CHAPTER-6

STRUCTURE OF GLYCOSIDE-AR_4 AND COMP. B

Isolation of glycoside-AR_4 and comp.B. are described in chapter-2. The present chapter deals with their structure determination. Due to paucity of the samples, they could not be studied in detail. However, spectral studied were carried out to arrive at their tentative structures.

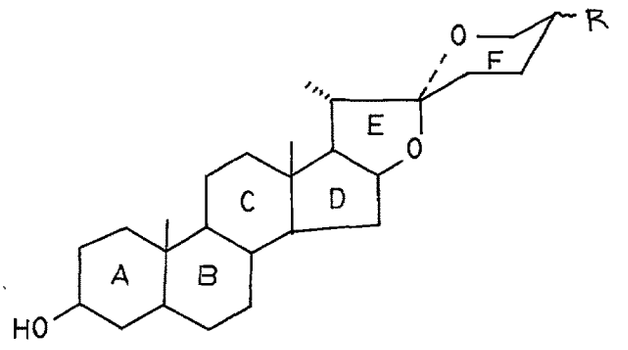
STRUCTURE OF GLYCOSIDE-AR_4

Glycoside-AR_4 was obtained as a white crystalline powder. It was insoluble in most organic solvents, soluble only in hot alcohol. When spotted on TLC, it was found to have the same R_f value as that of shatavarin-IV.

APPLICATION OF IR SPECTROSCOPY

In case of steroidal saponins, analysis of IR spectrum helps in determination of the aglycone. Examination of IR spectrum in the range 850-1000 cm^{-1} gives information about the sapogenin.

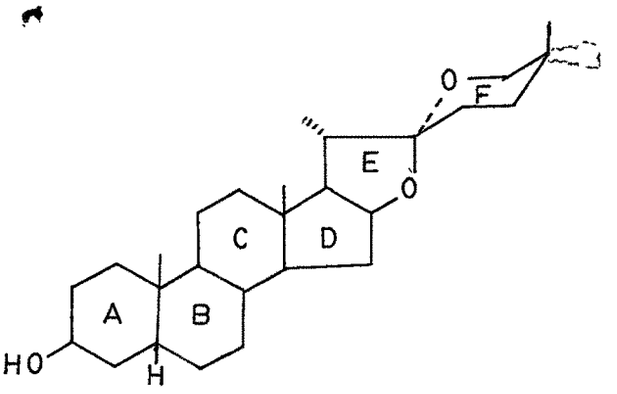
The study of IR spectrum of steroidal sapogenins, sapogenin acetates and saponins was systematically carried out by Wall *et al.*[1]. According to Wall, for a steroidal sapogenin, four absorption bands in the region of 850-1000 cm^{-1} are characteristic of F ring of the genin (I).



I

In case of spirostanol saponins with F ring the four bands appear at 850, 900, 922 and 990 cm^{-1} (11.75, 11.1, 10.85 and 10.14 μ), with 922 band absorptivity stronger than 900 cm^{-1} band indicate a normal sapogenin[2]. Bands near 866, 900, 922 and 982 cm^{-1} (11.55, 11.1, 10.85 and 10.18 μ), with 900 band absorptivity stronger than 922 band indicate an iso sapogenin[2]. The same relationship is observed in crude saponin mixture and sapogenin and saponin acetates[3].

Sarsasapogenin (spirostane-3β-ol, II) is the most commonly occuring normal sapogenin and can be identified easily from IR spectrum.



II

It shows four absorption bands at 852, 900, 922 and 987 cm^{-1}, with 922 band absorptivity stronger than that of 900 cm^{-1} band.

In the IR spectrum of glycoside-AR_4 (Fig.1) four bands at 850, 892, 912 and 987 cm^{-1} are observed which are characteristic of a steroidal saponin with F ring. The absorptivity of 912 cm^{-1} band is stronger than that of 892 cm^{-1} band, indicating that the aglycone in glycoside-AR_4 is a normal sapogenin. Since sarsasapogenin[4-8] is the most commonly occuring normal sapogenin and has been isolated from most asparagus species, it can be assumed that the aglycone in glycoside-AR_4 must be sarsasapogenin.

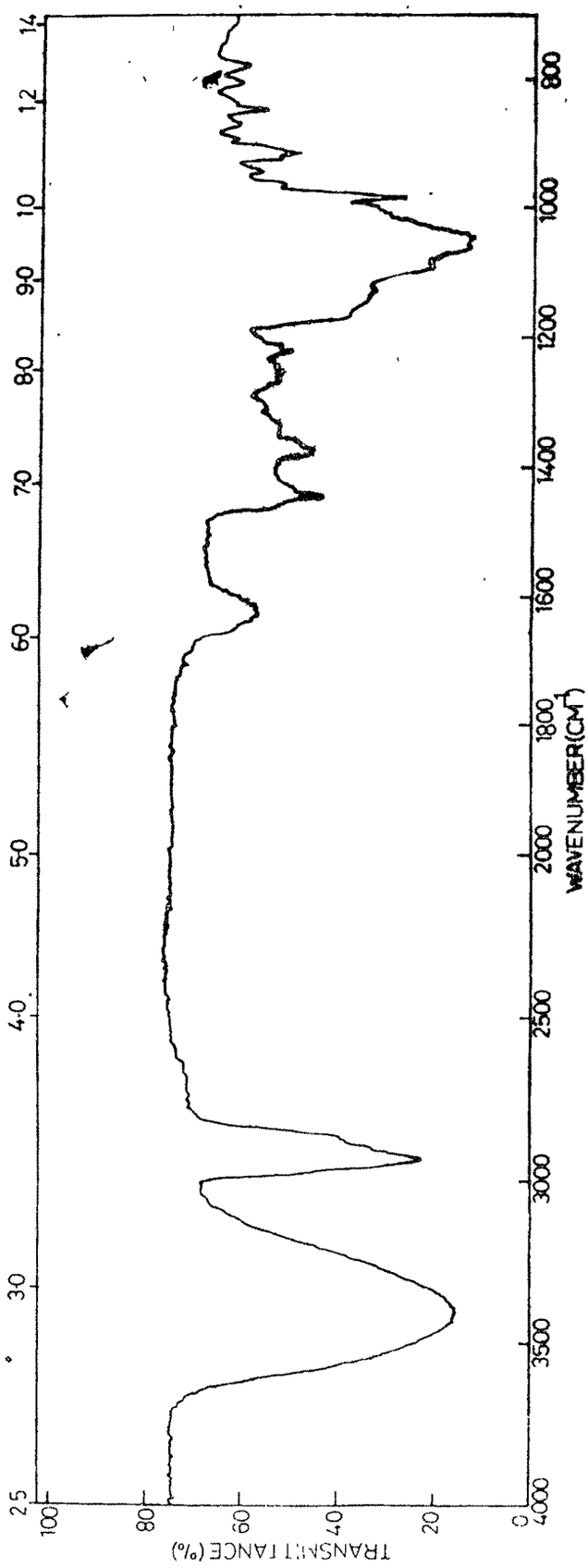


FIG. 1 : IR SPECTRUM OF GLYCOSIDE-AR_4

Comparison of the IR spectrum of glycoside-AR_4 with that of shatavarin-IV showed that they are similar. The other bands in the IR spectrum of glycoside-AR_4 are at 3400 for -OH, 2930 for saturated carbons and 1000-1200 cm^{-1} for C-O-C absorptions.

In the FABMS of glycoside-AR_4 (Fig.3) pseudomolecular ion is observed at m/z 857 with 12.5% intensity, which can be assigned to $[M^+ + H]$. Molecular weight is thus 856. All the fragments are derived from the molecular ion $[M^+ + H]$. Glycoside-AR_4 shows a signal at m/z 726 with 2.3% intensity, due to loss of 131 mass units from $[M^+ + H]$, which is the result of the loss of the terminal sugar unit from the molecule by a protonation mechanism analogous to acidic solvolysis. In this proton-induced cleavage of the glycosidic oxygen and charge localization can induce a characteristic electron shift from the ring oxygen resulting in the formation of a glycosyl ion. This type of ion is almost always found in the FD-MS of sugars and their derivatives.

This loss of 131 mass units corresponds to the loss of an unknown sugar having formula $C_5H_8O_4$ i.e. it is a pentose with molecular weight 132.

