I. Carbon-13 Nuclear Magnetic Resonance Spectroscopy of Erythromycin Derivatives

II. Algebraic Chemistry

Thesis by

James Gregory Nourse

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1974

(Submitted September 12, 1973)

ACKNOWLEDGMENT

To my research advisors, Dr. John D. Roberts and Dr. Robert E. Ireland for helpful discussions and their encouragement of independent work.

To the National Institute of Health and California Institute of Technology for the traineeship and teaching assistantships.

To the investigators at Abbott Laboratories, particularly Dr. Richard S. Egan and Dr. James B. McAlpine for collaborative help and donation of most of the erythromycin samples.

To Dr. A. K. Mallams of Schering for the loan of the megalomycin A sample; Dr. Norbert Neuss of Eli Lilly for the erythromycin A sample; and Dr. K. Heusler and Professor E. F. Jenny of CIBA for the Lankamycin sample.

To Mrs. Ruth Stratton for her excellent typing and her tolerance of the gradual completion of the manuscript.

Dedication . . . . . most of the time.

ABSTRACT

I. Carbon-13 Nuclear-Magnetic Resonance Spectroscopy of Erythromycin Derivatives

CMR spectra were taken and assigned for a series of erythromycin derivatives. Interpretation was based on known effects on cmr spectra and conformational changes among the derivatives. While the results were consistent with a gross conformational homogeneity, some subtle differences were observed. Proposals were made concerning the conformational differences between erythromycin A and B and among the series of aglycones. A conformation was proposed for 5,6-dideoxy-5-oxoerythronolide B consistent with cmr and pmr results. It was also proposed that in the natural antibiotics the desosamine side-chain sugar is free to swing around to the side of the aglycone ring, while the cladinose side-chain sugar is relatively fixed. The results suggested and were consistent with this situation. In this series it appeared that cmr was most useful in ascertaining intramolecular hydrogen bonding and orientation of and around carbonyls. Diagnostic usefulness for the various ketal structures was also suggested. The results of an enriched propionate feeding experiment confirmed the propionate biosynthesis hypothesis and gave direct evidence for the origin of the individual aglycone carbons.

II. Algebraic Chemistry

Several mathematical-chemical problems were treated and general considerations discussed.

It was shown that the Polya method could be used to count stereoisomers in systems with chiral ligands. An iterative procedure was given to determine the distribution and symmetry of all isomers possible in any given system. An extension of the method was made to count possible one-step substitution reactions in benzene systems.

The algebraic nature of the synthetic design problem was discussed, based on some simple concepts from algebraic topology. It was shown that the symmetry group of a chemical structure relevant to the synthetic design problem can be considerably larger than the usual point group. Examples of this group were constructed along with some methods for taking advantage of this symmetry in hypothetical cases.

A complete algebraic description of pseudochirality was given. It was shown that pseudochirality results when a structure lacks a particular kind of symmetry based on the point group and operations which invert configurations of chiral ligands. Other related forms of stereoisomerism were similarly described. A derivation of the possible pseudochirality groups based on the usual point groups was given.

The treatment of through-space orbital interactions given by Goldstein and Hoffmann was described abstractly and extended to reaction problems. A ribbon topology was assigned a homomorphism from a group of mode change operations to a group of interaction change operations. The concept of topologically equivalent transition state geometries was shown to correspond to group isomorphisms. The action of the mode change group on a dynamic symmetry group for 1,5-hexadiene was given which interchanged allowed and forbidden reactions. It was suggested that the existence of allowed and forbidden

reactions could be considered as a symmetry of a chemical structure.

The applicability of concepts of categorical algebra to biological, physical and chemical problems was discussed. Basic categorical concepts were discussed and it was shown that the center and inner automorphism groups are functorial constructions on the subcategory of Grp including only epi arrows. General features of application of categorical concepts outside pure mathematics were compared. The functorial nature of the problems treated in earlier sections was discussed. The similarity between isomerization of triaryl methanes and tris-metal chelates observed by Gust and Mislow was expressed as an isomorphism of the dynamic symmetry groups. The use of groupoids in problems of isomerizations between different geometric forms was demonstrated. It was shown that groupoids can be considered as representations of wreath products. A dynamic symmetry group for cyclobutene-butadiene interconversions was derived.

CONTENTS

Table 0

Macrolide Derivatives Studied

(1) Erythromycin A

(2) Erythromycin B

(3) Erythromycin C

(4) 6-deoxyerythronolide B

(5) Erythronolide B

(6),(13) 3-O-α-L-mycarosylerythronolide B

(7) 5-O-β-D-desosaminyl-3-O-α-L-mycarosylerythronolide B

(8) 5-deoxy-5-oxoerythronolide B

(8h) 5-deoxy-5-oxoerythronolide B-6,9-hemiketal

(9) 5,6-dideoxyerythronolide B

(10) oleandomycin

(11) 11-acetylerythronolide B

(12) 11-acetyl-6-deoxyerythronolide B

(14) 5-O-β-D-desosaminylerythronolide B

(15) 3,5-diacetylerythronolide B

(16) 3,5,11-triacetylerythronolide B

(17) 9-S-9-dihydroerythronolide B

(18) de-N-methylerythromycin A

(19) de-N-methylerythromycin B

(20) 3'-dedimethylamino-3',4'-dehydroerythromycin A

(21) anhydroerythromycin A

(22) erythralosamine

(23) 8,9-anhydroerythromycin B-6,9-hemiketal

(24) 8,9-anhydroerythronolide B-6,9-hemiketal

(25) megalomycin A

I. CARBON 13 NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY OF ERYTHROMYCIN DERIVATIVES

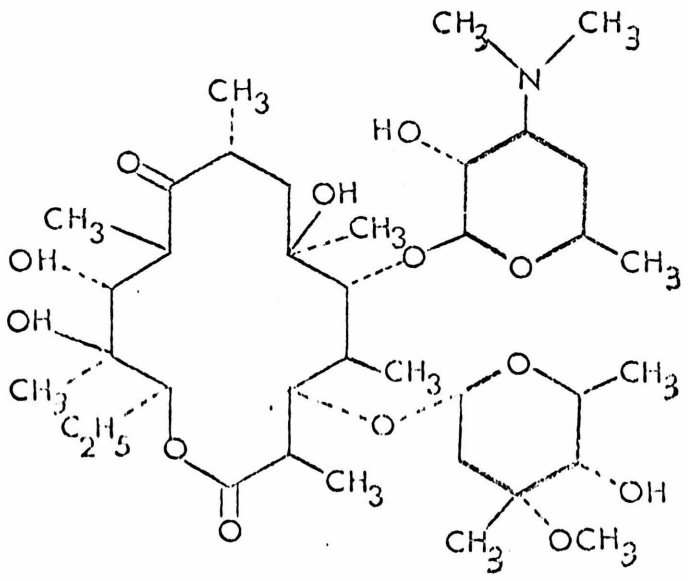
A. Introduction

A study of the ^{13}C nmr of erythromycin and several of its derivatives was undertaken for two reasons. First, carbon chemical shifts could be useful in further determinations of the conformations of these antibiotics and the changes in these shifts could be indicative of the conformational changes in going from one derivative to another. Secondly, assigned spectra are necessary for biosynthesis studies using ^{13}C enriched precursors.

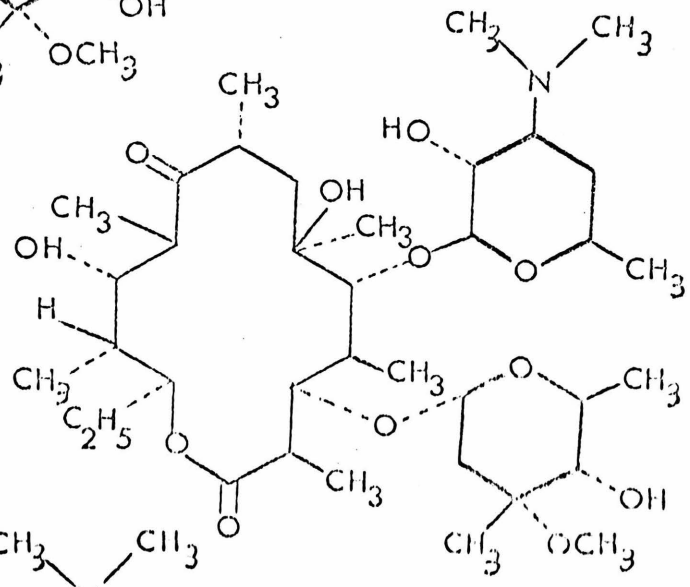
Erythromycin A(1) is one of the macrolide family of antibiotics which also includes oleandomycin, picromycin, and magnamycin among others. The first isolation of "erythromycin" (actually a mixture of erythromycins A, B and C) and discovery of its antibiotic properties was in 1952[(1)]. Separation of the mixture and determination of the absolute structure came much later due to the complexity of the molecule[(2)] (Figure 1). Final proof came in 1965 with the results of a single-crystal x-ray determination of erythromycin A hydroiodide[(3)]. Erythromycin A is the major component of the antibiotic produced for commercial use. Erythromycin B differs from Erythromycin A by the lack of the hydroxy group at the 12-position. Erythromycin C lacks the O-methyl group at the 3" position.

The structures of the other 12- and 14-membered ring macrolides are all similar to erythromycin. This observation was made by W. D.

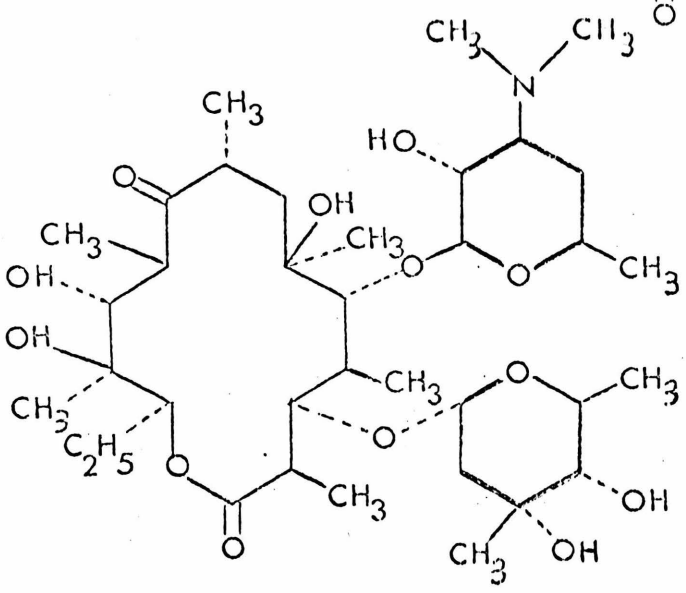
FIGURE 1



Erythromycin A (1)



Erythromycin B (2)



Erythromycin C (3)

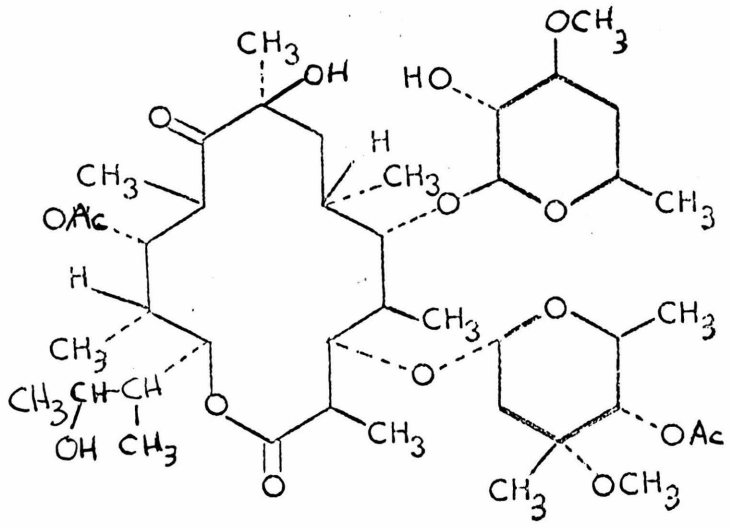
Celmer before many of the final structural determinations had been completed and prompted him to propose a configuration model for the series[4]. This model was successful in predicting and correcting configurational assignments of macrolides discovered before and after the model was proposed. In Figures 2 and 3 are standard and projected drawings of some of these macrolides.

B. Biosynthesis

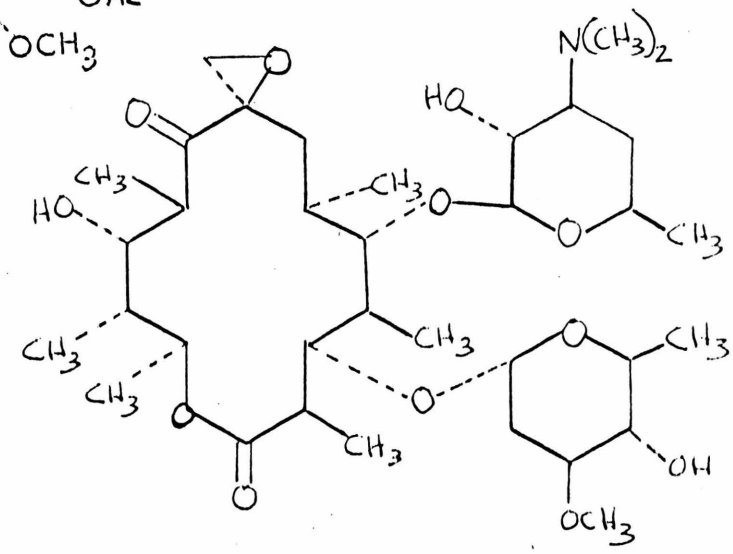
Erythromycin has been shown to originate biosynthetically from seven propionate units[5], verifying a prediction by Gerzon et al. in 1956[6] (Figure 4). The available evidence indicates that the lactone originates from a chain built from a starter propionate unit and six extending methyl-malonate units. This conclusion was based on the results of ^{14}C feeding experiments and subsequent degradation and analysis[5].

More recent work has led to the elucidation of some of the final steps in the biosynthetic sequence leading to the three known erythromycins (Figure 5). 6-Deoxyerythronolide B (4) is converted to erythronolide B (5) [7] by appropriate oxidation. Erythronolide B is then converted to 3-O-α-L-mycarosylerythronolide B (6) by glycosidation. From here it is postulated that the major change is a conversion to 5-O-β-D-desosaminyl-3-O-α-L-mycarosylerythronolide B (7), an intermediate as yet unobserved. This intermediate can then be converted to erythromycin B (2) or C (3), both of which can be converted to erythromycin A (1). These studies[8] have also led to the isolation and characterization of several shunt metabolites, two of which,

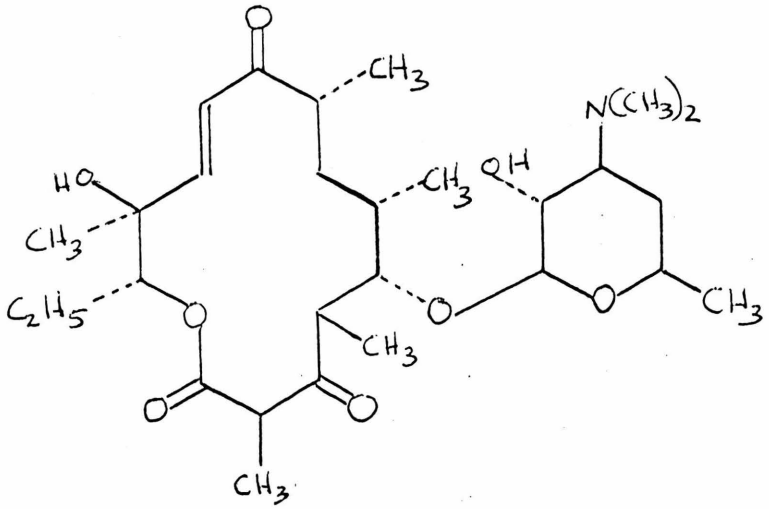
FIGURE 2



Lankamycin



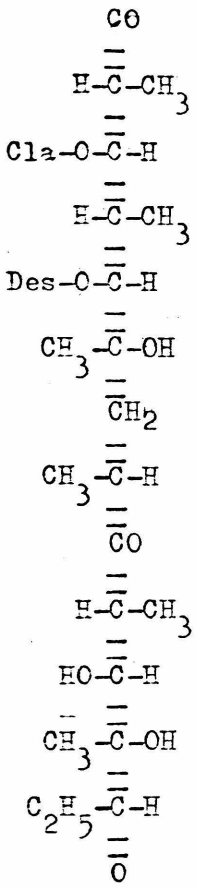
Oleandomycin



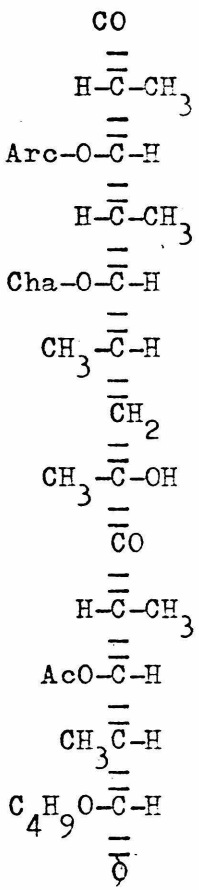
Picromycin

FIGURE 3

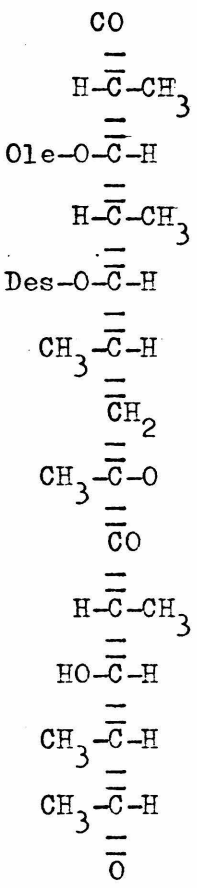
Projected Macrolides.



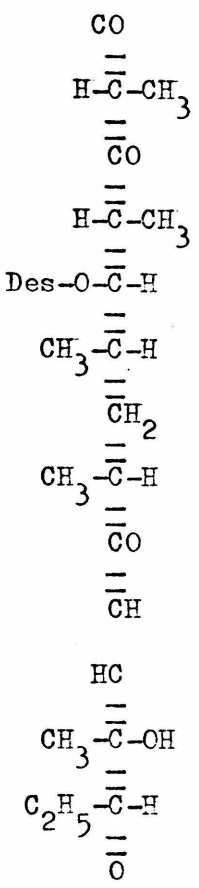
Erythromycin



Lankamycin



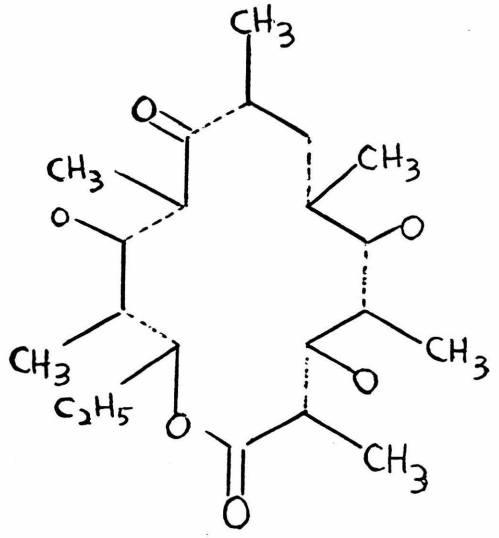
Oleandomycin



Picromycin

FIGURE 4

Propionate Units in Erythronolide Biosynthesis



Final Steps in Erythromycin Biosynthesis

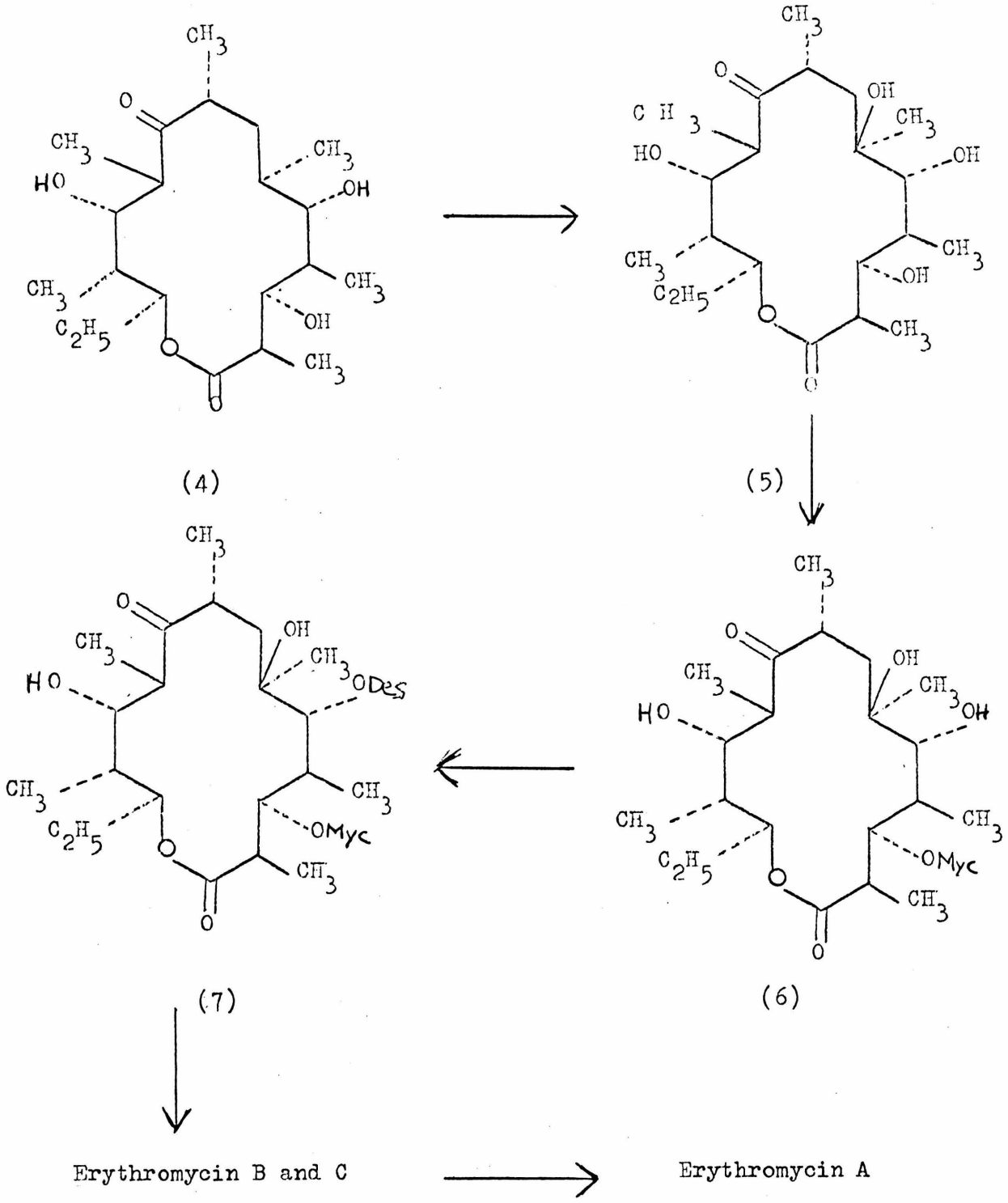


FIGURE 5

5-deoxy-5-oxoerythronolide B (8) and 5,6-dideoxy-5-oxoerythronolide B (9) are shown in Figure 6. Compounds (1-2,4-6,7-9) make up part of the collection on which the ^{13}C nmr studies were done.

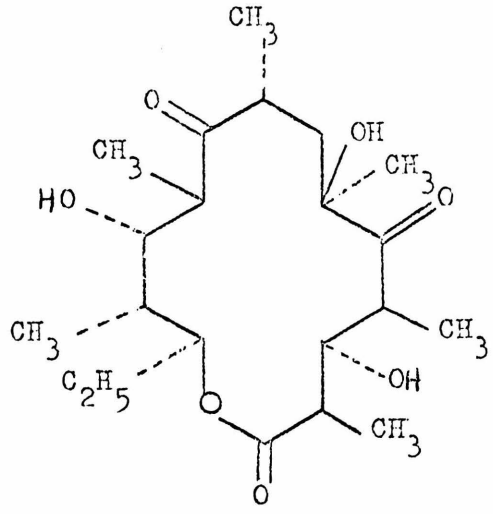
C. Conformational Studies

A considerable amount of work has been done to determine the conformation of erythromycin antibiotics and derivatives in solution. The principal methods used have been proton nmr[9] (pmr) and circular dichroism (CD)[10].

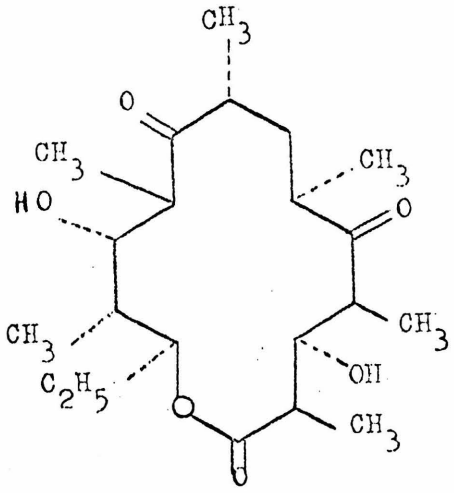
The first conformation proposed for a macrolide was for oleandomycin (10) by Celmer[11]. He fit the diamond lattice conformation proposed for cyclotetradecane by Dale[12] to the lactone of oleandomycin minimizing unfavorable interactions (Figure 7). As it turned out, the analogy to the hydrocarbon was good, but the wrong diamond lattice conformation was predicted.

The bulk of the conformational studies were done by investigators at Abbott[13]. They examined an extensive series of erythromycin derivatives by pmr in deuteriochloroform and pyridine solutions. Generally, the vicinal couplings between adjacent protons on the aglycone ring were found to be invariant to both temperature and solvent changes over the full range of derivatives studied. Most of the differences that were observed could be explained by conformational changes caused by addition or removal of side chain sugars or acetoxy groups in the various compounds. With this strong evidence for a rigid rather than mobile system, a conformation was proposed based on the Karplus relation for vicinal proton couplings[14]. The proposed conformations

Shunt Metabolites from Erythromycin Biosynthesis



5-Deoxy-5-oxoerythronolide B (8)

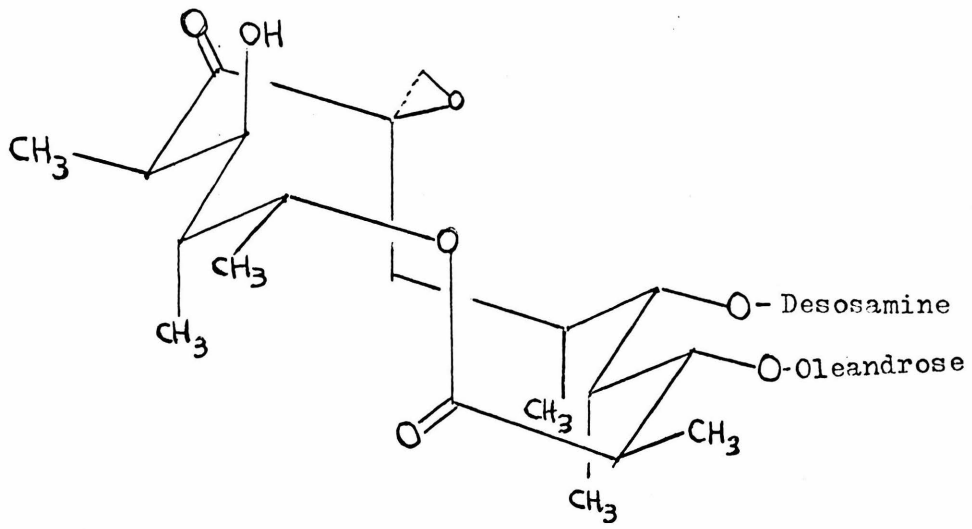


5,6-dideoxy-5-oxoerythronolide B (9)

FIGURE 6

FIGURE 7

Proposed Oleandomycin Conformation

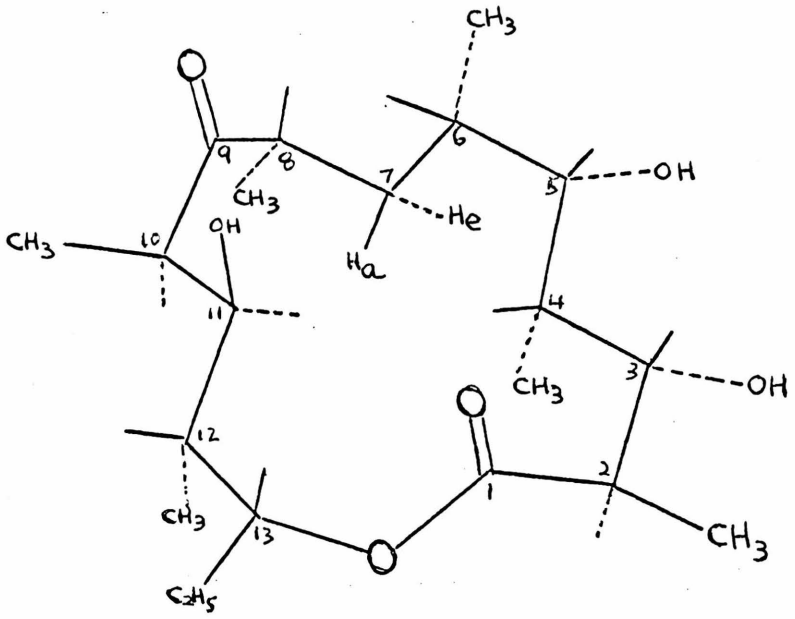


for the aglycones (sugars removed) and the parent glycosidated antibiotics are shown in Figure 8. In Table 1 is a list of representative vicinal couplings in a series of derivatives. These are split into those which exist in the aglycone conformation and those which exist in the erythromycin conformation. These conformations resemble an alternant diamond lattice arrangement as shown in Figure 9. This conformation would be of higher energy in the parent hydrocarbon, cyclotetradecane, due to steric interactions of protons directed into the center of the ring. The proposed conformation also resembles very closely the solid-state conformation as determined by x-ray crystallography[3].

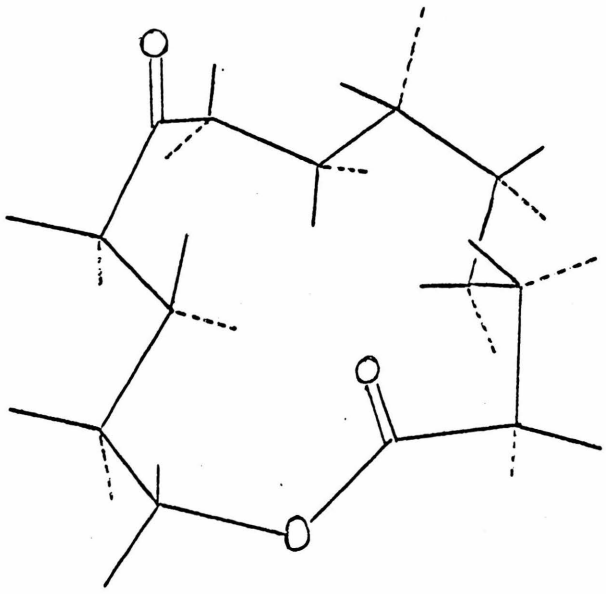
A few deviations from this conformational model were observed, however. Egan[15] noted some significant, unexpected differences in the chemical shifts and vicinal couplings between erythronolide B and 11-acetylerythronolide B (Figure 10 and Table 2). Of particular interest were the chemical shifts of H-8 and the 7-8 vicinal couplings. Further, a temperature dependence of the 7-8 couplings was noted. Egan[15] postulated a conformational change in this region of the ring since a direct shielding or deshielding effect of the distant 11-acetyl group seemed unlikely. All of the data were consistent with the following proposal. There are two conformations of these aglycones which differ in the 6-9 region of the ring, as shown in Figure 11 (conformation A is the one shown earlier in Figure 8). They are populated differently in the various derivatives. Presumably, 11-acetylerythronolide B exists in conformation B to a greater extent than do any of the other derivatives. It was further noted that conformation B has the 6-9 region in a pseudo five-membered ring which resembles very closely the

FIGURE 8

Proposed Conformations[a]



A. Aglycone Conformation



B. Erythromycin Conformation

a. Reference 15

Table 1

Representative Vicinal Proton-Proton Couplings[b]

	erythromycin conformation			aglycone conformation		
	(1)	(2)	(15)	(4)	(5)[a]	(6)
$J_{2,3}$	9.0	8.3	10.9	10.5	10.2	10.5
$J_{3,4}$	2.0	ca.1	1.4	1	1.3	ca.0
$J_{4,5}$	7.5	7.0	4.6	2.5	2.9	2.5
$J_{5,6}$	-	-	-	4.7	-	-
$J_{6,7a}$	-	-	-	4.7	-	-
$J_{6,7e}$	-	-	-	10.2	-	-
$J_{7a,7e}$	15.0	15.0	14.6	15.0	14.5	14.6
$J_{7a,8}$	10.0	10.0	9.4	13.0	6.7	11.4
$J_{7e,8}$	3.0	3.0	3.3	4.0	7.4	2.2
$J_{10,11}$	1.5	ca.1	1.5	2.0	2.0	ca.1
$J_{11,12}$	-	9.8	10.3	10.2	9.8	10.0
$J_{12,13}$	-	ca.1	0.9	1.5	1.2	1
$J_{13,14a}$	10.6	9.0	8.4	8.9	8.8	9.5
$J_{13,14e}$	2.4	5.5	5.6	4.6	6.6	4.8
$J_{14a,14e}$	ca.15			14.0	14.0	14.0

[a]Pyridine solvent

[b]Data taken from Ref. 15.

Compound names are in Table 0

FIGURE 9

Erythronolide B on Alternant Diamond Lattice Conformation

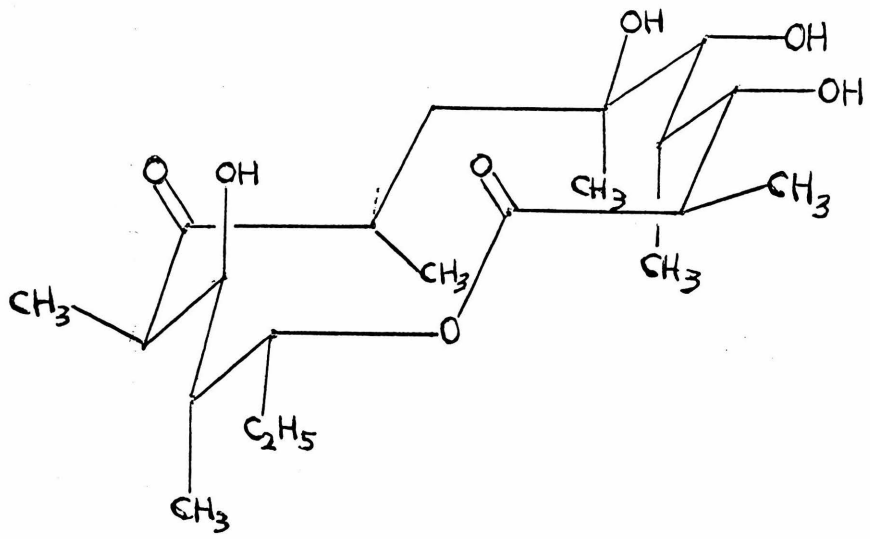
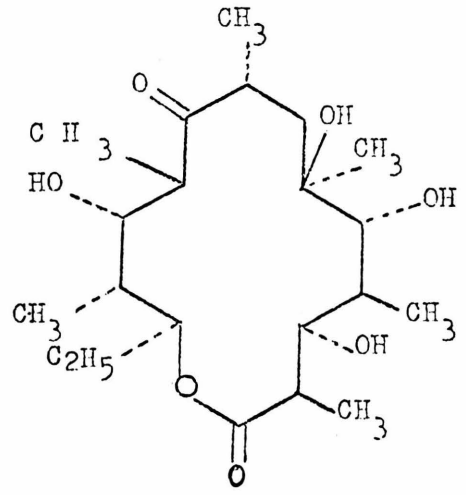
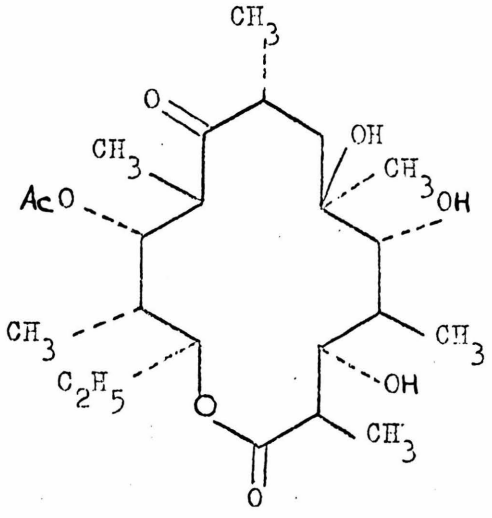


Figure 10



Erythronolide B (5)



11-acetylerythronolide B (11)

Table 2

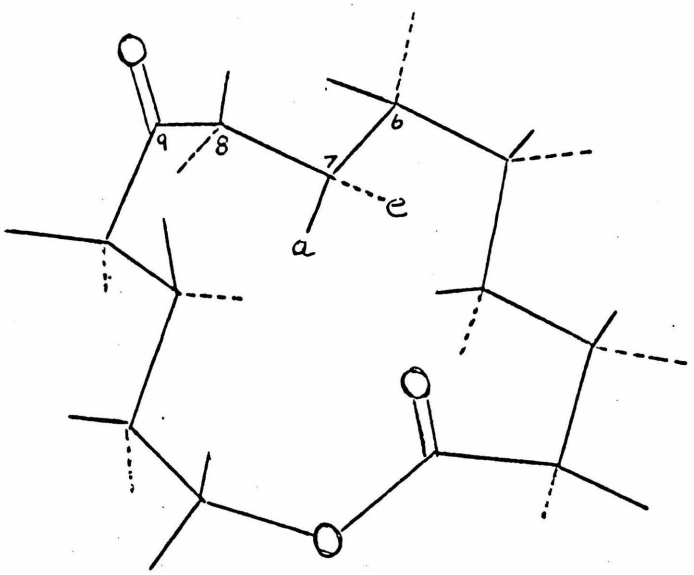
Comparison of pmr Data for Erythronolide B(5) and 11-Acetylerythronolide B(11)

	(5)[a]	(11)[a]		(5)[a]	(11)[a]
H-2	2.94	2.94	$J_{2,3}$	10.2	10.1
H-3	4.08	4.24	$J_{3,4}$	1.3	1.5
H-4	2.45	2.33	$J_{4,5}$	2.9	2.0
H-5	4.07	4.34	$J_{7a,7e}$	14.5	14.5
H-7a	2.23	2.17	$J_{7a,8}$	6.7	3.4
H-7e	1.64	1.69	$J_{7e,8}$	7.4	11.5
H-8	3.09	3.59	$J_{10,11}$	2.0	1.5
H-10	3.08	3.14	$J_{11,12}$	9.8	9.6
H-11	4.25	5.27	$J_{12,13}$	1.2	1.4
H-12	1.81	2.04	$J_{13,14a}$	8.8	8.0
H-13	5.70	5.25	$J_{13,14e}$	6.6	5.7
H-14a	1.74		$J_{14a,14e}$	14.0	
H-14e	1.55				

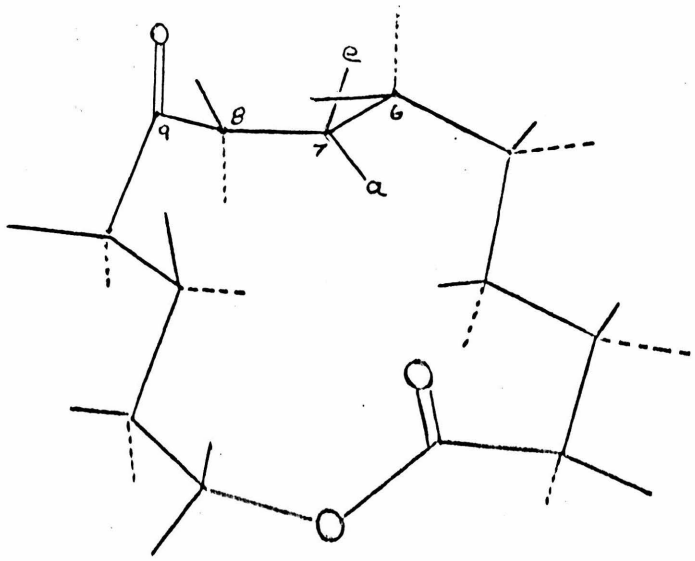
[a]Pyridine solution. All data from Ref. 15.

FIGURE 11

Equilibrating Aglycone Conformations[a]



CONFORMATION A



CONFORMATION B

a. reference 15

hemiketal that a number of erythromycin derivatives form. In Figure 12 are shown 5-deoxy-5-oxoerythronolide B and its hemiketal form which exist in a 3:1 ratio. Much of the observed chemistry of erythromycin appears to proceed through such a hemiketal[16]. In the case of 11-acetylerythronolide B, the driving force for the conformational change is unclear; however, steric effects or lack of hydrogen bonding between the 11-acetyl group and the 9-ketone may contribute. In most of the other derivatives any unexpected deviations from the proposed model conformations were confined to the 6-9 region.

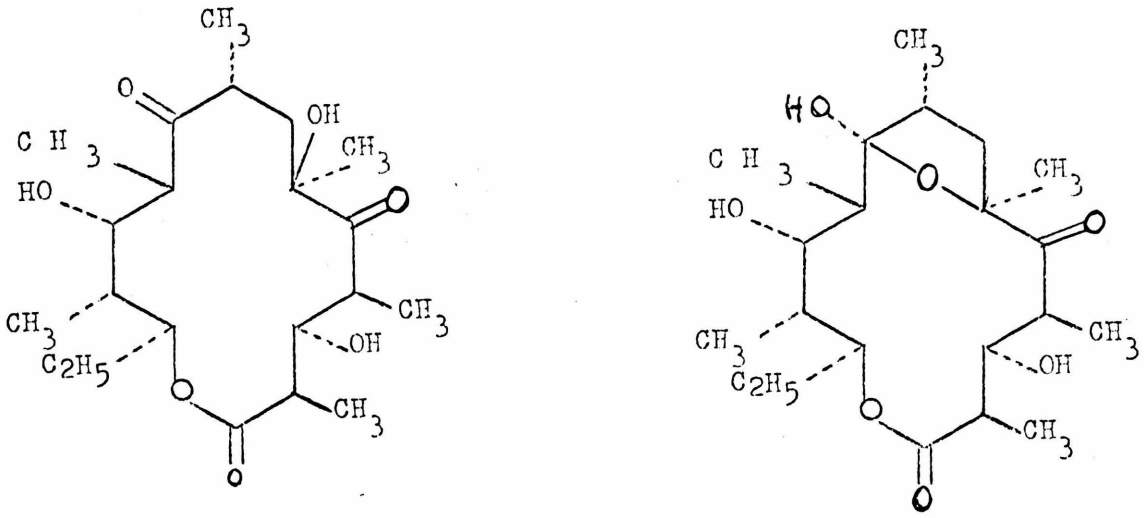
D. ^{13}C Nuclear Magnetic Resonance Spectroscopy

Carbon-13 nuclear magnetic resonance spectroscopy capabilities have increased enough over the last few years to the point where a study of compounds as complicated as macrolide antibiotics is feasible and practical. The most significant advance has been the development of pulsed Fourier-transform techniques. Samples with concentrations as low as 0.1 M can be run in reasonable times.

There is now a large body of ^{13}C chemical shift literature[17]. In most typical organic molecules carbon shifts and comparisons between chemical shifts in similar molecules can be rationalized in terms of substituent effects and steric effects. Substituent effects generally have the following pattern. In a linear alkane, substitution for hydrogen causes a downfield shift at the α-carbon, a downfield shift at the β-carbon and an upfield shift at the γ-carbon. Beyond that the shifts are variable but usually small. The size (and occasionally the sign) of these effects vary from substituent to substituent.

