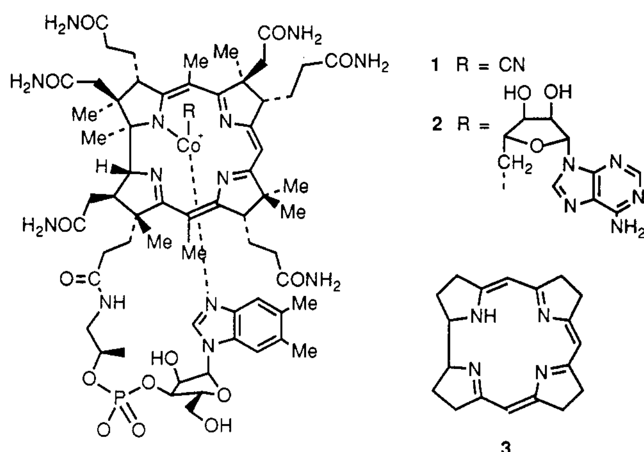
Biosynthesis of Vitamin B_{12}

ALAN R. BATTERSBY†

University Chemical Laboratory, Cambridge, CB2 1EW, U.K.

Received June 29, 1992

No chemist worth his salt can fail to be fascinated by the chemistry carried out with such exquisite control in living systems. That a humble microorganism can construct a molecule having the complexity of vitamin B_{12} (1), the antipernicious anemia vitamin, is quite amazing. No less remarkable are the metabolically important rearrangement reactions[1] catalyzed by a family of enzymes that use the coenzyme form 2 of the parent macrocycle 1. But it is the question "How is vitamin B_{12} (1) built in Nature" that will dominate this Account; such is the complexity of the vitamin that it has often been referred to as the Everest of biosynthetic problems. The molecule is based on the unique *corrin* macrocycle for which 3 is the parent.



1 R = CN

2 R =

3

The scale of the problem can be judged from the fact that the first experiments on B_{12} biosynthesis were initiated about 40 years ago,[2] yet we still do not have complete knowledge of all the biosynthetic stages on the long pathway. Nevertheless, there has been enormous progress, and the analogy of climbing Everest is apt for the pattern of this forward movement. This has been characterized by periods where the research workers have been lodged at different "staging camps" each followed by an upward surge. The residence times at some of the "staging camps" for B_{12} have sometimes been agonizingly long! But recently there have been dramatic new developments leading not only to rapid forward and upward progress but also to a striking change of direction up the "mountain". This Account will focus sharply on these latest advances. Readers interested in the details of the earlier work can consult reviews in previous Accounts[3,4] and elsewhere.[1,5] The more general readers will first be guided along those parts of the biosynthetic pathway that had been elucidated before the very recent advances. This treatment must be brief and logical (rather than necessarily chronological); it concentrates on the essential minimum information, and the large number of original references available in the quoted reviews are not repeated here, though more recent ones are covered. Some sensitivity and imagination are needed in reading the next two sections in order to feel the excitement by putting the flesh of real life onto the bare bones of the skeleton. In this way, one can also get some appreciation of the massive effort that was involved by the groups of Arigoni, Battersby, Blanche, Bykhovski, Crouzet, Müller, and Scott; Bernhauer and Shemin also made some early contributions.

The Pathway as Far as the Trimethylated Intermediate, Precorrin-3. Living systems construct the huge variety of substances they contain by a system of branching pathways. So one starting material can serve to form many different final products. A striking example is that the oxygen carrier, heme, the photosynthetic pigment, chlorophyll, and vitamin B_{12} all share a common origin in uroporphyrinogen III (7), shortened to uro'gen III. Uro'gen III in turn is biosynthesized[6] from 5-aminolevulinic acid (4, ALA) by way of porphobilinogen (5, PBG) and (hydroxymethyl)bilane (6), Scheme I. Since it was early shown that cobyrinic acid (11) is a late precursor of vitamin B_{12} (1) in the anaerobic organism *Propionibacterium shermanii*, one can compare structure 11 with that of uro'gen III (7) to pinpoint the necessary changes. These are (a) at least seven C-methylations (we shall see later that the true number is eight); (b) ring contraction to generate the corrin macrocycle; (c) decarboxylation of the 12-acetate residue; (d) cobalt insertion; (e) possible redox changes. Clearly the number of sequences by which these many transformations could be carried out is enormous, and the researchers' task is to discover the actual one that

