mixture was cooled and poured into 150 mL of saturated $NaHCO_3$. The organic layer was separated, washed twice with 100-mL portions of saturated $NaHCO_3$ and once with brine, and dried over anhydrous Na_2SO_4, and the solvent was removed under reduced pressure to yield 3.03 g of crude product. Chromatography on silica gel (1:1 ether–hexane) afforded 644 mg (20%) of diketal **55**, 1.23 g (44%) of **30**, and 791 mg (33%) of recovered **29**. The required monoketal was crystallized from ether: mp 139–141 °C; IR ($CHCl_3$) 1660 cm^{-1}; ^1H NMR ($CDCl_3$) δ 0.93 (d, J = 6 Hz, 3), 1.23 (s, 3), 1.3–2.9 (m, 13), 3.77 (s, 4), 5.37 (t, 1).

Anal. Calcd for $C_{18}H_{24}O_3$: C, 74.97; H, 8.39. Found: C, 74.72; H, 8.27.

Diketal **55**, a clear oil, has no carbonyl absorption (IR) and has the following: ^1H NMR (CCl_4) δ 0.95 (d, J = 6 Hz, 3), 1.22 (br s, 3), 1.0–3.0 (m, 13), 3.82 (br s, 8), 5.8 (br s, 1). No elemental analysis was obtained.

Photoaddition of Ethylene and 30. The cycloaddition was conducted as previously described (**25** → **26**) by using 1.39 g (4.83 mmol) of **30** in 1.1 L of methylene chloride. Reaction was complete after 1 h of irradiation. Chromatography of the crude product (1.54 g) on Florisil with 1:1 ether–hexane gave 1.26 g (81%) of cycloadduct **31**: mp 93–93.5 °C; IR (CCl_4) 1690 cm^{-1}; ^1H NMR ($CDCl_3$) δ 0.75 (s, 3), 0.93 (d, J = 6 Hz, 3), 1.1–2.8 (m, 17), 3.73 (s, 4), 5.55 (m 1).

Anal. Calcd for $C_{20}H_{28}O_3$: C, 75.91; H, 8.92. Found: C, 76.02; H, 8.88.

Tetracyclic Hydroxy Enone 33. To a slurry of 2.47 g (65 mmol) of lithium aluminum hydride in 100 mL of THF was added 2.94 g (77 mmol) of **31** in 50 mL of THF over a period of 30 min. The resulting mixture was heated at reflux for 8 h and cooled to 10 °C, and the excess LAH was decomposed by the successive addition of 2.5 mL of water, 2.5 mL of 15% NaOH, and 7.5 mL of water. The resulting mixture was filtered and washed with ether, and the organic filtrate was dried (Na_2SO_4). Removal of solvent gave 2.62 g (88%) of crude alcohols **32**: IR (neat) 3440 cm^{-1}, no carbonyl absorption.

A solution of crude **32** in 150 mL of THF was combined with 25 mL of 2 N HCl and stirred at 25 °C for 5 h, washed with brine, saturated $NaHCO_3$, and brine, and dried (Na_2SO_4). Removal of solvent gave 2.26 g of a mixture of axial and equatorial alcohols **33a** and **33e** (1:2, respectively). Chromatography on silica gel (1:4 ethyl acetate–hexane) gave 958 mg (42%) of **33e** and 459 mg (20%) of **33a**. Recrystallization of **33e** from ether gave an analytical sample: mp 123–124 °C; IR ($CHCl_3$) 3440, 1655 cm^{-1}; ^1H NMR ($CDCl_3$) δ 1.18 (s, 3), 1.25–2.80 (m with s at 1.75, 22), 3.58 (unresolved t, 1).

Anal. Calcd for $C_{18}H_{26}O_2$: C, 78.79; H, 9.55. Found: C, 78.93; H, 9.72.

Keto alcohol 33a, an oil, had the following: IR (neat) 3450, 1655 cm^{-1}; ^1H NMR ($CDCl_3$) δ 1.32 (s, 3), 1.40–2.90 (m with s at 1.75, 22), 3.57 (unresolved d, 1).

Reductive methylation of 33e as described above (**7** → **23**) gave 312 mg of keto alcohol **34** (from 360 mg of **33e**). Chromatography on silica gel (1:4 ethyl acetate–hexane) gave 174 mg of pure **34e** as an oil: IR (neat) 3450, 1710 cm^{-1}; ^1H NMR ($CDCl_3$) δ 0.87 (s, 3), 1.10 (s, 3), 1.17 (s, 3), 1.20–2.65 (m, 20), 3.20–3.70 (m, 1). No elemental analysis was obtained.

Oxidation of 34e (150 mg, 0.52 mmol) as described above (**26** → **27**) gave 131 mg of solid [90% **35** and 10% **27** (?) by ^1H and ^{13}C NMR] which was recrystallized from ether to give pure dione **35**: mp 149–150 °C; IR ($CHCl_3$) 1695 (br) cm^{-1}; ^1H NMR ($CDCl_3$) δ 0.90 (s, 3), 1.08 (s, 3), 1.18 (s, 3), 1.20–2.70 (m, 19).

Anal. Calcd for $C_{19}H_{28}O_2$: C, 79.12; H, 9.79. Found: C, 79.16; H, 9.53.

Acknowledgment. It is a pleasure to thank Professor A. L. Wilds for sharing his experience related to Scheme VI.

Registry No. 1, 38966-21-1; 7, 40266-30-6; 8, 38312-61-7; 8 α-phenylselenenyl derivative, 77827-76-0; 8 trimethylsilyl enol ether, 77827-77-1; 9, 77827-78-2; 10a, 77827-79-3; 10a α-phenylselenenyl derivative, 77827-80-6; 10b, 77827-81-7; 10c, 77827-82-8; 10d, 77827-83-9; 12, 77827-84-0; 13, 18631-96-4; 14, 77827-85-1; 14 α-phenylselenenyl derivative, 77827-86-2; 14 trimethylsilyl enol ether, 77827-87-3; 15, 77827-88-4; 16, 22118-00-9; 17, 77827-89-5; 18, 77827-90-8; 19, 77827-91-9; 20, 77827-92-0; 22a, 77827-93-1; 22b, 77846-76-5; 23, 77827-94-2; 24a, 77827-95-3; 24e, 77827-96-4; 25a, 77827-97-5; 25e, 77827-98-6; 26a, 77827-99-7; 26e, 77880-87-6; 27, 77828-00-3; 28, 77828-01-4; 29, 77828-02-5; 30, 77828-03-6; 31, 77828-04-7; 32a, 77828-05-8; 32e, 77880-88-7; 33a, 77846-77-6; 33e, 77881-62-0; 34e, 77846-78-7; 35, 77881-63-1; 39, 77828-06-9; 40, 38312-62-8; 41, 77828-07-0; 42, 77828-08-1; 43, 77828-09-2; 44 (epimer 1), 77828-10-5; 44 (epimer 2), 77880-89-8; 45 (epimer 1), 77828-11-6; 45 (epimer 2), 77880-90-1; 46, 77828-12-7; 47, 77828-13-8; 48, 77828-14-9; 49, 77828-15-0; 50, 59633-34-0; 51, 77828-16-1; 52, 77828-17-2; 53, 38301-79-0; 55, 77828-18-3; 3,3-dibromotricyclo[4.4.2.0]dodecan-2-one, 77828-19-4; 3-bromotricyclo[4.4.2.0]dodec-3-en-2-one, 77828-20-7; 4-methyl-3-bromotricyclo[4.4.2.0]dodecan-2-one, 77828-21-8; 11,12-dichlorotricyclo[4.4.2.0]dodecan-2-one, 77880-91-2; 11,12-dichlorotricyclo[4.4.2.0]dodecan-2-ol, 77880-92-3; cyclohexane-1,3-dione, 504-02-9; ethyl vinyl ketone, 1629-58-9; 2-(3-oxopentyl)cyclohexane-1,3-dione, 77828-22-9.