FIGURE 12

Keto-hemiketal Equilibrium of 5-Deoxy-5-oxoerythronolide B



(8) 3 : 1 (8h)

Typical examples are given in Table 3. It has been shown by Grutzner et al.[18] that substituent effects are not additive but rather decrease with increasing substitution. In a heavily substituted molecule the β-effect can be reduced by as much as 75%.

The steric effect is the upfield shift that is observed for carbons which are sterically compressed. It can actually be considered a remote substituent effect and in fact the γ-effect mentioned above appears to actually be a steric shift[19]. The origin of this shift has been discussed from a theoretical standpoint by Grant[20]. The magnitude of this effect has been determined for a number of systems. Some representative values for cyclic structures are given in Table 4. The major contribution to these steric shifts seems to be 1,3 diaxial interactions.

One other observed effect on carbon shifts is relevant to this study. It has been noted that hydrogen bonding causes a downfield shift in carbonyl derivatives. This has been observed as a solvent effect[22] and intramolecularly[23].

These are the known effects on carbon chemical shifts that seem to play the largest role in interpretation of macrolide spectra. The body of available knowledge on cmr is certainly much larger.

E. Model Systems

There do not seem to be any reported cmr studies of systems which model macrolides in complexity. However, a number of simpler systems have been studied and the results are relevant to this one.

Table 3

Typical Substituent Effects in Linear Alkanes

	Me[a]	OH[b],1°	OH[b],2°
α-effect	-9.1 ± 0.1	-48.3	-44.5
β-effect	-9.4 ± 0.1	-10.2	-9.7(C1) -7.4(C3)
γ-effect	+2.5 ± 0.1	+ 5.8	+3.3
δ-effect	-0.3 ± 0.1	-0.3	-0.2
ε-effect	-0.1 ± 0.1	-0.1	-0.2

[a] Ref. 17a, p.58

[b] Ref. 17a, p.142.

A negative shift is <u>downfield</u> (internal CS_2 reference)

Table 4

Typical Substituent Effects in Cyclohexanes

	eq-Me[c]	ax-Me[c]	eq-OH[d]	ax-OH[d]
α-effect	-5.6 ± 0.2	-1.1 ± 0.4	-41 to 44	-35 to -40
β-effect	-8.9 ± 0.1	-5.2 ± 0.3	-6.5 to -8	-7.6 to -5.3
γ-effect	0.0 ± 0.6	+5.4 ± 0.2	+1 to 1.5	+6.6 to +7.8
δ-effect	+0.3 ± 0.2	+0.1 ± 0.1	+1 to 1.8	+0.7 to +2.6

[c] Ref. 17a, p.65

[d] Ref. 17a, p.106

An extensive series of methylated aliphatic and alicyclic hydrocarbons, alcohols and ketones have been studied by several groups[24]. In Table 5 are listed data for cyclohexane derivatives. Besides the effects mentioned above a number of comparisons are relevant. In derivatives with predominantly axial hydroxyl groups (cis-2-methylcyclohexanol and trans-3-methylcyclohexanol) the carbon with the hydroxy group is shifted upfield relative to one with an equatorial hydroxy group. A similar effect is noted in cis-1,2-dimethylcyclohexane and trans-1,3-dimethylcyclohexane. The methyl carbons are also shifted upfield in these compounds due to a steric shift. Comparing cyclohexanols with the corresponding cyclohexanones shows a downfield shift at all the saturated carbons except the methyls and the position directly across from the oxygenated carbon when the hydroxy group is oxidized to a ketone. In a temperature study, Anet[25] has determined the chemical shifts of axial and equatorial methylcyclohexane. There is a 5-6 ppm upfield shift in the resonances of the 3 and 5 positions and of the methyl carbon on going from the equatorial to the axial conformer. This provides another measurement of the steric shift mentioned earlier.

A number of carbohydrate spectra have also been reported. Dorman and Roberts[26] studied a series of pentose and hexose aldopyranoses for which some representative data appear in Table 6. Two consistent effects deserve to be mentioned. First, the anomers with an axial hydroxyl (α-glucose and α-galactose) have all carbons except C-4 and C-6 shifted upfield relative to the β-anomer. This is another example

Table 5

^{13}C Shifts of Cyclohexane Derivatives

	1	2	3	4	5	6	M_e
cyclohexane	165.1	165.1	165.1	165.1	165.1	165.1	-
methyl-	159.1	156.4	165.7	165.8	165.7	156.4	169.5
cis-1,2-dimethyl-	157.8	157.8	160.7	168.5	168.5	160.7	176.5
trans-1,2-di-	152.7	152.7	156.2	165.4	165.4	156.2	172.0
cis-1,3-di-	159.4	147.5	159.4	156.8	165.8	156.8	169.4
trans-1,3-di-	165.2	150.8	165.2	158.3	171.5	158.3	171.7
cyclohexanol	123.0	157.0	168.1	166.6	168.1	157.0	-
1-methyl-	123.5	152.8	169.7	166.5	169.7	152.8	163.0
trans-2-methyl-	115.9	152.8	158.5	166.7	167.1	157.4	173.7
cis-2-methyl-	121.4	156.7	163.2	168.3	171.0	160.7	176.3
trans-3-methyl-	126.0	151.3	165.9	158.1	172.3	159.7	172.3
cis-3-methyl-	122.0	148.5	160.8	157.7	168.1	158.1	170.0
cyclohexanone	-16.0	152.1	166.0	168.7	166.0	152.1	-
2-methyl-	-17.5	148.5	157.3	168.3	165.5	151.9	179.0
3-methyl-	-15.6	148.7	159.3	160.3	168.3	152.7	171.7

Data from Ref. 17a, pp. 64, 163, 173, 290

Table 6

^{13}C Shifts for Some Pentose and Hexose Aldopyranoses[b]

	1	2	3	4	5	6	OM_e
2-D-glucose	100.5	119.7	121.0[a]	122.9	121.2[a]	131.7	-
-methyl glycoside	93.4	119.3	121.0[a]	127.8	121.2[a]	131.8	137.7
3-O-methyl-	100.4	121.4	109.8	123.4	121.2	131.8	132.6
β-D-glucose	96.7	118.3	116.8	122.9	116.8	131.7	-
-methyl glycoside	89.3	119.4	116.6[a]	127.8	116.7[a]	131.6	135.6
3-O-methyl-	96.5	118.8	107.3	123.6	116.8	131.8	132.9
α-D-galactose	100.3	123.3	124.1	123.3	122.3	131.4	-
β-D-galactose	96.0	120.5	119.7	123.8	117.6	131.6	-
α-D-fucose	100.3	123.1	124.3	120.7	126.5	176.9	-
β-D-fucose	96.2	120.7	119.5	121.1	122.0	176.9	-

[a] interchangeable with others noted on same row

[b] data from Ref. 26

Table 7

^{13}C Shifts for Fructose

	1	2	3	4	5	6
Fructopyranose	94.6	122.7	123.3	124.7	128.4	129.3
Fructofuranose	91.0	111.8	116.7	117.8	129.4	130.1

of a steric effect. Secondly, the methyl glycosides have the anomeric carbon shifted downfield relative to the free sugar. This is just the β-substituent effect.

Dorman and Roberts also studied a series of oligosaccharides[27]. Aside from the expected and observed effects of glycosidation, a particularly relevant comparison of the spectra of the furanose and pyranose forms of fructose was given (Table 7). All of the carbons common to both rings are shifted significantly downfield in the furanose form.

Finally, there are reports of the carbon spectra of cyclotetradecanone and cyclotetradecane. Stothers and Lauterbur[28] reported a chemical shift of -15.3 (downfield from internal CS_2) for the ketone of cyclotetradecanone. Burke and Lauterbur[29] reported a value of 167.0 for the mobile hydrocarbon. Anet[30] has done a variable temperature cmr study of cyclotetradecane and obtained a low temperature spectrum consistent with the proposed lowest energy diamond lattice conformation mentioned earlier. The chemical shifts were not reported or assigned; however, they could be estimated from the drawing of the spectrum in the paper. It seems reasonable that the most upfield carbons would be those designated D since they have protons directed into the ring. The carbon designated A can be assigned by the observed intensities leaving B and C for the downfield peak. It has already been established, however, that the erythronolide ring does not exist in this conformation.

F. ^{13}C Spectra and Interpretation

1. Experimental

All spectra were taken on a Varian DFS-60 or HR-220 spectrometer. Most of the DFS-60 spectra and all of the HR-220 spectra were taken using a pulse-Fourier transform system. Unless otherwise noted, spectra were taken in a 4 to 1 mixture of $CDCl_3$ and CH_2Cl_2. Chemical shifts were measured relative to internal CH_2Cl_2 and converted to internal CS_2 using the relationship: $d(CS_2) = d(CH_2Cl_2) + 138.8$. The resulting value can be converted to the currently favored internal TMS scale by the relationship $d(TMS) = 192.8 - d(CS_2)$. The estimated error in the chemical shift, unless otherwise noted, is plus or minus 0.1 ppm. The DFS-60, as currently equipped, has a Bruker probe and pulse system with a Varian 620i computer. A typical spectrum is taken in approximately 0.5 - 1.0 ml solvent at ambient probe temperature (30-35^{o} centigrade) using either noise, off-resonance, or single frequency decoupling (SFD). The magnet is locked on the deuterium signal from $CDCl_3$. Spectra can usually be obtained at concentrations greater than 0.1 molar, but concentrations of 0.3 - 0.4 molar are more typical.

A couple of general statements can be made about what might be expected in macrolide cmr spectra with regard to the effects discussed above. The effect of adding or removing a substituent should be smaller than in most of the model systems because of the observed attenuation on increased substitution. In particular the β-effect should be smaller and the γ-effect variable. When a substituent is removed the expected upfield (negative) β-shift could be opposed by a coincident downfield

shift caused by conformational reorganization which reduces steric compression. The observed β-effect which is already reduced due to extensive substitution could be even further reduced. Steric shifts would be expected to be much harder to predict due to the many possible steric interactions around and inside the large ring.

2. Aglycone spectra

Spectra were taken and assigned for nine aglycone derivatives. An example of a noise-decoupled spectrum for 6-deoxyerythronolide B is shown in Figure 13. This spectrum was obtained with a 0.5 - 1.0 molar solution in less than thirty minutes on the DFS-60 as currently equipped.

Assignments for the spectra of these nine derivatives plus the hemiacetal of 5-deoxy-5-oxoerythronolide B are given in Tables 8-10. The structures are shown in Figure 14. Ambiguous assignments are indicated on the tables; the more probable assignment is the one given. By and large these assignments were very tediously made. The only reliable method appeared to be low-power single frequency decoupling of the directly bonded protons. Discussion of how the assignments were made will proceed by region of the spectrum. The letter A followed by a number refers to "aglycone, position __". Later on the letter D refers to desosamine, the letter C to cladinose, and the letter M to mycarose.

The oxygenated carbons are the ketone and lactone carbonyls A9 and A1, the three secondary alcohols A3, A5, A11, the tertiary alcohol A6, and the lactone ester A13. The assignment of A9 was generally easy

FIGURE 13

Noise-decoupled CMR spectrum of 6-Deoxyerythronolide B

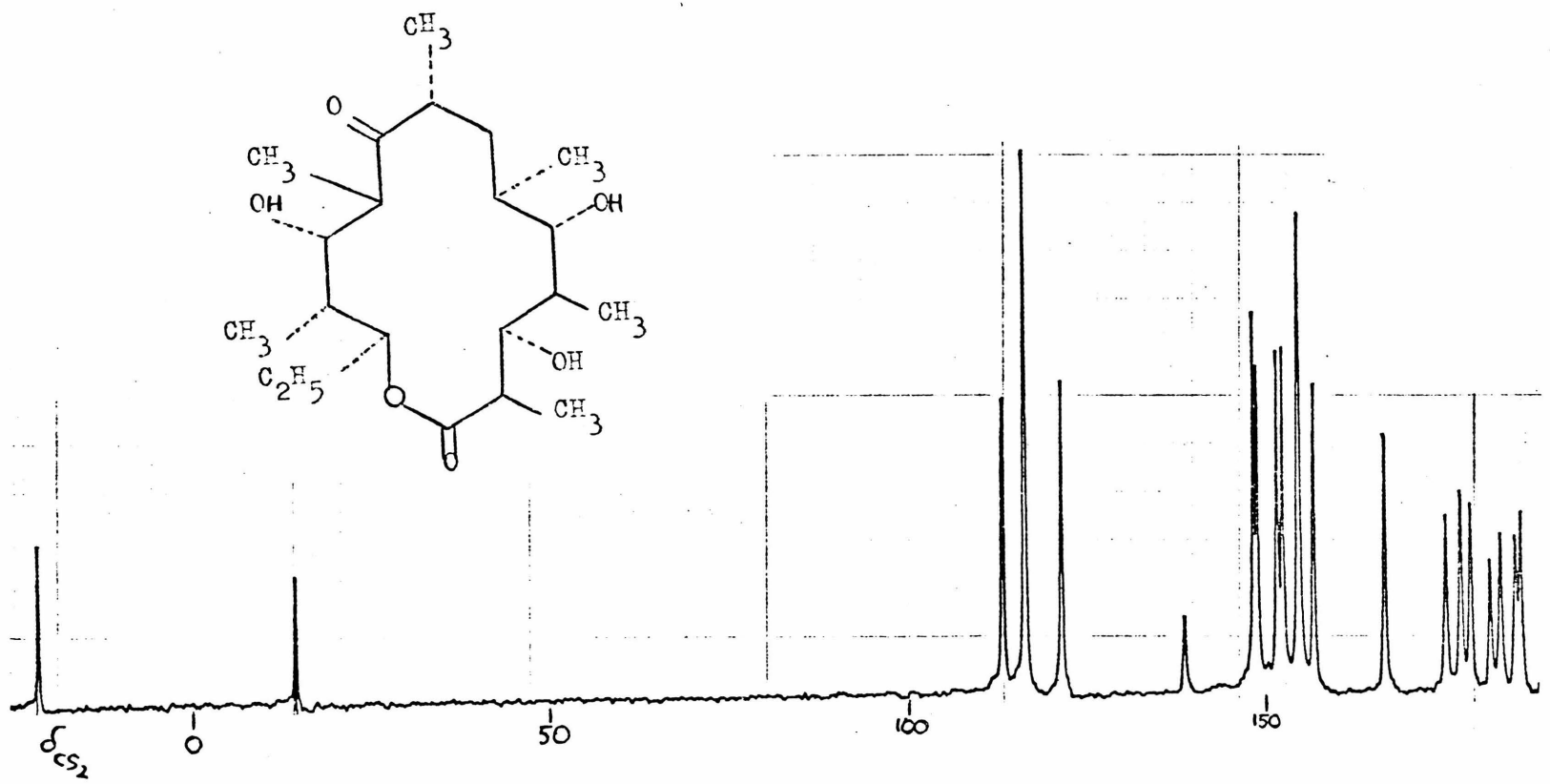


Table 8

Aglycone Assignments for Oxygenated Carbons

	(4)	(5)[a]	(8)	(8h)	(9)[c]	(11)	(12)	(15)	(16)	(17)
A9	-22.1	b	-23.7	81.2 ± 1	-23.0	-20.5 ± 1.5	-20.7	-27.2	-22.5	110.9
A1	14.1	14.8 ± 1	16.8	ca.15	16.8	17.2	17.4	17.9	19.5	14.1
A3	113.2	113.4	120.5	121.2	120.9	112.7[d]	113.6	115.1	115.9	112.1[d]
A5	116.3	111.5	-26.5	ca.-25	-25.0	113.0[d]	116.4	113.1	113.7	112.4[d]
A6	157.2	117.8	113.4	105.9	147.2	116.6	156.9	118.0	118.4	116.5
A11	121.6	127.6	122.8	122.4	122.0	117.6	118.2	122.7	119.9	121.8
A13	116.3	117.8	116.5	116.0	117.1	118.0	118.2	117.2	118.4	116.5

[a] dioxane solvent

[b] not investigated

[c] CH_2Cl_2 solvent

[d] interchangeable in same column.

See Table 0 for compound names.

Table 9

Aglycone Assignments for Methine and Methylene Carbons

	(4)	(5)[a]	(8)	(8h)	(9)[c]	(11)	(12)	(15)	(16)	(17)
A6	157.2[b]	117.8	113.4	105.9	147.2	116.6	156.9	118.0	118.4	116.5
A2	149.1	148.8	148.9	148.2	148.9	149.2[b]	149.4[b]	149.9	149.7[d]	148.2[b]
A8	152.8[d]	150.6[b]	147.0	141.0	145.9	149.9[b]	152.5[b]	147.1	150.3[d]	158.2
A10	148.6[d]	153.1[b]	154.8[b]	155.1	153.4[b]	153.4[b]	150.2[b]	153.8[b]	152.5[b]	160.5
A4	155.0[b]	156.4	151.4[b]	149.2	145.9	155.8[b]	155.2[b]	156.6[b]	156.5[b]	156.3
A12	152.0[b]	152.2	151.9[b]	e	152.2[b]	153.4[b]	153.5[b]	152.2[b]	153.4[b]	152.2
A7	155.0[b]	151.6	151.9[b]	150.2	156.0[b]	150.6[b]	155.2[b]	155.0[b]	155.8[b]	151.4

[a] Dioxane solvent

[b] Positively confirmed by SFD

[c] CH_2Cl_2 solvent

[d] Interchangeable with others in same column, most likely assignment shown

[e] buried

Table 10

Aglycone Assignments for Methyl Carbons

	(4)	(5)[a]	(8)	(8h)	(9)[c]	(11)	(12)	(15)	(16)	(17)
A14	167.2	166.7[d]	166.6	167.2	167.0	166.5[d]	166.6	166.4	166.4[d]	167.3
6-M_e	176.0	167.0[d]	171.3	167.7	176.4	165.9[d]	175.7	166.4	166.8[d]	165.9
12-M_e	183.6	183.9	183.6	184.3	183.7[d]	182.8	182.8	183.5[d]	182.9	184.1
4-M_e	185.8[d]	186.2	182.8[d]	b	183.7[d]	186.1	185.1	183.5[d]	183.7[e]	187.2
10-M_e	179.4[e]	176.3[e]	174.8	b	178.4	176.7[e]	178.5	177.5	178.0	175.4
2-M_e	178.0[e]	178.1[e]	178.4	177.6	178.4	178.1[e]	178.5	183.0[d]	182.4[e]	178.2
8-M_e	186.5[d]	185.0	183.1[d]	b	183.9[d]	184.8	185.1	183.5[d]	183.7[e]	182.7
A15	182.1	182.7	182.3	b	182.4	182.1	182.2	181.9	182.4	182.2
Ac-M_e	-	-	-	-	-	171.8	171.5	171.5(A3)	171.7	-
	-	-	-	-	-	-	-	173.8(A5)	all 3	-

[a] Dioxane solvent

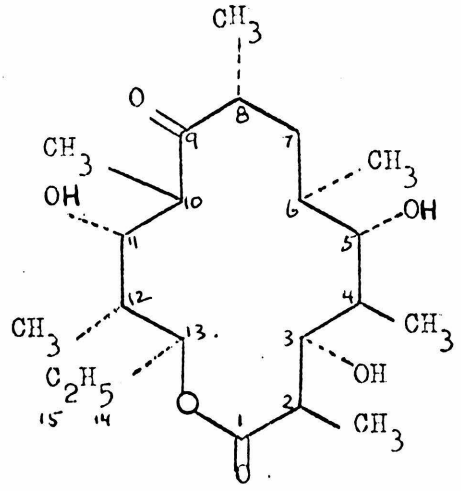
[b] buried

[c] CH_2Cl_2 solvent

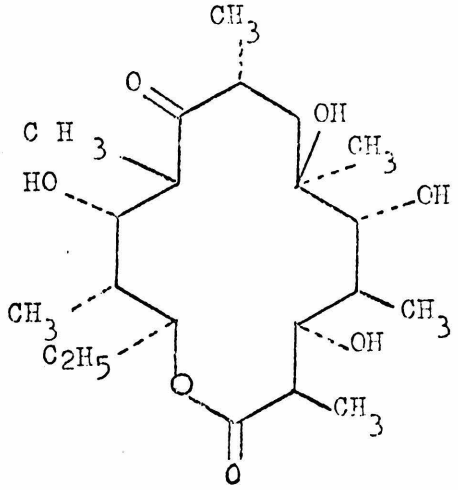
d,e Interchangeable with others in the same column.

FIGURE 14

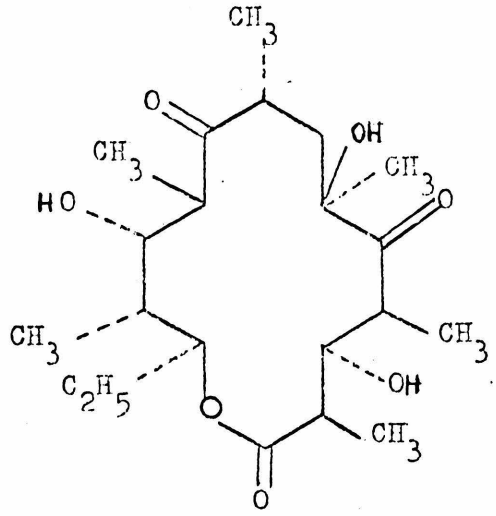
Aglycone Structures



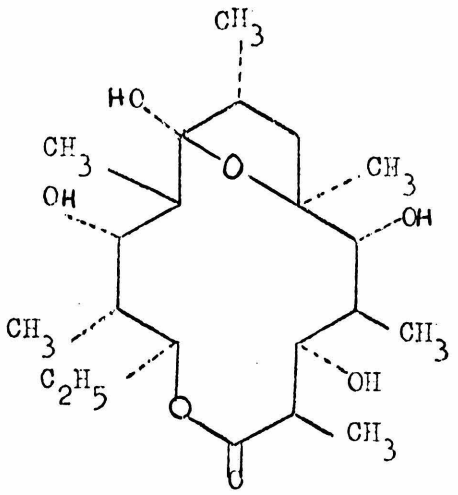
(4) 6-Deoxyerythronolide B



(5) Erythronolide B

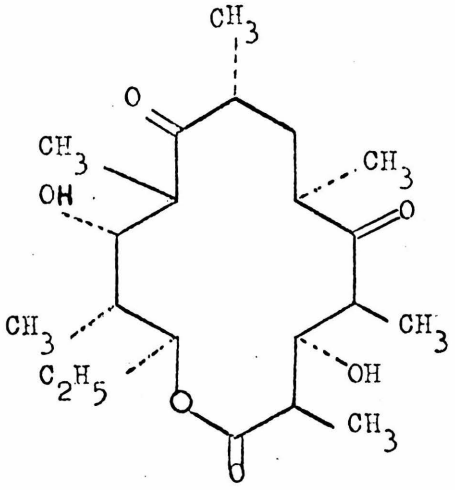


(8) 5-Deoxy-5-oxoerythronolide B keto form

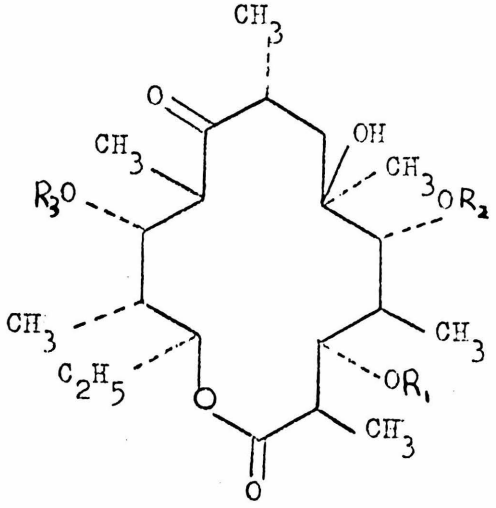


(8h) 5-Deoxy-5-oxoerythronolide B hemiketal form

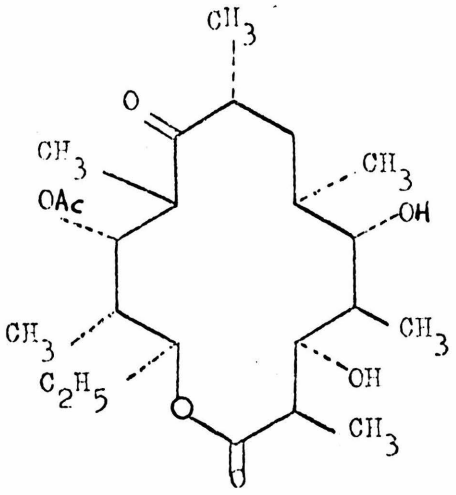
FIGURE 14 (Cont.)
Aglycone Structures



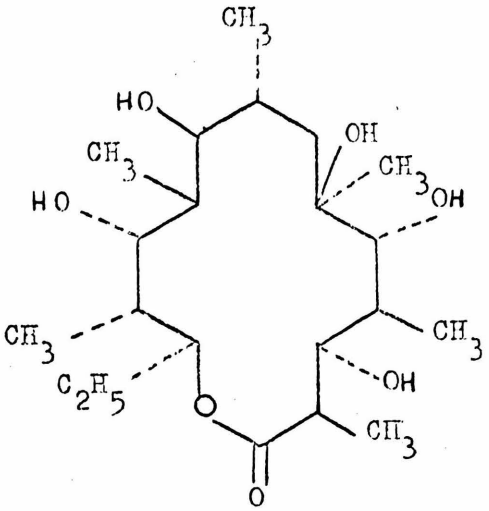
5,6-dideoxy-5-oxoerythronolide B (9)



	R_1	R_2	R_3
(11)	OH	OH	OAc
(15)	OAc	OAc	OH
(16)	OAc	OAc	OAc



11-acetyl-6-deoxyerythronolide B
(12)



9S-9-dihydroerythronolide B
(17)

as ketones appear at lower field than any other type of carbon in question here. Only in the case of the 5-oxo derivatives was ambiguity possible. The differentiation in this case could be made by the observation of a (negative) β-shift in going from 5-deoxy-5-oxoerythronolide B to 5,6-dideoxy-5-oxoerythronolide B and the comparison of the former with its hemiketal spectrum. Assignment of A1 was also easy. The only possible ambiguities here were in the acetoxy derivatives; however, the actoxy carbonyls are generally 3-6 ppm upfield from the lactone.

Most of the assignments of A3, A5, A6, A11 and A13 were made by single-frequency decoupling (SFD) and off-resonance decoupling (ORD). Unequivocal assignments could usually be made if the proton signals differed by more than 0.1 ppm. Most of the proton data came from Reference 15. The spectrum of erythronolide B was taken on a very dilute sample in dioxane since it is nearly insoluble in deuteriochloroform. For this reason the comparison of this sample with the others is less reliable. Single frequency decoupling was not feasible but the pattern of resonances is the same as for the others so assignment was made by comparison. The spectrum of 5,6-dideoxy-5-oxoerythronolide B was taken in methylene chloride and again no proton data were available. Assignments were made by comparison with the 5-oxo derivatives. Little solvent shift was observed in other spectra taken in $CDCl_3$ and CH_2Cl_2. The resonances in the spectrum of the hemiketal of 5-deoxy-5-oxoerythronolide B were not all found in this relatively insoluble sample. Assignments were made by comparison with the ketone spectrum and with the other ketal spectra. The changes observed on

going from the keto to hemiketal from will be discussed later.

The remaining aglycone ring carbons A2, A8, A10, A4, A12 and A7 proved to be the most difficult to unambiguously assign. For this region only, unambiguous assignments based on SFD experiments are noted. Proton data were available in most cases, but the resonances were closer together and less well characterized due to the complexity of the proton spectra. In all cases in which proton data were available, the peaks representing A2, A8 and A10 could be separated from those representing A4, A12,and A7. Further, A7 could often be assigned since it is a methylene rather than a methine. The spectra of 11-acetyl-erythronolide B and 11-acetyl-6-deoxyerythronolide B could be completely assigned in this region by SFD. Beyond this some comparative assignments had to be made. A2 was generally assigned by SFD but for those not so indicated in Table 10, assignments were based on the fact that this region appears to be conformationally homogeneous in all derivatives without a substituent on the adjacent (A3) hydroxy group. This is the conclusion of Egan[15] based on his study of vicinal proton couplings. In the 3-acetoxy derivatives an upfield shift would be expecte d due to steric interaction with the freely rotating acetoxy group.

The assignments of A8 and A10 were the most difficult. The assignments for 6-deoxyerythronolide B were indicated by repeated SFD experiments. In this case the proton signals differ by 0.12 ppm[15]. For 9S-9-dihydroerythronolide B no SFD was possible but the assignments of A8 and A10 rest on the observed upfield shift on reducing a ketone to an alcohol of the adjacent carbons. This is one substituent effect

which was consistently observed at near its expected value in the entire series. The assignments of A8 and A10 for erythronolide B are based on comparison. The uncertainty of the solvent change makes these assignments even less meaningful. The assignment of A8 in 3,5-diacetylerythronolide B is by elimination. The proton frequencies of A8 and A2 were too close in this compound; however, comparison with 3,5,11-triacetylerythronolide B indicates the assignment is correct. Egan's[15] results show that the only differences in the vicinal couplings between these two compounds are in the A8 region. No proton data (in $CDCl_3$) was available for 5,6-dideoxy-5-oxo-erythronolide B but in light of the surprisingly different carbon shifts, SFD experiments were done in the following manner. Comparison of proton results for a number of compounds in pyridine (including the one in question) and $CDCl_3$ show that H10 would be furthest downfield; H8 and H2 would be next; H6 would be next most upfield; and H12 would be the furthest upfield. The carbon peaks responded significantly to a "scanning" of the proton region so assignments were made on this basis.

The assignments of A4, A12 and A7 are almost all unequivocal by SFD. Only for 9S-9-dihydroerythronolide B and erythronolide B were comparative assignments made.

The spectrum of the hemiketal form of 5-deoxy-5-oxoerythronolide B presented a problem because of the low intensity of the peaks. The peak at 141.0 corresponding to A8 is expected by comparison with other derivatives of this type. The others that could be found were assigned by comparison.

In the methyl region the only completely assigned spectrum is of 9S-9-dihydroerythronolide B. This was done by SFD using data from Demarco[31] and from the observed effect of reducing a ketone on the adjacent methyl as mentioned earlier. In most cases proton data were unavailable or the proton peaks were too close together. A14 and 6-Me were assignable in most cases based on their chemical shift. A15 was usually assigned by SFD since its proton is generally the furthest upfield and the carbon shift varied little. Some of the assignments of 12-Me were made by SFD since the methyl protons generally appear at higher field than all the others except A15. The remaining assignments of 12-Me were made assuming small changes. This region of the ring was also found to be conformationally homogeneous by Egan[15]. Furthermore, the carbon shifts for A12 already assigned are nearly constant except in 11-acetyl derivatives in which a 1.0- 1.5 ppm upfield shift is observed. The corresponding shifts of 12-Me in these derivatives seems to be downfield. The remaining methyls fall into a reasonable pattern based on comparison with 9S-9-dihydroerythronolide B and known effects. The two highfield methyls in all but the 3,5-diacetyl derivatives correspond to 4-Me and 8-Me. These are axial in the expected conformation. In particular, 4-Me is held in a very unfavorable axial position because of the apparent hydrogen bonding between the hydroxyls on A3 and A5 (Figure 8). Differentiation between 4-Me and 8-Me is made by comparison of the various derivatives and the most likely assignments are shown. The remaining methyls, 2-Me and 10-Me are therefore downfield in the various derivatives. The only exceptions are in the 3,5-diacetyl compounds in which 2-Me is shifted upfield. The effect of acetoxy

substitution on the gamma-carbon will be discussed more fully later. Differentiation of 2-Me and 10-Me is made on the same comparative basis with the more likely assignment shown. Further support for these assignments is given in the discussion of the erythromycin spectra.

Most of the peaks for the hemiketal of 5-deoxy-5-oxoerythronolide B were not found. Those that were are assigned as indicated. The acetoxy methyl assignments are also indicated.

3. Amino-sugar spectra

Samples of cladinose and desosamine were obtained and cmr spectra run. The assignments are shown in Table 11. No SFD was done on the free sugars, but a number of assignments were made by comparison with glycoside spectra which were assigned by SFD. The assignments of the anomers not observed naturally were made using the results of Dorman and Roberts[26]. In the alpha-anomers the resulting 1,3-diaxial interactions with the 3- and 5-carbons results in upfield shifts of all three. The shift is 5-6 ppm for D3 and D5 of desosamine and for C5 of cladinose. For C3 the shift was 1 ppm; however, this is a quaternary carbon. No sugars with such quaternary carbons were studied by Dorman and Roberts[26]. The smaller upfield shift for the quaternary center may be due to the fact that steric compression is felt more strongly by a carbon with a directly bonded proton. The corresponding upfield shifts of C1, C2, D1, and D2 are also observed (comparing anomers).

4. Monoglycoside spectra

Two monoglycosides, 3-O-α-L-mycarosylerythronolide B (13) and 5-O-β-D-desosaminylerythronolide B (14) were available. The structures

Table 11

^{13}C Assignments for Cladinose and Desosamine

	C1	C2	C3	C4	C5	C6	C3-OM_e	C3-M_e
α-cladinose	101.1	156.0	116.9	114.9	128.3	174.6	142.5	171.9
β-cladinose	100.6	153.3	117.8	114.7	121.8	174.6	143.8	171.9

	D1	D2	D3	D4	D5	D6	D-N-Me_2
α-desosamine	99.6	122.6	132.4	163.4	127.5	171.6	152.4
β-desosamine	94.4	120.8	127.5	163.4	123.5	171.6	152.1

are shown in Figure 15 and the assignments in Table 12. The sample of (14) was somewhat difficult to work with due to partial decomposition.

The assignments in the oxygenated region were made relatively easily by SFD. The pattern is the same as in the aglycone derivatives with the exception of the carbon retaining the sugar substituent. The assignments of the sugar peaks were also made by SFD.

The methine region was again the most difficult to assign. Assignments of (13) were made by SFD. However, it was here that the trouble with (14) was greatest. The assignments shown are the most likely and are consistent with the repeated SFD experiments done. The signals for A12, D-NMe_2 and A10 were often unresolved.

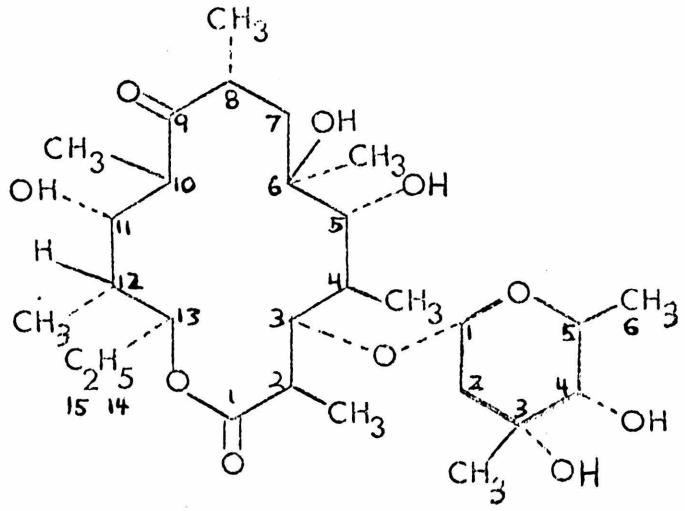
The methyl peaks were assigned by comparison with the aglycone spectra. The only significant difference between the two are the assignments for 2-Me which would be expected to be upfield in (13). These monoglycosides were found to be largely conformationally homogeneous with the aglycone derivatives by Egan[15].

5. Erythromycins' spectra

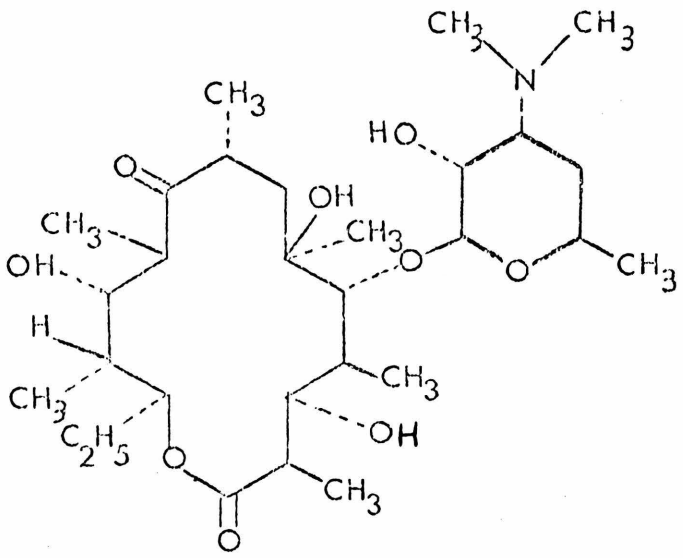
Included in this set are the parent antibiotics erythromycin A and B and several derivatives made by acid degradation or modification of the glycosidated desosamine. Structures are shown in Figure 16 and the assignments in Tables 13-16. A sample noise-decoupled spectrum is shown in Figure 17. Complete proton data were available for all but the modified desosamine derivatives. The 3'-de-dimethylamino-3',4'-dehydroerythromycin A sample was only slightly soluble.

The oxygenated carbons were assigned by ORD and SFD where proton data were available and by comparison otherwise. The only

FIGURE 15



3-O-α-L-Mycarosylerythronolide B (6)



5-O-β-D-desosaminylerythronolide B (14)

Table 12

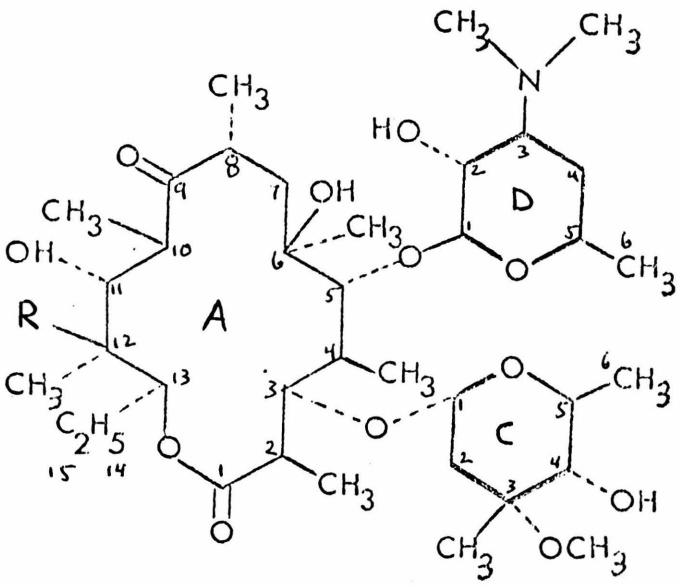
Monoglycoside Assignments

	(6)	(7)		(6)	(7)		(6)	(7)
A9	-26.3	-22.1	A2	148.2	147.9	A14	166.8	167.3
A1	17.4	14.5	A8	147.1	149.8	6-M_e	167.2	165.1
A3	102.9	115.0	A10	153.6	152.7[d]	12-M_e	183.7[c]	183.8
A5	110.9	99.6	A4	156.4	154.2[d]	4-M_e	183.5[c]	185.7[b]
A6	117.3	117.8	A12	152.5	152.3	10-M_e	177.0	176.9
A11	122.6	122.0	A7	151.4	151.6[d]	2-M_e	183.7[c]	176.9
A13	117.3	116.7	A15	182.4	182.3	8-M_e	184.1[c]	185.3[b]

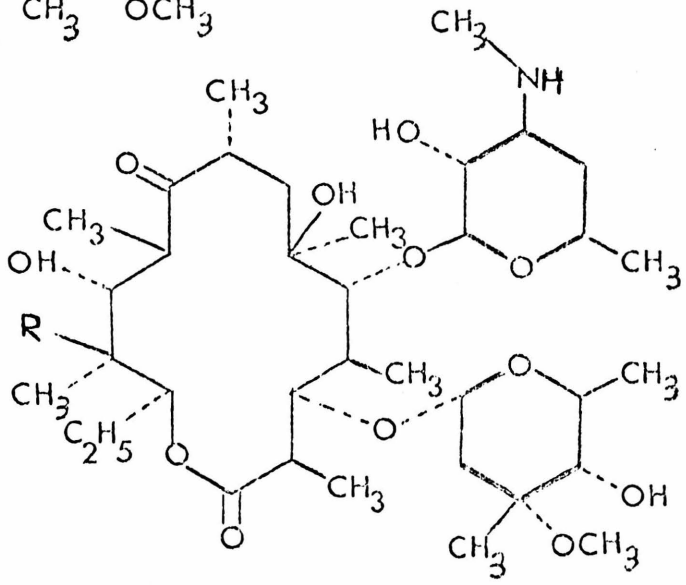
	(7)		(6)
D1	86.4	M1	91.8
D2	122.0	M2	155.8
D3	127.1	M3	122.1
D4	164.7	M4	116.0
D5	122.6	M5	125.7
D6	171.7	M6	174.4[a]
D-NM_e	152.5	M3-M_e	174.9[a]

a,b,c,d interchangeable

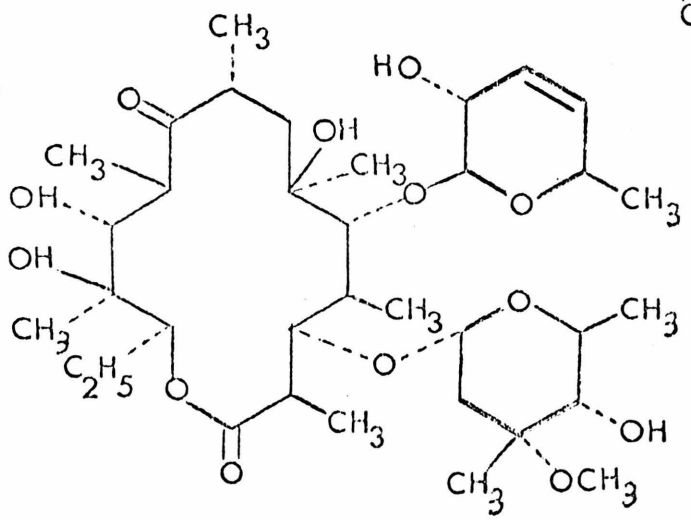
FIGURE 16
Erythromycin Structures



Erythromycin A
R= OH (1)
Erythromycin B
R= H (2)



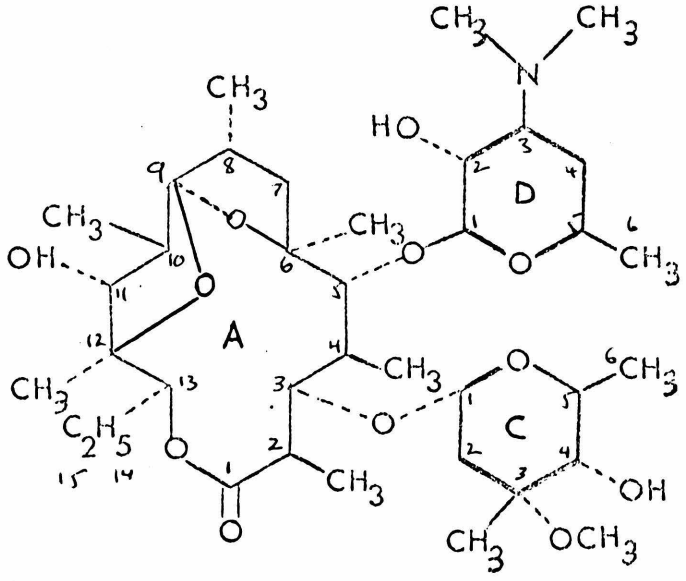
de-N-methylerythromycin A
R = OH (18)
de-N-methylerythromycin B
R = H (19)



3'-dedimethylamino-
3',4'-dehydro-
erythromycin A (20)

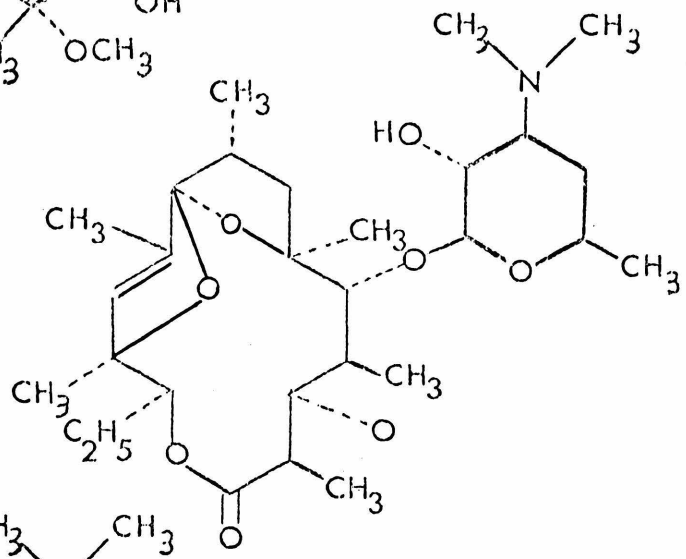
FIGURE 16 (Cont.)