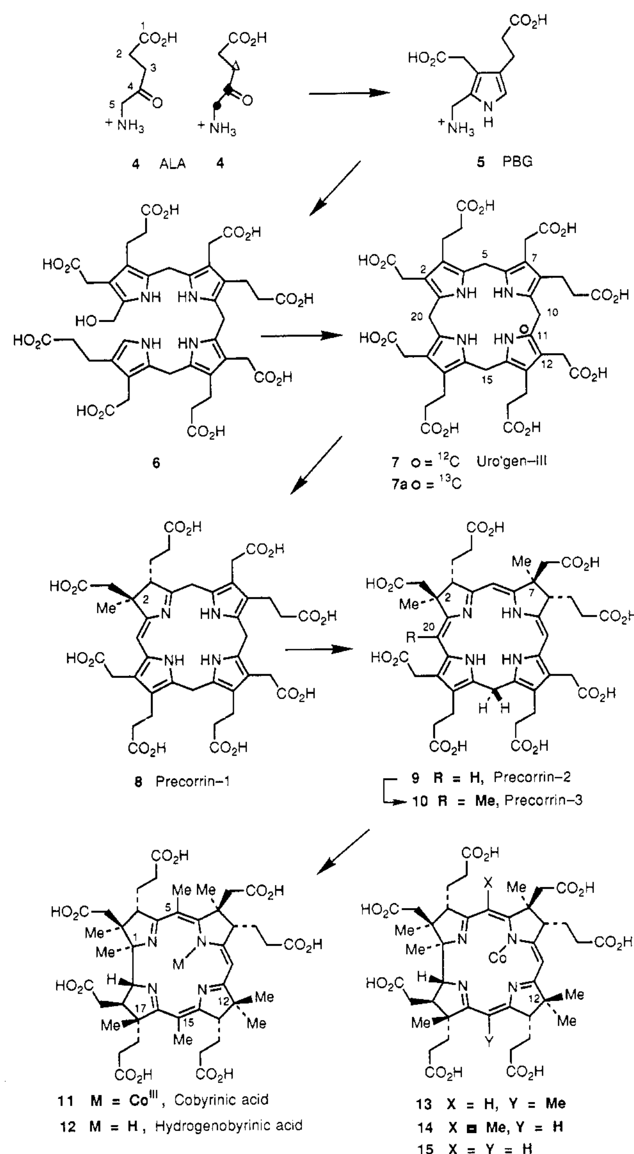
Alan Battersby was born in England in 1925, and he now holds the 1702 Chair of organic chemistry at the University of Cambridge, having previously held a chair at the University of Liverpool. He is a Fellow of the Royal Society and received the Davy and Royal Medals of that society. He worked in the United States during 1950–1952 (Rockefeller Institute and University of Illinois), and in 1982 he received the Roger Adams Award from the ACS. The Feltrinelli and Wolf Prizes came to him in 1986 and 1989, respectively, and a knighthood was conferred in 1992. Professor Battersby's present research interests are concerned with the biosynthesis of the tetrapyrrolic pigments of life, especially hemes and vitamin B_{12}.

† The author is the spokesman for a team of close colleagues who have generously contributed their different skills and knowledge in making the major advances described in this Account. They are, in Paris, Francis Blanche, Beatrice Cameron, Joel Crouzet, Laurent Debussche, Denis Thibaut and, in Cambridge, Fumiyuki Kiuchi, Masahito Kodera, Finian Leeper, and George Weaver.

(1) For a full review, see: B_{12}; Dolphin, D., Ed.; Wiley: New York, 1982; Vols. 1, 2 and especially Battersby, A. R.; McDonald, E., in Vol. 1, p 107.
(2) Shemin, D.; Corcoran, J. W.; Rosenblum, C.; Miller, I. W. *Science* **1956**, *124*, 272.
(3) Battersby, A. R. *Acc. Chem. Res.* **1986**, *19*, 147.
(4) Scott, A. I. *Acc. Chem. Res.* **1990**, *23*, 308.
(5) Leeper, F. J. *Nat. Prod. Rep.* **1989**, *6*, 171 and references therein.
(6) For a recent review of the biosynthesis of uro'gen III, see: Battersby, A. R.; Leeper, F. J. *Chem. Rev.* **1990**, *90*, 1261.

0001-4842/93/0126-0015$04.00/0 © 1993 American Chemical Society

Scheme I



may have small variations in different organisms, e.g., the timing of cobalt insertion.

Most of the major advances in knowledge of B_{12} biosynthesis until recent times came from experiments on *P. shermanii*. By exclusion of cobalt from the medium, three new pigments, one green and two violet, were isolated, which proved to be mono-, di-, and trimethylated derivatives of uro'gen III (**7**). It turned out that these deeply colored materials had been formed by adventitious dehydrogenation of the true biosynthetic intermediates for cobyrinic acid (**11**), which were highly sensitive to air. But determination of the structures of these isolated substances gave the important information that the first C-methylation of uro'gen III (**7**) en route to B_{12} occurs at C-2, the second at C-7, and the third at C-20. The numbering system is easily remembered by recognizing that the one-carbon bridges connecting the pyrrole rings (or their derivatives) are always numbered C-5, C-10, C-15, and C-20. Later it was rigorously proved that the actual dimethylated intermediate is the yellow precorrin-2[7] having structure[9] **9**, and structure[9] **10** was assigned[5,10] to the closely related (also yellow) precorrin-3, the trimethylated precursor of cobyrinic acid (**11**); the latter structure has recently been confirmed.[11] In addition, the oxidation level of the monomethylated intermediate, precorrin-1, was shown[12] to be as in structure[9] **8**.

By this stage, the early part of the biosynthetic pathway in *P. shermanii* was well delineated (Scheme I) with uro'gen III (**7**) undergoing stepwise methylation by way of precorrin-1 (**8**) and precorrin-2 (**9**) to give precorrin-3 (**10**) and then forward to cobyrinic acid (**11**). It should be noted that all four intermediates carry *eight* double bonds, that is, their oxidation level is the same, a reassuring finding since C-methylation should not change the oxidation state of the substrate.

Naturally, many searches were made to find precorrin-4 and later intermediates, but despite enormous efforts, no higher methylation products could be detected prior to 1990. However, though new intermediates were not forthcoming during this period, very important progress was being made in other ways now to be outlined.

Signposts on the Pathway beyond Precorrin-3. The ring-contraction step necessary to generate the corrin macrocycle of B_{12} from the foregoing precursors was extensively studied by multiple labeling techniques. As a result, it was shown that C-20 of precorrin-3 (**10**) and the methyl group attached to it are extruded at some later stage on the biosynthetic pathway. Interestingly, the ejected fragment was *acetic acid* with its methyl group derived from the original C-20 methyl and the carboxyl group from C-20 of the macrocycle. Possible mechanisms for the ring-contraction process must accommodate these findings.

Other experiments made use of so-called "pulse labeling" with ^{13}C, a technique that allows information to be gained about the *order* of sequential events, in this case, the order of introduction of the last five methyl groups beyond the trimethylated stage. Two studies involving *Clostridium tetanomorphum* and *P. shermanii* as the organisms led to C-17 being identified as the site for the fourth methylation followed by C-12α, then C-1, with C-5 and C-15 being the last to be methylated (see **11**).

