# SYNTHETIC AND BIOSYNTHETIC STUDIES ON VITAMIN B$_{12}$[1]

ALAN R. BATTERSBY[2]

University Chemical Laboratory, University of Cambridge, Cambridge CB2 1EW, England

ABSTRACT.—Vitamin B$_{12}$ is the anti-pernicious anemia vitamin. It functions as the main part of a coenzyme for a range of remarkable rearrangement reactions. The macrocycle for vitamin B$_{12}$ is smaller by one carbon than that for chlorophyll and protoheme. Research is described involving numerous skills at the microscale level. In addition, synthesis of new pigments that have been encountered has been undertaken. A large part of the biosynthetic pathway to vitamin B$_{12}$ has been revealed.

This second Vario Tyler lecture covers our research on the biosynthesis of vitamin B$_{12}$, the anti-pernicious anemia vitamin that has quite a complex structure and also a beautiful deep color. Its structure, shown in Scheme 1, is related to those such as heme and chlorophyll, which were considered in the first paper. Here again the macrocycle is composed of four nitrogenous five-membered rings joined in three places by a one-carbon bridge. But, there is a major difference from the structure of heme in that vitamin B$_{12}$ has a direct link between rings A and D. A fascinating feature of the B$_{12}$ molecule is that it is an organo cobalt complex, and finally there are many methyl groups around the periphery that make the system highly chiral. The benzimidazole ligand, which is coordinated to the cobalt, is carried by the so-called nucleotide loop, and prior to the start of our work in 1968, it was known that this loop was added very late in the whole biosynthetic process. Moreover, all the amide groups around the periphery are also added late in the sequence (1). So the problem can be simplified by mentally removing those groups added late, which leads to the material called cobyrinic acid (Scheme 1). It
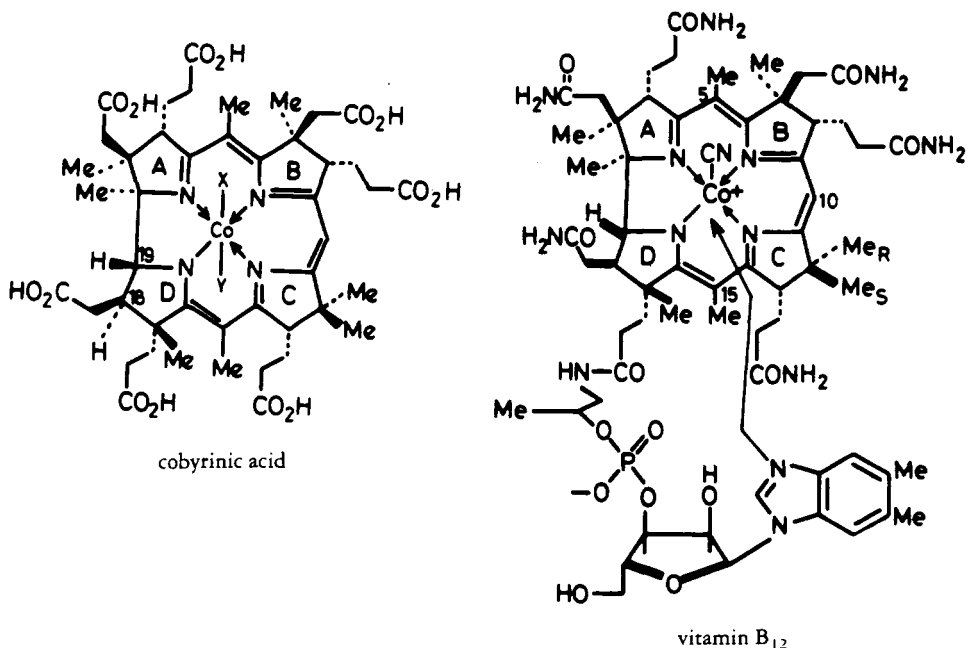

cobyrinic acid

vitamin B$_{12}$

SCHEME 1.

was already known (1) by the mid-1960s that cobyrinic acid is a biosynthetic precursor of vitamin B_{12}, and so our problem resolves itself into asking how cobyrinic acid is biosynthesized in a living system.

Scheme 2 shows what was already known by the end of the 1960s as a result of the admirable efforts of several research groups (1). Given that knowledge, the question was asked whether cobyrinic acid might be derived from uro'gen-III, which, as described in the first lecture, is true for chlorophyll on the one hand and protoheme on the other. This possibility was studied in several laboratories (1), and the specific labelling

methyl groups from

cobyrinic acid

vitamin B_{12}
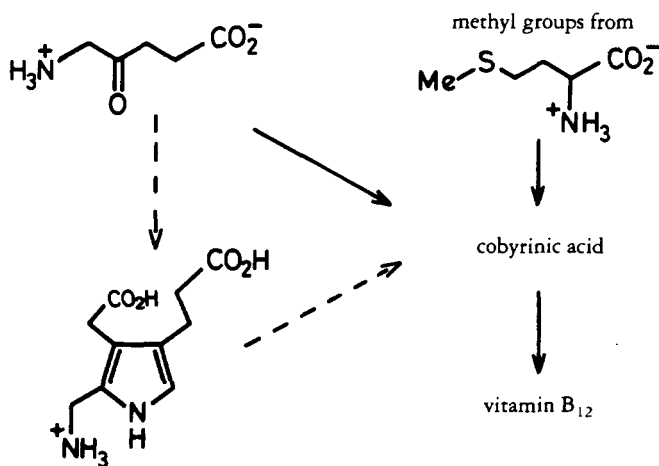
SCHEME 2.

experiment using ^{14}C that was carried out in Cambridge (2,3) is shown in Scheme 3; the labelled uro'gen-III was produced synthetically. A broken-cell (cell-free) enzyme system prepared from the B_{12} producer *Propionibacterium shermanii* was used for the incorporation work. As can be seen from Scheme 3, the cobyrinic acid formed carried a single ^{14}C label demonstrating clearly a precursor-product relationship.

It is instructive to look at the variety of chemical transformations that have to be done to change uro'gen-III into cobyrinic acid. Many methyl groups have to be added, the acetate residue at C-12 has to be decarboxylated to give a methyl group, a direct link has to be joined between rings A and D, and finally cobalt has to be inserted at some stage. It is conceivable that the first thing to happen is the decarboxylation step, and

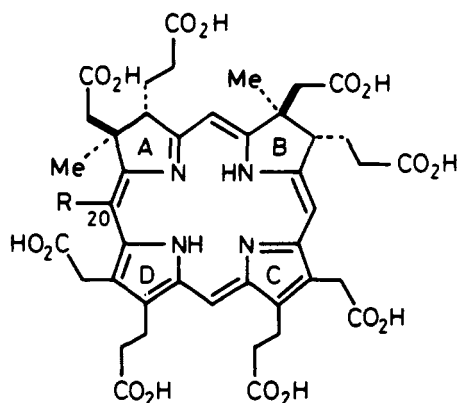uro'gen-III   cobyrinic acid
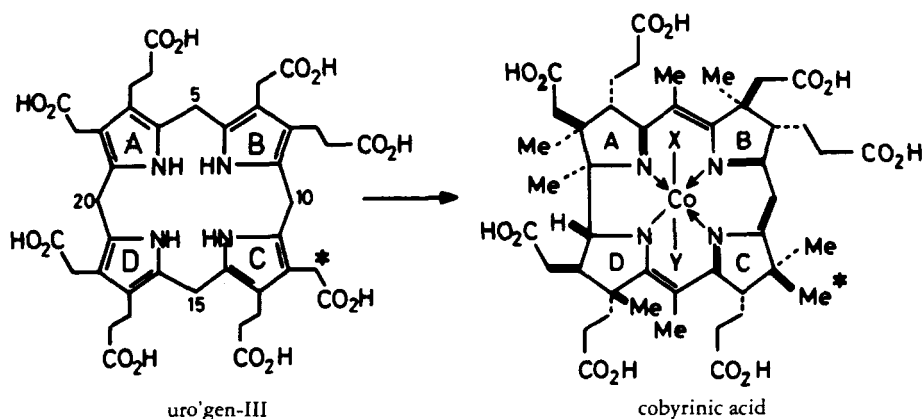
SCHEME 3.

this possibility can be tested by synthesizing the 12-decarboxy material to check whether it acts as a precursor of cobyrinic acid in the cell-free enzyme system (Scheme 4). In fact, it was not a precursor of cobyrinic acid (2,3). So decarboxylation is not the first step on the biosynthetic pathway after generation of uro'gen-III.
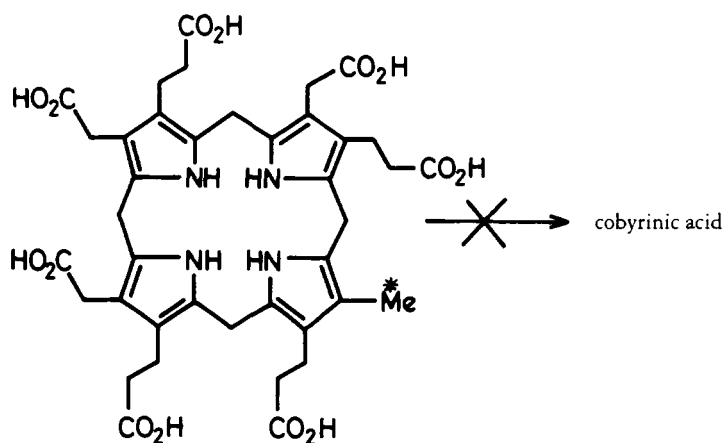
SCHEME 4.