Erythromycin Structures



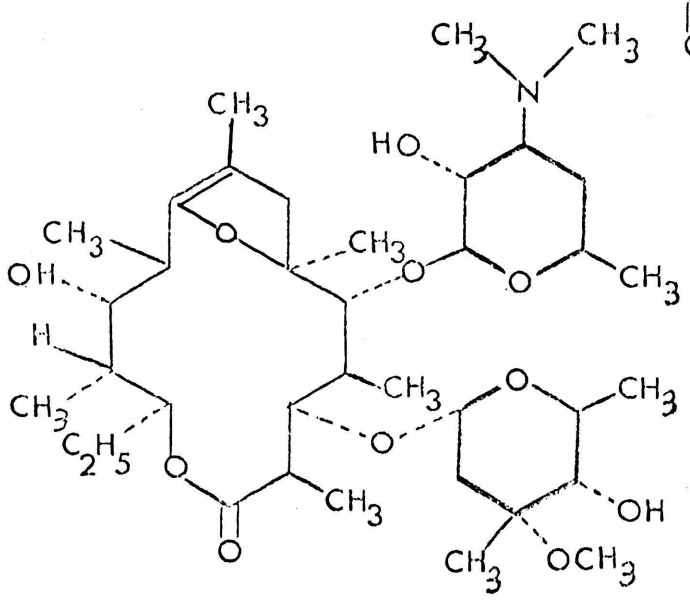
anhydroerythromycin A

(21)



erythralosamine

(22)



8,9 anhydroerythro-
mycin B-6,9-hemiketal

(23)

Table 13

Erythromycin Assignments for Oxygenated Aglycone Carbons

	(1)	(2)	(18)[a]	(19)[a]	(20)	(21)	(22)	(23)
A9	-29.4	-27.2	-28.2	-26.3	-29.1	76.4	72.7	41.4
A1	16.2	16.4	16.7	16.6	16.4	13.2	13.9	14.5
A3	112.3	112.1	112.4	112.1	112.2	116.5	121.9	115.5
A5	108.7	108.7	108.2	108.1	108.3	105.5	105.6	112.2
A6	117.5	117.5	117.8	117.7	117.3	110.3[b]	110.6	106.5
A11	123.5	123.1	123.6	123.2	123.4	106.5	64.0	121.3
A13	115.3	117.5	115.6	117.7	115.3	117.3	113.7	115.1
A12	117.5	153.2	117.8	153.1	117.3	111.0[b]	103.8	148.2

[a] CH_2Cl_2 solvent

[b] interchangeable

See Figure 16

Table 14

Erythromycin Assignments for Sugar Carbons

	(1)	(2)	(18)[a]	(19)[a]	(20)	(21)	(22)	(23)
D1	89.1	89.4	89.7	89.7	89.6	89.5	88.0	89.8
D2	121.4	121.5	118.1	118.2	122.0[b]	123.0[d]	122.5	121.5
D3	126.8	126.9	132.4	132.3	60.2[c]	127.4	126.7	126.7
D4	163.6	163.6	160.1	160.0	66.1[c]	163.7	163.7	163.6
D5	123.5	123.6	124.0	124.0	123.4[b]	122.8	123.2	123.7
D6	171.0	171.2	171.6	171.4	171.0	171.2	171.0	171.5
DN-M_{e2}	152.1	152.2	155.7	155.6	-	152.3	152.4	152.3
C1	96.0	95.9	96.1	95.9	95.8	97.8	-	98.0
C2	157.4	157.5	157.6	157.5	157.4	157.9	-	157.6
C3	119.7	119.9	119.9	119.9	119.7	119.7	-	119.3
C4	114.3	114.6	114.5	114.6	114.4	114.2	-	114.2
C5	126.8	126.9	127.0	126.9	126.8	126.7	-	126.7
C6	173.8	174.0	174.1	174.2	174.1	174.9	-	174.2
C3-OM_e	142.9	143.1	143.2	143.3	143.0	143.4	-	142.9
C3-M_e	171.0	171.2	171.3	171.4	171.0	171.6	-	171.0

[a] CH_2Cl_2 solvent

b,c,d interchangeable

Table 15

Erythromycin Assignments for Methine Carbons

	(1)	(2)	(18)[a]	(19)[a]	(10)	(21)	(22)	(23)
A2	147.4	147.6	147.5	147.5	147.3	146.7	152.0[b]	147.9
A8	147.4	147.6	147.9	148.1	147.1	141.6	145.8	91.4
A10	154.3	153.2	154.1	153.1	153.8	151.0	53.4	158.4
A4	152.9	152.9	153.1	152.5	152.7	149.6	147.8	149.2
A12	117.5	153.2	117.8	153.1	117.3	111.0	103.8	148.2
A7	153.9	154.6	154.1	154.4	153.8	151.0	149.6[b]	150.1

[a] CH_2Cl_2 solvent

[b] interchangeable

Table 16

Erythromycin Assignments of Methyl Carbons

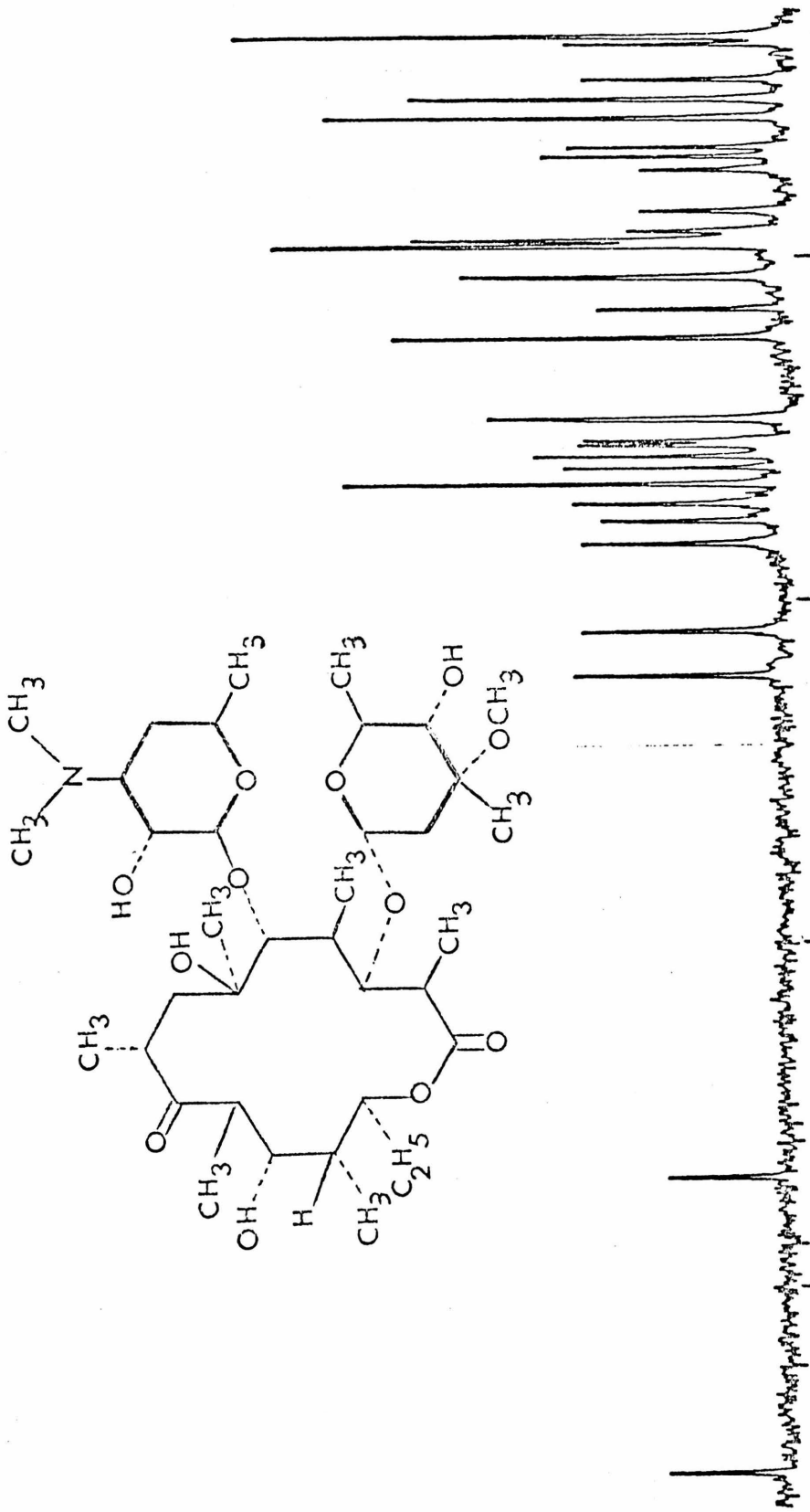
	(1)	(2)	(18)[a]	(19)[a]	(20)	(21)	(22)	(23)
A14	171.0	166.8	171.3	166.9	171.0	168.9	168.9	166.4
6-M_e	165.6	165.5	166.1	165.7	165.7	164.9	163.2	167.5
12-M_e	174.0	183.3	174.3	183.6	174.1	167.8	169.2	183.8
4-M_e	183.2	183.3	183.1	183.6	183.0	180.6[d]	179.9[e]	183.5
10-M_e	176.4[b]	174.0	176.4[c]	174.2	176.4[b]	178.3	178.9	179.4[f]
2-M_e	176.1[b]	176.8	176.8[c]	177.0	176.1[b]	175.2[d]	179.0[e]	177.3[f]
8-M_e	180.3	183.3	180.7	183.0	180.6	179.2[d]	180.8	180.4
A15	181.7	182.2	182.1	182.4	181.8	181.9	182.5	182.1

[a] CH_2Cl_2 solvent

b,c,d,e,f interchangeable

FIGURE 17

Noise-decoupled CMR Spectrum of Erythromycin B



difficulty in this region was determining that the quaternary carbons A6 and A12 are coincident in erythromycin A.

Except for the modified desosamine derivatives assignment of the sugar peaks was straightforward. These assignments were independently verified by SFD. In the de-N-methyl derivatives the assignments were made based on the expected substituent and steric effects. In these pyranose sugars these effects tend to be more predictable based on others' results[26,27]. The downfield shift of D2 in going to the de-N-methyl derivatives is a negative gamma effect and the upfield shifts of D3 and D-NMe are negative β-effects. The upfield shift of D5 is a negative remote steric effect.

The methine region was assigned primarily by SFD and ORD. Some of the difficulties encountered in the aglycone spectra were also encountered here. The problems caused by overlapping peaks were acute. in this region; however, repeated SFD experiments allowed the assignments to be made independent of comparison of spectra.

The overlapping of peaks is acute in the methyl region. As mentioned before, all the proton data were available but many of the proton resonances were too close together to allow differentiation in the carbon spectra. One particularly critical assignment was the differentiation of 2-Me and 10-Me in erythromycin B. In this case the proton resonances happened to be separated by 0.20 ppm and assignment could be made. The result is also the intuitive comparative one based on the observed invariance of vicinal proton couplings in the A2-A5 region[15].

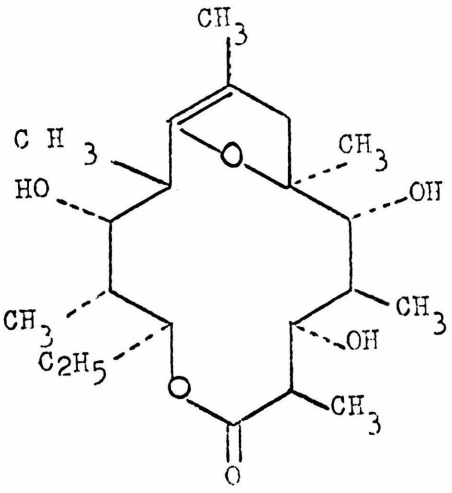
Spectra were taken and assigned of two other samples, 8,9-anhydroerythronolide B -6,9-hemiketal and megalomycin A. The former is obtained from erythronolide B in acid[32]. Megalomycin A is a naturally occurring antibiotic recently discovered[33]. The structures and spectral assignments are shown in Figure 18 and Tables 17 and 18. Proton data were available for 8,9-anhydroerythronolide B-6,9-hemiketal [34] and the spectrum was assigned by SFD. The only signficant differences between assigned peaks for this sample and 8,9-anhydroerythronycine B-6,9-hemiketal are in the A2-A6 region and will be discussed later. The assignments for megalomycin A are based on an ORD experiment and comparison with the erythromycins.

G. Discussion

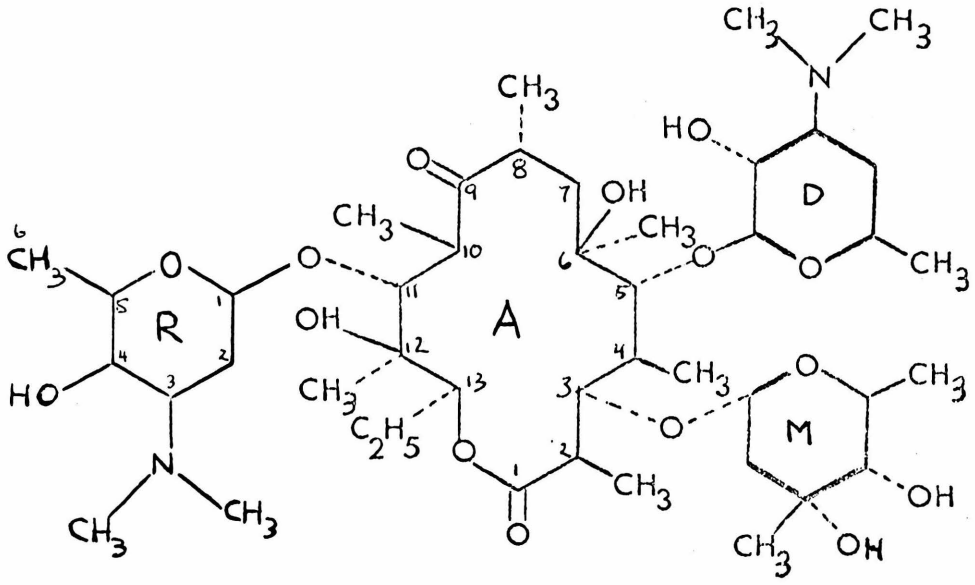
1. Substituent and steric effects

A number of regularities are discernible based on commonly observed effects on carbon spectra. The effect of adding or removing a hydroxyl group in aliphatic and alicyclic hydrocarbons was discussed earlier. Several of the samples studied differ by the presence or absence of hydroxy groups at A6 or A12. Removing the hydroxyl group causes a 35-40 ppm upfield shift at the directly bonded carbon and a 9-10 ppm upfield shift at the geminal methyl except in the 5-oxo derivatives. These are the expected magnitudes. The effects at the other β-carbons are considerably diminished. Removing the 6-hydroxy causes a 3-5 ppm upfield shift of A5 (when hydroxylated) and A7. The effect on A5 as a ketone is only 1.2 ppm. Substituent effects on a ketone are expected to be smaller[35]. The effect of removing the

FIGURE 18



8,9-anhydroerythronolide B-6,9-hemiketal



Megalomycin A

Table 17

Assignments for 8,9-anhydroerythronolide B-6,9-hemiketal

A9	41.4	A2	148.2	A14	166.7
A1	16.0	A8	90.9	6-M_e	164.3
A3	110.5	A10	158.7	12-M_e	184.3
A5	110.5	A4	157.6	4-M_e	186.5
A6	108.6	A12	146.7	10-M_e	177.5[a]
A11	121.9	A7	150.3	2-M_e	178.6[a]
A13	114.6	A15	182.3	8-M_e	180.6

[a] interchangeable

Table 18

Assignments for Megalomycin A

A9	-28.9	R4	123.2	D4	163.8
A1	16.9	D5	123.5	R2	164.4
D1	87.9	M5	125.1	6-M_e	166.6
M1	93.9	D3	126.8	A14	170.9
R1	102.0	R3	126.8	D6	171.2
A5	107.3	R5	132.9	M3-M_e	173.6
A3	109.4	A8	146.6	M6	173.8
A12	112.0	A2	147.6	12-M_e	174.0
A13	115.5	R-NM_{e2}	149.9	10-M_e	176.0
M4	115.7	A4	151.3	R6	176.0
A6	117.9	D-NM_{e2}	152.2	2-M_e	177.4
A11	118.9	A7	153.6	8-M_e	180.2
D2	121.2	A10	154.7	A15	181.9
M3	122.8	M2	155.2	4-M_e	182.9

12-hydroxy is only +2 ppm at A13, ca. 0 ppm at A11, and -4 ppm at A14. Smaller β-effects are expected in more highly substituted molecules as discussed earlier, but it is likely that conformational changes associated with removing or adding these hydroxy groups contribute also. These changes will be discussed later.

The effect of interconverting a hydroxy and a ketone appears to be larger and comparable to the magnitude observed in simpler systems. Reducing the ketone at A9 shifts the β-methines upfield 7-8 ppm and the γ-methyls downfield 1-2 ppm. The β-effect here is comparable to that observed in cyclohexane systems while the γ-effect is smaller (Table 5). The effect of oxidizing A5 to a ketone is 4-5 ppm downfield at A6 when hydroxylated and 10 ppm downfield in the 6-deoxy derivatives. The β-effect at A4 is 4 ppm downfield. The effects at the γ-methyls are variable. Conformational changes are more important here because the flexible part of the ring (A6-A9) is affected.

The effects of acetoxy and methoxy groups relative to the free alcohols have been studied in several systems[36]. The general trend for acetoxy substitution seems to be a downfield shift at the α-carbon, an upfield shift at the β-carbon, and little or no effect elsewhere. When A11 is acetylated there is a 2-3 ppm downfield shift at A11, a 1-1.5 upfield shift at A12 and a 1 ppm upfield shift at A13. These are reasonable based on accepted considerations. An upfield shift in the lactone A1 is also apparent in 11-acetoxy derivatives. This is most likely a remote steric effect. The effects at A9 and A10 vary considerably. These inconsistencies (based on the simple models) are almost certainly caused by the conformational changes which accompany

11-acetylation and will be discussed more fully later. It is worthwhile to note that the A10-A13 region does not change throughout the erythromycin B series as indicated by the constant vicinal proton-proton couplings[15]. The success of the simple model in predicting the acetylation shifts at A11-A13 is consistent with this conclusion. The effect of acetylation at A3 and A5 is complicated by the conformational change that occurs in this region. Egan[15] has shown that the conformation of the 3,5-diacetyl derivatives resembles the conformation of the natural antibiotics in which the substituent at A3 moves above the plane of the ring and the substituent at A5 moves down (Figure 8). In the 3,5-diacetyl derivatives the shifts of A3 and A5 are upfield 1-4 ppm. There are also upfield shifts of 1-2 ppm at A2, A4 and A6. The effects at the γ-carbons are more profound. A1 is shifted upfield 1-2 ppm, 2-Me is shifted upfield 5-6 ppm, 4-Me is shifted downfield 3-4 ppm, 6-Me is essentially unchanged and A7 is shifted upfield 3-4 ppm. As noted earlier, acetylation generally has little or no effect at γ-carbons so an explanation is in order. These shifts are consistent with the conformational changes and the fact that the free hydroxyls in the non-acetylated derivatives are diaxial and hydrogen bonded. Since no hydrogen bonding can exist in the diacetyl derivatives, the acetyl groups are more free to rotate and the steric effects of this substituent are felt at the gamma carbons. The upfield shifts at 2-Me and A7 are consistent with the nearest acetyl group moving closer as in the proposed conformation. The downfield shift at 4-Me is a result of the relief of 1,3-diaxial steric interactions in the non-acetylated derivatives. As shown in Figure 8a, the 4-Me group is

held in an unfavorable axial position by the hydrogen bonded hydroxy groups. In Figure 8b the 4-Me has moved away from the center of the ring. This conclusion is consistent with the assignments given earlier. In fact, the chemical shift of 4-Me in the non-substituted (at A3 and A5) derivatives is among the highest reported for any methyl in any hydrocarbon.

The effect of methyl substitution at oxygen and nitrogen can be assessed by comparing mycarose with cladinose and the de-N-methyl derivatives with desosamine respectively. Mycarose differs from cladinose by methylation of the 3-hydroxy group (see Figure 19 and Table 19).

Dorman, Angyal and Roberts[37] determined in a series of inositols that O-methylation of axial hydroxyls causes downfield shifts at the α- and β-carbons and little or no effect elsewhere. The only significant deviation from this pattern is at C5 where an upfield shift is observed in going from mycarosyl to either cladinose or cladinosyl. The change is only +1.2 ppm in the glycosidated sugars, however.

The expected changes caused by de-N-methylation of desosamine were used to assign the spectra and no further comment is justified or necessary.

Most of the steric effects observed are associated with the conformational changes to be discussed later. One change that is satisfactorily explained by a simple argument is the upfield shift at A2 in the non-glycosidated derivatives. It has already been pointed out that 4-Me is forced down by the hydrogen bonding in these

Table 19

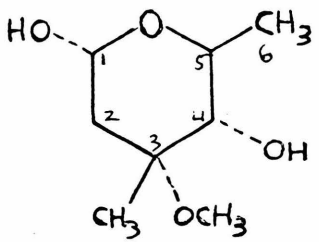
Comparison of Glycosidated Cladinose and Mycarose with Free Cladinose

	C1	C2	C3	C4	C5	C6	C3-OM_e	C3-M_e
Glycosidated mycarose	91.8	155.8	122.1	116.0	125.7	174.4[b]	-	174.9[b]
Glycosidated cladinose[a]	96.0	157.5	119.7	114.3	126.9	174.2	143.0	171.2
Free cladinose	101.1	156.0	116.9	114.9	128.3	174.6	142.5	171.9

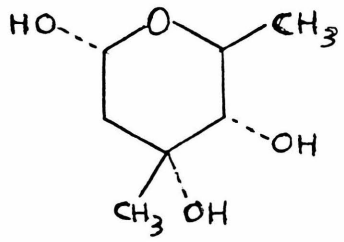
[a]Average position

[b]Interchangeable

FIGURE 19



α-cladinoe



α-mycarose

derivatives. This increases the 1,3-diaxial interaction between A2 and A4 and causes the observed upfield shift.

It is generally accepted that carbons bearing axial hydroxyl groups appear upfield from those with equatorial hydroxyl groups in cyclohexane derivatives[38]. These considerations seem to be only partially applicable to this study. A11 is always axial and appears at higher field (121-123). A3 and A5 are always equatorial and appear at low field in unmodified derivatives (111-115). However, A6 which is primarily axial and A12 which is equatorial when hydroxylated are coincident in erythromycin A. Worse yet, A9 in 9S-9-dihydroerythronolide B has been shown to be axial[39] yet appears at 110.9 ppm. A possible explanation is that this position is a "corner" on the alternate diamond lattice conformation and in the proposed conformation is free of diaxial interactions except with the hydroxy at A11. Differentiation of axial and equatorial methyls by chemical shift is much clearer. The axial methyls 4-Me, 8-Me and 12-Me are always upfield from the equatorial methyls 2-Me, 6-Me and 10-Me. The average difference is about 6 ppm. A priori use of these simplified considerations on molecules of this complexity is probably not justified; however, a reasonable agreement is obtained.

2. Hydrogen bonding

It appears that cmr is extremely useful in assessing intramolecular hydrogen bonding in these molecules. The observed downfield shift of carbonyl carbons involved in hydrogen bonding has already been mentioned[22,23]. Examination of the ketone shifts in

the various derivatives shows a variation of about 9 ppm. The shift of A9 for erythromycin A is -29.4 ppm and is one of the most deshielded (if not the most deshielded in a solvent other than sulfuric acid) organic carbonyls reported to date. By comparison, bis-isopropyl ketone appears at -22.8 ppm[40]. Since it is unlikely that ring strain (which can cause a downfield shift[41]) is much of a problem here, and since steric effects should cause an upfield shift, intramolecular hydrogen bonding provides the most reasonable explanation. Examination of the proposed conformations (Figure 8) reveals that hydrogen bonding to A9 could be from either the hydroxyl group at A11 or A6. It appears, however, that the dominant bonding is from A6. Comparison of erythromycin A with megalomycin A, which has a sugar substituent at A11, shows ketone shifts of -29.4 and -28.9 ppm respectively. In all 6-deoxy derivatives the ketone appears at -20.7 to -23.0 ppm. In lankamycin (Figures 2,3) which lacks a 6-hydroxy group and has an 11-acetoxy group (and also an 8-hydroxy) the ketone appears at -21.9 ppm. These latter values are comparable to what would be expected for an unstrained tetra-substituted ketone. Further, 11-acetylation causes a 1-2 ppm upfield shift in the 6-deoxy derivative which could be attributed to loss of hydrogen bonding or a steric effect.

It has already been noted that the carbons bearing the sugar substituents in the monoglycosides appear at very low field (Table 12). This is also most likely a result of hydrogen bonding from the corresponding free hydroxyl (that is, the hydroxyl at A3 is hydrogen bound to the glycosidated oxygen at A5 in 5-O-β-D-desosaminylerythronolide B). In the 5-oxo derivatives the A5 ketone appears at low field (Table 8)

again, probably due to hydrogen bonding from A3. It has been determined that these derivatives (5-oxo) are conformationally homogeneous to the others in the A2-A5 region[42]. Also significant is the 8 ppm upfield shift of A3 in the 5-oxo derivatives (Table 8). Available evidence indicates a downfield shift is expected for the hydrogen donor in hydrogen bonding situations. In particular, the carbon of chloroform was found to shift downfield in oxygenated solvents[43]. Downfield shifts are also observed in hydrogen bonding phenolic systems such as salicylaldehyde and salicylic acid derivatives[44]. A cmr spectrum of 4-hydroxy-4-methylpentanone was taken. The shifts for this sample and some other reported alcohols are shown in Table 20. At best, a 1 ppm upfield shift is observed in the hydrogen bonded situation. The upfield shift of A3 in the 5-oxo derivatives relative to 5-hydroxy derivatives is most probably caused by the lack of hydrogen bonding to A3 and an apparent conformational change at the lactone. A1 is shifted upfield in these derivatives also.

3. Conformational considerations

The cmr data described are consistent with several subtle conformational changes among the various derivatives.

Comparison of the spectra for erythromycin A and erythromycin B show a number of differences that have not already been accounted for by simpler considerations. In particular, the changes at A9, A10, 10-Me and 8-Me are significant. Going from erythromycin A to erythromycin B (removal of the 12-hydroxy group) the changes are A9 upfield 2.2 ppm, A10 downfield 1.1 ppm, 10-Me downfield 2.1-2.4 ppm, and 8-Me upfield 3.0 ppm. Comparison of the pmr spectra for these two indicates

Table 20

Comparison of Some Alcohol Spectra

	C1	C2	C3	C4	C5	M_e
4-hydroxy-4-methyl-2-pentanone	161.0	-16.7	137.8	123.3	163.2	163.2
2-butanol[a]	169.9	123.8	160.5	182.6	-	-
2-methyl-2-butanol[a]	163.9	122.2	156.0	184.0	-	163.9
2-pentanol[a]	169.2	125.5	150.9	173.4	178.5	-

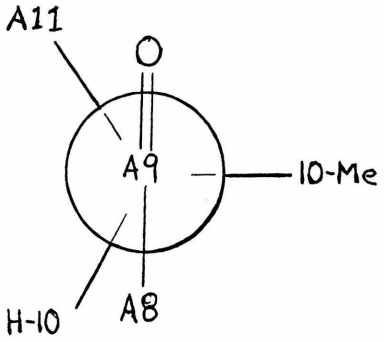
[a]Data from Ref. 17a, p.141

no significant change in any of the vicinal proton-proton couplings on the ring[15]. Further, it was demonstrated earlier that the dominant interaction at the A9 ketone is the hydrogen bonding with the hydroxyl group at A6. All of this is consistent with the following conformational difference between erythromycins A and B. In erythromycin A the 10 methyl moves down to avoid the 1,3-diaxial interaction with the 12-hydroxyl by a rotation of the 9-10 bond. This moves the ketone into the ring closer to the 6-hydroxyl. There is no change in the 10-11 bond but a slight rotation around the 11-12 bond (Figure 20). Further evidence for this change can be obtained from the CD data for these two ketones as determined by Mitscher et al.[45]. The effect of this change is to drive the 10-Me into a positive quadrant of the ketone in erythromycin A. The observed rotations are -6600 degrees for erythromycin A and -10800 for erythromycin B. Since the 12-hydroxy is in a negative quadrant, its removal should cause an increase rather than a decrease in the observed rotations going from erythromycin A to B. This change is also consistent with the cmr data. The A9 ketone is closer to the 6-hydroxy in erythromycin A and hence can be more strongly hydrogen bound. This explains the observed downfield shift. The 10-Me and A10 shifts move downfield as a result of reduced steric compression. The 8-Me moves upfield because of increased diaxial interaction with A10 across the ketone. Alternatively, a rotation around the 8-9 bond could also occur. Since pmr and cmr indicate no significant differences between these two erythromycins anywhere else, ascertaining the nature of this change could be important in explaining the different antibiotic activities observed.

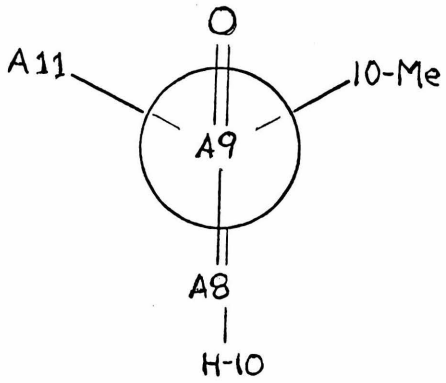
FIGURE 20

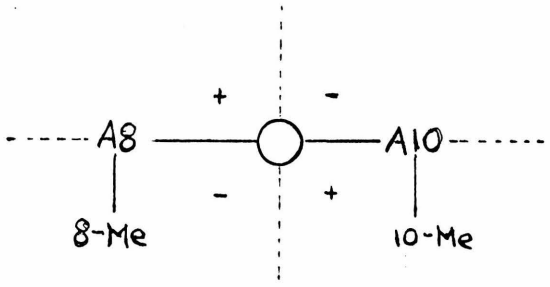
Conformational Difference between Erythromycins A and B

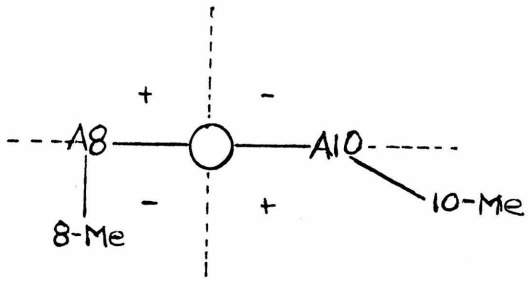
Erythromycin A



Erythromycin B







It would seem that conformational differences are also responsible for the bewildering scatter of the methine peaks in the aglycone derivatives. As discussed earlier, Egan[15] has considered this problem using the observed differences in vicinal proton-proton couplings. The differences between the proton spectra of erythronolide B and 11-acetylerythronolide B were explained by assuming that the two conformers shown in Figure 11 are populated differently in the two derivatives. In particular, 11-acetylerythronolide B was assumed to exist in conformer B more than the other derivatives[15]. Examination of the relevant carbon spectra offers evidence in favor of this argument. A7 in 11-acetylerythronolide B appears at 150.6 ppm, shifted considerably downfield from the other aglycone derivatives (Table 9). A similar downfield shift of A7 is observed in the spectra of anhydroerythromycin A and erythralosamine relative to erythromycin A and in the spectrum of the hemiketal of 5-deoxy-5-oxoerythronolide B relative to the ketoform. It has already been pointed out that the carbons in the furanose form of fructose appear downfield from those in the pyranose form[27]. The proposed conformation for 11-acetylerythronolide B has this region of the ring in a pseudo five-membered ring and the analogy to the hemiketals was also made by Egan[15]. Furthermore, this conformation has the A6 hydroxy group directed into the ring and poorly oriented for hydrogen bonding with the ketone at A9. The observed highfield shift of A9 in 11-acetylerythronolide B offers further evidence for this conformation.

It appears, however, that the spectra of the other derivatives which could be conformationally mobile may not follow this pattern. In Table 21 are shown the relevant data for several derivatives which differ only by their oxygen substituents. The strongest correlations among the first six examples are between A8, A9, H-8, and the vicinal couplings. The proposed equilibration[15] requires the first three to exist in conformation A (Figure 11) and the second three to exist in a mixture of conformations A and B. This is based on the observed proton chemical shifts of H-8 and the vicinal couplings. Those in the second group of three have values intermediate between those of the first group of three and those of 11-acetylerythronolide B (Table 21). Further evidence for this proposal is seen in the observed shifts of A9. The same averaging pattern is apparent. However, the data for A8, A7, 8-Me and A10 are not consistent in any obvious way. In particular, a downfield shift of A7 would be required in the second group of three relative to the first (independent of already ascertained steric effects at A7 due to substitution at A5). An alternative conformational change is consistent with the observed data. The change going from the first group of three to the second group of three in Table 22 is shown in Figure 21. The principal effects of this change would be to move A9 away from A6, thus weakening the hydrogen bond and causing the observed upfield shift; increase the diaxial interaction of A8 with the hydroxy at A6 causing an upfield shift in A8; make the vicinal 7-8 proton couplings more nearly equal; and move H-8 away from the ketone causing a downfield shift of the proton. The effects at other centers are harder to ascertain, but a slight increase in the

Table 21

Comparison of cmr and pmr Data for Several Aglycone Derivatives

	A8	A7	A9	A10	8-M_e	H-8[b]	$J_{7a,8}$[a]	$J_{7e,8}$[a]
(2)	147.6	154.3	-26.4	153.2	183.6	2.76	10	3
(6)	147.1	151.4	-26.3	153.6	183.5-184	2.70	9	4
(15)	147.1	155.0	-27.3	154.0	183-183.5	2.74	9	3
(5)	150.6	151.6	-	153.1	185.0	-	7	7
(16)	150.3	155.8	-22.5	152.5	182.4-183.7	2.88		
(7)	149.8	151.6	-22.1	152.7	185.3	2.80	7	8
(11)	149.9	150.6	-20.5 ±1	153.4	184.8	3.18	3	12

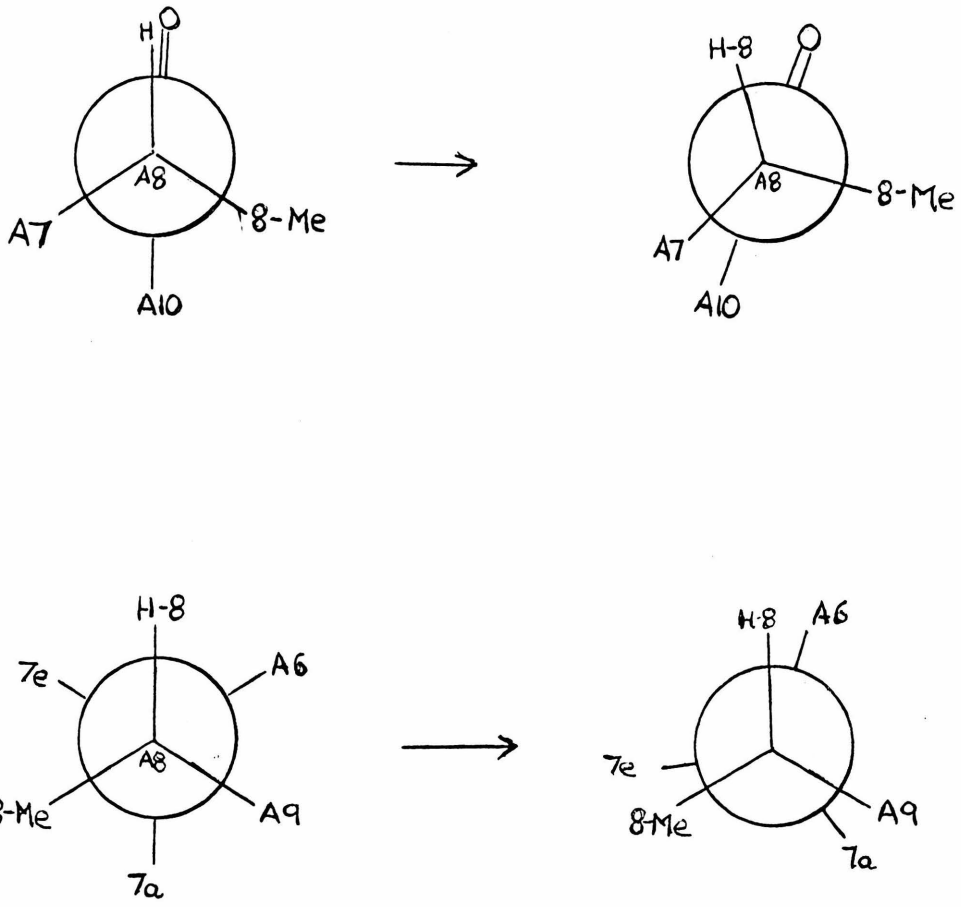
[a]Approximate values, taken from Ref. 15

[b]From Ref. 15

(2) erythromycin B
(6) 3-O-α-L-mycarosylerythronolide B
(15) 3,5-diacetylerythronolide B
(5) erythronolide B
(16) 3,5,11-triacetylerythronolide B
(7) 5-O-β-D-desosaminylerythronolide B
(11) 11-acetylerythronolide B

FIGURE 21

Aglycone Conformational Change



6-Me, 8-Me interaction is likely. The resulting conformation resembles a cis-1,3-dimethylcyclopentane at A6-A8, but the "axial" proton on A7 still points into the aglycone ring.

A number of changes other than those already discussed occur in the 6-deoxy derivatives. Relative to the 6-oxygenated derivatives, 6-deoxyerythronolide B and 11-acetyl-6-deoxyerythronolide B have A8 upfield 3 ppm and A10 downfield 4-5 ppm (Table 10). The upfield shift of A8 is most probably caused by an increased 1,3-diaxial interaction of its proton with 6-Me. This is consistent with the proposed conformation and observed vicinal couplings. The reason for the downfield shift of A10 is not apparent but probably involves a change in the diaxial interaction across the ketone A9, since no changes are observed in the A11-A13 region in either the proton or carbon spectra. The change in going to the 5-oxo derivatives (Table 10) is amazingly large. A8 moves downfield 7 ppm and A10 moves upfield 5 ppm. The conformation proposed[42] for 5,6-dideoxy-5-oxoerythronolide B does not seem consistent with either the proton or carbon data. In Figure 22 is shown a conformation compatible with the observed data in Table 22. This conformation is homogeneous to the other aglycones except that A6, A7 and A8 are eclipsed as in a rigid cyclopentane with the A6 proton pointing into the ring. The reasons for observed changes in the carbon spectra are now apparent. The 1,3-diaxial interaction between the 6-methyl and the 8-proton has been relieved while the interaction between the 8-methyl and 10-hydrogen has been increased. This could cause the observed shifts. This conformation is also consistent with the observed vicinal 6-7 and 7-8 proton-proton couplings.

Table 22

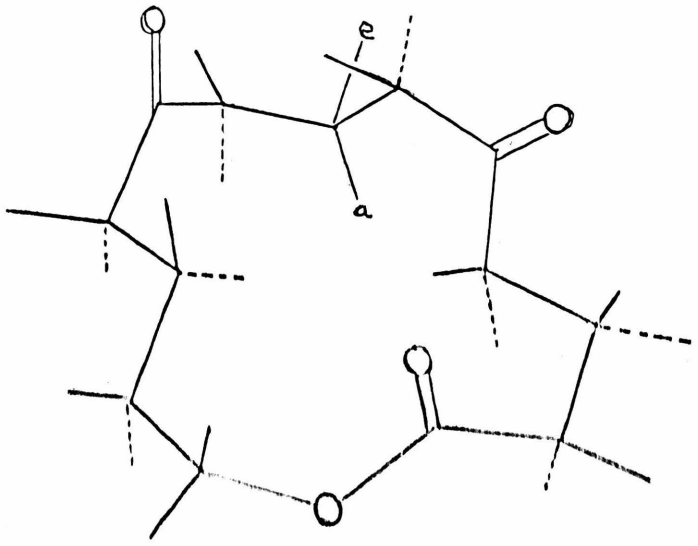
cmr and pmr Data for 5,6-dideoxy-5-oxoerythronolide B

	A8	A10	$J_{6,7a}$	$J_{6,7e}$	$J_{7a,8}$	$J_{7e,8}$
(9)[a]	145.9	153.4	8.2	8.2	3.0	13.0
(4)[a]	152.8	148.6	4.7	10.2	13.0	4.0

[a]Proton data from Ref. 42

[b]Proton data from Ref. 15

FIGURE 22



One conformational problem that has received little attention is the orientation of the 6-deoxy sugars with respect to the aglycone ring. The cmr results seem to shed some light on this. In the crystal structure of erythromycin A [3] the sugars are roughly perpendicular to the ring with the 6-methyl groups above the plane of the ring.

A comparison of the free and glycosidated chemical shifts of the sugars is given in Table 23. Glycosidation of the β-anomer of desosamine has little effect as might be predicted[27]. Glycosidation of the α-anomer of cladinose has a greater effect, again as would be predicted. The interesting change here is the 2.8 ppm upfield of C3 on glycosidation. The 3-carbon on an α-hexose would be expected to move downfield after glycosidation with a non-freely rotating group due to a partial relief of the 1,3-diaxial interaction. In the crystal structure conformation, the C3-OMe group is closest to desosamine and is the only group that seems to have a significant interaction with anything else in the molecule. C3-OMe also moves upfield, but only slightly, on glycosidation. The expected change again would be a downfield one. All of this is consistent with a solution conformation similar to that in the solid-state in which the observed upfield shifts at C3 and C3-OMe are caused by the remote (through many bonds) steric compression.

