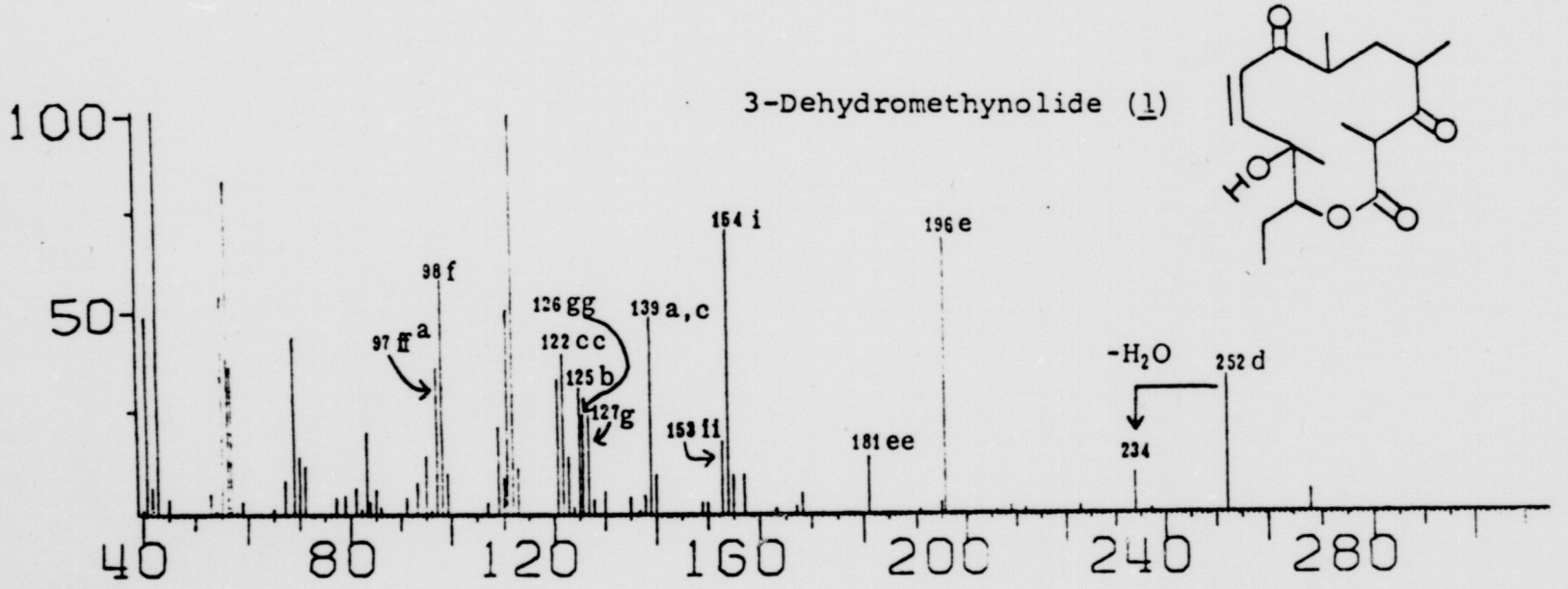
Figure 2. Schematic representation of the mass spectral process A. The arrow indicates the charge carrying fragment.