Naturally, knowing that the 5-methyl and 15-methyl are added last, the question was asked as to whether the corrin macrocycles lacking methyl groups at C-5 (**13**) or C-15 (**14**) or at both these sites (**15**) could act as precursors of cobyrinic acid (**11**). Two independent sets of enzymic experiments[13,14] showed that none of these nor-compounds in labeled form acted as a

(7) Biosynthetic intermediates that precede the first formed corrin macrocycle on the pathway are given the name *precorrin* together with a number corresponding to the number of C-methyl groups that have been introduced from SAM into uro'gen III (7) to generate that intermediate. If two or more intermediates carry the same number of SAM-derived methyl groups, then letters are added to the number to allow discrimination. See ref 8 for details.

(8) Uzar, H. C.; Battersby, A. R.; Carpenter, T. A.; Leeper, F. J. *J. Chem. Soc., Perkin Trans. 1* **1987**, 1689.

(9) Though specific positions are illustrated for the double bonds and NH groups around rings A and B of structures 8–10, the proviso must be made that, for any or all of the three cases, the enzymic product may be a close tautomer that undergoes change during isolation.

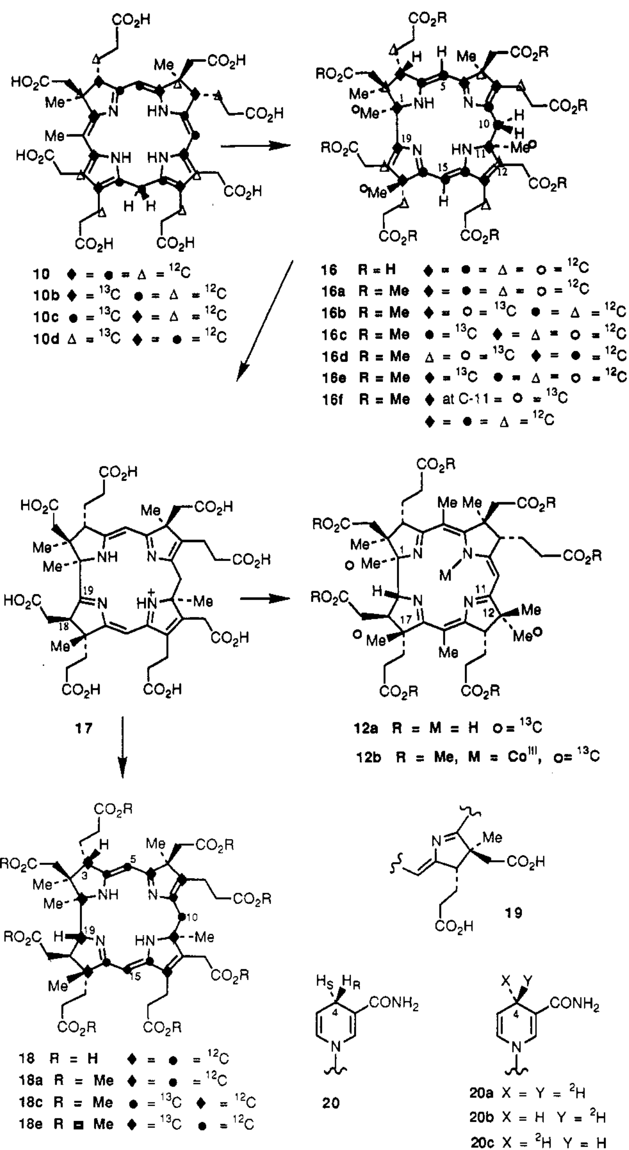
(10) Thibaut, D.; Debussche, L.; Blanche, F. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8795.
(11) Warren, M. J.; Roessner, C. A.; Ozaki, S.-I., Stolowich, N. J.; Santander, P. J.; Scott, A. I. *Biochemistry* **1992**, *31*, 603.
(12) Brunt, R. D.; Leeper, F. J.; Grgurina, I.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1989**, 428.
(13) Kulka, J.; Nussbaumer, C.; Arigoni, D. *J. Chem. Soc., Chem. Commun.* **1990**, 1512.
(14) Grgurina, I.; Handa, S.; Weaver, G.; Cole, P. A.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1990**, 1514.

precursor of **11**, a result that will later be seen to drop neatly into place.

Progress has also been made on the timing of cobalt insertion, and it is becoming apparent that this can differ in different B_{12}-producing organisms. Thus, the aerobic organism *Pseudomonas denitrificans*, which is used for commercial production of B_{12} and will be a major contributor to the main section of this Account, carries out its biosynthetic steps largely with metal-free macrocycles to yield hydrogenobyrinic acid (**12**), which can be isolated.[15] Notice that this material (**12**) is simply the cobalt-free form of cobyrinic acid (**11**); importantly, pulse-labeling experiments[15] with the *P. denitrificans* system established that the seven SAM-derived methyl groups of **12** are set in place following the *same sequence* as was described earlier for cobyrinic acid (**11**) in *P. shermanii*; i.e., 2, 7, 17, 12α, 1, 5/15. Presumably cobalt is inserted late in this organism. The present indications are that this is not so for the anaerobic *P. shermanii*. The observation already mentioned that strict exclusion of cobalt from the growing medium causes precorrin-2 (**9**) and precorrin-3 (**10**) to accumulate (and undergo oxidation by air) hints that cobalt is inserted at, or not far ahead of, the third methylation stage. Two other studies point that same way, one[16] involving incorporation experiments including the use of cobalt derivatives of the dehydrogenated (aromatic) forms of precorrin-2 and precorrin-3, while the other[17] employed pulse labeling with ^{60}Co and [*methyl*-^{14}C]SAM. Neither study on its own gives a fully reliable fix on the exact point of cobalt insertion. But the fact that these two very different approaches both point to cobalt being inserted at or close to precorrin-3 increases one's confidence that in *P. shermanii* the cobalt is set in place early on the pathway rather than late. Also, enzyme preparations from this organism will not convert[15] hydrogenobyrinic acid (**12**) into cobyrinic acid (**11**), and a reasonable explanation is that cobalt is present before the corrin macrocycle is generated in *P. shermanii*.

