HIGH-PRESSURE LIQUID CHROMATOGRAPHY, NUCLEAR MAGNETIC RESONANCE AND MASS SPECTRA OF BIOSYNTHETIC SOYASAPOGENOLS

ERICH HEFTMANN[1], ROBERT E. LUNDIN, WILLIAM F. HADDON

Plant Biochemistry and Instrumental Analysis Research Units, Western Regional Research Center, Science and Education Administration, U.S. Department of Agriculture, Berkeley, California, U.S.A. 94710

and

IRENA PERI, UZI MOR, ARON BONDI

Faculty of Agriculture, The Hebrew University of Jerusalem, Rehovot, Israel 76-100

ABSTRACT.—Soybean plants were grown in the presence of mevalonic acid-[2-^{14}C], and the radioactive soyasapogenols were extracted. Soyasapogenols A, B, C, and E were isolated by high-pressure liquid chromatography and characterized by nuclear magnetic resonance and mass spectrometry.

Soybeans, which contain saponins (1), are an important food source for farm animals (2, 3). On hydrolysis, these saponins yield sugars and aglycones. Five aglycones have been isolated from soybeans and designated as soyasapogenols A, B, C, D, and E. The structure of soyasapogenol D, which may be an artifact of methanolysis, is still uncertain. The remaining four structures are shown in fig. 1. The soyasapogenols are pentacyclic triterpenoids differing from each other in the nature and number of oxygen functions.

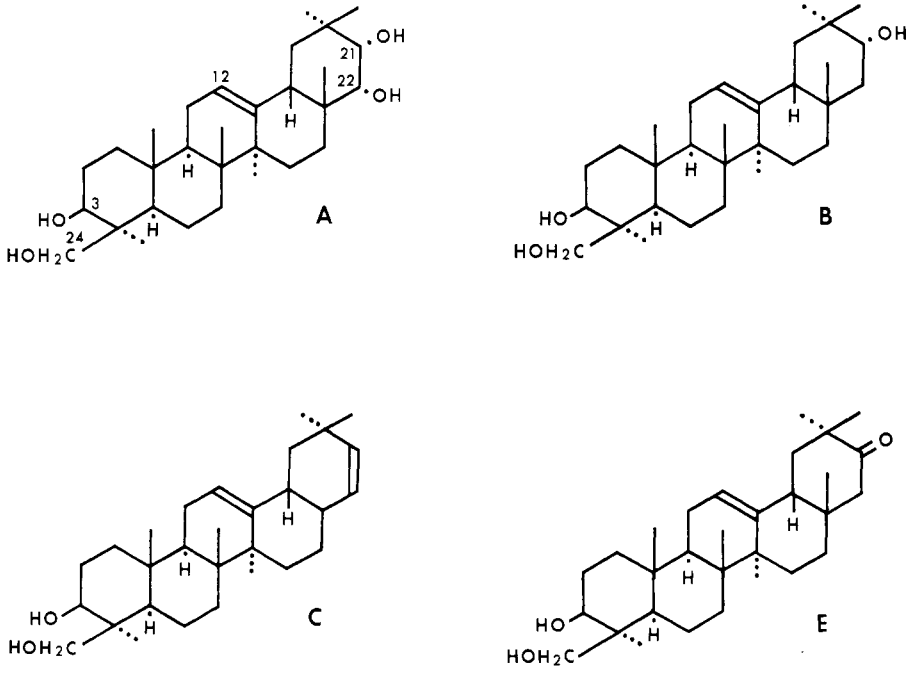

FIG. 1. Soyasapogenols.

[1]To whom inquiries are to be addressed.

In connection with our work on the biosynthesis of soyasapogenols in soybean seedlings (4), we have felt the need for identification of sapogenols beyond the comparison with reference material by thin-layer chromatography. Finding the supply of reference material and of reference spectra wanting, we have devised a preparative method for the isolation of radioactive sapogenols by high-pressure liquid chromatography (hplc) and have determined their nuclear magnetic resonance (nmr) and mass spectra.

EXPERIMENTAL

Sterilized germinating seeds of soybean (*Glycine max* var. Lincoln) were incubated with 1.48×10^7 cpm D,L-mevalonic acid-[2-^{14}C][2] for 10 days at 22°, for the first 3 days in the dark and for the following 7 days illuminated with white fluorescent light.[3] The seedlings were extracted by homogenization with 80% ethanol. After removal of ethanol from the filtrate under reduced pressure, the lipids and pigments were separated from the watery residue by partition in chloroform-acetone-water (4:3:2).