Supplementary Material Available: Tables of coordinates and anisotropic temperature factors for nonhydrogen atoms, hydrogen coordinates, distances, and angles (4 pages). Ordering information is given on any current masthead page.

Computer-Assisted Carbon-13 Nuclear Magnetic Resonance Spectrum Analysis and Structure Prediction for the C_{19}-Diterpenoid Alkaloids

Janet Finer-Moore, Naresh V. Mody, and S. William Pelletier*

Institute for Natural Products Research and the Department of Chemistry, The University of Georgia, Athens, Georgia 30602

Neil A. B. Gray, Christopher W. Crandell, and Dennis H. Smith*

Department of Chemistry, Stanford University, Stanford, California 94305

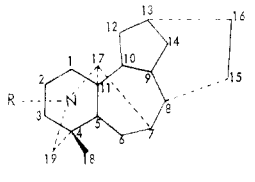
Received January 30, 1981

Computer programs for the analysis of ^{13}C NMR spectra and for structure prediction from the structural information thus inferred have been used successfully for structure elucidation of C_{19}-diterpenoid alkaloids. A data base created from the structures and the ^{13}C NMR spectra of 93 alkaloids and their derivatives was used by the programs to interpret the spectra of several unknown structures. Three examples are described in detail. The examples demonstrate that the programs can quickly limit possible structures for even complicated C_{19}-diterpenoid alkaloids to two or three when an aconitine-type skeleton is assumed for the unknown. The efficiency of the programs is based in part on their ability to utilize structural constraints during both spectrum analysis and structure generation.

The C_{19}-diterpenoid alkaloids have been known for their extreme toxicity for hundreds of years and studied by chemists for over a century. Elucidating the structures of these compounds was a formidable challenge until the

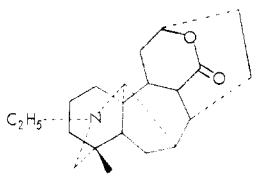
0022-3263/81/1946-3399$01.25/0

advent of sophisticated X-ray crystallographic methods.[1] Since then, the number of known C_{19}-diterpenoid alkaloid structures has grown rapidly. A recent catalogue of C_{19}-diterpenoid alkaloids includes structures of 73 compounds.[2] The majority of these alkaloids have a highly substituted aconitine skeleton (1). Their toxicity is related



R = CH_3 or C_2H_5

1 Aconitine Skeleton



2 Heteratisine Skeleton

to the degree of substitution, particularly by ester groups. Inspection of the aconitine-type alkaloid structures reveals regularities in their complex substituents patterns which may be described by a few empirical rules: (1) C(1), C(14), and C(16) are usually oxygenated, C(16) with a β-methoxyl, C(1) with a hydroxyl or methoxyl group, and C(14) frequently with a complex ester. (2) There is always an oxygen substituent or a double bond at C(8). The oxygen substituent may be a methoxyl, hydroxyl, acetyl, or methylenedioxyl group. (3) When C(7) is substituted, C(6) nearly always has a β-substituent and C(13) is never oxygenated. On the other hand, when no substituent is present at C(7), both C(13) and C(6) may be oxygenated. (4) No substitution occurs at C(2), C(5), C(12), C(17), or C(19).

These rules are not based on biogenetic theory, and occasionally exceptions are found. The recently isolated alkaloid gadesine, for example, violates the fourth rule, being oxygenated at C(19).[3] Nonetheless, the rules are useful for guiding the search for a solution to a structural problem by limiting the number of likely structures to relatively few. In practice, exceptional solutions are considered only when likely possibilities are inconsistent with observed data.

A few C_{19}-diterpenoid alkaloids with the heteratisine skeleton (2) have been isolated.[2] Though possibly biogenetically related to aconitine-type alkaloids, these relatively simple, nontoxic compounds are easily distinguished from those alkaloids by their lactone ring. The substitution rules listed above do not apply to heteratisine-type structures. In particular, these structures lack a C(16) substituent. Our attempts at computer-assisted structure elucidation have so far been restricted to alkaloids presumed to have the aconitine skeleton, and thus we have excluded heteratisine-type alkaloids from the ^{13}C NMR data base described later.

Jones and Benn were the first to recognize the value of ^{13}C NMR spectroscopy for the determination of the sites and degree of oxidation in C_{19}-diterpenoid alkaloids.[4] They measured ^{13}C NMR spectra for seven closely related C_{19}-diterpenoid alkaloids and were impressed by the similarities in the patterns of shifts each exhibited. They also noted that the resonance shifts for quaternary carbons showed a sensitive and consistent response to substitution of neighboring carbons. With the help of additivity relationships and SFORD techniques and making extensive use of correlation tables, they were able to make internally consistent assignments of the ^{13}C NMR resonances for the seven alkaloids. They used their assignments to interpret the ^{13}C NMR spectra of two new alkaloids, and the structures for the new alkaloids were deduced largely from the results of this analysis. Though the proposed structures were later revised,[5,6] their major features were correct, and the study demonstrated the potential of ^{13}C NMR spectroscopy for structure elucidation of natural products.

Following this pioneering work, Pelletier and co-workers began a systematic study of the ^{13}C NMR spectra of all available C_{19}-diterpenoid alkaloids and their derivatives.[7] As more ^{13}C NMR data accumulated, ^{13}C NMR spectroscopy emerged as one of the most useful tools available for the determination of these alkaloid structures. Now, even complicated structures are solved without the aid of X-ray crystallography or lengthy chemical methods.[8] However, the expansion of the ^{13}C NMR data base has made it difficult for structural chemists, especially novices in the field of diterpenoid alkaloids, to consider the data base in its entirety when determining a new structure or assigning the spectrum of a known compound. This has resulted in a lack of consistency in the spectral assignments of different research groups.[9] These facts suggested to us the potential utility of a computer-based method of structure elucidation which could (1) process a larger quantity of data than the mind could objectively handle and (2) use the problem-solving approach which has led so efficiently to structure solutions in the past. In particular, the method should make use of the valuable information about the regularities of the C_{19}-diterpenoid alkaloid substitution patterns.