The compound shows a signal at m/z 711 with 2.1% intensity, which arises by the loss of 146 mass units from the molecular ion $[M^+ + H]$, which corresponds to loss of terminal sugar unit. This loss of 146 mass units corresponds to the loss of $C_6H_{10}O_4$ unit, which by comparison with the FABMS of shatavarin-IV must be a loss of rhamnose. The loss of pentose and loss of rhamnose from the molecular ion indicate that pentose and rhamnose form the two terminal sugars. Thus, $[M^+ + H$ - pentose$]$ and $[M^+ + H$ - rhamnose$]$ are the two ions which are consistent with the branched sugar sequence. The signal at m/z 579 arises due to loss of both pentose and rhamnose from the molecule. i.e. $[M^+ + H$ - (pentose+rhamnose)$]$. The ion at m/z 417 is formed directly from 579 by loss of 162 mass units, which can be accounted for the loss of glucose unit. i.e. it is observed due to loss of pentosyl-rhamnosyl-glucoside from $[M^+ + H]$ and this reveals the aglycone Mol. wt. 416 , sarsasapogenin. Formation of 417 from 579 implies that a glucose is connected at C-3 of sarsasapogenin. Water elimination from the aglycone (m/z 417) produces a peak at m/z 399.

Signals observed at m/z 399, 397, 285, 255 and 139 are characteristic of the sapogenin.[9,10] The formation of fragment ions can be explained as in Fig.2. The formation

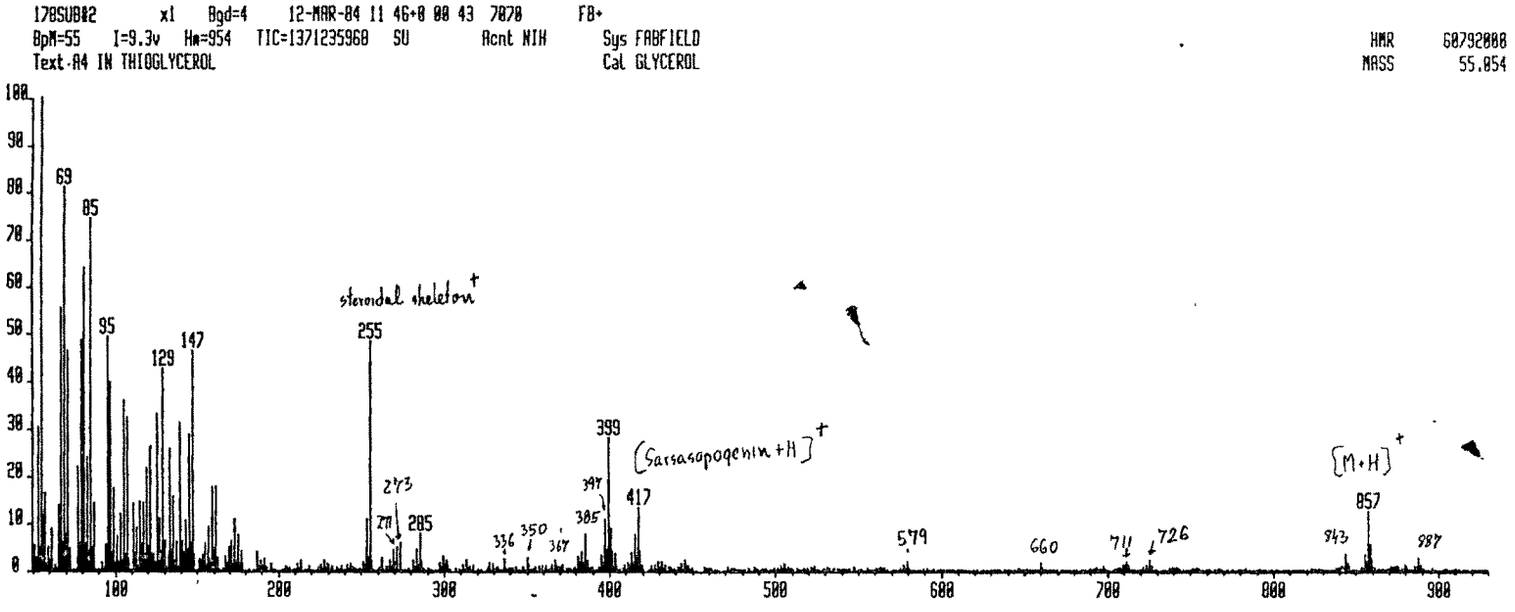


FIG.3 : FABMS OF GLYCOSIDE-AR$_4$

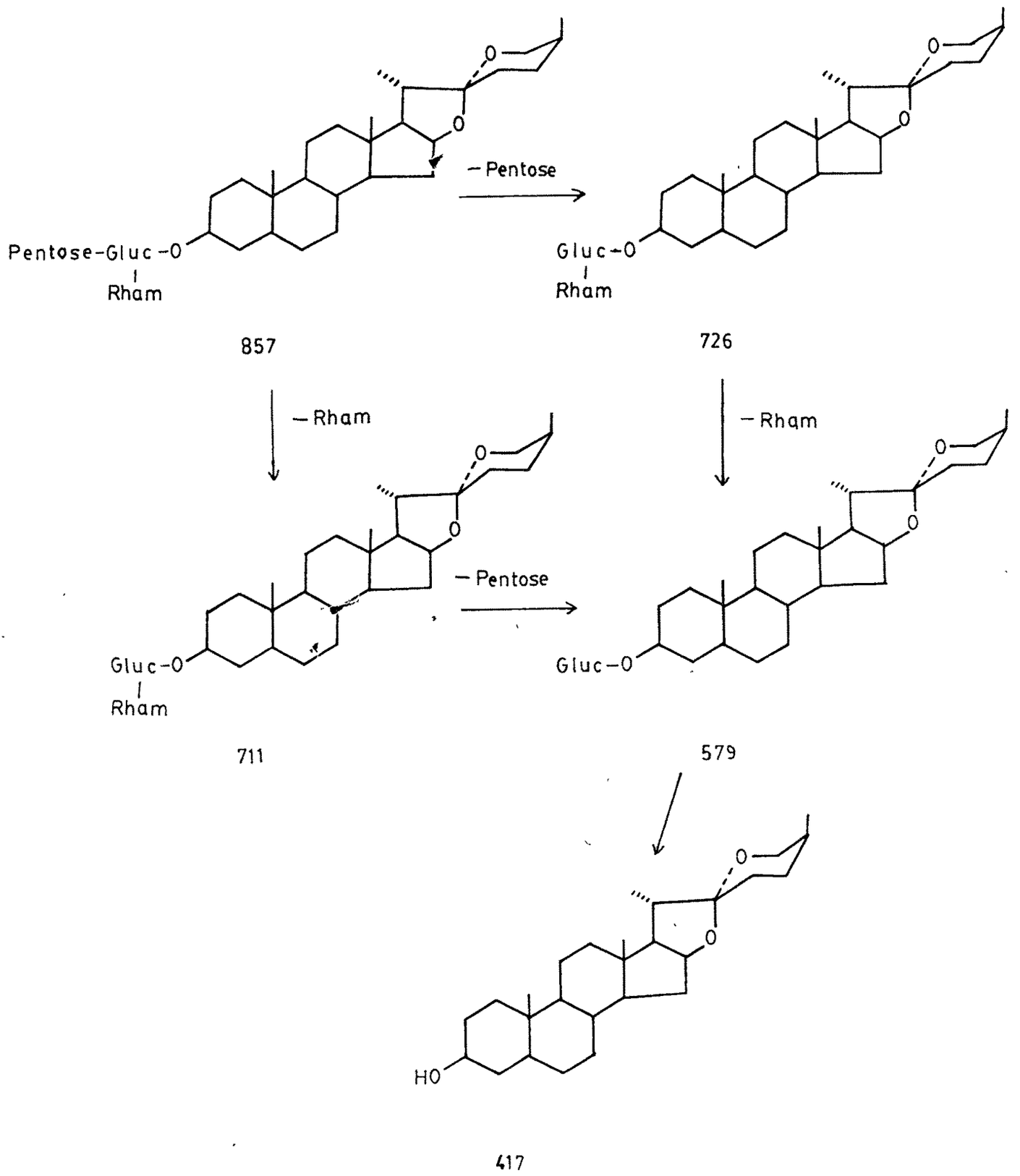


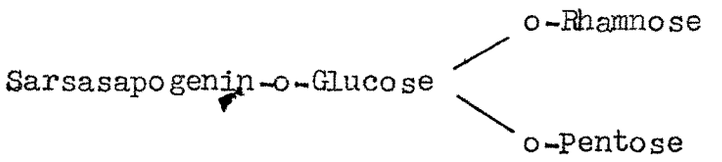
FIG.4 : FORMATION OF IONS IN THE FAB^{+}MS OF GLYCOSIDE-AR$_4$.

of important ions is shown in Fig.4. Presence of glucose is confirmed by ions appearing at m/z 163, 145 and 127, that of rhamnose is confirmed by the ions at m/z 147 and 129. The ion at m/z 132 corresponds to the pentose. Only one mole of glucose, rhamnose and pentose are removed from the molecule, which proves that they are in the ratio of 1:1:1.