Egan[15] briefly mentioned observations that the proton shift of the methoxy group on cladinose was sensitive to the substitution at the 3-position of desosamine. These comparisons can also be made in cmr spectra. C3-OMe is found at 142.9, 143.2 and 143.0 in erythromycin

Table 23

Comparison of Free and Glycosidated Sugars

	C1	C2	C3	C4	C5	C6	C3-M_e	C3-OM_e
α-cladinose	101.1	156.0	116.9	114.9	128.3	174.6	171.9	142.5
α-cladinosyl[a]	96.0	157.3	119.7	114.4	127.0	174.1	171.6	143.1

	D1	D2	D3	D4	D5	D6	D-NM_{e2}
β-desosamine	94.4	120.8	127.5	163.4	123.5	171.6	152.1
β-desosaminyl[a]	89.6	121.5	126.8	163.9	123.5	171.6	152.1

[a]Average values

A, de-N-methylerythromycin A (CH_2Cl_2 solvent) and 3'-de-dimethylamino-3',4'-dehydroerythromycin A respectively. The differences between these shifts are too small to be significant. The reverse comparison can also be made. Megalomycin A (Figure 18) has a mycarosyl substituent at A3 and desosaminyl at A5. Mycarose differs from cladinose by the lack of the 3-O-methyl group. The shift of D3-NMe_2 in megalomycin A is 152.2 compared to 152.1 for erythromycin A. Further comparison of the de-N-methyl spectra with the natural antibiotics does yield interesting results. Neglecting effects already discussed on the desosamine shifts and changes which appear to result from overlapping peaks, a number of small but significant differences are apparent (Table 24). The effect at A5 may be dismissible as a remote substituent or steric effect but the others require an explanation. Working with models of the proposed information (Figure 8b) indicates that it is possible for the desosamine to move around to the side of the aglycone ring by rotation principally around the A5-O bond. This places the D-NMe_2 group near the 8-Me and the D2 proton near A7. Since most of the changes occur in the A7-A10 region, this seems to be a viable explanation for the observed changes on going to the de-N-methyl derivatives. Further evidence for the existence of this freedom for desosamine can be cited. Comparison of the desosamine peaks in the various anhydro derivatives (Table 14) shows significant variation in the desosamine peaks, particularly D2 and D3. Meanwhile, almost no changes are observed in the cladinose spectra. Furthermore, A7 is upfield in erythromycin A and B relative to the unsubstituted (at A5) aglycones by about 3 ppm. It has already been observed that

Table 24

Comparison of cmr Data for Several Modified Desosamine Derivatives

	A9	A5	A3	4-M_e	A8
(1)[a]	-29.0	109.2	112.7	183.6	147.6
(18)[a]	-28.2	108.2	112.4	183.1	147.9

	A9	A5	A3	4-M_e	A8	8-M_e	A7	A10
(2)[b]	-27.2	108.7	112.1	183.3	147.6	183.3	154.6	153.3
(19)[b]	-27.1	107.4	111.9	183.1	147.8	183.1	154.0	152.9

	A9	A5	A8	A10	8-M_e	D1
(1)[b]	-29.4	108.7	147.4	154.3	180.3	89.1
(20)[b]	-29.1	108.3	147.1	153.8	180.6	89.6

[a] CH_2Cl_2 solvent

[b] $CDCl_3$ solvent

(1) erythromycin A
(2) erythromycin B
(18) de-N-methylerythromycin A
(19) de-N-methylerythromycin B
(20) 3'-dedimethylamino-3',4'-dehydroerythromycin A

acetylation at A5 causes an upfield shift at A7. Thus the observed upfield shifts of A7 are consistent with the steric interaction that would be caused by movement of the desosamine ring. All of this evidence seems consistent with a situation in which cladinose is in the crystal state conformation and relatively fixed, while desosamine can move in the manner described.

Comparison of the spectra of the various anhydro and ketal derivatives (Tables 9-11, 14-17) verifies the pronounced downfield shifts of the carbons which end up in five-membered rings. Some examples are given in Table 25.

Going from erythromycin A to anhydro erythromycin A, a 17 ppm downfield shift of A11 is observed. While this may be only due to the five-membered ring formation, it is possible that this shift is indicative of steric relief at A11 in forming this derivative. As such, this may be part of the driving force for this reaction. The downfield shifts of A8 and A7 seem characteristic for the formation of the hemiketal and the saturated spiroketal. Since derivatives of this type play a major role in macrolide chemistry, these shifts could be used diagnostically. Certainly, other physical methods are capable of detecting these changes. The upfield shifts of A3 are almost certainly due to the steric interaction that can occur between the A6-oxygen and and A3-hydrogen.

H. Enriched Studies

It was mentioned earlier that erythromycin is believed to originate biosynthetically from seven propionate units. The

Table 25

Comparison of Some cmr Data for Anhydro Derivatives

	A3	A11	A8	A7
(1)	112.4	123.5	147.4	153.9
(21)	116.5	106.5	141.6	151.0
(22)	121.9	64.0	145.8	149.6
(2)	112.0	123.0	147.6	154.3
(23)	116.5	121.9	91.4	150.1

(1) erythromycin A

(21) anhydroerythromycin A

(22) erythralosamine

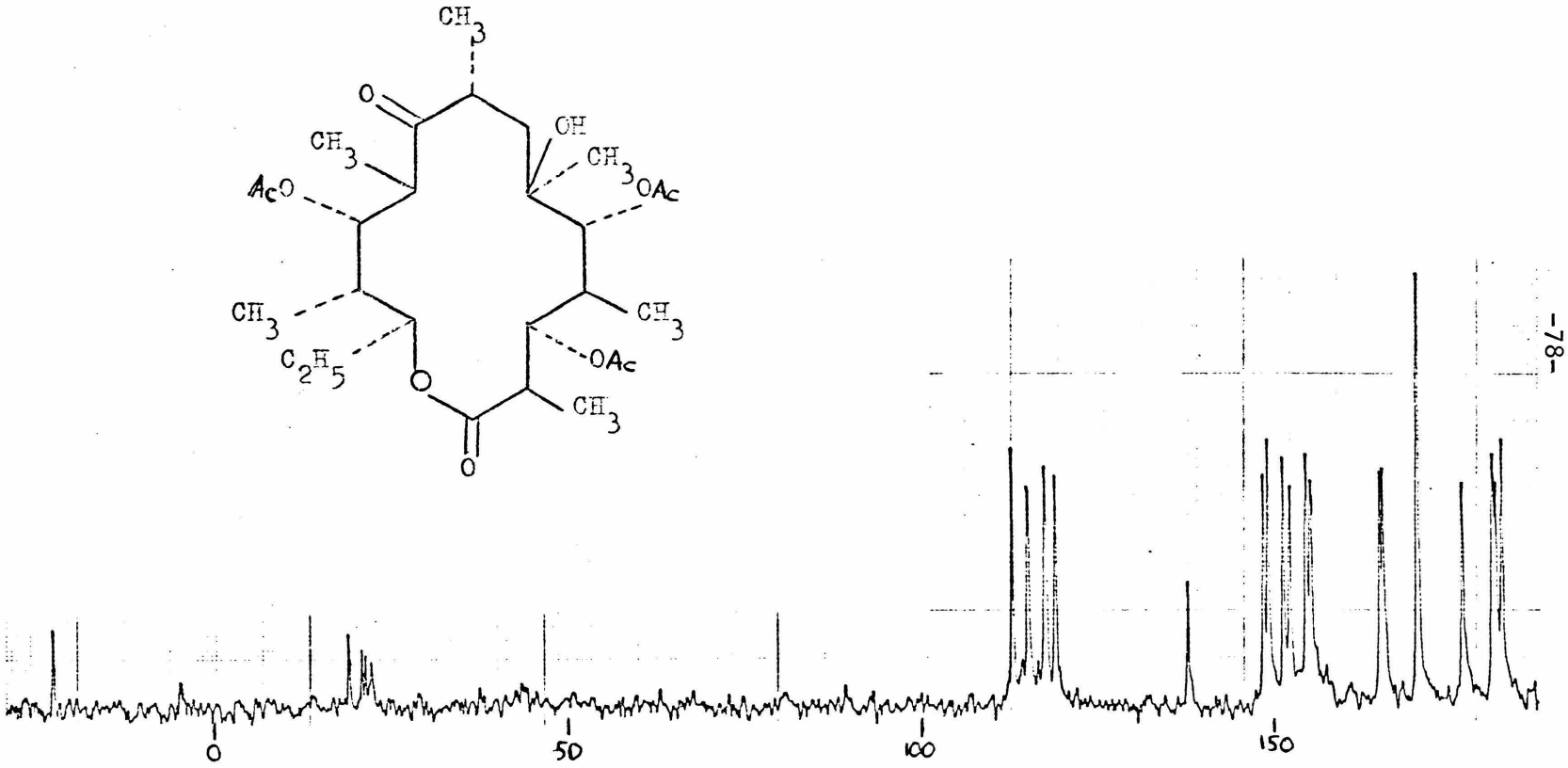
(2) erythromycin B

(23) 8,9-anhydroerythromycin B-6,9-hemiketal

experimental evidence for this consists of ^{14}C labelling studies[5]. These studies did not dissect the molecule into its actual precursors, but rather, convenient degradations were done which mix individual enriched carbons. As a test of the propionate theory and of the usefulness of cmr in biosynthesis studies, feeding experiments were done by Dr. James B. McAlpine at Abbott. A blocked mutant of streptomyces erythreus which accumulates erythronolide B was fed C1 labelled propionate. The resulting enriched erythronolide B was converted into the 3,5,11-triacetate to improve solubility. Spectra of an unenriched and the enriched samples are shown in Figures 23-24. It is evident from the assignments given earlier that A1, A3, A5, A7, A9, A11 and A13 are enriched as expected. That A13 is enriched rather than A6 with which it is coincident is evident from the intensity. A6 is quaternary and would be of much lower intensity. Also, undecoupled spectra were run. From what can be seen of the remaining unenriched part of the spectrum in the enriched sample, no differences in intensity are evident compared to the unenriched sample. The enrichment appears to be about twenty-five times over natural abundance. While the biosynthetic potential is dramatically indicated, the very efficient, specific enrichment suggests that these samples could be used to determine carbon-hydrogen and carbon-carbon coupling constants. In particular, vicinal carbon-hydrogen couplings have an angular dependence[46] and if resolvable could be used in conformational studies. For example, the sugar orientations would be determined by the two dehedral angles A3-O-C1-C1H and A3H-A3-O-C1 for cladinose and A5-O-D1-D1H and A5H-A5-O-D1 for desosamine. A3 and A5 are easily enriched as shown

FIGURE 23

Noise-decoupled CMR Spectrum of Unenriched 3,5,11-triacetylerythronolide B



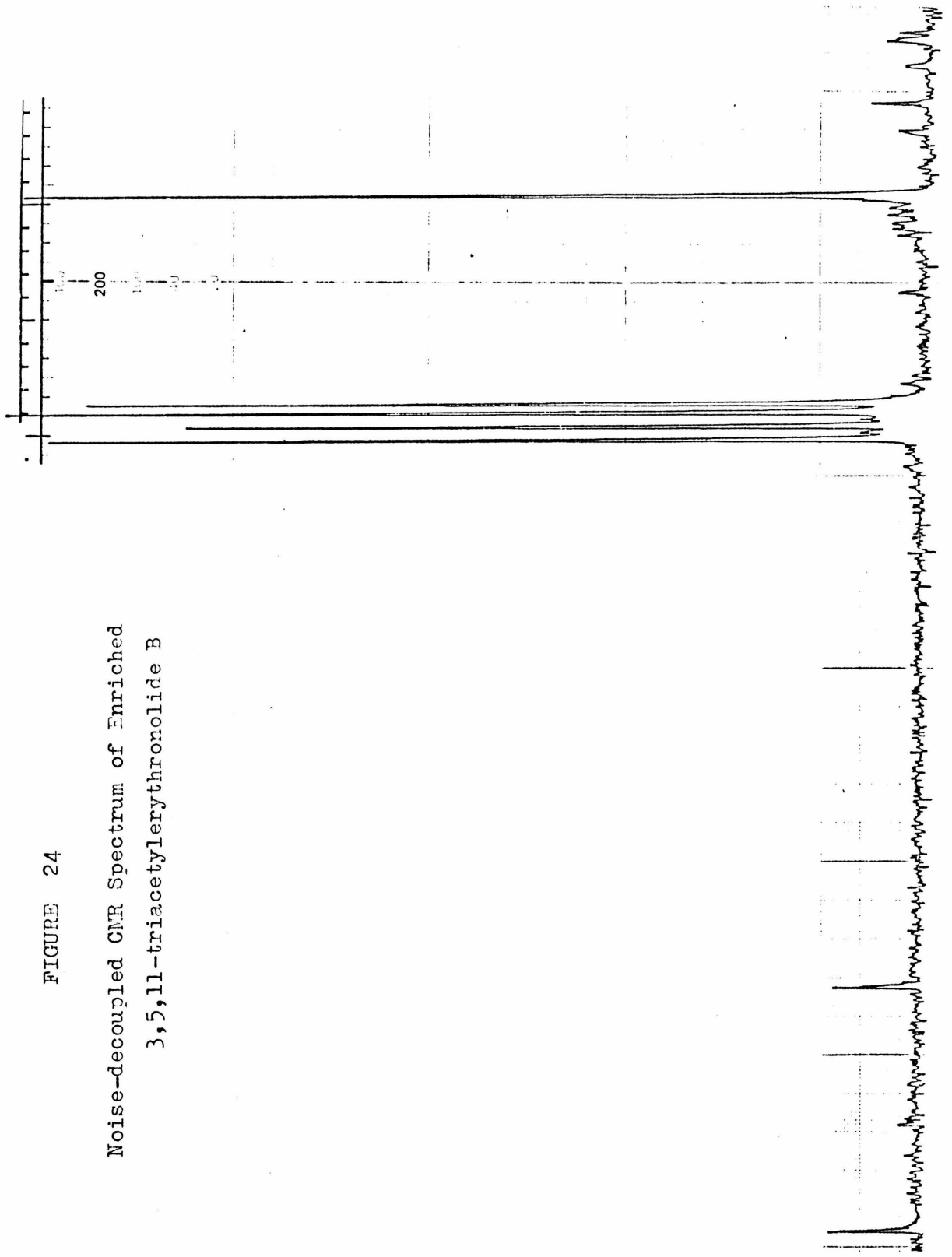


FIGURE 24

Noise-decoupled CMR Spectrum of Enriched
3,5,11-triacetylerythronolide B

above. D1 and C1 could be enriched using glucose (in a separate experiment of course). The resolution of the desired couplings would be nontrivial because of the many geminal and vicinal couplings possible.

I. Summary

The assigned cmr spectra of this series of erythromycin derivatives have been interpreted in light of known effects on carbon spectra and conformational differences among the derivatives. Some new proposals were made concerning these differences and evidence offered. In these compounds it appears that conformationally, natural abundance cmr is most useful in questions of intramolecular hydrogen bonding and orientation across and around carbonyls and other quaternary centers. Diagnostic usefulness for the various ketal structures is also suggested. Enriched cmr is shown to be useful biosynthetically and its potential for conformational studies based on carbon spectra fine structure is suggested.

II. ALGEBRAIC CHEMISTRY

A. Isomer Counting

The problem of counting the number of isomers of a chemical structure has received chemists' attentions for many years[47]. Aside from the fascination of the combinatorial problem involved, isomer counting is a necessity in any problem involving isomerizations. In particular, polytopal and carbonium ion rearrangement problems have received considerable attention recently[48].

The combinatorial formulas required for isomer counting problems were derived by Polya[49]. More recently, a paper appeared describing the adaption of Polya's permutation group method to the more familiar point group formalism[50]. The method will be illustrated by means of an example. Consider the problem of the number of isomers of an octahedrally coordinated metal with two different substituents $m(A)_n(B)_{6-n}$. The relevant symmetry group is the octahedral rotation group O (Figure 25). Each element of the group permutes the 6 vertices of the octahedron. In this way the group is represented by the permutation of the vertices. Each member of a conjugacy class is represented by a permutation of the same cyclic type. That is, the sizes of the sets of vertices which are closed with respect to the permutation (called orbits or cycles) are the same for each member of a conjugacy class. For example, the C_4 rotations all have cyclic type (a)(b)(cdef).

From this information the cycle index Z(G) can be constructed. For an arbitrary group G:

FIGURE 25

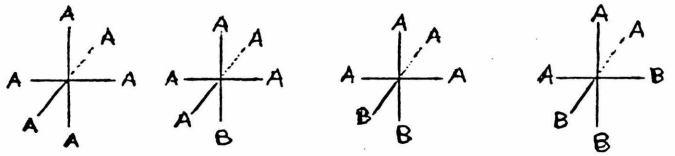
Example of Isomer Counting Calculation

	E	$6C_4$	$6C_2$	$3C_2'$	$8C_3$
cyclic type		(a)(b)(cdef)	(ab)(cd)(ef)	(a)(b)(cd)(ef)	(abc)(def)

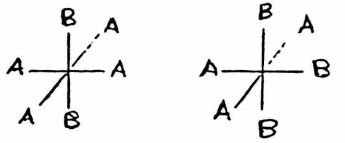
cycle index $1/24(f_1^6 + 6f_1^2f_4 + 3f_2^3 + 6f_1^2f_2^2 + 8f_3^2)$

no. isomers $1/24((A+B)^6+6(A+B)^2(A^4+B^4)+3(A^2+B^2)^3+6(A+B)^2(A^2+B^2)^2+8(A^3+B^3)^2)$

$= A^6 + A^5B + 2A^4B^2 + 2A^3B^3 + 2A^2B^4 + AB^5 + B^6$



etc.

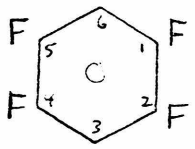


$$Z(G) = \frac{1}{|G|} \sum_{g \in G} f_1^{n_1(g)} f_2^{n_2(g)} \cdots f_m^{n_m(g)}$$

The symbol $|G|$ refers to the order (number of elements) of the group The symbol $f_p^{n_p(g)}$ means that for the permutation g there are $n_p(g)$ cycles of length p. In the example a C_4 rotation has two cycles of length one and one cycle of length four, thus the term $f_1^2 f_4$ is contributed to the cycle index. Since all C_4 rotations have the same cyclic type, the coefficient is the number of members of the conjugacy class. To determine the number of isomers of the desired type, the expression $(A^p + B^p)$ is substituted for f_p in the cycle index. The resulting polynomial is expanded and terms collected. The coefficient of the term $A^p B^{m-p}$ is the number of isomers with (p) A substituents and (m-p) B substituents.

This method seems to have been used only for determining the number of permutation isomers of a structure with a designated number of different achiral ligands. Application to dissymetric molecules seems to be straightforward. The problem is treated in the same way except that the permuted substituents A and B now refer to the two absolute configurations of constitutionally identical chiral ligands. As an example, consider a benzene substituted in the 1,2,4, and 5 positions with identical chiral ligands (symbolized F; F written backwards refers to the enantiomeric configuration, Figure 26). This structure has D_2 rotation symmetry. The cycle index is computed in the same way as before, the substitution $(A^p + B^p)$ is made for f_p. The only difference is that A and B are dummy substituents which refer

FIGURE 26

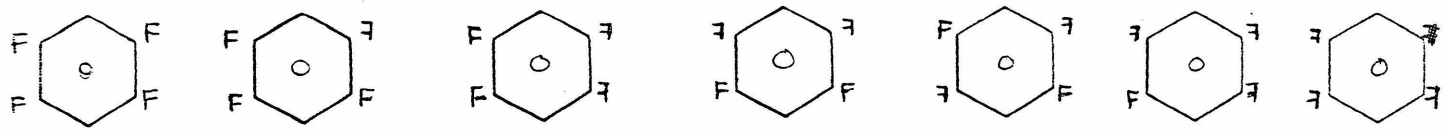


$$Z(D_2) = \frac{1}{4}\left(f_1^4 + f_2^2 + f_2^2 + f_2^2 \right)$$

E C_{2x} (12)(45) C_{2y} (15)(24) C_{2z} (14)(25)

$$= \frac{1}{4}\left((A+B)^4 + 3(A^2+B^2)^2\right)$$

$$= A^4 + A^3B + 3A^2B^2 + AB^3 + B^4$$



A^4 A^3B A^2B^2 A^2B^2 A^2B^2 AB^3 B^4

to absolute configurations. The number of stereoisomers is verified by the drawings. No general method seems to have been given for determining the number of stereoisomers in a dissymetric system. Some formulas which work in special cases have been described[51]. Further, it has been stated that Polya's theorem is inadequate in cases with chiral ligands[52]. This statement was based on the generally accepted concept that there are two causes for the reduced number of isomers in dissymmetric systems called <u>enantiomeric reduction</u> and <u>diastereomeric reduction</u>. A diastereomeric reduction is termed one caused by a rigid rotation. The three isomers shown in Figure 27 are equivalent by a C_3 rotation (open circles and filled circles represent enantiomeric configurations of the same ligand). An enantiomeric reduction is termed one caused by symmetrical substitution of ligands of opposite chirality to give a meso compound (Figure 28). These are, in fact, exactly the same thing. The reduction in the number of isomers in both cases is caused by the presence of a rotation axis of symmetry in the original structure without chiral ligands. If there were no rotation symmetry in this latter case (Figure 28) then <u>two</u> diastereomeric meso structures would be possible and there would be no reduction in the number of isomers. Polya's method was stated to be inadequate for systems exhibiting enantiomeric reduction[52]. Several examples of systems said to exhibit enantiomeric reduction were described and are correctly treated using the Polya method. The abbreviated calculations are shown in Figure 29. Note that substituting 2 for (A+B) gives the number of isomers without the distribution.

FIGURE 27

Example of Diastereomeric Reduction

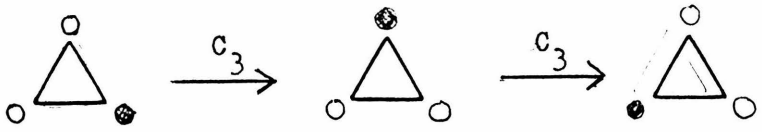


FIGURE 28

Example of Enantiomeric Reduction

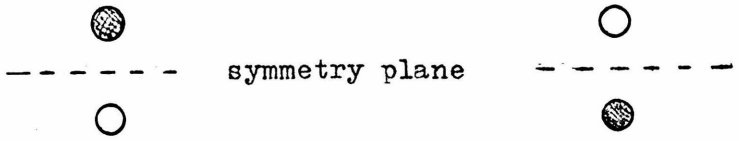
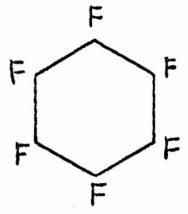


FIGURE 29

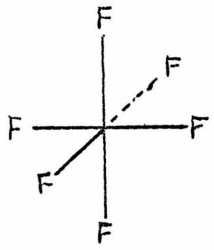
Examples of Isomer-Counting Calculations on Dissymetric Systems



D_6 E $2C_6$ $2C_3$ C_2 $3C_2'$ $3C_2''$

$$Z(D_6) = 1/12(f_1^6 + 2f_6 + 2f_3^2 + f_2^3 + 3f_2^3 + 3f_1^2f_2^2)$$

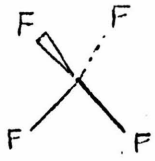
$$\text{no. isomers} = A^6 + A^5B + 3A^4B^2 + 3A^3B^3 + 3A^2B^4 + AB^5 + B^6$$



O E $8C_3$ $6C_2$ $6C_4$ $3C_2$

$$Z(O) = 1/24\ (f_1^6 + 8f_3^2 + 6f_2^3 + 6f_1^2f_4 + 3f_1^2f_2^2)$$

$$\text{no. isomers} = A^6 + A^5B + 2A^4B^2 + 2A^3B^3 + 2A^2B^4 + AB^5 + B^6$$



T E $8C_3$ $3C_2$

$$Z(T) = 1/12\ (f_1^4 + 8f_1f_3 + 3f_2^2)$$

$$\text{no. isomers} = A^4 + A^3B + A^2B^2 + AB^3 + B^4$$

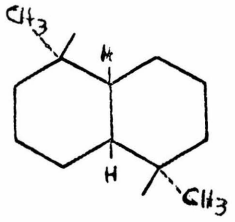
This method can be applied to any dissymmetric system, not just those which resemble permutation problems. Three more examples are shown in Figure 30. In the first, the decalin system, the permutations of the four asymmetric centers by the C_2 rotation has to be determined on the structure in which all constitutionally identical centers have the same absolute configurations. In the second example, α-onocerin, a naturally occurring triterpene, is treated. Chiral planes and axes[53] can also be handled. The third example shows a structure with two planes and two centers of chirality.

The effect of changes in chemical structures on the number of isomers can also be readily determined. In Figure 31 four related problems are treated. Structure A can be thought of as a benzene substituted with 6 constitutionally identical amino acids. Using the rotation group D_6, the number of possible stereoisomers is calculated to be thirteen. Structure B is the same as structure A except that the amide bonds have been formed. The rotation group is now C_6 and fourteen isomers are found. The extra isomer results because of the formation of a cycloenantiomeric pair[54]. Structure C is the same as A except for the indicated metal coordination. Again the rotation symmetry is C_6 and fourteen isomers result. Structure D combines structures B and C. Here a plane of chirality has been added since the faces of the ring are different after cyclization. Again the rotation symmetry is C_6 and twenty-eight isomers are found.

The number of permutation isomers of a structure including chiral ligands can be determined in the same way as the number of

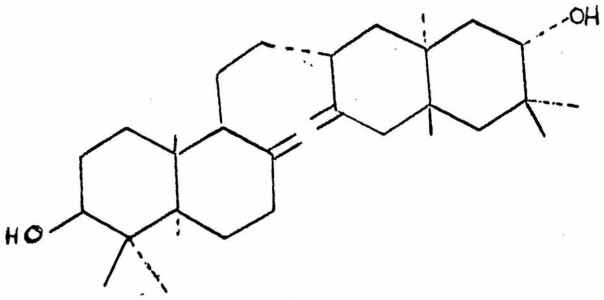
FIGURE 30

Examples of Isomer-Counting Calculations



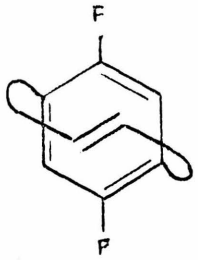
$Z(C_2) = 1/2 (f_1^4 + f_2^2)$

no. isomers = $1/2 (2^4 + 2^2) = 10$



$Z(C_2) = 1/2 (f_1^8 + f_2^4)$

no. isomers = $1/2 (2^8 + 2^4) = 136$

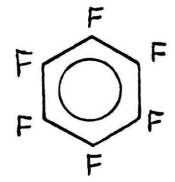


$Z(C_2) = 1/2 (f_1^4 + f_1^2 f_2)$

no. isomers = $1/2 (2^4 + 2^3) = 12$

FIGURE 31

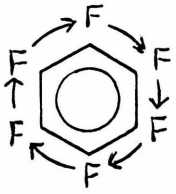
Effect of Change of Symmetry on Isomer-Counting Calculations



A

$Z(D_6) = 1/12(f_1^6 + 2f_6 + 2f_3^2 + f_2^3 + 3f_2^3 + 3f_1^2f_2^2)$

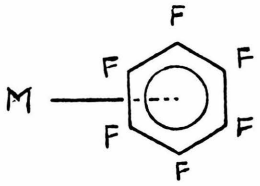
no. isomers = 1/12 (156) = 13



B

$Z(C_6) = 1/6\ (f_1^6 + 2f_6 + 2f_3^2 + f_2^3)$

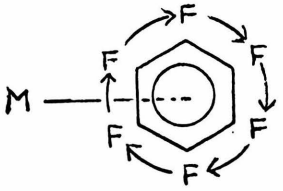
no. isomers = 1/6 (84) = 14



C

$Z(C_6) = 1/6\ (f_1^6 + 2f_6 + 2f_3^2 + f_2^3)$

no. isomers = 1/6 (84) = 14



D

$Z(C_6) = 1/6\ (f_1^7 + 2f_1f_6 + 2f_3^2f_1 + f_2^3f_1)$

no. isomers 1/6 (168) = 28

permutation isomers with achiral ligands were determined above. Another system[52] said to exhibit enantiomeric reduction is the cyclobutane structure shown in Figure 32 with four achiral ligands A, and four constitutionally identical chiral ligands with two in each absolute configuration. The relevant symmetry group is D_4. The cycle index is computed as usual and the substitution $(A+B+C)^p$ for f_p is made. The coefficient of the terms $A^4B^2C^2$ is needed. Isolation of a single term can be done in several ways. The fully generalized Polya theorem considers this[55]. The resulting coefficient is 480 which means there are 480/8 = 60 isomers. This total is verified by the structure drawings. Isomers designated with the number 2 are chiral and only one enantiomer is shown.

Further information about an isomeric system can be obtained by noting that the terms arising from the cycle index for each symmetry element represent the permutation isomers which have that particular kind of rotation symmetry. Returning to the octahedral example in Figure 26, the cycle index for a C_4 rotation is $f_1^2f_4$. Substitution yields the polynomial $A^6+2A^5B+A^2B^4+A^2B^4+2AB^5+B^6$. These terms correspond to the permutation isomers (not distinct isomers) shown in Figure 33. Each of these structures has a C_4 axis of symmetry, and in this manner all isomers with at least C_4 symmetry can be determined. To find isomers with more than one rotation axis, a <u>combined cycle index</u> is created. As shown in Figure 34 the indicated C_4 and C_2 axes are considered. The cycles used to construct the cycle index are those that are intact after <u>both</u> symmetry operations. The resulting term in this

Figure 32

Cyclobutane Isomers with four Chiral Ligands

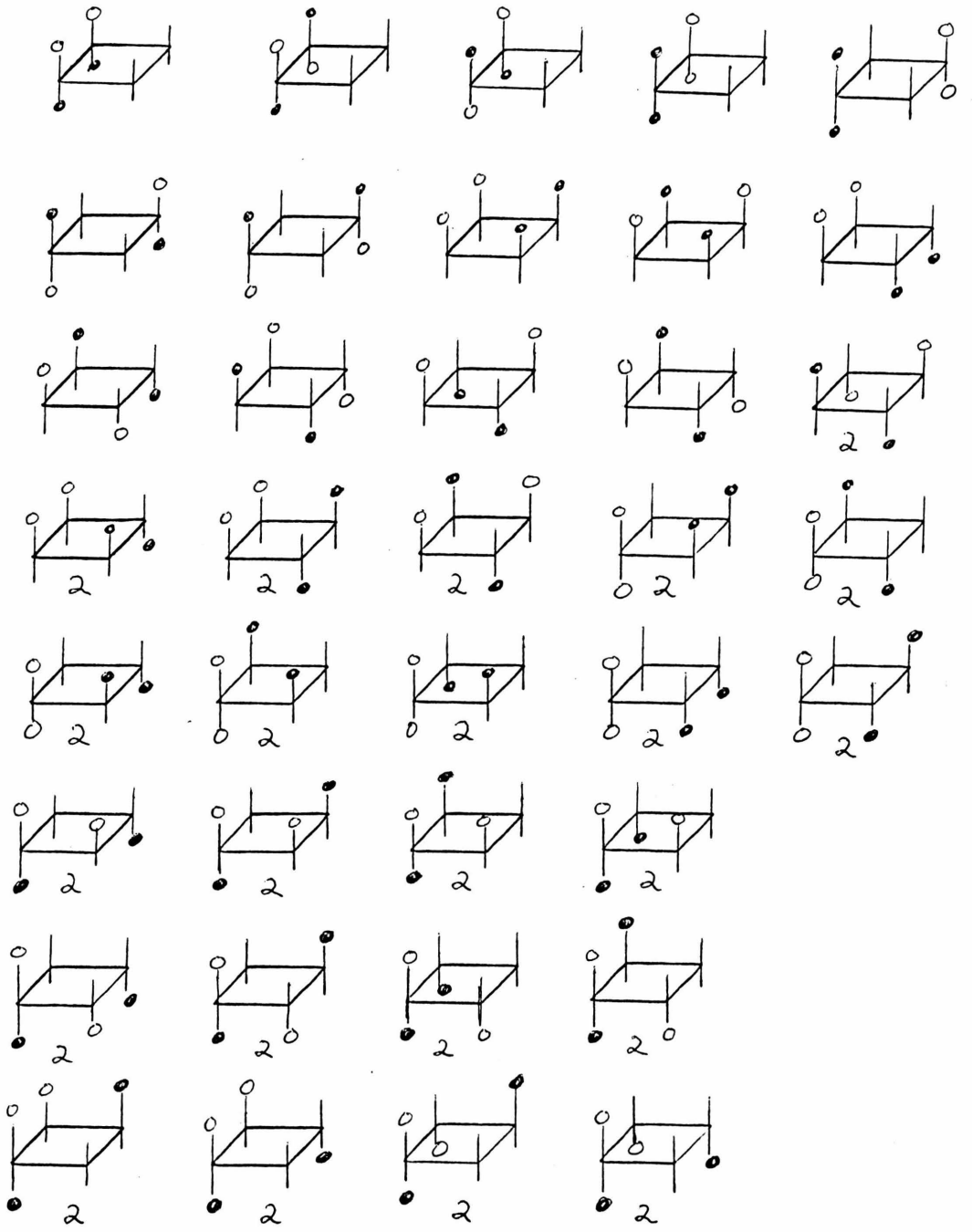
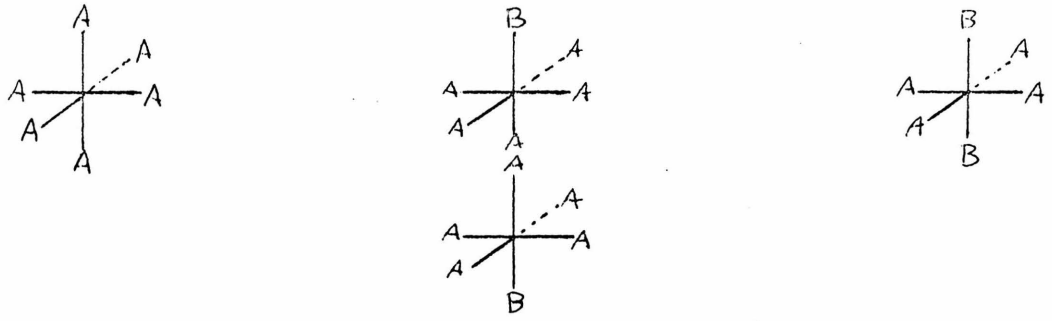


FIGURE 33

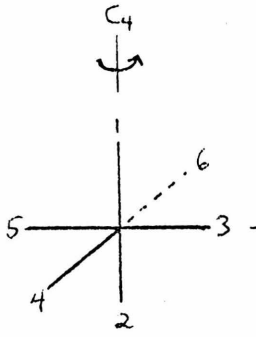
Permutation Isomers with C_4 Symmetry



A^6 + $2A^5B$ + A^4B^2 etc.

FIGURE 34

Illustration of Combination of Symmetry Elements

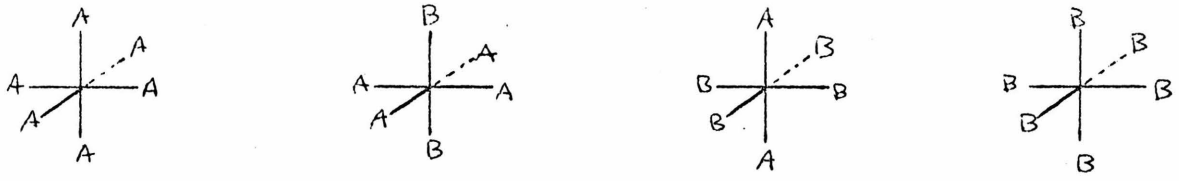


$C_4 = (1)(2)(3456)$

$C_2 = (12)(3)(5)(46)$

overall, $(12)(3456) \longrightarrow f_2 f_4$

A^6 + A^4B^2 + A^2B^4 + B^6



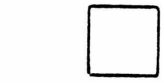
case is $f_4 f_2$ and substitution yields the polynomial $A^6 + A^4B^2 + A^2B^4 + B^6$. The permutation isomers are shown in Figure 34. This procedure can also be used to determine the reflective (improper rotation) symmetry of isomers but only when achiral ligands are involved. The cycle index is computed in the same way using the permutation representation.

The problem with chiral ligands is that they change configuration on reflection and that no planes or alternating axes of symmetry can pass through them. Therefore, determining the reflective symmetry in such cases must be done differently. Clearly a structure with chiral ligands which is itself achiral must have an equal number of ligands in opposite configurations. Therefore only symmetry operations represented by cycles of even length are possible. All of this can be taken into account by substituting $(A)^n + (-A)^n$ for f_n in the cycle index for the reflective elements of symmetry. If n is odd, this term will go to zero. As an example, the number of meso forms possible in a square planar structure with two constitutionally different chiral ligands will be computed in Figure 35.

A simple relationship between the order of the rotation subgroup of an isomer and its permutational degeneracy can be derived. As was pointed out earlier, substitution into the cycle index for a particular symmetry element counts the number of permutation isomers with that symmetry. Any permutation isomer will be counted once for each element of rotation symmetry it possesses. Since a distinct isomer must be counted $|G_r|$ times (where $|G_r|$ is the order of the rotation symmetry group used in the calculation), its permutation degeneracy times the order of its rotation subgroup must equal r. Note that the

FIGURE 35

Counting Isomers with Constitutionally Different Chiral Ligands



permute F, -F, G, -G

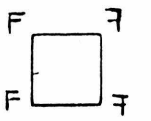
D_{4h} E $2C_4$ C_2 $2C_2'$ $2C_2''$ i $2S_4$ σ_h $2\sigma_v$ $2\sigma_d$

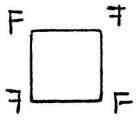
reflective cycle index $Z(RD_{4h}) = f_2^2 + f_1^4 + 2f_4 + 2f_1^2 f_2 + 2f_2^2$

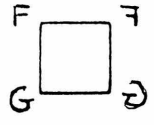
no. isomers $= 1/8\ (3(F^2+(-F)^2+G^2+(-G)^2)+2(F^4+(-F)^4+G^4+(-G)^4))$

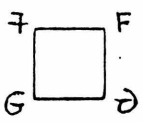
$= 1/8\ (3(2F^2+2G^2)^2 + 2(2F^4+2G^4))$

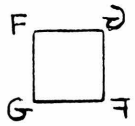
$= 2F^4 + 3F^2G^2 + 2G^4$

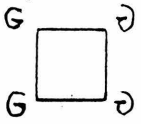


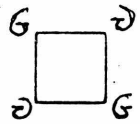












symmetry group of any isomer must be a subgroup of this group used in the calculation. Then the relationship is obtained:

$$|G_r| = (P_i)\ |G_{r_i}|$$

P_i = permutation degeneracy of isomer i

$|G_{r_i}|$ = order of rotation symmetry group of isomer i

Combining all the methods described so far allows one to do a complete analysis of an isomer problem. That is, one can determine the number, ligand distribution, and symmetry of all possible isomers. The only additional piece of information needed is a knowledge of the lattice of subgroups of the principal symmetry group. The method will be illustrated by means of an example. The number and distribution of chlorobenzenes will be determined as shown in Figure 36. First, the number and distribution of isomers are found as described earlier. There are 13 isomers distributed as shown. Next it will be determined if there are any isomers of D_{6h} symmetry (the most possible). All the symmetry operations of this group yield the combined cycle index f_6. Substitutions in the usual manner gives $A^6 + B^6$. As mentioned above, the permutation degeneracy times $|G_{r_i}|$ must equal $|G_r|$. In this case $|G_{r_i}| = |G_r| = 12$ and the permutation degeneration is one (the coefficient of the term in the resulting polynomial). Hence there are two isomers with D_{6h} symmetry, benzene and hexachlorobenzene.

Isomers of lower symmetry are now located. The largest subgroups of D_{6h} are of order 12. These are C_{6v}, C_{6h}, D_6, D_{3h} (twice), and D_{6d} (twice). C_{6v} yields the combined cycle index f_6. However, any isomer

FIGURE 36

Calculation of Possible Chlorobenzenes

D_6 E $2C_6$ $2C_3$ C_2 $3C_2'$ $3C_2''$

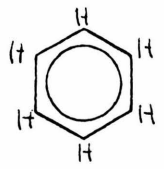
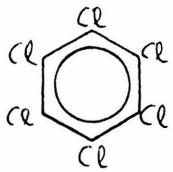
$$Z(D_6) = f_1^6 + 2f_6 + 2f_3^2 + f_2^3 + 3f_1^2f_2^2 + 3f_2^3$$

$$\longrightarrow A^6 + A^5B + 3A^4B^2 + 3A^3B^3 + 3A^2B^4 + AB^5 + B^6$$

13 possible isomers

D_{6h}: combined cycle index, $Z_c(D_{6h}) = f_6$

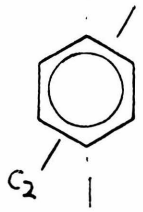
substitution gives $A^6 + B^6$, thus two isomers:

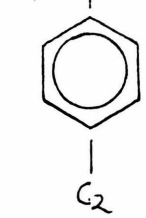



C_{6v}: $Z_c(C_{6v}) = f_6$; $f_6 - f_6 = 0$

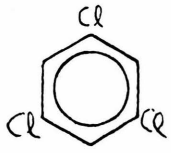
C_{6h}: $Z_c(C_{6h}) = f_6$; $f_6 - f_6 = 0$

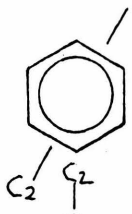
D_6: $Z_c(D_6) = f_6$; $f_6 - f_6 = 0$

D_{3h}:  $Z_c(D_{3h}) = f_6$; $f_6 - f_6 = 0$

D_{3h}:  $Z_c(D_{3h}) = f_3^2$; $f_3^2 - f_6 \neq 0$

$(A^3+B^3)^2-(A^6+B^6) = 2A^3B^3$

one isomer 

D_{3d}:  $Z_c(D_{3d}) = f_6$; $f_6 - f_6 = 0$

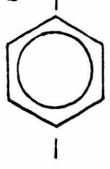
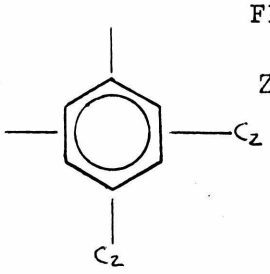
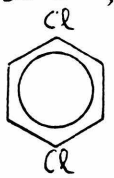
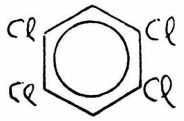
D_{3d}:  $Z_c(D_{3d}) = f_6$; $f_6 - f_6 = 0$

FIGURE 36 (Cont.)

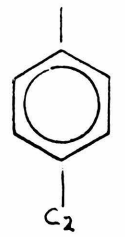
D_{2h}: 

$Z_c(D_{2h}) = f_2 f_4 \quad 3(f_2 f_4 - f_6) \neq 0$

$3A^4B^2 + 3A^2B^4$; two isomers

D_{3h} D_{2h} D_{6h}

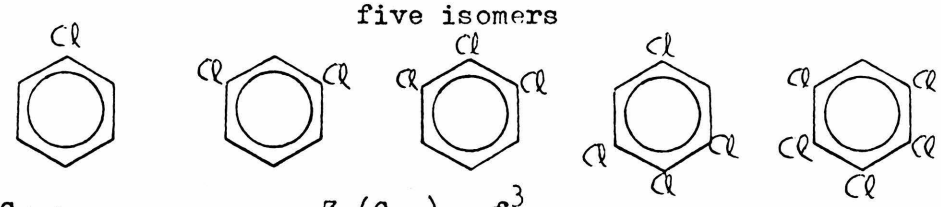
C_{2v}: 

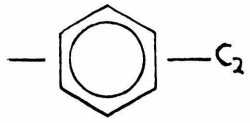
$Z_c(C_{2v}) = f_1^2 f_2^2$

$3 (f_1^2 f_2^2 - (f_3^2 - f_6) - (f_2 f_4 - f_6) - f_6)$

$3(2A^5B + 2A^4B^2 + 2A^3B^3 + 2A^2B^4 + 2AB^5)$

five isomers

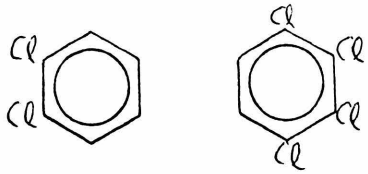


C_{2v}: 

$Z_c(C_{2v}) = f_2^3$

$3(f_2^3 - (f_2 f_4 - f_6) - f_6)$

$3(2A^4B^2 + 2A^2B^4)$; two isomers



C_s:

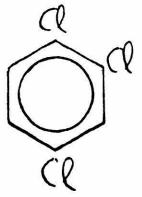
$Z_c(C_s) = f_1^6$

$f_1^6 - (f_6 + (f_3^2 - f_6) + 3(f_2 f_4 - f_6) +$

$3(f_2^3 - (f_2 f_4 - f_6) - f_6) + 3(f_1^2 f_2^2 -$

$(f_3^2 - f_6) - (f_2 f_4 - f_6) - f_6))$

$12A^3B^3$; one isomer



with D_{6h} symmetry will also have C_{6v} symmetry, hence the expression for determining the number of C_{6v} isomers is $f_6 - f_6 = 0$. Thus there are no C_{6v} isomers. Similarly there are no C_{6h} or D_6 isomers. The two possible D_{3h} subgroups yield different combined cycle indices. The one with the C_2 axes between the carbon atoms yields the combined cycle index f_6 and therefore there are no isomers of this symmetry. The other D_{3h} subgroup has its C_2 axes through the carbon atoms and the cycle index is f_3^2. Again any isomer with D_{6h} symmetry also has D_{3h} symmetry, so the expression is $f_3^2 - f_6$. As shown in Figure 36, substitution and cancellation gives the term $2A^3B^2$. Here the permutation degeneracy is two and $|G_{r_i}|$ is six, so there is one chlorobenzene (1,3,5-trichlorobenzene) with D_{3h} symmetry. No D_{3d} isomers are found.

