Milligram Scale Automatic Preparative GLC of the Steranes and Triterpanes Isolated from Green River Formation Oil Shale

E. Gelpi, P. C. Wszolek, E. Yang, and A. L. Burlingame

Space Sciences Laboratory
University of California, Berkeley, California 94720

Introduction

To the organic geochemist the study of the cyclic hydrocarbon fraction extractable from the Green River Formation Oil Shale is both a multifaceted challenge and a promising reward: A challenge in that to secure meaningful information from the very complex mixture of isomeric hydrocarbons in the Green River bitumens (1), he must make use of the most efficient and sophisticated separation and identification techniques; and a reward in that the complete structural characterization of such a unique array of alicyclic hydrocarbons will allow, through the concept of "molecular fossils"(2), a better understanding of the ecological setting at the time the sediment was laid down (3). Although much work has been done in this area in the past five years (1, 4-8) only a few of the major constituents in the range C_{27} - C_{31} have been identified. Included in the list are the 5α- and 5β- isomers of the C_{27}, C_{28} and C_{29} tetracyclic steranes (1,4,5,7,8) and a few isomeric pentacyclic triterpanes with empirical formulas, $C_{30}H_{52}$, $C_{30}H_{56}$, $C_{29}H_{50}$, $C_{31}H_{54}$ (4-7). However, most of these identifications are still tentative and as a consequence the stereochemical configurations of the molecules are not known. This is due to the difficulties encountered in resolving the whole mixture, especially each group of isomers into individual components. The most versatile technique in this field has been the combination of the gas chromatograph with the mass spectrometer (1,7), but even in cases where GLC gives the necessary resolution, structural assignment from the MS data alone can be rather ambiguous particularly among the different isomeric forms.

The approach followed in this work (9) rests first of all on the determination of the optimum operating GLC parameters for preparative collection, in the highest purity possible, of any preselected compound. This will allow complete structural characterizations by other spectrometric methods, such as C^{13} NMR or x-ray diffraction.

Experimental

A. Sample description and treatment. The sample used in this work originated at the Colony Mine, 15 miles N.W. of Grand Valley, Colorado. About 17.5 Kg were pulverized in a ball-mill for 10 days and then extracted ultrasonically twice (15-30 minutes each time), with a mixture of benzene-methanol (2:1). The bitumen fraction thus extracted weighed 250 g which represents a recovery of 1.4% of the total.

B. Separation of the acidic fraction. The extract was sonicated and saponified with 1.5 liters of 2M KOH/methanol for 3 hours. The resulting basic solution was extracted with n-heptane (300-450 ml) and then washed with water to remove the acids.

C. Separation of the basic fraction. The heptane fraction was washed with 700 ml of 6NHCl and then again with water to remove the basic components.

D. Column Chromatography. After removal of the acidic and basic fractions the remaining heptane extract was chromatographed on a column of activated Al_2O_3 (37.5 cm high by 4.5 cm i.d.). Heptane, benzene and methanol were used sequentially in order to separate the non-polar aliphatic-alicyclic hydrocarbons, from the aromatic and other polar lipid components. Three cuts of 300 ml each were obtained for the heptane eluate. Cuts No. 2 and 3 were found to be enriched in the aliphatic and alicyclic fractions respectively.

E. Digestion with molecular sieves. Both cuts 2 and 3 (dissolved in benzene) were treated with 5 Å molecular sieves (10) for three days under reflux, to separate the straight chain from the branched and cyclic hydrocarbons.

F. Thiourea adduction. The resulting branched-cyclic fraction was adducted with thiourea according to the following procedure (5). The branched-cyclic hydrocarbons were dissolved in 700 ml of chloroform before the addition of a saturated solution of thiourea in 700 ml of methanol. The resulting mixture was heated until it was homo-

genous, cooled, and kept at room temperature for one day. Thiourea adduct needles fell out of solution on cooling, and these were later collected, rinsed with cold chloroform and dissolved in water. The water solution was then extracted three times with heptane. The thiourea adduction was repeated once on the supernatant solution so that the overall treatment yielded a combined adduct fraction enriched in the 5α- steranes and a non-adduct fraction enriched in the 5β- steranes and the triterpanes (9).

G. Liquid gel-filtration chromatography. About two gr of the total non-adduct were chromatographed on Sephadex LH:20 to partially separate the 5β- steranes from the triterpanes. Details of this step have been presented elsewhere (9).