Then, in collaboration with Dr. Bykhovsky in Moscow, many new pigments were isolated from *P. shermanii*; we will concentrate on two of them in this lecture. They were lovely violet pigments with a striking fluorescence in ultraviolet light. Their composition was established by field desorption mass spectrometry, and the first pigment had the composition corresponding to uro'gen-III plus two methyl groups minus two hydrogens. The second pigment made us even more excited because its composition corresponded to uro'gen-III plus three methyl groups minus two hydrogens. We (4–7) and others (1) worked out the structures of both these materials, but that work is outside the scope of this lecture. Suffice to say that the structures that were revealed are shown in Scheme 5, which also gives their names. There is no question that the isolation of these two substances and their structure determination completely transformed the research on B$_{12}$ biosynthesis. Indeed, suitable labelling experiments (1, 4–7) demonstrated that both these substances are converted into cobyrinic acid by the enzymes present in the

R = H sirohydrochlorin
R = Me trimethylisobacteriochlorin

SCHEME 5.

broken cell enzyme system from *P. shermanii*. Thus, it was clear that the next biosynthetic steps beyond uro'gen-III on the pathway to cobyrinic acid involve *C*-methylations.

There was, however, an inconsistency. Uro'gen-III contains eight double bonds, whereas sirohydrochlorin and the trimethylisobacteriochlorin of Scheme 5 have nine. Yet *C*-methylation does not affect the oxidation level of a substance. So it seemed highly probable that the true biosynthetic intermediates are not the materials in Scheme 5 but dihydro derivatives of these substances. This important problem of oxidation level was studied by carrying out the enzymic methylation of uro'gen-III in a glove box with oxygen levels below 5 ppm with the enzyme completely degassed. After mild basic esterification (Meerwein reagent), a yellow product was isolated as its octamethyl ester. This on oxidation with iodine gave sirohydrochlorin octamethyl ester (Scheme 5). Mass spectroscopy confirmed that the yellow product was a dihydro derivative of sirohydrochlorin (8).

The structure proof for this dihydro derivative had to be carried out in an interesting, somewhat indirect way because so little material was available. However, the evidence was unambiguous that the dihydro macrocycle shown in Scheme 6 correctly rep-
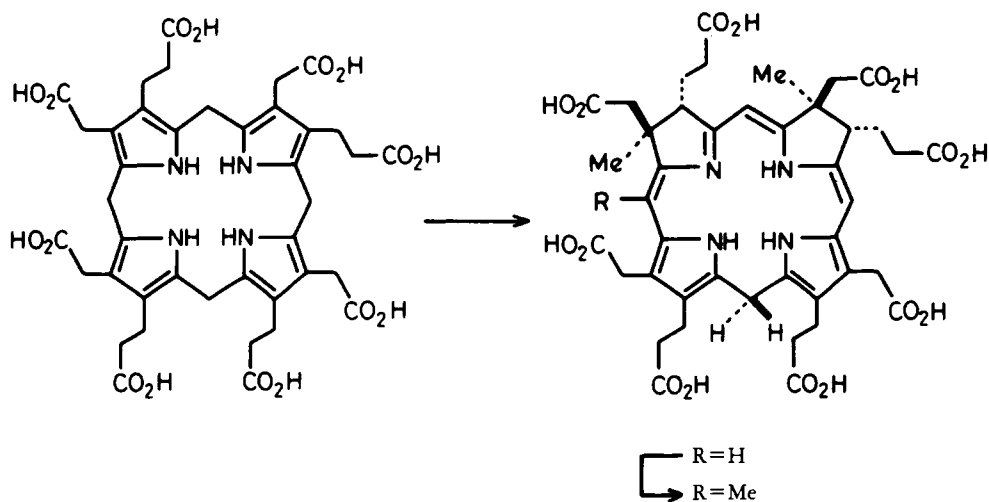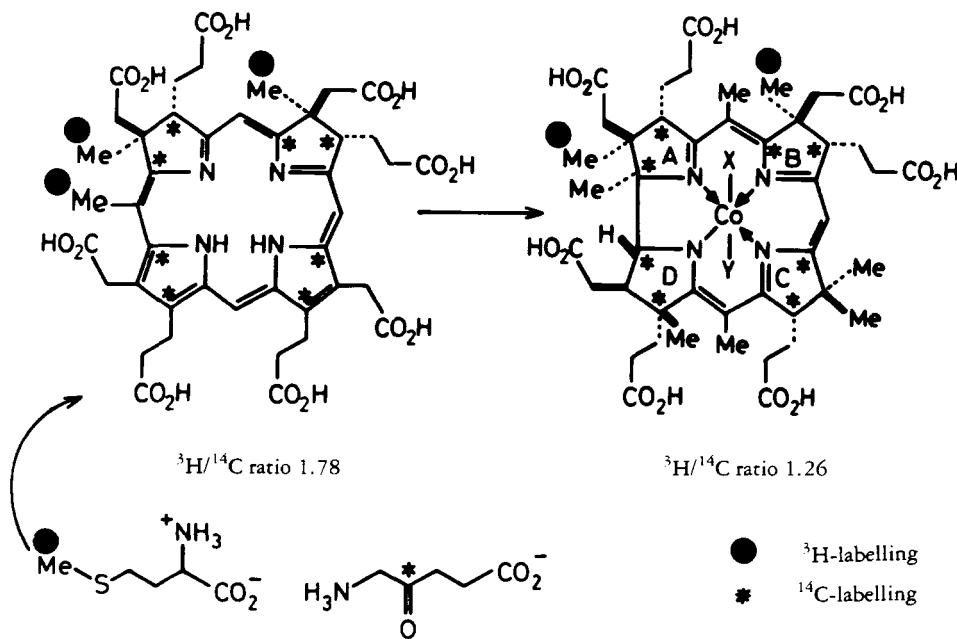

SCHEME 6.

resents the biosynthetic intermediate that is known as precorrin-2. Again, the appropriate (and experimentally highly demanding) labelling experiments were carried out to prove that precorrin-2 is a specific precursor of cobyrinic acid (8). Moreover, it is essentially certain that the trimethylated intermediate is produced and further transformed at the dihydro level of oxidation; this material (Scheme 6) is called precorrin-3.

At this stage one can ask whatever is Nature doing placing a methyl group at C-20. Perhaps, we thought, the methyl group rearranges to C-1 in the final cobyrinic acid, and that possibility we wanted to examine. Scheme 7 shows the development of the method that involved generation of sirohydrochlorin biosynthetically from methionine labelled with tritium in the methyl group and from aminolaevulinic acid labelled with ^{14}C at C-5. These two precursors lead to the illustrated labelling pattern with 8 starred labels and 2 dotted labels. This product can now be transformed enzymically into cobyrinic acid, and one would expect all the labels to appear in the final product. So the ratio of tritium to ^{14}C at the start should be maintained in the product, and, indeed,

$^3H/^{14}C$ ratio 1.26 $^3H/^{14}C$ ratio 1. 18

● ^{3}H-labelling
✳ ^{14}C-labelling
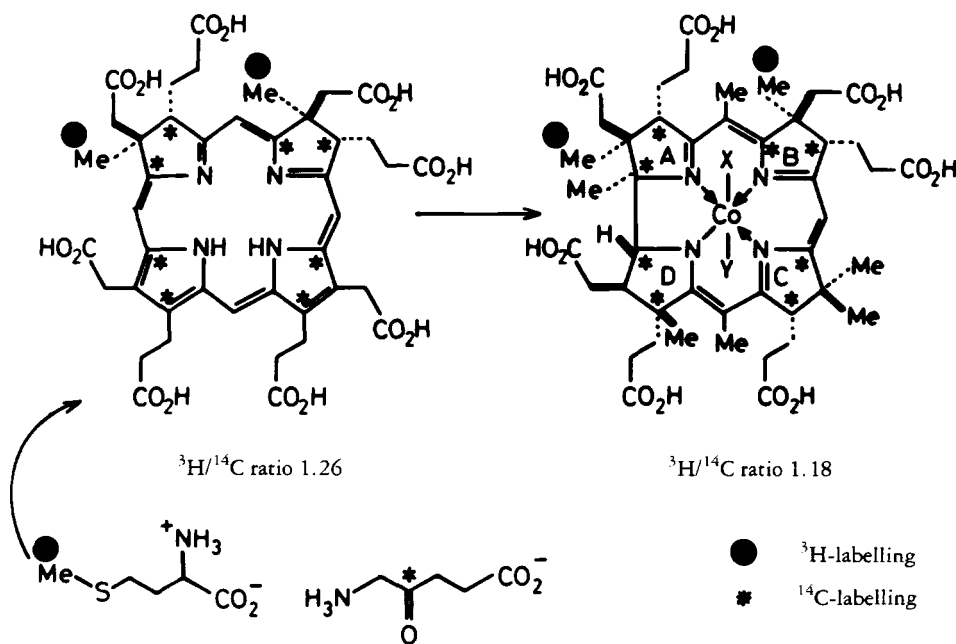
SCHEME 7.

within experimental error, that is the result that was found (9). So this method of test-
ing for retention (in this case) or loss of methyl groups by measurement of tritium
against an internal standard of ^{14}C works well. Now the real case can be studied
(Scheme 8). The trimethylated system was biosynthesized in labelled form as for the
previous case; now there are 3 dotted methyl groups and 8 starred carbon labels. When
this product is transformed into cobyrinic acid, then, if all the methyl groups survive,

$^3H/^{14}C$ ratio 1.78 $^3H/^{14}C$ ratio 1.26

● ^{3}H-labelling
✳ ^{14}C-labelling

SCHEME 8.

there will be no change in the tritium-to-carbon ratio. If, on the other hand, that at C-20 is lost at some stage, the ratio will change. Scheme 8 shows that almost exactly one third of the tritium label was lost, and so it was proved that the C-20 methyl group is extruded in the overall biosynthetic process (9).