Ions		m/z
$[M^+ + H]$		857
$[M^+ + H\text{-pentose}]$	= 857-131	726
$[M^+ + H\text{-rhamnose}]$	= 857-146	711
$[M^+ + H\text{-(pentose+rhamnose)}]$	= 857-278	579
$[M^+ + H\text{-(pentose+rhamnose+glucose)}]$	= 857-440	417
$\text{Sarsasapogenin}^+ + H$	= 416+1	417

FIG.2 : Fragments formed in the FAB^+MS of glycoside-AR_4.

The molecular weight 856 corresponds to the molecular formula $C_{44}H_{72}O_{16}$ which is consistent with its tentative structure (II), assigned to it on the basis of its IR spectrum and FABMS.



(II)

The FABMS can be successfully used in case of sugars with different chemical structures. e.g. rhamnose and glucose can be distinguished but glucose and galactose which have same molecular weight cannot be distinguished. Also it cannot give the information about the positions of linkages and the stereochemistry of linkages.

Since it was now known that glycoside-AR_4 contains glucose, rhamnose and an unknown pentose as its sugars, it was hydrolysed to find out the structure of the pentose and to confirm presence of glucose and rhamnose.

HYDROLYSIS OF GLYCOSIDE-AR_4

Glycoside-AR_4 was hydrolysed by refluxing it with 2N sulfuric acid in dioxan.[11] The aglycone and the sugars were separated. The aglycone was identified as sarsasapogenin (by mixed TLC).

The aqueous part was neutralised by passing through an ion exchange resin.[12] The sugars on paper chromatogram showed three spots with R_f values 0.20, 0.25 and 0.36. When

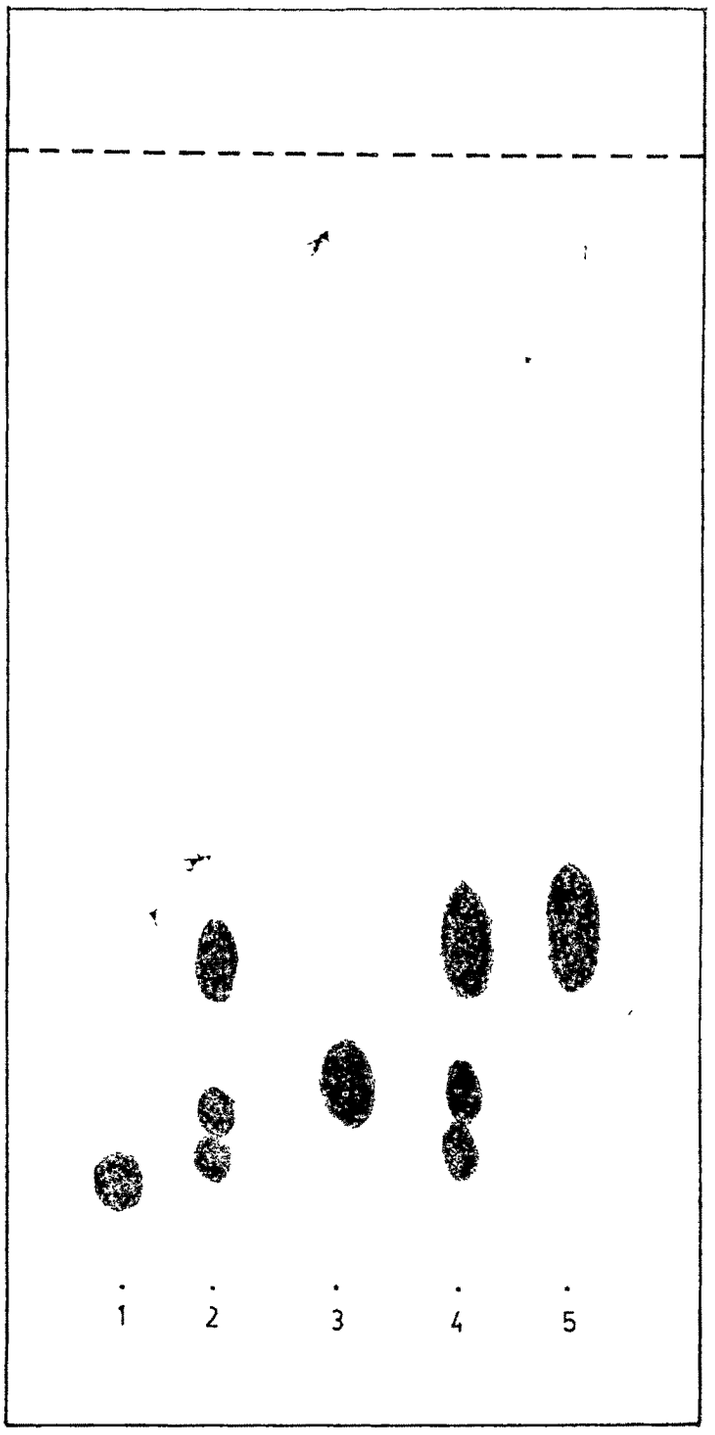


FIG.5 : PAPER CHROMATOGRAM OF SUGARS OF GLYCOSIDE-AR_4

SOLVENT SYSTEM : n-BuOH-HOAc-H_2O (4:1:5)(UPPER PHASE)

SPRAY REAGENT : ANILINE HYDROGEN PHTHALATE

TIME : 4 HOURS.

SPOTS : 1) GLUCOSE 2) SUGARS OF GLYCOSIDE-AR_4
3) ARABINOSE 4) STANDARD MIX. OF SUGARS
5) RHAMNOSE

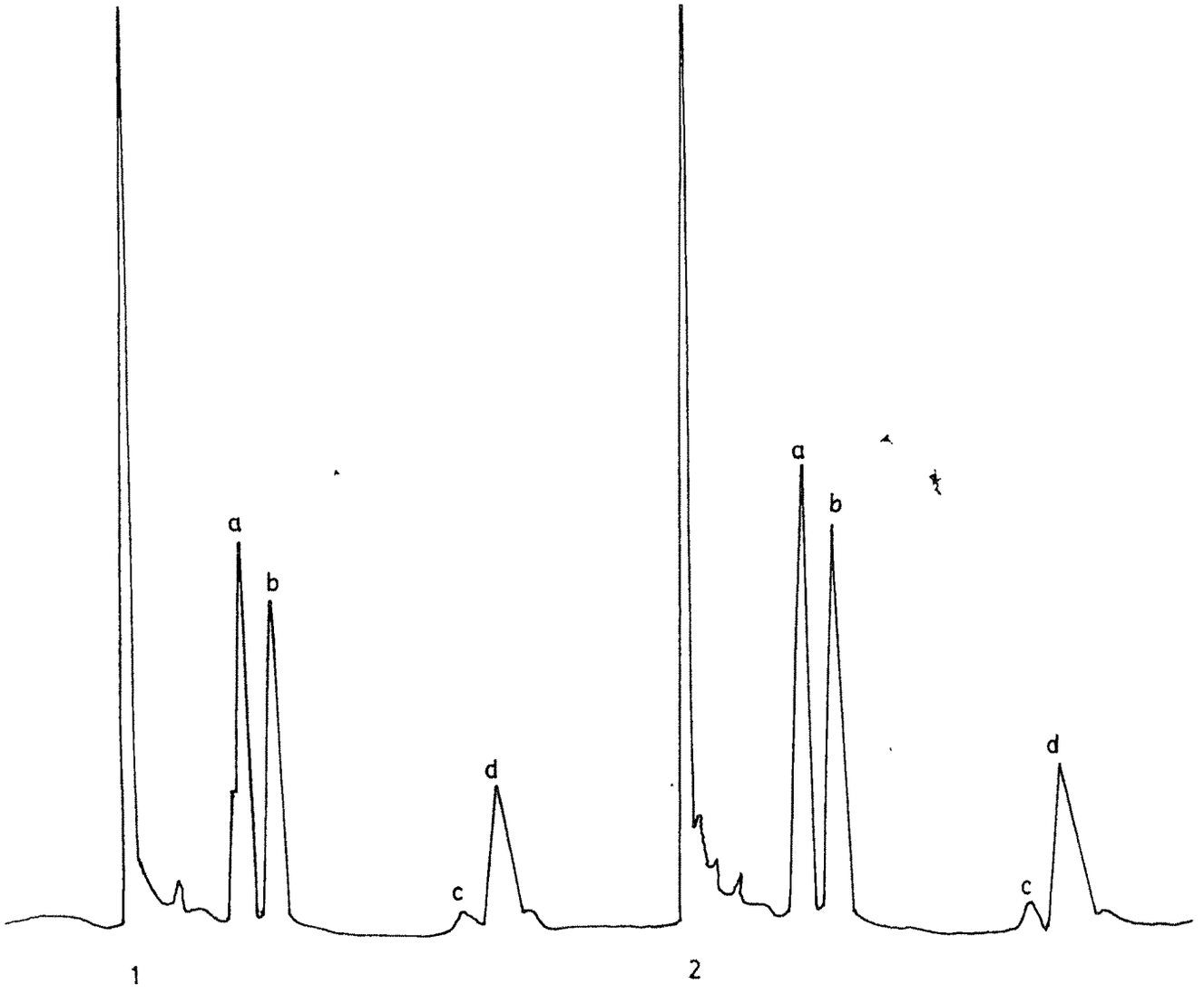


FIG.6 : GLC OF SUGARS OF GLYCOSIDE-AR_4

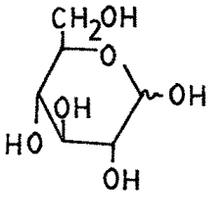
COLUMN : 10% DCQF-1 ON CHROMOSORB W HP 60/80

