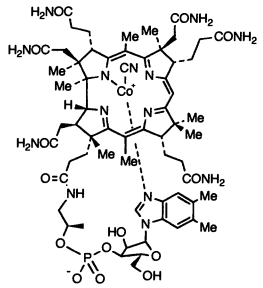
How Nature Builds the Pigments of Life: The Conquest of Vitamin B_{12}

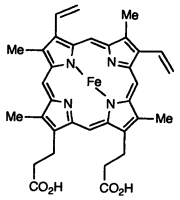
Alan R. Battersby

In part because humans cannot synthesize vitamin B_{12} and must obtain it from organisms that produce it and because B_{12} deficiency leads to pernicious anemia, it has been important to understand how microorganisms build this quite complex substance. As shown here, an interdisciplinary attack was needed, which combined the strengths of genetics, molecular biology, enzymology, chemistry, and spectroscopy. This allowed the step-by-step synthetic pathway of B_{12} to be elucidated, and this approach has acted as a model for future research on the synthesis of substances in living organisms. One practical outcome of such an approach has been the improved availability of B_{12} for animal feedstuffs and human health.

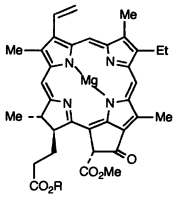
There are those wonderful times in scientific research when a formidable problem is solved: Discoveries are made that not only fit new pieces into the difficult puzzles we study but, by so doing, transform our view of entire areas of research. The past few years have witnessed a dramatic surge in knowledge of how a complex vitamin, B_{12}, is constructed by living systems. This vitamin (**1**) is a cobalt complex that in its coenzyme form is required by enzymes catalyzing several metabolically essential rearrangement reactions. Deficiency in B_{12} leads to pernicious anemia, a serious problem for those afflicted but easily remedied by treatment with B_{12}. Now, the detailed biosynthetic pathway to vitamin B_{12} can be drawn, showing the structures of essentially all of its intermediates, from its simple starting material all the way through to the complex molecular assembly that is vitamin B_{12}. My aim here is to describe the sometimes dramatic developments that led to discovery of its biosynthetic pathway; it is a remarkable detective story.



1



2



3

The Pigments of Life

Evolution has provided living organisms with a family of brightly colored and important organometallic systems that are derived by often extensive structural modification from one parent molecule called uroporphyrinogen III (Fig. 1). By allowing a fine-tuning of their properties, these modifications have led to a group of substances able to carry out a variety of tasks. The lovely colors of these molecules and the vital roles they play in living organisms have led to their being known as the pigments of life. For example, oxygen transport in blood is dependent on the iron complex heme (**2**), whereas plants use the closely related chlorophyll a (**3**), a complex with magnesium, to achieve photosynthesis. Other similar complexes act to shuttle electrons; yet more are cofactors for the reduction of nitrite and sulfite, whereas a nickel complex mediates the natural production of methane (marsh gas). Finally, there is vitamin B_{12}.

The author is in the University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK.

Biosynthesis of B_{12}: The First Phase

As mentioned above, all the pigments of life, including vitamin B_{12}, are derived biosynthetically from uroporphyrinogen III. Uroporphyrinogen III is built from eight molecules of 5-aminolevulinic acid (ALA), which pair two-by-two to yield four molecules of porphobilinogen (PBG) (Fig. 1). These porphobilinogen molecules are assembled through several enzyme-controlled steps to generate uroporphyrinogen III (*1*). When one compares this structure (**6**) with that of heme (**2**) and that of vitamin B_{12} (**1**), two major differences stand out for B_{12}. First, the macrocycle of B_{12} is smaller because of the direct link between rings A and D. Second, there is heavy methylation of carbons around the periphery of the molecule, which sets up many chiral centers. This series of methylations, the ring contraction, cobalt insertion, and several other steps could be carried out in an enormous number of possible sequences. The problem was to discover the sequence of steps used naturally and to elucidate the structures of all the intermediates on the pathway; it was truly the Everest of biosynthetic problems.

Before the availability of the radioactive isotopes ^{14}C and ^{3}H, chemists could only speculate about the natural biosynthetic pathways for the construction of steroids, alkaloids, or porphyrins. But when substances labeled with ^{14}C and ^{3}H were in hand during the 1950s and 1960s, many biosynthetic problems could be solved by enzymatic incorporation into the natural substance of postulated precursors, labeled, for example, with ^{14}C. Then, controlled degradations of the radioactive product had to be devised to pick out individual carbon atoms for radioassay. In this way, the labeling patterns were laboriously worked out, although the chemistry involved was fun.

