# Sulfolipid I of *Mycobacterium tuberculosis*, Strain H37RV. Nature of the Acyl Substituents*

Mayer B. Goren,† Olga Brokl, Bhupesh C. Das, and Edgar Lederer

ABSTRACT: Sulfolipid I of *Mycobacterium tuberculosis*, strain H37Rv, was previously characterized as a 2,3,6,6′-tetraacyltrehalose 2′-sulfate.

The structures of the acyl functions have been elucidated largely by mass spectrometry, and are reported herein. Three principal (and related) series of carboxylic acids were found: palmitic–stearic acids with minor amounts of other homologs; a multibranched series, the "average" member of which is 2,4,6,8,10,12,14-heptamethyltriacontanoic acid; and a second, related, oxygenated multibranched group consisting principally of 17-hydroxy-2,4,6,8,10,12,14,16-octamethyldotriacontanoic acid.

Homology by 42 mass units is prominent in both series and suggests a biogenesis involving successive incorporations of propionate onto a palmitate residue. All representatives of the two methyl-branched series are dextrorotatory; they are therefore very likely of the L configuration and related to the phthienoic(mycolipenic) acids.

Virulent strains of *Mycobacterium tuberculosis*, var. *hominis*, elaborate an unusual sulfated glycolipid when grown in surface culture (Middlebrook *et al.*, 1959). Earlier reports in this series (Goren, 1970a,b) described studies in which the principal sulfolipid (SL-I) of strain H37Rv was shown to be a complex 2,3,6,6′-tetraacyltrehalose 2′-sulfate. Analysis of the ammonium salt suggested the approximate empirical formula $C_{145}H_{275}NO_{20}S$ ±(5C, 10H) molecular weight about 2384. Additional sulfolipids were also seen in small amounts during the course of chromatography of the crude lipids on diethylaminoethylcellulose, almost homogeneous samples of these being occasionally obtained. Alkaline solvolysis of the principal (and of some of the minor) sulfolipids allowed the recognition of apparently four carboxylic acids. We report here the isolation, separation, and structural analysis, chiefly by mass spectrometric methods, of the acids on the trehalose core of these glycolipid sulfates.

## Experimental Section

Purification of solvents, column and thin-layer chromatographic procedures, bacterial culturing, and preparation and solvolysis of SL-I were as described in our earlier studies (Goren, 1970a,b).

Methylation of lipid products obtained by spontaneous desulfation of the ammonium sulfolipids (Goren, 1970b) was carried out most satisfactorily with diazomethane–boron trifluoride etherate (Caserio *et al.*, 1959; Goren, 1970b): at about 10-min intervals, 5-ml portions of about 1% diazomethane followed after 1–2 min by 0.1-ml portions of catalyst, were added to stirred solutions of the glycolipid at ice temperature for a prolonged period. Very large excesses of diazomethane were used; this effected essentially complete methylation as indicated by infrared examination. For instance, for 10 mg of a lipid of mol wt 2200 with six hydroxyl functions, as much as 0.3 g of $CH_2N_2$ would be added in 8–10 portions along with eight 0.1-ml portions of 2% boron trifluoride etherate in $CH_2Cl_2$. By contrast, it is suggested in recently reported work of Brennan *et al.* (1970) that a one-shot but prolonged methylation with $CH_2N_2$ and $BF_3$ is less effective.

Methylations were also conducted according to the procedure of Brimacombe *et al.* (1966): about 2 mg of substance in a small volume of anhydrous dimethylformamide is treated with an excess of sodium hydride and dry methyl iodide. With certain sterically hindered branched-chain hydroxy esters encountered in the present study, methylation was usually incomplete even after several hours of reaction; the methylated substance was easily separable from unconverted starting material by preparative thin-layer chromatography on silica gel.

Recovery of methylated trehalose products, HCl-catalyzed hydrolysis, and recovery of the individual monosaccharides and identification by thin-layer and gas–liquid chromatography of trimethylsilyl derivatives were done as described in earlier reports (Goren, 1970b). Deuterium exchange involving activated ($\alpha$) hydrogens in certain branched-chain esters of the present study was conducted by maintaining about 2 mg of ester at reflux 18 hr in 3 ml of $CH_3OD$ containing catalytic amounts of $NaOCH_3$. The reaction mixture was evaporated dry *in vacuo*, mixed with a little $D_2O$ and sufficient acetic acid to neutralize the base, and the ester recovered by extraction into hexane. For oxidation of certain hydroxy esters, about 6 mg of $CrO_3$ was slowly added (1 hr) with stirring to 0.08 ml of pyridine (Poos *et al.*, 1953) maintained at 15–20°. Hydroxy ester (10 mg; about 0.015 mmole) in 0.2 ml of pyridine was added; the mixture was maintained at 20° for 18 hr, diluted with 5 ml of $H_2O$, and the keto ester was extracted with several portions of hexane, which were combined, washed with $H_2O$, dried with $Na_2SO_4$, and evaporated *in vacuo* for recovery of the product. The keto ester was separable from traces of contaminating starting material by preparative thin-layer chromatography, the keto ester having

* Joint contribution from Division of Research, National Jewish Hospital and Research Center, Denver, Colorado (M. B. G. and O. B.) and Institut de Chimie des Substances Naturelles, Gif-sur-Yvette, France (B. C. D. and E. L.). *Received July 30, 1970.* This investigation was supported largely by Grant No. AI-08401 of the U. S. Japan Cooperative Medical Science Program administered by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health, Department of Health, Education, and Welfare. A part of these funds provided the support allowing M. B. G. to participate personally in the joint effort at the Institut de Chimie des Substances Naturelles at Gif-sur-Yvette, France.

† To whom to address correspondence.

about twice the mobility of the hydroxy ester on silica gel (hexane–ether–acetic acid, 90:10:2, v/v). Overall yield was about 85%.

Gas chromatography of fatty acid esters and trimethylsilyl ethers of carbohydrates and carbohydrate derivatives was done with a Barber Colman series 5000 gas chromatograph as described earlier (Goren, 1970b). Infrared spectra on microgram quantities of products were obtained with a Beckman IR-7 spectrophotometer equipped with a beam condenser. The samples were mostly studied as thin smears on a Beckman KBr micro demountable cell. Nuclear magnetic resonance spectra were measured at 60 MHz on a Varian 60A instrument. Mass spectra were obtained with a MS9 (AEI) mass spectrometer.

## Results

*Distribution of Acyl Substituents in Different Sulfolipids.* In our earlier studies, we were concerned with methodology for preparing essentially homogeneous SL-I from the mixture of crude lipids extracted with hexane–0.1% decylamine from the harvested tubercle bacilli. This was achieved by column chromatography on DEAE-cellulose. A contaminant lipid designated SL-II was occasionally obtained essentially homogeneous in fractions just preceding the principal sulfolipid. SL-II also has somewhat higher mobility in thin-layer chromatography. Still a third lipid, designated SL-I′, trails SL-I in DEAE chromatography and also in thin-layer chromatography mobility, and was recognized as a contaminant of the final SL-I fractions collected in the column chromatography.

The lipid products recovered from alcoholysis of SL-I or desulfated SL-I appear to contain four classes of carboxylic esters as revealed by silica gel thin-layer chromatography (hexane–ether–acetic acid, 90:10:1 or 90:10:3, v/v). In decreasing order of mobility, these are designated as A, B, C, and D. Component B was early identified as a mixture consisting principally of methyl palmitate and stearate (Goren, 1970b). Studies with the "minor" sulfolipids II and I′ showed that component A is apparently missing in SL-II;[1] but B, C, and D seem to be present in the same proportions as they are in SL-I. SL-I′ contains the same four substances as SL-I, but apparently has a *double* complement of A and *half* the complement of component *C* (Figure 1).

Permethylation analysis showed that SL-I′, like SL-I, is also a 2,3,6,6′-tetraacyltrehalose. Like the principal sulfolipid, SL-I′ desulfates spontaneously in ether. Moreover, alkaline solvolysis of SL-I′ yields trehalose 2-sulfate, characterized as the potassium salt. Accordingly, it resembles SL-I in all respects save in the distribution of components A and C.