The saponins were hydrolyzed by refluxing the aqueous phase with 2*N* H_2SO_4 for 10 hr, and afterwards the sapogenols were extracted with chloroform-ether (1:1). The residue obtained after evaporation of these solvents was purified by boiling with 6% (w/v) KOH in ethanol for 1 hr and was again extracted with chloroform-ether (1:1). This extract, containing the sapogenols, gave 1.77×10^5 cpm. An aliquot of this extract, having a radioactivity of 3.3×10^4 cpm, was dissolved in 1 ml ethanol and injected into the hplc apparatus.

The hplc apparatus was assembled from commercially available components. The pump was of the single-piston reciprocating type,[4] operated at 1 ml/min. The detector was a Hitachi variable-wavelength spectrometer, equipped with a flowcell having a 10-mm pathlength and a 20-μl volume,[5] which was set to 208 nm. One channel of a dual-channel recorder[6] was connected to the output of the detector. The sample injector was a six-port rotary valve[7] with a 2-ml sample loop.

A 16-ft chromatographic column was assembled from four 2-ft and two 4-ft sections of 7-mm I.D. stainless-steel tubes, $\frac{3}{8}$ in. O.D., capped with 10-μm end fittings.[8] Each section was packed with Porasil A[8], 37–75 μm, by pouring small portions into the tube and tapping its end against the table top. The sections were interconnected by short 0.009-in. I.D. stainless-steel tubes with Swagelok lock nuts and ferrules, and a forecolumn, 6 cm in length and likewise packed with Porasil A, was installed between the injector and the main column.

The eluent was a mixture of *n*-hexane-ethanol-acetonitrile (13:2:1).[9] The effluent stream passed from the detector into a 5-ml syphon of a fraction collector[10], equipped with an event marker. At the completion of each collection period, the event marker produced a signal which was fed into the second channel of the recorder.

A 0.5-ml aliquot of each fraction was added to 10 ml of a scintillation cocktail, prepared by dissolving 6 g of 2,5-diphenyloxazole (PPO) and 150 mg of 1,4-bis[2-(5-phenyloxazoyl)]benzene (POPOP)[11] in 1 liter of toluene, and counted in a liquid scintillation counter.[12]

The nmr spectra were obtained on a spectrometer[13] operating at a frequency of 90 MHz with 10% chloroform as the internal lock and approximately 0.01% tetramethylsilane (TMS) as internal reference. Chemical shifts were determined in ppm downfield from TMS (δ values). For all nmr data the sample temperature was 34° and concentrations were in the range of 1–2% by weight. Temperature variation and D_2O exchange were employed in some cases to confirm spectral assignments.

Electron ionization (ei) mass spectra at 70 eV were obtained on the underivatized soyasapogenols by use of the direct insertion probe on an MM-70/70F double-focusing mass spectrometer.[14] The ion-source temperature was 200°. Accurate masses of molecular ions were

---

[2]DBED salt, 42 μCi/mg, Radiochemical Centre, Amersham, England. Reference to a company and/or product named by the Department is only for purposes of information and does not imply approval or recommendation of the product to the exclusion of others which may also be suitable.

[3]Cool White, photon fluence rate (Ph AR) 400–700 μEinstein m^{-2} sec^{-1}.

[4]Model 110, Altex Scientific Inc., Berkeley, California.

[5]Model 155-30, Altex Scientific Inc., Berkeley, California.

[6]Model 385, Linear Instruments Inc., Irvine, California.

[7]Model 706, Disc Instruments Inc., Costa Mesa, California.

[8]Waters Associates, Milford, Massachusetts.

[9]"Distilled in Glass" quality, Burdick & Jackson, Muskegon, Michigan.

[10]Radirac Model 3400B, LKB Produkter AB, Bromma, Sweden.

[11]Amersham/Searle, Arlington Heights, Illinois.

[12]Tricarb Model 3003, Packard Instrument Co., Downers Grove, Illinois.

[13]Model EM-390, Varian Instrument Division, Palo Alto, California.

measured by peak matching on a CEC 21-110A mass spectrometer.[15] The complete high-resolution spectrum of soyasapogenol A was determined with the MM-70/70F in conjunction with a model 2035 data system.[16]

RESULTS AND DISCUSSION

HPLC

Fig. 2 shows the elution pattern obtained by hplc of the radioactive soybean extract. The curve records the absorbance of the eluate. The radioactivity of the fractions is represented by a bar graph. Fractions from individual zones were pooled, rechromatographed under the same conditions, and then recrystallized to constant specific activity (4). The identity of the soyasapogenols was confirmed by nmr and mass spectrometry (see below).