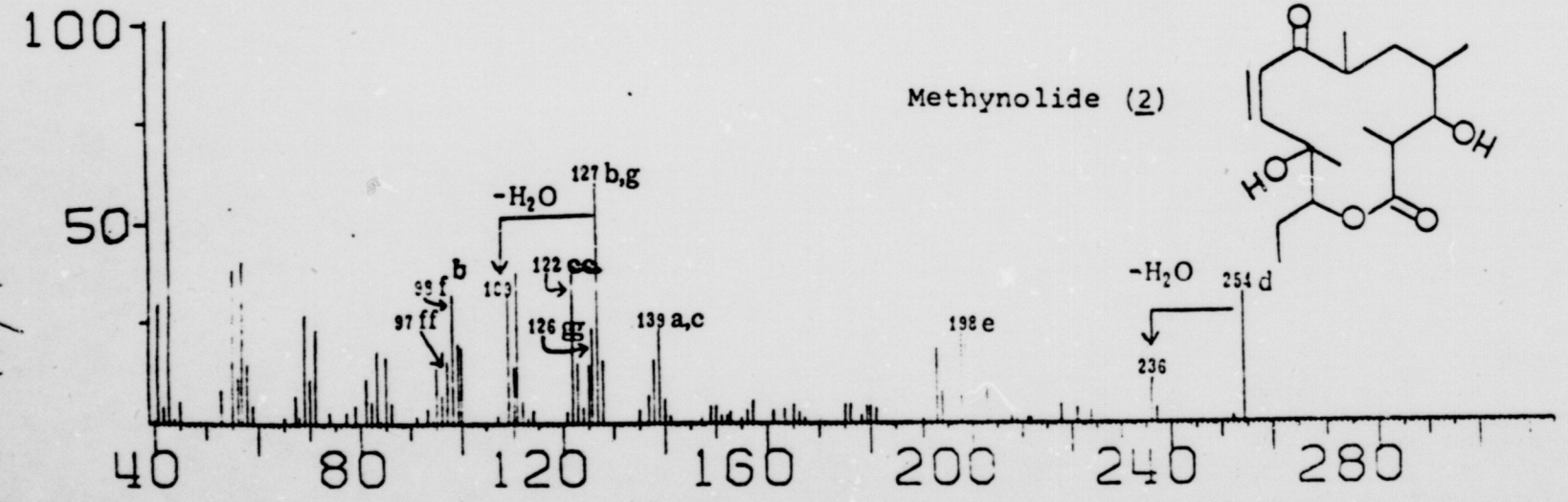
Figure 3. Possible mechanisms for the observed metastable pathways leading to the m/e 226 ion in the mass spectrum of dihydromethynolide (4).

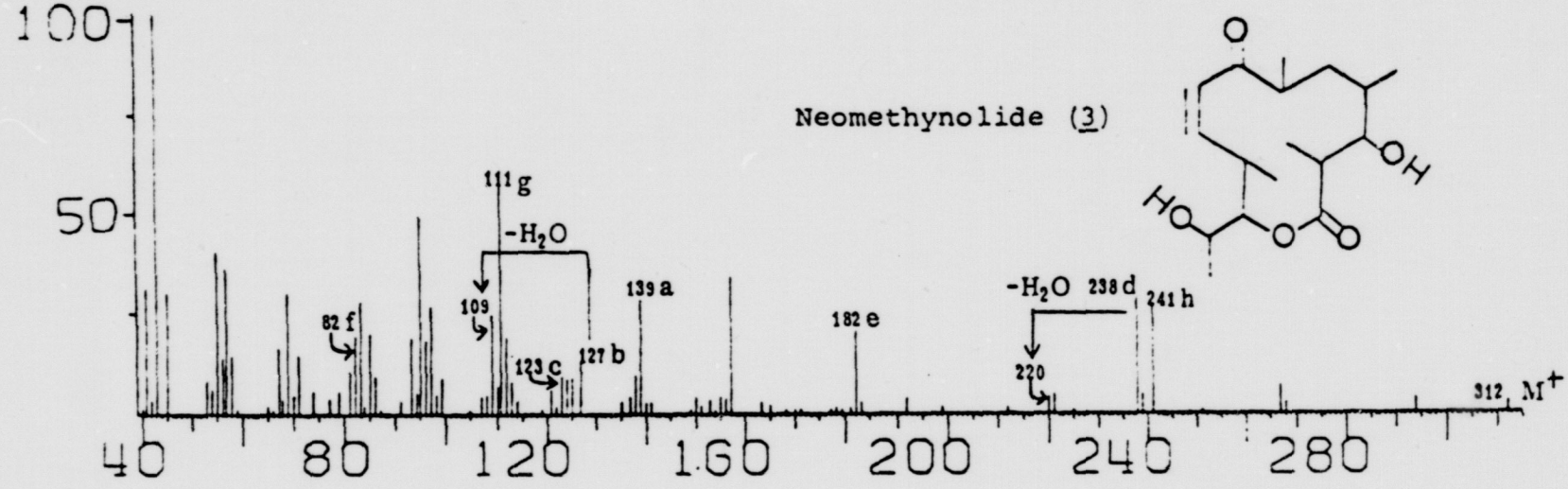
Figure 4. Unit resolution mass spectra of 14- and 16-membered macrolides 6-7.