H. Gas-liquid chromatograph. All of the packed columns were prepared with short lengths of 316 stainless steel tubing purchased from the Perkin-Elmer Corp., Norwalk, Connecticut. The columns were packed under vibration and beating (11) and conditioned at 290°C until ready for use. The stationary phases and supports were purchased from: 5% SE-30 (methyl silicone) on 100-120 Varaport 30 and 3% SE-30 on 80/100 Aeropak: Varian Aerograph, Walnut Creek, California; 3% SE-30 on 100/120 Gas Chrom Q: Applied Science, State College, Pa.; OV-101 (dimethylsilicone fluid): Supelco, Inc., Supelco Park, Bellefonte, Pa. All analytical gas chromatographic data were obtained on a Perkin-Elmer Model 900 gas chromatograph equipped with flame ionization detectors. In general, the detector temperature was kept at 370 ± 10°C. Helium was used as the carrier gas and all data were obtained under isothermal conditions.

I. Preparative GLC. One of the standard Perkin-Elmer Model 900 gas chromatographs available in our laboratory was coupled to the model 900 preparative attachment (Cat. No. 009-8002) for automatic cyclic preparative work. The combined system can be seen in Figure 1, and its principle of operation will be described below.

J. Electrostatic precipitaiton. The unit Model 850 Prepkromatic electrostatic precipitator was purchased from Nester Faust Mfg. Corp., Newark, Delaware, 18711, and was used without modification.

Results and discussion

1. Analytical GLC

The complexity of the alicyclic hydrocarbon mixtures in the molecular weight range of interest to us is shown in Figure 2. The adduction of the heptane eluates with thiourea, as described above, is very useful in that it segregates almost quantitatively the tetracyclic sterane isomers according to the stereochemical configuration (cis or trans) of the 5, 10 ring fusion. Only the steranes (cholestane, C_{27}; ergostane, C_{28}; sitostane, C_{29}) are labelled in the figure. Other peaks correspond to the triterpanes. Thus the pattern obtained after digestion with molecular sieves and before adduction with thiourea would be the sum of the two GC traces shown in Figure 2A.

Although the GC patterns of the thiourea non-adduct fractions (Figures 2A, B, C) cannot be considered too complex at first glance, the difficulties associated with the isolation and identification of each component can be understood by inspection of the chromatographic trace in Figure 2D. This is the chromatogram of the same mixture on a short capillary column with an efficiency about five times greater (9) than, for instance, that of the packed column giving the separation shown in Figure 2B. Three changes in the GC pattern of Figure 2D are most evident. The peak corresponding to the tetracyclic sterane 5α- C_{29} in chromatograms A, B and C (Figure 2) is actually a mixture of two components and the same is true for the two large peaks eluting after the major component.

With these facts in mind, it is evident that prior to attempting any preparative isolation or GC/MS identification of the different components in these mixtures, special attention must be paid to the chromatographic parameters required for maximum separation. This point (optimization of GLC parameters) has been treated in detail in the literature (9), and only a few general ideas will be added here.

Due to their total non-polar character and close structural similarities, the choice of chromatographic methods for these types of separations becomes rather limited. Although we are presently investigating the possibilities of a high pressure

liquid chromatography system, GLC seems to be the most practical and efficient technique. However, the types of columns that can be used are limited to the long and narrow bore analytical columns (high efficiencies) due to the complexity of the mixtures. One is also limited to high operating temperatures by the molecular weight of these compounds, which in turn restrict the choice of stationary phases to only those most thermally stable. Because of the characteristically large elution volumes of these compounds the support size in analytical packed columns must be less than 100 mesh to achieve fast enough flow rates.

In terms of the separation of individual compounds in substantial amounts, their preparative GLC is restricted to small sample sizes if high column efficiencies are to be maintained (9). Recovery yields can be extremely low if standard trapping systems are used, due to the high tendency that these hydrocarbons have to form aerosols.

For the purpose of this work chromatographic resolution becomes an all decisive factor. Since resolution (R) is directly proportional to both the efficiency (N) and the selectivity (α) of the column, it follows that R could be increased by a favorable change in any of the two factors. However, the choice of liquid phases is restricted by high operating temperatures as stated above, so that not much can be done in terms of choosing the proper selectivity factors. This spells out the requirement for the highest possible efficiencies of the chromatographic system. For instance, considering the doublet of peaks eluting between 42 and 43 minutes in Figure 2D, it is possible to calculate the required minimum efficiency to separate them with a given resolution by means of the formula (12): $Nr = 16R^2 (\alpha/\alpha - 1)^2 (k_2 + 1/k_2)^2$.