A set of three computer programs which seemed to meet these needs has recently been developed.[10,11] The programs utilize a data base of ^{13}C NMR resonances together with the substructures which characterize the constitutional and configurational stereochemical environment of the resonating carbons. A ^{13}C NMR spectrum is analyzed in the first program by matching the observed resonances to resonances in the data base and retrieving the associated substructures. The information contained in the total set of retrieved substructures is refined by using an iterative interpretation procedure which eventually deduces detailed portions, or substructures, of the unknown structure.[10] Piecing together these substructures into complete structures is accomplished by using the structure-generating program GENOA.[12] These "candidate" structures represent the complete set of structural possibilities for the unknown. A third program evaluates the candidate structures by predicting their spectra and comparing each predicted spectrum with the observed spectrum of the unknown and rank ordering the candidates accordingly.[10] Besides performing the final step in the structure elucidation procedure, the third program can be run independently of the

(1) M. Przybylska and L. Marion, *Can. J. Chem.*, **34**, 185 (1956).
(2) S. W. Pelletier and N. V. Moody, in "The Alkaloids", Vol. XVIII, Manske and Rodrigo, Eds., Academic Press, New York, 1979, Chapter 1, pp 1–103.
(3) A. G. Gonzalez, G. de la Fuente, R. Diaz, J. Fayos, and M. Martinez-Ripoll, *Tetrahedron Lett.*, 79 (1979).
(4) A. J. Jones and M. H. Benn, *Can. J. Chem.*, **51**, 486 (1973).
(5) S. W. Pelletier, N. V. Mody, A. J. Jones, and M. H. Benn, *Tetrahedron Lett.* 3025 (1976).
(6) P. W. Codding, K. A. Kerr, M. H. Benn, A. J. Jones, S. W. Pelletier, and N. V. Mody, *Tetrahedron Lett.* 127 (1980).
(7) S. W. Pelletier, N. V. Mody, and L. C. Schramm, unpublished results.
(8) S. W. Pelletier, N. V. Mody, A. P. Venkov, and N. M. Mollov, *Tetrahedron Lett.* 5054 (1978).
(9) (a) S. Sasaki, H. Takayama, and T. Okamoto, *Yakugaku Zasshi*, **98**, 1376 (1978); (b) Y. Tsung-Ren, H. Xiao-Jiang, and C. Jun, *Yun-nan Chih Wu Yen Chiu*, **1**, 41 (1979).
(10) N. A. B. Gray, C. W. Crandell, J. G. Nourse, D. H. Smith, M. L. Dageforde and C. Djerassi, *J. Org. Chem.*, **46**, 703 (1981).
(11) N. A. B. Gray, J. G. Nourse, C. W. Crandell, D. H. Smith, and C. Djerassi, *Org. Magn. Reson.*, **15**, 375 (1981).
(12) R. E. Carhart, D. H. Smith, N. A. B. Gray, J. G. Nourse, and C. Djerassi, *J. Org. Chem.*, **46**, 1708 (1981).

others to check spectrum assignments for a compound with a known structure.[11]

These programs were expected to be effective for the elucidation of C_{19}-diterpenoid alkaloid structures for several reasons. First, they are designed to exploit the information which is already known about a structure in much the same way that we would in manually solving a structural problem. Such information, which for our problems would include the aconitine-type skeleton and its likely substitution sites, can be used both in the iterative interpretation procedure of the spectrum analysis program to derive more elaborate pieces of an unknown structure and in the form of additional constraints on structure generation in GENOA. A second advantage of the programs is that their empirical approach to spectrum analysis and prediction obviates complicated additivity relationships. Simple additivity relationships derived from data on smaller systems have been found inadequate for predicting the magnitudes of the chemical shifts of substituted carbons in the C_{19}-diterpenoid alkaloids, presumably because of steric interactions in the compact diterpenoid alkaloid ring system.[4] Finally, the data base can incorporate all available ^{13}C NMR data for the C_{19}-diterpenoid alkaloids and their derivatives in the form of resonance-substructure pairs, and each pair is given equal weight during spectrum analysis and prediction. Thus the biases imposed on structure elucidation by neglect of certain ^{13}C NMR data, which may occur when a ^{13}C NMR spectrum is manually interpreted, are avoided.

We have been investigating the application of these programs to diterpenoid alkaloid structural problems and here present examples of how they were used successfully for testing structure assignments against ^{13}C NMR data and for determining the structures of new compounds. An example illustrating the limitations of the programs is also described.

Experimental Section

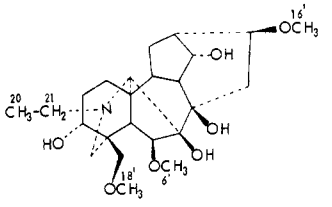
All computer programs were run at Stanford University on a DEC PDP-10 computer, accessed remotely by a computer terminal from Athens, GA, via the TYMNET computer network. ^{13}C NMR data, consisting of noise-decoupled spectra and single-frequency off-resonance decoupled (SFORD) spectra, were measured on JEOL FX-60 and PFT-100 spectrometers.

The structures and ^{13}C NMR spectra of 93 diterpenoid alkaloids and their derivatives were used to create a data base of four-shell substructure codes[11] and associated resonances. Each four-shell substructure code represented the environment of a particular resonating carbon out to at least a four-bond radius. The method of coding substructures from the input structures, which was carried out by computer, has been described previously.[11]

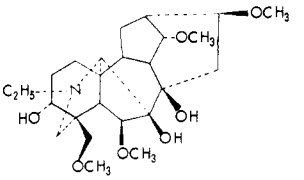
The ^{13}C NMR spectra of the alkaloids used to create the data base had generally been assigned by comparing the spectra of closely related alkaloids and noting the effects of specific structural differences. The criterion for the correctness of a spectrum assignment had been its consistency with the spectrum assignments for similar diterpenoids alkaloids. However, a computer-generated summary of the data base revealed some unusually broad ranges of chemical shifts associated with carbons in similar environments, indicating that spectrum assignments were not completely internally consistent. A number of reassignments were made in order to narrow these ranges and impose internal consistency. In the final data base the ranges of chemical shifts for particular four-shell environments averaged 0.7 ppm (maximum observed range of 3.9 ppm for a methylene group bonded to nitrogen, potentially influenced more strongly by solvent effects). Ranges for three-shell environments averaged 1.1 ppm, with the maximum observed range of 7.5 ppm for a class of methines, reflecting continued uncertainty in correct assignments. This class of methines included carbons at C(5), C(9), C(10), and C(13) of the aconitine skeleton. The resonances for these carbons, all of which fall in the same region of the spectrum, show a complex and unpredictable response to substitution of other carbons in the molecule. Consequently, in many studies they were only tentatively assigned.[4,9a,13] The resonances for tertiary carbons at these four locations in the aconitine skeleton have been of limited usefulness in structure determination. In spectrum analysis by the computer methods discussed here, the substructures retrieved for such resonances generally had little in common, and it was difficult to deduce detailed structural fragments from them.

Results and Discussion

(1) Identification of an Incorrect Structure Assignment. The unusual structures **3** and **4** were assigned to the alkaloids "iliensine" and "acomonine", respectively, on the basis of chemical and spectral analysis.[14] These are the only published structures of C_{19}-deterpenoid alkaloids which lack a C(1) substituent.