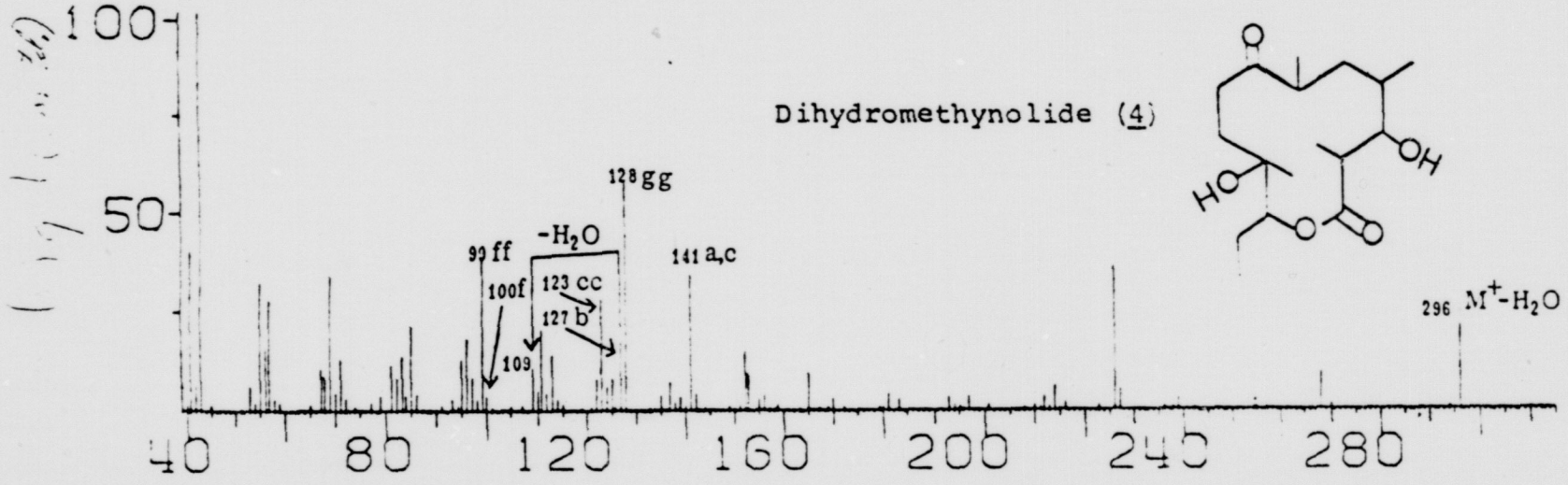
Figure 5. Structure ranked higher than 3-dehydromethynolide (1).

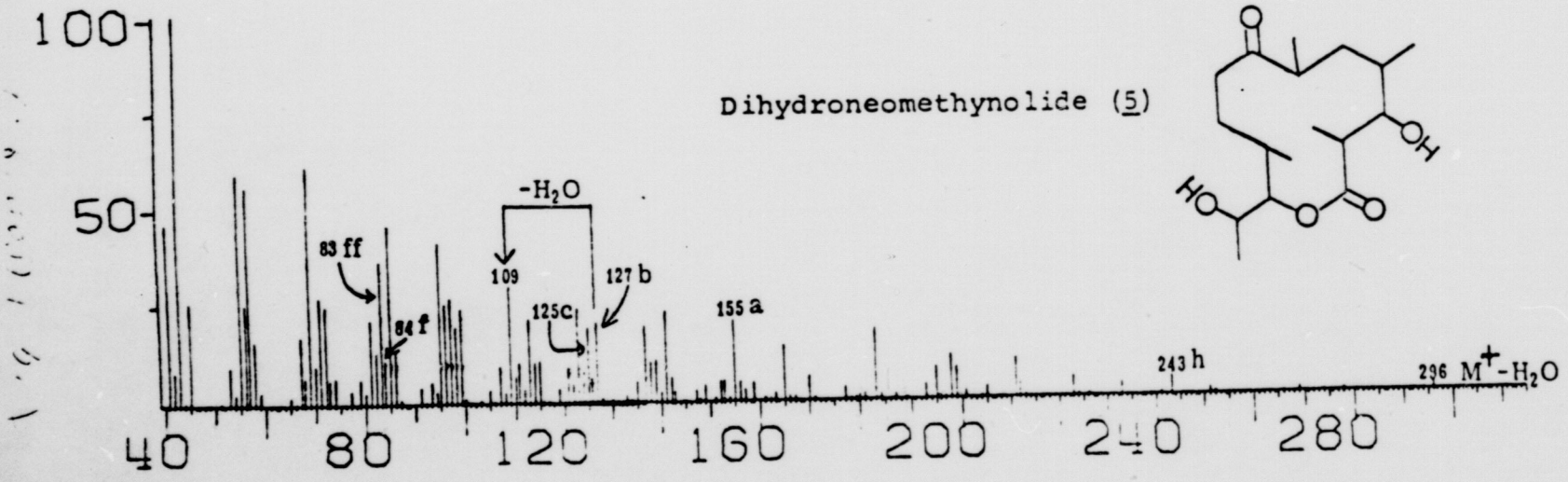
Figure 6. Structures ranked higher than kromycin (6).