NITROGEN FLOW : 60 ML/MIN. TEMP.: 210°

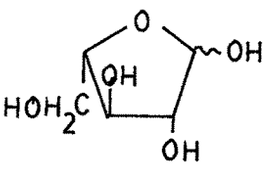
1) SUGARS OF GLYCOSIDE-AR_4. 2) MIXTURE OF GLUCOSE+RHAMNOSE+ARABINOSE (1:1:1)

a) RHAMNOSE b) ARABINOSE c, d) GLUCOSE.

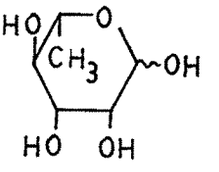
samples of standard sugars were spotted, the three correspond to D-glucose*(III), L-arabinose*(IV) and L-rhamnose*(V).(Fig.5). All the three spots appeared to be of equal intensity indicating that they are possibly in the ratio of 1:1:1. To confirm this, the carbohydrate part of glycoside-AR_4 hydrolysate was converted into the acetate derivatives by treatment with acetic anhydride and pyridine. GLC of this was recorded on DCQF-1 column.[13] The GLC (Fig.6) showed four peaks. Peak 1 with smallest retention time was identified as due to authentic L-rhamnose, the second peak was identified as due to L-arabinose and last two peaks are from the acetate derivative of D-glucose. Calculation of areas under the peaks showed proportion of D-glucose, L-arabinose and L-rhamnose to be 1:1:1. For comparison purpose authentic D-glucose+, L-rhamnose** and L-arabinose° in proportion of 1:1:1 were acetylated and analyzed by GLC (Fig.6).



III



IV



V

---

* Naturally occuring glucose, rhamnose and arabinose occure in D, L and L form respectively.

+ BDH, GR; ** SD fine chemicals, ° Loba Chem. LR.

Due to paucity of the sample, ^{13}C-NMR could not be recorded hence positions of linkages of sugars could not be obtained.

^{1}H-NMR spectrum of glycoside-AR_4 was recorded in C_5D_5N. Due to broad anomeric signals, stereochemistry of linkages could not be assigned.

From the above results, it was confirmed that glucose, arabinose and rhamnose in the proportion of 1:1:1 are the sugars connected to sarsasapogenin.

STRUCTURE OF COMP.B

Comp.B was obtained from both A. racemosus and shatavari. It was deposited as a white powder from methanol solution, insoluble in almost all organic solvents like chloroform, benzene, petroleum ether, solvent ether, etc. also insoluble in water, sparingly soluble in cold methanol, soluble in hot methanol and ethanol. It was crystallized from a mixture of methanol+ethanol (1:1). It developed a pink coloured spot when sprayed with vanillin/ phosphoric acid.

Since a number of steroidal saponins were isolated from shatavari, it was assumed that comp.B also may be a

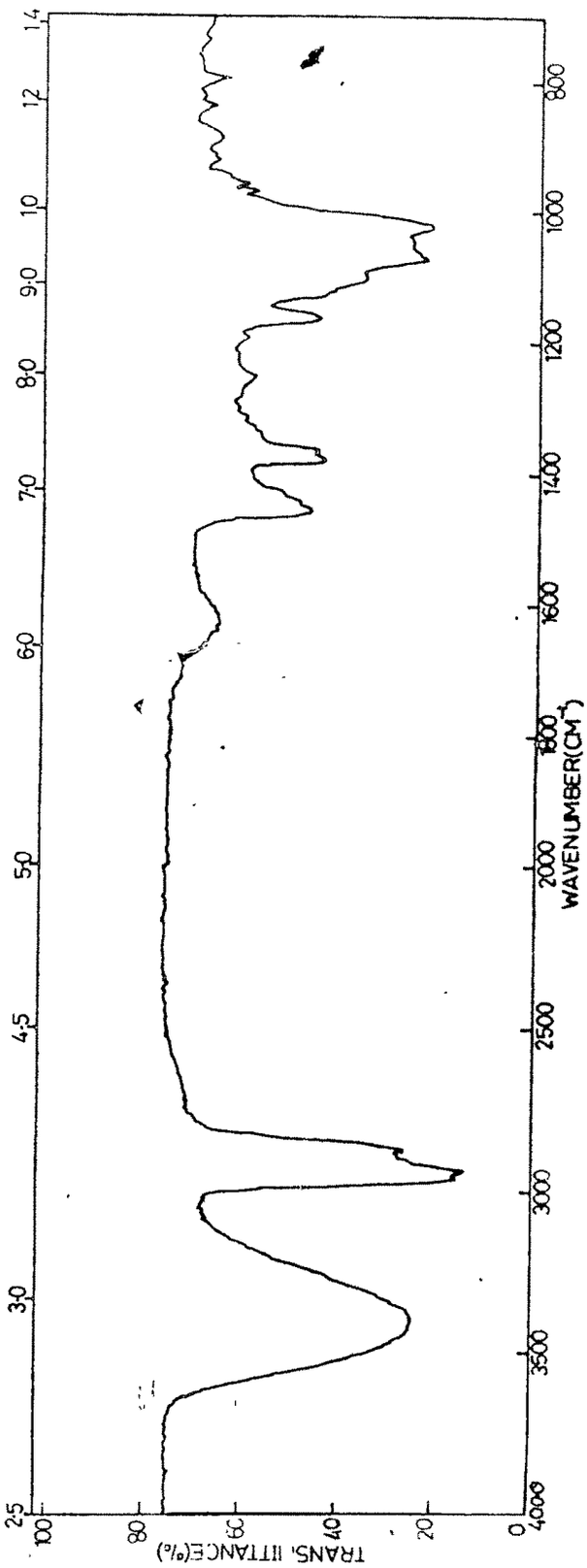


FIG.7 : IR SPECTRUM OF COMP. B

steroidal saponin. As already mentioned earlier, preliminary examination of an IR spectrum gives an indication whether the compound is a steroidal saponin or not. IR spectrum of comp.B (Fig.7) showed that it is not a steroidal saponin as the characteristic absorption band pattern (bands at 850, 900, 922, 987 cm^{-1}) is missing in the spectrum. Bands are observed at 3400, 2960, 2940, 2870, 1455, 1370, 1040-1160 cm^{-1} and C-H absorption bands at 834 and 797 cm^{-1} for trisubstituted olefin.[14]

In the FABMS of comp.B. (Fig.8), a peak is observed at m/z 579, which can be assigned to $[M^+ + H]$, which implies that molecular weight of the compound is 578.

From the biogenesis point of view, this will correspond to any saturated steroidal alcohol which is glycosylated. Presence of glucose is indicated by the ions at m/z 163, 145 and 127, but $[M^+ + H$ - glucose$]$ ion is not observed in the spectrum.

From the FABMS, it was evident that comp.B. is not a single compound, but is a mixture of two compounds, which do not separate on TLC or separate by crystallization. In the mass spectrum, the difference can be observed clearly. Formation of some of the fragments can be explained by considering m/z 579 as the molecular ion, but other

fragments can be explained only by considering a double bond in the molecule. In the FABMS (Fig.8) important ions are observed at m/z 579, 551, 397, 395, 383, 369, 325, 297, 259, 254, 217 and 188. The fragments at m/z 579, 551, 383, 369, 325, 297, 259, 217 and 188 can be explained from the saturated compound whereas the ions at m/z 397, 395, 383 and 254 originate from the unsaturated compound. These ions are characteristic of a steroidal alcohol with a saturated side chain at C-17.[15] Considering comp.B to be a mixture of two compounds, the ions can be explained as follows.