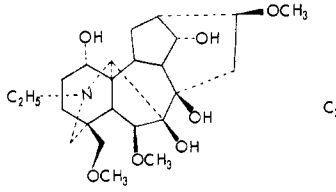


3 Iliensine

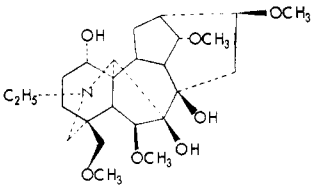


4 Acomonine

Our analysis of the ^{13}C NMR spectra,[15] however, showed "iliensine" and "acomonine" to be identical with delcosine (**5**) and delsoline (**6**), compounds whose structures have been known for many years. The result can be demonstrated quickly and easily by computer methods as shown by the following experiment.



5 Delcosine



6 Delsoline

The proposed structure for "iliensine" (**3**) was entered into the program[10] for the prediction of ^{13}C NMR spectra. By matching the environment of each carbon in the structure with substructures in the data base created from standard C_{19}-diterpenoid alkaloid structures, the program was able to predict a range for the chemical shifts of each carbon. It also attempted to correlate the predicted spectrum with the observed spectrum after accepting the measured resonances, including multiplicities, as input: the lists of observed and predicted chemical shifts were each grouped according to multiplicity, the shifts in each multiplicity group were ordered by increasing magnitude, and the corresponding values in each list were matched. The results are shown in Table I. Note that C(11) of the "iliensine" structure has a unique environment which could not be matched at even a one-shell level with substructures in the data base. The range of shifts for quaternary carbons in the data base is so wide (60 ppm) that the resonance average predicted for C(11) is practically meaningless, and the fact that it is not close in value to any of the

(13) S. W. Pelletier and Z. Djarmati, *J. Am. Chem. Soc.*, **98**, 2626 (1976).

(14) (a) V. E. Nezhevenko, M. S. Yunusov, and S. Y. Yunusov, *Khim. Prir. Soedin.*, **10**, 409 (1974); (b) *ibid.*, **11**, 389 (1975); (c) M. S. Yunusov, V. E. Nezhevenko and S. Y. Yunusov, *ibid.*, **11**, 107 (1975); (d) *ibid.*, **11**, 770 (1975).

(15) S. W. Pelletier and N. V. Mody, *Tetrahedron Lett.* **22**, 207 (1981).

Table I. Observed and Predicted Chemical Shifts[a] for "Iliensine" (3)

atom	mult	shell[b]	av reson[c]	σ(av)[d]	obsd reson
C(1)	t	1	33.0	5.0	29.4
C(2)	t	2	29.0	2.3	29.4
C(3)	d	1	71.7	5.0	72.7
C(4)	s	2	43.4	2.3	37.6
C(5)	d	1	46.1	5.0	45.3
C(6)	d	2	90.7	2.3	90.1
C(7)	s	2	88.4	2.3	87.9
C(8)	s	3	76.9	1.0	78.1
C(9)	d	2	48.1	2.3	45.3
C(10)	d	1	40.3	5.0	39.4
C(11)	s	0	61.8	18.8	48.9
C(12)	t	2	28.9	2.3	27.5
C(13)	d	3	44.6	1.5	44.0
C(14)	d	3	75.4	1.0	75.8
C(15)	t	4	33.5	0.8	34.5
C(16)	d	4	81.8	0.4	82.0
C(17)	d	1	64.9	5.0	66.3
C(18)	t	3	76.4	1.0	77.4
C(19)	t	2	48.9	2.3	50.4
C(20)	q	4	14.1	0.4	13.7
C(21)	t	3	50.7	1.0	57.1
C(6′)	q	4	57.9	0.7	57.4
C(16′)	q	4	56.3	0.4	56.4
C(18′)	q	4	59.0	0.4	59.1

[a] In parts per million downfield from Me_4Si. [b] Maximum shell out to which substructures in the data base could be matched to the environment of the resonating carbon. [c] Average resonance value of the substructures matching at the maximum shell level. [d] Standard deviation of the resonance average.

observed singlets is not very informative. Of much more importance is the discrepancy between the observed and predicting shifts for C(4). The predicted shifts for this carbon fell within a narrow range, and the standard deviation of the average shift is fairly small. None of the observed singlet shifts were within the range of predicted shifts or within 2 standard deviations of the average predicted resonance. Also significant is the fact that the observed triplet at 57.1 ppm did not fall within the ranges of the predicted shifts of any of the secondary carbons. According to these results, the structure proposed for "iliensine" is incorrect.

More likely structures for "iliensine" (3) were determined from its observed ^{13}C NMR spectrum. The molecular formula for the compound, $C_{24}H_{39}NO_9$, and the observed ^{13}C resonances with their multiplicities were entered into the program[10] for spectrum analysis. The program was directed to retrieve substructures from the data base whose resonances matched the input resonances to within 1.5 ppm. The common portions of the substructures remaining for each observed resonance after interpretation, truncated to the second shell where necessary, are shown in Figure 1. The 19 fragments which included at least two nonhydrogen atoms were marked by the program as possible substructure constraints for GENOA and were automatically passed to GENOA in coded form.

The aconitine skeleton, with possible substitution sites indicated by specifying a range for the hydrogens attached to each carbon, was defined and used as the first constraint on structure generation in GENOA. In formulating this constraint, the empirical rules for substitution of the C_{19}-diterpenoid alkaloids were used conservatively. Substitution was allowed, but not required, at many atoms normally substituted in these alkaloids, and no information on the likely substituents for any substitution point was included. Substitution or a double bond was required at C(8) and C(14) and only one substituent per carbon was

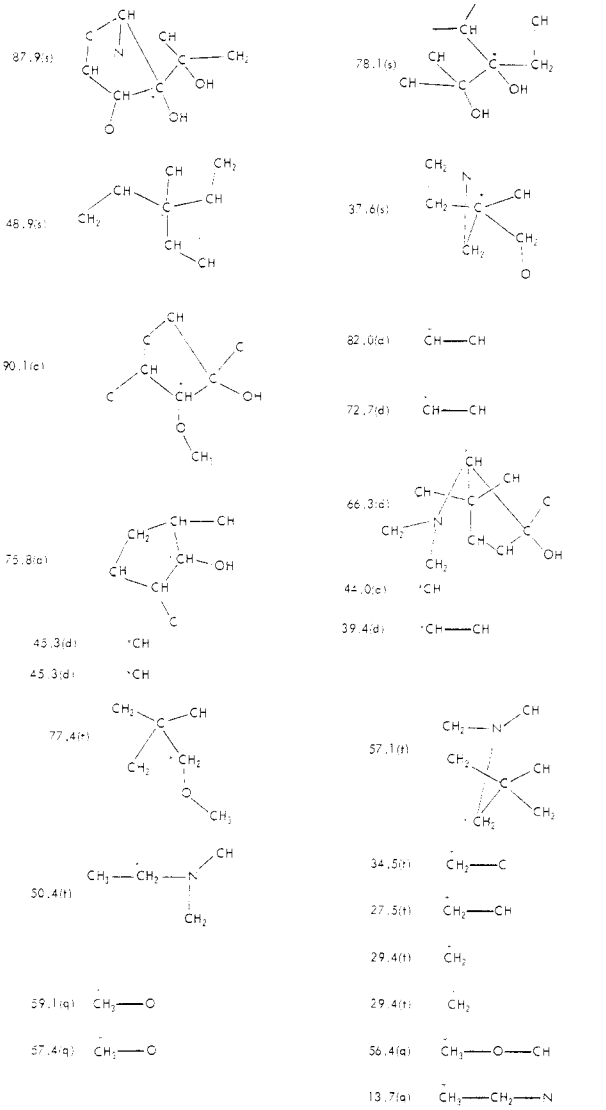
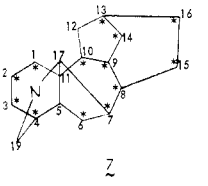