Infrared examination of the individual methyl esters separated by preparative thin-layer chromatography showed A (Figure 2a) and B to be considerably different esters, but nevertheless free of other functional groups; C (Figure 2b) and D, on the other hand, appeared to be very similar hydroxylated esters, with a prominent OH band at 3550 $cm^{-1}$; a deep carbonyl absorption band at about 1745 $cm^{-1}$ was accompanied by a shallower peak (about half the intensity) at about 1725 $cm^{-1}$. (The *mycolic acids* exhibit the same two absorption bands, but with reversed intensity; Etemadi, 1965.) Accordingly, our infrared spectra for a time seemed to support an unusual mycolic structure which was once considered for components C and D (*vide infra*).)

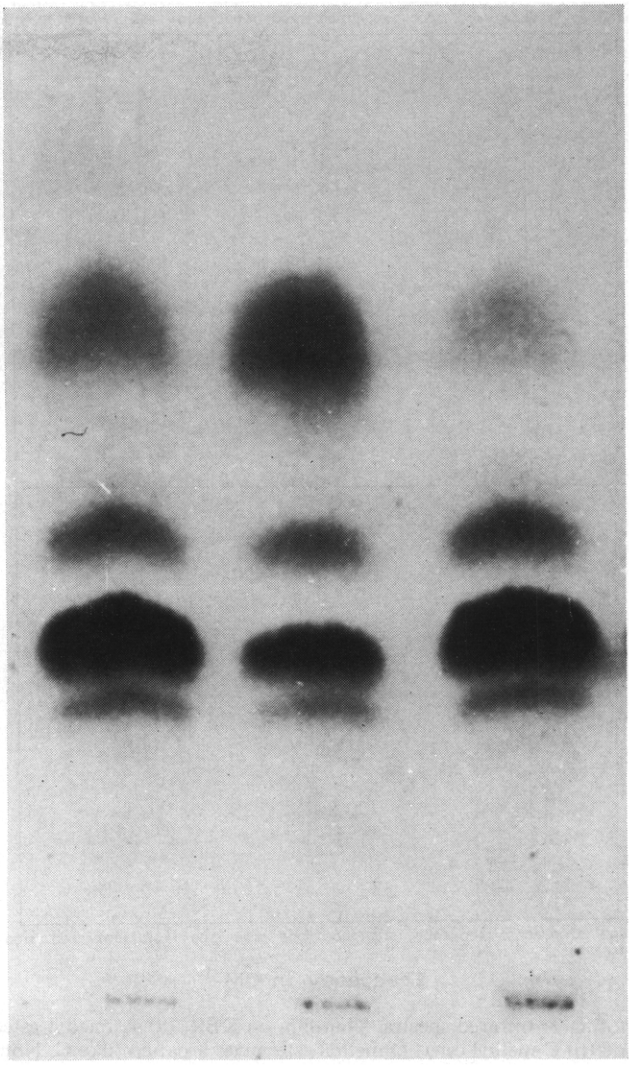


FIGURE 1: Thin-layer chromatographic pattern of carboxylic acids derived from (left to right) SL-I, I′, and II; in descending order of mobility the acids are designated A, B, C, and D.

If the four seemingly different acyl substituents are present to the extent of *at least* 1 equiv each per mole of a given sulfolipid, then SL-I and I′ must have five substituents, a conclusion in conflict with that from permethylation analysis. The latter showed that only four positions of these lipids are acylated (even if five substances are involved).

In a first effort to reconcile the inconsistency, we reasoned that if any of the hydroxy acids were esterified through the hydroxyl function with any of the other acids, four positions on the trehalose could then accommodate five or more acyl groups. We have fragmentary evidence that a complex of this sort may in fact exist. However, the hydroxy acids in the sulfolipid are in the main *not* esterified at the OH group: when desulfated SL-I or SL-I′ is permethylated ($CH_2N_2$, $BF_3$), essentially *no* hydroxy acids (C and D) are recovered; instead only *O*-methyl ethers are obtained. Thus, with only minor exceptions, the hydroxyl functions in C and D are free in the parent sulfolipids.

The inconsistency was nevertheless resolved by the finding that D, which gives a very weak spot on thin-layer chromatography, is present only in small quantities; moreover C is, in

[1] The trace of A which is seen is probably due to slight contamination of the sample with SL-I.

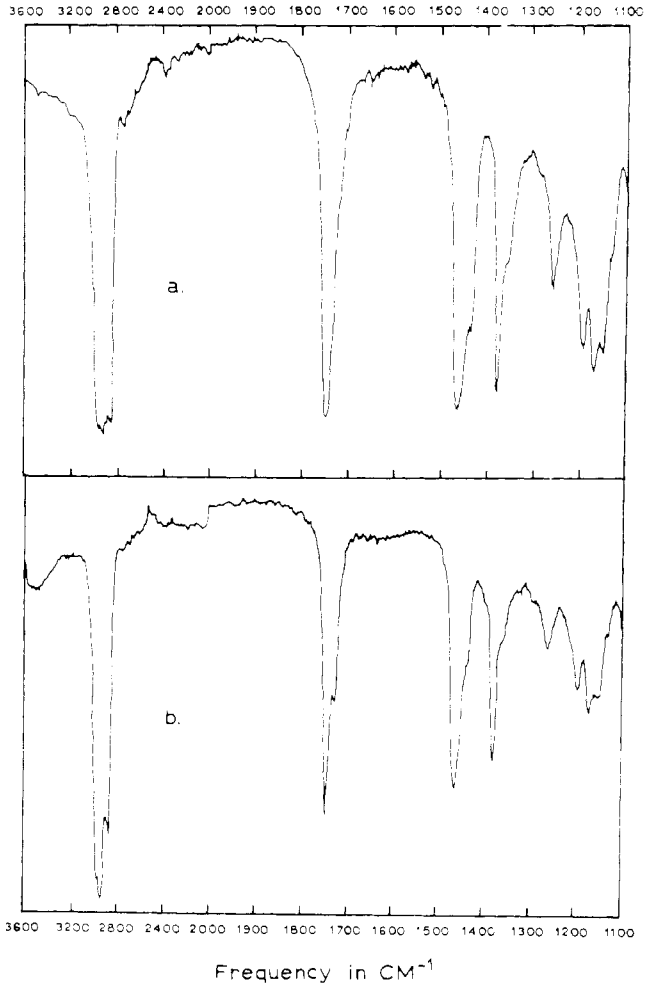



FIGURE 2: Infrared spectra. Thin film on KBR. (a) A methyl ester and (b) C methyl ester. D methyl ester gives a pattern like C. Note the split carbonyl band (1725 and 1745 $cm^{-1}$). Both spectra indicate a multiplicity of methyl branches (1380 $cm^{-1}$).

fact, a continuum of homologous branched-chain hydroxy acids in which an abrupt discontinuity leads to D as the lowest homolog.

Examination of the infrared spectra in Figure 2 suggests that A is also probably intimately related to C (and D). Thus, save for the absence of hydroxyl absorption and the split carbonyl band, the infrared spectrum of A is essentially superimposable on that of C. All three substances clearly have an abundance of methyl branches as indicated by the deep absorption band at 1380 $cm^{-1}$ (confirmed subsequently in other studies). On the basis of evidence reported here, we arrived at structural assignments for the four groups of carboxyl substituents, summarized in Table I. From this we see further that the intimate relationship not only involved A, C, and D; B, too, is included, and is undoubtedly a biosynthetic precursor of the other three substances. The biogenesis of the present multibranched acids is only now under investigation. However, it seems quite probable that the A and C series are both initiated through the condensation of successive propionate units ("$C_3$ homology") onto a palmitate (stearate) receptor. For propionate incorporation into mycobacterial lipids, a number of examples have, of course, already been well documented (Polgar and Robinson, 1951; Gastambide-Odier *et al.*, 1963a,b).