The next group considered is D_{2h}, of which there are three identical subgroups. The combined cycle index for one of these is f_2f_4 so that the expressions for the number of isomers is $3(f_2f_4 - f_6)$. This yields the polynomial $3A^4B^2 + 3A^2B^4$. Since $|G_{r_i}| = 4$, there are two chlorobenzenes with this symmetry. The rest of the calculation is done in the same manner. Only subgroups yielding isomers are shown.

When chiral ligands are present the procedure has to be modified slightly. The rotation symmetry of all isomers is calculated in the same way. The reflective symmetry of meso forms is calculated by substituting ±A, ±B, etc. for each constitutionally different chiral ligand into the cycle index for the reflective symmetry elements of each subgroup. The same iterative procedure is used.

An alternative formulation of Polya's isomer counting method indicates further uses of the procedure. The set of permutation isomers form a permutation representation of the rotation symmetry group of the unsubstituted skeletons. This reducible representation can be decomposed into its irreducible components in the usual manner[56]. The number of times the totally symmetric representations appears is equal to the number of distinct isomers. This statement can easily be proved. All the permutation isomers of a distinct isomer are permuted among themselves by the action of the skeletal symmetry group. Further, each permutation isomer can be converted into any other by the action of some member of the group. Therefore the permutation isomers of each distinct isomer yield a transitive representation of the group. From standard group representation theory it is known that a transition permutation representation contains the totally symmetric representation exactly once[57]. The reducible representation obtained in this manner is identical to the cycle index after substitution of n^p for f_p where n is the number of substituents permuted. However, in other counting problems the relationship between the permutation representations and the cycle index is not as apparent. For example, in a recent paper, J. B. Hendrickson[58] discusses the design of syntheses of benzene derivatives. A relevant problem is the calculation of the number of possible one-step conversions of benzene with one kind of substituent. This can be accomplished by constructing a representation of D_6 using the number of permutation isomers times the number of remaining unsubstituted sites on each isomer. The resulting

reducible representation is shown in Figure 37. The totally symmetric representation is included twenty times, which is the number of possible one-step conversions. This representation can also be constructed by taking the derivative of the cycle index with respect to f_1 and substituting 2^p for f_p . The validity of this method can be easily verified. On an unsubstituted benzene ring there are six possible sites for substitution and five remaining sites which may already be substituted or not. This yields the term $6 \times 2^5 = 192$ as shown in Figure 38. The process of designating a site to be substituted destroys all symmetry except the C_2 axes which go through the carbon atoms. Hence all other terms go to zero as shown. Further, if $A^p + B^p$ is substituted for f_p , the distribution of possible conversions is obtained. This is shown in Figure 37 along with a graph of the possible changes. Similarly the second derivative yields the number of two-step substitutions with different substituents in a prescribed order.

FIGURE 37

Calculation of One-Step Substitutions of Benzene Derivatives

D_6	E	$2C_6$	$2C_3$	C_2	$3C_2'$	$3C_2''$
permutation representation	192	0	0	0	16	0

$= 20A_1 + \ldots$

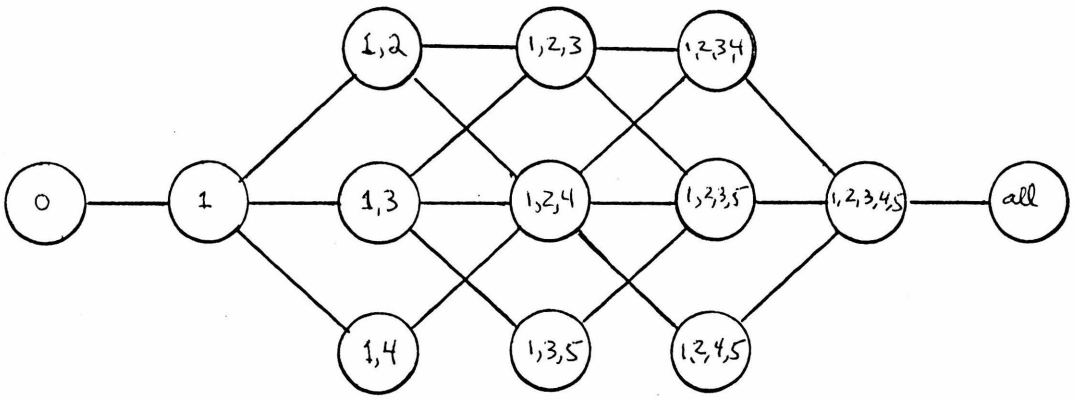
$$Z(D_6) = f_1^6 + 2f_6 + 2f_3^2 + f_2^3 + 3f_1^2f_2^2 + 3f_2^3$$

$$dZ(D_6)/df_1 = 6f_1 + 0 + 0 + 0 + 6f_1f_2^2 + 0$$

substitute $A^p + B^p$ for f_p,

distribution: $A^5 + 3A^4B + 6A^3B^2 + 6A^2B^3 + 3AB^4 + B^5$

1 3 6 6 3 1



B. Synthetic Design

One unsolved and largely unstructured chemical problem is that of designing a synthesis of a given target molecule. The usual strategy is to "work backwards" from this target structure to generate a synthesis tree. Likely intermediates can then be selected. Convergent syntheses, in which the lengths of consecutive chains of reactions are minimized, are generally preferable to non-convergent syntheses. Beyond these considerations strategies generally apply to individual cases.

Some of the simpler aspects of the general design problem have an underlying algebraic structure generated by the bonds to be made. For each bond a change of bond status operation is defined which acts on the possible intermediate structures in the synthesis. This operation forms a bond in a structure if it is unformed and breaks it if it is formed. Repeating the operation returns the structure to its original bonded status. This construction gives a group of order 2^n where n is the number of bonds considered. The group is isomorphic to the direct product of n cyclic groups of order 2. The set of all synthetic intermediates form a regular representation of the group by the action of the group on the set[59]. The construction of this group is well known in algebraic topology. The chemical structure defines a one-dimensional simplical complex (actually a graph). The group described is called the (one-dimensional) mod 2 chain group of the complex[60].

This group can be applied in at least three ways to the design problem. First, the action of this group on the target structure generates all possible intermediates. These would of course be needed for any complete analysis of the design problem for a given target molecule. The other applications involve the use of subgroups. Corey and Petersson[61] have considered the problem of the number of rings (cycles) and their disconnection in a polycyclic molecule. There are organic structures sufficiently complicated to make such a problem non-trivial. They found that a group could be used to generate all the cycles in a given target structure. This group is a subgroup of the mod 2 chain group and is called the mod 2 cycle group of the complex[60]. Evans[62] has considered the problem of the separation of functional pairs of functional groups (substituents) on the target structure. Briefly, he classifies the possible functional groups on the basis of their ability to impart nucleophilicity or electrophilicity on the adjacent sites. He then shows that nontrivially different situations arise depending on whether the functional groups are separated by an odd or even number of carbon atoms. All the fragments of a chemical structure with designated ends will generate a subgroup of the chain group. This group, in turn, can be used to find all these fragments which need to be considered if Evans' considerations are to be used.

Another group can be described by considering the carbon atoms and their absolute stereochemistry. Each appropriately substituted carbon atom can exist in two possible absolute configurations. An operation is defined which "inverts" the configuration of each center. All of these operations generate a group of order 2^n in which n is

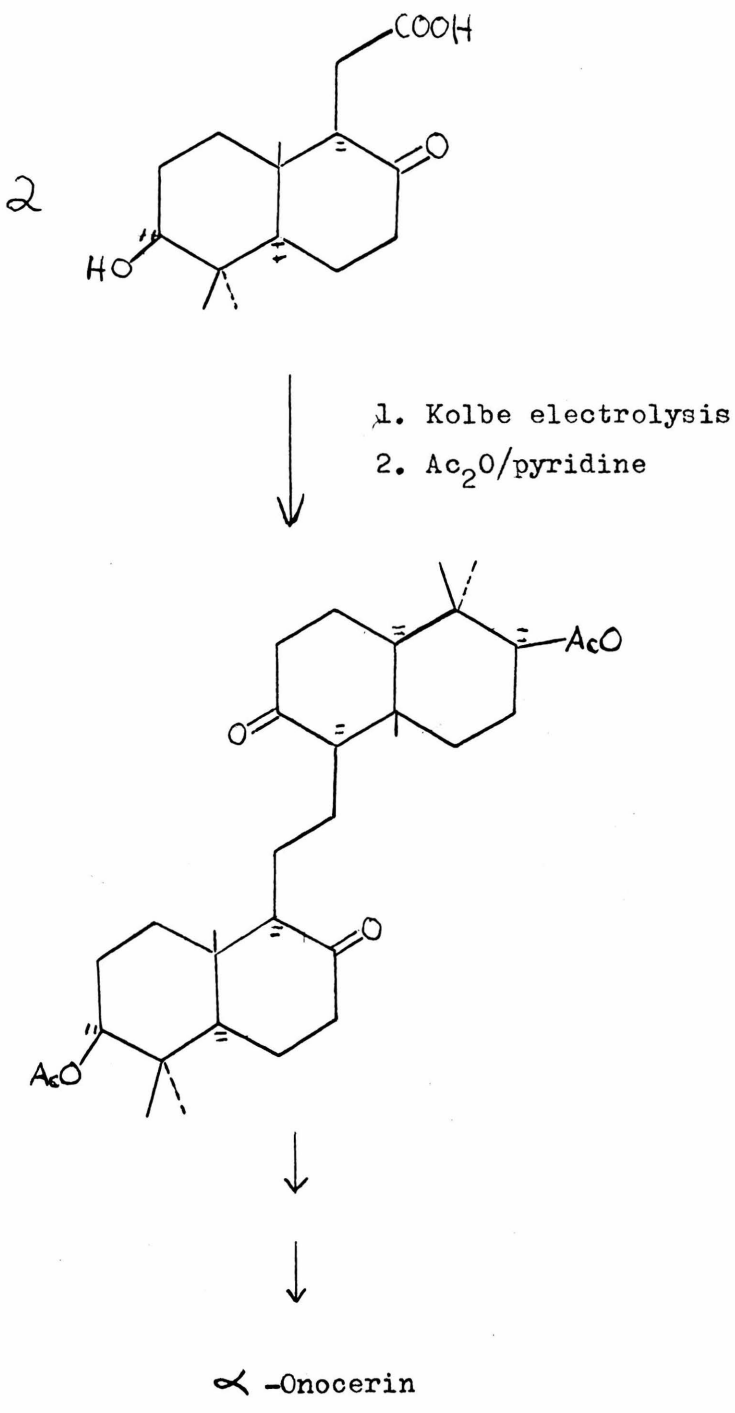
the number of asymmetric centers. The set of all stereoisomers form a regular representation of this group and action of the group on one stereoisomer generates all the others. This group, hereafter referred to as the stereochemistry group, is a subgroup of the (zero-dimensional) mod two chain group of the simplical complex defined by the target molecule.

Another strategy of synthetic design is to take advantage of symmetry whenever any is present. Clearly, whenever a molecule can be broken into two identical pieces, it would be advantageous to design a synthesis which requires their combination in a late step. The identical pieces can be made together in the earlier steps. An example is the synthesis of α-onocerin[63] (Figure 38). The action of the symmetry group (in most cases the point group of the target molecule) on the groups mentioned above can be described. For example, consider the stereochemistry group. Each symmetry operation determines an automorphism of the stereochemistry group by its permutation of and action on the asymmetric centers. The symmetry group and stereochemistry group can then be combined as a semidirect product[64]. This enlarged group contains all the symmetry operations, stereochemical inversions and their combinations. So far this enlarged group has no particular significance.

For a given design problem it is possible to define symmetry operations other than the usual point group operations which turn out to be useful. As an example, the abstract structure shown in Figure 39 will be considered. As drawn, this structure has a symmetry point group that includes only two operations, the identity and reflection

FIGURE 38

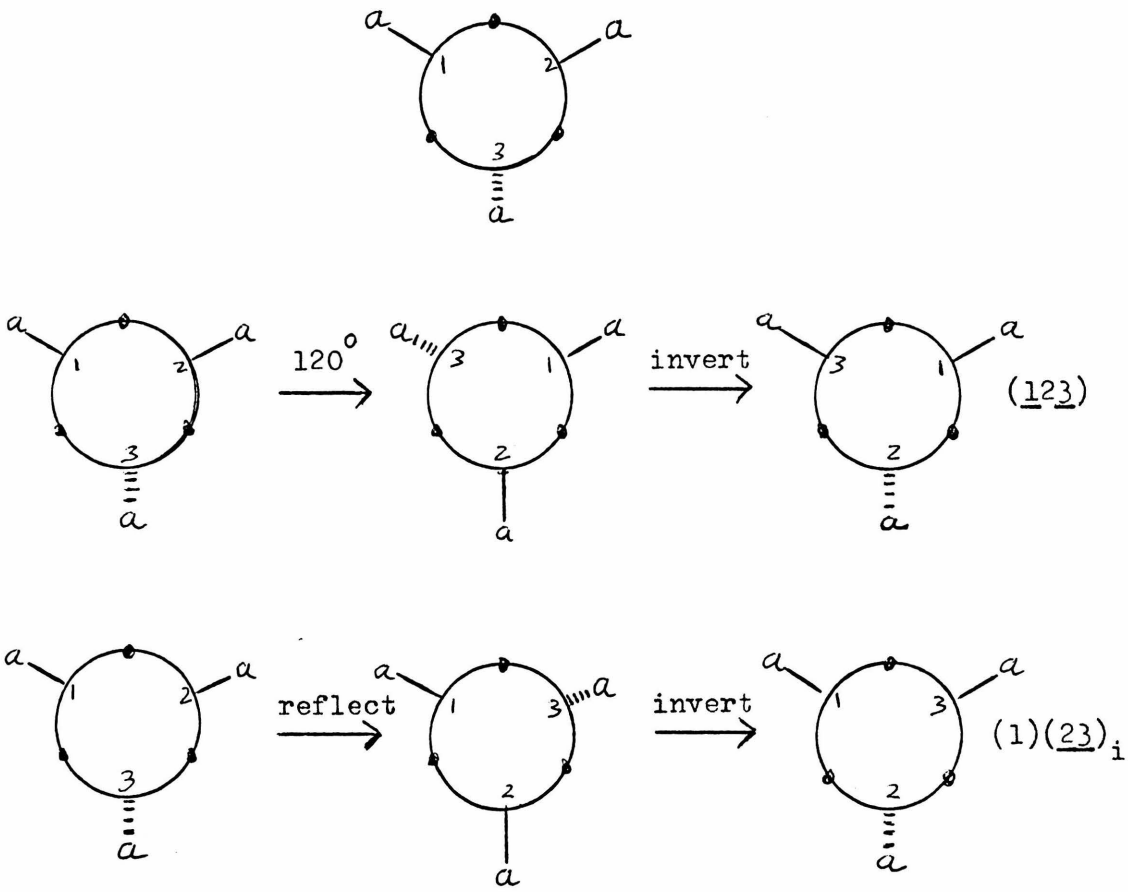
Synthesis of ∝-Onocerin



∝ -Onocerin

FIGURE 39

Construction of the Enlarged Symmetry Group



rotations	reflections
(1)(2)(3)	$(\underline{1})(\underline{2})(\underline{3})_i$
$(\underline{123})$	$(\underline{123})_i$
$(\underline{132})$	$(\underline{132})_i$
$(\underline{12})(\underline{3})$	$(12)(3)_i$
$(\underline{1})(23)$	$(1)(\underline{23})_i$
$(13)(\underline{2})$	$(\underline{13})(2)_i$

in the vertical plane between atoms 1 and 2 and including atoms 3. However, a number of other operations can be defined. If the structure is rotated 120^{o} clockwise and atoms 2 and 3 inverted, the structure is left unchanged. This sequence of operations is shown in Figure 40 and is symbolized as (123). This symbol is read: move the atom in site 1 to site 2 and leave it in the same configuration; move the atom in site 2 to site 3 and invert its configuration; move the atom in site 3 to site 1 and invert its configuration. Another type of sequence of operations will leave the structure unchanged. The structure is reflected in the plane going through atom 1 and between atoms 2 and 3. Following this, the configurations of the atoms in sites 2 and 3 are inverted. This sequence is also shown in Figure 39 and is symbolized $(1)(23)_i$ where the i refers to the fact that the original operation was an improper rotation. There turns out to be twelve such operations which leave this structure unchanged, which are listed in Figure 39. These twelve operations form a group which is isomorphic to the symmetry point group D_{3h}. The "reason" for the existence of this larger group is that all three centers on the structure are constitutionally identical and differ only in configuration.

The precise mathematical description of this group can be briefly stated. The semidirect product of the stereochemical inversion group $C_2xC_2xC_2$ and the permutation group D_3 is formed as determined by the action of D_3 on $C_2xC_2xC_2$. This is also the wreath product $D_3[C_2]$ [65]. and has order 48. The group derived above is a subgroup of this group. Alternatively, the semidirect product of C_2xC_2 by D_{3h} is formed where

C_2xC_2 is isomorphic to the factor group of $C_2xC_2xC_2$ by the operation which inverts all three centers at once. This can be symbolized as:

$$0 \rightarrow D_2 \rightarrow D_2x_sD_{3h} \underset{F}{\overset{\rightarrow}{\leftarrow}} D_{3h} \rightarrow 0$$

where the arrows are group homomorphisms and the zeroes represent the trivial group of one element. This diagram represents a splitting extension which means that the product group has at least one subgroup isomorphic to D_{3h} satisfying certain criteria. The group derived above is one of these subgroups.

Other strategies useful in design problems can be defined as symmetries of the target structure. For example, there are numerous known interconversions of functional groups so that in the synthesis of a functionalized molecule it is usually possible to take care of this functionality after the carbon skeleton is formed. Stated differently, a symmetry is defined which permutes functionalized parts of a chemical structure based on known interconversions. As an example, a ketone and an alcohol can usually be interconverted and are often thought of as just an oxygenated carbon. An arbitrary designation for functionality is often used in the crude design of the carbon skeleton synthesis[67]. In this group derived above, the atom permutations were based on a familiar point group. Since synthesis presupposes that parts of the target molecule will be separate in the earlier stages, they need not have any obvious relationship to each other in the final structure. Hence any permutation of constitutionally identical atoms in the target structure could represent a useful symmetry. The main point here is that the symmetry group of a target structure that is of

potential use to the synthetic chemist can be considerably larger than the usual point group. In fact an upper limit on the size of this group can be given as

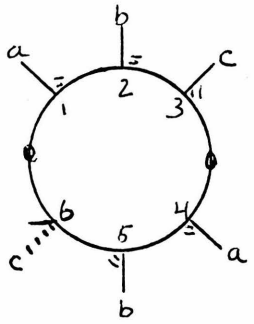
$$\prod_i (n_i!)(2^{n_i})$$

where n_i is the number of atoms of constitution i. The formula gives the order of the corresponding wreath product.

Existence of symmetries of the type discussed above does not determine how or even whether they are used in the design of the synthesis. Generally this will depend on the individual case. Two hypothetical examples indicate some strategies that may be used. First, consider the abstracted structure shown in Figure 40. The heavy dots are junction points which include the part of the molecule not being considered for symmetry. This structure is left invariant by the four indicated symmetry operations. A synthesis which would take advantage of this symmetry would proceed as follows. The pure enantiomer (1) is stereoselectively synthesized with appropriate functionality for creation of the third chiral center and formation of the junction. The third chiral center is formed non-stereoselectively. This will naturally give two diastereomers, each of which has the proper configuration for one-half of the desired target structure. These diastereomers are separated and the junction is formed. The inherent advantage of this route is apparent. A molecule with six chiral centers is synthesized requiring only one stereoselective reaction, a resolution of enantiomers and a separation of diastereomers. Normally such a synthesis

FIGURE 40

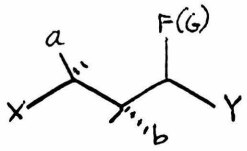
Illustration of the Use of Symmetry in a Hypothetical Synthesis

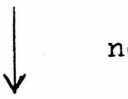


left invariant by:

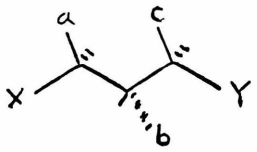
$(1)(2)(3)(4)(5)(6)$
$(14)(25)(\underline{36})$
$(1)(2)(3)(4)(5)(6)_i$
$(\underline{14})(\underline{25})(36)_i$

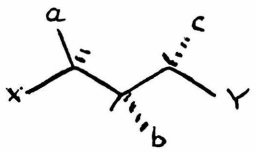
Projected synthesis:

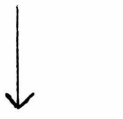




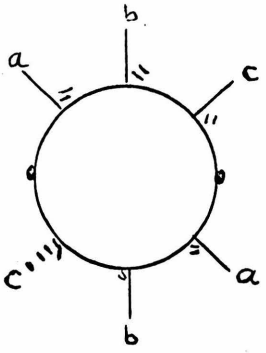
non-steroselective







1. separate diastereomers
2. form junction



would require five stereoselective reactions, a resolution of enantiomers and separations of diastereomers depending on the extent of stereoselectivity of the previous reactions.

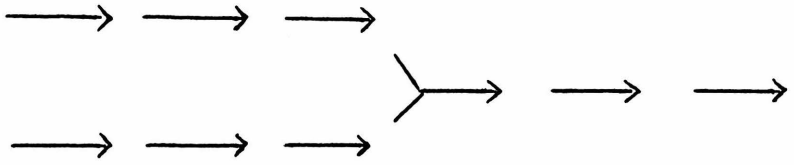
The synthesis proposed here can be thought of as intermediate between a completely symmetric synthesis such as that reported for α-onocerin[63] (Figure 38) and the usual convergent synthesis. The comparison is pictured in Figure 41.

As a second example consider the target structure trans-decalin shown in Figure 42. This structure is left invariant by the four operations indicated. Advantage could be taken of this symmetry if the synthesis could be projected through a cis-decalin intermediate as shown. The chiral centers could be created symmetrically in this intermediate. The cis-decalin would then be converted into the desired trans-decalin. Two racemic cis-decalin intermediates are possible, but both give the same desired racemic trans-decalin product. One of the cis-decalin intermediates might be preferable, however if a mixture were obtained, both would give the desired product.

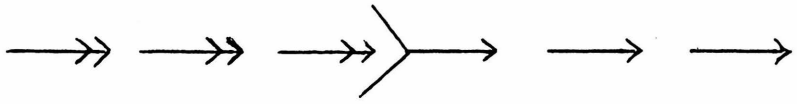
Other examples have to be treated individually, but it is likely that some advantage can be taken of symmetry in most cases.

FIGURE 41

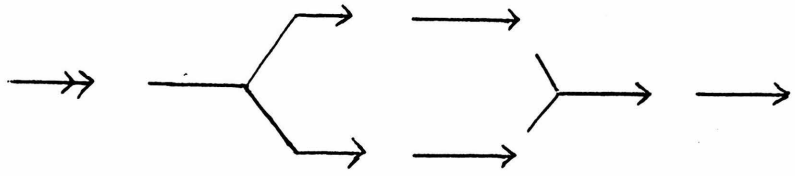
Comparison of types of Syntheses



Usual Convergent Synthesis



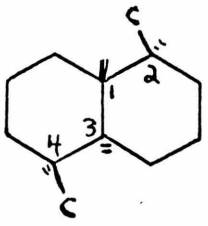
Completely Symmetric Synthesis



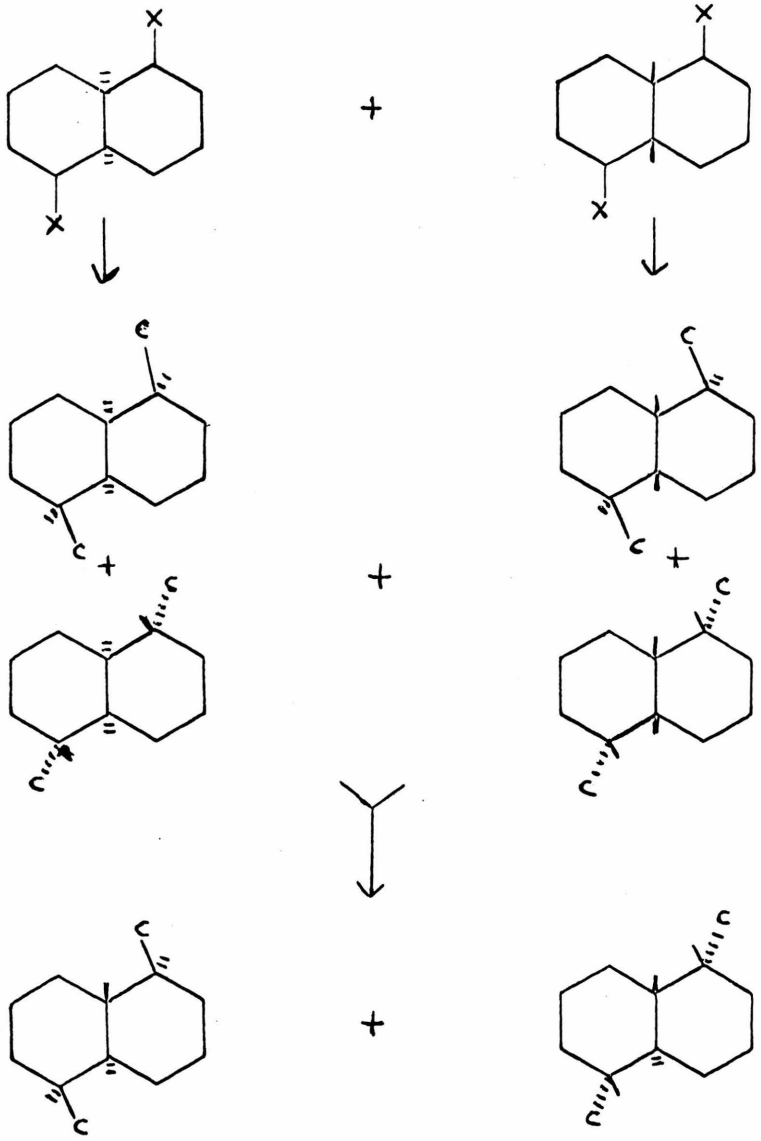
Intermediate Case

FIGURE 42

Example of the Use of Symmetry in a Hypothetical Synthesis



invariant to: (1)(2)(3)(4)
(13)(24)
$(1)(2)(3)(4)_i$
$(13)(24)_i$



C. Pseudochirality

The concept of pseudoasymmetry is reasonably well known as a special case of stereoisomerism[68]. It is well known that carbon atoms with four different substituents are chiral and that the enantiomeric forms can be interconverted by reflection in any mirror plane. However, if two of the substituents are themselves enantiomeric chiral ligands, two isomeric forms result which are meso and cannot be interconverted by reflection in a mirror plane. Such a pair of stereoisomers have been termed a pseudoasymmetric pair[69]. A symbolic and a real example are given in Figure 43. These are diastereomers and differ in physical properties such as melting point, etc.

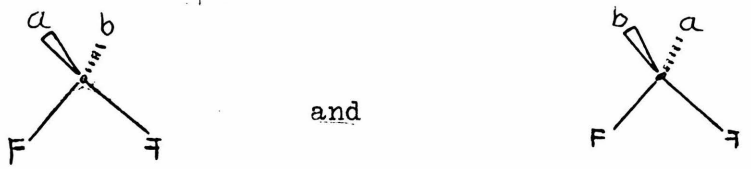
An important distinction must be made between pseudoasymmetric structures typically termed "meso" and structures such as (2) shown at the bottom of Figure 43. For this structure there exists only one meso form. Such a structure will be designated meso from now on while the pseudoasymmetric structures will be termed pseudoasymmetric or pseudochiral. This is a departure from standard terminology which designates all these structures as meso. The intrinsic difference between the pseudochiral and meso situations will be established later in this section.

Prelog[70] has synthesized compounds which are pseudoasymmetric based on axes and planes of chirality, rather than on centers of chirality like the example in Figure 43. An axial pseudoasymmetric pair is shown in Figure 44. The reported[70] melting points of these diastereomers are given underneath the structures.

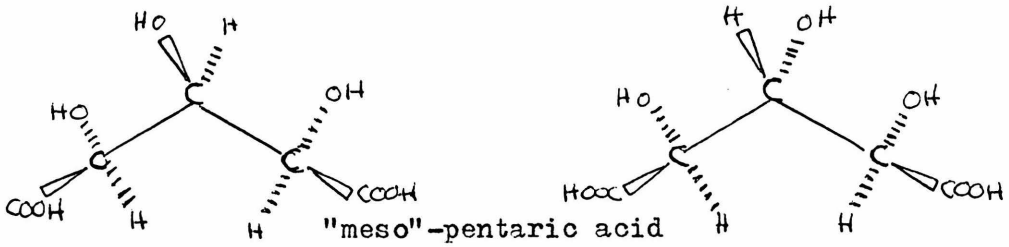
FIGURE 43

Illustration of Pseudoasymmetric and Meso Structures

Pseudoasymmetric Structures

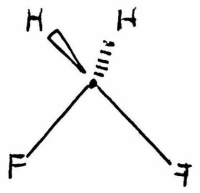


F is a chiral ligand, Ꟊ is its enantiomer



"meso"-pentaric acid

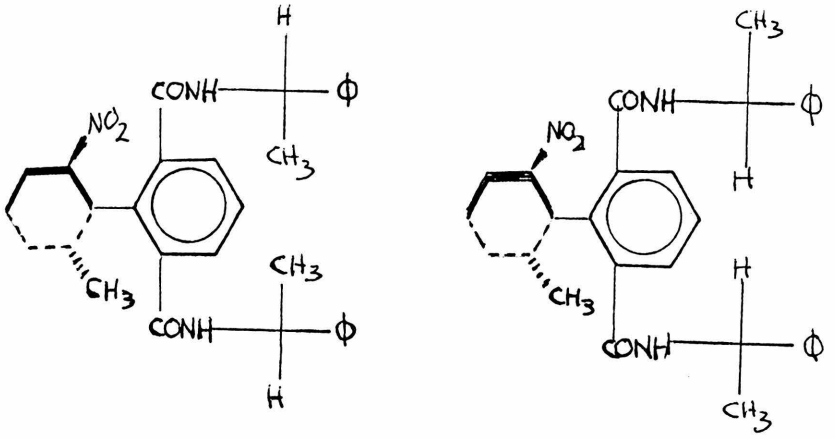
Meso Structure



(2)

FIGURE 44

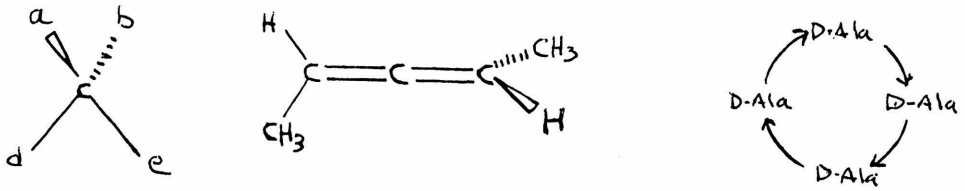
An Axial Pseudoasymmetric Pair



151° 181°

FIGURE 45

Chiral Structures with rotation symmetry



C_1 C_2 C_4

The purpose here is to give an algebraic description of this phenomenon analogous to that for regular chirality. A chemical structure is chiral if it lacks an alternating axis of symmetry. It may, however, have pure rotation symmetry. In Figure 45 are shown three chiral structures (one enantiomer only) with their rotation symmetry designated underneath.

To get at the desired algebraic description an apparent digression must be made. In Figure 46 are shown a number of configurations of triangles and the effect of symmetry operations on them. Configuration (1) is invariant to a C_2 rotation as shown. Configuration (2) is invariant to reflection in a vertical plane. These are the two types of symmetry operations traditionally encountered. Two more will now be defined. Configuration (3) is not invariant to a C_2 rotation, since the triangles are colored differently. However, this configuration is invariant to the sequence of operations shown. First, a C_2 rotation is done followed by a color change operation which colors the white triangle black and the black triangle white. This combined operation will be designated $\underline{C}_2$. Similarly, the sequence of reflection and color change leaves configuration (4) unchanged. This operation will be designated $\underline{S}_v$. Thus a total of <u>four</u> symmetry operations are defined. This situation has been well studied and is termed <u>dichromatic antisymmetry</u>[71].

Now a similar sequence will be done on the structures shown in Figure 47. Structure (1) is invariant to a C_2 rotation and structure (2) is invariant to a S_v reflection. However, structure (1) is also

FIGURE 46

Illustration of Dichromatic Symmetry Operations

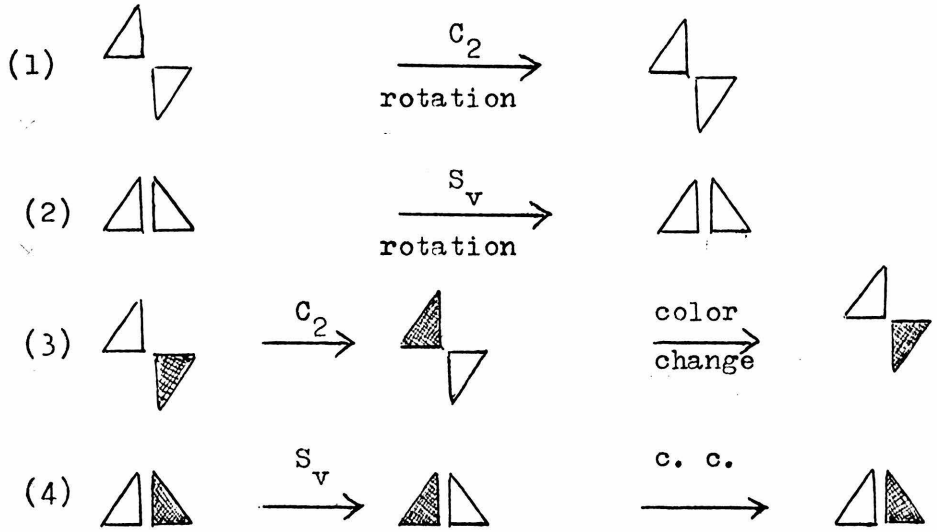
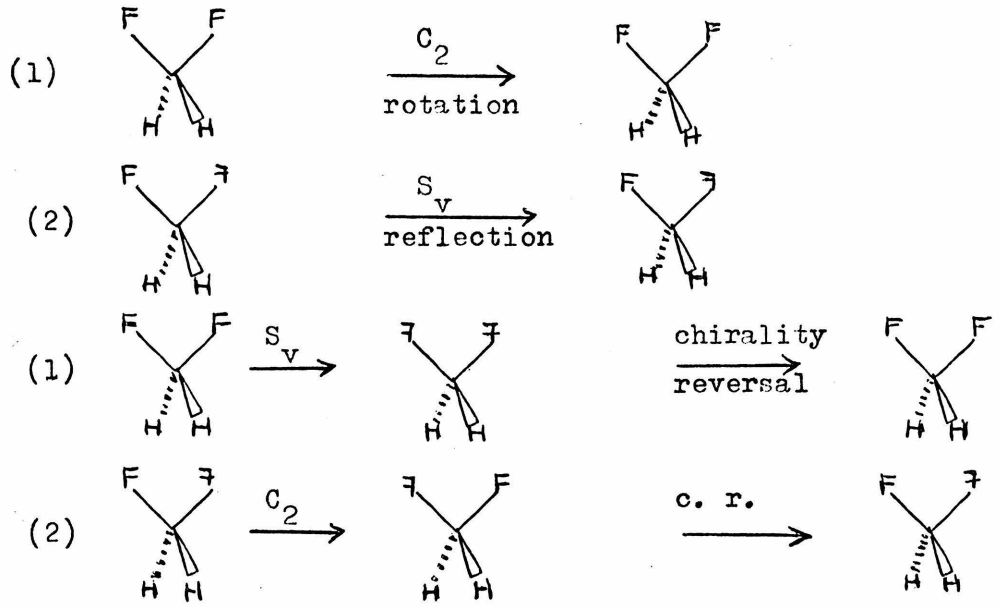


FIGURE 47

Illustration of Pseudochirality Operations



invariant to the sequence of operations shown. First (1) is reflected in the plane S_v which reverses the configuration of the chiral ligands. The operator, chirality reversal, is performed which changes the ligands back to their original configuration. This sequence of operations will be symbolized S_{vp}. Similarly the sequence of rotation and chirality reversal leaves (2) unchanged. This combination will be symbolized C_{2p}. These two combination operations will be termed pseudochirality operations.

Some further clarification of the effect of these pseudochirality operations may be helpful. A C_{np} operation in its simplest form reverses the chirality of all chiral ligands in a structure, while it is simply a C_n rotation for the rest of the structure. Therefore a structure with a C_{np} element of symmetry is invariant to the C_n rotation except that all (+)-chiral ligands are taken to (-)-chiral ligands of the same constitution. A C_{np} axis (for a set of chiral ligands) cannot pass through any of the ligands in the set and any structure with a C_{np} element of symmetry has an equal number of (+)- and (-)-chiral ligands. An S_{np} rotation-reflection operation leaves all (+)-chiral ligands as (+)-chiral ligands, while it is simply an S_n rotation-reflection for the rest of the structure. A structure with an S_{np} element of symmetry is invariant to the S_n operation except that all (+)-chiral ligands are taken into (+)-chiral ligands of the same constitution. An S_{np} axis can pass through chiral ligands and there are no restrictions on the relative number of (+)- and (-)-ligands.

Comparison of the color change and chirality change operations shows one significant difference. Reflection does not change color, while it does change chirality. This means that the reflection operation in the color change case corresponds to the reflection plus chirality change operations in the chirality change case.

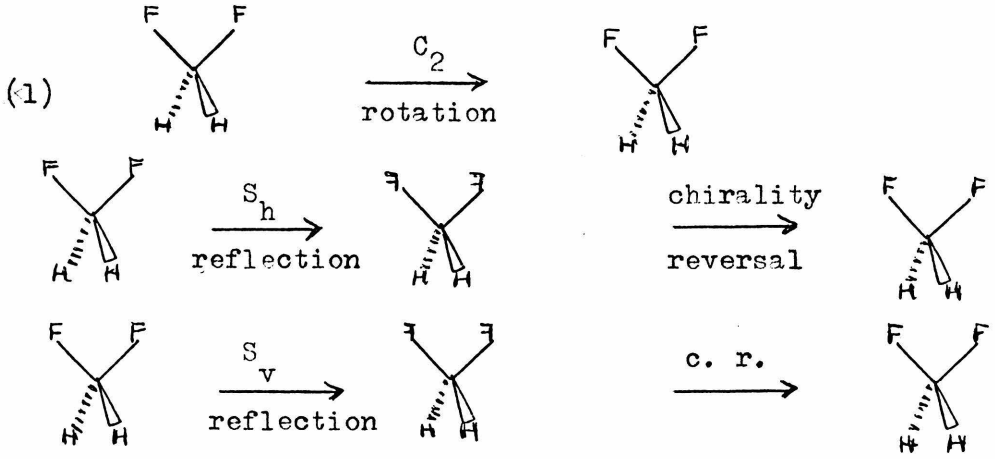
It is well known that all the symmetry operations of a structure form a group, hence the group properties of these pseudochirality operators will be investigated.

In Figure 48 are shown three structures and the symmetry operations to which they are invariant. It is a simple matter to verify that these operations form the groups indicated (hereafter termed pseudochirality groups). For structures (1) and (2) the group is isomorphic to the symmetry point group C_{2v}. In fact, these structures would have C_{2v} symmetry if the ligands were not chiral. Structure (3) would have C_s symmetry without chiral ligands. The problem now is to determine how these groups can be derived and what relation they have to the problem at hand.

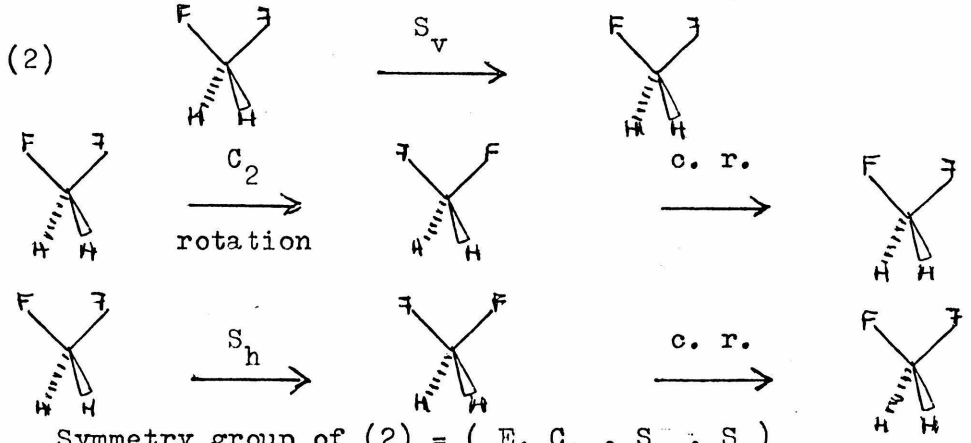
The mathematical problem here can be concisely stated. The group that includes <u>all</u> of the symmetry and pseudochirality operators is the direct product $S \times C$. S is the symmetry point group and C is the chirality change group. In the present example C is isomorphic to the cyclic group of order two, C_2. In the general case C will be the direct product of cyclic groups of order 2. The direct product $C \times S$ has the property shown diagrammatically in Figure 49. The zeroes refer to the trivial group of one element. The homomorphisms i_c and i_s

FIGURE 48

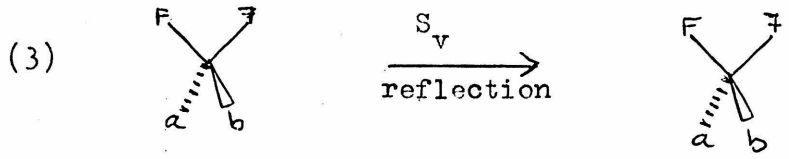
Illustration of Symmetry Operations and Groups



Symmetry group of (1) = (E, C_2, S_{hp}, S_{vp})



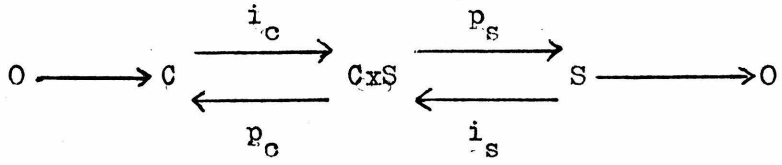
Symmetry group of (2) = (E, C_{2p}, S_{hp}, S_v)



Symmetry group of (3) = (E, S_v)

FIGURE 49

Direct Product Diagram



are not unique. Composing the homomorphisms shown in Figure 49 gives the indicated results. Injecting C into C x S (i_c) followed by the projection p_c gives the same result as the identity mapping of C onto itself. Similary, i_s followed by p_s gives 1_s. The significance of all this to the problem at hand can now be stated. The possible pseudochirality groups correspond to the possible injections i_s. The desired group is the image in C x S of i_s. Now the composition of i_s with p_c gives a projection of S onto C. It has already been noted that C is isomorphic to C_2, hence each pseudochirality group determines a homomorphism of S onto C_2. This property suggests an easy way of deriving these groups using chemically familiar group-theoretic methods. Each homomorphism of a group onto C_2 determines a one-dimensional irreducible representation of the group. Thus the desired pseudochirality groups correspond to the one-dimensional, real irreducible representations of the symmetry point group.