The nonspecialist reader is now equipped with all the essential knowledge of the biosynthetic pathway to vitamin B_{12} as it was known early in 1990; also some more recent results concerning cobalt insertion have been included. The foundation is thus in place on which to build our survey of recent exciting events. These new successes have come entirely from research on the commercial organism *P. denitrificans*, and the remainder of this Account will be restricted almost entirely to this organism.

Discoveries since 1990. There can be no doubt that genetics and molecular biology jointly formed the springboard from which fresh progress was made. A major effort by Crouzet, Blanche, and their teams[18-20] during the years preceding 1990 led to the detection and sequencing of the genes coding for the various enzymes that carry out the biosynthesis of vitamin B_{12} in *P. denitrificans*. It will be seen later that many of these enzymes have been purified, characterized, and overproduced. One can hardly overemphasize how important this phase of the work was in allowing the subsequent detection of new biosynthetic intermediates on the B_{12} pathway. But details of the genetics[21] are outside the scope of this Account; rather, our purpose is to survey the more chemical, structural, and spectroscopic studies on the newly discovered intermediates together with some of the enzymology. Accordingly, just the key information about the various genes is collected in Table I. Table I is a monument to the



(15) Blanche, F.; Thibaut, D.; Frechet, D.; Vuilhorgne, M.; Crouzet, J.; Cameron, B.; Hlineny, K.; Traub-Eberhard, U.; Zboron, M.; Müller, G. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 884.

(16) Müller, G.; Zipfel, F.; Hlineny, K.; Savvidis, E.; Hertle, R.; Traub-Eberhard, U.; Scott, A. I.; Williams, H. J.; Stolowich, N. J.; Santander, P. J.; Warren, M. J.; Blanche, F.; Thibaut, D. *J. Am. Chem. Soc.* **1991**, *113*, 9893.

(17) Monaghan, S. M.; Vishwakarma, R. A.; Balachandran, S.; Prelle, A.; Stamford, P. J.; Leeper, F. J.; Battersby, A. R. Manuscript in preparation.

(18) Cameron, B.; Briggs, K.; Pridmore, S.; Brefort, G.; Crouzet, J. *J. Bacteriol.* **1989**, *171*, 547.

(19) Crouzet, J.; Cauchois, L.; Blanche, F.; Debussche, L.; Thibaut, D.; Rouyez, M.-C.; Rigault, S.; Mayaux, J.-F.; Cameron, B. *J. Bacteriol.* **1990**, *172*, 5968.

(20) Crouzet, J.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez, M.-C.; Blanche, F.; Thibaut, D.; Debussche, L. *J. Bacteriol.* **1990**, *172*, 5980.

(21) Many genes concerned with B_{12} biosynthesis have also been identified in *Bacillus megaterium* (ref 22) and in *Salmonella typhimurium* (ref 23); see also ref 24. For *S. typhimurium*, the important work of Roth's group has provided many sequences for the B_{12} genes of this organism (ref 23). So far, no new intermediates have been structurally identified in either of these organisms.

(22) Wolf, J. B.; Brey, R. N. *J. Bacteriol.* **1986**, *166*, 51. Brey, R. N.; Banner, C. D. B.; Wolf, J. B. *J. Bacteriol.* **1986**, *167*, 623.

(23) Jeter, R. M.; Olivera, B. M.; Roth, J. R. *J. Bacteriol.* **1984**, *159*, 206. Jeter, R. M.; Roth, J. R. *J. Bacteriol.* **1987**, *169*, 3189.

(24) Scott, A. I. *Tetrahedron* **1992**, *48*, 2559.

Table I. The Genes and Enzymes for Corrin Biosynthesis in *P. denitrificans*[a]

gene	function of enzyme	ref
cobA	methylation at C-2 and C-7 of uro'gen III	40
cobI	methylation at C-20 of precorrin-2	41
cobK	reduction precorrin-6x	29
cobL	methylation at C-5 and C-15 of precorrin-6y and C-12 acetate decarboxylation	34
cobH	rearrangement of precorrin-8x to hydrogenobyrinic acid	35
cobF *cobJ* *cobM*	probably methyl transferases	38
cobG	?	20
cobE	?	19

[a] Many genes involved with the stages beyond hydrogenobyrinic acid (12) have been identified; ref 38 and references therein.

research efforts that provided it, but also it illustrates and emphasizes the following message: that any chemist working on the biosynthesis of a natural product, especially a complex one, must embrace genetics and molecular biology.

The present scientific trail starts with the development[20] of a genetically engineered strain of *P. denitrificans*, in which eight of the genes involved in the biosynthesis of hydrogenobyrinic acid (**12**) had been overexpressed. This means that the eight corresponding enzymes are overproduced by this strain. Let it be emphasized that this strain is in no way odd or aberrant; it simply produces *more of the normal enzymes* and so allows greater turnover of material. Since ^{13}C-NMR spectroscopy was to be used for the structural studies, amounts of material were critically important. The collection of enzymes in the cells could be released by disrupting them, and this cell-free solution was able to convert precorrin-3 (**10**) into hydrogenobyrinic acid (**12**) in high yield on a scale suitable for NMR spectroscopy, provided various cofactors[25] were added to the incubation mixture. One of these cofactors was reduced nicotinamide adenine dinucleotide phosphate, NADPH, part structure **20**. The seminal observation was made[10] that when this cofactor was omitted from the incubation mixture, not only was there no hydrogenobyrinic acid (**12**) formed but a previously unobserved yellow pigment was produced in its place. Importantly, this new pigment in doubly-labeled form (^{3}H:^{14}H) was converted into the corrin **12** in high yield without significant change in the ^{3}H:^{14}C ratio when it was incubated with the enzyme system with the full set of cofactors including NADPH. *So a new biosynthetic intermediate had been isolated.* This new substance was a fascinating one that was to have as profound an effect on the field of B_{12} biosynthesis as did the earlier detection of precorrins-1, -2, and -3; it was named *precorrin-6x* for reasons that will become clear later.