Naturally, the next question is: In what form is the C-20 methyl group lost? The approach was to synthesize uro'gen-III with a ^{14}C label at C-20 and to transform it biosynthetically first into the dimethylated system and forward to the trimethylated system (precorrin-3) having a high specific activity. Now this product can be transformed enzymically into cobyrinic acid checking for the fate of the ^{14}C label (Scheme 9). The labelled extruded product was found to be HOAc, which was degraded by the Schmidt reaction to give CO_2 and methylamine. This showed that all the ^{14}C label was in the carboxyl group of the original HOAc. The complementary experiment in which the C-20 methyl group was labelled was also carried out; now, all the label was present in the methyl group of the HOAc. All this work rigorously established that C-20 and its attached methyl group of precorrin-3 (Scheme 9) are extruded from the molecule at some stage as HOAc such that C-20 itself becomes the carboxyl group of HOAc (10). The same conclusion was reached independently in Zürich (11). This knowledge is important in understanding the process that produces the direct link between rings A and D of cobyrinic acid.
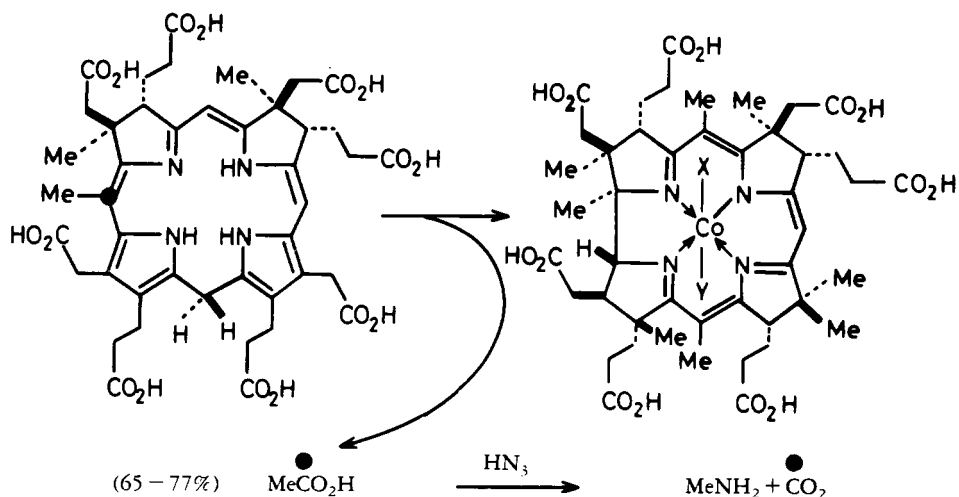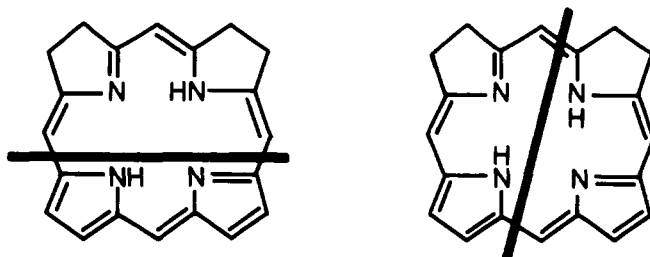

SCHEME 9.

The intermediates precorrin-2 and precorrin-3 are highly important materials for experiments aimed at exploring the biosynthetic steps that follow the trimethylated stage (precorrin-3). Yet, both are difficult to obtain in quantity, especially precorrin-3, and so we decided to launch a synthetic attack to clear this bottleneck in the supplies. The actual targets were to be the aromatized forms of the biosynthetic intermediates, that is, sirohydrochlorin and the trimethylisobacteriochlorin (Scheme 5). The decision to undertake this task was not an easy one because the synthesis is obviously going to be demanding. Furthermore, the syntheses are being carried out because the materials are needed to solve important problems. A successful synthesis is not the end; it opens the way to a new beginning on the biosynthetic problem posed by vitamin B_{12}.

The two synthetic targets above are both based on the isobacteriochlorin macrocycle, and at the outset of our studies, no method was available to construct even simple members of this class. Then a number of valuable methods were devised (12-14) to ob-

tain isobacteriochlorins carrying methyl groups around the periphery of the macrocycle, but these approaches ran into difficulties when the natural acetate and propionate groups were present. So a different approach was undertaken.

The first synthetic methods for the isobacteriochlorins were based on the retro-synthetic cleavage shown on the left of Scheme 10. The different approach used the retro-

SCHEME 10.

cleavage on the right of Scheme 10. Gradually, the final plan emerged as a result of many exploratory experiments, and it is illustrated in Scheme 11. The aim was to synthesize an open-chain precursor (what we call the seco system) in ways to be discussed later, and inspection shows that the conjugated tautomer is an 18π-system (9 double bonds). By the Woodward-Hoffman rules, such a system should ring-close photochemically in an antarafacial fashion. When the illustrated seco system was synthesized and irradiated with light of the appropriate wavelength, the chemistry worked superbly, and the cyclized isobacteriochlorin was produced (15). Notice that even in this model
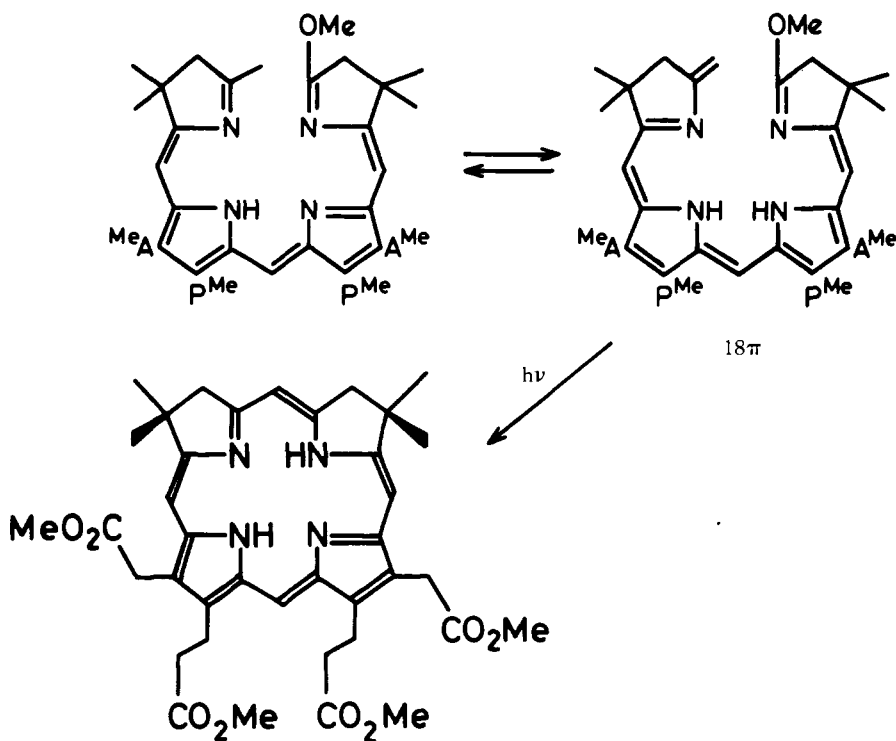

found [M]$^+$ 686.3333
C$_{38}$H$_{46}$N$_4$O$_8$ 686.3315

SCHEME 11.

series, acetic and propionic side chains are present, and the photochemical approach was mild enough to allow one to deal with these reactive systems. The way was open to go for the synthesis of the natural pigments.