The derivation of the pseudochirality groups obtainable from C_{2h} is sketched in Figure 50 along with examples. The desired groups are obtained by "imposing" the one-dimensional irreducible representation on the point group. For example, consider the B_g representation. Imposition maps E and i to +1 and C_2 and S_h to -1. This corresponds to the pseudochirality group (E,C_{2p},i,S_{hp}). This group is given the unique designation $C_{2h}:B_g$. The structure with this symmetry is meso (one isomeric form). Similarly the B_u representation yields a pseudochirality group and a meso structure. Now imposition of the A_g representation maps all the group elements to +1. However, it must be

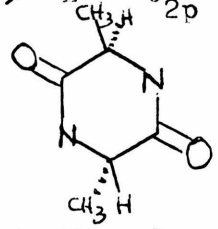
FIGURE 50

Derivation of Pseudochirality groups obtained from C_{2h}

C_{2h} character table

C_{2h}	E	C_2	i	S_h
B_g	1	-1	1	-1
B_u	1	-1	-1	1
A_g	1	1	1	1
A_u	1	1	-1	-1

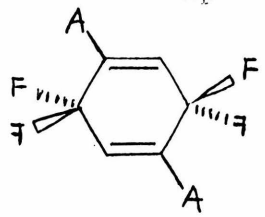
B_g 1 -1 1 -1 ⟶ E C_{2p} i S_{hp} C_{2h}: B_g



≡

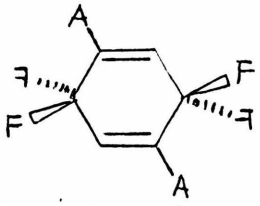
meso

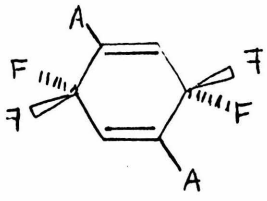
B_u 1 -1 1 -1 ⟶ E C_{2p} i_p S_h C_{2h}: B_u



meso

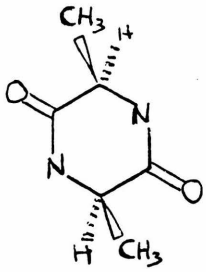
A_g 1 1 1 1 ⟶ E C_2 i S_h C_{2h}: A_g

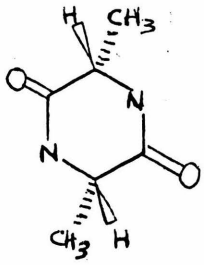




pseudochiral

A_u 1 1 -1 -1 ⟶ E C_2 i_p S_{hp} C_{2h}: A_u





chiral A

remembered that chiral ligands are present. This group includes no pseudochirality operators and describes the symmetry of a pseudochiral structure with C_{2h} point group symmetry. The two pseudoenantiomeric forms of such a structure are shown. The reason for using the term pseudochiral rather than pseudoasymmetric is now apparent. This example satisfies the requirement for being "pseudoasymmetric" yet has non-trivial rotation symmetry, hence the term pseudochirality is preferred. This parallels the reason for using the term chirality rather than asymmetry to describe traditional enantiomerisms.

The final representation is A_u and imposition yields the group (E,C_2,i_p,S_{np}). This group includes no alternating axis of symmetry, hence any structure with this symmetry must be chiral and exist in two enantiomeric forms, as does the example.

At this point it is possible to describe the types of isomerism possible. The type of isomerism of a structure is determined by the kinds of symmetry element a structure has. This information is contained in Table 27. Examples of the first three types are included among those in Figure 50. Examples of the latter two are given in Figure 51. These structures can be thought of as octahedrally coordinated metals with two tridentate ligands, themselves containing chiral ligands. In all cases the isomeric forms are interconverted by operations which are not included among the symmetry operations. This gets complicated in the latter case and the interconversions of the four isomers are indicated.

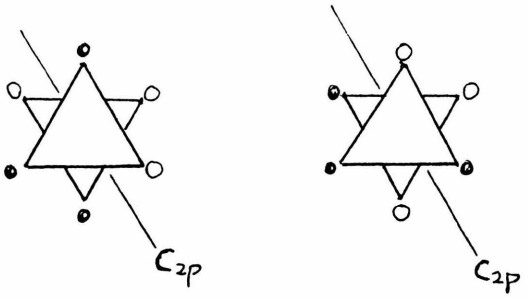
Table 27

Types of Isomerism

NAME	C_n	C_{np}	S_n	S_{np}	No. of Isomers
meso	yes	yes	yes	yes	1
pseudochiral	yes	no	yes	no	2
chiral A	yes	no	no	yes	2
chiral B	yes	yes	no	no	2
chiral C	yes	no	no	no	4

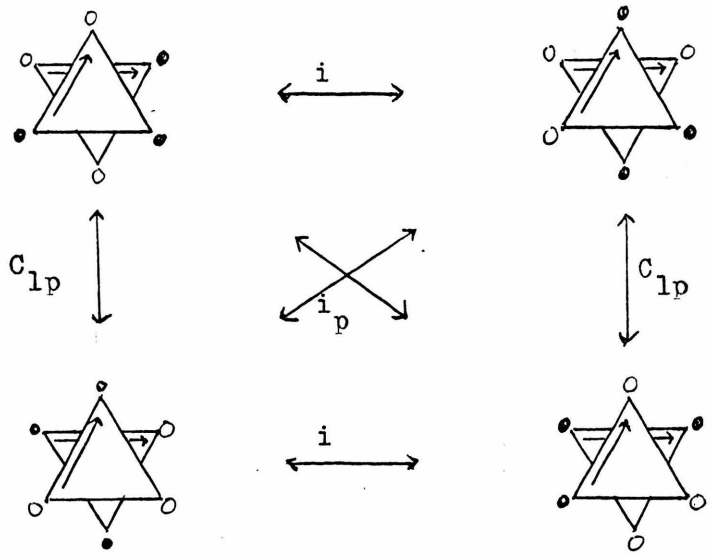
FIGURE 51

Illustration of remaining chiral forms



chiral B, C_2: B

chiral C, C_1: A



A generalization is possible by allowing more than one set of chiral ligands. Such sets can be constitutionally different or be situated in sites that are never interchanged. As might be expected, the number of possibilities increases significantly. However, the method of determining the pseudochirality groups and types of isomerism is easily derived from the sample case. In the case of n sets of chiral ligands the group C in Figure 49 is the direct product of n cyclic groups of order two. The desired pseudochirality groups are therefore in correspondence with all the possible sums of n one-dimensional irreducible representations of the symmetry point group. The possible double pseudochirality groups for C_{2h} are derived in Figure 52 in an abbreviated form, except for the first one. The two representations A_g (for the set of F ligands) and B_g (for the G set) are imposed on C_{2h}. The resulting group $C_{2h}: A_g+B_g$ is shown and an example given. The operator C_{2pf} is a C_2 rotation followed by a chirality reversal operation on the F ligands only.

Using these methods the possible single and double pseudochirality groups have been derived and are listed in Table 28. Only one representative group from each class of symmetry point groups is used. For example, C_{2v} is representative of all C_{nv}, n even. The information given for each group is its name, type (m = meso, p = pseudochiral, c = chiral), number of isomers, and the location or name of any examples. There is a simple relationship:

$$ne = 2^{s+1}$$

FIGURE 52

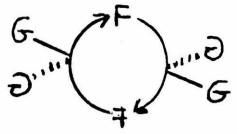
Derivation of Double Pseudochirality Groups from C_{2h}

C_{2h}	E	C_2	i	S_h
A_g	1	1	1	1
B_g	1	-1	1	-1
	E	C_{2pF}	i	S_{hpF}

C_{2h}: $A_g + B_g$ pseudochiral, 2 isomers

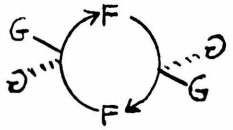
set G

set F



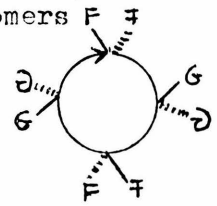
C_{2h}: $A_g + A_u$, E C_2 i_{pF} S_{hpF}

chiral, 4 isomers



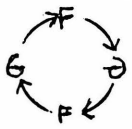
C_{2h}: $A_g + B_u$, E C_{2pF} i_{pF} S_h

pseudochiral, 2 isomers



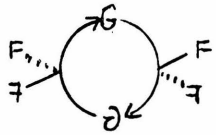
C_{2h}: $B_g + A_u$, E C_{2pG} i_{pF} S_{hpFG}

chiral, 2isomers



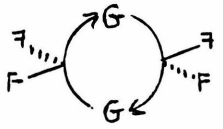
C_{2h}: $B_g + B_u$, E C_{2pFG} i_{pF} S_{hpG}

chiral, 2 isomers



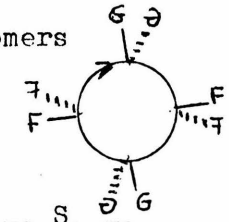
C_{2h}: $A_u + B_u$, E C_{2pF} i_{pFG} S_{hpG}

chiral, 2 isomers



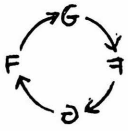
C_{2h}: $2A_g$, E C_2 i S_h

pseudochiral, 4 isomers



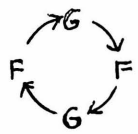
C_{2h}: $2B_g$, E C_{2pFG} i S_{hpFG}

pseudochiral, 2 isomers



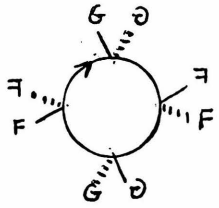
C_{2h}: $2A_u$, E C_2 i_{pFG} S_{hpFG}

chiral, 4 isomers



C_{2h}: $2B_u$, E C_{2pFG} i_{pFG} S_h

pseudochiral, 2 isomers



where n is the number of isomers, e is the number of types of symmetry elements, and s is the number of sets of chiral ligands. Some of the infinite groups yield no derived groups or groups that cannot be realized in any chemical structure and are included only for completeness. Since chiral ligands are being considered as two-valued points, the group $K_h:A_n$ represents a single asymmetric carbon atom.

A number of tables of single and double antisymmetry groups exist (also called magnetic symmetry groups)[72]. The present derivation is necessary because of some important differences between the two problems, which change the number and kind of groups obtained. It has already been noted that the operations are labelled differently. Furthermore, double antisymmetry groups of the type $C_{2h}: A_n+B_n$ and $C_{2h}: B_u+A_u$ are considered to be different. In the present case such a distinction leads to constitutional isomerism. Antisymmetry groups resulting from the totally symmetric representation are trivial. In the present case they yield pseudochiral structure, the major form of isomerism here. Finally, degeneracies which occur in the antisymmetry groups are removed in the present case. For example, the groups $C_{2v}: B_1$ and $C_{2v}: B_2$ (Figure 53) are considered as a degenerate pair, since the planes of symmetry differ only by labels. However, these two groups can be represented by different chemically realizable examples as shown and therefore must be considered different. The assignment of each example to its pseudochirality group is arbitrary, however.

FIGURE 53

Illustration of Non-degenerate Pseudochirality Groups

C_{2v}	E	C_2	$S_{(xz)}$	$S_{(yz)}$
B_1	1	-1	1	-1
B_2	1	-1	-1	1

C_{2v}: B_1 E C_{2p} S S_p

C_{2v}: B_2 E C_{2p} S_p S

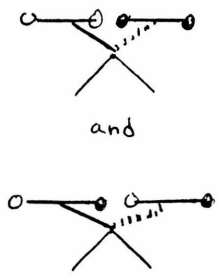


FIGURE 54

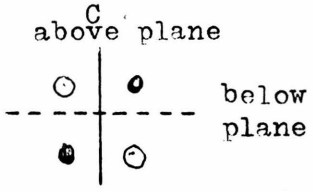
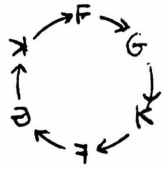


FIGURE 55



C_{6h}: $B_g + E_{1g}$

FIGURE 56

Other Group Extensions

$0 \longrightarrow C \longrightarrow Cx_sS \underset{\longleftarrow}{\overset{\longrightarrow}{}} S \longrightarrow 0$ semidirect product

$0 \longrightarrow C \longrightarrow CextS \longrightarrow S \longrightarrow 0$ general extension

Groups with three or more sets of chiral ligands would be derived similarly. There would be about 400 such groups with three sets of chiral ligands. Such chemical structures are easily realizable, particularly as cyclic peptides. It is interesting that there can be no meso structure with three or more sets of chiral ligands. This can be easily proved by noting that meso structures occur only for D_{nh} groups with two sets of chiral ligands. A meso structure can result only if the homomorphism of the symmetry group onto the chirality change group is onto with respect to both rotation and reflection symmetry. Since D_n has only two generators, it cannot be mapped (epimorphically) onto $C_2xC_2xC_2$. Hence no meso structure is possible.

An overview of this work is hopefully ascertainable from the following discussion. Consider the three structures with chiral ligands drawn in Figure 54. The symmetry of these structures can be specified in several ways. The first and simplest is to specify just the rotation symmetry. A has D_4 rotation symmetry, B has C_2 rotation symmetry, and C has D_2 rotation symmetry. Knowing only rotation symmetry is sufficient to describe some properties of chemical structures. For example, the number of isomers is determined using just the rotation symmetry as discussed in an earlier section. The π-molecular orbitals of benzene can be determined from just the rotation group C_6 [73]. Specification of only rotation symmetry is, however, insufficient to determine other properties such as chirality. This leads to the next higher level of symmetry specification, namely point group symmetry. A has D_4 symmetry, B has C_{2h} symmetry, and C has D_{2d} symmetry. From

these specifications it is apparent that A exists in two enantiomeric forms and B and C are superimposible on their mirror images. The point group description is a great deal more useful than the rotation group description and most applications of group theory to chemistry require it. The point group designation is, however, inadequate to ascertain pseudochirality.

The present work extends this sequence by one. Of the examples given, A is D_{4h}: A_{1u}, B is D_{2h}: B_{1g}, and C is D_{2d}: A_1 . This indicates that B is meso and C is pseudochiral and exists in two pseudoenantiomeric forms. The information gained (pseudochirality) by specifying the pseudochiral symmetry and point group symmetry instead of just the point group symmetry is analogous to the information gained (chirality) by specifying point group symmetry instead of just rotation symmetry.

This sequence can be extended and the work generalized in any number of ways. What is needed is a designation of the group C in Figure 49. For example, the members of this group could include permutations of constitutionally different ligands. The structure shown in Figure 55 yields a representation for C_{6h}: $B_g + E_{1g}$. The development here is somewhat similar to the treatment of polychromatic symmetry[74]. Other generalizations are possible by modifying the combination of the groups. The diagram shown in Figure 49 is for the direct product which is a rather trivial extension of C by S [75]. More generalized extensions are represented by the diagrams in Figure 56. The semidirect product was already used in the section on

synthetic design. Also, some nonrigid symmetry groups can be expressed as semidirect products[76].

Table 28

Possible Pseudochirality Groups

Supergroup	Symmetry elements		
C_1	E		
1) One set of chiral ligands (single)			
derived group name	type	isomers	example
C_1:A	c	4	Figure 51
2) two sets of chiral ligands (double)			
C_1:2A	c	8	
C_s	E σ_h		
1) Single			
C_s:A'	p	2	All known (experimentally) pseudochiral structures
C_s:A"	c	2	
2) Double			
C_s:2A'	p	4	
C_s:2A"	c	4	
C_s:A'+A"	c	4	
C_i	E i		
1) Single			
C_i:A_g	p	2	
C_i:A_u	c	2	
2) Double			
C_i:2A_g	p	4	
C_i:2A_u	c	4	
C_i:A_g+A_u	c	4	

Table 28 (continued)

Supergroup	Symmetry elements				
C_2	E	C_2			
derived group name		type	isomers	example	
1) Single					
C_2:A		c	4		
C_2:B		c	2	Figure 51	
2) Double					
C_2:2A		c	8		
C_2:2B		c	4		
C_2:A+B		c	4		
D_4	E	$2C_4$	C_2	$2C_2'$	$2C_2''$
1) Single					
D_4:A_1		c	4		
D_4:A_2		c	2		
D_4:B_1		c	2		
D_4:B_2		c	2		
2) Double					
D_4:$2A_1$		c	8		
D_4:$2A_2$		c	4		
D_4:$2B_1$		c	4		
D_4:$2B_2$		c	4		
D_4:A_1+A_2		c	4		
D_4:A_1+B_1		c	4		
D_4:A_1+B_2		c	4		

Table 28 (continued)

Supergroup derived group name	Symmetry elements type	isomers	example
$D_4:A_2+B_1$	c	4	
$D_4:A_2+B_2$	c	2	
$D_4:B_1+B_2'$	c	2	

C_{2v}	E	C_2	$\sigma_{v(xz)}$	$\sigma'_{v(yz)}$	
1) Single					
$C_{2v}:A_1$			p	2	
$C_{2v}:A_2$			c	2	
$C_{2v}:B_1$			m	1	Figure 53
$C_{2v}:B_2$			m	1	Figure 53
2) Double					
$C_{2v}:2A_1$			p	4	
$C_{2v}:2A_2$			c	4	
$C_{2v}:2B_1$			p	2	
$C_{2v}:2B_2$			p	2	
$C_{2v}:A_1+A_2$			c	4	
$C_{2v}:A_1+B_1$			p	2	
$C_{2v}:A_1+B_2$			p	2	
$C_{2v}:A_2+B_1$			c	2	
$C_{2v}:A_2+B_2$			c	2	
$C_{2v}:B_1+B_2$			c	2	

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
C_{3v}	E $2C_3$ $3\sigma_v$		
1) Single			
$C_{3v}:A_1$	p	2	
$C_{3v}:A_2$	c	2	
2) Double			
$C_{3v}:2A_1$	p	4	
$C_{3v}:2A_2$	c	4	
$C_{3v}:A_1+A_2$	c	4	
C_{2h}	E C_2 i σ_h		
1) Single			
$C_{2h}:A_g$	p	2	Figure 50
$C_{2h}:B_g$	m	1	Figure 50
$C_{2h}:A_u$	c	2	Figure 50
$C_{2h}:B_u$	m	1	Figure 50
2) Double			
$C_{2h}:2A_g$	p	4	Figure 52
$C_{2h}:2B_g$	p	2	"
$C_{2h}:2A_u$	c	4	"
$C_{2h}:2B_u$	p	2	"
$C_{2h}:A_u+B_u$	c	2	"
$C_{2h}:A_g+B_g$	p	2	"
$C_{2h}:A_g+A_u$	c	4	"
$C_{2h}:A_g+B_u$	p	2	"

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
C_{2h}:B_g+A_u	c	2	Figure 52
C_{2h}:B_g+B_u	c	2	"

C_{3h}	E	C_3	C_3^2	h	S_3	S_3^5
1) Single						
C_{3h}:A'	p	2				
C_{3h}:A"	c	2				
2) Double						
C_{3h}:2A'	p	4				
C_{3h}:2A"	c	4				
A'+A"	c	4				

D_{3h}	E	$2C_3$	$3C_2$	σ	$2S_3$	$3\sigma_v$
1) Single						
D_{3h}:A_1'	p	2				
D_{3h}:A_2'	p	2				
D_{3h}:A_1''	c	2				
D_{3h}:A_2''	m	1				
2) Double						
D_{3h}:2A_1'	p	4				
D_{3h}:2A_2'	p	2				
D_{3h}:2A_1''	c	4				
D_{3h}:2A_2''	p	2				
D_{3h}:A_1'+A_2'	p	2				

Table 28 (Continued)

Supergroup	Symmetry elements		
derived group name	type	isomers	examples
$D_{3h}:A_1'+A_1''$	c	4	
$D_{3h}:A_1'+A_2''$	p	2	
$D_{3h}:A_2'+A_1''$	c	2	
$D_{3h}:A_2'+A_2''$	c	2	
$D_{3h}:A_1''+A_2''$	c	2	

D_{4h}	E	$2C_4$	C_2	$2C_2'$	$2C_2''$	i	$2S_4$	σ_h	$2\sigma_v$	$2\sigma_d$
1) Single										
$D_{4h}:A_{1g}$			p		2					
$D_{4h}:A_{2g}$			m		1					
$D_{4h}:B_{1g}$			m		1					
$D_{4h}:B_{2g}$			m		1					
$D_{4h}:A_{1u}$			c		2					
$D_{4h}:A_{2u}$			m		1					
$D_{4h}:B_{1u}$			m		1					
$D_{4h}:B_{2u}$			m		1					
2) Double										
$D_{4h}:2A_{1g}$			p		4					
$D_{4h}:2A_{2g}$			p		2					
$D_{4h}:2B_{1g}$			p		2					
$D_{4h}:2B_{2g}$			p		2					
$D_{4h}:2A_{1u}$			c		4					
$D_{4h}:2A2u$			p		2					
$D_{4h}:2B_{1u}$			p		2					

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
$D_{4h}:2B_{2u}$	p	2	
$D_{4h}:A_{1g}+A_{2g}$	p	2	
$D_{4h}:A_{1g}+B_{1g}$	p	2	
$D_{4h}:A_{1g}+B_{2g}$	p	2	
$D_{4h}:A_{1g}+A_{1u}$	c	4	
$D_{4h}:A_{1g}+A_{2u}$	p	2	
$D_{4h}:A_{1g}+B_{1u}$	p	2	
$D_{4h}:A_{1g}+B_{2u}$	p	2	
$D_{4h}:A_{2g}+B_{1g}$	m	1	
$D_{4h}:A_{2g}+B_{2g}$	m	1	
$D_{4h}:A_{2g}+A_{1u}$	c	2	
$D_{4h}:A_{2g}+A_{2u}$	c	2	
$D_{4h}:A_{2g}+B_{1u}$	m	1	
$D_{4h}:A_{2g}+B_{2u}$	m	1	
$D_{4h}:B_{1g}+B_{2g}$	m	1	
$D_{4h}:B_{1g}+A_{1u}$	c	2	
$D_{4h}:B_{1g}+A_{2u}$	m	1	
$D_{4h}:B_{1g}+B_{1u}$	c	2	
$D_{4h}:B_{2g}+A_{1u}$	c	2	
$D_{4h}:B_{1g}+B_{2u}$	m	1	
$D_{4h}:B_{2g}+A_{2u}$	m	1	
$D_{4h}:B_{2g}+B_{1u}$	m	1	

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
$D_{4h}:B_{2g}+B_{2u}$	c	2	
$D_{4h}:A_{1u}+A_{2u}$	c	2	
$D_{4h}:A_{1u}+B_{1u}$	c	2	
$D_{4h}:A_{1u}+B_{2u}$	c	2	
$D_{4h}:A_{2u}+B_{1u}$	m	1	
$D_{4h}:A_{2u}+B_{2u}$	m	1	
$D_{4h}:B_{1u}+B_{2u}{}^{q}$	m	1	
D_{2d}	E $2S_4$ C_2 $2C_2'$ $2\sigma_d$		
1) Single			
$D_{2d}:A_1$	p	2	
$D_{2d}:A_2$	m	1	
$D_{2d}:B_1$	c	2	
$D_{2d}:B_2$	m	1	
2) Double			
$D_{2d}:2A_1$	p	4	
$D_{2d}:2A_2$	p	2	
$D_{2d}:2B_1$	c	4	
$D_{2d}:2B_2$	p	2	
$D_{2d}:A_1+A_2$	p	2	
$D_{2d}:A_1+B_1$	c	4	
$D_{2d}:A_1+B_2$	p	2	
$D_{2d}:A_2+B_1$	c	2	

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
$D_{2d}:A_2+B_2$	c	2	
$D_{2d}:B_1+B_2$	c	2	
D_{3d}	**E $2C_3$ $3C_2$ i $2S_6$ $3\sigma_d$**		
1) Single			
$D_{3d}:A_{1g}$	p	2	
$D_{3d}:A_{2g}$	m	1	
$D_{3d}:A_{1u}$	c	2	
$D_{3d}:A_{2u}$	m	1	
2) Double			
$D_{3d}:2A_{1g}$	p	4	
$D_{3d}:2A_{2g}$	p	2	
$D_{3d}:2A_{1u}$	c	4	
$D_{3d}:2A_{2u}$	p	2	
$D_{3d}:A_{1g}+A_{2g}$	p	2	
$D_{3d}:A_{1g}+A_{1u}$	c	4	
$D_{3d}:A_{1g}+A_{2u}$	p	2	
$D_{3d}:A_{2g}+A_{1u}$	c	2	
$D_{3d}:A_{2g}+A_{2u}$	c	2	
$D_{3d}:A_{1u}+A_{2u}$	c	2	
S_4	**E S_4 C_2 S_4^3**		
1) Single			
$S_4:A$	p	2	
$S_4:B$	c	2	

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
2) Double			
S_4:2A	p	4	
S_4:2B	c	4	
S_4:A+B	c	4	

T_d	E	$8C_3$	$3C_2$	$6S_4$	$6\sigma_d$			
1) Single								
T_d:A_1	p		2					
T_d:A_2	c		2					
2) Double								
T_d:$2A_1$	p		4					
T_d:$2A_2$	c		4					
T_d:A_1+A_2	c		4					

T_h	E	rC_3	$4C_3^2$	$3C_2$	i	$4S_6$	$4S_6^2$	$3\sigma_d$
1) Single								
T_h:A_g	p		2					
T_h:A_u	c		2					
2) Double								
T_h:$2A_g$	p		4					
T_h:$2A_u$	c		4					
T_h:A_g+A_u	c		4					

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples

O_h	E	$8C_3$	$6C_2$	$6C_4$	$3C_2$	i	$6S_4$	$8S_6$	$3\sigma_h$	$6\sigma_d$
1) Single										
$O_h:A_{1g}$			p	2						
$O_h:A_{2g}$			m	1						
$O_h:A_{1u}$			o	2						
$O_h:A_{2u}$			m	1						
2) Double										
$O_h:2A_{1g}$			p	4						
$O_h:2A_{2g}$			p	2						
$O_h:2A_{1u}$			o	4						
$O_h:2A_{2u}$			p	2						
$O_h:A_{1g}+A_{2g}$			p	2						
$O_h:A_{1g}+A_{1u}$			c	4						
$O_h:A_{1g}+A_{2u}$			p	2						
$O_h:A_{2g}+A_{1u}$			c	2						
$O_h:A_{2g}+A_{2u}$			c	2						
$O_h:A_{1u}+A_{2u}$			c	2						

I_h	E	$12C_5$	$12C_5^2$	$20C_3$	$15C_2$	i	$12S_{10}$	$12S_{10}^3$	$20S_6$	15σ
1) Single										
$I_h:A_g$		p		2						
$I_h:A_u$		c		2						

Table 28 (Continued)

Supergroup derived group name	Symmetry elements					
	type	isomers	examples			
2) Double						
I_h:$2A_g$	p	4				
I_h:$2A_u$	c	4				
I_h:A_g+A_u	c	4				
T	E	$8C_3$	$3C_2$			
1) Single						
T:A	c	4				
2) Double						
T:2A	c	8				
O	E	$8C_3$	$6C_2$	$6C_4$	$3C_2$	
1) Single						
O:A_1	c	4				
O:A_2	c	2				
2) Double						
O:$2A_1$	c	8				
O:$2A_2$	c	4				
O:A_1+A_2	c	4				
I	E	$12C_5$	$12C_5^2$	$20C_3$	$15C_2$	
1) Single						
I:A	c	4				
2) Double						
I:2A	c	8				

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
$C_{\infty v}$	$E \quad 2C_\infty \cdots \infty\sigma_v$		
1) Single			
$C_{\infty v}:\Sigma^+$	This group is not physically possible because there cannot be $\infty\,\sigma$ planes in a structure containing discrete chiral units. This is another difference from dichromatic antisymmetry in which infinite symmetry planes are possible.		
$C_{\infty v}:\Sigma^-$	c	2	
2) Double			
$C_{\infty v}:2\Sigma^-$	c	4	
The other combinations are not possible.			
$D_{\infty h}$	$E \quad 2C_\infty \cdots \infty\sigma_v \quad i \quad 2S_\infty \cdots \infty C_2$		
1) Single			
$D_{\infty h}:\Sigma_g^+$	Not possible		
$D_{\infty h}:\Sigma_g^-$	m	1	meso-tartaric acid
$D_{\infty h}:\Sigma_u^+$	Not possible		
$D_{\infty h}:\Sigma_u^-$	c	2	d-tartaric acid
2) Double			
$D_{\infty h}:2\Sigma_g^-$	p	2	
$D_{\infty h}:2\Sigma_u^-$	c	4	
$D_{\infty h}:\Sigma_g+\Sigma_u^-$	c	2	

The other combinations are not possible.

Table 28 (Continued)

Supergroup derived group name	Symmetry elements type	isomers	examples
D_∞	E $2C_\infty$ ∞C_2		
1) Single			
$D_\infty:A_1$	c	4	
$D_\infty:A_2$	c	2	
2) Double			
$D_\infty:2A_1$	c	8	
$D_\infty:2A_2$	c	4	
$D_\infty:A_1+A_2$	c	4	
C_∞	E $2C_\infty$		
1) Single			
$C_\infty:A_1$	c	4	
2) Double			
$C_\infty:2A_1$	c	8	
C_{∞_h}	E $2C_\infty$ i $2S_\infty$		
1) Single			
$C_{\infty_h}:A_g$	p	2	
$C_{\infty_h}:A_u$	c	2	
2) Double			
$C_{\infty_h}:2A_g$	p	4	
$C_{\infty_h}:2A_u$	c	4	
$C_{\infty_h}:A_g+A_u$	c	4	

Table 28 (Continued)

Supergroup	Symmetry elements			
derived group name	type	isomers	examples	
K_h	E ∞C_∞ i ∞S_∞			
1) Single				
K_h:A_g	Not possible.			
K_h:A_u	c	2	Single asymmetric carbon	
2) Double				
K_h:$2A_u$	c	4		
K	E ∞C_∞	Yields no derived groups.		

Table 28 (Continued)

Summary of Derived Groups

	Single					Double						
	M	P2	C2	C4	Tot	M	P2	P4	C2	C4	C8	Tot
C_1	0	0	0	1	1	0	0	0	0	0	1	1
C_s	0	1	1	0	2	0	0	1	0	2	0	3
C_i	0	1	1	0	2	0	0	1	0	2	0	3
C_2	0	0	1	1	2	0	0	0	0	2	1	3
C_3	0	0	0	1	1	0	0	0	0	0	1	3
D_3	0	0	1	1	2	0	0	0	0	2	1	3
D_4	0	0	3	1	4	0	0	0	3	6	1	10
C_{2v}	2	1	1	0	4	0	4	1	3	2	0	10
C_{3v}	0	1	1	0	2	0	0	1	0	2	0	3
C_{2h}	2	1	1	0	4	0	4	1	3	2	0	10
C_{3h}	0	1	1	0	2	0	0	1	0	2	0	3
D_{3h}	2	1	1	0	3	0	3	1	2	]	0	10
D_{4h}	6	1	1	0	8	12	12	1	9	2	0	36
D_{2d}	2	1	1	0	4	0	4	1	3	2	0	10
D_{3d}	2	1	1	0	4	0	4	1	3	2	0	10
S_4	0	1	1	0	2	0	0	1	0	2	0	3
T_d	0	1	1	0	2	0	0	1	0	2	0	3
T_h	0	1	1	0	2	0	0	1	0	2	0	3
O_h	2	1	1	0	4	0	4	1	3	2	0	10
I_h	0	1	1	0	2	0	0	1	0	2	0	3
T	0	0	0	1	1	0	0	0	0	0	1	1

Table 28

Summary of Derived Groups (Continued)

	Single					Double						
	M	P2	C2	C4	Tot	M	P2	P4	C2	C4	C8	Tot
O	0	0	1	1	2	0	0	0	0	2	1	3
I	0	0	0	1	1	0	0	0	0	0	1	1
C_∞	0	0	1	0	1	0	0	0	0	1	0	1
D_{∞_h}	1	0	1	0	2	0	1	0	1	1	0	3
C_∞	0	0	0	1	1	0	0	0	0	0	1	1
D_∞	0	0	1	1	2	0	0	0	0	2	1	3
C_{∞_h}	0	1	1	0	2	0	0	1	0	2	0	3
K_h	0	0	1	0	1	0	0	0	0	1	0	1
K	0	0	0	0	0	0	0	0	0	0	0	0
Total	19	16	26	10	71	12	37	16	31	49	10	155

D. Through-Space Orbital Interactions

In a recent paper Goldstein and Hoffmann[77] have considered the problem of through-space orbital interactions and their stabilization or destabilization of chemical structures. The purpose here will be to give an algebraic structure to this problem and describe some consequences and extensions of the idea.

Goldstein and Hoffmann's purpose in their work was to establish the topological requirements for aromaticity. Their approach was to consider interactions based on simple second-order perturbation theory between conjugated polyene fragments (called ribbons) in various orientations or topologies. Some of these topologies for one, two or three ribbons are shown in Figure 57. The familiar simple Huckel molecular orbitals were used to determine the interactions. Ribbon interactions were assumed to result primarily from the interactions of the termini. Each orbital was classified as pseudo-p or pseudo-d based on the relationship of the termini to each other as shown in Figure 58. It was noted that this designation alternates with increasing orbital energy in an acyclic polyene, also shown in Figure 58. When two ribbons interact, it was assumed the stabilization was principally a result of the HOMO-LUMO (Highest Occupied Molecular Orbital - Lowest Unoccupied Molecular Orbital) interaction. They further noted that there were only four possible patterns of orbital occupancy for ground states of linear polymers with respect to the pseudo-symmetry designations of the HOMO and LUMO. These were designated by mode numbers calculated as (n-z) modulo 4 where n is the number of carbons in the

FIGURE 57

Ribbon Interaction Topologies

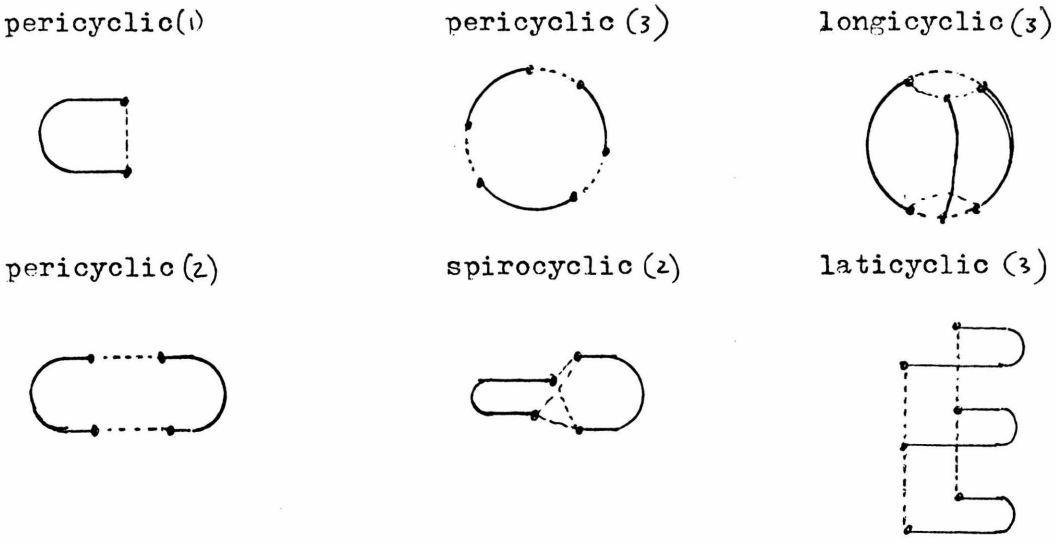
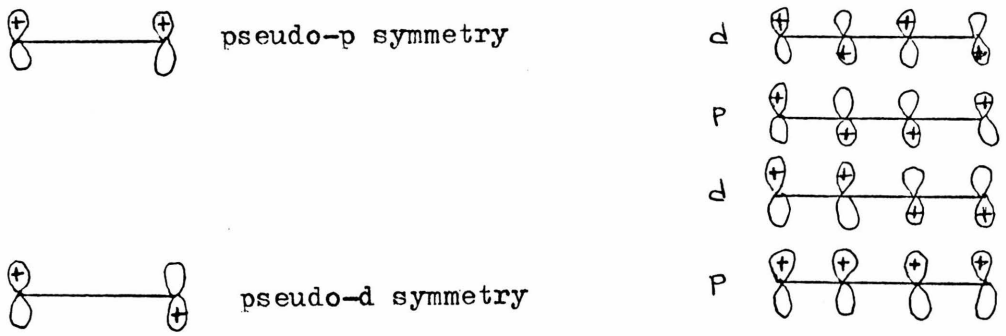


FIGURE 58



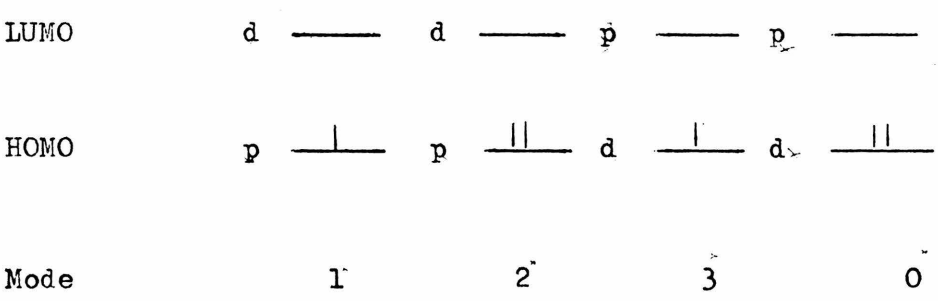
ribbon and z is the charge. These mode numbers had the following algebraic property. When ribbons were added by extending one ribbon by another, the mode of the resultant ribbon was equal to the sum of the modes of the two components (modulo 4). This information is given pictorally in Figure 59.

They further assumed that HOMO's and LUMO's be at the same absolute level in different ribbons (universal energies)[78]. They also restricted this work to ribbons of mode 0 and 2. As an example they considered the conditions for stabilization of two ribbins in a spirocyclic topology. The only orbital interaction which is non-zero is the d-d interaction. Hence, there will be net stabilization only when one ribbon has a LUMO with d pseudosymmetry and the other has a HOMO with d pseudosymmetry. This situation arises when one ribbon has mode 0 and the other has mode 2. Thus a spirocyclic is stabilized when there are 4n+2 electrons. Where both ribbons are mode 0 or mode 2 (4n electrons) there is no stabilization and in fact a net destabilization as noted in their paper. They then proceeded to other topologies and determined favorable and unfavorable mode combinations.

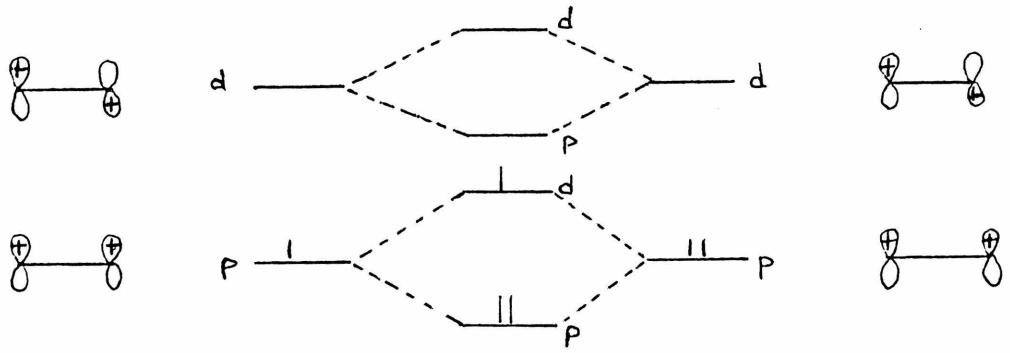
Before giving an algebraic description of this problem it is necessary to formulate a mode change operator. Considering only cases involving mode 0 and mode 2, such an operator will change a ribbon from mode 0 to mode 2 or vice-versa. This operation can be accomplished in any of several ways. For example, lengthening or shortening the ribbon by two carbon atoms will cause this change. Twisting the ribbon by 180° or adding or moving two electrons will also change the mode. These changes are shown in Figure 60.

FIGURE 59

Illustration of Ribbon Modes



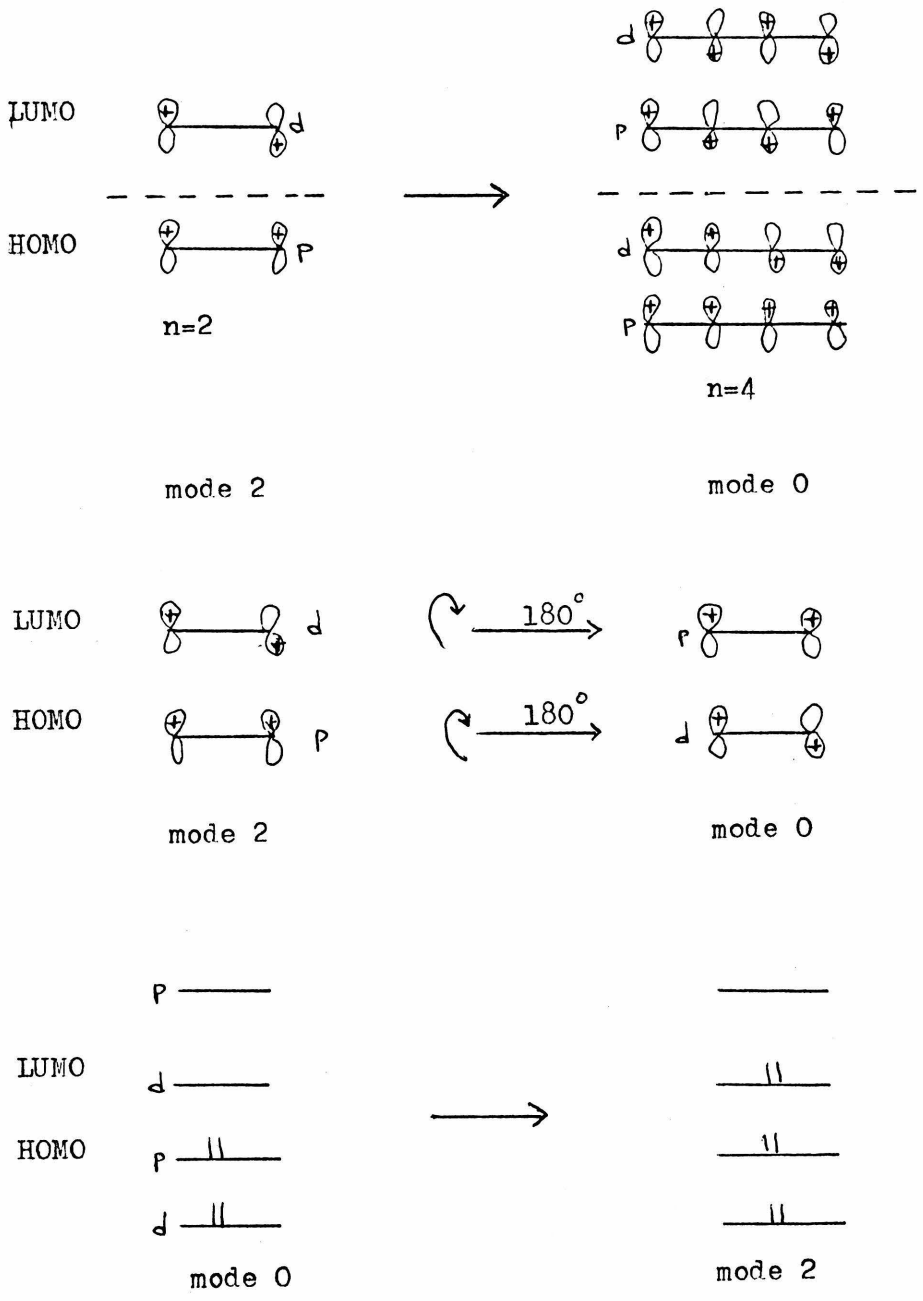
Sum of Mode 1 and Mode 2 by acyclic extension



1 + 2 = 3

FIGURE 60

Illustration of Mode Change operations



In the ensuing discussion it is most useful to think of this operator (symbolized M_c) as changing the length of a neutral ribbon (first example in Figure 60).