From the chromatogram in Figure 2D: $R = 0.7$; $\alpha = 1.02$; and $k_2 = 10.9$. (The factor k_2, known as the retardation ratio (12), is expressed as the ratio of the corrected retention time of peak 2 in a binary mixture to the retention time of the air peak). To double the resolution from 0.7 to 1.4, Nr must then be equal to 96249 theoretical plates. Since the efficiency of the 30 m long capillary column (Figure 2D) can be calculated, also from the GC trace, to be equivalent to 15,000 plates, which sets the height of a theoretical plate at 2 mm, the length required to obtain 96249 plates for an equivalent column would be 193.5 m. Thus, just to double the resolution in this particular case one needs to increase the number of plates and the length about 6.4 times. The magnitude of these numbers, well above those characteristic of large diameter preparative columns, forced us to investigate the potential use of small bore analytical columns for preparative purposes, giving special attention to the effect of sample size on resolution and efficiency. This effect is shown graphically in Figures 3 and 4. As predicted, the sharpest decline of resolution with the increase in sample size (Figure 3) occurs in the 0.1 cm i.d. column, while the 0.45 cm i.d. column is less affected and the 0.20 cm i.d. columns are in an intermediate position. The broken line corresponds to two 3 m x 0.10 cm i.d. columns coupled together and, although higher resolutions can thus be obtained, the slope is not any better than that obtained for each one alone. The parallel loss of resolution and efficiency with sample size is shown in Figure 4. Note the simultaneous change of slope in both parameters at 50 µg sample size (9).

2. Preparative GLC

Considering the narrow sample size range (1 - 150 g) to which one is limited by the requirements of resolution and efficiency, a repetitive automatic injection and collection unit becomes a necessity, especially if milligram amounts of each substance are to be collected in a reasonable amount of time. For this purpose, one of the two Perkin-Elmer Model 900 gas chromatographs available in the laboratory was fitted with the Model 900 preparative attachment as shown in Figure 1. A schematic diagram of the complete system is given in Figure 5. The carrier gas flows from 1 to 9 through the solenoid and non-return valves 2 and 3. In the injection block a small split takes place. Most of the carrier gas goes into the GC column 10, but a small part (3 - 8 ml/min) flows out through 8, the dosing capillary and into the sample vessel. This constant backflow is regulated by valve 5. Thus, at any given

time the pressure at the head of the column equals the pressure in the sample vessel. Dosing is triggered by the electronic programmer unit which contains the control electronics for automatic cyclic operation and controls solenoid valve 2. When a dosing command is given by the programmer the entire gas flow is diverted through 4 into the sample vessel. In consequence the equilibrium in pressures is momentarily upset so that sample flows from 7 to 9. The quantity of sample that is thus injected into the column is fixed by the total time that the solenoid valve 2 is open, the dimensions of the dosing capillary and the viscosity of the sample. When 2 is closed by the programmer the original column pre-pressure is re-established in 9 and the capillary is flushed free of sample.

At the exit of the column the effluent is split so that from 1 to 10%, or more in some cases, goes to the flame ionization detector and the rest goes on through the distributor assembly to the traps which are opened and closed by solenoid valves placed on their exit sides. This eliminates moving parts in the flow path of the sample before it is collected. The traps are opened and closed according to preset threshold values and according to a pre-selected peak sequence which is controlled by the electronic programmer. One of them (waste trap) is always on as a bypass trap to avoid losing the part of the effluent not directed to the other traps. Both the injection and collection systems have been modified in our laboratory to allow the use of analytical size columns in the preparative mode and to insure maximum recovery of the compounds emerging from the column. This will be described in detail elsewhere. The results obtained with this system are shown in Figure 6. Both chromatograms correspond to those shown in Figure 2A for the thiourea adduct and non-adduct fractions.