The molecular ion $[M^+ + H]$ at m/z 579 ($C_{35}H_{62}O_6$) corresponds to the saturated compound. The ion at 551 ($C_{33}H_{58}O_6$) is formed by the loss of an ethyl group from $[M^+ + H]$. The ion corresponding to the peak 383 ($C_{28}H_{47}$) arises by the loss of glucose and a water molecule with simultaneous loss of methyl group from the side chain. The ion at 369 ($C_{27}H_{45}$) is formed by the loss of glucose, a water molecule and an ethyl group from the side chain, subsequent degradation of the side chain forms the ions at 325 ($C_{24}H_{37}$) and 297 ($C_{22}H_{33}$). The ions at 259 ($C_{18}H_{27}O$), 217 ($C_{16}H_{25}$) and 188 ($C_{14}H_{20}$) are formed by fragmentation of ring D.

The ions originating from the unsaturated compound can be explained as follows. The ions indicate that the double

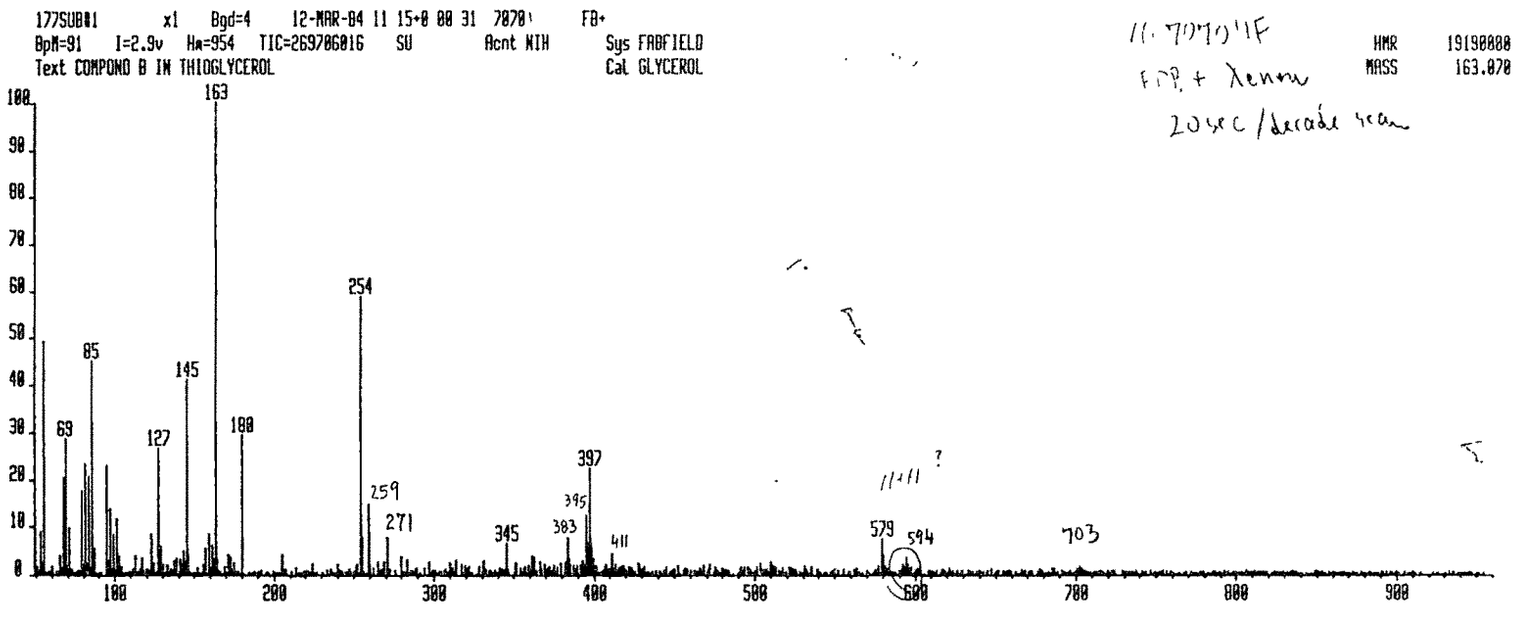


FIG.8 : FABMS OF COMP.B.

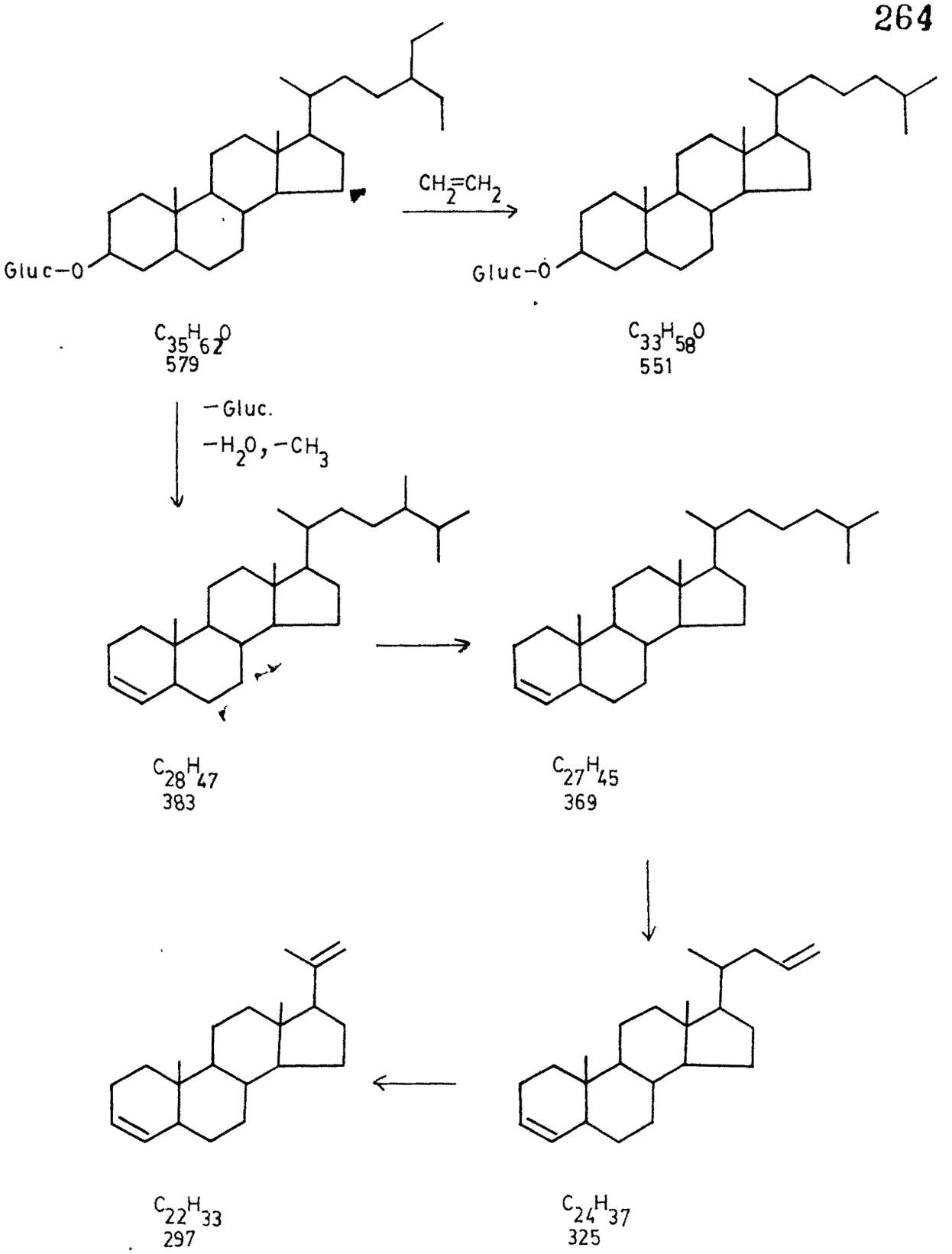
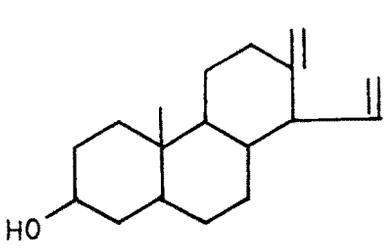
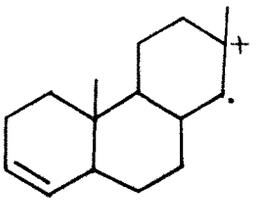


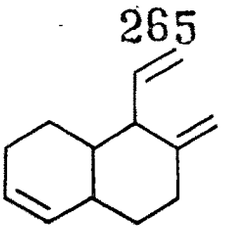
FIG.9 : FORMATION OF IONS IN THE FAB$^+$MS OF COMP. B.