The Structure of Precorrin-6x. How many *C*-methyl groups does the new intermediate contain? This was answered[10] by preparing it from [2,7,20-*methyl*-^{14}C]precorrin-3 using [*methyl*-^{3}H]SAM, and the results showed that three methyls had been added to precorrin-3. So the new material had arisen by hexamethylation of uro'gen III (**7**) and therefore was named[26] precorrin-6x. Also, the ^{14}C values indicated, very surprisingly, that ring contraction had occurred because the C-20 methyl group of [2,7,20-*methyl*-^{14}C]-precorrin-3 was no longer present in precorrin-6x. The locations of the three added methyl groups were studied by preparing precorrin-6x from unlabeled precorrin-3 (**10**) using [*methyl*-^{13}C]SAM and then enzymically converting this product into hydrogenobyrinic acid (**12a**). Nonenzymic cobalt insertion and esterification afforded heptamethyl cobyrinate (**12b**) for NMR analysis. This showed that the *C*-methyl groups that appear in the final ester **12b** at C-17, C-12α, and C-1 are present in precorrin-6x. This result was also in accord with the earlier findings that the 5- and 15-methyl groups of hydrogenobyrinic acid (**12**) and of cobyrinic acid (**11**) are the last to be added.

But the surprises were not yet over. Precorrin-6x proved to be an octacarboxylic acid, and hence the 12-acetate had not undergone decarboxylation. Yet if the methyl group at 12α in the final corrin **12** is at the same position in its precursor precorrin-6x, then subsequent decarboxylation would surely be blocked; see **19**. Also, precorrin-6x was shown to have seven double bonds and so to be at the oxidation level of a dehydrocorrin rather than of a corrin (six double bonds). Hence a reduction step is needed before the corrin **12** is reached, and this fits with the method used to accumulate precorrin-6x (omission of NADPH).

All the experiments so far were carried out entirely by the French group,[10] and it was abundantly clear that precorrin-6x was a most exciting intermediate standing roughly at the middle of the unknown part of the pathway from **10** to **11** and **12**. Accordingly, a joint effort was launched between the teams in Paris and Cambridge to determine its complete structure.[27] The approach involved ^{13}C-labeling of every carbon in the macrocycle, and also the three introduced methyl groups, in a series of interlocking experiments. Let us outline just one set starting with the enzymic preparation of precorrin-3 (**10b**) from [4-^{13}C]ALA (see structure **4**). This was converted into precorrin-6x using [*methyl*-^{13}C]SAM as earlier, and its octamethyl ester **16b** was isolated for study by the full panoply of modern NMR techniques. [5-^{13}C]ALA and [3-^{13}C]ALA (see structure **4**) were starting materials for two further preparations of labeled precorrin-3, **10c** and **10d**, which as before yielded precorrin-6x ester, **16c** and **16d**, respectively, for NMR analysis. The assembled mass of data led to the quite startling structure **16a** for precorrin-6x octamethyl ester.[27,30]

In summary, the foregoing studies gave rigorous confirmation that in precorrin-6x (**16**) ring contraction had most surprisingly already occurred, showed that the additional double bond was C-18/C-19 and the C-12

(25) These cofactors are added routinely for such enzymic conversions and were also included for the earlier enzymic transformations (ref 5) of precorrin-2 (**9**) and precorrin-3 (**10**) into cobyrinic acid (**11**) by cell-free systems from *P. shermanii* though in this case on a very small scale.

(26) Only later will it be known how many different hexamethylated intermediates exist. When precorrin-6x has been shown to be, for example, the second or third member of the precorrin-6 family in the biosynthetic sequence, then the name will be changed respectively to precorrin-6B or precorrin-6C.

(27) Thibaut, D.; Blanche, F.; Debussche, L.; Leeper, F. J.; Battersby, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8800.

(28) Blanche, F.; Kodera, M.; Couder, M.; Leeper, F. J.; Thibaut, D.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1992**, 138.

(29) Blanche, F.; Thibaut, D.; Famechon, A.; Debussche, L.; Cameron, B.; Crouzet, J. *J. Bacteriol.* **1992**, *174*, 1036.

(30) As for earlier cases, the proviso is needed that, though the double-bond locations are established for the ester, the initial octaacid produced enzymically may be a close tautomer.

acetate was still present, and most striking of all, proved that C-methylation had occurred not at C-12α but at C-11. This last feature was confirmed[28] by biosynthesis of precorrin-6x from [*methyl*-^{13}C]SAM and [11-^{13}C]-uro'gen III (**7a**) that had been unambiguously synthesized. It was satisfying to find that the ^{13}C-NMR signals from precorrin-6x ester **16f** for C-11 and for the methyl group previously assigned as that at C-11 were coupled. Notice that methylation at C-11 generates two separate short chromophores, and this is why precorrin-6x (**16**) is only pale yellow. Methylation at C-11 was totally unexpected, yet once known, it allowed one to understand previously puzzling features such as why it is the 12-acetate that specifically undergoes decarboxylation in B_{12} biosynthesis; we shall return to this point.