Our basic strategy for assembly of the building blocks is illustrated in Scheme 12 with the imide at the top of the scheme being converted into the two isomeric monothioimides. Then, it was envisaged that these might be connected to the pyrrolic building block destined to become rings C and D of the final macrocycle by a type of thio-Wittig reaction or some related process.
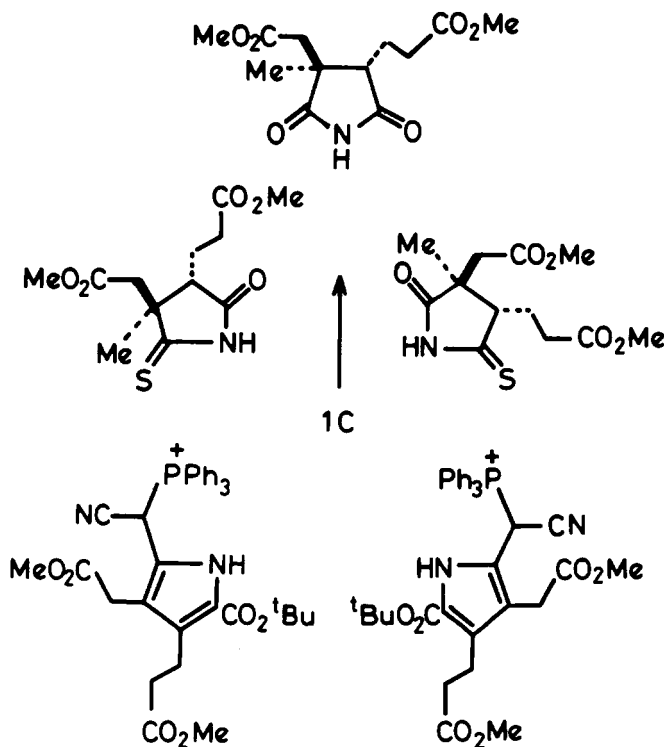
SCHEME 12.

We have developed several ways to synthesize the key imide in Scheme 12 with high enantioselectivity; one route is shown in Scheme 13, which uses the bis-anion chemistry developed by Seebach for the steps that set up the quaternary center. This sequence has been run to provide 10 g of the final imide in crystalline and optically pure form.[3] The two monothioimides can now be produced as in Scheme 14, and the structures of both products were rigorously confirmed by X-ray analysis. For the crucial C-C bond forming step, it was found that the Wittig reagent of necessity had to carry the illustrated nitrile group (see Scheme 15). Then the coupling step proceeded very smoothly to generate the eastern building block. However, the unwanted nitrile had now to be removed, and this proved to be a real stumbling block. However, success came by reducing the nitrile to an aminomethyl group which was eliminated by a reverse Mannich reaction as shown in Scheme 16. This process was only effective, however, when the eliminated $CH_2 = N^+H_2$ fragment was removed from the reaction center by a suitable trapping reagent (16).

[3]A.R. Battersby and M.K. Ellis, unpublished work.
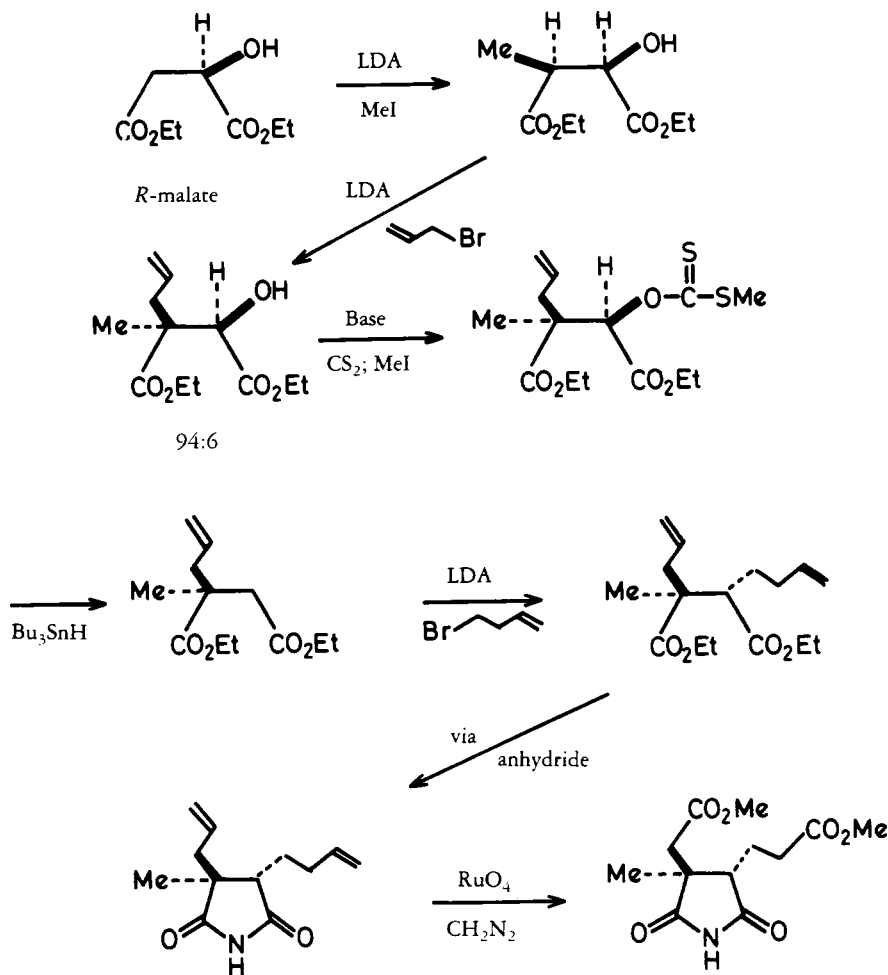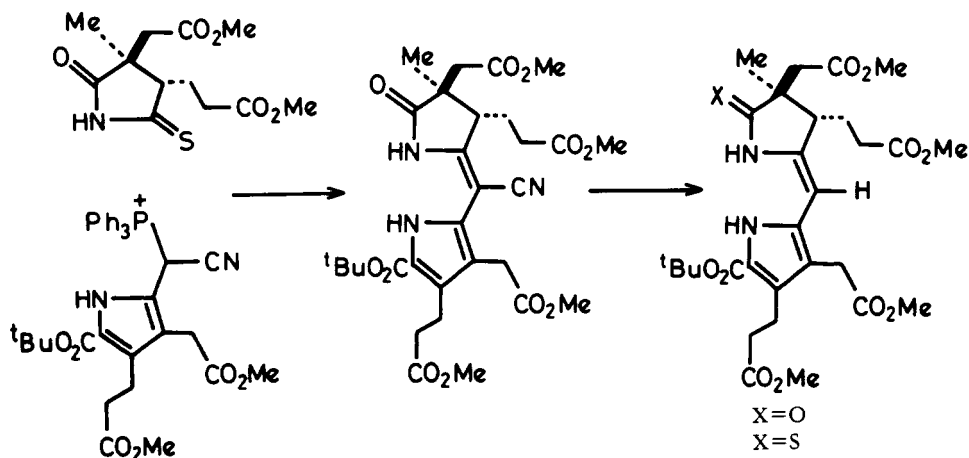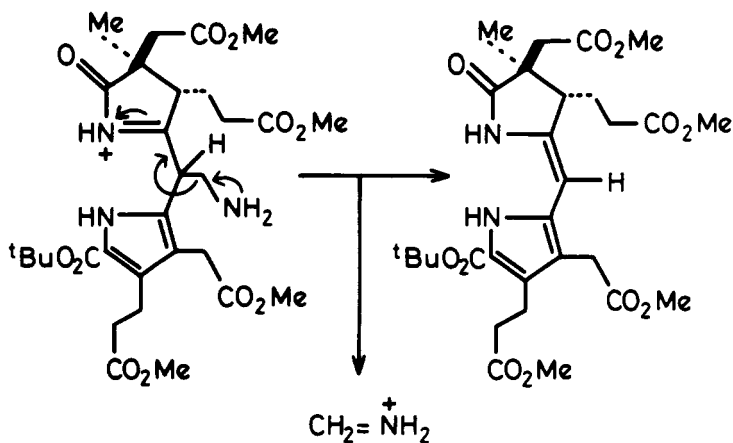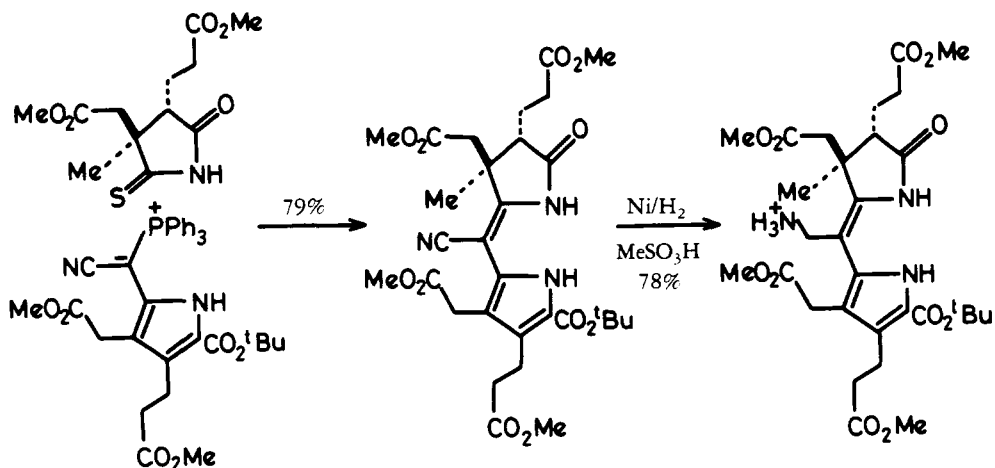
SCHEME 13.
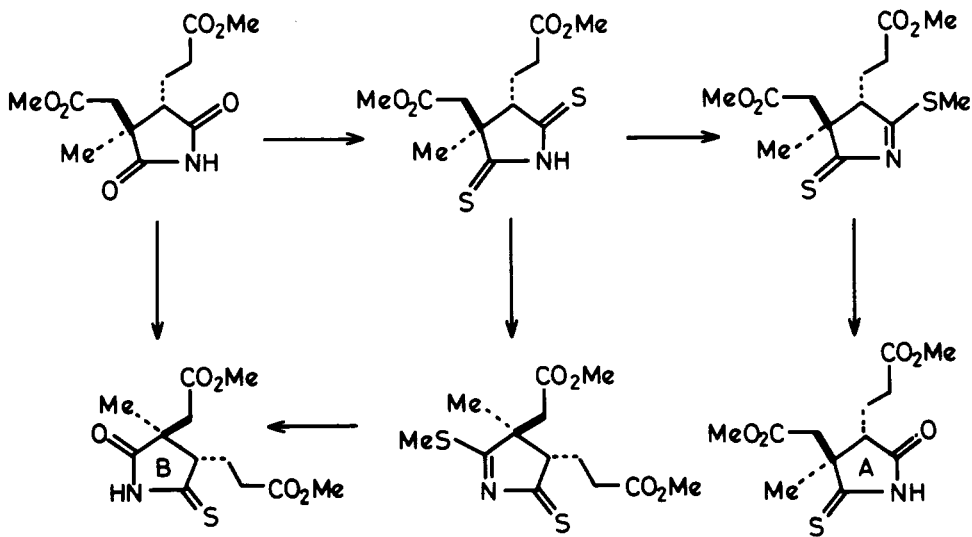
SCHEME 14.