The preparative isolation of both the adduct and non-adduct fraction was achieved by taking into account the proper parameters to insure optimum resolution (9), and minimum analysis time in accordance with the considerations discussed above. The operational parameters selected for the collection of the two major peaks in the thiourea adduct (Figure 6, top), 5α-cholestane and 5α-ergostane, are given in Table 1. The data in Table 1 is also representative of other longer runs. In one of them the unit operated continuously in the automatic mode for a total of 116 hours and 20 minutes. The top chromatogram in Figure 6 is taken from one of the actual automatic cycles. From this chromatogram, the operating efficiency of the column under the conditions selected for the run can be calculated as equivalent to 3047 theoretical plates, and the resolution for a mixture of 5α- and 5β-cholestane as 1.79 for sample sizes of 100 μg/peak. Common threshold values for the opening and closing of the traps were set for the front (F) and rear (R) of each one of the two peaks. Peak 1, collected in trap I, corresponds to 5α-cholestane and peak 2, collected in trap II, corresponds to 5α-ergostane. The shaded areas indicate how much of the peak was directed to the corresponding trap and the length of time for which each trap was opened. The purity of bands collected in this fashion was assayed by re-chromatographing them on a more efficient column (N = 6789 theoretical plates). The results are shown in Figure 7. The chromatogram of the total adduct is given in Figure 7A for purposes of comparison with that shown on Figure 2A. Figure 7B shows what was collected in the waste trap. Note the practical disappearance of the two major peaks from the pattern. The band of 5α-cholestane, collected in trap I, is shown in Figure 7C and that of 5α-ergostane collected in trap II is shown in Figure 7D and Figure 7E. Both steranes can be estimated chromatographically at a purity greater than 99%. Quantitative data are given in Table II. The contribution of column bleed to the figures given in Table II has not been deducted; however, it can be estimated to be negligible in traps I and II and less than 10 mg in the waste trap. Assuming a total contribution of 10 mg of column bleed the recovery yield is still 94.3%. It is obvious from these chromatograms any mixing or cross contamination of traps in the distributor assembly is practically negligible with this system.

Similar parameters as those shown in Table I were set up for the collection of the four major multi-component peaks in the non-adduct fraction (Figure 6, bottom). However, in this case the diameter of the column was reduced to half (from 0.45 cm i.d. to 0.20 cm i.d.) to approach the minimum resolution factors required for the

preliminary separations of these substances. The flow rate (56.6 cc/min) selected was a compromise between the optimum for maximum resolution and the fastest allowable for minimum analysis time. However, analysis time was favored over resolution in this case, since it could be calculated that this column would not be able to resolve the multi-component peaks under any conditions.

The split ratio was readjusted by changing the stream splitter so that only 2.1% of the total effluent was directed to the FID. The variable threshold programmer allowed the collection of each band with individually set thresholds as indicated on Figure 6, bottom. The purity of the collected bands was again assayed by GLC on the same column used to check out the bands collected from the adduct fraction. The total GC pattern of the pentacyclic triterpane fraction is shown in Figure 8A. The GLC trace obtained from the material collected in the waste trap shows the relative degree of removal of each of the peaks which were collected in traps I through IV (Figure 8C - F). The peak collected in trap I shows the presence of a minor component underneath its front slope, and furthermore, it is known from the separation on capillary columns (see Figure 2D, doublet of peaks at about 35 minutes) to be a 50:50 mixture of two components, one of them corresponding to the 5α - C_{29} sterane in retention time.

The two small peaks following the major compound collected in trap II can be easily explained considering the threshold levels 2R and 3F in Figure 6, bottom. During the time trap II stayed open (2F - 2R) it was a small amount of component 3, which is itself a mixture of at least two components as indicated by the chromatogram in Figure 8E. The same is true for component 4 in trap 4 (Figure 8F). The cross contamination of trap 3 with the components collected in trap 4 is consequence of an accidental error in trap sequencing due to a faulty injection. This is of little importance in this case because it was known that each of these peaks would have to be re-chromatographed on capillary columns of an efficiency and length at least equivalent to those calculated above in order to resolve them completely.

At the time of this writing, less than 10mg of the non-adduct fraction have been run through the system. Full quantitative data will be presented at a later date, together with the results obtained with the use of capillary columns in the preparative system.

Another major problem encountered in this work deals with the tendency that these substances have in forming aerosols, thus decreasing the collection efficiency. Extensive work in this area with labelled cholesterol and other standards showed this to be a significant factor. Many trap configurations, packing materials and coolant systems were tested, unsuccessfully for the most part. The most effective system was determined to be that of electrostatic precipitation. For this purpose a commercial unit was bought from Nester Faust and the standard Perkin-Elmer traps designed for aerosol forming substances were modified to fit the slip over double lead probes. The unit when turned on sets a field of 5000 volts AC at 0.040 amps between the two electrodes, Without the electrostatic precipitator, collection efficiencies of the order of 60% were difficult to achieve, while they went up to about 95% with the use of the precipitator. No change at all was ever observed in any of the GC patterns even when this possibility was checked with standards, which indicates that compound breakdown does not take place with the AC voltage, and that the electrostatic precipitator is safe to use in these separations.