The Reductase Enzyme and the Structure of Precorrin-6y. The foregoing summary brings out that precorrin-6x (**16**) can be accumulated by blocking a reductive step. The relevant reductase enzyme has been identified in *P. denitrificans*, it is coded by the *cobK* gene, and the reductase was isolated and characterized.[29] The enzyme is specific in its requirement for NADPH. Protein preparations from *P. denitrificans* containing this reductase and, in later experiments, samples of highly enriched reductase were used to reduce precorrin-6x (**16**) with NADPH as *the only cofactor*. By using this SAM-free system, steps beyond the reduction were blocked, so yielding[29] the next biosynthetic intermediate on the pathway, precorrin-6y. This was converted into hydrogenobyrinic acid (**12**) in high yield when incubated under normal conditions. The isolated precorrin-6y was even paler yellow than precorrin-6x.

The joint Paris–Cambridge structural work on precorrin-6y started with the preparation from [4-^{13}C]ALA (see structure **4**) of ^{13}C-labeled precorrin-6x (**16e**), which was enzymically reduced to afford precorrin-6y. This sample and a second sample of precorrin-6y similarly biosynthesized from [5-^{13}C]ALA were isolated as the corresponding octamethyl esters, **18e** and **18c**, respectively. It turned out that partial epimerization at C-3 had occurred during the isolation process, but the epimers were separable. As a result of full NMR analysis of these labeled products, the structure[30] **18a** was established[31] for precorrin-6y octamethyl ester.

Two further questions needed answers before a reasonably complete picture of the foregoing reduction process was available: (a) To which site in precorrin-6x (**16**) is the hydride equivalent delivered from NADPH? (b) From which face of NADPH (**20**) is it removed? The answer to the first question was gained by using [4-$^2H_2$]NADPH (**20a**) as cofactor for the reduction of precorrin-6x in both ^{13}C-labeled **16e** and unlabeled forms. The two sets of experiments led clearly to the same conclusion,[32] that C-19 is the site of reduction, and presumably it is the C-18-protonated form of precorrin-6x (**17**) to which hydride transfer from NADPH occurs. The preparation of [4(*R*)-$^2H_1$]NADPH (**20b**) and [4(*S*)-$^2H_1$]NADPH (**20c**) then allowed proof[33] that the reductase catalyzes the transfer of H_R from NADPH (**20**).

(31) Thibaut, D.; Kiuchi, F.; Debussche, L.; Leeper, F. J.; Blanche, F.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1992**, *139*.
(32) Weaver, G. W.; Leeper, F. J.; Battersby, A. R.; Blanche, F.; Thibaut, D.; Debussche, L. *J. Chem. Soc., Chem. Commun.* **1991**, 976.
(33) Kiuchi, F.; Thibaut, D.; Debussche, L.; Leeper, F. J.; Blanche, F.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1992**, 306.

The way was now open, at least in principle, to detect the next biosynthetic intermediate by seeking the enzyme that transforms precorrin-6y. Comparison of precorrin-6y (**18**) with hydrogenobyrinic acid (**12**) shows that three different reaction types are involved: (i) methylation at C-5 and C-15, (ii) decarboxylation of C-12 acetate, and (iii) rearrangement of the C-11 methyl group to C-12, not necessarily in that order. The successful approach depended on screening protein fractions from the engineered strain of *P. denitrificans* for methyltransferase activity. Remarkably, a single protein was isolated[34] that not only carried out the C-5 and C-15 methylations but also decarboxylated the acetate at C-12. The structural gene for this enzyme was shown to be *cobL*, and the enzyme itself was purified to homogeneity.[34] Now it was possible to incubate ^{14}C-labeled precorrin-6y, as **18**, with SAM together with this pure enzyme to produce a small amount of the next intermediate on the corrin pathway, which was identical with a sample that had already been obtained in a different way, to be described below. This new intermediate is called precorrin-8x since it will be shown below that two more methyl groups have been added to precorrin-6y (**18**) from SAM.

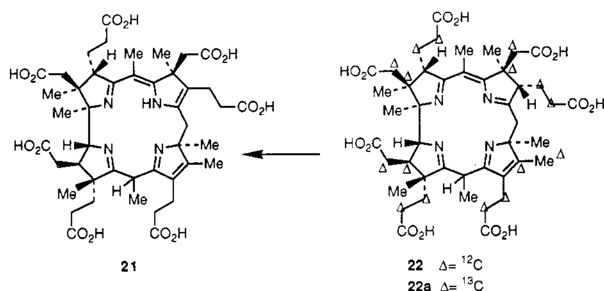
It is fascinating that a single protein catalyzes such different reactions as methyl transfer and decarboxylation. There are indications[34] from the amino acid sequences that the *cobL* gene probably results from a fusion of two ancestral genes, one initially encoding just the C-5/C-15 methylase activity and the other the decarboxylase activity.

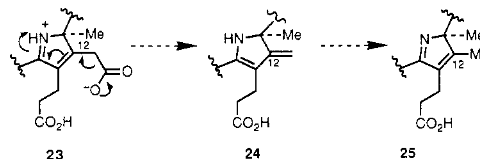
The Structure of Precorrin-8x. Though precorrin-8x could be produced using the purified *cobL* enzyme, a better preparative method was based on the observation[35] that the final step of corrin biosynthesis, viz., precorrin-8x $\rightarrow$ hydrogenobyrinic acid (**12**), is subject to strong product inhibition.[36] That is, added hydrogenobyrinic acid (**12**) inhibits its own formation from precorrin-8x; so addition of **12** to an enzymic incubation of precorrin-3 (**10**), which would normally give **12**, caused precorrin-8x to be accumulated in its place, and it could be isolated.[35] It was then established that equal numbers of methyl groups were added to the three already present when ^{14}C-labeled precorrin-3 (prepared from [4-^{14}C]ALA) was converted in the presence of [*methyl*-^{3}H]SAM either into precorrin-8x or separately into hydrogenobyrinic acid, the number of added methyls necessarily being five for the latter known structure **12**. This result proved[35] that the new product was an octamethylated derivative of uro'gen III (**7**) and hence justifies the name we have been using. Finally, the doubly-labeled (^{3}H:^{14}C) precorrin-8x was shown to be converted very efficiently into hydrogenobyrinic acid, as **12**, without significant change in the labeling ratio; the status of precorrin-8x as a new intermediate was thus secure.