SCHEME 15.

At this point, it seemed that the western building block ought to be available by following steps analogous to those outlined above for the eastern block. Indeed, this was true for all the steps up to the formation of the aminomethyl intermediate (Scheme 17). But, as shown in Scheme 18, attempted reverse-Mannich chemistry led to disaster; the illustrated stable lactam was the only product. The way through this blockage was to switch to the sulfonamide intermediate shown in Schemes 18 and 19, which allowed the equivalent of a reverse Mannich reaction to be achieved (16). It was a day for celebration!

It then remained to add a one-carbon unit to the eastern block and another one-carbon residue to the western block. These steps were achieved as in Scheme 20 to allow final assembly of the seco system with all the required functionality in place and with the correct absolute stereochemistry (Scheme 21). Photochemical cyclization (Scheme 22) then produced sirohydrochlorin octamethyl ester identical in all respects to the natural ester (16); you will recall that sirohydrochlorin is the aromatized form of the biosynthetic intermediate precorrin-2, and the latter can be generated non-enzymatically from the former. It will be clear that this synthesis allows the introduction of single and multiple labels into the final product, an advantage that can be built upon in future biosynthetic research.
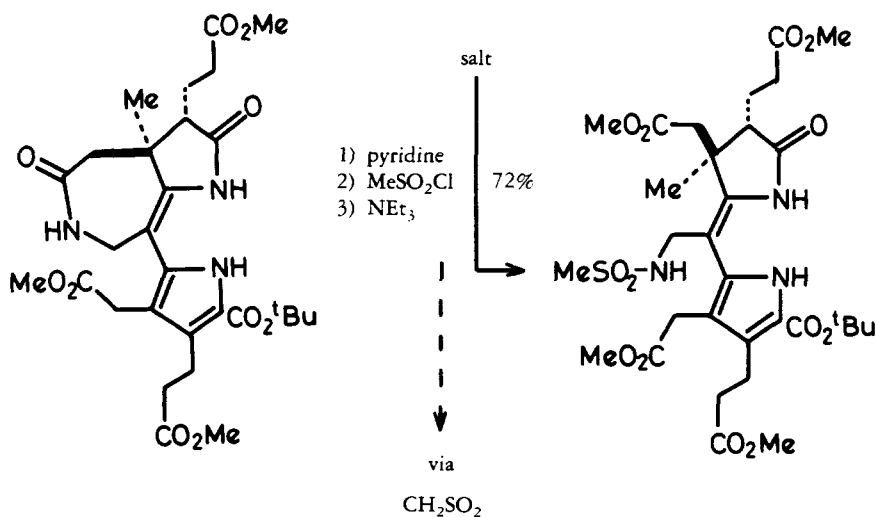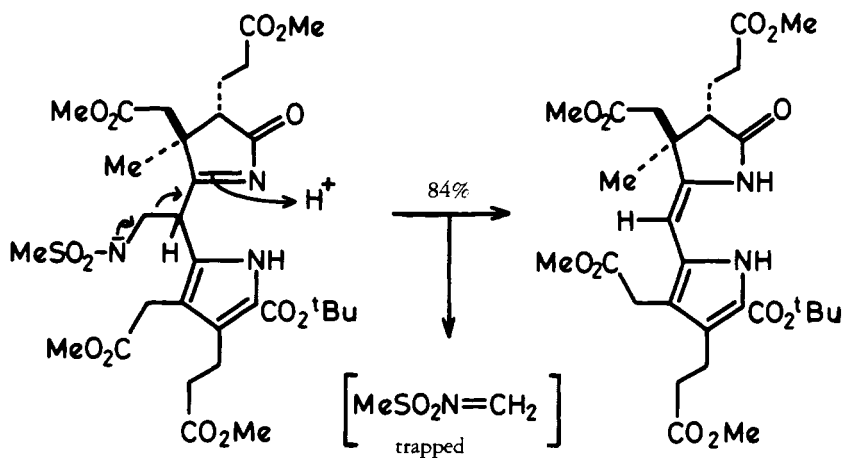
SCHEME 16.

SCHEME 17.

The synthesis just described is very well suited to its purpose. Only two starting materials are used, and both can be prepared on a large scale. The synthesis is entirely convergent with control of the absolute stereochemistry from the outset.

Finally, let us look at the biosynthetic pathway to vitamin B$_{12}$ beyond the trimethylated stage. We and others looked very carefully in organisms that produce B$_{12}$ for intermediates that carry four or five or six methyl groups, and none was found. So we changed tack completely to try to discover the order of methylation using pulse labelling. Scheme 23 illustrates the plan (17). The dimethylated system is to be treated with the cell-free enzyme system from a B$_{12}$-producing organism in the presence of a small amount of unlabelled S-adenosylmethionine. One can expect that during this incubation some intermediates will be produced beyond the trimethylated stage. Having generated this pool of intermediates, the biosynthesis is completed by adding a large amount of SAM carbon-13 labelled at its methyl group. This way of carrying out the experiment is called the Normal Pulse. By this approach, methyl groups added early in the biosynthetic process will have a low ^{13}C content, and the ones that are added late will be
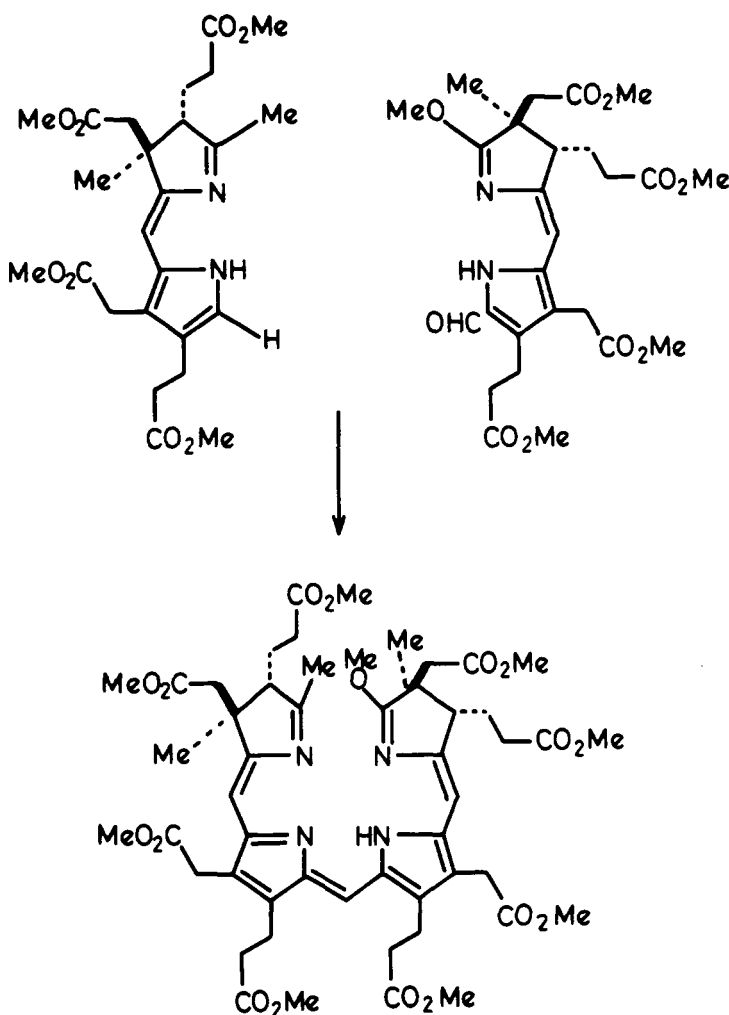

SCHEME 18.