3-Dehydromethynolide (1)
100
50
40
80
120
160
200
240
280
98 f
97 ff
a
126 gg
122 cc
125 b
127 g
139 a,c
154 i
153 ii
181 ee
196 e
-H2O
234
252 d

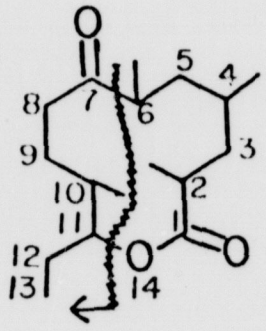

Methynolide (2)
100
50
40
80
120
160
200
240
280
-H2O
127 b,g
99 f
b
97 ff
122 cc
126 gg
139 a,c
198 e
-H2O
254 d
236

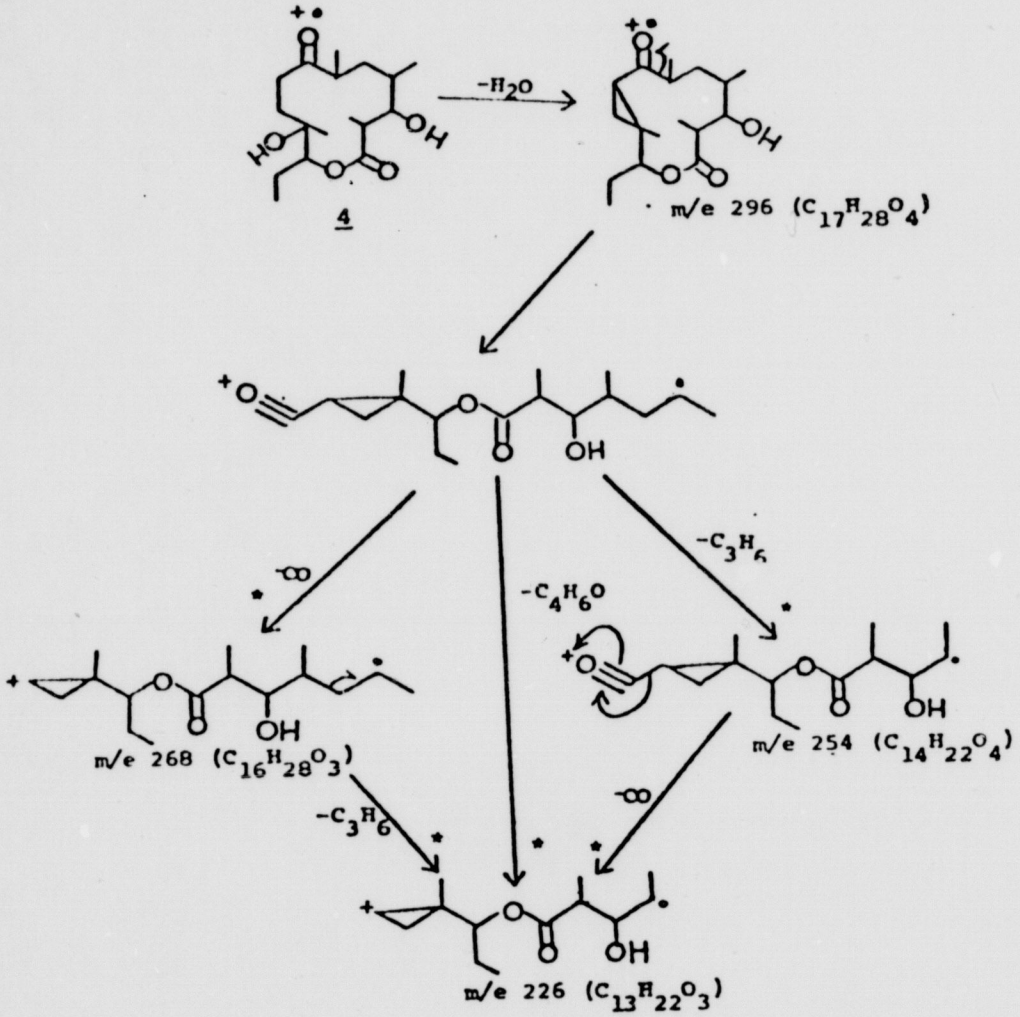

100
50
Neomethynolide (3)
111 g
-H2O
109
82 f
123 c
127 b
139 a
182 e
-H2O 238 d
241 h
220
312 M+
40
80
120
160
200
240
280

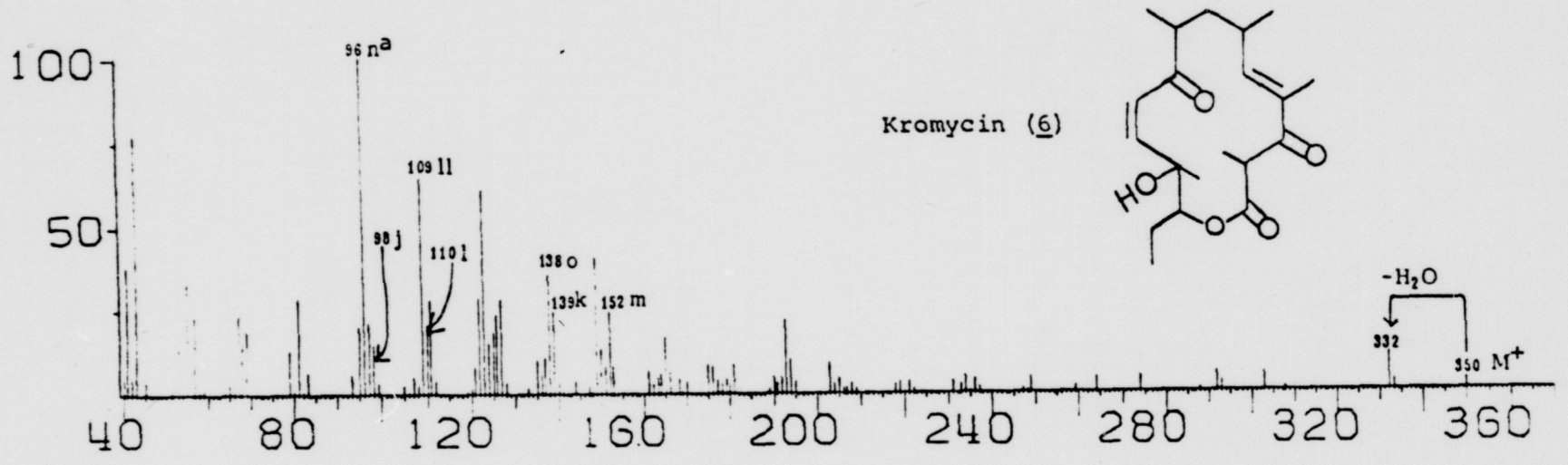

100
50
Dihydromethynolide (4)
128 gg
-H2O
99 ff
100 f
123 cc
127 b
109
141 a,c
296 M+-H2O
40
80
120
160
200
240
280