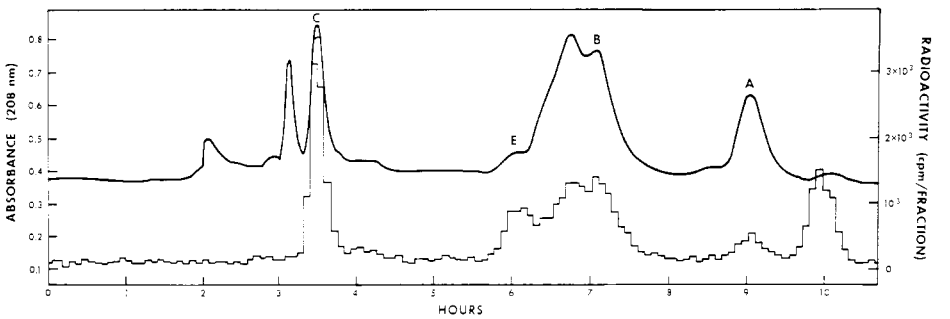

FIG. 2. HPLC or Radioactive Soybean Extract. Column, 16 ft x 7 mm I.D.; sorbent, Porasil A, 37-75 μm; eluent, *n*-hexane-ethanol-acetonitrile (13:2:1); flow rate, 1 ml/min; detector at 208 nm, range at 1.0, time constant at 1.0; dual-pen recorder with event marker at 5 min, span 10mV; chart speed, 6 cm/hr; 5 ml fractions collected. Curve, absorbance at 208 nm; bar graph, radioactivity in cpm/fraction. A, B, C, and E are the radioactive soyasapogenols.

To our knowledge, the separation of soyasapogenols by hplc has not previously been reported. The method we have used is a modification of the preparative methods we have devised for the separation of steroidal alkaloids (5), free sterols (6), and gibberellin esters (7). Their main advantages are that the capacity of the columns is large enough for crude extracts and preparative work and that the sorbent is cheap and easy to pack. Their main disadvantages are that the separation is slow and less efficient than chromatography with the finer sorbents.

We have succeeded in separating the four known soyasapogenols A, B, C, and E. A reference sample of soyasapogenol D was eluted in an area of the chromatogram without appreciable radioactivity, near soyasapogenol C. Apparently no soyasapogenol D was formed under the conditions of our experiment. The order of elution is as expected: Soyasopogenol C, having 2 hydroxyl groups, is followed by soyasapogenol B with 3 hydroxyl groups, and then soyasapogenol A with 4 hydroxyl groups. Soyasapogenol E, the keto analog of soyasapogenol B, precedes the latter. Other radioactive metabolites, isolated by hplc, are still unidentified.

NMR

The proton magnetic resonance (pmr) data obtained are shown in tables 1 and 2, together with available literature values. Most of the limited pmr data

[14]VG-Organic Ltd., Altricham (Ches.), England.
[15]DuPont, Wilmington, Delaware.
[16]VF-Data Systems Ltd., Altrincham (Ches.), England.

TABLE 1. Chemical shifts of angular methyl protons in soyasapogenols.

Soyasapogenol	Solvent	Chemical shift (δ)→H_0						
B	$CDCl_3$	1.17	1.10	1.03	0.93	0.89	0.89	0.86
C	"	1.24	1.11	0.98	0.94	0.94	0.91	0.88
E	"	1.23	1.23 (1.21)[a]	0.98	0.98	0.92	0.89	0.84
E[b]	"	1.23	1.21	0.98	0.98	0.92	0.89	0.86
A	Py-d_5	1.50	1.36	1.23	1.21	1.18	1.00	0.93
B[b]	"	1.56	1.28	1.28	1.23	1.08	0.99	?
E	"	1.50	1.24	1.13	0.92	0.92	0.92	0.84

[a]Minor peak.
[b]Reference 8.

on soyasapogenols in the literature are on derivatives. We did not employ derivatization to avoid any possibility of incomplete conversion.

As is the case with steroids, the sharp angular methyl resonances can be clearly distinguished from the complex overlapping mass of aliphatic methylene and methyne resonances, which provide no structural information at 90 MHz, though it is possible that some assignments could be made at the much higher magnetic fields now becoming available. Most of the relatively small number of protons located in non-aliphatic electronic environments can be straight-forwardly assigned on the basis of empirical shielding correlations.