SCHEME 19.
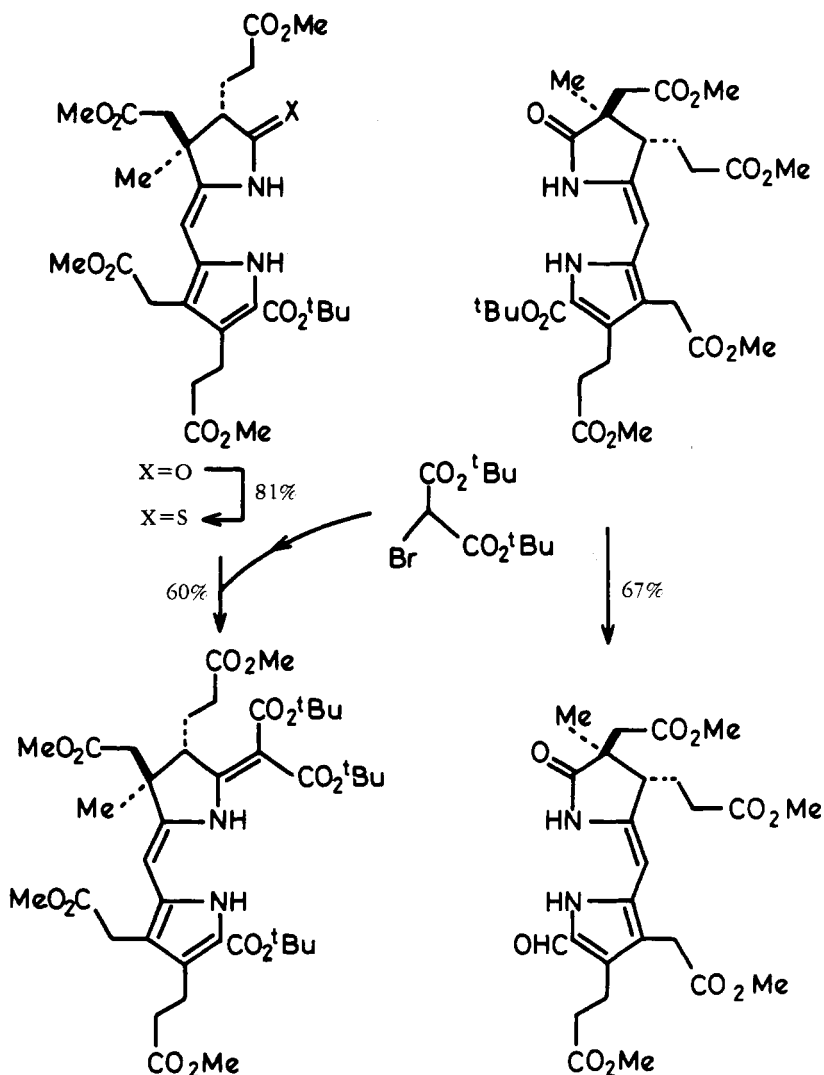
SCHEME 20.

SCHEME 21.

largely derived from the labelled SAM and should have a high ^{13}C content. The ^{13}C content in the various methyl groups can be measured by nmr, and obviously this demands assignments of all the signals for cobyrinic acid. In practice, all the nmr is carried out on the heptamethyl ester of cobyrinic acid, called cobester.

Figure 1 shows the comparison of the ^{13}C spectrum of cobester from the pulse labelling experiment (upper trace) with the spectrum of a sample of this material having all its SAM-derived methyl groups uniformly labelled (lower trace). It is strikingly clear that the signal from the *C*-methyl group at C-17 is much smaller than the others, showing that the fourth methyl is added at that site (17). This leads to knowledge of the structure of the tetramethylated intermediate. *C*-Methylation of precorrin-3 at C-17 generates the illustrated pyrrocorphin carrying an acetate residue at C-12 (Scheme 24). However, the possibility must be borne in mind that the C-12 acetate residue of precorrin-3 is decarboxylated before C-17 methylation. If that is so, then the pyrrocorphin generated by the C-17 methylation step will also carry a methyl group at C-12.

It should now be added that the pulse experiment can be carried out the other way around, that is, carry out the initial stage with a limited amount of ^{13}C-labelled SAM
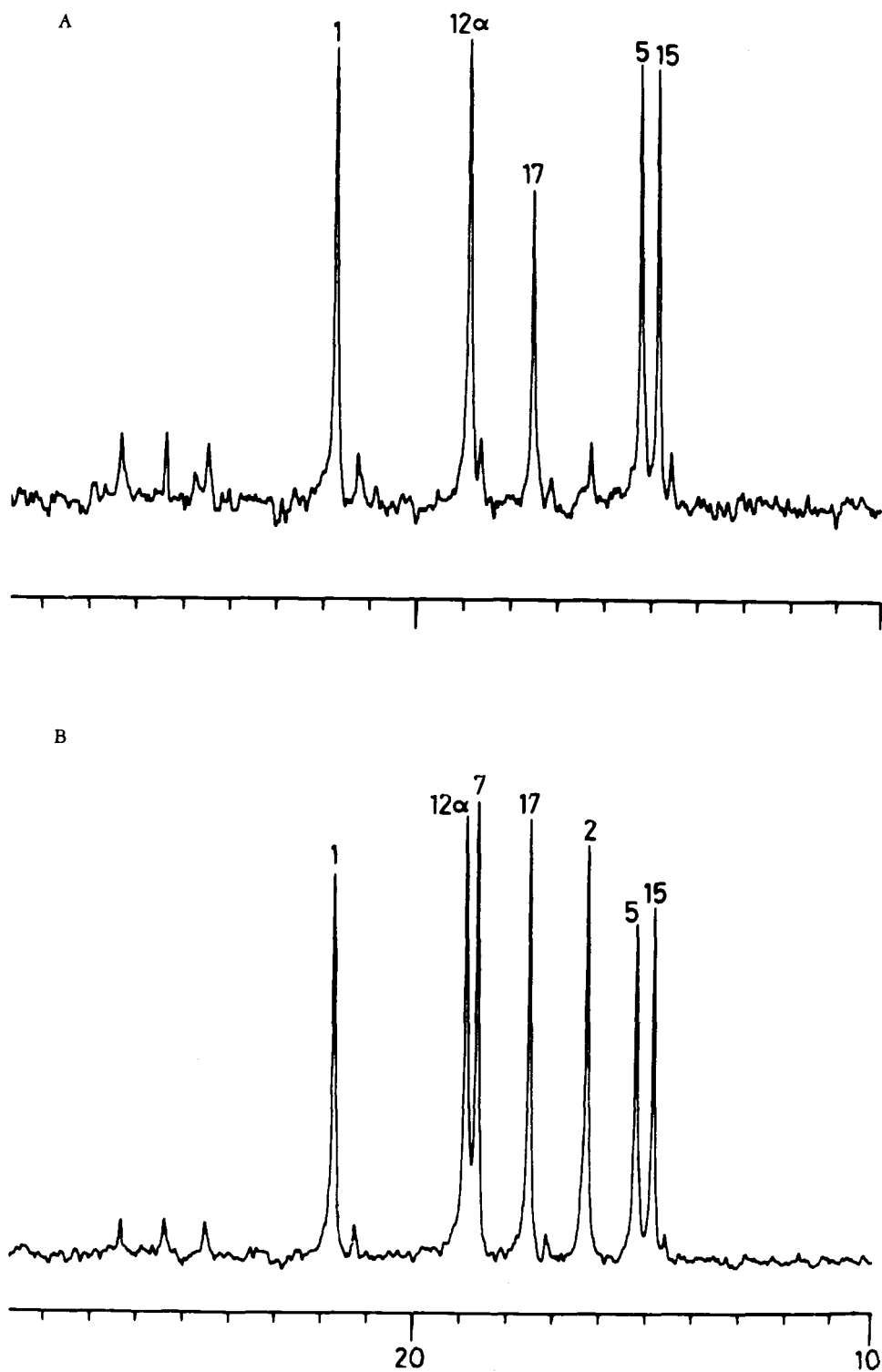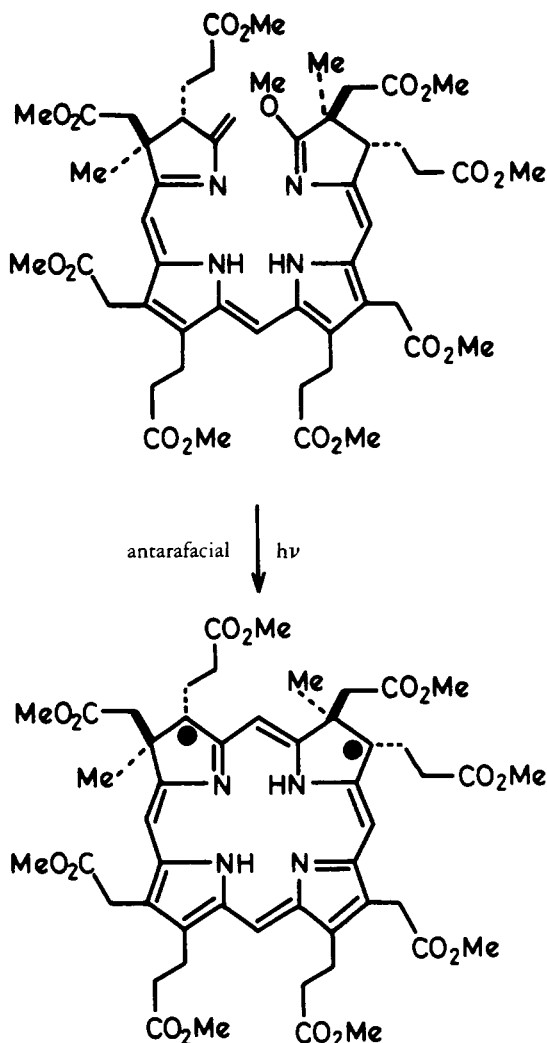
FIGURE 1.   Comparison of the ^{13}C-nmr spectrum of cobester from the Normal Pulse labelling experiment (A) with the spectrum of a sample of cobester with all SAM-derived methyl groups uniformly labelled (B).
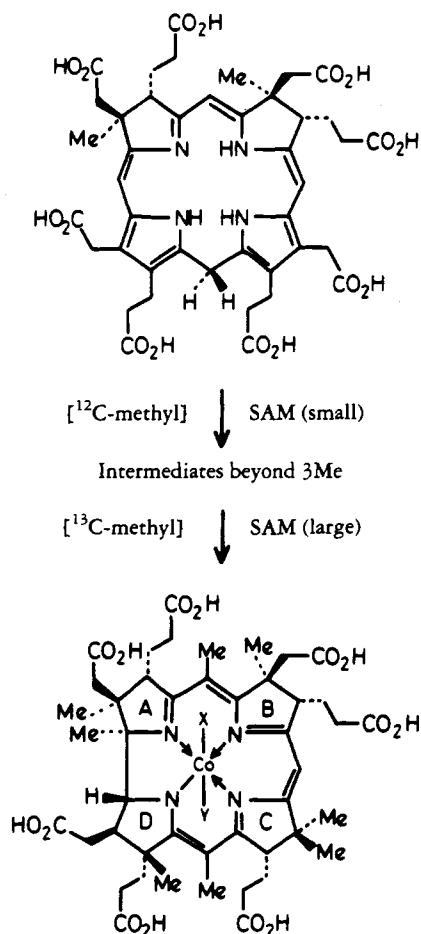
SCHEME 22.