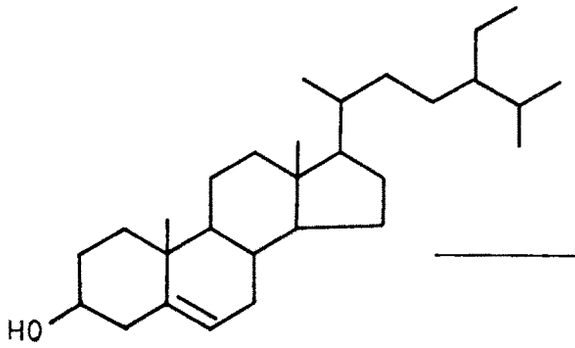
$C_{18}H_{27}O$
259



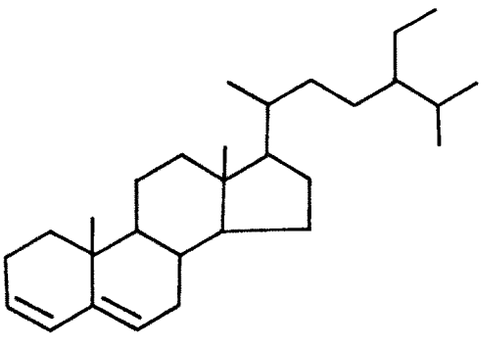
$C_{16}H_{25}$
217



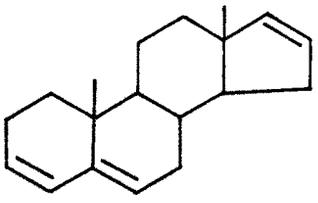
$C_{14}H_{20}$
188



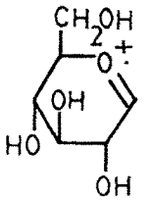
$C_{35}H_{60}O$
576



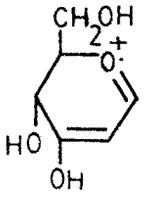
$C_{29}H_{49}$
397



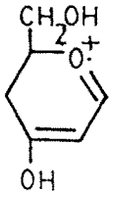
$C_{19}H_{26}$
254



163



145

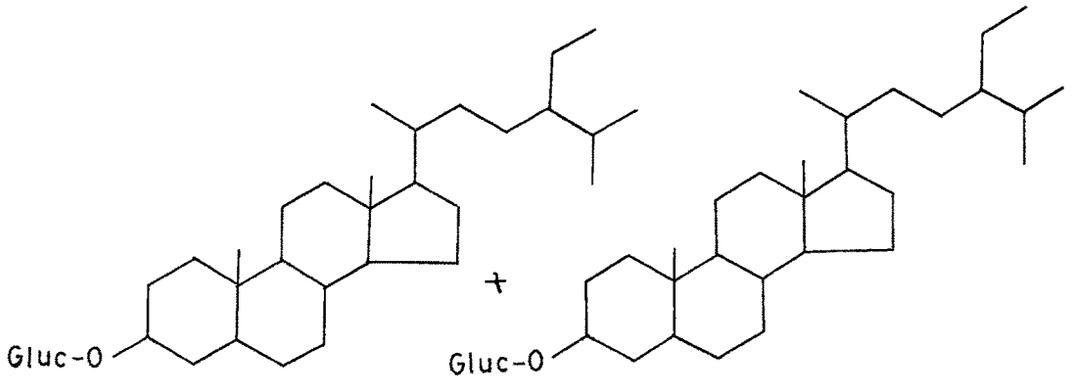


129

FIG.10 : FORMATION OF IONS IN THE FAB$^+$MS OF COMP.B.

bond is not in the side chain but is in the rings. Since Δ^5 alcohol (β-sitosterol) occure commonly in nature, it was assumed that the compound must be having a double bond at C_5-C_6. The calculated molecular weight of the compound is then 576. The ion at 397 ($C_{29}H_{49}$) is formed by the formation of $[M^+ + \text{gluc} - H_2O]$, which loses a methylene group to form the ion at 383. The ion at 395 is formed by the formation of a double bond. Cleavage of the side chain at C_{17} results in the formation of the ion at 254 ($C_{19}H_{26}$). Formation of fragment ions is shown in Figs.9 and 10.

From the IR spectrum and FABMS, it can be concluded that the compound may be a mixture of two compounds, a sterol glucoside and a stanol glucoside (V).



(V)

To get the aglycone, comp.B was hydrolysed.

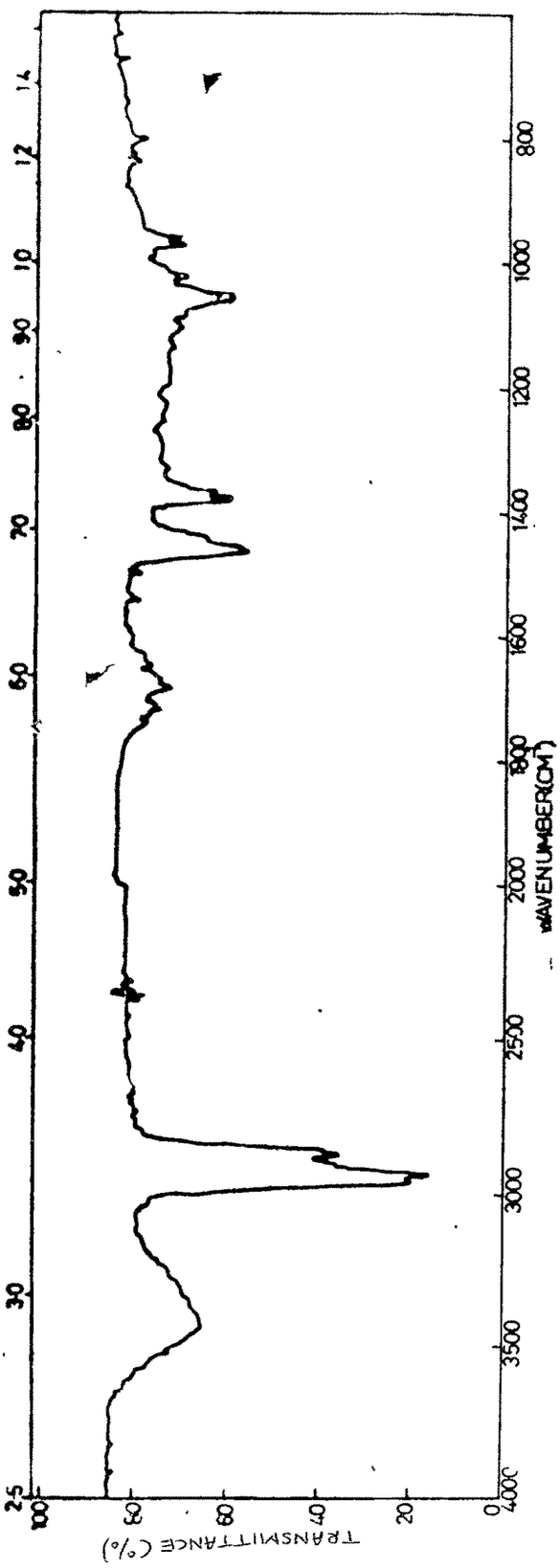


FIG.11 : IR SPECTRUM OF AGLYCONE OF COMP.B.