TABLE 2. Methyne and methylene PMR data for soyasapogenols.

Soyasapogenol	Solvents	Site of chemical shift (structure, splitting)[a]					
		12	21	22	24a	24b	3
B	$CDCl_3$	5.24(t,3)	3.41(m)		4.20(d,11)	3.41(m)	
C	"	5.29(t,4)	5.22(s)		4.20(d,11)	3.40(m)	
E	"	5.26(t,3)			4.18(d,11)	3.38(m)	
E[b]	"	5.29(t,3)			4.33(br)	3.36(br)	
A	Py-d_5	5.31(t,3)	3.69(m)		4.45(d,11)	3.69(m)	
B[b]	"	?	3.73(br,m)		?	3.73(br,m)	
E	"	5.25(m)			4.45(d,11)	3.64(d,11)	3.58(br)

[a]δ values; $|J|$ in Hz.
[b]Reference 8.

The close agreement between the published data for soyasapogenol E and those obtained in our work confirms its identification. The difference in position and structure of the H24a signal presumably arises from a change in the conformation of the methylol group as a result of competition between intra- and intermolecular hydrogen bonds (9), which would be expected at higher sample concentrations. The 11 Hz doublet in the vicinity of 4.2 ppm in $CDCl_3$ is a distinctive feature of all of our soyasapogenol spectra. In no case was sufficient sample available for a study of the effect of concentration on this signal.

As is the case for steroid spectra, pyridine produces a greater dispersion of the methyl resonances, which often aids in the resolution of overlapping signals. Variation of the sample temperature also facilitates resolution. Proton counting by integration often generates ambiguous data because of overlapping ring proton signals.

Table 1 shows that the presence of several methyl groups was confirmed for each of the four sapogenols we have examined. Katsui *et al.* (8) have reported only six for soyasapogenol B. The regularities evident in the methyl signal positions suggest that, when more data are available, it may be possible to assign some of these signals to specific sites and then to derive site- and group-specific shielding factors, which could then be used to locate known substituents on ring systems of this type. This technique has proven very useful with steroid spectra.

Mass Spectra

All four soyasapogenols give well-defined EI mass spectra, as shown in figs. 3–6. Exact mass measurement of the molecular ions confirms their elemental compositions, as shown in table 3. The underivatized compounds all yield molecular ions, although their abundance decreases markedly when one or more hydroxyl groups are present on Ring E.

A structurally diagnostic reverse Diels-Alder reaction is the source of the base peak in the mass spectra of all four sapogenols. Fig. 7 shows the mechanism of

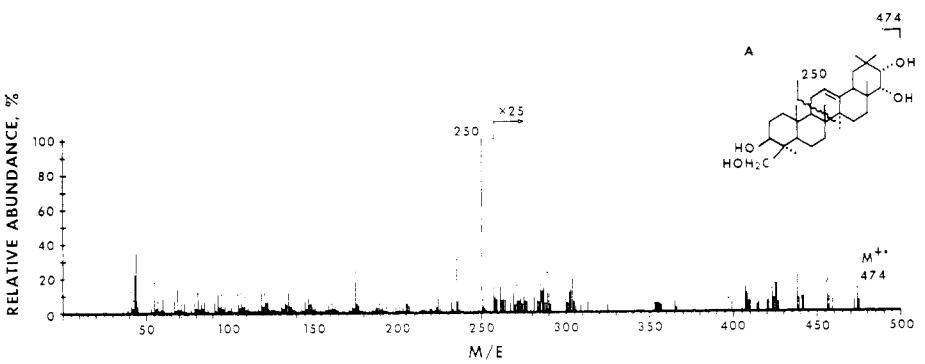

FIG. 3. Mass Spectrum of Soyasapogenol A.

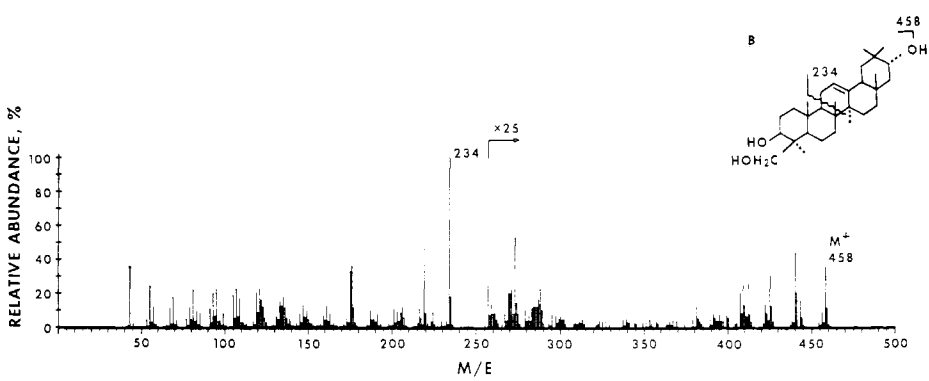

FIG. 4. Mass Spectrum of Soyasapogenol B.