Conclusions

The isolation of relatively large quantities of each one of the steranes and triterpanes from the Green River Shale bitumens with this system will enable us to undertake a full study of their molecular structures. The data obtained so far indicates that with proper consideration of the basic parameters of the chromatographic process it becomes possible not only to separate very complex mixtures, but also to collect each component on an individual basis and in high purities for further study. Thus, it appears that if analytical specifications can be incorporated into a standard preparative system, new ways will be open in the field of natural product analysis toward the isolation of the significant components within any given

mixture. The manufacturers of the existing preparative systems should give some attention to the many chemists handling mixtures that simply cannot be resolved on standard preparative gas chromatographs. These kinds of problems, similar to the one discussed here, will ultimately call for reliable automatic units with injection systems capable of going down to the 1 μg level or less and with efficient trapping systems for collection of submilligram amounts.

Acknowledgments: We thank Mr. G. Johanson and Mr. S. Gaskell for technical assistance. Financial support was provided by the U. S. National Aeronautics and Space Administration Grants NA 9-7889 and NGL 05-003-003.

References:

1. E. J. Gallegos, "Identification of new steranes, terpanes and branched paraffins in Green River Shale by combined capillary gas chromatography and mass spectrometry." Presented ACS meeting, Div. Petrol. Chem. Vol. 15, No. 3, Chicago, Illinois, September 13-18, 1970.
2. M. Calvin, "Chemical evolution", Oxford University Press, New York, 1969.
3. W. H. Bradley, Geol. Soc. Am. Bull., 81, 985 (1970).
4. A. L. Burlingame, P. Haug, T. Belsky, and M. Calvin, Proc. Nat. Acad. Sci., 54, 1406 (1965).
5. T. J. Murphy, A. McCormick, and G. Eglinton, Science, 151, 1040 (1967).
6. I. R. Hills, E. V. Whitehead, D. E. Anders, J. J. Cummins, and W. E. Robinson, Chem. Communications 752 (1966).
7. W. Henderson, V. Wollrab, and G. Eglinton, "Advances in organic geochemistry", P. A. Schenck and I. Havenaar, eds., Pergamon Press, 1968.
8. P. C. Anderson, P. M. Gardner, E. V. Whitehead, D. E. Anders, and W. E. Robinson, Geochim. Cosmochim. Acta, 38, 1304 (1969).
9. E. Gelpi, P. Wszolek, E. Yang and A. L. Burlingame, J. Chrom. Science (in press). Presented at the 6th International Symposium on Advances in Chromatography, Miami Beach, Florida, June 2-5, 1970.
10. C. Fenselau and M. Calvin, Nature, 212, 889 (1966).
11. A. B. Littlewood, "Gas chromatography", Academic Press, New York, 1962.
12. I. Halász and E. Heine, "Progress in gas chromatography", J. H. Purnell, ed., Interscience, New York, 1968.

Table I

Green River Shale Adduct Fraction - Typical Operating Parameters for Automatic Preparative GLC.

Column:	6m x 0.45 cm i.d. 5% OV-1 on 60/80 mesh supelcoport
Carrier gas:	He, 163.5 cc/min
Temperatures:	Injector 290°C Column 270°C Detector 380°C Distributor 370°C
No. of injections:	42
Total Time/cycle:	59' 22"
Total time (42 cycles):	41 hours 33'
Average volumn injected:	17.4 µl
Range in vol. injected:	7 µl to 21.3 µl
Average amount injected:	79.0 µg (5α-cholestane)
Max. amount injected:	96.9 µg (5α-cholestane) 194 µg (5α-ergostane)
Split ratio:	11% to FID
Dosing time:	3 seconds
Trigger levels:	See Figure 6.

Table II

Quantitative Estimate of Recovery Yields from Total Adduct Fraction.