It is worthwhile to consider the (lack of) commutativity of this operator with the energy operators involved here. Clearly M_c does not commute with total energy operator for a single ribbon

$$M_c H \neq H M_c$$

Stated in more familiar terms, the total energy of a polyene is different from the total energy of the polyene with two more or two less carbon atoms. Similarly, any power of M_c (M_c^n) will not commute with H. Now consider the case of two interacting ribbons. The total energy operators here can be expressed as $H_1 + H_2 + P$ where H_1 and H_2 are the energy operators for the isolated ribbon and P is the perturbation operator which will be considered as the operator determining the (de)stabilization caused by the interaction. It has already been shown that M_c for each ribbon will not commute with H for that ribbon. Furthermore, M_c for one ribbon will not commute with P. This can be seen in the spirocyclic example presented earlier. The combination 0+2 was stabilized while the combination 0+0 and 2+2 (which are obtained from 0+2 by action of M_c) were destabilized. However, under the assumptions given earlier, M_c^2 <u>does</u> commute with P. Specifically, since only the interactions of ribbon termini are considered and a universal energy for the LUMO and HOMO is assumed, the effect of M_c^2 on one ribbon is to leave it unchanged.

Similarly M_c^n (n=even) will commute with P while M_c^{n+1} will not. It is therefore convenient to think of M_c as an operator of order two. (Actually, the homomorphic image of M_c on the cyclic group of order two is considered).

With the operator M_c as defined, it is possible to give a group-theoretic description of this problem. For each ribbon in an n-interacting ribbon topology the group M_c (isomorphic to the cyclic group of order 2) is assigned. Hence for the set of n ribbons, the direct product group $(M_c)^n$ (isomorphic to $(C_2)^n$) is assigned. This group is then mapped to an interaction change group I_c. These operators change a two-orbital interaction from stabilizing to destabilizing or vice-versa. Further consideration of the spirocyclic example should make these concepts clearer. In this case there are two interacting ribbons so that the mode change group is isomorphic to D_2. The interaction change group consists of operators which change all possible binary orbital interactions. Such a group would be quite large; however, it has already been pointed out that only certain binary orbital interactions are significant, and that they can be classified on the basis of the pseudosymmetry of the interacting orbitals. In the case of the spirocyclic topology, only the d-d interaction is non-zero, hence I_c is isomorphic to C_2. (More precisely, the image of the mapping is isomorphic to C_2). From the previous discussion it is apparent that changing the mode of one ribbon will change the interaction. However, changing the mode of both ribbons leaves the interaction unchanged. This homomorphism is depicted in Figure 61. The operators are named by the ribbon or interaction they change.

FIGURE 61

Illustration of Homomorphism for Spirocyclic Topology

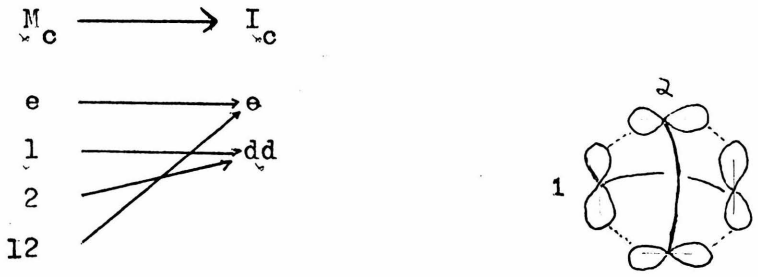
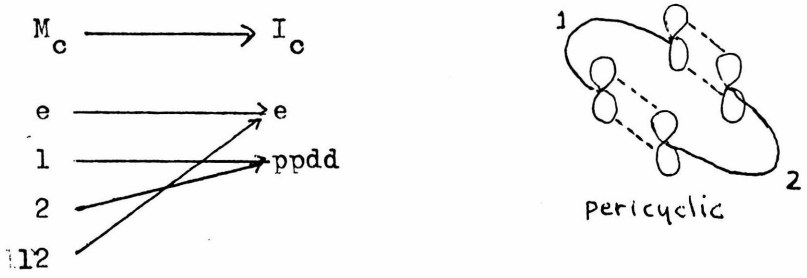
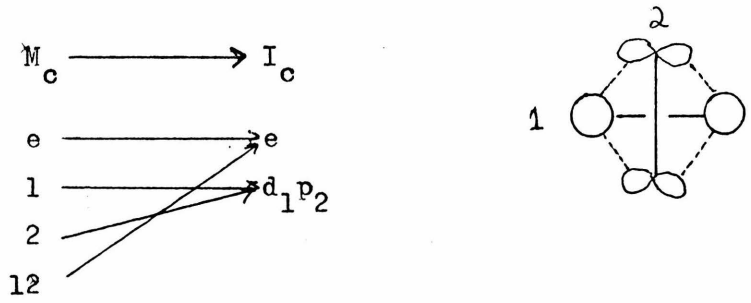


FIGURE 62

Illustration of Homomorphism for other two-ribbon Topologies





Overall, this process assigns to each topology of n interacting ribbons a homomorphism $M_c \rightarrow I_c$. Other interaction topologies of two ribbons yield different homomorphisms as shown in Figure 62. Only the image of the mappings is shown. The actual homomorphism considered is into the large group I_c discussed earlier.

In the pericyclic case both the p-p and d-d interactions change when changing the mode of one ribbon. In the unnamed case which is the same as the transition state for the allowed 2+2 cyclo addition, the d_1-p_2 interaction is the only non-zero interaction and changes as indicated in Figure 62.

So far nothing has been said about which mode combination is the favored one in the various topologies. This preference is determined by the dominant interaction in the image of the homomorphism. For example, in the spirocyclic and pericyclic examples, the dominant interactions are between orbitals of the same pseudosymmetry which means that the 0+2 and 2+0 combinations are favored (4n+2 electrons). In the third example the dominant (and only) interaction is between orbitals of different pseudosymmetry which means the 0+0 and 2+2 combinations are favored (4n electrons).

Cases with three or more interacting ribbons can be considerably more complicated. Only when the n-ribbon interactions can be broken up into binary interactions will a group homomorphism of the type described above result. The three ribbon topologies considered in detail by Goldstein and Hoffmann satisfied this requirement, although their formulation of the orbital interactions in these cases used ternary interactions. They remarked that annelating ribbons caused additional

pericyclic-like interactions and that therefore stabilization could only come from 0+2 combinations. The resulting group homomorphism for the longicyclic topology is shown in Figure 63. Changing the mode of one ribbon causes a change in the p-p and d-d interactions with adjacent ribbons. Goldstein and Hoffmann also noted that any formulations of this example must be invariant to any permutation of the ribbons. This property can be made precise in this group-theoretic formulation. The homomorphism must be invariant to the automorphism of the groups induced by the ribbon permutations. Stated differently, the square diagram shown at the bottom of Figure 63 must commute. The horizontal arrows refer to the induced automorphisms. The vertical arrows are the same homomorphism. Commutativity means that going around the square either way from upper left to lower right must yield the same composed homomorphism.

The laticyclic topology yields a homomorphism in the same way as shown in Figure 64. Clearly the image of the homomorphism in I_c is different in the laticyclic and longicyclic cases. Although it was not mentioned in their paper, the formulation in this case must be invariant to permutation of the outer ribbons.

Although it is aside from the main point here, it is interesting to look at the algebraic structure resulting from a topology not satisfying the simplifying assumptions mentioned at the beginning of the preceding paragraph. Such a case was briefly mentioned by Goldstein and Hoffmann[79]. For the topology shown in Figure 65 they gave a table of the doubly stabilized, singly stabilized and non-stabilized combinations, also shown in Figure 65. It is apparent

FIGURE 63

Illustration of Homomorphism for Longicyclic Topology

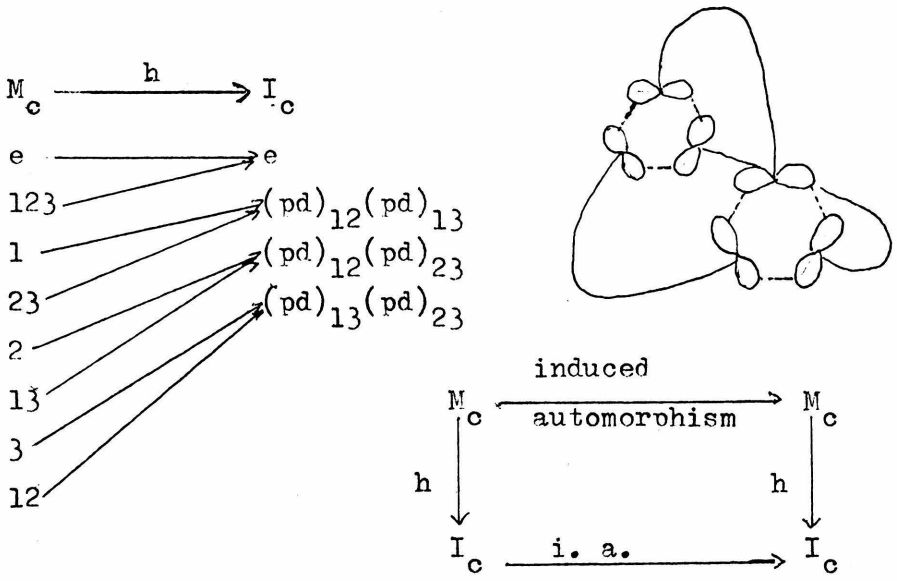


FIGURE 64

Illustration of Homomorphism for Laticyclic Topology

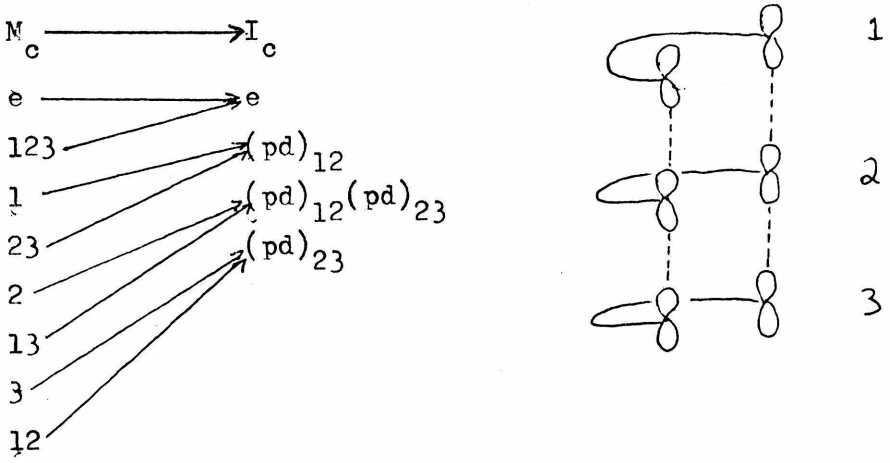
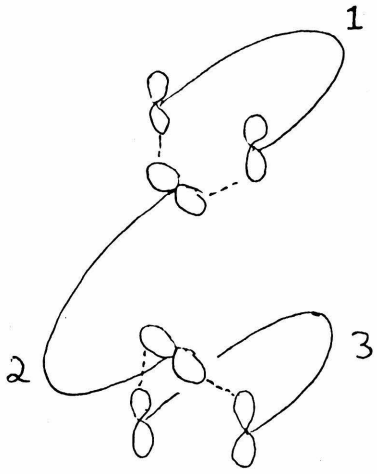
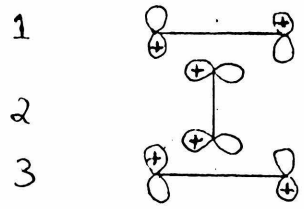


FIGURE 65

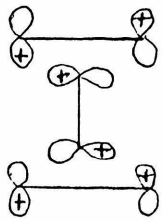
Interactions for a Three-Ribbon Topology

	1	2	3
Doubly stabilized	2	0	0
	0	0	2
Singly stabilized	0	0	0
	0	2	0
	2	0	2
	2	2	2
Not stabilized	0	2	2
	2	2	0

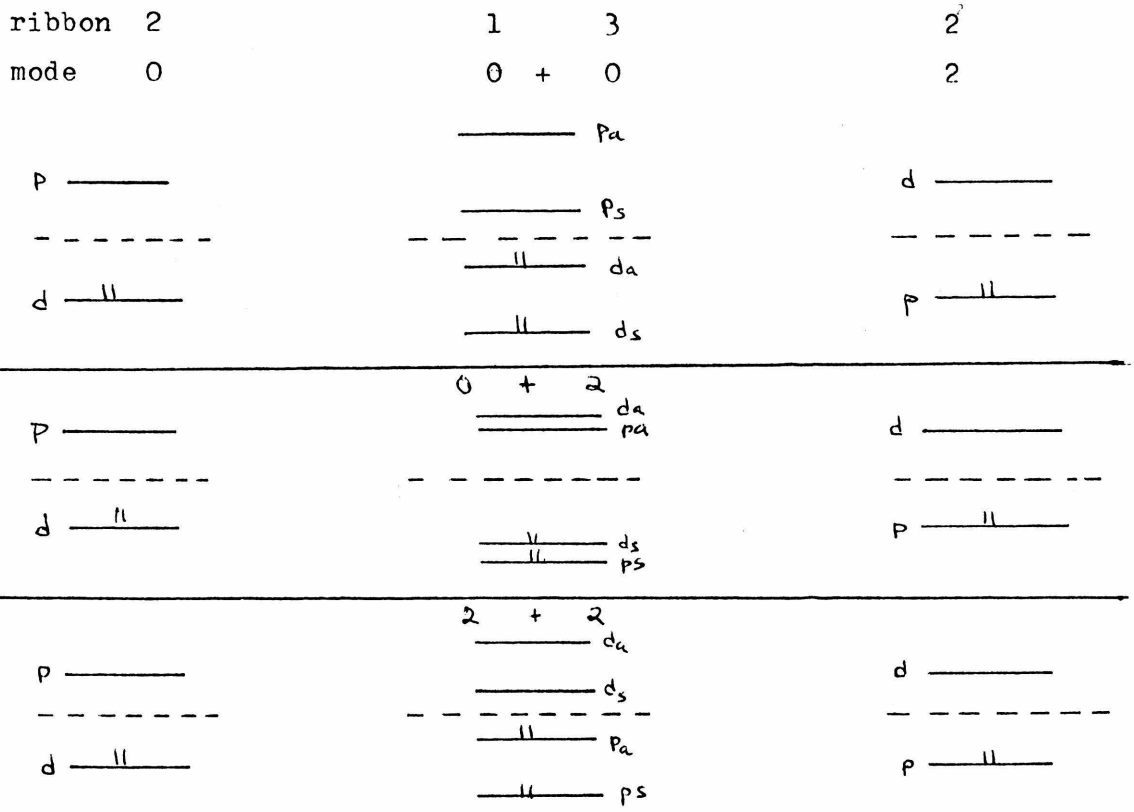




p_2–ds_{13} interaction



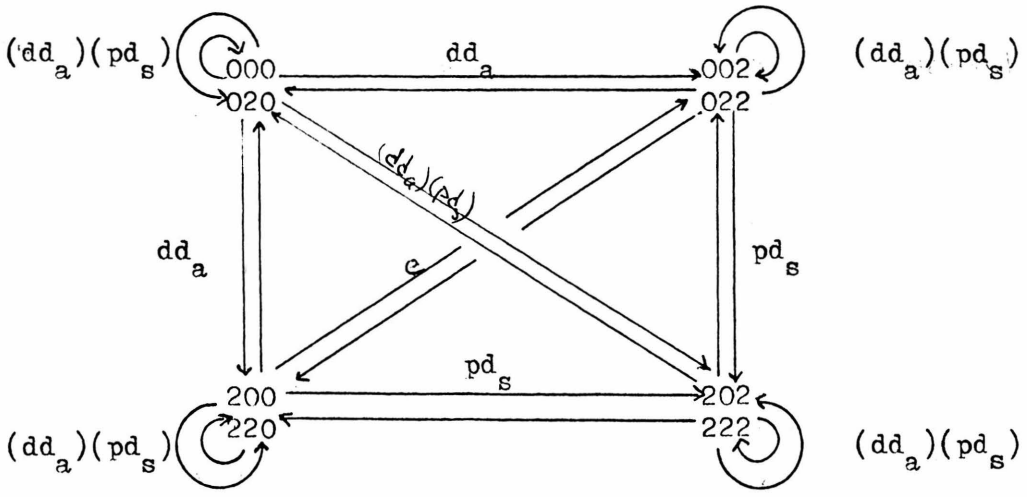
d_2–da_{13} interaction



that the rule concerning 0+2 combinations fails here. In this case the only non-zero orbital interactions are between the pseudo-p orbitals of the center ribbon and the pseudo-d_s orbitals of the outer two and between the pseudo d orbital of the center ribbon and the pseudo-d_a orbital of the outer two, also shown in Figure 65. No analysis of binary orbital combinations suffices to define the interaction. Examination of the possible ternary interactions shown in Figure 65 indicates that a simple pattern of alternation of the interactions with ribbon mode change does not exist. The effects of changing the mode of a ribbon depend on which combination is considered. This problem did not occur in any of the previous cases. In particular, changing the mode of either of the outer ribbons on the 000 or 020 combinations causes a change in the d-d_a interaction, while the same mode change in the 202 or 222 combinations causes a change in the p-d_s interaction. This difference makes the group description impossible since the same mode change operators must be mapped to two interaction change operations. The proper algebraic description is depicted in Figure 66. This construction is called a groupoid and is mapped homomorphically onto the interaction change group in the same way the groups were earlier. The image of each arrow of the groupoid in the interaction change group is indicated by the designations on the arrows in the figure. This point will not be developed further here. Groupoids and their applications to chemical problems will be discussed more completely in the next section.

FIGURE 66

Illustration of Groupoid described in Text



Remaining arrows obtained by composition of those shown

Some uses and extensions of these ideas can now be described. Goldstein and Hoffmann mentioned that the direction of isomerizations which interconvert topologies could be predicted. Such a prediction is based only on stabilization or destabilization caused by these through-space interactions. Other reasons for predicting reaction directions may agree with or oppose this prediction. The set of ribbon mode combinations for any topology yields a regular representation of the group. Likewise the set of favorable interactions forms a regular representation of the interaction change group. The homomorphism $M_c \rightarrow I_c$ induces a mapping of these sets dependent upon an arbitrary choice of one member of the mode combination set. All of the other combinations are then compared to this arbitrary combination. This process can be made clearer pictorially as in Figure 67 for the laticyclic case. The arbitrary combination chosen is the destabilized $0+0+0$ combination. The relative stabilities of the other combinations are then completely determined by the induced set mapping. Repeating this procedure for the longicyclic case (Figure 68) using the same choice of the $0+0+0$ combination allows a mapping from the laticyclic set to the longicyclic set to be made. From this mapping it is apparent which combinations prefer which topology (Figure 69). No account is taken of the strengths of the interactions.

Further use of these concepts can be made by recalling the similarity of this problem of through-space interaction to that of transition state stabilization in concerted reactions[80]. This

FIGURE 67

Illustration of Induced Set Mapping Showing Stabilization of Ribbon Combinations in the Laticyclic Topology

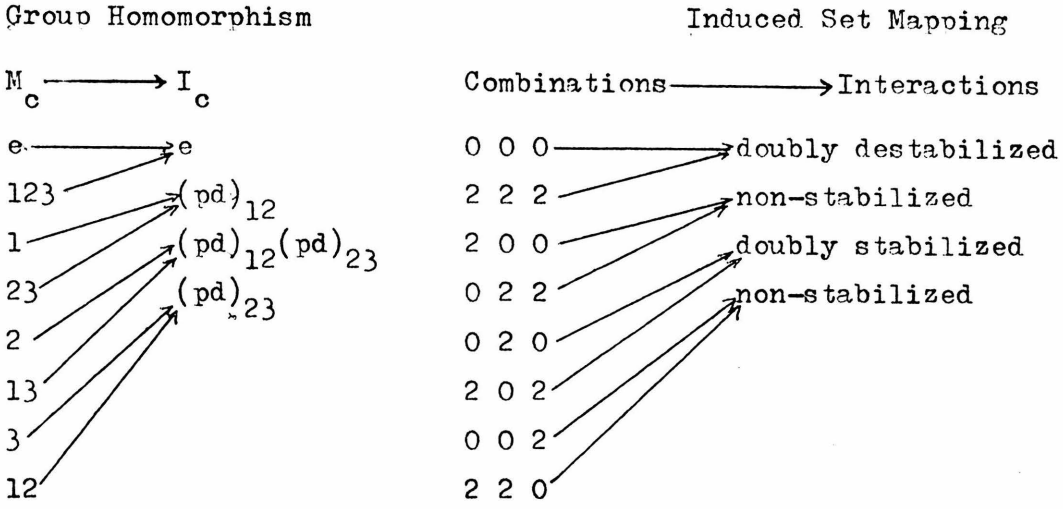


FIGURE 68

Illustration of Induced Set Mapping for Longicyclic Topology

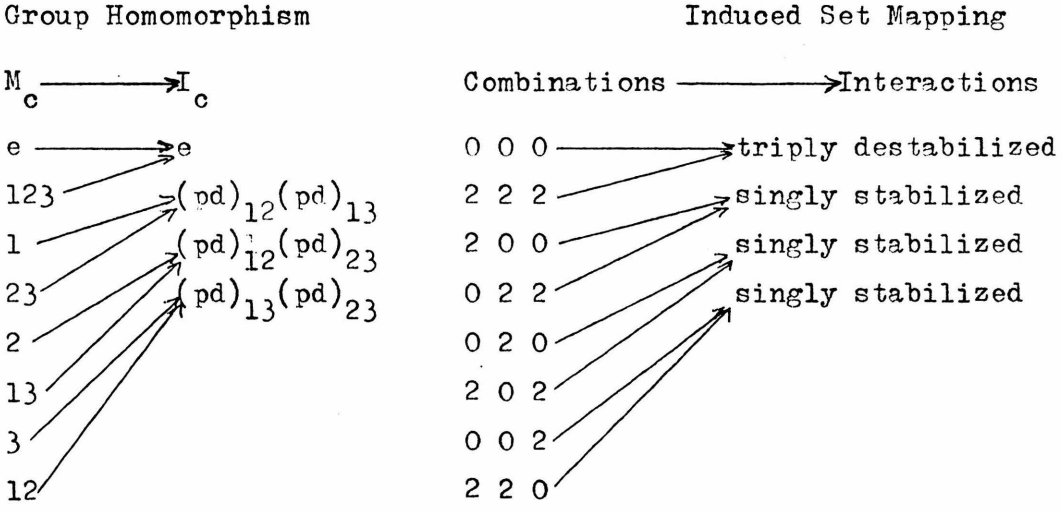


FIGURE 69

Illustration of Induced Set Mapping from Laticyclic to Longicyclic

Stabilization	Laticyclic Combination		Longicyclic Combination	Stabilization	
(-2)	0 0 0	⟶	0 0 0	(-3)	disfavored
(-2)	2 2 2	⟶	2 2 2	(-3)	disfavored
(0)	2 0 0	⟶	2 0 0	(+1)	favored
(0)	0 2 2	⟶	0 2 2	(+1)	favored
(+2)	0 2 0	⟶	0 2 0	(+1)	disfavored
(+2)	2 0 2	⟶	2 0 2	(+1)	disfavored
(0)	0 0 2	⟶	0 0 2	(+1)	favored
(0)	2 2 0	⟶	2 2 0	(+1)	favored

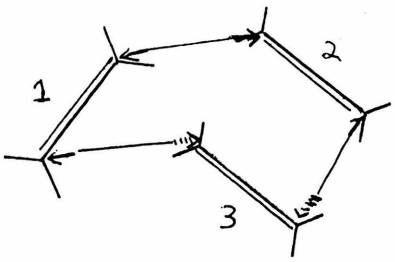
analogy was alluded to earlier when it was noted that the second example in Figure 62 corresponds to the transition state for the allowed $2_s + 2_a$ cycloaddition. Woodward and Hoffmann[81] make use of the concept of a "continuum of topologically equivalent . . . processes" in describing the relationship of the two transition state geometries shown in Figure 70. This concept can be made precise using the group homomorphisms derived above. These cases can be analyzed as the pericyclic interaction of two ribbons since ribbons can be added acyclically[77]. The resulting homomorphisms are shown underneath the respective drawing in Figure 70. The mathematical condition for a topological equivalence based on continuous deformations can be stated precisely. If there is a continuum of topologically equivalent transition state geometries going from B to C then there will be isomorphisms between the respective M_c and I_c groups which makes the diagram shown at the bottom of Figure 70 commute. This is apparent because the interactions in all the intermediate geometries must be the same. The converse is not true; the existence of an isomorphism does not establish the existence of a topological equivalence. The contrapositive: no commutative diagram implies no equivalence can be used to establish the topological difference between the "topologies" used as examples so far. This sort of mathematical operation will be discussed in greater detail in the next section.

The next application requires the development of an independent concept. Consider the degenerate rearrangements of 1,5-hexadiene. These can proceed by a 3,3-sigmatropic shift (Cope rearrangement) or

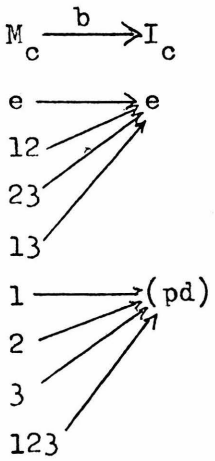
FIGURE 70

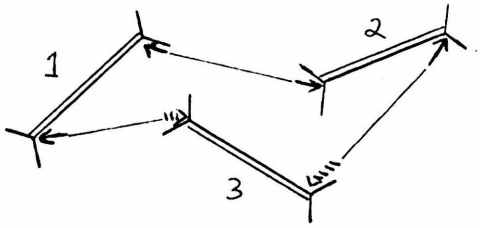
Topologically Equivalent Transition State Geometries for a

$\pi_2 s + \pi_2 s + \pi_2 s$ cycloaddition

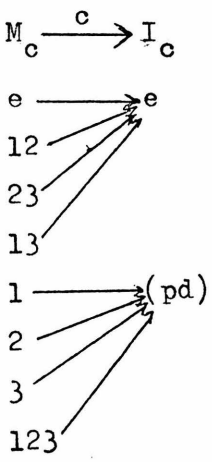


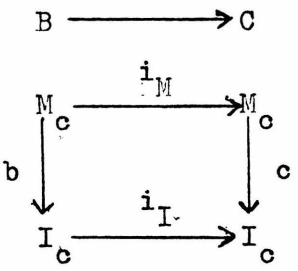
B





C





by one of two possible 1,3-sigmatropic shifts as shown in Figure 71. A "reaction group" G_r can be constructed which includes all these degenerate rearrangements. Such a group resembles the non-rigid symmetry groups derived by various authors[82] to assign microwave spectra. A convenient representative of this group is by the permutations of the substituents (most likely protons) since these changes could in principle be detected experimentally by appropriate labelling. For example, the 3,3 shift with the stereochemical outcome shown in Figure 72 is represented by the indicated permutations of the numbered hydrogens. The resulting group is of order 64 and can be conveniently decomposed as the direct product $D_4 \times D_4$. Each D_4 component represents one of the possible 1,3 sigmatopic isomerizations. The 64 permutations represent all the possible changes of 1,5-hexadiene by single or successive [3,3] or [1,3] sigmatopic rearrangements.

Now the effect of the mode change group on this reaction group will be demonstrated. Some of the permutations in the reaction group represent symmetry-allowed reactions, while others represent symmetry forbidden reactions. Choose an allowed reaction r . A correspondence between allowed and disallowed reactions can be made with the following sequence of operations: (mode change)(preferred reaction corresponding to r)(mode change). Stated differently, if the 1,5-hexadiene could be "ribbon extended" to a 1,3,7-octatriene, the preferred reaction allowed to occur, and then converted back into a 1,5-hexadiene, the overall permutation would be different from r . This process determines the <u>action</u> of the mode change group M_c on the reaction group G_r . This

FIGURE 71

Degenerate Rearrangements of 1,5-Hexadiene

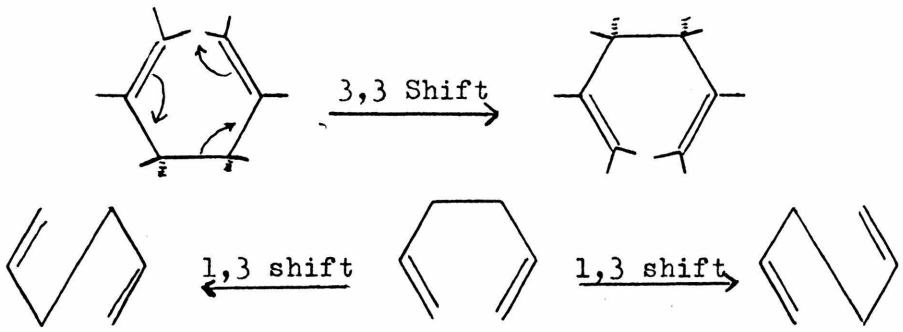
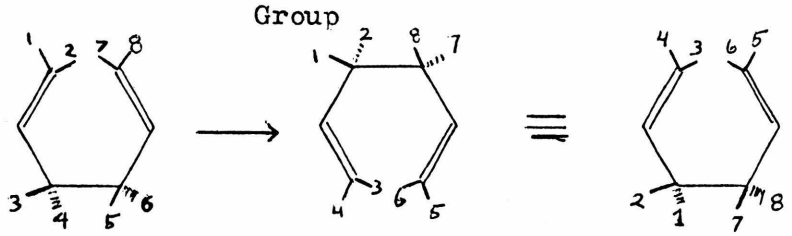


FIGURE 72

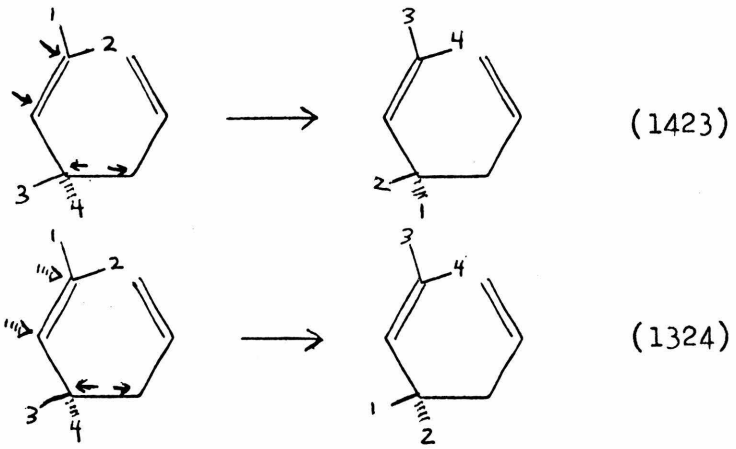
Illustration of Permutation Representation of 1,5-Hexadiene



overall (14)(23)(5867)

FIGURE 73

Illustration of Symmetrically Equivalent 1,3 Shifts



(1423)

(1324)

operation permutes the members of G_r. However, an automorphism of G_r does not result, since there is no obvious correspondence of symmetrically equivalent reactions. That is, two 1,3-shifts like those shown in Figure 73 are equivalent by reflections in the plane of the paper. This problem can be alleviated by forming a factor group of G_r which collects these equivalent reactions into cosets. The factor group is of order 16 and isomorphic to the direct product group $C_2xC_2xC_2xC_2$. A complete listing is given in Table 26 with meaningless names for the cosets which will be used later. The action of the mode change group M_c on this group F_r induces automorphisms of F_r. Stated more precisely, there is a homomorphism from M_c into the automorphism group of F_r. These automorphisms relate allowed and disallowed reactions and are depicted in Figure 74. The effect of the mode change operation is to exchange the roles of these reactions just like the effect of the mode change operation was to interchange stabilizing and destabilizing interactions in the earlier part of this section. This analogy is not surprising in light of the observed similarity of the through-space ribbon interaction problem to the symmetry allowed and disallowed transition state problem[77]. In fact, the mode change group can be regarded as defining a further symmetry of these chemical systems. This notion can be made more precise by observing that the homomorphism from M_c into $Aut(F_r)$ determines a semi-direct product of the two groups[83]. Since F_r is actually a dynamic symmetry group for the 1,5-hexadiene, this process has enlarged its symmetry group. Hence the existence of allowed and disallowed reactions for a chemical structure can be thought of as a symmetry of the structure (at least when considering

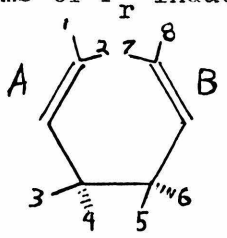
Table 26

Listing of Factor Group of Reaction Group

E	R	RS	S
e	(58)(76)	(56)	(5867)
(12)(34)	(12)(34)(58)(67)	(12)(34)(56)	(12)(34)(5867)
(56)(78)	(57)(68)	(78)	(5768)
(12)(34)(56)(78)	(12)(34)(57)(68)	(12)(34)(78)	(12)(34)(5768)
P	**PR**	**PRS**	**PS**
(13)(24)	(13)(24)(58)(67)	(13)(24)(56)	(13)(24)(5867)
(14)(23)	(14)(23)(58)(67)	(14)(23)(56)	(14)(23)(5867)
(13)(24)(56)(78)	(13)(24)(57)(68)	(13)(24)(78)	(13)(24)(5768)
(14)(23)(56)(78)	(14)(23)(57)(68)	(14)(23)(78)	(14)(23)(t768)
PQ	**PQR**	**PQRS**	**PQS**
(12)	(12)(58)(67)	(12)(56)	(12)(5867)
(34)	(34)(58)(67)	(34)(56)	(34)(5867)
(12)(56)(78)	(12)(57)(68)	(12)(78)	(12)(5768)
(34)(56)(78)	(34)(57)(68)	(34)(78)	(34)(5768)
Q	**QR**	**QRS**	**QS**
(1324)	(1324)(58)(67)	(1324)(56)	(1324)(5867)
(1423)	(1423)(58)(67)	(1423)(56)	(1423)(5867)
(1324)(56)(78)	(1324)(57)(68)	(1324)(78)	(1324)(5768)
(1423)(56)(78)	(1423)(57)(68)	(1423)(78)	(1423)(5768)

FIGURE 74

Automorphisms of F_r induced by M_c



Change A	Change B	Change A and B
E ⟶ E	E ⟶ E	E ⟶ E
R ⟶ R	R ⟶ S	R ⟶ S
RS ⟶ RS	RS ⟶ RS	RS ⟶ RS
S ⟶ S	S ⟶ R	S ⟶ R
P ⟶ Q	P ⟶ P	P ⟶ Q
PR ⟶ QR	PR ⟶ PS	PR ⟶ QS
PRS ⟶ QRS	PRS ⟶ PRS	PRS ⟶ QRS
PS ⟶ QS	PS ⟶ PR	PS ⟶ QR
PQ ⟶ PQ	PQ ⟶ PQ	PQ ⟶ PQ
PQR ⟶ PQR	PQR ⟶ PQS	PQR ⟶ PQS
PQRS ⟶ PQRS	PQRS ⟶ PQRS	PQRS ⟶ PQRS
PQS ⟶ PQS	PQS ⟶ PQR	PQS ⟶ PQR
Q ⟶ P	Q ⟶ Q	Q ⟶ P
QR ⟶ PR	QR ⟶ QS	QR ⟶ PS
QRS ⟶ PRS	QRS ⟶ QRS	QRS ⟶ PRS
QS ⟶ PS	QS ⟶ QR	QS ⟶ PR

degenerate isomerizations; more general cases will be discussed in the next section.

E. General Chemical-Mathematical Considerations

The purpose here is to describe the applicability of some concepts of categorical algebra to problems in chemistry. This will be done by briefly covering the concepts of category theory, discussing general problems in its application, and treating several chemical examples.

Category theory is a relatively new field of mathematics, having been born in 1945[84] and developed most extensively after 1954. A number of comprehensive introductory treatments of the subject have recently appeared. An incomplete listing is given in Reference (85). Any of these may be consulted for a more detailed explanation of the basic concepts. A category consists of a class of objects $(A,B,C,\cdots)$ and a class of morphisms or arrows $(f,g,h,\cdots)$ subject to several axioms. Each arrow relates two (not necessarily different) objects which are called its domain and codomain. A is the domain and B is the codomain of f.

$$A \xrightarrow{f} B$$

Furthermore, each arrow uniquely determines its domain and codomain. That is, the arrow f cannot relate any other objects and relates A and B uniquely. Whenever the codomain of one arrow is the same as the domain of a second arrow, a composite arrow is defined which relates the domain of the first arrow and the codomain of the second.

$$A \xrightarrow{f} B \xrightarrow{g} C \qquad\qquad A \xrightarrow{g \cdot f} C$$

The composite is written $g \cdot f$ rather than $f \cdot g$ by convention. This composition must be associative, so that $(h \cdot g) \cdot f = h \cdot (g \cdot f)$ if the domain of h matches the codomain of g. For each object there must be an identity arrow which serves as a left and right identity for the appropriate compositions

$$B \xrightarrow{i} B \qquad if = f \qquad and \qquad gi = g$$

Familiar examples of categories are the class of all sets and mappings (Set), all groups and homomorphisms (Grp) and all topological spaces and continuous functions (Top). Similarly any class of structural sets and structure preserving maps will be a category. These are termed concrete categories because they are based on underlying sets and mappings. More abstract categories are possible. A set with any binary relation, a lattice, and a partially ordered set are three more examples. A group is a category with one object and arrows the elements of the group acting on the object. Composition is multiplication of group elements. Finally a diagram of objects and arrows is a category if compositions are properly defined.

$$0 \xrightarrow{f} 1 \xrightarrow{g} 2 \xrightarrow{h} 3$$

If the composites gf, hg and hgf and four identity arrows are included, this diagram is a category.

A number of familiar set-theoretic mapping concepts are abstracted in category theory. An arrow f is termed epi if $gf = hf$ implies $g = h$, that is, f is right cancellable. In familiar

categories such as Set or Grp, epi arrows are the surjective, epimorphic or onto mappings. Similarly (actually dually), an arrow k is termed monic if $kg = kh$ implies $g = h$, that is, k is left cancellable. Monic arrows are the injective, monomorphic or into mappings. An isomorphism (bijection) is an arrow which has a right and left inverse, and in Set or Grp corresponds to an arrow which is both monic and epi. A category in which all arrows are invertable, that is all arrows are isomorphisms, is called a groupoid.

A subcategory of a category is a subclass of the objects and a subclass of the arrows which themselves satisfy the axioms. For example, the category of abelian groups (Ab) is a subcategory of the category of all groups.

A structure preserving mapping between categories is called a functor and consists of two parts, an object map and an arrow map. Any arrow which is an identity in the domain of a category must be mapped to an identity in the codomain. Composition of arrows must be preserved, so that if F is a functor from the category C to category D and if f and g are composable arrows in C, $F(g)\ F(f) = F(g \cdot f)$ in D. An example of a functor from Top to Ab is the assignment of a homology group to a topological space. Continuous functions go over to group homomorphisms. As another example, consider the functor from Grp to Set which assigns to a group its underlying set of elements but "forgets" the multiplicative structure. Homomorphisms go over to the underlying set mappings. This type of functor is called a forgetful functor. Another example is the identity

functor, symbolized 1_C for the category $\underline{C}$ which assigns all objects and arrows to themselves.

A mapping of one functor to another is called a natural transformation. This mapping is formulated as follows. Consider two functors F and G from $\underline{C}$ to $\underline{D}$. A natural transformation N is a function which assigns to each object of $\underline{C}$ an arrow from its image by F in $\underline{D}$ to its image by G in $\underline{D}$. That is, an arrow, $N(c)\colon F(c) \to G(c)$ in the category $\underline{D}$ is assigned to c, an object of $\underline{C}$. This assignment is subject to the following condition. If h is an arrow of $\underline{C}$ taking c to c', the square diagram in $\underline{D}$ must commute

$$\begin{array}{ccc} & \overset{N(c)}{} & \\ F(c) & \to & G(c) \\ {\scriptstyle F(h)}\downarrow & & \downarrow{\scriptstyle G(h)} \\ & \underset{N(c')}{} & \\ F(c') & \to & G(c') \end{array}$$

The natural transformation is symbolized $N\colon F \dot{\to} G$. An example of a natural transformation is determined by the projection of the fundamental group of a topological space (with a base point) onto its first singular homology group. The kernel of the projection is the commutator subgroup of the fundamental group. The natural transformation is from the fundamental group functor to the singular homology group functor, both of which go from the category of pointed topological spaces to the category of groups. All functors from one category to another form a category themselves with natural transformations as arrows. This construction is appropriately called a functor category.

This admittedly sketchy outline of the basis of categorical algebra can be expanded and clarified by referring to the more rigorous treatments given in Reference (85).

Now the subject of functors from subcategories, which is generally unemphasized in most discussions of category theory, will be treated in greater detail, because of its apparent relevance to the applications discussed below. Most categories, in particular Set and Grp, can be decomposed into subcategories in the following fashion:

$$\begin{array}{ccccc} \underline{D} & \rightarrow & \underline{G} & \rightarrow & \underline{M} \\ & & \downarrow & & \downarrow \\ & & \underline{E} & \rightarrow & \underline{C} \end{array}$$

$\underline{C}$ is the category being decomposed. $\underline{E}$ is the subcategory which consists of all the objects of $\underline{C}$ and all the epi arrows. The arrow from $\underline{E}$ to $\underline{C}$ represents the functor or embedding of $\underline{E}$ into $\underline{C}$. $\underline{M}$ is the subcategory which consists of all the objects and monic arrows, and is similarly embedded into $\underline{C}$. $\underline{G}$ is the subgroupoid of $\underline{C}$ which consists of all the objects and all isomorphisms. Since isomorphisms are both monic and epi, $\underline{G}$ is contained in $\underline{M}$ and $\underline{E}$. $\underline{D}$ is the discrete subcategory of $\underline{C}$ and consists of all the objects and identity arrows only.

Often in discussions of categories and functors examples are given of nonfunctorial constructions, such as the center of a group or the automorphism group of a group or set[86]. It will be shown that these can be functorial construction if the domain of the functor is restricted to the appropriate subcategory. Consider the assignment of

the symmetric group S_n to the finite set with n elements. This is clearly not a functorial construction for all mappings of finite sets. Any non-bijective, surjective set mapping cannot be assigned a group homomorphism of the appropriate symmetric groups. There is no homomorphism of S_n onto S_{n-1} for $n \geq 5$. However, if the assignment is restricted to monic arrows only, a functorial construction results. Injective set mappings go to the corresponding injective group mappings since S_n has a subgroup S_m for all $m \leq n$. Permutations of a set go to the inner automorphism of the symmetric group induced by that permutation by conjugation. Thus a functor from the monics-only subcategory of the category of finite set to the monics-only subcategory of Grp has been constructed.