Major progress in elucidating the biosynthetic pathway of B_{12} started in the late 1960s as nuclear magnetic resonance (NMR) spectrometers capable of observing ^{13}C nuclei were becoming available. It was then possible to study simultaneously many labeled sites in a complex molecule. Labeling with ^{14}C and ^{3}H still contributed important information, but it was ^{13}C that opened the door to substantial progress.

Most of the research on B_{12} until the mid-1980s was carried out with *Propionibacterium shermanii* (*2–4*). This bacterium is normally grown on a cobalt-containing medium, and its cells can be broken to yield a cell-free preparation of soluble enzymes that, with the appropriate cofactors, will synthesize cobyrinic acid (**10**) from simpler precursors (Fig. 1). Cobyrinic acid has the corrin macrocycle of vitamin B_{12} fully built and is a late intermediate for B_{12} itself. However,

when the cells were grown with strict exclusion of cobalt, they no longer synthesized cobyrinic acid but instead produced three new pigments. These turned out to be dehydrogenated forms (by aerial oxidation) of the intermediates, namely precorrin-1 (**7**), precorrin-2 (**8**), and precorrin-3 (**9**) (*5*). The biosynthetic conversion of **7**, **8**, and **9** into cobyrinic acid was demonstrated by decisive labeling experiments.

The extensive structural work on these new pigments and on the precorrins **7**, **8**, and **9** depended heavily on ^{1}H and ^{13}C NMR together with ^{14}C labeling. Importantly, this structural information revealed that the first three steps beyond uroporphyrinogen III on the pathway to B_{12} all involve C methylation, first at C-2, then at C-7, and as the forerunner of many surprises, third at C-20. Clearly it was methylation that distinguished the pathways to B_{12} (**1**), heme (**2**), and chlorophyll a (**3**). The methyl groups are transferred from the methylating agent (*S*)-adenosylmethionine (SAM).

An intense effort was made in several laboratories over many years to isolate new precorrins lying beyond precorrin-3. Before 1990, none was found. It had to be accepted that further progress by this approach was unlikely; happily, other methods gave important results.

One study focused on the fate of C-20 and its attached methyl group present in precorrin-3; these are lost somewhere along the pathway before cobyrinic acid is formed with its direct ring A–ring D link. It was shown by multiple labeling experiments (*6*, *7*) that these two carbon atoms are ejected as acetic acid, with C-20 providing the carboxyl group with the methyl group passing through unchanged. Success depended on radioactive labels because about 9 μg of acetic acid in a substantial volume of aqueous buffer was sought.

The second study used pulse labeling (*8*), a method designed to reveal the order of sequential events. For cobyrinic acid and B_{12}, researchers were interested in pinning down the order of attachment of the remaining five C methyl groups added after precorrin 3. Experiments based on ^{13}C-labeled methyl groups (*9*) proved that C-17 is the fourth methylation site. The method was then extended (*8*, *10*, *11*) to show that C-12α is the fifth site followed by C-1, with C-5 and C-15 being last with roughly similar timing. The 12β-methyl of cobyrinic acid (**10**) is not SAM-derived; it is formed by decarboxylation of the 12-acetate residue still present in precorrin-3 (**9**).

During this first phase of the research on B_{12} biosynthesis, several key facts (*12*) were established: (i) the pathway from ALA (**4**) to uroporphyrinogen III (**6**) was the same for heme (**2**), chlorophyll a (**3**), and vitamin B_{12} (**1**); (ii) the B_{12} pathway then branched away, with methylation at C-2 of **6** that initiated the switch; (iii) the precorrin-1 (**7**) thereby formed was followed by precorrin-2 (**8**) and precorrin-3 (**9**) and the last five methyl groups were added in the following order: C-17, C-12α, C-1, and C-5 and C-15; and (iv) the steps that make the direct ring A–ring D connection, thereby contracting the macrocycle, eliminate C-20 of **9** and its attached methyl group as acetic acid.