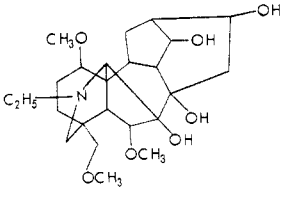
Figure 1. Substructures inferred for the resonances of "iliensine" (3). An asterisk marks the resonating carbon. Substituents not explicitly shown may be any nonhydrogen atom.

allowed. C(18) was not included in the skeleton since several natural products had been isolated in which this carbon was not present.[2] The skeleton as defined (7) is shown with each possible substitution site marked with an asterisk.

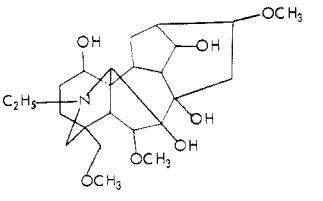


7

Subsequent application of the constraints passed to GENOA from the spectrum analysis program served to limit the generated structures to 8 and 9. Each of these two



8



9

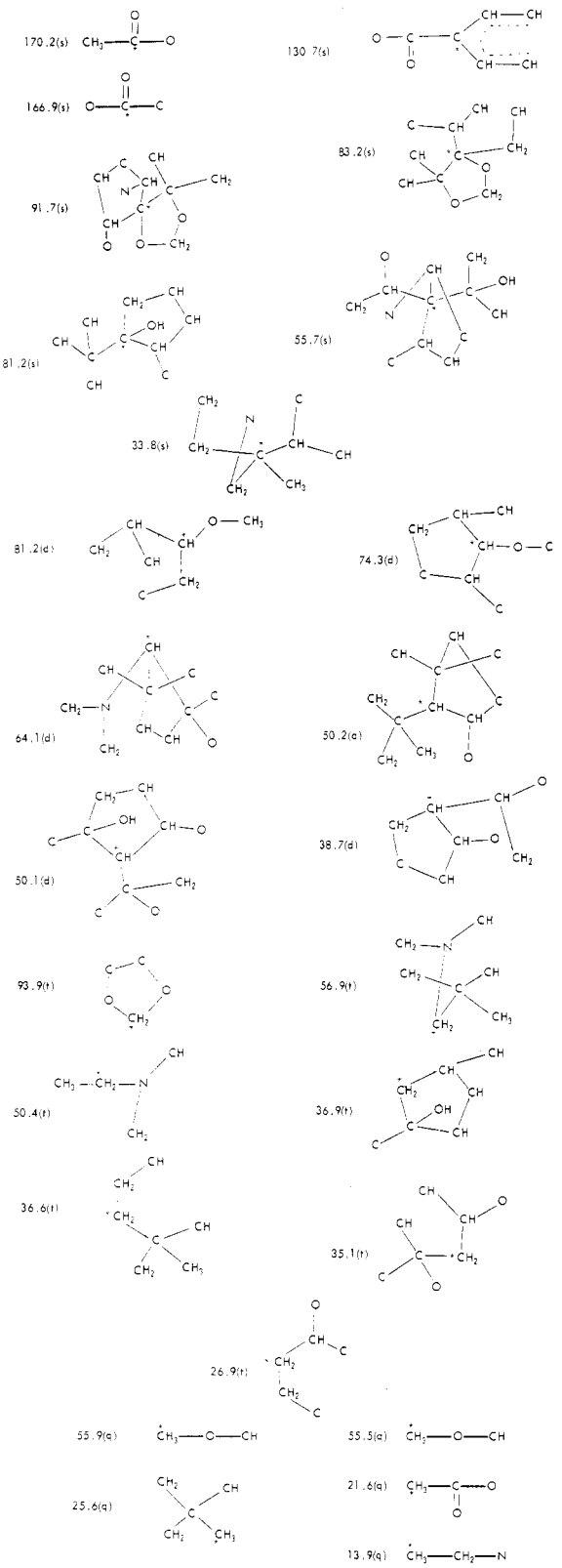


Figure 2. Two-shell substructure constraints derived for glaucephine (**10**). An asterisk marks the resonating carbon. Substituents not explicity shown may be any nonhydrogen atom.

structures was evaluated by predicting its ^{13}C NMR spectrum (disregarding stereochemistry) and correlating it with the observed spectrum for "iliensine" (Table II). The structures were ranked according to the agreement between observed and predicted resonances. The constitution of the higher ranked structure (**9**, Table II) matched that of delcosine (**5**). At this point, all stereocenters in the higher ranked structures except for C(1) and

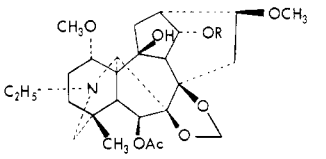
Table II. Observed and Average Predicted Chemical Shifts[a] for the Two Candidate Structures, 8 and 9, for "Iliensine" (3)

	predicted	
obsd	8	9
87.8 (s)	88.4 (s)	87.8 (s)
78.1 (s)	77.4 (s)	76.8 (s)
48.9 (s)	48.1 (s)	48.9 (s)
37.6 (s)	38.2 (s)	37.5 (s)
90.0 (d)	90.6 (d)	90.2 (d)
81.9 (d)	84.0 (d)	81.8 (d)
75.8 (d)	75.5 (d)	75.4 (d)
72.7 (d)	64.7 (d)	72.6 (d)
66.3 (d)	49.6 (d)	66.1 (d)
45.3 (d)	49.1 (d)	45.3 (d)
45.3 (d)	45.8 (d)	44.9 (d)
44.0 (d)	43.2 (d)	43.8 (d)
39.4 (d)	36.6 (d)	39.4 (d)
77.4 (t)	77.8 (t)	77.8 (t)
57.1 (t)	54.6 (t)	54.6 (t)
50.4 (t)	50.9 (t)	50.9 (t)
34.5 (t)	36.4 (t)	33.6 (t)
29.4 (t)	32.1 (t)	29.5 (t)
29.4 (t)	28.9 (t)	29.4 (t)
27.5 (t)	26.0 (t)	27.3 (t)
59.1 (q)	59.1 (q)	59.1 (q)
57.4 (q)	57.8 (q)	57.8 (q)
56.4 (q)	55.9 (q)	56.3 (q)
13.7 (q)	14.0 (q)	14.0 (q)
compd	rank	score[b]
8	2	52.3
9	1	4.8

[a] In parts per million downfield from Me_4Si; multiplicities are given in parentheses. [b] The score is a measure of the degree of mismatch between the observed and predicted spectra.[10]

C(6) were assigned their usual configurations, and the four possible stereoisomers remaining were generated.[16] Spectra were predicted for each stereoisomer, this time with the complete stereochemical substructure codes in the data base,[10,11] and the structures were rank ordered as before. The highest ranked structure had an α-hyroxyl group at C(1) and a β-methoxyl group at C(6) and was identical with the structure of delcosine (**5**). Since "iliensine" (**3**) and "acomonine" (**4**) have been chemically correlated,[14] it followed that "acomonine" (**4**) had the structure of delsoline (**6**).