TABLE I: Structures of the Acyl Components of SL-I.

| Substance | Structure |
|---|---|
| A | $CH_3(CH_2)_{14}CH_2[CH(CH_3)CH_2]_nCH(CH_3)$-COOH<br>Average constituent, $n$ = 5–6; homology by $C_3$ units is seen from $n$ = 4–9. Some $C_2H_4$ homology suggested in the $C_{16}$ portion. |
| B | Palmitic acid principally, followed by stearic. Small amounts of homologs down to octanoic acid. Possible traces of tuberculostearic and eicosanoic acids. |
| C | $CH_3(CH_2)_{14}CHOH[CH(CH_3)CH_2]_nCH(CH_3)$-COOH<br>Principal component, $n$ = 7. Homology by $C_3$ units is seen from $n$ = 4 to $n$ = 8, 9. |
| D | $CH_3(CH_2)_{14}CHOH[CH(CH_3)CH_2]_2CH(CH_3)$-COOH<br>The parent homolog of the *C* series. Some $C_2H_4$ homology suspected in the $C_{16}$ portion. |

We propose the name "phthioceranic acids" for the multibranched acids represented in A, and "hydroxyphthioceranic acids" for the C–D series.[2] With recognition that the straight-chain moieties are predominantly derived from palmitic acid and that the hydroxyl group is *always* adjacent to the methyl branch most remote from the carboxyl function, the parent member of both series would be the $C_{25}$ acid corresponding to *D*—which is designated "$C_{25}$-hydroxyphthioceranic acid"; a principal component of A is the "$C_{31}$-phthioceranic acid", and the principal component of C is "$C_{40}$-hydroxyphthioceranic acid". These are comprised of a $C_{16}$ straight-chain portion condensed with three, five, and eight propionate units, respectively.

*Isolation of the Carboxylic Substituents from SL-I.* Carboxylic substances combined from several alkaline solvolyses of SL-I or desulfated SL-I were converted into the methyl esters, which were separated by column chromatography on silicic acid–Celite (2:1) with the results summarized in Table II. Figure 3 is a thin-layer chromatogram of the various fractions. For the chromatoplate, the sample numbers correspond to the fractions of Table II.

Components A and B of fraction 2 were separated by preparative thin-layer chromatography on silica gel (hexane–ether–acetic acid, 90:10:1, v/v); the products were eluted with $CHCl_3$–$CH_3OH$ (2:1, v/v) with final recovery of 17 and 13 mg of A and B, respectively. Examined in the context of Table I and the estimated mean molecular weights given there, the relative quantities of the methylcarboxylic esters recovered are summarized in Table III. Thus, the molar distributions of the acyl substituents in SL-I are probably A:B:C,D = 1:1:2. The amount of A accounted for is about 70% of that to be expected; this may be due to incomplete recovery of A, or to substitution of some D for A; but most probably to contami-

[2] The term "phthioceranic" was originally proposed by C. Coleman, who, in collaboration with Middlebrook *et al.* (1959) first described and then carried out purification and degradative studies on the sulfolipid fraction from H37Rv (unpublished data). In these studies, fatty acids recovered from saponification of "sulfolipid" were recognized as being branched, but were at that time only minimally characterized.

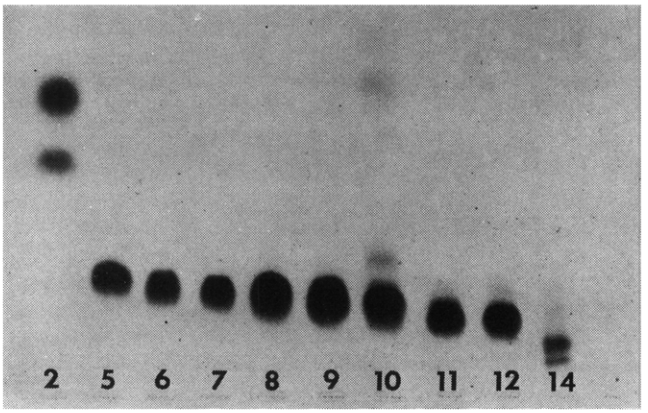



FIGURE 3: Thin-layer chromatogram of methyl esters recovered in silicic acid column chromatography. For details, see text and Table II.

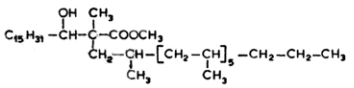
nation of SL-I samples by small amounts of the lipid designated SL-II. This is the most persistent contaminant of the principal sulfolipid, and lacks component A. We have not had sufficient homogeneous SL-II at hand to determine with confidence whether the deficiency in A is alleviated by an increase in one of the other components or if SL-II is indeed a triacyltrehalose 2′-sulfate.

*Mass Spectrometry of the Methyl Esters.* COMPONENT B. An earlier assignment of structure based on gas–liquid chromatography examination was confirmed. The fraction consists largely of methyl palmitate and stearate, with the $C_{16}$ acid predominant ($M^+$ peaks at 270 and 298, respectively). Minor peaks were found at *m/e* 312 and 326, possibly interpretable as being due to molecular ions from methyl tuberculostearate and eicosanoate, respectively. The remainder of the spectrum faithfully reproduced that to be expected from these simple esters.

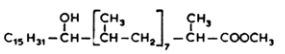
TABLE II: Silicic Acid–Celite Chromatography of Methyl Esters.[a]

| Solvent | Vol (ml) | Wt Eluted (mg) | Composition |
|---|---|---|---|
| (1) Hexane–ether, 98:2 | 75 | 0 | |
| (2) Hexane–ether, 98:2 | 25 | 30.8 | A + B[b] |
| (3) Hexane–ether, 98:2 | 600 | 0 | |
| (4) Hexane–ether, 95:5 | 50 | 0 | |
| (5) Hexane–ether, 95:5 | 50 | 1.4 | C |
| (6) Hexane–ether, 95:5 | 50 | 22.6 | C |
| (7) Hexane–ether, 95:5 | 50 | 18.8 | C |
| (8) Hexane–ether, 95:5 | 50 | 3.3 | C |
| (9) Hexane–ether, 95:5 | 50 | 1.9 | C |
| (10) Hexane–ether, 95:5 | 50 | 0.9 | C |
| (11) Hexane–ether, 95:5 | 50 | 2.0 | C[c] |
| (12) Hexane–ether, 95:5 | 150 | 2.4 | C[c] |
| (13) Hexane–ether, 95:5 | 150 | 0 | |
| (14) Hexane–ether–acetic, 90:10:1 | 150 | 2 | D |

[a] 88 mg of esters; 54 g of adsorbent. [b] A and B esters are subsequently separated by preparative thin-layer chromatography. [c] Mass spectrometry confirmed these to be among the lower homologs in the C ester series, but containing no D.



C-I



C-II

FIGURE 4: Structures considered for C methyl ester. C-I was ultimately rejected.

COMPONENT C. The Hydroxyphthioceranic Acids. The structure of the C ester series was revealed through infrared, mass, and proton magnetic resonance spectrometry and selective chemical transformations. With reference to Table II, the two major C fractions (6 and 7) and several minor ones (including 8 and 12) were examined. The mass spectrum (not shown) of the somewhat inhomogeneous fraction 6 and that of fraction 7 presented a cleavage pattern readily interpretable as being due to a series of methyl esters having a common, essentially constant $C_{16}$ portion bearing the hydroxyl group and a variable multibranched fragment having a homology pattern progressively increasing by 42 mass units (*i.e.* three-carbon units) and which included the carbomethoxyl function. The principal component gave a molecular ion peak ($M^+$) of 622 ($C_{41}H_{82}O_3$). To account for the very prominent peaks at *m/e* 88 and 101, an $\alpha$-methyl branch is required (*cf.* in particular, Ryhage and Stenhagen, 1963), and thus two structures, *viz.*, C-I and C-II of Figure 4 can be proposed, which are in accord with the observed fragmentation pattern.