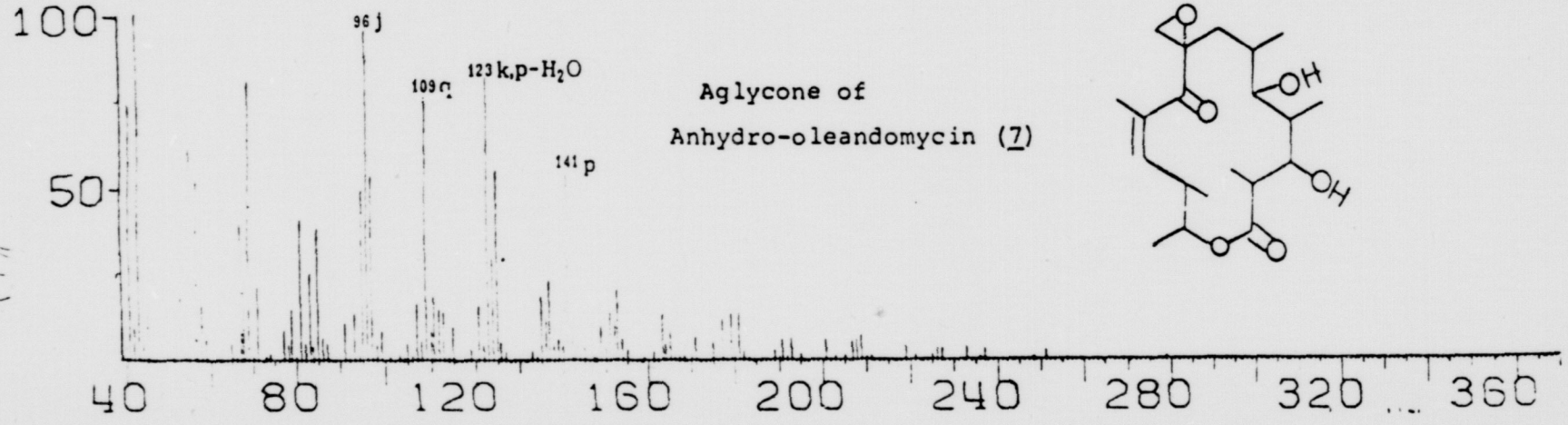
Dihydroneomethynolide (5)