Precorrin-8x proved to be a will-o'-the-wisp. On the one hand, precorrin-8x heptamethyl ester could be

(34) Blanche, F.; Famechon, A.; Thibaut, D.; Debussche, L.; Cameron, B.; Crouzet, J. *J. Bacteriol.* **1992**, *174*, 1050.
(35) Thibaut, D.; Couder, M.; Famechon, A.; Debussche, L.; Cameron, B.; Crouzet, J.; Blanche, F. *J. Bacteriol.* **1992**, *174*, 1043.
(36) Cobyrinic acid (11) also strongly inhibits its own formation by enzyme preparations from *P. shermanii*: Uzar, H. C.; Monaghan, S.; Battersby, A. R. Unpublished work, Cambridge, 1982–1987.
(37) Thibaut, D.; Kiuchi, F.; Debussche, L.; Blanche, F.; Kodera, M.; Leeper, F. J.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.*, in press.

prepared, but it underwent rapid oxidative lactone formation.[35] On the other, the corresponding heptaacid changed as it was handled to give a mixture of at least five closely related forms that could be separated. Imagine the frustrations generated by such instability especially when one is dealing with ca. 300 μg of precious ^{13}C-labeled material. Eventually it was observed that these different forms over time (days) all yielded the same final stable form.[37] At last, structural work could begin, and using the approach of multiple ^{13}C-labeling starting with [4-^{13}C]- and [5-^{13}C]ALA (**4**) followed by NMR as described already for precorrin-6x and precorrin-6y, the structure **21** was established for this stable form of precorrin-8x. Which of the five closely related forms of precorrin-8x is the true biosynthetic intermediate? One component stood out from the rest in the speed and efficiency of its enzymic conversion into **12**; the final stable form **21** of precorrin-8x was not so converted into **12**. When it was found that the biologically active form could be stabilized at high pH, the way was open to show that structure **22** best fitted all the ^{13}C- and ^{1}H-NMR data together with the results summarized earlier. Importantly, incorporation experiments were also carried out in this case starting from [2,3-$^{13}C_2$]ALA (see structure **4**) to label the carbon attached to C-12 (see **22a**) and so to establish that in precorrin-8x this carbon is a methyl attached to an sp^2 carbon and not a methylene group; see probable decarboxylation mechanism **23** → **24** → **25**. Some points