Structure C-I is an unusual highly branched mycolic acid in which the $\alpha$ branch is longer than the $C_{18}$-straight-chain portion. If one postulates for this structure an initial mycolic type cleavage to palmitaldehyde and the heptamethyl heptadecanoate (Etemadi, 1965), all of the important peaks in the mass spectrum of the C esters can be accommodated (Ryhage and Stenhagen, 1963). Though this unusual mycolic acid with its quaternary carbon requires an almost unacceptable biogenetic scheme, we were unable to reject this alternative on the basis of mass spectrometric evidence alone. However, a combination of nuclear magnetic resonance spectra, deuterium exchange and chemical transformations of C ester ultimately showed that C-II was valid and that C-I was

TABLE III: Distribution of Acyl Substitutents in SL-I.

| Substance | Wt (mg) | Approx Av Mol Wt (Table I) | Mole |
|---|---|---|---|
| A[a] | 17.4 | 550–592 | 0.03 |
| B | 13 | 280 | 0.046 |
| C[a] | 53 | 662 | 0.085 |
| D | 2-3 | 412 | 0.005–0.007 |

[a] The combined sample of A esters had $\alpha_D^{21}$ +7.9° ($CHCl_3$, *c* 2.03). The relatively homogeneous portion of C esters (fraction 7 of Table II) had $\alpha_D^{21}$ +21.8° ($CHCl_3$, *c* 1.62). All chromatographically separated fractions were dextrorotatory.

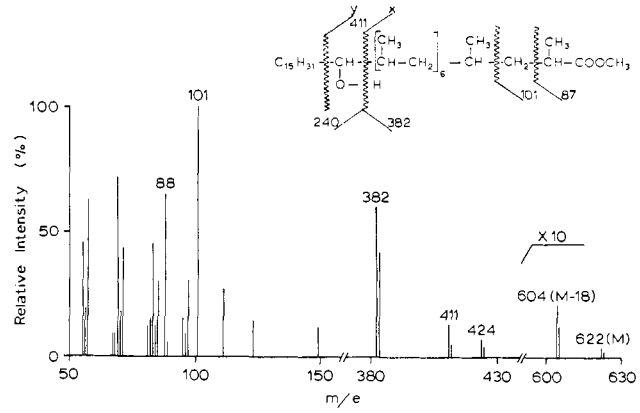



FIGURE 5: Mass spectrum of C ester (fraction 7, Table II).

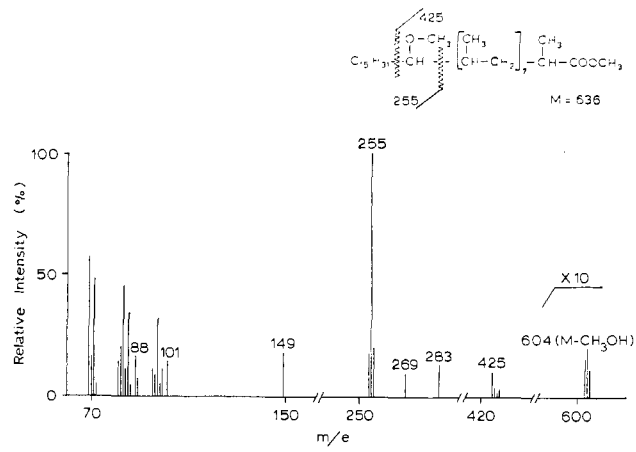



FIGURE 7: Mass spectrum of *O*-methyl ether of fraction 7, Table II.

unacceptable. The analysis of the C-series mass spectra (Figures 5, 6, 7, and 9) will accordingly be made on the basis of the phthioceranic acid structure (C-II) we have proposed.

The base peaks seen for any member of the C ester series are at *m/e* 88 and 101, trivial with respect to the more important features of the structures, but nevertheless fixing the environment of the ester function as bearing an $\alpha$-methyl branch. Transesterification of fraction 7 with absolute ethanol ($NaOC_2H_5$) transformed these peaks to the expected *m/e* 102 and 115, respectively. Referring to Figure 5, the most *prominent characteristic* of the C series spectrum is due to the cleavage of the molecule at *x* with hydrogen transfer to the charged right-hand fragment. The residue at the left is almost without exception of mass 240—*i.e.*, palmitaldehyde. Thus, the principal charged fragments seen for the series are always *M − 240*. Moreover, as part of the definitive pattern characteristic of the C series, a companion peak of almost equal intensity is observed at *M − 239*, presumably by transfer of a second hydrogen to the charged right-hand fragment. In Figure 5 these two peaks are at *m/e* 382 and 383. The molecular ion is at 622, but more prominent is the peak at M − 18 (604) arising from loss of a water molecule. Thus, fraction 7 is almost exclusively the $C_{40}$ homolog as depicted, although a small amount of the $C_{48}$ homolog is evident.

Associated with the M − 240 peak in the C series is a peak 29 mass units higher (*e.g.*, at *m/e* 411 for the present case) which is attributed to cleavage at *y*. For fraction 7, examination of the critical peaks at high resolution confirmed the elementary compositions assigned to the indicated fragments.

On conversion into the ethyl ester the appropriate changes in the mass peaks were obtained: the 382/383 and 411 peaks are shifted to 396/397, 425, whereas the molecular ion (622) and M − 18 (604) are shifted to 636 and 618.

The $C_3$ homology in the hydroxyphthioceranic ester series was discerned from the mass spectra of less homogeneous C-series fractions. The major fraction of C methyl ester (6 of Table II) contains the higher homologs, and gave prominent $M^+$ peaks at 622, 664, and 706 and the associated M − 18 peaks 604, 646, 688; *i.e.*, $C_{40}$-, $C_{43}$-, and $C_{46}$-hydroxyphthioceranic methyl esters. The characteristic M − 240 (and the associated M − 240 + 29) peaks corresponding to these homologs were obtained as the principal fragments: *m/e* 382:411, 424:453, and 466:495.

The lower members of the homologous C esters were found in the succeeding chromatographic fractions: fraction 8 gave the characteristic M − 240/M − 239 peak pairs at *m/e* 298/299, 340/341, and 382/383 (trace), corresponding to the $C_{34}$-, $C_{37}$-, and $C_{40}$-hydroxyphthioceranic acids. All the appropriate mass peaks characterizing this range of the C ester series were obtained. The mass spectrum of fraction 12, the last of the apparent C ester series obtained in the preparative chromatography is given in Figure 6. This is again an almost homogeneous fraction; the identifying peak pair at *m/e* 256/257 (and molecular ion peak at 496) establishes fraction 12 as the $C_{31}$-hydroxyphthioceranic ester. This fraction gives no evidence of a $C_{28}$ homolog. However, a very weak molecular ion peak at $M^+$ 524 and M − 18 peak at *m/e* 506 are taken as evidence for a trace of $C_2$ homology in the unbranched left-hand fragment (*i.e.*, transition from a palmitic to a stearic residue).

High-resolution spectrometry of this fraction confirmed the assignment given the peaks at *m/e* 256, 257, and 285 and of the molecular ion. Thus, the *continuum* of the C acid series seems to be established from the $C_{31}$- to the $C_{46}$-hydroxyphthioceranic acid (and perhaps beyond). Judged from the distribution of homologs discerned in thin-layer chromatography and mass spectrometry, the most abundant member is the $C_{40}$ acid.

*Methoxy $C_{40}$-Phthioceranic Ester.* Additional confirmatory evidence for the proposed C ester structures was derived from O methylation of the homogeneous fraction 7: the base peak for methylmethoxy $C_{40}$-phthioceranate is at *m/e* 255, arising from fragmentation as indicated in Figure 7. Though the cleavage pattern of the parent hydroxy ester afforded minimal information about the $C_{16}$ portion of the molecule.