O
1
2
3
4
5
6
7
8
9
10
11
12
13
O
14
O

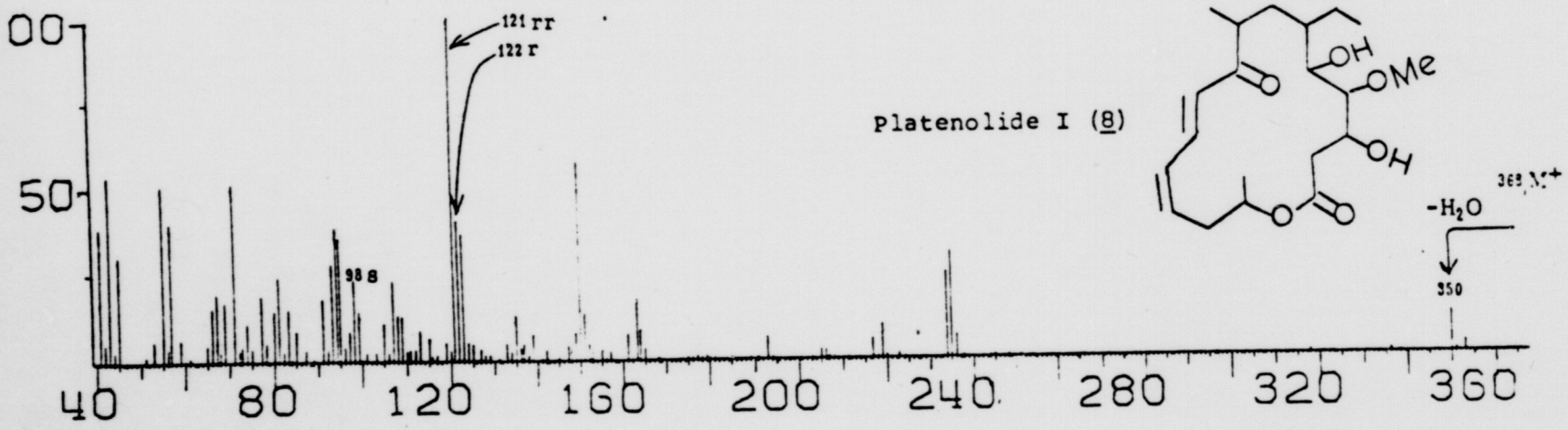

-H2O
4
m/e 296 (C17H28O4)
-C3H6
-C4H6O
-C3H6
m/e 268 (C16H28O3)
m/e 254 (C14H22O4)
m/e 226 (C13H22O3)



Kromycin (6)



Aglycone of
Anhydro-oleandomycin (7)


100
50
121 rr
122 r
98 s
Platenolide I (8)
OH
OMe
OH
369 M+
-H2O
350
40
80
120
160
200
240
280
320
360

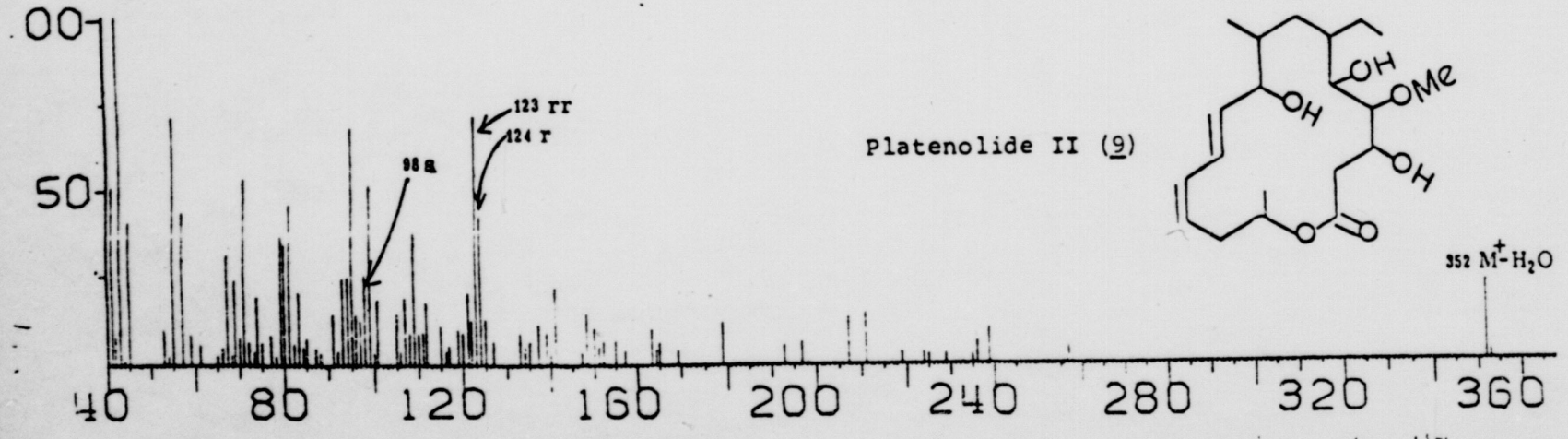

100
50
123 rr
124 r
98 s
Platenolide II (9)
OH
OMe
OH
352 M+-H2O
40
80
120
160
200
240
280
320
360

fig 5

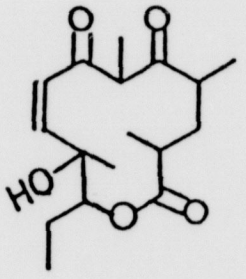
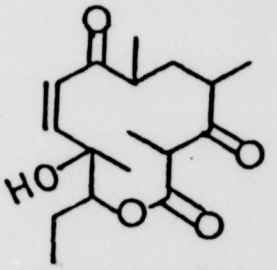
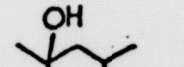
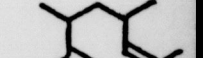
Structure		 3-Dehydromethynolide (1)
Score	6013	4721

fig 6

Structure					 Kromycin (6)
Score	4386	4178	4027	3992	3806