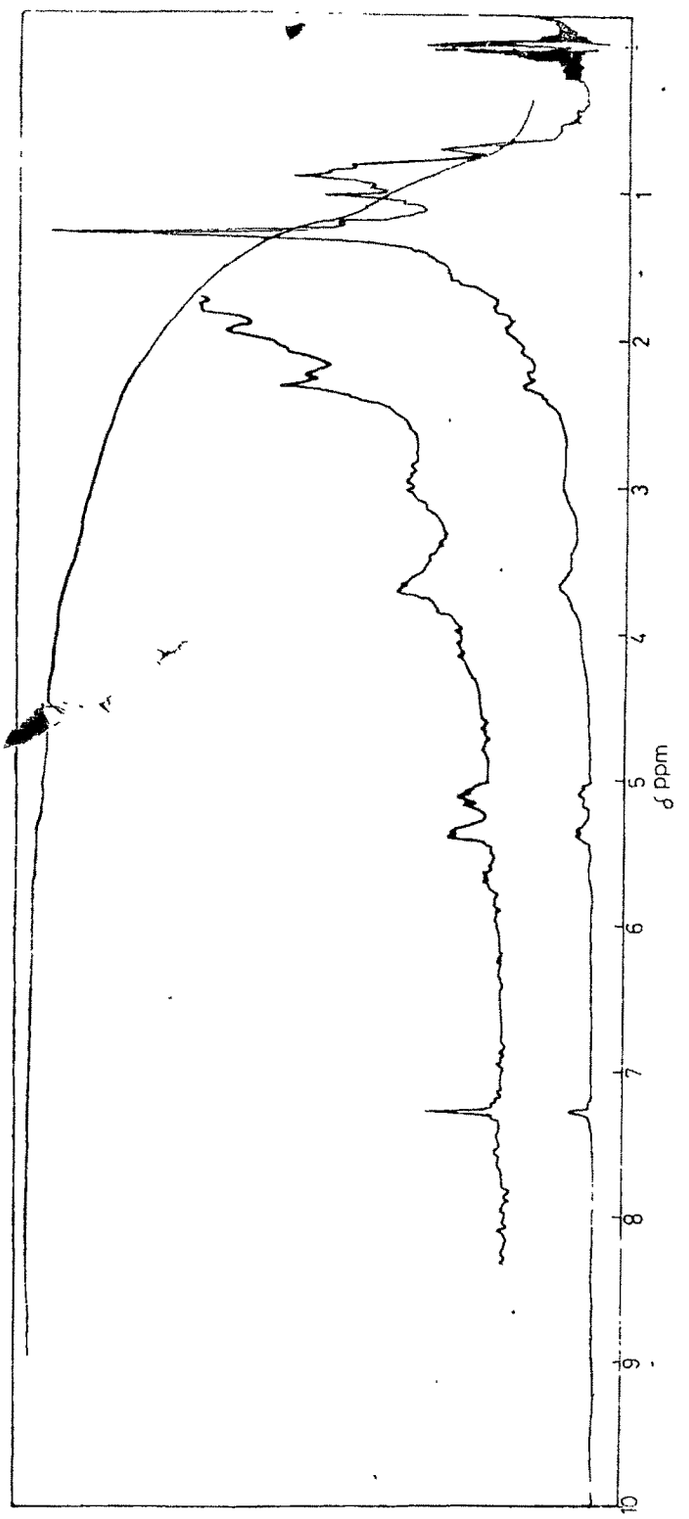


FIG.12 : ^{1}H-NMR SPECTRUM OF AGLYCONE OF COMP.B.

HYDROLYSIS OF COMP.B

Comp.B was hydrolysed by refluxing it with 2N sulfuric acid with continous stirring for 6 hours. It was cooled and the insoluble compound was filtered and taken up in chloroform. 50% of the unreacted comp.B was recovered back. The chloroform soluble aglycone was crystallized from methanol to get white powder m.p. 128-129.5°. (β-sitosterol 137-139°). The aglycone was negative to the TNM[15] test, but with Liebermann-Burchard reagent it gave stable green colour[16].

In the mass spectrum (Fig.13) two compounds can be detected which differ only by a double bond. For the molecular ion corresponding to the unsaturated compound at 414 ($C_{29}H_{50}O$) there is an ion at 416 ($C_{29}H_{52}O$) corresponding to the saturated compound. Principal peaks are observed in the range 205-245.[15] For the unsaturated compound, the peak is observed at 213 by fragmentation of ring D with the side chain. The presence of an intense peak at 119 suggested it to be Δ^5-3β-sterol, m/z 273 [M^+- side chain], 255 [M^+- side chain - H_2O], 231 [M^+- side chain -42(ring D fragment)] are also observed. For the corresponding saturated analog, the ions are observed at m/z 233 [M^+- side chain - H_2O], 275 [M^+- side chain], 215 [233 - H_2O], 257 [M^+- side chain- H_2O], 383 [M^+-CH_3-H_2O], 401 (M^+-CH_3) and 416 M^+ are observed in the spectrum.