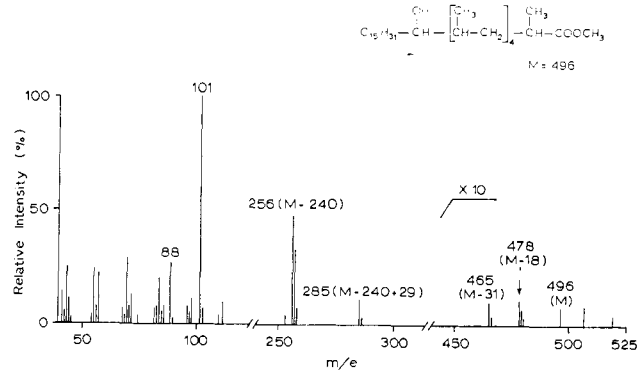



FIGURE 6: Mass spectrum of C ester (fraction 12, Table II).

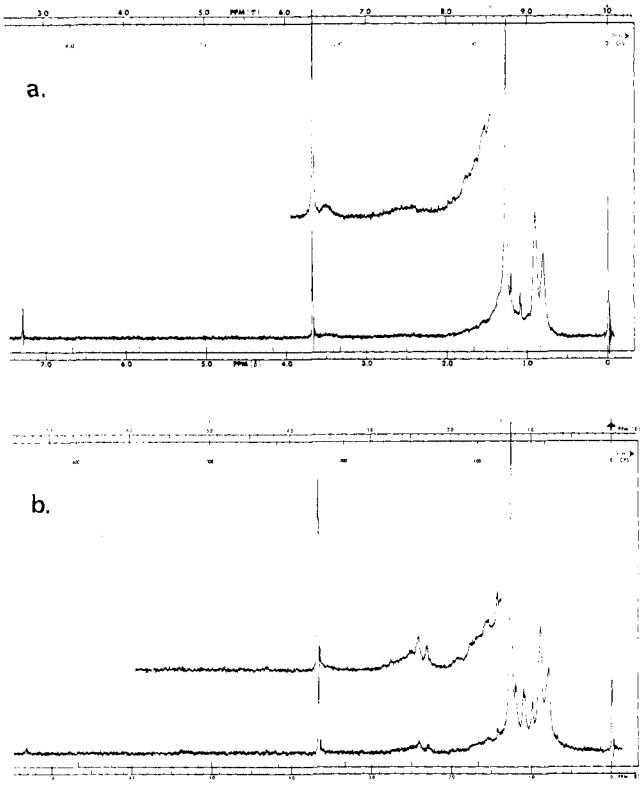



FIGURE 8: Nuclear magnetic resonance spectra (relative to tetramethylsilane). (a) $C_{40}$-Hydroxyphthioceranate and (b) $C_{40}$-oxophthioceranate.

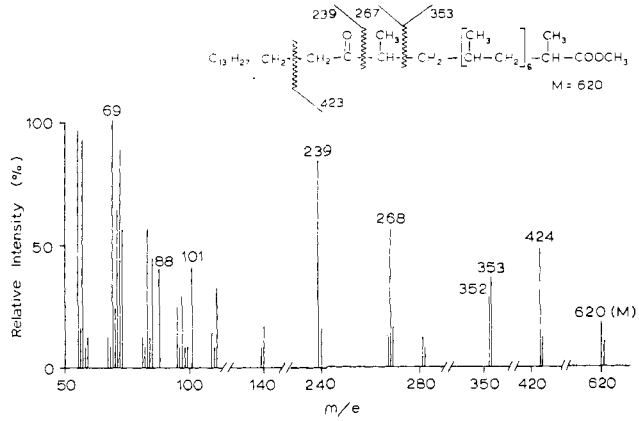



FIGURE 9: Mass spectrum of $C_{40}$-oxophthioceranate.

the influence of this moiety was *dominant* in the cleavage of the methyl ether. Some $CH_2$ homology in the palmitate moiety is suggested by small peaks at *m/e* 269 and 283. The characteristic 382/383 peak pair of the hydroxy ester is no longer evident. Instead it is replaced by a prominent but less intense peak at *m/e* 425, to which the indicated structure is assigned. The molecular ion (636) is not seen; instead M − 31 ($OCH_3$) and M − 32 ($CH_3OH$) are prominent at *m/e* 605 and 604.

It is appropriate to record at this point that the mixed methyl esters of carboxylic acids recovered from hydrolysis of *permethylated* (desulfated) sulfolipid also gave a similar spectrum, the base peak being at *m/e* 255, with a prominent peak at 425 as in the purified *O*-methyl C. Other peaks at *m/e* 383, 341, and 299, belonging to the same series with *m/e* 425 attested to the propionate homology. Prominent contributions due to the A ester series were easily recognized (*vide infra*).

*Nuclear Magnetic Resonance Spectra.*[3] The nuclear magnetic resonance spectrum of the $C_{40}$-hydroxyphthioceranate ester in $CDCl_3$ is shown in Figure 8a. The broad signal centered at about 150 cps and integrating to two protons is contributed by the hydroxyl function and by the tertiary hydrogen $\alpha$ to the ester carbonyl (Cason *et al.*, 1962). The signal from the hydroxyl proton is markedly attenuated after brief equilibration with $D_2O$. The band at 210 cps, unaltered by the equilibration is assigned to the single proton on the carbinol carbon. Of greatest interest are the large doublet at 48 and 54 cps and the smaller doublet downfield at 65 and 72 cps; both are due to secondary methyl groups, the larger doublet also embracing the signal from the primary terminal methyl group. The doublet displaced downfield is ascribed to a methyl group differing from the others and probably located on an $\alpha$ carbon. Estimation of the area under the methyl bands associated with the $\alpha$ carbon is not entirely accurate owing to the overlapping of the adjacent CH band. (In this respect, *cf.* Cason *et al.*, 1962, 1964.) The integration of the 48/54-cps doublet indicates it to represent about eight to nine methyl groups. According to the earlier disclosures of Cason with somewhat simpler systems, this integral would sum to a value which is in fact too high, the departure from linearity occurring after about four methyl groups (Cason *et al.*, 1962). A mean of about eight methyl groups is in good accord with the proposed structure: a terminal methyl group in the palmitic moiety, a terminal propionate containing the $\alpha$-methyl, and seven additional equivalent methyl branches. The alternative structure (C-I of Figure 4) has no secondary $\alpha$-methyl group to which the downfield doublet can be attributed, and accordingly the nuclear magnetic resonance spectrum gives good evidence for rejecting this alternative.

*Other Transformations with $C_{40}$-Hydroxyphthioceranate.* Still more evidence for rejecting the unusual mycolic acid structure C-I was obtained from deuterium-exchange studies and from selective oxidation of the hydroxy function. Although the deuterium exchange was incomplete, the mass spectrum of $C_{40}$-hydroxy ester which had been equilibrated with $CH_3OD$, $NaOCH_3$ showed the necessary transformations. The characteristic companion peaks *m/e* 382/383 of Figure 5 were shifted to 383/384; the associated 411 peak was also shifted to *m/e* 412; *m/e* 88 was transformed to 89. Since the *m/e* 88 and 382 peaks were both shifted by only one mass unit, it is clear that the deuteron is not present as a deuteroxy function but is on the $\alpha$ carbon. The M − $H_2O$ peak also shifts to *m/e* 605 instead of 604, further indicating the deuteron not to be on the carbinol oxygen. For reasons not yet understood, the major peak at *m/e* 101 was retained, a phenomenon we also observed with $\alpha$-deuterio A esters and in deuterium-equilibrated keto *C* (*cf.* below).

Sarett oxidation (Poos *et al.*, 1953) of the $C_{40}$-hydroxyphthioceranate gave the $C_{40}$-keto compound in high yield. The mass spectrum of Figure 9 was the most compelling argument for rejecting the multibranched mycolic ester structure of Figure 4, since it is incompatible with many of the important fragments obtained. When the keto ester was equilibrated with

[3] We acknowledge with thanks the kind advice of Dr. A. Gaudemer toward the interpretation of the nuclear magnetic resonance spectra described herein.