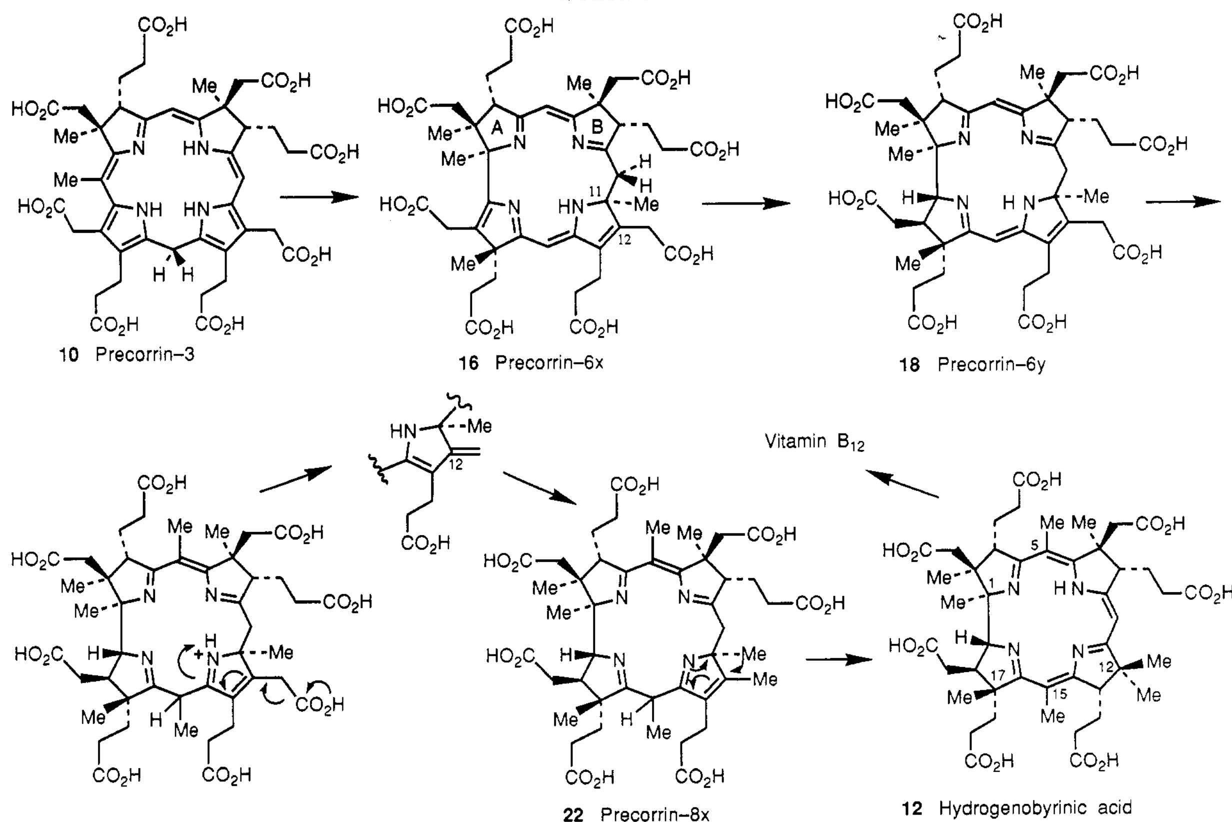




of stereochemistry (e.g., at C-15) remain to be established in structure **22**; nevertheless, all the main features of precorrin-8x, the final intermediate for the corrin system **12**, are now clear, and it is a fascinating molecule. That such a structure can undergo tautomeric change in the two separated chromophores is not surprising, and there could be epimerization at sp^3 centers. Different combinations of such changes can explain the various structural forms all eventually equilibrating to the stable system **21**. It should be noted that precorrin-8x has all the *C*-methyl groups of **12** and **11** and vitamin B_{12} (**1**) attached to the macrocycle, and only rearrangement of the 11-Me to C-12 is needed to yield **12**, a rearrangement catalyzed in *P. denitrificans* by the next enzyme to be considered.

The Final Rearrangement Step Leading to the Corrin System. With precorrin-8x (**22**) in hand, the search was on for the enzyme catalyzing the conversion of this final intermediate into hydrogenobyrinic acid (**12**). As a result, the relevant enzyme was isolated in pure state, and its N-terminal sequence showed it to be encoded by the *cobH* gene.[35] In agreement, a strain of

Scheme II[a]



[a] Rings A and B of **16**, **18**, and **22** have been illustrated as being in the same tautomeric form; see ref 30.

P. denitrificans was available in which *cobH* was not overexpressed, and the enzyme system from this strain stopped the forward transformation of precorrin-6x (**16**) at precorrin-8x (**22**), 45% yield.[35] This "rearrangase" enzyme is a relatively small protein (M_r 22 000), and it merits intensive study.

The Genes and Enzymes. Those genes and enzymes that have played such a crucial role in the foregoing biosynthetic steps are collected in Table I. There too are other key genes and enzymes of *P. denitrificans* concerned with the conversion of uro'gen III (**7**) into precorrin-3 (**10**), a section of the pathway that in this Account has taken second place to the post-1990 discoveries. Finally, other genes are listed that are proven[38] to be involved in the biosynthesis of **12** in *P. denitrificans*, and elucidation of their role cannot be long delayed.

Present Knowledge of the Biosynthetic Pathway. We can now draw together all the foregoing knowledge. Scheme I shows the steps leading forward from ALA (**4**) via uro'gen III (**7**), precorrin-1 (**8**), precorrin-2 (**9**), and on to precorrin-3 (**10**). Now the post-1990 discoveries can be added by joining Scheme II onto Scheme I. The next known intermediate after precorrin-3 (**10**) in *P. denitrificans* is precorrin-6x (**16**), which is reduced to precorrin-6y (**18**). The latter undergoes the remarkable double methylation at C-5 and C-15 and decarboxylation by a process shown as the equivalent of proton-catalyzed decarboxylation of pyrrole-3-acetic acids to afford precorrin-8x (**22**). Finally, precorrin-8x is rearranged, which must involve methyl migration from C-11 to C-12, illustrated as a [1,5] sigmatropic suprafacial step, to form hydrogenobyrinic acid (**12**). In *P. denitrificans*, cobalt insertion must occur at some stage beyond **12** since this organism finally produces vitamin B_{12} (**1**). But in *P. shermanii*, the strong pointers are that cobalt is in place at a much earlier stage, and so several of the later steps presumably involve cobalt-containing macrocycles. Much has been discovered recently, especially by the French group, about the genes[38] and enzymology[1,38] for the amidation steps and the attachment of the nucleotide loop for vitamin B_{12} (**1**), but our focus in this Account has been on the pathway to the corrin macrocycle of the vitamin. The gaps in Schemes I and II are presently at precorrin-4 and precorrin-5 leading up to precorrin-6x (**16**), and this is a measure of the massive progress which has been made. Interesting speculations have been offered[27,39] concerning some of the presently unknown steps.

Prior to 1990, a reasonable guess was that we would be into the next century before the main intermediates on the pathway to corrins and vitamin B_{12} had been mapped out. The situation since 1990 is dramatically different, and it now seems very unlikely that we will have to wait so long before at least the main biosynthetic intermediates are known. Delineation of the main pathway will correspond to reaching the summit of the mountain. Then more work will be needed, e.g., on the mechanism of the key steps. Our joint aim is to complete the climb to the summit and then to contribute strongly to the subsequent consolidation phase.

It has been my immense good fortune to be associated with the outstanding young scientists listed on the first page of this Account. Particular acclaim should go to the French group, who made the latest breakthrough in 1990, which initiated a new surge up the B_{12} "mountain". I am extremely grateful to all my colleagues in Paris and Cambridge for enthusiastically dedicating their skills, knowledge, and effort to the B_{12} problem. I also thank the Science and Engineering Research Council for financial support.

(38) For a scheme showing all genes, see: Crouzet, J.; Levy-Schil, S.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez, M.-C.; Blanche, F.; Debussche, L.; Thibaut, D. *J. Bacteriol.* **1991**, *173*, 6074.

(39) Eschenmoser, A. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 6.

(40) Blanche, F.; Debussche, L.; Thibaut, D.; Crouzet, J.; Cameron, B. *J. Bacteriol.* **1989**, *171*, 4222.

(41) Thibaut, D.; Couder, M.; Crouzet, J.; Debussche, L.; Cameron, B.; Blanche, F. *J. Bacteriol.* **1990**, *172*, 6245.