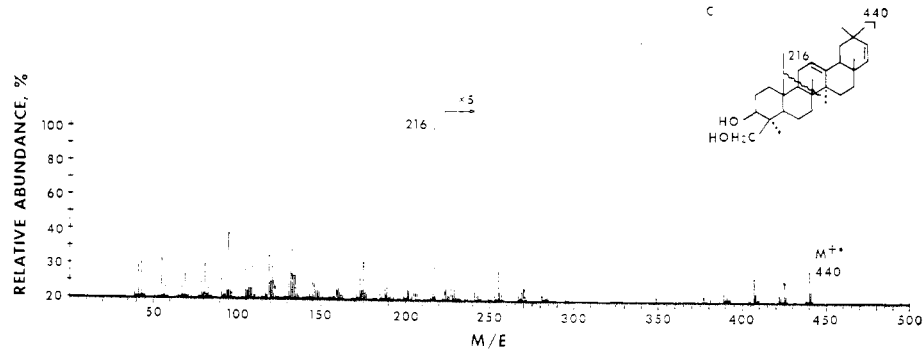

FIG. 5. Mass Spectrum of Soyasapogenol C.

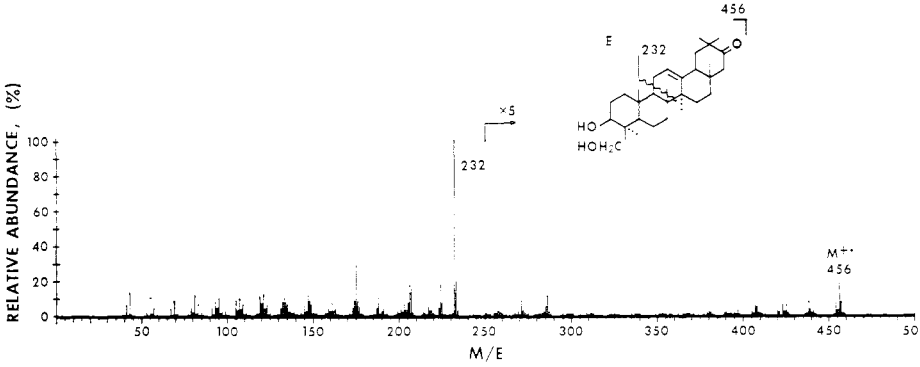

FIG. 6. Mass Spectrum of Soyasapogenol E.

TABLE 3. Accurate mass measurements of soyasapogenols.

Soyasapogenol	Empirical formula	Calculated mass	Measured mass	Error, ppm
A	$C_{36}H_{50}O_4$	474.3708	474.3746	+7.8
B	$C_{30}H_{50}O_3$	458.3759	458.3763	+0.7
C	$C_{30}H_{48}O_2$	440.3654	440.3642	−2.7
E	$C_{30}H_{48}O_3$	456.3603	456.3602	−0.2

this reaction, which is a characteristic feature in the mass spectra of Δ^{12}-unsaturated pentacyclic triterpenes (10). The occurrence of **a** has been noted previously in the mass spectra of the soyasapogenols and their derivatives (8, 11). Ion **a** is always observed at mass (M−224) for the soyasapogenols. The mass of the (M−224) fragment, together with the elemental composition derived from accurate mass measurement of the molecular ion, therefore defines the heteroatom content of the substituents on Rings C, D, and E. Subsequent loss of methyl from **a** always occurs and gives a peak 15 mass units lower (ion **5** in fig. 7) (10). For soyasapogenols A and B additional ions derived by loss of H_2O from **a** also occurs, giving rise to peaks at *m/e* 232 in the spectrum of soyasapogenol A (fig. 3) and at *m/e* 216 in the spectrum of soyasapogenol B (fig. 4). These peaks provide partial information on the identity of functional groups present on Ring E.