TABLE IV: Principal Fragments from Natural and Deuterated Keto C.

| Fragment | *m/e* Natural | *m/e* Deuterated |
|---|---|---|
| $C_{13}H_{27}CH_2CH_2COCH(CH_3)CH_2[CH(CH_3)CH_2]_6CH(CH_3)COOCH_3$ | 620 | 623, 624 |
| $C_{13}H_{27}CH_2CH_2CO^+$ | 239 | 241 |
| $C_{13}H_{27}CH_2CH_2COCH_2CH_3$ | 268[a] | 271 |
| $CH_2{=}C(CH_3)CH_2[CH(CH_3)CH_2]_5CH(CH_3)COOCH_3$ | 352, 353[a] | 352, 353, 354 |
| $CH_3COCH(CH_3)CH_2[CH(CH_3)CH_2]_6CH(CH_3)COOCH_3$ | 424 | 427, 428 |
| $CH_3COCH_2CH_3$ | 72 | 75 |

[a] The 268 and 352 fragments derive from a McLafferty rearrangement (Ryhage and Stenhagen, 1963) at the $\alpha$ carbon to the right of the keto carbonyl, with both fragments ultimately becoming charged. A fragmentation without hydrogen transfer gives rise to a prominent peak at *m/e* 353. A similar pattern is seen in the spectrum of keto D ester, a considerably lower homolog.

$CH_3OD$, $NaOCH_3$, the fragmentation pattern changed in the expected manner, although the conversion was again not complete. However, the sluggish deuterium exchange affected only the position $\alpha$ to the ester function; the exchanges adjacent to the keto group were complete. Table IV summarizes the fragmentation pattern obtained from the keto-ester and its deuterated analog.

The nuclear magnetic resonance spectrum of the $C_{40}$-oxophthioceranic ester is reproduced in Figure 8b. A comparison with the spectrum of the hydroxy ester (Figure 8a) indicates two new broad signals at 138 and 145 cps; these merge with the smaller broad signal centered at 150 cps seen before in the spectrum of the hydroxy ester and are attributed to the protons alpha to both carbonyl functions. In accord with this assignment, summation indicates these signals to be due to about four protons. Still another difference distinguishing the keto ester from the hydroxy compound is the new doublet downfield at 58 and 65 cps attributable to the secondary methyl group $\alpha$ to the keto function. This merges in part with the doublet at 65 and 72 cps previously seen in the hydroxy ester. The signals ascribed to the remaining methyl groups, *i.e.*, the doublet at 47 and 53 cps, integrate to somewhat more than seven methyl groups, in agreement with that demanded by the proposed structure: one terminal methyl in the palmitoyl portion and six substantially identical methyls in the branched moiety. This completes the evidence in support of the hydroxyphthioceranate acid structures for the C ester series.

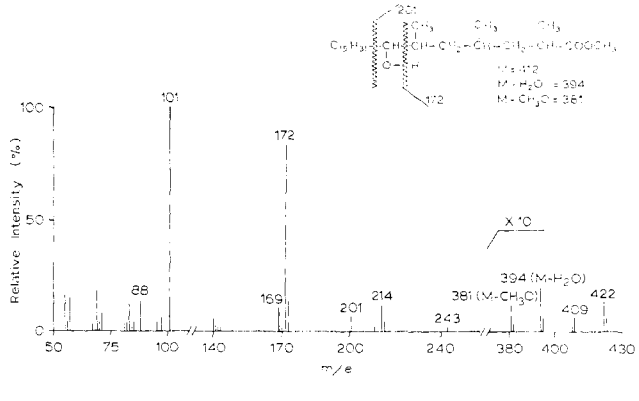



FIGURE 10: Mass spectrum of D ester.

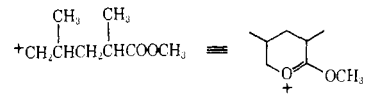
COMPONENT D. The minor fraction (14 of Table II) is composed of a number of substances separable by thin-layer chromatography. The principal component, however, is clearly recognizable as a member of the C ester series. The separation from the remainder of the C series in thin-layer chromatography is attributed to an abrupt discontinuity in the homology: *i.e.*, in the lowest member of the C series (Table II, fraction 12), the branched moiety of the molecule contains *five* propionate groups with only meager evidence of a molecule with four methyl branches. On the other hand, mass spectrometry indicates that the D ester possesses *three* methyl branches, only a trace of the next higher $C_3$ homolog being discernible; this explains the separation of D from C in thin-layer chromatography.

The mass spectrum of D ester (Figure 10) is not simply a displacement of the C series by 84 mass units. Certain significant changes occur, which may relate to the closer proximity of the two functional groups to one another. For the structural assignment as a $C_{25}$-hydroxyphthioceranate a molecular ion of 412 is expected. However, only M − 18 and M − 31 are seen at *m/e* 394 and 381, respectively. The dominant peaks found at M − 240/M − 239 which characterized the *C* series are replaced in D by a single M − 240 peak at *m/e* 172 with only a minor peak one unit higher. The M − 240 + 29 companion peak characteristic of the C series is present here at *m/e* 201. An additional fragment, not seen in the regular C series, apparently derives from loss of $CH_3OH$ from the latter, giving an important peak at *m/e* 169. Significant peaks at *m/e* 409 and 422 are interpreted as M − 18 and M − 31 peaks for a $C_{18}$ homolog in the unbranched moiety of D ester.

Sarett oxidation of the D ester afforded the corresponding $C_{25}$-oxophthioceranic ester, the mass spectrum of which further substantiates the structural assignment. The principal features are summarized in Table V.

COMPONENT A. We have described in part the isolation of A ester in the chromatographic procedure. The mass spectrum of this product (Figure 11) as compared to the other spectra is the least revealing. Viewed, however, in the context of the B and C ester structures, the implications seem considerably clearer. The immense base peaks at *m/e* 88 and 101 confirm the $\alpha$-methyl branching. The principal molecular ion is at *m/e* 480, interpretable as a $C_{31}$-phthioceranic ester (a palmitic moiety condensed with five propionic acids). An apparent homology by $CH_2$ units is suggested in the series; however, the 42 mass intervals are more prominent at $M^+$ 522 ($C_{34}$),

TABLE V: Principal Fragments in Mass Spectrum of Keto D.

| *m/e* | Fragment |
|---|---|
| 410 (M) (minor peak) | $C_{14}H_{29}CH_2COCH(CH_3)CH_2CH(CH_3)CH_2CH(CH_3)COOCH_3$ |
| 268 | $C_{14}H_{29}CH_2COCH_2CH_3$ |
| 267 | 268 − H |
| 239 | $C_{14}H_{29}CH_2CO^+$ |
| 214 | $CH_3CO[CH(CH_3)CH_2]_2CH(CH_3)COOCH_3$ |
| 199 | $^+CO[C(CH_3)HCH_2]_2CH(CH_3)COOCH_3$ |
| 143 (base peak) |  |

$M^+$ 564 ($C_{37}$), and $M^+$ 606 ($C_{40}$). As the probe temperature is raised, the lower members of the A series disappear and the spectrum shifts toward the higher homologs, the 606 member becoming the most prominent along with other molecular ions at 648 and even 690. An accurate assessment of the mean molecular weight is therefore not warranted on the basis of these data.

Rechromatography of a combined A–B ester mixture on silicic acid–Celite, using only hexane as elutrient allowed separation of the A from B and in addition, a partial separation of homologs within the A series. The first lipid effluents contained the higher homologs of A ($M^+$ principally 606, 648, and 690). These were followed by fractions containing the lower homologs, some of which were examined by mass spectrometry. All of the fractions gave evidence of $CH_2$ homology which can be explained by participation of both palmitic and stearic acid to give the more abundant members of the series ($C_{31}$, $C_{33}$, $C_{34}$), following which the palmitic contribution clearly predominates. The intermediate $C_{32}$, $C_{35}$, etc., components might be contributed by small amounts of myristate or eicosanoate condensed with the appropriate number of propionic units. The elemental composition of the principal molecular ions (based on the palmitic fragment) was confirmed by high-resolution mass spectrometry.