As another example, consider the assignment of the center of a group to the group (the center Z_G of a group G is the subgroup of all Z in G such that $zg = gz$ for all g in G.) This assignment fails to be functorial for monic arrows. For example, consider the injection of D_2 into the symmetric group S_4. D_2 is abelian, hence is its own center. S_4 has a trivial center, so that the injection of D_2 into S_4 does not induce a corresponding map of the centers. However, if the assignment is restricted to the subcategory of epi arrows, a functorial construction results. If $f: A \to B$ is the epimorphic group homomorphism, $a_1a_2 = a_2a_1$ in A implies $f(a_1)f(a_2) = f(a_2)f(a_1)$ in B. Thus if a_1 commutes with everything in A, $f(a_1)$ must commute with everything in B since all of B is in the image of f. This defines a mapping of the center of A to

the center of B, which need not be epi. Thus a functor Z from $\underline{E}_{Grp}$ to $\underline{Grp}$ has been defined which assigns each group its center. A natural transformation $Z \dot{\to} I$ can be defined where I is the functor which embeds $\underline{E}_{Grp}$ into $\underline{Grp}$. This natural transformation simply embeds the center of a group into the group so that the diagram commutes:

$$Z \dot{\to} I \qquad \begin{array}{ccc} Z(G) & \to & G \\ {\scriptstyle Z(f)}\downarrow & & \downarrow{\scriptstyle f} \\ Z(H) & \to & H \end{array}$$

f is the epi arrow in $\underline{Grp}$.

Another functor from $\underline{E}_{Grp}$ to $\underline{Grp}$ is defined by the assignment of the inner automorphism group of a group G to G. (The inner automorphism group is the group of automorphisms of a group G caused by conjugation by elements of G.) For an epimorphism $f: A \to B$, $a_1^{-1}a_2a_1$ is taken to $f(a_1^{-1})\, f(a_2)\, f(a_1)$, thus mapping the inner automorphism of A induced by a_1 to the inner automorphism of B induced by $b(a_1)$. A natural transformation from I to Aut_i is defined by the projection of a group onto its inner automorphism group

$$\begin{array}{ccc} G & \to & Aut_i(G) \\ {\scriptstyle f}\downarrow & & \downarrow \\ H & \to & Aut_i(H) \end{array} \qquad\qquad I \dot{\to} Aut_i$$

The kernel of this projection is just the center of the group so the two natural transformations can be composed

$$\begin{array}{ccccc} Z(G) & \rightarrow & G & \rightarrow & Aut_i(G) \\ \downarrow & & f \downarrow & & \downarrow \\ Z(H) & \rightarrow & H & \rightarrow & Aut_i(H) \end{array}$$

(The natural transformation $Z \dot{\rightarrow} I$ is the kernel of $I \dot{\rightarrow} Aut_i$ in the functor category $\underline{E}_{Grp}$ to $\underline{Grp}$).

It is now apparent that any object assignment from one category to another can be made functorial by sufficient restriction of the arrows in the domain category. Clearly this can result in a completely trivial case in which the domain is the discrete subcategory. Two points can be made about the relevance of all this to applications of categorical methods. First, deciding which arrows are to be included in the domain category is of utmost importance. Too restricted a domain will give trivial constructions while too large a domain restricts the constructions possible. There is even no a priori reason why the arrows have to be structure-preserving. Secondly, changing the domain is itself a potentially useful variable. This is apparent from the above example in which the center and automorphism group assignments are not recognized as functorial constructions until the domain is sufficiently restricted.

Next, the question of why any of this should be relevant to chemistry will be discussed. First, treating any problem in chemistry or anything else by mathematical methods constitutes an assignment of the problem to a mathematical object and hence a potential functorial construction. Secondly, a category is itself a rather weak algebraic structure and can be thought of as a generalized group. To show this,

only the arrows of a category are considered. Axiomatically, a group is a set on which a two-sided multiplication is defined for all pairs of elements (composability). Further, a two-sided identity, inverses of all elements, and associativity of the multiplication is required. Consider the two axioms of composability and invertibility and the effect of their removal. This is shown in Table 29.

Table 29

	composable?	invertible?
group	yes	yes
semigroup	yes	no
groupoid	no	yes
category	no	no

The three algebraic objects, semigroups, groupoids and categories can be regarded as generalized groups. Now an arrow representing a chemical reaction or any change can also have these properties. A reaction that has the same reactant and product is a degenerate reaction and can certainly be reversible. An example of the applications of group theory to such a problem was discussed in the previous section. Degenerate irreversible reactions would yield a semigroup. Reactions with different reactants and products are not generally composable, yet can certainly be reversible. This structure is described by a groupoid and will be discussed more fully later in this section. Finally, noncomposable, irreversible reactions describe a category.

An analogy of the two-level applicability of categorical ideas to chemistry can be made to the applicability to topology. Consider the assignment of a collection of atoms to its potential energy surface. Arrows would be mappings or some comparison of these surfaces. This corresponds to the consideration of topological spaces and continuous functions between them. For a given space, the fundamental groupoid can be constructed by considering a set of disjoint points in the space and all paths between them. This would correspond to the consideration of chemical reactions as paths on the potential energy surface. Treating topology functorially is done in the field of algebraic topology. It was for this reason that the term algebraic chemistry was chosen.

Some work on the application of categorical ideas outside of pure mathematics has appeared. In an extensive series of papers, Rosen[87] has described abstract biological systems as sequential machines. Briefly, he was able to describe the organizational properties (metabolic and genetic) of a biological system as a sequence of mappings. In the simplest case this corresponded to[88]:

$$A \longrightarrow B \longrightarrow H(A,B)$$

in which A represents the set of inputs to the various components. of the system and B represents the set of outputs. H(A,B) is the set of mappings of A to B , typically called a hom-set. The mapping f refers to the metabolic parts of the system which relate

outputs to the various metabolic components represented as mappings in $H(A,B)$. A sequential machine is a quintuple $L = (S,M,N,g,h)$ where S is the set of states of the machine L, M is the set of inputs to L, N is the set of outputs of L, $g: M \times S \to S$ is the next state function and $L: M \times S \to N$ is the output function[89]. M generates a monoid (also called a semigroup) by its action on the set of states. Rosen was able to define his biological systems as sequential machines by the following assignment[89]

$$S \equiv H(A,B)$$

$$M \equiv A$$

$$N \equiv B$$

$$g: A \times H(A,B) \to H(A,B) \equiv g(a,f) = \phi_F(f(a))$$

$$h: A \times H(A,B) \to B \equiv h(a,f) = f(a)$$

h is an evaluation map, while g is determined by the original mapping.

A morphism of sequential machines $L \to L'$ is a triple of mappings

$$S \overset{k_1}{\to} S'$$

$$M \overset{k_2}{\to} M'$$

$$N \overset{k_3}{\to} N'$$

subject to the commutavity conditions

$$\begin{array}{ccc} S \times M & \xrightarrow{k_1 \times k_2} & S' \times M' \\ g \downarrow & & \downarrow g' \\ S & \xrightarrow{k_1} & S' \end{array} \qquad \begin{array}{ccc} S \times M & \xrightarrow{k_1 \times k_2} & S' \times M' \\ h \downarrow & & \downarrow h' \\ N & \xrightarrow{k_3} & N' \end{array}$$

With composition defined in the obvious manner, the class of sequential machines and arrows forms a category. Rosen was able to show that biological systems as sequential machines formed a subcategory if the triple of arrows was restricted in the following manner. k_1 must be onto (epi) while k_2 and k_3 must be one-to-one (monic)[90].

Give'on[91] has considered categories of transition systems. Briefly, a transition system A is composed of a fixed monoid M, a set of states S(A) and a transition function B f: $S(A) \times W \to S(A)$, where f describes the action of the monoid on the set of states. A category of transition systems over a fixed monoid is constructed with homomorphisms of the type $A \xrightarrow{g} B$ such that the diagram commutes

$$\begin{array}{ccc} S(A) \times W & \xrightarrow{g \times 1} & S(B) \times W \\ f \downarrow & & \downarrow f' \\ S(A) & \xrightarrow{g} & S(B) \end{array}$$

This category resembles the category of right actions of a monoid.described by MacLane[92]. Give'on shows that for a certain broad class of monoids the properties of a given transition system are "retrievable" by studying arrows to and from the systems. Stated differently, study of the category of transition systems is in principle equivalent to the

study of individual systems, although the former may certainly not be preferred.

Hajek[93] has considered the application of categorical concepts to dynamical system theory. A dynamical system (also a transformation group) is composed of a topological space X (the phase space), a topological group (phase group) and a mapping $\pi: X \times T \to X$ subject to certain conditions[94]. Hajek[93] considers the situation in which T is R_1, the group of real numbers. Considerable attention is given to the problem of which morphisms should be included in the category. Morphisms $P \times R^1 \to P' \times R^1$ can be phase-space maps combined with a group identity homomorphism, an identity phase space map combined with a group homomorphism, or any of several combinations of phase-space maps and group homomorphisms. That is, either the phase-space is changed, the group is changed or both are changed. The different results, depending on the class of morphisms chosen, are cited.

Finally, a possible application to theoretical physics can be formulated. A quantum mechanical system can be described as a dynamic system in which a Lie algebra of observables L acts on a Hilbert space V of states, subject to a number of conditions[95]. This can be represented as $L \times V \to V$. Morphisms can map the space, the algebra or both. In particular, physical systems with isomorphic Lie algebras will have identical spectra. This is shown for the case of angular momentum and isospin by Lipkin[96]. However, another type of arrow includes more significant changes in physical systems. The intuitive process of changing a physical system by letting a constant go to zero or infinity can be mathematically described. Well known examples are

the passage from relativistic mechanics to classical mechanics by letting the speed of light go to infinity and the passage from quantum mechanics to classical mechanics by letting Planck's constant go to zero. These changes can be described by Lie algebra contractions. A contraction is a non-continuous non-homomorphism which changes a Lie algebra into another Lie algebra with a smaller or equal dimensional derived algebra. As formulated by Inonu and Wigner[97] and elaborated by Saletan[98] and Conatser[99], a contraction is defined on a Lie algebra V, decomposed as $V = W + U$. The linear transformation:

$$A_t = \begin{vmatrix} I_1 & 0 \\ 0 & 0 \end{vmatrix} + t \begin{vmatrix} D & 0 \\ 0 & I_2 \end{vmatrix}$$

on $W + U$ is defined where I_1 and I_2 are identity transformations and D is any non-singular operator on W. For any two operators X,Y in the original algebra, the new algebra is defined by the relation:

$$[X,Y]_c = \lim_{t \to 0} A_t^{-1}[A_tX, A_tY]$$

If $[X,Y] = 0$, then $[X,Y]_c = 0$; however, if $[X,Y] \neq 0$, $[X,Y]_c$ may be zero, thus the derived algebra of the new algebra can be only smaller or of the same dimension. Composition of contractions does not necessarily give a contraction, but does give a non-continuous non-homomorphism. Such arrows yield a category if it is noted that W can be all of V, thus giving identity arrows. As a functorial construction, assignment is restricted to the subcategory of physical

systems and non-continuous changes in which the system becomes "less complicated."

All of these possible applications of categorical concepts seem to fall into the same pattern. The objects are defined by the action of some algebraic structure on some (un)structured set. Arrows can relate the algebra, the set or both and need not be structure preserving, although they generally are. In particular, for any given problem different classes of arrows are most relevant. Observation of this pattern suggests a justification for the claim of applicability of categorical concepts. Category theory has proved useful in the development of homological algebra, which is largely a study of the category of R-modules and homomorphisms[100], R-Mod. An R-module M_R is defined by the action of a unitary ring R on an abelian group, $M: R \times M \to M$ [101]. (This is a left module, a right module is defined by $M \times R \to M$) . The analogy to the above examples is clear. However, the category R-Mod has a number of "nice" properties which make it into an abelian category, on which the standard methods of homological algebra are applicable. None of the above examples were abelian categories. Morphisms of R-modules map the abelian group and leave the ring unchanged. Hence the category studied, R-Mod, consists of all the modules over a particular ring. Change of rings is handled functorially, that is, a functor R-Mod $\to$ R'-Mod is induced by a ring homomorphism $R' \to R$ [102].

Some applications of these and earlier discussed concepts to specific chemical problems will now be described. The original

formulation of the Woodward-Hoffmann rules considered a chemical reaction as a mapping of orbitals, their correlation diagram[103]. The condition for an allowed reaction can be stated as the existence of bijective mapping r of orbitals so that diagram commutes:

$$\begin{array}{ccc} (b,a) & \rightarrow & (b,a) \\ \uparrow & & \uparrow \\ O_x & \rightarrow & O_y \\ \downarrow & & \downarrow \\ (A,S) & \rightarrow & (A,S) \end{array}$$

O_x is the set of orbitals for the reactant X , (b,a) refers to bonding or antibonding, and (A,S,) refers to symmetric or antisymmetric with respect to the symmetry operation(s) in question.

Functorially, each chemical structure is assigned the diagram: $(a,b) \leftarrow 0 \rightarrow (S,A)$, and each allowed concerted reaction the commutative diagram. This can be described as a functor $\underline{\text{Chem}} \rightarrow \underline{\text{Set}}^{\bullet\leftarrow\bullet\rightarrow\bullet}$. The former category is the one described above, the latter is a functor category as described by MacLane[104].

At this point the concept of a forgetful functor becomes relevant. The above process assigns to each chemical structure a set of molecular orbitals and in a sense forgets every other property of the structure. It is well known that errors are made by over-simplification. This corresponds to finding arrows in the codomain category which are not in the image of the functor from the domain category. As another example, the existence of an isomorphism between the fundamental groups of two topological spaces does not insure the existence

of a continuous function between them, only that they are of the same homotopy type.

The construction described in Section IID can be treated functorially. In particular, each ribbon topology is assigned the construction:

$$\begin{array}{ccc} M_c \times M & \rightarrow & M \\ \downarrow & & \downarrow \\ I_c \times I & \rightarrow & I \end{array}$$

M_c is the mode change group (or groupoid in more complicated cases), M is the set of modal combinations, I_c is the interaction change group, and I is the set of interactions. It was already noted that M_c acts on M as a regular representation, as does I_c on I. Thus an object assignment is defined. If the arrows between the ribbon topologies correspond to continuous deformations (Woodward and Hoffmann's topological equivalence[81]) the arrows between the diagrams are those which change the groups only. As was noted earlier, topological equivalences go over to commutative diagrams of groups by this functor. However, if the arrows are reactions which change topology, they are assigned the set mappings which determine preferred reactions based on these considerations (Figure 69). Thus two different functorial constructions result depending on the information desired.

It is interesting to look at this problem in the light of a more rigorous description to topology. These ribbon topologies could, in principle, be assigned standard topological invariants such as fundamental groups or homology groups. However, one counter-example

shows that these will not suffice to determine chemically different cases. If the ribbons are thought of as sheets with corners tied together to represent interactions, it can be shown that the pericyclic and longicyclic "topologies" are topologically identical (rigorously, from a mathematical standpoint). This is shown in Figure 75. Hence the standard algebraic topological methods are inadequate for this problem. On the other hand, it has not been shown in any way that the group construction derived yields a (real) topological invariant. It only suffices to describe the chemical problem at hand.

The group construction derived in Section IIB can also be described functorially. To each chemical structure is assigned the group describing the symmetry relevant to the synthetic design problem. The only relevant arrows are monics, which correspond to considering only a part of the target structure for symmetry. This consideration is useful, since most syntheses proceed through smaller intermediates.

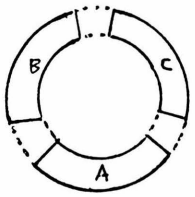
For the next example, an observation of Gust and Mislow[105] will be discussed. In their analysis of the isomerizations of generalized polyaryl methanes they noted a similarity between the analysis for a triphenyl carbonium ion and a transition metal tris chelate. Structures are shown in Figure 76. The purpose here is to make this notion more precise and suggest extensions.

The possible degenerate isomerizations of a triaryl carbonium ion form a group as might be expected. Four changes are possible (Figure 77): flip ring A, flip ring B, flip ring C, and change helicity. These are all of order two and all commute, hence the direct

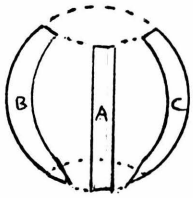
FIGURE 75

Illustration of Pericyclic-Longicyclic Equivalence

Pericyclic

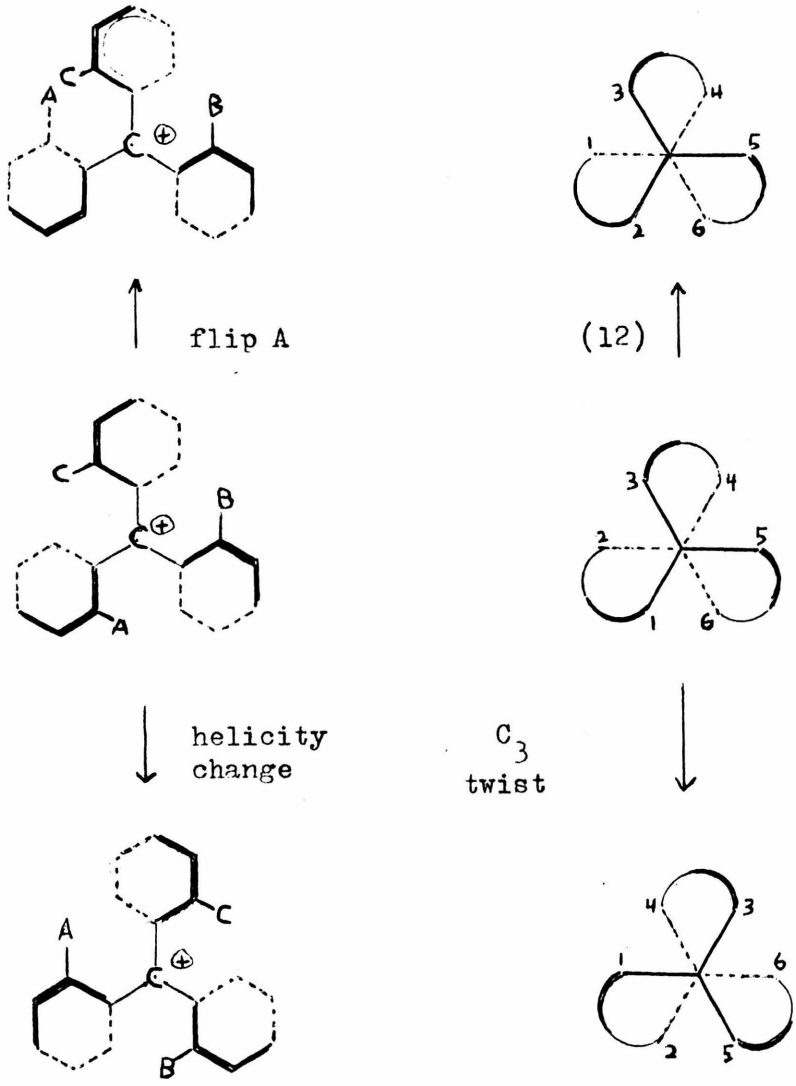


Longicyclic



FIGURES 76, 77, 78

Degenerate Isomerizations of Tri-Phenyl Carbonium Ions and Tris-Metal Chelates



product group $C_2 \times C_2 \times C_2 \times C_2$ of order 16 describes these isomerizations. (The ring flip operation differs from that described by Gust and Mislow). It has been shown by Eaton and Eaton[106] that the degenerate isomerizations of a tris metal chelate likewise form a group. A convenient representation is by permutation of the ligands and twisting along the C_3 axis as shown in Figure 78. This group is also isomorphic to $(C_2)^4$. The isomorphism between the group for the triaryl case and the group for the tris-chelate case describes the desired similarity. This isomorphism is listed in Table 30 along with names for the isomerizations in the tris-chelate case[107,108].

Some mathematical considerations of this case will now be described. A large number of isomorphisms between these two groups are mathematically possible so that the chemical considerations which distinguish this one are of interest. First of all, isomerizations which change chirality must be distinguished from those which do not. These are the operations which change the helicity of the structures. This condition can be stated as a requirement of commutivity of the diagram

$$\begin{array}{ccc} (C_2)^4 & \rightleftarrows & (C_2)^4 \\ \downarrow & & \downarrow \\ C_i & \rightleftarrows & C_i \end{array}$$

where C_i is the inversion group. Similarly, diagrams which relate ring flip operations to ligand permutation operations can be constructed

Table 30

Isomorphism of Isomerization Groups for

Triaryl Methanes and Tris-Chelates

ϕ_3C	$M(chel)_3$
E	E
Flip A	(12)
B	(56)
C	(34)
AB	(12)(56)
AC	(12)(34)
BC	(34)(50)
ABC	(12)(34)(56)
change helicity (H)	C_3 twist (T) Bailar twist
AH	(12)(T) Ray and Dutt twist
BH	(56)(T)
CH	(34)(T)
ABH	(12)(56)(T)
ACH	(12)(34)(T)
BCH	(34)(56)T
ABCH	(12)(34)(56)T

$$\begin{array}{ccc} (C_2)^4 & \rightleftarrows & (C_2)^4 \\ \downarrow & & \downarrow \\ C_{2a} & \rightleftarrows & C_{2(12)} \end{array}$$

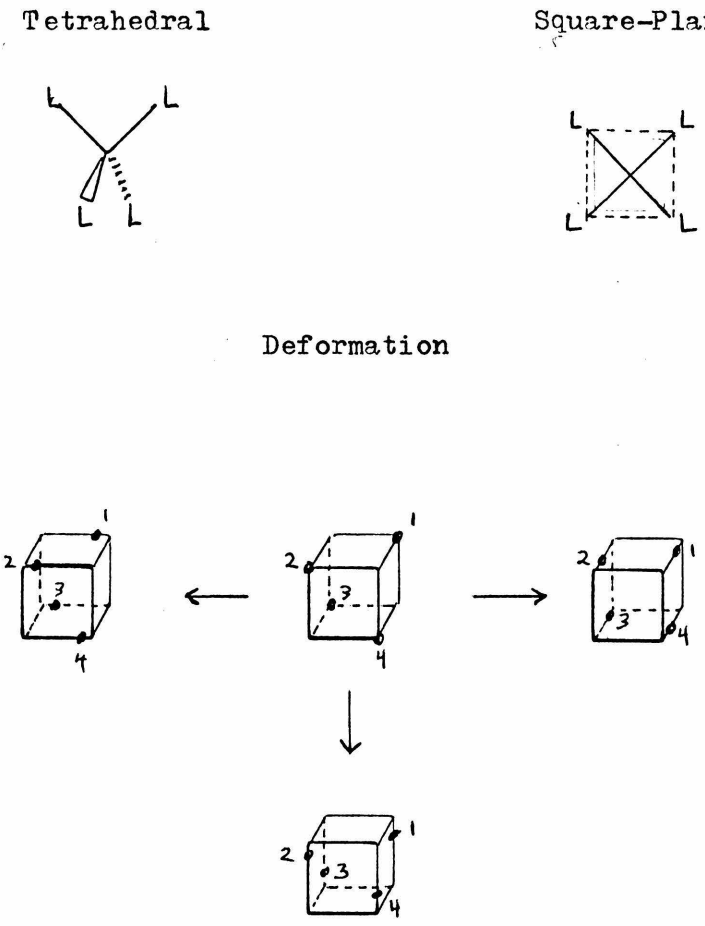
Secondly, these structures have D_3 point group symmetry. A semi-direct product of the point group and the isomerization group can be constructed by considering the action of the point group on the isomerization group. An isomorphism of the resulting groups is required

$$(C_2)^4 \; X_s D_3 \; \rightleftarrows \; (C_2)^4 \; X_s D_3$$

Use can be made of groupoids[109] in treating problems of isomerization between different geometries. The applicability of group theory to degenerate isomerization has already been exploited. Consider the interconversion of the tetrahedral and square-planar geometries for 4-coordinate structures. The drawings in Figure 79 are considered to be in a fixed orientation. There are twenty-four ways of mapping the four ligands on the tetrahedral structure to the square-planar structure. Furthermore, there are twenty-four degenerate isomerizations for each geometry, corresponding to the symmetric group on four objects. This situation is described by a groupoid with two vertex groups isomorphic to S_4 and twenty-four invertible arrows between the two vertices. Functorially, each geometry is assigned a vertex, each degenerate isomerization is assigned to a member of the corresponding vertex group, and each non-degenerate isomerization is assigned to an arrow between the two vertices. With this groupoid it

FIGURE 79

Interconversion of Tetrahedral and Square-Planar Geometries



is possible to determine the differentiable degenerate and non-degenerate isomerizations using the methods described by Klemperer[110].

Further use can be made of this construction by first noting that a groupoid is actually a representation of the wreath product of the vertex group with the symmetric group S_n, where n is the number of vertices. Consider the collection of all n-tuples of arrows of the following type. There must be one and only one arrow with any particular vertex as its domain and similarly, one and only one with this vertex as its range. The collection of all these arrows forms a group of order $|G|^n|S_n|$. Now define an isomorphism of the vertex objects. In the chemical examples considered this is done most conveniently by numbering all the ligands with the same numbers for each geometry. The above group can now be represented by $n\ell^n \times n\ell^n$ permutation matrices where ℓ is the number of members of the vertex set or, in the chemical examples, the number of ligands. The isomorphism to the wreath product is now apparent since the wreath product $GWrS_n$ can be represented by permutation matrices with entries from G[111]. The particular element of g is the one which yields the permutation of ligands caused by a non-degenerate arrow (relative to the numbering). It should be possible to do this proof without consideration of the vertex objects, but such generality is not needed here.

It is now possible to determine the differentiable degenerate cycles of isomerization of any isomer which goes through intermediates of different geometry by forming the double cosets in the wreath

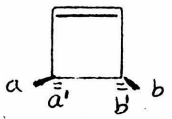
product by a group which is the direct product of the symmetry point groups of all the involved geometries. In the example shown in Figure 79, the degenerate isomerization cycles for either geometric form would be the double cosets of $T \times D_4$ in the wreath product. In principle, all differentiable paths and cycles can be determined by reference to the groupoid or wreath product.

Gilles and Phillipot[82c] have mentioned the possibility of obtaining groups which describe isomerizations between different geometries. The wreath product obtained above is such a group, but its representation by permutation of the structures is not regular, a property shared by point groups and non-rigid symmetry groups. Such groups describing isomerizations between different geometric forms are possible. The requirement for this is that every reaction be defined on every structure. Such a group would be a subgroup of the corresponding wreath product. However, wreath products are very large and enumeration of possible subgroups would be formidable. Hence, a more intuitive construction of an example of this type of group will be described. Consider the butadiene-cyclobutene rearrangement and all its stereochemical consequences (Figure 80). A stereochemical inversion group, isomorphic to D_2, acts on each structure by inverting the substituted centers. This group will be extended by C_2 to include the reactions. This type of group extension is described by Hall[112]. The factor set is shown in Figure 80. The chemical "reason" for this construction is that two consecutive reactions of

FIGURE 80

Construction of Reaction Group for Cyclobutene-Butadiene Interconversions

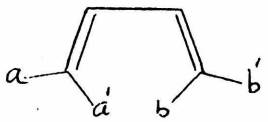
Cyclobutene



D_2 = (e, a, b, ab)

C_2 = (E, R)

Butadiene



Stereochemical inversion Reaction

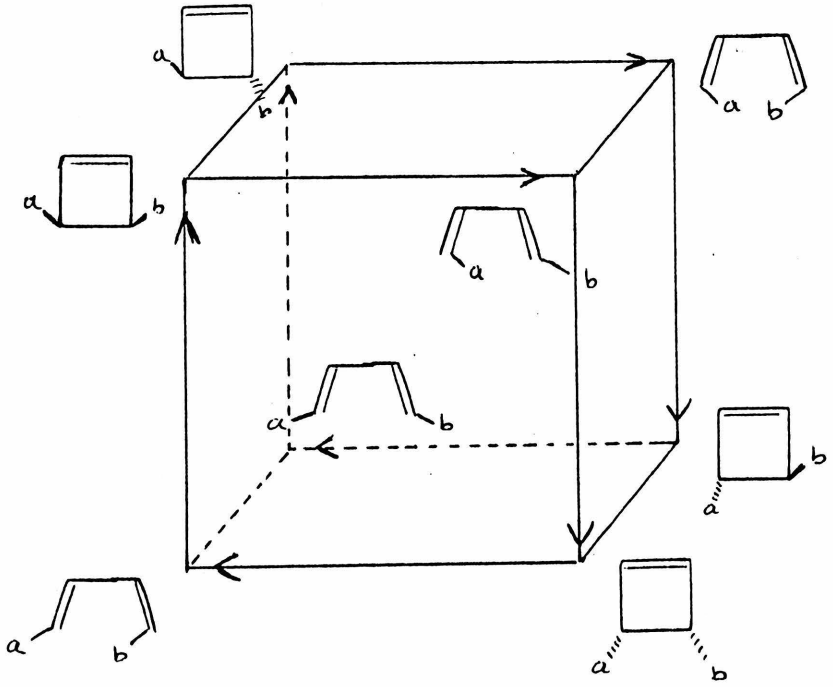
Factor set $C_2 x C_2 \longrightarrow D_2$

$(E,E) \longrightarrow e$

$(E,R) \longrightarrow e$

$(R,E) \longrightarrow e$

$(R,R) \longrightarrow ab$



the same type result in stereochemical inversion at both centers. The resulting extended group is abelian since the automorphisms of D_2 associated with C_2 are all trivial and the factor set mapping is symmetric with respect to permutation of the factors. This extended group is isomorphic to $C_4 \times C_2$ which is isomorphic to the symmetry point group C_{4h}. For this reason, a very convenient graphical representation of this group can be constructed as shown in Figure 80. This graph has C_{4h} symmetry. Rotation clockwise represents the conrotatory reaction, while S_4 rotation represents the disrotatory reaction. This group can also be extended by point group symmetry or by the mode change group as in Section IID. The group obtained by point group extension can be used to determine differentiable reactions[110].

The final example considers the problem of symmetry change in chemical reactions. If the domain category Chem is restricted to those reactions which do not lower symmetry, a functor to the lattice of point groups (thought of as a category) is readily constructed. In any (single-step) reaction which increases symmetry, the symmetry group of the reactant is a subgroup of the symmetry group of the product. Another construction yields a functor from the category of all one-step reactions which change symmetry. Consider the deformation of tetrahedral structure as shown in Figure 79. Each of the three deformed structures is represented by a left coset in T_d by D_{2d}. The permutation of these cosets by right multiplication represents the effect (permutation) of one of the original symmetry operations on

the deformed structures. Cosets fixed by any operation represent structures which retain or conserve that symmetry. The assignment of any chemical structure to this permutation group (which is isomorphic to the least upper bound of the symmetry groups of all structures that can be interconverted) allows all reactions which change symmetry to be expressed as isomorphisms. These isomorphisms are useful in determining symmetry conserved in any reactions, as required for use of the Woodward-Hoffman rules in their original form[103]. These isomorphisms also determine symmetry operations which change effect during the course of a reaction such as those which interconvert reactants and products but are symmetry operations for a transition state. These considerations are useful in using selection rules which limit the size of the symmetry group of a transition state relative to that for the reactants or products[113].

In all these examples the full force of category theory is hardly being used. In fact it may turn out that categorical concepts will only be an intuitive language for formulating problems. MacLane[114] has noted that for the first ten years of its existence "category theory was just a language", even for mathematicians.

REFERENCES

1. a) J. M. McGuire, R. L. Bunch, R. C. Anderson, H. E. Boaz, E. H. Flynn, H. M. Powell and J. W. Smith, Antibiot. Chemotherapy 2, 281 (1952). b) H. Welch, Antibiot. Chemotherapy 2, 279 (1952).

2. P. F. Wiley, K. Gerzon, E. H. Flynn, M. V. Sigal, Jr., O. Weaver, U. C. Quarck, R. R. Chauvette and R. Monahan, J. Amer. Chem. Soc. 79, 6062 (1957). Also other papers in series.

3. D. R. Harris, S. G. McGeachin and H. H. Mills, Tetrahedron Letters, 679 (1965).

4. a) W. D. Celmer, J. Amer. Chem. Soc. 87, 1801 (1965). b) W. D. Celmer, Antimicrobial Agents and Chemotherapy, 144 (1965).

5. T. Kaneda, J. C. Butte, S. Taubman and J. W. Corcoran, J. Biol. Chem. 237, 322 (1962). H. Grisebach, W. Hofheinz and H. Achenbach, Zeit. Naturforsch 17B, 64 (1962). Z. Vanek, M. Puza, J. Majer and L. Dolezilova, Folia Microbiologica 6, 408 (1961).

6. K. Gerzon, E. H. Flynn, M. V. Sigal, Jr., P. F. Wiley, R. Monahan, U. C. Quarck, J. Amer. Chem. Soc. 78, 6396 (1956).

7. a) Jerry R. Martin and William Rosenbrook, Biochemistry 6, 435 (1967). b) Jerry R. Martin and Alma W. Goldstein, Proc. 6th Int. Congress of Chemotherapy (Tokyo), II, 1112 (1970).

8. Jerry R. Martin and Thomas J. Perun, Biochemistry 7, 1728 (1968). Jerry R. Martin and Richard S. Egan, Biochemistry 9, 3439 (1970).

9. a) Thomas J. Perun and Richard S. Egan, Tetrahedron Letters, 387 (1969). b) T. J. Perun, R. S. Egan, P. H. Jones, J. R. Martin, L. A. Mitscher and B. J. Slater, Antimicrobial Agents and Chemotherapy, 116 (1969).

10. L. A. Mitscher and B. J. Slater, Tetrahedron Letters, 4505 (1969).

11. Reference 4b.

12. J. Dale, J. Chem. Soc., 93 (1963).

13. Reference 9.

14. M. Karplus, J. Chem. Phys. 30, 11 (1959).
M. Karplus, J. Amer. Chem. Soc. 85, 2870 (1963).

15. Richard S. Egan, Thesis, University of Illinois, 1971.

16. T. J. Perun, J. Org. Chem. 32, 2324 (1967).

17. a) J. B. Stothers, Carbon-13 NMR Spectroscopy, Academic Press, New York, 1972.
b) George C. Levy and Gordon L. Nelson, Carbon-13 Nuclear Magnetic Resonance for Organic Chemists, Wiley-Interscience, New York, 1972.

18. J. B. Grutzner, M. Jautelat, J. B. Dence, R. A. Smith and J. D. Roberts, J. Amer. Chem. Soc. 92, 7107 (1970).

19. Reference 17a.

20. D. M. Grant and B. V. Cheney, J. Amer. Chem. Soc. 89, 5315 (1967).

21. This reference is inoperative.

22. Reference 17a, p. 494.

23. Reference 17a, p. 287-288.

24. Reference 17a for a review.

25. F. A. L. Anet, C. H. Bradley and G. W. Buchanan, J. Amer. Chem. Soc. 93, 258 (1971).

26. Douglas E. Dorman and John D. Roberts, J. Amer. Chem. Soc. 92, 1355 (1970).

27. Douglas E. Dorman and John D. Roberts, J. Amer. Chem. Soc. 93, 4463 (1971).

28. Reference 17a, p. 289.

29. J. J. Burke and P. C. Lauterbur, J. Amer. Chem. Soc. 86, 1870 (1964).

30. F. A. L. Anet, A. K. Cheng and J. J. Wagner, J. Amer. Chem. Soc. 94, 9250 (1972).

31. Paul V. Demarco, The Journal of Antibiotics XXII, 327 (1969).

32. P. Kurath and R. S. Egan, Helv. Chim. Actg 54, 523 (1971).

33. Alan K. Mallams, Robert S. Jaret and Hans Reimann, J. Amer. Chem. Soc. 91, 7506 (1969).

34. R. S. Egan, private communication.

35. Ref. 17a, Chapter 8.

36. a) Reference 17a. b) Reference 26. c) Manfred Christl, Hans J. Reich and John D. Roberts 93, 3463 (1971).

37. Douglas E. Dorman, S. J. Angyal and John D. Roberts, J. Amer. Chem. Soc. 92, 1351 (1970).

38. John D. Roberts, Frank J. Weigert, Jacqueline I. Kroschwitz and Hans J. Reich, J. Amer. Chem. Soc. 92, 1338 (1970).

39. Reference 9.

40. Reference 17a, p.282.

41. Reference 17a, p.289.

42. J. R. Martin, T. J. Perun and R. S. Egan, Tetrahedron 28, 2973 (1972).

43. Robert L. Lichter and John D. Roberts, J. Phys. Chem. 74, 912 (1970).

44. a) Reference 17a, p. 287. b) Leroy F. Johnson and William C. Jankowski, Carbon-13 NMR Spectra, Wiley-Interscience, New York, 1972.

45. Reference 10.

46. a) A. A. Bothner-By and R. H. Cox, J. Phys. Chem. 73, 1830 (1969). J. A. Schwarcz and A. S. Perlin, Can. J. Chem. 50, 3667 (1972). c) R. Wasylishen and T. Schaefer, Can. J. Chem. 50, 2710 (1972).

47. A. C. Lunn and J. K. Senior, J. Phys. Chem. 33, 1027 (1929).

48. E.L. Muetterties, Record of Chemical Progress 31, 51 (1970).

49. G. Polya, Acta Math. 68, 145 (1937).

50. B. A. Kennedy, D. A. McQuarrie and C. H. Brubaker, Jr., Inorg. Chem. 3, 265 (1964).

51. Kurt Mislow, Introduction to Stereochemistry, W. A. Benjamin, Inc., New York, New York, 1966, p. 88.
J. K. Senior, Chem. Ber. 60B, 73 (1927).

52. J. E. Leonard, thesis submitted May 24, 1971, California Institute of Technology.

53. R. S. Cahn, Sir Christopher Ingold and V. Prelog, Angew.Chem. Int. Ed. Eng. 5, 385 (1966).

54. V. Prelog and H. Gerlach, Helv. Chim. Acta, 2288 (1964).

55. C. Berge, Principles of Combinatorics, Academic Press, New York, 1971.

56. F. Albert Cotton, Chemical Applications of Group Theory, Wiley-Interscience, New York, 1963.

57. Marshall Hall, The Theory of Groups, Macmillan, New York, 1959, p.280.

58. J. B. Hendrickson, J. Amer. Chem. Soc. 93, 6854 (1971).

59. Reference 57, p. 9.

60. E. M. Patterson, Topology, Oliver and Boyd, New York, 1963.

61. E. J. Corey and George A. Petersson, J. Amer. Chem. Soc. 94, 460 (1972).

62. David A. Evans, private communication.

63. G. Stork, et al., J. Amer. Chem. Soc. 81, 5516 (1959); 85, 3419 (1963).

64. Reference 57, p.88-90.

65. a) Reference 57, p.81. b) Frank Harary, Graph Theory, Addison-Wesley, Menlo Park, California, 1969.

66. Joseph J. Rotman, The Theory of Groups, An Introduction, Allyn and Bacon, Boston, 1965, p. 134.

67. Robert E. Ireland, Organic Synthesis, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

68. Reference 51, p. 91.

69. Reference 54.

70. a) V. Prelog and G. Helmchen, Helv. Chim. Acta 55, 2581 (1972). b) G. Helmchen and V. Prelog, Helv. Chim. Acta 55, 2599, 2612 (1972).

71. a) A. V. Shubnikov, N. V. Belov et al., Colored Symmetry, MacMillan, New York, 1964. b) V. L. Indenborn, Soviet Physics-Cryst. 4, 578 (1959). c) E. I. Galyarskii and A. M. Zamorzaev, Soviet Physics-Cryst. 8, 68 (1963). d) A. M. Zamorzaev, Soviet Physics-Cryst. 12, 717 (1965).

72. Reference 71.

73. Reference 56, p. 125.

74. V. L. Indenbom , N. V. Belov, and N. N. Neronova, Soviet Physics-Cryst. 5, 477 (1960).

75. Reference 57, Chapter 15.

76. C. M. Woodman, Molecular Physics 19, 753 (1970).

77. M. J. Goldstein and Roald Hoffmann, J. Amer. Chem. Soc. 93, 6193 (1972).

78. This term is used in Reference 77, p. 6201.

79. Reference 77, p. 6201.

80. Reference 77, p. 6196.

81. R. B. Woodward and R. Hoffmann, The Conservation of Orbital Symmetry, Verlag Chemie, Academic, 1970.

82. a) Reference 76. b) H. C. Longuet-Higgins, Molecular Physics 6, 445 (1963). c) J.-M. F. Gilles and J. Phillipot, International Journal of Quantum Chemistry VI, 225 (1972).

83. Reference 66, p.134-135.

84. S. Eilenberg and S. MacLane, Trans. Am. Math. Soc. 58, 231 (1945).

85. a) Saunders MacLane, Categories for the Working Mathematician, GTM5, Springer-Verlag, New York, 1971.
b) Bodo Pareigis, Categories and Functors, Academic Press, New York, 1970.
c) Horst Schubert, Categories, Springer-Verlag, New York, 1972.
d) Horst Herrlich and George E. Strecker, Category Theory, An Introduction, Allyn and Bacon, Boston, 1973.

86. a) Philip J. Higgins, Notes on Categories and Groupoids, VNR, New York, 1971, p. 2.
b) M. Barr in Reports of the Midwest Category Seminar, III, ed. by Saunders MacLane, Lecture Notes in Mathematics 106, Springer-Verlag, New York, 1969.

87. a) Robert Rosen, Bull. Math. Biophys. 20, 245 (1958).
b) Robert Rosen, Bull. Math. Biophys. 20, 317 (1958).
c) References 88-90.

88. Robert Rosen, Bull. Math. Biophys. 33, 303 (1971).

89. Robert Rosen, Bull. Math. Biophys. 26, 103 (1964).

90. Robert Rosen, Bull. Math. Biophys. 28, 141 (1966).

91. Yehoshafat Give'on in Proceedings of the Conference on Categorical Algebra, La Jolla, 1965, Springer-Verlag, New York, 1966. p. 317.

92. Reference 85a, p. 170.

93. Otomar Hajek in Topological Dynamics, An International Symposium ed. by Joseph Auslander and Walter H. Gottschalk, Benjamin, New York, 1968, p. 243.

94. Robert Ellis, Lectures on Topological Dynamics, Benjamin, New York, 1969.

95. Robert Hermann, Lie Groups for Physicists, Benjamin, New York, 1966.

96. Harry J. Lipkin, Lie Groups for Pedestrians, North Holland, Amsterdam, 1966, p. 23.

97. E. Inonu and E. P. Wigner, Proc. Nat. Acad. Sci. 39, 510 (1953).

98. Eugene J. Saletan, J. Math. Phys. 2, 1 (1961).

99. C. W. Conatser, J. Math. Phys. 13, 196 (1972).

100. P. J. Hilton and U. Stammbach, A Course in Homological Algebra GTM4, Springer-Verlag, New York, 1970.

101. B. Hartley and T. O. Hawkes, Rings, Modules and Linear Algebra, Chapman and Hall, London, 1970.

102. Reference 100, p. 162.

103. Reference 81.

104. Reference 85a, p. 66.

105. Devens Gust and Kurt Mislow, J. Amer. Chem. Soc. 95, 1535 (1973).

106. S. S. Eaton and G. R. Eaton, J. Amer. Chem. Soc. 95, 1825 (1973).

107. J. C. Bailar, Jr., J. Inorg. Nucl. Chem. 8, 165 (1958).

108. P. Ray and N. K. Dutt, J. Indian Chem. Soc. 20, 81 (1943).

109. Reference 86a.

110. W. G. Klemperer, J. Amer. Chem. Soc. 95, 380, 2105 (1973), and earlier papers cited therein.

111. Adelbert Kerber, "Representations of Permutation Groups I", Lecture Notes in Mathematics 240, Springer-Verlag, New York, 1971.

112. Reference 57, Chapter 15.

113. James W. McIver Jr. and Richard E. Stanton, J. Amer. Chem. Soc. 94, 8618 (1972).

114. Reference 85a, p. 103.