Adduct fraction; total:	251.000 mg
Not used:	115.760 mg
Trap I (5α-cholestane):	10.800 mg
Trap II (5α-ergostane):	25.561 mg
Waste trap:	70.350 mg
Split to FID 11%:	24.471 mg
Total accounted for:	246.942 mg
Recovery yield:	98.3%

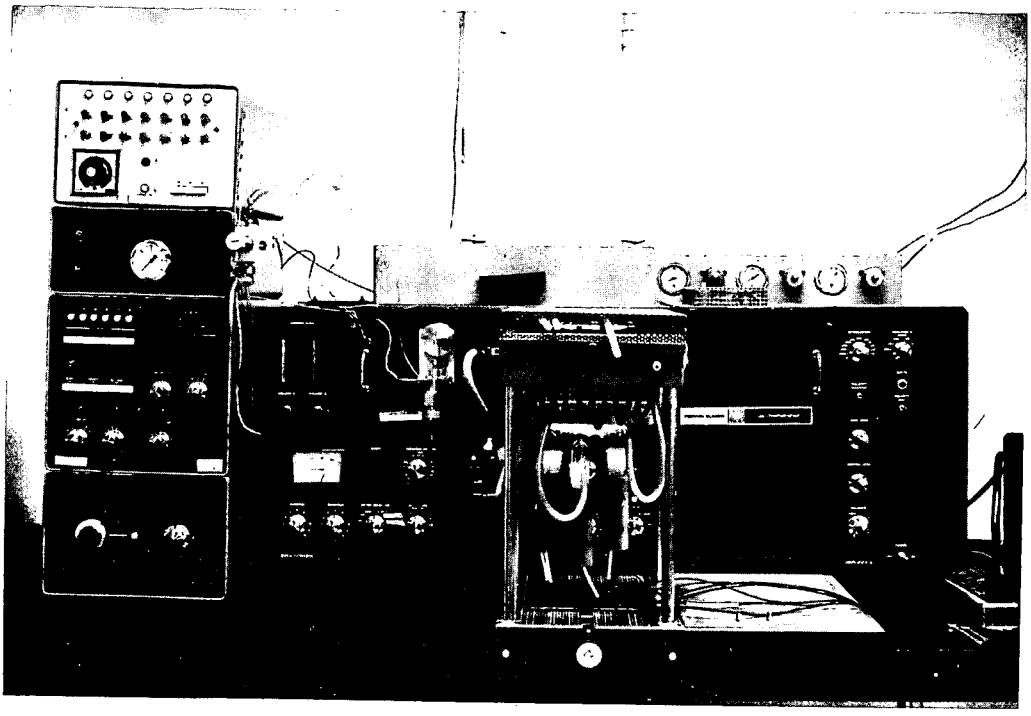


Figure 1