The nuclear magnetic resonance spectrum of the mixture of A esters confirmed many of the features implicit in the mass spectrum and suggests a mean structure as indicated in Table I. Thus, as in the C series, a single hydrogen (on the $\alpha$ carbon) was indicated by a broad shallow signal centered at about 160 cps. A downfield doublet at 65 and 72 cps was identical with the doublets seen in the C ester and keto C. The principal doublet at 47 and 53 cps integrated to somewhat more than *seven* methyl groups, suggesting a higher mean molecular weight than is surmised from examination of the mass spectrum. Deuterium exchange as described previously for the C ester gave the expected results: the base peak *m/e* 88 was shifted to *m/e* 89; but as in the C series, the *m/e* 101 peak was retained. All of the remaining principal peaks and molecular ions were shifted upward by one mass unit to give $M^+$ 481, 523, 565, and 607.

To summarize, mass spectrometry clearly demonstrates a simple relationship between the A and C series, the most prominent molecular ions characterizing the two series differing only by an oxygen atom. The nuclear magnetic resonance spectra, save for the hydroxyl signal in the C series, are essentially identical, as are the infrared spectra. Both series are dextrorotatory; and we have described earlier that a change in the molecular ratios of A and C appears to be the sole distinction between SL-I and SL-I′. These are compelling arguments for the structural similarity we propose.

The relation was unequivocally established through converting a portion of C ester (sample 6 of Table II; *i.e.*, consisting principally of the $C_{40}$-hydroxyphthioceranate) into the corresponding homologs in the A series. The C ester was dehydrated at reflux in acetic anhydride–$KHSO_4$ (Pudles and Lederer, 1951) with some difficulty. The olefinic product was separated from unreacted (but acetylated) starting material by vigorous alkaline saponification, and preparative thin-layer chromatography; it was then reesterified with diazomethane. The mass spectrum confirmed the transformation: the principal molecular ion had *m/e* 604 with no evidence of a 622 peak. Significant peaks ascribed to $C_2H_4$ and $C_3H_6$ homology were seen at *m/e* 632 and 646.

Catalytic hydrogenation of the anhydro ester gave the highest homologs in the phthioceranate series: mass spectrom-

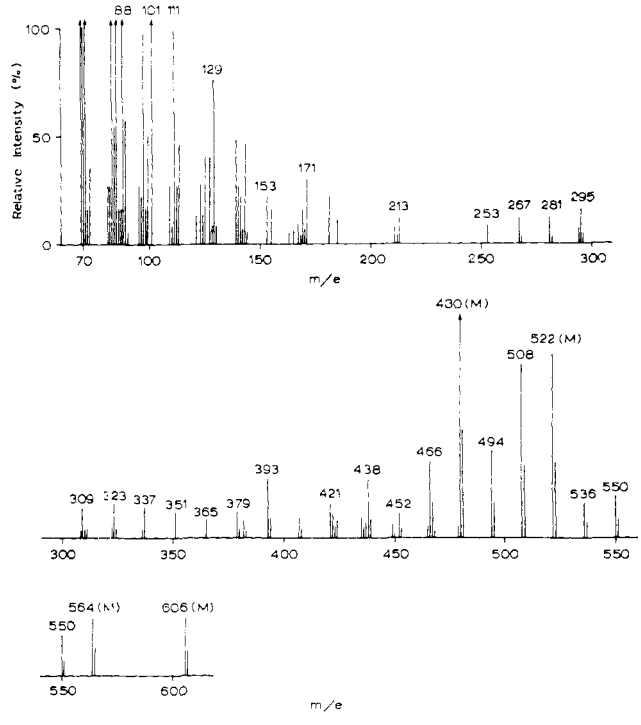



FIGURE 11: Mass spectrum of A ester.

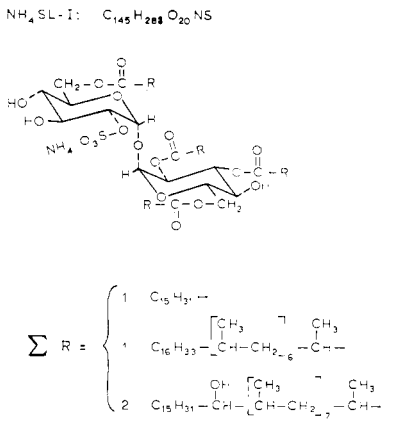


FIGURE 12: Revised structure of $NH_4SL$-I from *M. tuberculosis*, strain H37Rv.

etry showed that beyond the characteristic massive peaks at *m/e* 88 and 101, the critical regions for the two substances were identical in the positions and almost identical in the relative intensities of peaks at *m/e* 606, 620, 634, and 648, with the $C_3$ homology dominant. The authentic A ester had in addition the lower homologs down to $M^+$ 522 which disappeared while the spectrum was run several times. The latter were entirely absent in the derived A ester sample from the hydroxyphthioceranate. In addition, significant, but less prominent, identical peaks were duplicated in the spectrum of both samples at *m/e* 111, 129, 153, 171, 211, and 213. In completion of the comparison, the infrared spectra of the two preparations were entirely superimposable.

## Discussion

The present studies have led to structural expressions for the carboxylic acids substituted on the trehalose sulfate core of sulfolipid I of *M. tuberculosis*, strain H37Rv. These are principally palmitic acid, 2,4,6,8,10,12,14-heptamethyltriacontanoic acid, and 17-hydroxy-2,4,6,8,10,12,14,16-octamethyldotriacontanoic acid ($C_{37}$-phthioceranic and $C_{40}$-hydroxyphthioceranic acids). SL-I contains 1, 1, and 2 moles of the respective substances. Major homology, by 42 mass units, is seen in the two branched series; a minor homology of a second type is evidently associated with a substitution of $CH_3(CH_2)_{14}CHOH$ by $CH_3(CH_2)_{16}CHOH$.

*Infrared "Anomalies"*. We have mentioned a curious characteristic of the infrared spectra of the hydroxyphthioceranates: a split carbonyl band peaking at 1725 and 1745 $cm^{-1}$, the latter having about twice the intensity of the former. Etemadi (1965) has described a similar pattern (but with *reversed* intensities) in the spectra of the authentic mycolic acid esters and attributed the aberration to a chelation involving the ester carbonyl and the $\beta$-hydroxyl function (*cf.* also Asselineau, 1966). The peculiar pattern that we observed was for a time accepted as lending support to the very unusual mycolic acid structure C-I once considered (*vide supra*) for our hydroxy esters; namely, that the additional crowding on the $\alpha$ carbon interferes sufficiently with the intramolecular hydrogen bonding so as to reduce the contribution of the (chelate) form responsible for the 1725-$cm^{-1}$ absorption. With the rejection of the mycolic structure from consideration, we were, of course, compelled to seek another explanation, and found that *intermolecular* hydrogen bonding was responsible. The split carbonyl band was prominent in spectra of films or smears of the pure substances, where intermolecular interactions would be at a maximum; but the spectrum obtained of even relatively concentrated solutions in $CCl_4$ had only a single sharp band at 1745 $cm^{-1}$! Distortions in the carbonyl region seen in infrared spectra of the parent mycobacterial sulfolipids and the desulfated analogs (Goren, 1970a) are undoubtedly attributable to similar effects.

*Biogenetic Implications.* STEREOCHEMISTRY. The elaboration of the phthioceranic acids probably involves condensation of successive units of methylmalonyl-CoA (*cf.* Asselineau, 1966) onto palmitate and, to a lesser extent, stearate. This propionate incorporation is in accord with the hypothesis of Polgar and Robinson (1951) concerning the biogenesis of the phthienoic acids and was in fact clearly demonstrated for the mycocerosic acids (Gastambide-Odier *et al.*, 1963a). The most prominent member of the phthienoates is the $C_{27}$-acid, which is based on a stearate moiety condensed with three propionate groups. Moreover, a $C_{25}$-acid structured on a palmitate, *i.e.*, a 2,4,6-trimethyldocosenoic acid, was discerned by Asselineau (Ahlquist *et al.*, 1959) among the natural phthienoates, and was subsequently isolated and characterized by Cason *et al.* (1962). In resemblance to the phthienoates and their reduction products, but in contradistinction to the levorotatory mycocerosic acids, the branched carboxylic acids of the present study are all *dextrorotatory*. This was confirmed across the entire series of C esters, for D (the $C_{25}$ parent hydroxy acid), and for the poorly separated A series. Accordingly, the present branched acids are judged to be of the same configuration as the phthienoates; and thus the D ester (the $C_{25}$-hydroxyphthioceranate) and the $C_{25}$-phthienoic acid of Asselineau and of Cason (*vide supra*) are very likely convertible into the same $C_{25}$-phthioceranic acid.