The complementary product of the reverse Diels-Alder reaction, **c** in fig. 7, is

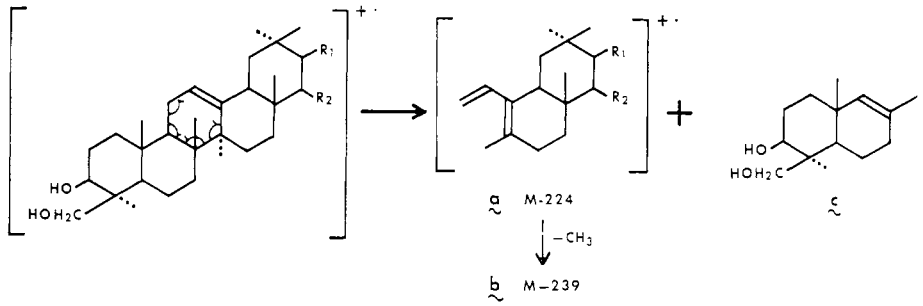

FIG. 7. Mechanism of Reverse Diels—Alder Reaction in the Mass Spectra of the Soyasapogenols.

also present, although its abundance is considerably less than that of **a**. The presence of **c** in the mass spectra of these compounds is in contrast to the behavior of structurally similar Δ^{12}-pentacyclic triterpenes (10), for which **c** is not observed at all. The subsequent decomposition of **c** by loss of H_2O and H_2O plus $\cdot CH_2OH$ gives rise to additional ions at *m/e* 206 ($C_{14}H_{22}O$) and 175 ($C_{13}H_{19}$), as indicated in fig. 8. The presence of ions **c-e** appears to be characteristic of the particular A and B ring substitution of the soyasapogenols, and may serve to differentiate them from other pentacyclic triterpenes having closely related molecular structures.

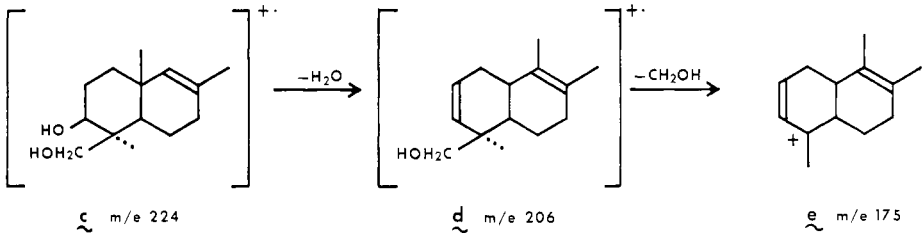

FIG. 8. Origin of *m/e* 206 and 175 Ions in Soyasapogenol Mass Spectra.

ACKNOWLEDGMENTS

This work was supported in part by a grant to A. Bondi by the U.S. Israel Binational Science Foundation. The authors thank Richard Young for constructing the event marker, Sue Witt for technical assistance with nmr, and Dale R. Black and Steve Poling for performing the accurate mass measurements. A gift of soyasapogenol D from D. Arigoni is gratefully acknowledged.

Received 12 February 1979.

LITERATURE CITED

1. P. R. Cheeke, *Nutr. Reports Intern.*, **13**, 315 (1976).
2. Y. Birk, "Toxic Constituents of Plant Food Stuffs," I. E. Liener, ed., Academic Press, New York, 1969, p. 169.
3. A. Bondi, Y. Birk, and B. Gestetner, "Chemistry and Biochemistry of Herbage," G. W. Butler and R. W. Bailey, eds., Academic Press, New York, 1973, p. 511.
4. I. Peri, U. Mor, E. Heftmann, A. Bondi, and Y. Tencer, *Phytochemistry*, in press.
5. I. R. HUNTER, M. K. Walden, J. R. Wagner, and E. Heftmann, *J. Chromatogr.*, **119**, 223 (1976).
6. I. R. Hunter, M. K. Walden, and E. Heftmann, *J. Chromatogr.*, **153**, 57 (1978).
7. E. Heftmann, G. A. Saunders, and W. F. Haddon, *J. Chromatogr.*, **156**, 71 (1978).
8. N. Katsui, H. Matsue, T. Hirata, and T. Masamune, *Bull. Chem. Soc. Japan*, **45**, 223 (1972).
9. A. Gaudemer, J. Polonsky, and E. Wenkert, *Bull. Soc. Chim. France*, No. 2, 407 (1964).
10. H. Budzikiewicz, J. M. Wilson, and C. Djerassi, *J. Amer. Chem. Soc.*, **85**, 3688 (1963).
11. B. Berrang, K. H. Davis, Jr., M. E. Wall, C. H. Hanson, and M. E. Pedersen, *Phytochemistry*, **13**, 2253 (1974).