and then complete the biosynthesis with unlabelled SAM in a large amount. This approach is called the Inverted Pulse. Now the methyl groups added early should be rich in ^{13}C, and the later ones should carry a lower level of ^{13}C. What was found for this Inverted Pulse was that the signal from the C-17 methyl was by far the biggest one (Figure 2), confirming (18,19) exactly the earlier conclusions that the fourth methyl group is inserted at position 17.

At this stage accurate determinations of signal intensities were available from both Normal and Inverted Pulse experiments. So difference values could be taken between the integrals for the ^{13}C levels at the various methyl groups from the two approaches to give the maximum sensitivity in discrimination between the labelling levels. The outcome was to extend our knowledge of the methylation sequence beyond the fourth methyl group, the order being 12 followed by 1, 15, and 5 in that order (18,19). This pulse labelling approach was adopted recently by others to conclude that beyond the already known fourth methyl group, the order runs 12, then 1, followed by 5 and 15 in unknown sequence (20).

The foregoing results allow some of the main features of the biosynthetic pathway for vitamin B_{12} to be picked out in order to help future research (18,19). The stage at

[^{12}C-methyl] $\quad\downarrow\quad$ SAM (small)

Intermediates beyond 3Me

[^{13}C-methyl] $\quad\downarrow\quad$ SAM (large)

SCHEME 23.

which cobalt is inserted is at present unknown, so the postulated intermediates in Scheme 25 are shown cobalt-free. It must be remembered, however, that on the biosynthetic pathway some or all of them may be cobalt-containing.

Scheme 25 shows the 12-methylpyrrocorphin being methylated at the fifth site (C-

R=CH$_2$CO$_2$H
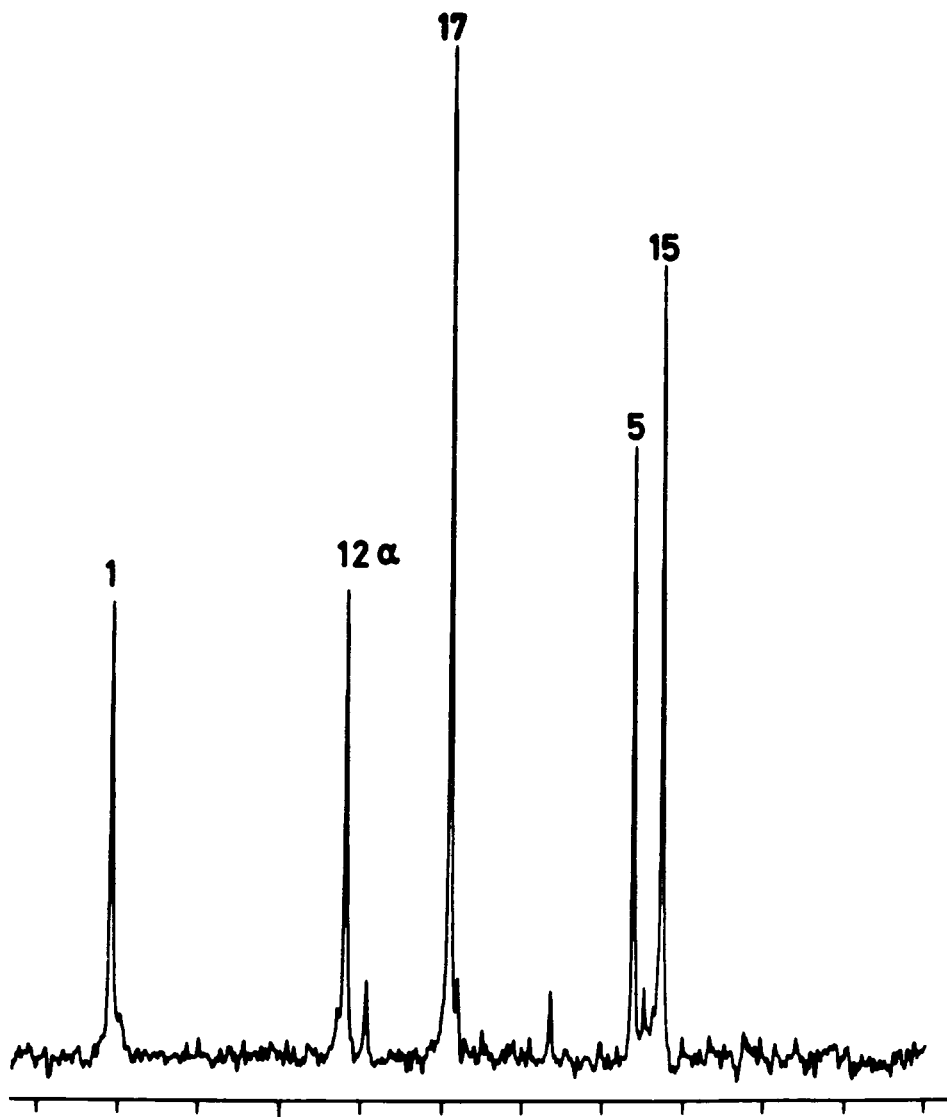R=Me

R=CH$_2$CO$_2$H
R=Me
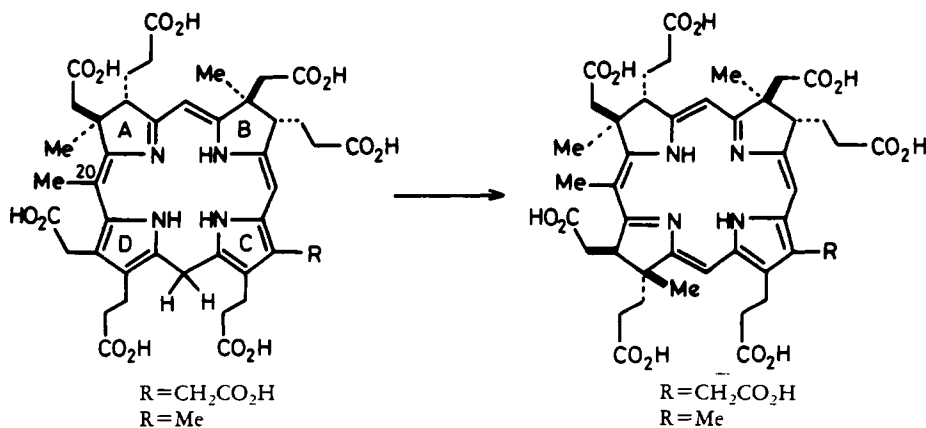
SCHEME 24.