Although Cason *et al.* (1962) prepared the hydrogenated $C_{25}$ acid, the stereochemistry at the newly generated asymmetric center involving the $\alpha$ methyl branch cannot be unequivocally assigned. It is not unreasonable to expect, however, that Cason's $C_{25}$-phthianoic acid[4] is probably a mixture and therefore in part diastereoisomeric with the present phthioceranic acids.

The stereochemistry of the phthienoic and mycocerosic acids has been summarized by Asselineau (1966): Ställberg-Stenhagen and Stenhagen (1947) related the levorotatory $\alpha$-methyl-branched acids to D-glyceraldehyde; degradative studies by Polgar (1953) showed that all the asymmetric centers of the levorotatory mycocerosic ("mycoceranic") acids are the D configuration. The asymmetric centers in the dextrorotatory phthienoic acids were established as being L (Ställberg-Stenhagen, 1954; Fray and Polgar, 1956). Our polarimetric evidence strongly suggests then that the phthioceranic acids are also of the L configuration—at most, if not at all of the methyl branches. The configuration of the hydroxyl function in the hydroxyphthioceranates is still unknown.

Implicit in the present derivation is a more complete molecular picture of the mycobacterial sulfolipids. In previous reports (Goren, 1970a,b) $NH_4SL$-I was characterized as a

[4] Traditionally, the (dextrorotatory) reduction products of the phthienoic acids have been designated phthianoates (Cason *et al.*, 1962; Asselineau, 1966). The nomenclature is somewhat confused, however, since a second class of acids has also been designated as *phthianoates* by Cason *et al.* (1964). These are also methyl-branched related substances, but levorotatory, and may in fact be enantiomeric with the phthioceranates (in the branched moiety).

2,3,6,6′-tetraacyltrehalose 2′-sulfate. The approximate molecular formula $C_{145}H_{275}NO_{20}S \pm$ (5C, 10H), molecular weight about 2384 was advanced for $NH_4$SL-I on the basis of microanalysis and molecular weight determination through utilization of the spontaneous desulfation which characterizes the ammonium sulfolipids. The apparent unsaturation was believed attributable in part to ketonic functions within the acyl groups; however, further study showed this not to be the case. The empirical formula is now revised to a more precise expression as indicated in Figure 12, and is in excellent accord with the now demonstrated distribution of the acids which have been described herein. The actual placement of the individual substituents has not yet been ascertained; this is an effort currently in progress.

References


Ahlquist, L., Asselineau, C., Asselineau, J., Ställberg-Stenhagen, S., and Stenhagen, E. (1959), *Ark. Kemi 13*, 543.

Asselineau, J. (1966), The Bacterial Lipids, San Francisco, Calif., Holden-Day, pp 102, 112, and 138.

Brennan, P. J., Lehane, D. P., and Thomas, D. W. (1970), *Eur. J. Biochem. 13*, 117.

Brimacombe, J. S., Jones, B. D., Stacey, M., and Willard, J. J. (1966), *Carbohyd. Res. 2*, 167.

Caserio, M., Roberts, J. D., Neeman, M., and Johnson, W. S. (1959), *Tetrahedron 6*, 36.

Cason, J., Lange, G. L., Miller, W. T., and Weiss, A. (1964), *Tetrahedron 20*, 91.

Cason, J., Tavs, P., and Weiss, A. (1962), *Tetrahedron 18*, 437.

Etemadi, A. H. (1965), Ph.D. Thesis, University of Paris.

Fray, G. I., and Polgar, N. (1956), *Chem. Ind. (London)*, 22.

Gastambide-Odier, M., Delaumeny, J.-M., and Lederer, E. (1963a), *Biochim. Biophys. Acta 70*, 670.

Gastambide-Odier, M., Delaumeny, J.-M., and Lederer, E. (1963b), *Chem. Ind. (London)*, 1285.

Goren, M. B. (1970a), *Biochim. Biophys. Acta 210*, 116.

Goren, M. B. (1970b), *Biochim. Biophys. Acta 210*, 127.

Middlebrook, G., Coleman, C. M., and Schaefer, W. B. (1959), *Proc. Nat. Acad. Sci. U. S. 45*, 1801.

Polgar, N. (1953), *Chem. Ind. (London)*, 353.

Polgar, N., and Robinson, R. (1951), *Chem. Ind. (London)*, 685.

Poos, G. I., Arth, G. E., Beyler, R. E., and Sarett, L. H. (1953), *J. Amer. Chem. Soc. 75*, 422.

Pudles, J., and Lederer, E. (1951), *Bull. Soc. Chim. Biol. 33*, 1003.

Ryhage, R., and Stenhagen, E. (1963), *in* Mass Spectrometry of Organic Ions, McLafferty, F. W., Ed., New York, N. Y., Academic Press, Chapter 9, p 399.

Ställberg-Stenhagen, S. (1954), *Ark. Kemi 6*, 537.

Ställberg-Stenhagen, S., and Stenhagen, E. (1947), *Ark. Kemi Miner. Geol. 24B*, No. 9.


# Protein Metabolism in Cultured Plant Tissues. Calculation of an Absolute Rate of Protein Synthesis, Accumulation, and Degradation in Tobacco Callus *in Vivo**

John D. Kemp† and Dennis W. Sutton

ABSTRACT: During exponential growth of cultured tobacco cells, amino acids present in the intracellular soluble pool were used to synthesize proteins. The mean pool size (0.1 μmole/g fresh weight for leucine) remained unchanged during growth. Radioactive leucine was taken up from the medium and incorporated into protein and the soluble leucine pool. Labeled leucine was not converted into other amino acids in the pool; all of the radioactivity found in protein was contained in leucine. When tissues were labeled on radioactive medium the specific radioactivity of the soluble leucine pool remained unchanged from at least 1 hr through 7-hr labeling and incorporation of radioactivity into protein was linear over the same time period. These data suggest that incorporation was a function of the specific radioactivity of the soluble leucine pool. A simple differential equation was developed to relate these data to observations of synthesis.

The rate of accumulation of total and soluble protein (7.1 mg of protein × g of protein $^{-1}$ × $hr^{-1}$) was measured directly and equaled the rate of growth of the tissue (7.1 mg of tissue × g of $tissue^{-1}$ × $hr^{-1}$). Kinetics of uptake and incorporation of radioactivity were used to calculate an absolute rate of protein synthesis. Comparing that rate (18 mg of protein × g of $protein^{-1}$ × $hr^{-1}$) to the rate of accumulation a rate of protein degradation was estimated (11 mg of protein × g of $protein^{-1}$ × $hr^{-1}$).

Although it is generally accepted that amino acids (not peptides) serve as precursors for protein synthesis, the role of the measurable amino acid pool within the cell remains unclear. The assumption that the measurable pool represents the protein biosynthetic precursor pool has been seriously challenged by findings in yeast (Cowie and McClure, 1959; Halvorson and Cohen, 1958), a slime mold (Wright and Anderson, 1960), mammalian cells (Kipnis *et al.*, 1961), and higher plant cells (Bidwell *et al.*, 1964). These studies indicate that only a small portion of the soluble amino acid pool is available for protein synthesis.

The objectives of this study were to determine: (1) whether

* From the Pioneering Research Laboratory, Crops Research Division, Agricultural Research Service, U. S. Department of Agriculture and Department of Plant Pathology, University of Wisconsin, Madison, Wisconsin 53706. *Received August 21, 1970.*

† To whom to address correspondence.