(2) Structure Determination of New Compounds. Glaucephine (**10**), glaucerine (**11**), and glaucenine (**12**) are

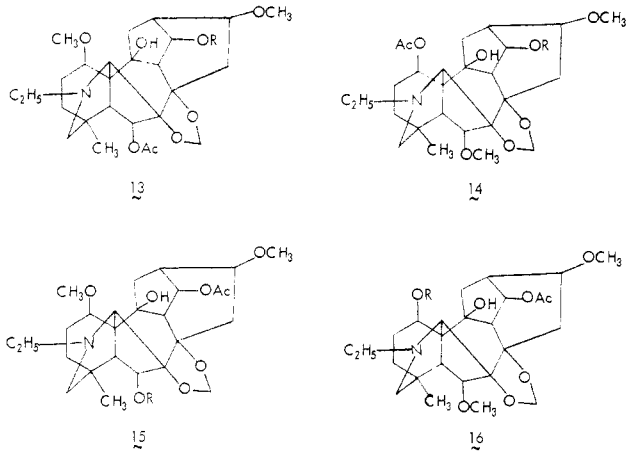


10 R = Bz Glaucephine
11 R = $COCH(CH_3)_2$ Glaucerine
12 R = $COCH(CH_3)C_2H_5$ Glaucenine
17 R = H Dictyocarpine

three new alkaloids recently isolated from *Delphinium glaucesens*.[17] An attempt was made to determine the structures for these compounds by using the computer. The observed resonances for glaucephine, $C_{33}H_{43}NO_9$, were entered into the program for spectrum analysis. The

(16) J. G. Nourse, D. H. Smith, R. E. Carhart, and C. Djerassi, *J. Am. Chem. Soc.*, **102**, 6289 (1980).
(17) S. W. Pelletier, O. D. Dailey, Jr., N. V. Mody, and J. D. Olsen, *J. Org. Chem.*, **46**, 3284 (1981).

Chart I[a]



[a] a, R = Bz (glaucephine); b, R = $COCH(CH_3)_2$ (glaucerine); c, R = $COCH(CH_3)CH_2CH_3$ (glaucenine).

aconitine skeleton with likely substitution sites (**7**) was defined as known. A benzene ring, identified by the characteristic pattern of five doublets and one singlet with resonances around 130 ppm, was also defined as a known fragment of the structure. Substructures with resonances within 1.5 ppm of those observed were retrieved from the alkaloid data base. This set of substructures was required to be internally consistent, as well as consistent with defined substructures, out to two shells. Including the aconitine skeleton and benzene ring as known fragments of glaucephine enabled the program to infer the two-shell enviroments of almost all of the resonating carbons, as shown in Figure 2. In addition, the program was able to associate several resonances with specific carbons in the aconitine skeleton.

Using the aconitine skeleton (**7**), benzene ring, and the substructures in Figure 2 as constraints, GENOA generated only four possible structures, **13a**, **14a**, **15a**, and **16a** (see Chart I). Structures **14a** and **16a**, which had the unprecedented features of an ester group at C(1) of the aconitine skeleton, seemed unlikely but were entirely consistent with the two-shell environments derived for the resonating carbons. After evaluation of the candidates, however, (Table III), the high mismatch scores for these two structures eliminated them from further consideration. The difference between the scores for **13a** and **15a** was not significant enough to distinguish which structure was correct. Ester substitution at C(14) is more common than at C(6) in the C_{19}-diterpenoid alkaloids, so the substitution pattern of **13a** was more likely than that of **15a**. Synthesis of glaucephine by C(14) esterification of dictyocarpine (**17**) proved the correct structure did have the substitution pattern of **13a**, and it also established the correct stereochemistry of the compound.

Analysis of the spectra of glaucerine, $C_{30}H_{45}NO_9$, and glaucenine, $C_{31}H_{47}NO_9$, gave results similar to those for glaucephine. Both alkaloids had ester substituents not represented in the diterpenoid alkaloid data base, and a larger, general data base,[11] which included all the substructures from the alkaloid data base, was used for spectrum analysis. The general data base contained over 12 000 distinct substructures, as compared to the 2000 or so in the alkaloid data base; therefore, it was necessary to be more selective in retrieving substructures during spectrum analysis. Only substructures whose resonances matched input resonances to within 1.25 ppm, or in the case of quartets, 0.75 ppm, were retrieved. The retrieved substructures were required to be self-consistent at the two-shell level and also to be consistent with the aconitine skeleton (**7**). For both compounds, detailed environment for most of the resonating carbons were deduced. GENOA generated structures **13b–16b** and **13c–16c** for glaucerine and glaucenine, respectively. Evaluation of the four candidates for each compound limited the likely structures to **13b** or **15b** for glaucerine and **13c** or **15c** for glaucenine, and syntheses from dictyocarpine showed **13b** and **13c** to be correct.

Table III. Observed and Predicted Spectral Data[a] for the Candidate Structures for Glaucephine (10)