FIGURE 2. ^{13}C-nmr spectrum of cobester after an Inverted Pulse experiment.

12) to produce a corphin that is pinpointed as an important intermediate on the biosynthetic pathway. Methylation at C-1 then sets up a cation at C-20, which is illustrated as being trapped by the acetate group at C-18; the cation could, however, simply be trapped by water, and there would be only slight changes to the mechanistic thinking. Either system so formed is well arranged for ring contraction, and the mechanism shown is based on Eschenmoser's important biomimetic model experiments (21). This step generates the corrin macrocycle, and, during the final steps, extrusion of the acetyl group in the form of HOAc finally leads to cobyrinic acid. You will remember that the proof that HOAc is indeed the form in which these two carbons is extruded was described earlier in this lecture.

The research on the biosynthesis of vitamin B_{12}, to which the Cambridge group has been contributing for about 20 years, has revealed not only the main shape of the biosynthetic pathway but also quite precise detail of several of the transformations (22). This work will continue until we know the entire pathway, step by step, which living
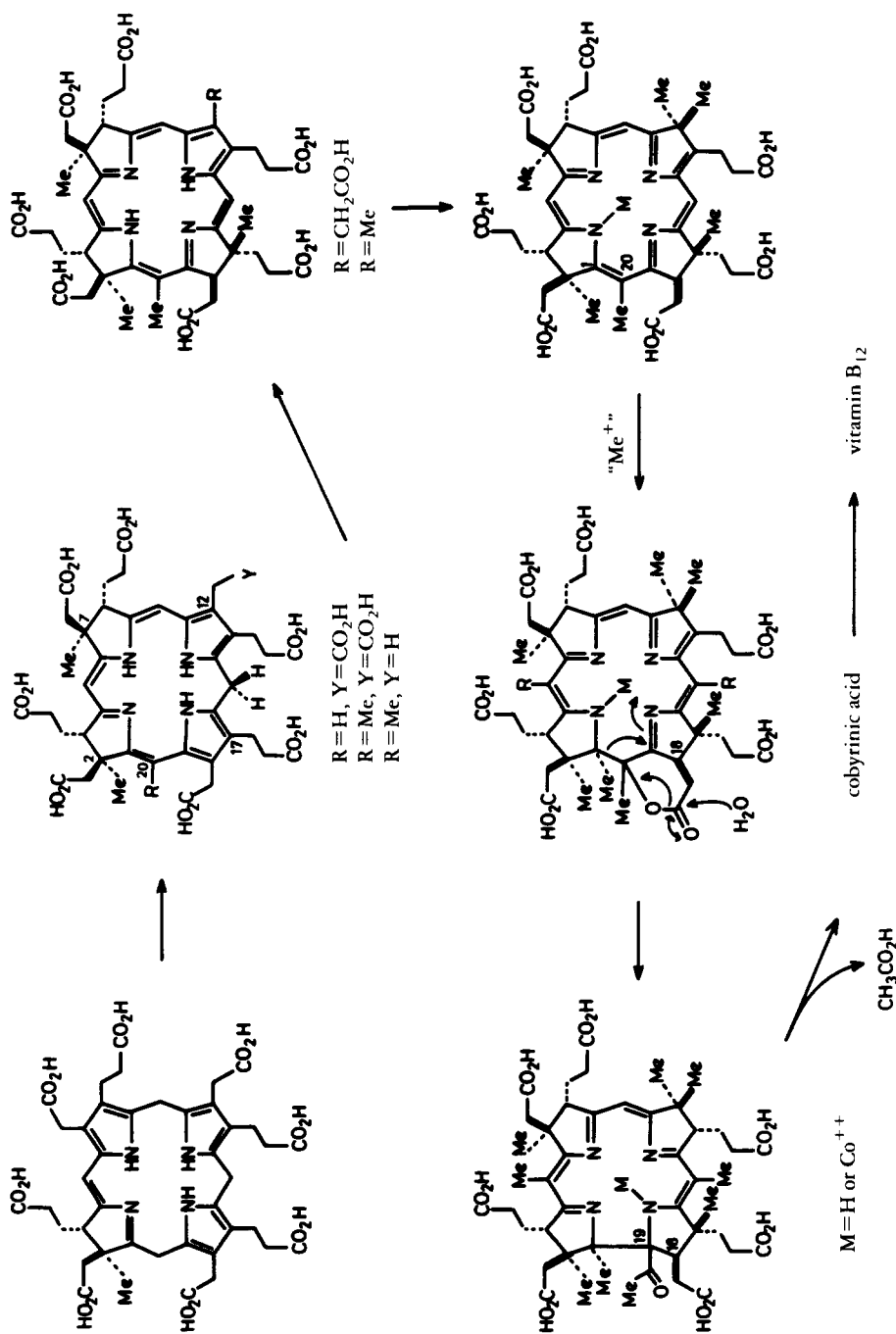
SCHEME 25.

systems use to build vitamin B_{12}; then the Everest of biosynthetic problems will have been scaled.

ACKNOWLEDGMENTS

LITERATURE CITED

1. A.R. Battersby and E. McDonald, in: "B_{12}." Ed. by D. Dolphin, Wiley, New York, 1982, p. 107.
2. A.R. Battersby, M. Ihara, E. McDonald, F. Satoh, and D.C. Williams, *J. Chem. Soc., Chem. Commun.*, 436 (1975).
3. A.R. Battersby, E. McDonald, R. Hollenstein, M. Ihara, F. Satoh, and D.C. Williams, *J. Chem. Soc., Perkin Trans. 1*, 166 (1977).
4. A.R. Battersby, K. Jones, E. McDonald, J. Robinson, and H.R. Morris, *Tetrahedron Lett.*, 2213 (1977).
5. A.R. Battersby, E. McDonald, H.R. Morris, M. Thompson, D.C. Williams, V. Bykhovsky, N. Zaitseva, and V. Bukin, *Tetrahedron Lett.*, 2217 (1977).
6. A.R. Battersby, E. McDonald, M. Thompson, and V. Ya Bykhovsky, *J. Chem. Soc., Chem. Commun.*, 150 (1978).
7. A.R. Battersby, G.W.J. Matcham, E. McDonald, R. Neier, M. Thompson, W.-D. Woggon, V. Ya Bykhovsky, and H.R. Morris, *J. Chem. Soc., Chem. Commun.*, 185 (1979).
8. A.R. Battersby, K. Frobel, F. Hammerschmidt, and C. Jones, *J. Chem. Soc., Chem. Commun.*, 455 (1982).
9. N.G. Lewis, R. Neier, G.W.J. Matcham, E. McDonald, and A.R. Battersby, *J. Chem. Soc., Chem. Commun.*, 541 (1979).
10. A.R. Battersby, M.J. Bushell, C. Jones, N.G. Lewis, and A. Pfenninger, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 13 (1981).
11. L. Mombelli, C. Nussbaumer, H. Weber, G. Muller, and D. Arigoni, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 9 (1981).
12. F.-P. Montforts, S. Ofner, V. Rasetti, A. Eschenmoser, W.-D. Woggon, K. Jones, and A.R. Battersby, *Angew. Chem., Int. Ed. Engl.*, **18**, 675 (1979).
13. P. Naab, R. Lattmann, C. Angst, and A. Eschenmoser, *Angew. Chem., Int. Ed. Engl.*, **19**, 143 (1980).
14. S. Ofner, V. Rasetti, B. Zehnder, and A. Eschenmoser, *Helv. Chim. Acta*, **64**, 1431 (1981).
15. P.J. Harrison, Z.-C. Sheng, C.J.R. Fookes, and A.R. Battersby, *J. Chem. Soc., Perkin Trans. 1*, 1667 (1987).
16. A.R. Battersby, M.H. Block, S.C. Zimmerman, G.B. Henderson, S.P.D. Turner, S.W. Westwood, and F.J. Leeper, *J. Chem. Soc., Chem. Commun.*, 1061 (1985).
17. H.C. Uzar and A.R. Battersby, *J. Chem. Soc., Chem. Commun.*, 1204 (1982).
18. H.C. Uzar and A.R. Battersby, *J. Chem. Soc., Chem. Commun.*, 585 (1985).
19. H.C. Uzar, T.A. Carpenter, F.J. Leeper, and A.R. Battersby, *J. Chem. Soc., Perkin Trans. 1*, 1689 (1987).
20. A.I. Scott, N.E. Mackenzie, P.J. Santander, P.E. Fagerness, G. Muller, E. Schneider, R. Sedlmeier, and G. Worner, *Bioorg. Chem.*, **12**, 356 (1984).
21. V. Rasetti, A. Pfaltz, C. Kratky, and A. Eschenmoser, *Proc. Natl. Acad. Sci. U.S.A.*, **78**, 16 (1981).
22. A.R. Battersby, *Acc. Chem. Res.*, **19**, 147 (1986).