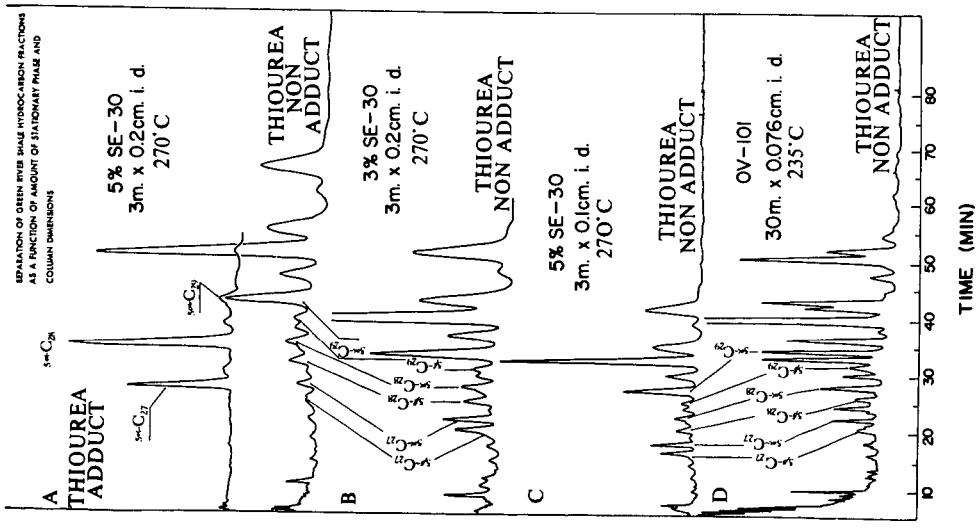


Figure 2

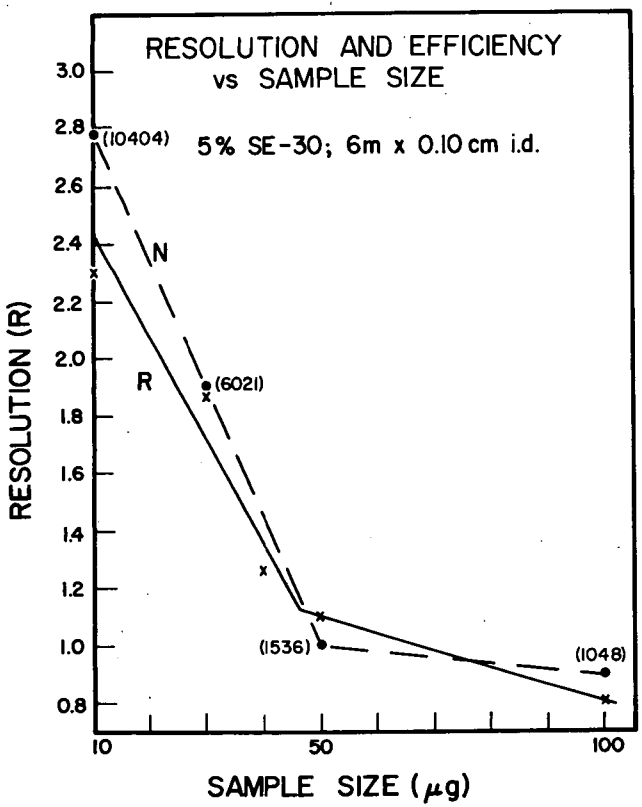


Figure 4

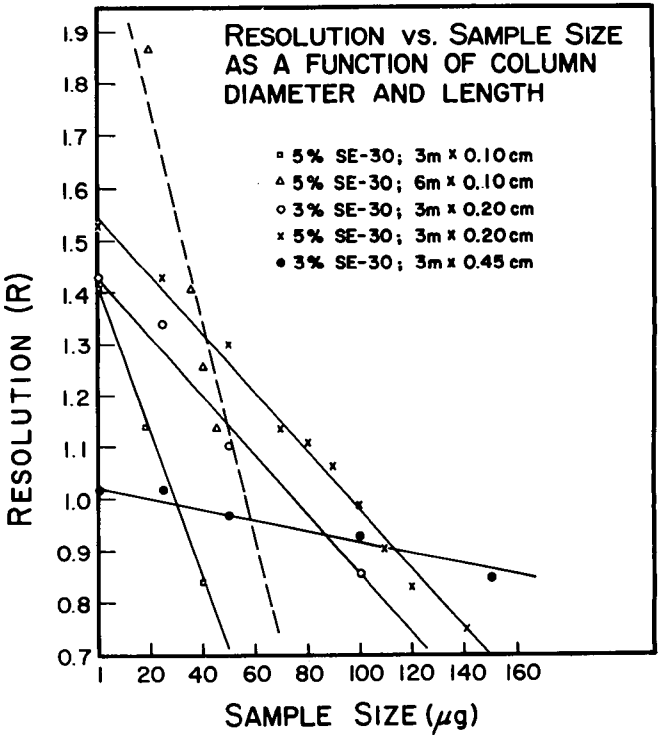


Figure 3

Schematic of Automatic Preparative Accessory

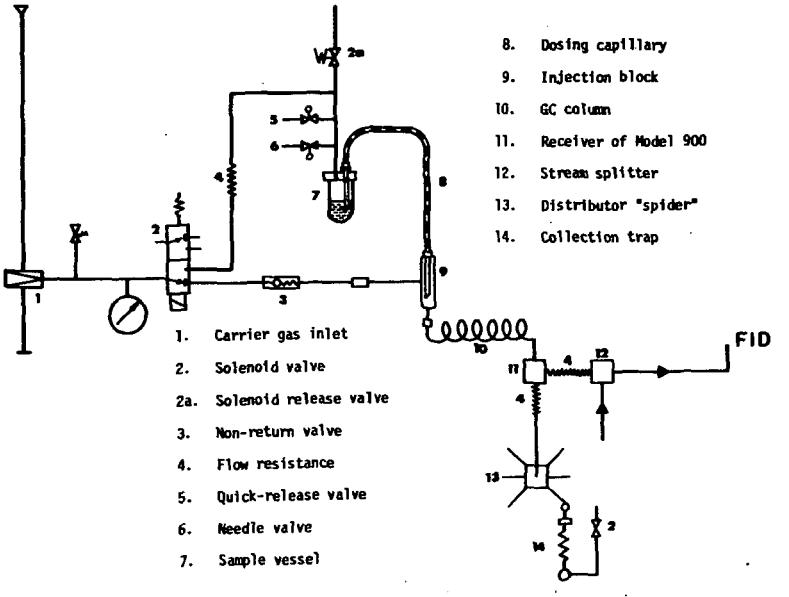
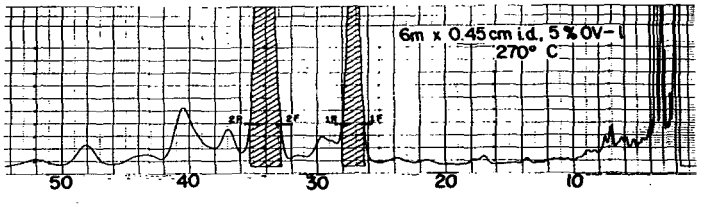