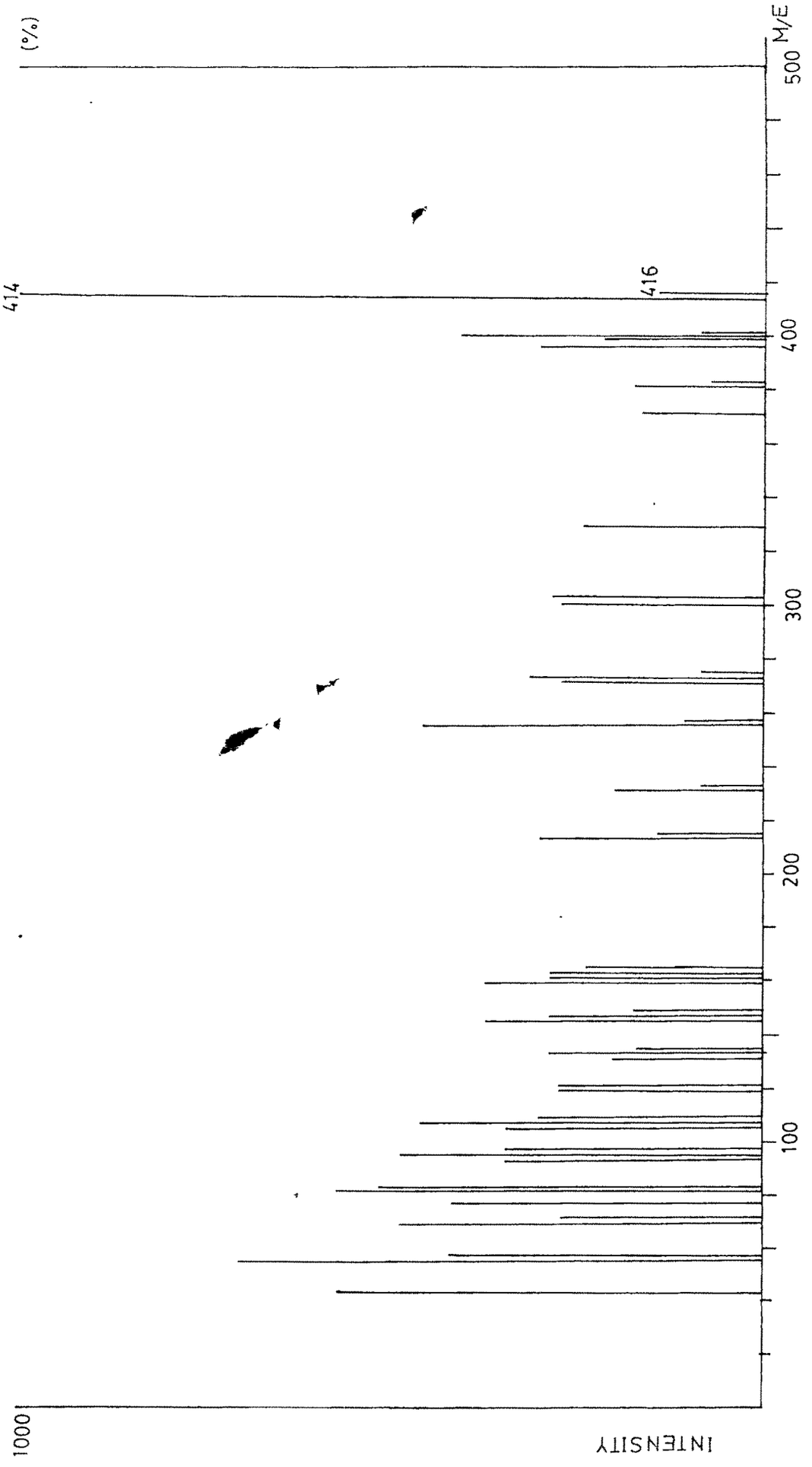


FIG.13 : MASS SPECTRUM OF AGLYCONE OF COMP.B

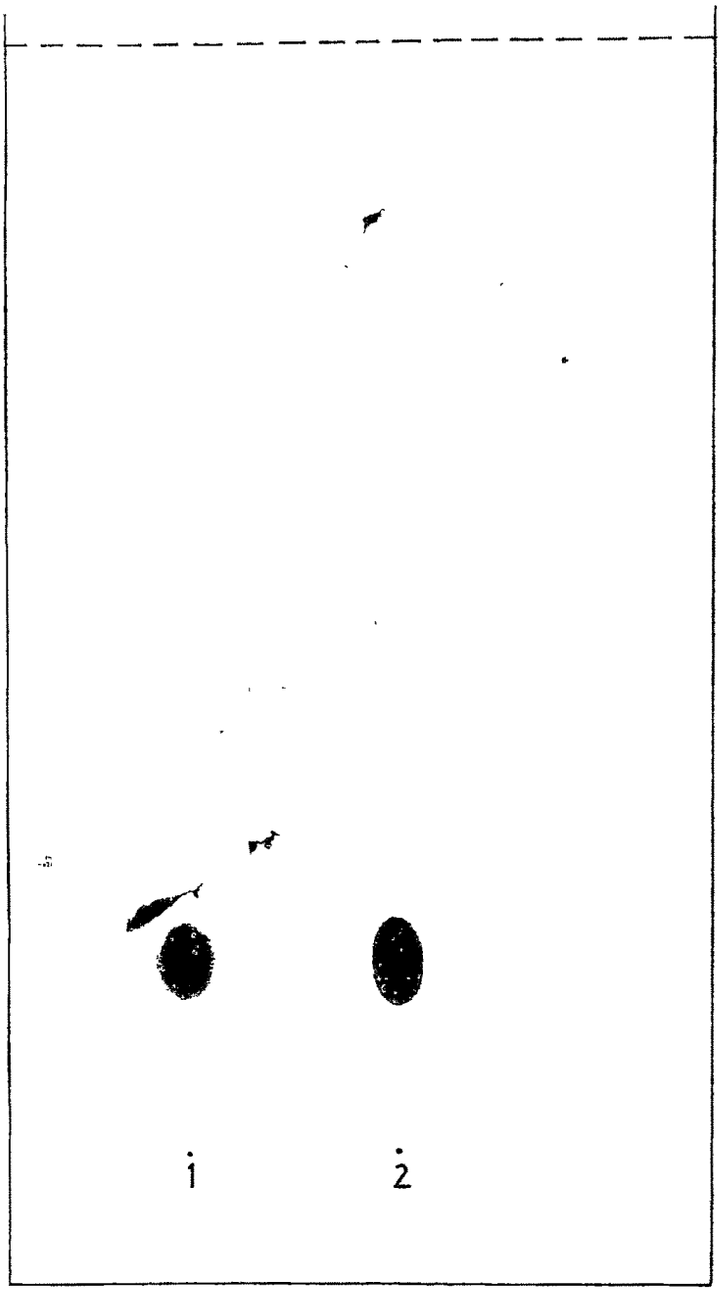


FIG. 14 : PAPER CHROMATOGRAM OF SUGARS OF COMP. B.

SOLVENT SYSTEM : n-BuOH-HOAc-H_2O (4:1:5)
(UPPER PHASE)

SPRAY REAGENT : ANILINE HYDROGEN PHTHALATE

TIME : 4 HOURS.

SPOTS : 1) GLUCOSE
2) SUGAR OF COMP. B

This clearly indicated that the aglycone is not a single compound, but a mixture of a saturated and unsaturated compound.

The aqueous part was neutralised with barium cabonate, which on paper chromatogram showed a single spot with R_f value of 0.20. Spotting alongwith standard samples of sugars showed (Fig.14) that it correspond to D-glucose.

From the above results, probable structure of comp.B was formulated as a mixture of sterol-glucoside and a stanol-glucoside.

EXPERIMENTAL

For general remarks please refer to page no. 100 and 160.

GLC were recorded on a Hewlett Packard h/p GC model 7624 A, equipped with a 1 m glass column (1/4" O.D.) of 10% $DCQF_1$ on chromosorb W HP 60/80 and operated at 210° was used. The carrier nitrogen gas flow was 60 ml/min.

Hydrolysis of glycoside-AR_4:

Glycoside-AR_4 (20 mg) was heated on a boiling waterbath with 2N sulfuric acid (7 ml) and dioxan (3 ml) for 4 hours. It was cooled to room temp. and diluted with water (10 ml) and extracted with benzene (25 ml x 3). The combined benzene extracts were washed with water (25 ml x 2) and dried with anhydrous sodium sulfate. Solvent was evaporated to get the genin, which correspond to sarsasapogenin. (TLC, IR).

Filtrate obtained after removing the aglycone was neutralized by passing through anion exchange resin** (50 ml) and filtered. The resin was washed with water (30 ml).

** Amberlite IRA-400, prewashed with 10% NaOH solution, followed by washing with water, till neutral.

Water was then removed from the combined extracts under suction to get 10 mg sugars.

Paper chromatography : This was carried out on Whatman No.1 filterpaper by upward irrigation with the organic phase of n-butanol-acetic acid-water (4:1:5) for 4 hours. Spray reagent used was aniline hydrogen phthalate. Three spots: One with value 0.20, 2nd with R_f value 0.25 and third with R_f value 0.36.

Acetylation : 10 mg of the sugar mixture from the hydrolysate and dry pyridine (0.5 ml) was acetylated with acetic anhydride (0.5 ml) by refluxing it at 145-150° for 4 hours. Acetic anhydride and pyridine were removed under vacuum (75-85°, 25 mm). The residue was taken up in chloroform (20 ml), washed with water (10 ml), N/2 hydrochloric acid (10 ml x 2), water (10 ml), 5% aqueous sodium carbonate solution (10 ml x 3) and finally with water (10 ml x 2), dried with anhydrous sodium sulfate and solvent removed. The residue was taken up in dry chloroform and used for GLC. under the mentioned conditions. Glucose, rhamnose and arabinose are found to be in the ratio of 1:1:1. Acetate of standard glucose, rhamnose and arabinose in the proportion of 1:1:1 were prepared and compared by GLC.

Hydrolysis of comp.B :

Comp.B (40 mg) was heated with 2N sulfuric acid (20 ml) and dioxan (10 ml) with continous stirring for 6 hours. It was cooled to room temp. and diluted with water (20 ml). The insoluble compound was filtered and taken up in chloroform. The chloroform insoluble part was identified as unreacted comp.B. The aqueous part was extracted with chloroform (20 ml x 3). The combined chloroform extracts were washed with water (20 ml x 2), dried with anhydrous sodium sulfate and solvent was evaporated to get the aglycone (20 mg), which was crystallized from methanol. m.p. 128-129°, IR (KBr), EI-MS and ^{1}H-NMR ($CDCl_3$) recorded.

Filtrate obtained after removing the aglycone was neutralised with barium carbonate. The insoluble barium sulfate was filtered off and the filtrate was concentrated and spotted on paper.

Paper chromatography : This was carried out on Whatman No.1 filterpaper under the conditions mentioned earlier for glycoside-AR_4. One spot: R_f value 0.25.

REFERENCES

1. M.E. Wall, C.R. Eddy, M.L. McCleannan and M.E. Klumpp, Anal. Chem. 24, 1337 (1952).

2. J. Elks, Rodd's chemistry of carbon compounds - IIE, pp. 1 (1971). Elsevier Publishing Co., Germany.

3. C.R. Eddy, M.E. Wall and M.K. Scott, Anal. Chem. 25, 266 (1953).

4. S.B. Rao, Indian J. Pharm. 14, 131 (1952).

5. M.E. Wall, M.M. Krider, E.S. Rothman and C.R. Eddy, J. Biol. Chem. 198, 533 (1952).

6. J. Sheer, R.B. Kostic and E. Mossetig, J. Am. Chem. Soc. 77, 641 (1955).

7. A.M. Smith and C.R. Eddy, Anal. Chem. 31, 1539 (1959).

8. D.H. Williams and N.H. Bhacca, Tetrahedron 21, 1641 (1965).

9. H. Budzikiewicz, C. Djerassi and D.H. Williams 'Structure Elucidation of Natural Products by Mass Spectrometry' - Vol.II, pp.110, Holden-Day Inc., 1964.

10. H. Budzikiewicz, J.M. Wilson and C. Djerassi, Monatsh. 93, 1033 (1962).

11. M. Kimura, Y. Haltori, I. Yoshizawa and M. Tohma, Chem. Pharm. Bull. 16, 613 (1968).

12. A. Okano, Chem. Pharm. Bull. 6, 173 (1958).

13. W. Sowa, A.C. Blackwood and G.A. Adams, Can. J. Chem. 41, 2314 (1963).

14. J.R. Dyer, Applications of Absorption Spectroscopy of Organic Compounds, pp. 33 (1978).

15. T.C. Jain and C.M. Banks, Can. J. Chem. 46, 2325 (1968).

16. K.H. Overton, Techniques of Organic Chemistry, Vol.XI, Part I, Editted by A. Weissberger - Interscience Publishers, Inc., New York, pp 42, 1963.

SUMMARY

Tentative structures of glycoside-AR_4 and comp.B are discussed. The compounds could not be studied in detaildue to paucity of the samples. IR spectrum of glycoside-AR_4 indicated that the aglycone is a normal sapogenin. FABMS was used for the determination of molecular weight and sugar sequence. Presence of glucose, rhamnose and a pentose could be detected. By acid hydrolysis sarsasapogenin was obtained as its aglycone and glucose, rhamnose and arabinose in the proportion of 1:1:1 as its sugars. A probable structure was assigned to it on the above mentioned results.

IR spectrum of comp.B confirmed that it was not a steroidal saponin. FABMS showed it to be a mixture of two compounds. By acid hydrolysis it gave a mixture of a sterol and a stanol as its aglycone and glucose as its sugar. It was concluded from the above data that comp.B is probably a mixture of sitostanol-glucoside and sitosterol glucoside.