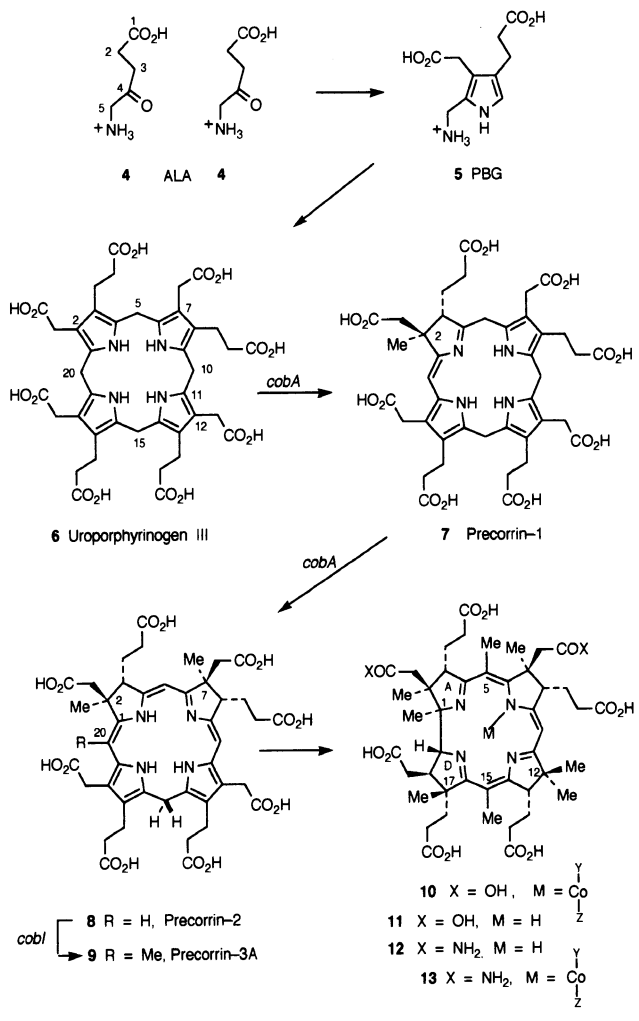


Fig. 1. Early part of the B_{12} biosynthetic pathway. Me, methyl group.

Biosynthesis of B_{12}: Recent Progress

The great progress made during the first phase of identifying the B_{12} biosynthetic pathway depended heavily on ^{13}C NMR. But researchers simply could not get their hands on the intermediates beyond precorrin-3 (**9**) with the methods in use at that time. Something new was needed, and genetics and molecular biology have subsequently opened up research on B_{12}. Chemists now have access to the individual enzymes catalyzing the conversion of one intermediate into the next. In addition, the genes encoding these enzymes can be overexpressed to produce large amounts of enzyme. This in turn allows production of biosynthetic intermediates in quantities sufficient for structural study by multiple isotopic labeling and ^{13}C NMR with the use of a full range of modern NMR techniques. These allow exploration of the molecule both at and around the ^{13}C-labeled sites, although in reality it is not nearly as simple as this outline might suggest. Initially, the scientists knew of 20 or so genes involved in B_{12} biosynthesis, but for only a few genes were the functions of the corresponding enzymes established reasonably quickly. In the unraveling of this puzzle that follows, I will bring out the often startling structures of the missing intermediates that at last could be elucidated with the newer techniques.

Almost all these latest advances have been made in two centers involving teams at Rhône-Poulenc-Rorer in France in genetics and biochemistry (led, respectively, by J. Crouzet and F. Blanche) and another team, more chemically inclined, being my group in Cambridge, England. (Where others have contributed, this will be clearly referenced.) Essentially all the outstanding work on genetics, molecular biology, and enzymology was done independently in France, some of the chemistry was carried out independently in Cambridge, and most, but not all, of the

structural work and chemistry described here has been carried out jointly. The genetics and molecular biology of the B_{12} pathway in *Pseudomonas denitrificans* were worked out first in a massive effort over many years (*13–16*), and this provided the launching pad for subsequent advances not only for the Paris and Cambridge groups but also for those who followed. I will describe initially what chemistry the various enzymes catalyze and control and then correlate each enzyme with its gene (*17*).

The Great Surprise

The B_{12} story has been full of surprises, starting even in the early days of its structure determination (for example, with the discovery that it was a cobalt complex with a contracted macrocycle) and carrying on through to the first phase of research on its biosynthesis [with the discovery that precorrin-3 (**9**) carries a methyl group at C-20 that is subsequently thrown away]. But other events caused the greatest surprise of all. The start was the French team's development of a genetically engineered strain of *P. denitrificans* in which eight of the genes required for biosynthesis of the late B_{12} precursor hydrogenobyrinic acid (**11**) (*15*) from precorrin-2 (**8**) were overexpressed. The former is the cobalt-free form of cobyrinic acid (**10**), referred to earlier as a product of *P. shermanii*; their difference arises simply because cobalt insertion occurs at different points on the pathways in the two organisms. However, the complex organic macrocycles of **10** and **11** are identical in every respect.