Figure 5

GREEN RIVER SHALE ADDUCT FRACTION

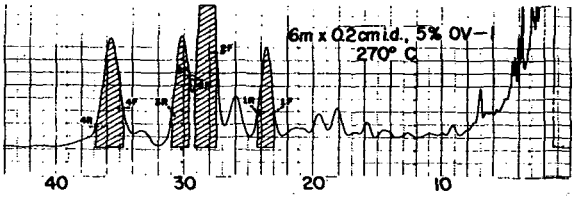
AUTOMATIC COLLECTION WITH A COMMON THRESHOLD



TIME (MIN)

GREEN RIVER SHALE NON-ADDUCT
PENTACYCLIC TRITERPANES

AUTOMATIC COLLECTION WITH INDIVIDUALLY SET THRESHOLDS



TIME (MIN)

Figure 6

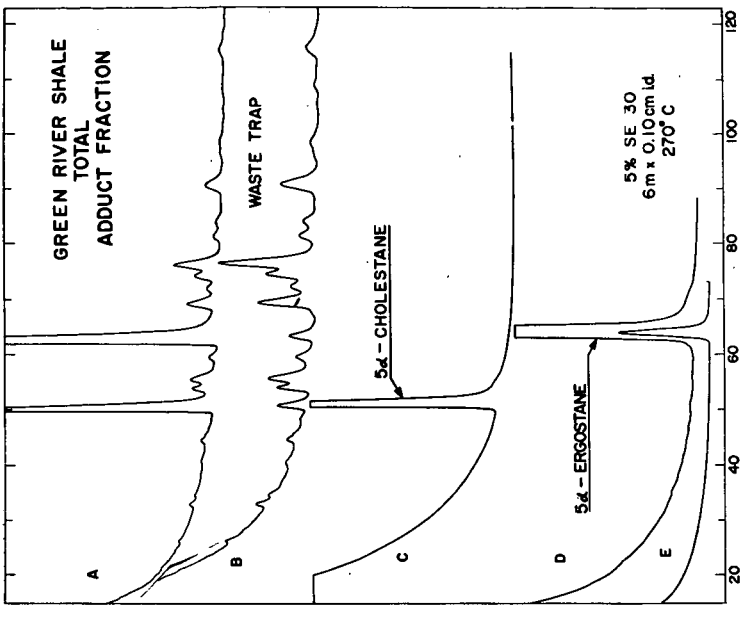


Figure 7

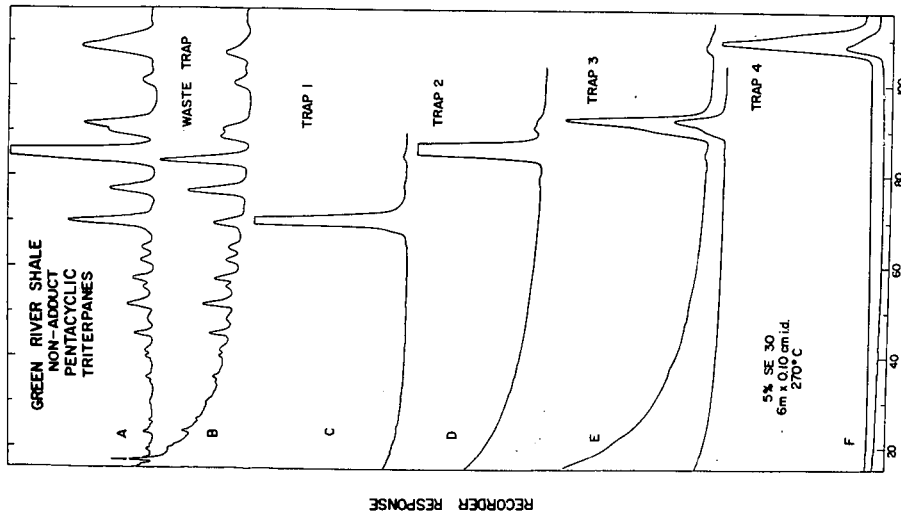


Figure 8