	predicted			
obsd	13a	14a	15a	16a
170.2 (s)	170.0 (s)	170.4 (s)	171.7 (s)	171.7 (s)
166.9 (s)	166.5 (s)	166.5 (s)	166.4 (s)	166.4 (s)
130.7 (s)	132.7 (s)	132.7 (s)	130.1 (s)	130.7 (s)
91.6 (s)	91.6 (s)	92.1 (s)	92.1 (s)	92.1 (s)
83.1 (s)	83.3 (s)	83.5 (s)	83.3 (s)	83.5 (s)
81.2 (s)	81.2 (s)	81.7 (s)	81.2 (s)	81.7 (s)
55.6 (s)	55.9 (s)	55.5 (s)	55.9 (s)	55.5 (s)
33.8 (s)	33.9 (s)	33.9 (s)	33.9 (s)	33.9 (s)
132.7 (d)	132.1 (d)	132.1 (d)	132.1 (d)	132.1 (d)
129.9 (d)	130.1 (d)	130.1 (d)	129.7 (d)	129.8 (d)
129.9 (d)	130.1 (d)	130.1 (d)	129.7 (d)	129.8 (d)
128.3 (d)	128.5 (d)	128.5 (d)	128.5 (d)	128.5 (d)
128.3 (d)	128.5 (d)	128.5 (d)	128.5 (d)	128.5 (d)
81.2 (d)	81.2 (d)	85.4 (d)	81.2 (d)	85.4 (d)
79.0 (d)	78.9 (d)	83.0 (d)	78.9 (d)	83.0 (d)
77.4 (d)	77.2 (d)	81.2 (d)	77.4 (d)	81.2 (d)
74.3 (d)	74.7 (d)	74.7 (d)	74.7 (d)	74.7 (d)
64.1 (d)	63.9 (d)	64.1 (d)	63.9 (d)	64.1 (d)
50.2 (d)	50.8 (d)	50.7 (d)	50.7 (d)	50.7 (d)
50.1 (d)	49.9 (d)	49.9 (d)	49.9 (d)	49.9 (d)
38.7 (d)	38.9 (d)	38.9 (d)	38.9 (d)	38.9 (d)
93.8 (t)	93.8 (t)	93.8 (t)	93.8 (t)	93.8 (t)
56.9 (t)	56.9 (t)	56.9 (t)	56.9 (t)	56.9 (t)
50.4 (t)	50.3 (t)	50.3 (t)	50.3 (t)	50.3 (t)
36.9 (t)	37.3 (t)	37.3 (t)	37.3 (t)	37.3 (t)
36.6 (t)	36.4 (t)	36.6 (t)	36.4 (t)	36.6 (t)
35.1 (t)	34.9 (t)	34.9 (t)	34.9 (t)	34.9 (t)
26.9 (t)	26.5 (t)	26.1 (t)	26.5 (t)	26.1 (t)
55.9 (q)	56.3 (q)	56.3 (q)	56.3 (q)	56.3 (q)
55.5 (q)	55.5 (q)	56.3 (q)	55.5 (q)	56.3 (q)
25.6 (q)	25.5 (q)	25.5 (q)	25.5 (q)	25.5 (q)
21.6 (q)	21.6 (q)	22.3 (q)	21.4 (q)	21.4 (q)
13.9 (q)	14.0 (q)	14.0 (q)	14.0 (q)	14.0 (q)

compd	rank	score[b]
13a	1	3.0
14a	3	18.1
15a	1	2.2
16a	3	16.8

[a] Chemical shifts are given in parts per million downfield from Me_4Si; multiplicities are given in parentheses. [b] See footnote in Table II for the definition of the mismatch score.

(3) The Problem of an Inadequate Data Base. In the examples discussed in the previous sections, the two-shell substructures of the unknown compounds were all represented in the ^{13}C NMR data base used for spectrum analysis. These examples represent ideal cases. Often one or more two-shell substructures of an unknown have no representatives in the data base, especially if the data base is small. In these cases, the programs will usually encounter inconsistencies during interpretation or structure generation which lead to no solutions or, possibly, lead to suggestions of incorrect structures, although we have not yet encountered the latter. As the following example demonstrates, it is sometimes possible for the user to resolve the inconsistencies resulting from an inadequate data base, thereby allowing the programs to successfully obtain structural candidates.

Table IV. Observed and Predicted Chemical Shifts[a] for Structure 18[b]

atom	mult	shell	av reson	σ(av)	obsd reson
C(1)	d	4	86.2	1.0	86.2
C(2)	t	3	26.0	1.0	25.8
C(3)	t	2	32.4	2.3	32.2
C(4)	s	3	37.7	1.0	39.1
C(5)	d	2	42.8	2.3	45.6
C(6)	t	3	25.0	1.0	24.6
C(7)	d	3	45.9	1.0	46.0
C(8)	s	2	73.9	2.3	73.0
C(9)	d	3	46.8	1.0	47.0
C(10)	d	2	38.8	2.3	37.6
C(11)	s	2	49.9	2.3	48.8
C(12)	t	2	28.7	2.3	27.7
C(13)	d	1	43.7	5.0	45.9
C(14)	d	2	75.5	2.3	82.3
C(15)	t	1	36.1	5.0	38.3
C(16)	d	0	63.0	17.7	75.6
C(17)	d	3	62.3	1.0	63.0
C(18)	t	4	79.1	0.4	68.8
C(19)	t	4	55.4	1.9	53.1
C(20)	q	4	13.2	0.4	13.7
C(21)	t	4	48.7	0.5	49.5
C(1′)	q	4	55.9	0.4	56.3
C(18′)	q	4	59.1	0.4	56.5

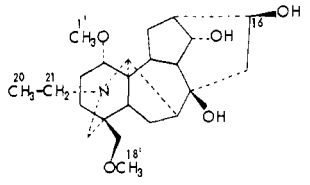
[a] In parts per million downfield from Me_4Si. [b] Score = 76.46; see footnote in Table II for a definition of score.

Table V. Observed and Predicted Chemical Shifts[a] for Structure 19[b]

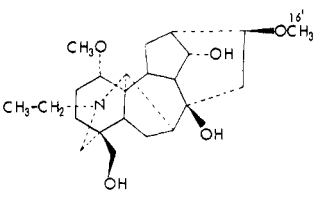
atom	mult	shell	av reson	σ(av)	obsd reson
C(1)	d	4	86.2	1.0	86.2
C(2)	t	3	26.0	1.0	25.8
C(3)	t	2	32.3	2.3	32.2
C(4)	s	1	38.6	5.0	39.1
C(5)	d	2	42.8	2.3	45.6
C(6)	t	3	25.0	1.0	24.6
C(7)	d	3	45.9	1.0	46.0
C(8)	s	4	73.8	2.3	73.0
C(9)	d	3	46.8	1.0	47.0
C(10)	d	2	38.4	2.3	37.6
C(11)	s	2	49.9	2.3	48.8
C(12)	t	3	28.5	1.1	27.7
C(13)	d	4	44.8	0.8	45.9
C(14)	d	4	75.6	0.4	75.6
C(15)	t	4	40.9	2.5	38.3
C(16)	d	4	82.3	0.4	82.3
C(17)	d	3	62.3	1.0	63.0
C(18)	t	2	67.3	2.3	68.8
C(19)	t	2	54.2	2.3	53.1
C(20)	q	4	13.2	0.4	13.7
C(21)	t	4	48.7	0.5	49.5
C(1′)	q	4	56.3	0.4	56.3
C(18′)	q	4	56.3	0.4	56.5

[a] In parts per million downfield from Me_4Si. [b] Score = 8.63; see footnote in Table II for a definition of score.

Cammaconine, a diterpenoid alkaloid isolated from *Aconitum variegatum*, was assigned structure **18** by



18



19

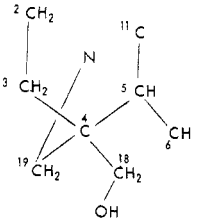
chemical correlation with another C_{19}-diterpenoid alkaloid and by spectral analysis.[18] Compound **18** is a very rare example of a C_{19}-diterpenoid alkaloid lacking a methoxyl substituent at C(16). The reported data, however, did not exclude another less unusual structure, **19**, for cammaconine.[19]

The molecular formula of cammaconine, $C_{23}H_{37}NO_5$, and its ^{13}C NMR shifts with their multiplicities were input to the program for spectrum analysis. Substructures associated with resonances within 1.5 ppm of the observed shifts were retrieved. Initially these were required to be self-consistent at the two-shell level and also to be consistent with the aconitine-type skeleton, but in the course of interpretation the program generated warning messages indicating that it could not obtain an internally consistent subset of two-shell substructures which were also compatible with the aconitine skeleton. Examination of the substructures which had, at that point, been inferred for the resonances showed that they accounted for six different substitution sites on the aconitine skeleton, yet the molecular formula for cammaconine allowed for, at most, five substituents. Some of the information derived by the analysis was therefore incorrect.

To help identify the weaknesses in the data base which were responsible for the wrong information, we examined the computer-predicted ^{13}C NMR spectra for the two hypothesized structures, **18** and **19**. Tables IV and V show the predicted spectra for **18** and **19**, respectively, correlated with the observed spectrum of cammaconine. Also presented in each table is the "disbelief" score obtained by comparing the predicted and observed spectra. A large mismatch score indicates that a particular hypothesized structure is probably incorrect. Comparison of the scores for **18** and **19** shows structure **19** to be more consistent with the observed spectrum. Significant for excluding **18** as a plausible structure for cammaconine are the discrepancies between the observed and predicted chemical shifts for its carbons at C(18) and C(18′). The differences between the predicted and observed chemical shifts for these two carbons are more that six times the standard deviations associated in the data base for carbons in such environments (σ_{av}). In contrast, the resonance averages for **19** could all be correlated with observed resonances to within 2 standard deviations.

Table IV also shows that the C(4) carbon of **19** has a two-shell environment not represented in the alkaloid data base. Only substructures matching C(4) at the one-shell level were retrieved during spectrum prediction. On the assumption that the environment of C(4) in **19** is correct, this weakness in the data base could account for the failure of the spectrum analysis program to retrieve internally consistent information for the observed spectrum.

The two-shell substructure derived for the C(4) singlet (39.1 ppm) during spectrum analysis is shown in **20**, num-



20

bered with the same scheme as used for the aconitine

(18) M. A. Khaimova, M. D. Palamareva, N. M. Mollov, and V. P. Krestev, *Tetrahedron*, **27**, 819 (1971).
(19) N. V. Mody, S. W. Pelletier, and N. M. Mollov, *Heterocycles*, **14**, 1751 (1980).

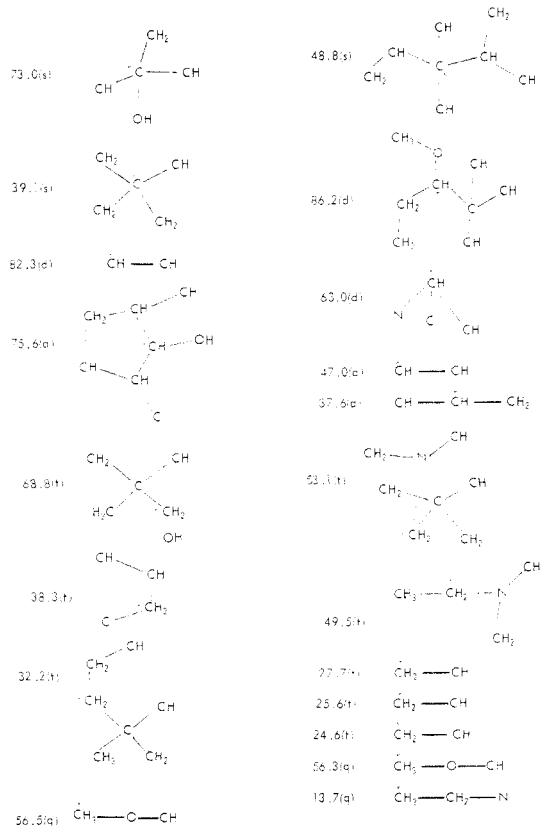


Figure 3. Substructure constraints derived for cammaconine (**19**). An asterisk marks the resonating carbon. Substituents not explicitly shown may be any nonhydrogen atom.

skeleton **7**. C(6) of the substructure has only one hydrogen substituent, implying oxygenation of this carbon in cammaconine. Substitution at C(6), however, contradicts the assumption that the C(4) environment of **19** is correct. The decision was made to use only the one-shell portion of this substructure. Substitution at C(6) was then no longer implied, and the inconsistencies in the inferred information were eliminated.

It is important to note that this decision effectively eliminated the possibility of substitution at C(6) of the aconitine skeleton, thereby biasing the outcome of structure elucidation. It was necessary to make some assumption about the final structure merely to resolve the inconsistencies and allow the program to obtain structural candidates. The assumption that the C(4) environment of **19** was correct was reasonable, considering the close match between the spectrum predicted for this structure and the observed resonances.

Given the above assumption, the interpretation procedure continued uneventfully, eventually inferring the substructure constraints shown in Figure 3. These substructural constraints plus the aconitine skeleton were incorporated into a single constitutional isomer, that with the same connectivity as structure **19**, by GENOA. Two stereoisomers, differing only in their configuration at C(1), were generated and evaluated by comparing their predicted spectra with the observed spectrum for cammaconine. The correct stereoisomer could not, however, be distinguished, as the spectrum mismatch scores for the two isomers were similar.

This example demonstrates that the major limitation of the programs is that their success largely depends on a data base with a complete set of substructures. As in this example, the problems arising from an inadequate data base can often be overcome by user intervention, but at the expense of objectivity. The programs are designed to allow extensive interactions with the user, and this interaction is at times crucial to obtaining successful results.

Conclusion

Computer programs for ^{13}C NMR interpretation and structure prediction have proven useful in elucidating the structures, including in some cases stereochemistry, of C_{19}-diterpenoid alkaloids. The programs also provide a method for storing ^{13}C NMR data in a convenient and easily accessible form. The stored data can be retrieved to check the consistency of spectral assignments with those made previously or to lend precision to arguments about structures which are based on ^{13}C NMR measurements.

The success of the programs in predicting new structures depends heavily on the availability of an adequate data base. Even the relatively small data base created for the C_{19}-diterpenoid alkaloids is adequate for solving many diterpenoid alkaloid structures because variation in the substitution patterns for this class of natural products is small. Success in predicting large structures such as the diterpenoid alkaloids also depends on judicious use of structural information which is either already known or which can be safely assumed about the compound. Of course, the more such information used, the easier a problem becomes, but including many assumptions about the structure also decreases the objectivity of structure determination.

The C_{19}-diterpenoid alkaloids are complicated structures, containing as many as 56 atoms. The straightforward application of the computer programs described in this paper to the diterpenoid alkaloids is, therefore, a striking demonstration of the power of ^{13}C NMR spectroscopy as a method of structure determination. The power of this technique has been enhanced by the ability to create, and utilize by computer, data bases of ^{13}C NMR spectra.

Acknowledgment. We thank the National Institutes of Health (Grants RR-00612, AM-04247, and CA-24180) for generous financial support. Computer resources were provided by the SUMEX facility at Stanford University under National Institutes of Health Grant RR-0785.

Registry No. 5, 545-56-2; **6**, 509-18-2; **10**, 78018-30-1; **11**, 78018-29-8; **12**, 78018-27-6; **19**, 32152-70-8.