The soluble protein preparation from this engineered bacterial strain contained large amounts of the enzymes encoded by the eight overexpressed genes. This cell-free system could convert precorrin-3 (**9**) into hydrogenobyrinic acid (**11**) in high yield provided that the necessary cofactors, including SAM and reduced nicotinamide adenine dinucleotide phosphate (NADPH) (**14**), were added. When this incubation was run without added NADPH, no hydrogenobyrinic acid was formed, although a yellow pigment was produced in its place. This yellow pigment was converted in high yield into **11** by incubation with the complete enzyme system including NADPH, and thus a new intermediate was in hand (*18*). Double-labeling experiments showed that formation of the new intermediate involved the addition of three SAM-derived methyl groups to precorrin-3 (**9**); its name is therefore (*5*) precorrin-6A. By preparing precorrin-6A from precorrin-3 so that the three added methyl groups were ^{13}C-labeled and then converting this sample enzymatically into hydrogenobyrinic acid, researchers found by NMR that the C methyl groups that appear in **11** at C-17, C-12α, and C-1 were the three new ones in precorrin-6A.

Further labeling and mass spectrometric work on the structure of precorrin-6A revealed features that were totally unexpected (*18*). First, the macrocycle was already contracted and C-20 with its attached methyl group [see precorrin-3 (**9**)] had been extruded. Second, precorrin-6A retained the acetate residue at C-12 and had seven double bonds, whereas the final corrin macrocycle (for example, that of **11**) had six. Thus, it was found that reduction is needed to produce vitamin B_{12}, which fits perfectly with the formation of precorrin-6A only when NADPH is omitted.

The great importance of precorrin-6A was now clear, and its structure was elucidated by a joint Paris-Cambridge effort. The plan was to synthesize precorrin-6A from the simple starting material ALA (**4**) by using a combination of the necessary overproduced enzymes. By starting with three forms of ALA (**4**) carrying ^{13}C at different sites (isotopomers), it was possible to synthesize three samples of precorrin-6A, each having ^{13}C atoms at known positions that together labeled every carbon of the macrocycle. Then NMR led us to the striking structure (**20**) for precorrin-6A (*19–21*).

Why did this structure cause such a stir? Largely because it revealed that B_{12} biosynthesis follows an entirely unexpected route. The methylation at C-11 (and not at C-12) of precorrin-6A (**20**) was the greatest surprise to be added to those already mentioned. Because this methyl group is undoubtedly at C-12 by the time the biosynthesis reaches hydrogenobyrinic acid (**11**), methyl migration (C-11 → C-12) has to occur at some later stage as **20** is converted into **11**.

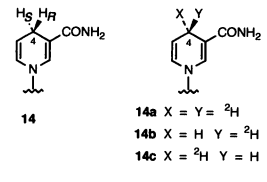
Multi-Enzyme Synthesis

Before the main story is continued here, a message of general importance in these experiments deserves to be highlighted: precorrin-6A (**20**), a complex molecule having seven chiral centers, was synthesized in homochiral form from readily available ALA (**4**) by the combined action of nine enzymes produced by overexpression of their corresponding genes. This synthesis (*18*) and several others described here and using even larger numbers of overproduced enzymes are early examples of preparative multi-enzyme synthesis by a cocktail of enzymes (six or more), like many that are now regularly done. Enzymatic synthesis has enormous potential, and the product can be further manipulated by nonenzymatic methods. When enzymes are available, the actual preparation takes a few hours or perhaps a few days. Otherwise, years of work by a large team would be needed to synthesize precorrin-6A (**20**) nonenzymatically (*22*).

The Reduction Step

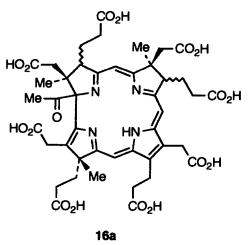
The enzyme that reduces precorrin-6A (**20**) was then purified: It is encoded by the *cobK* gene and is a reductase specifically dependent on NADPH (**14**). When the enzyme was incubated with precorrin-6A and NADPH as the only cofactor, the next biosynthetic intermediate in the sequence, precorrin-6B, could be isolated. Its status was confirmed when it was converted with high yield into hydrogenobyrinic acid (**11**) by incubation with the complete enzyme system and all its cofactors (*23*). With this, researchers were suddenly able to jump from one intermediate to the next by testing one after another of the palette of enzymes until the right one was found. Biosynthetic research had entered a new era.

Rapid progress was now possible in the elucidation of the B_{12} biosynthetic pathway. The structure of precorrin-6B was proven (*24*) to be structure **21** by the same multiple ^{13}C labeling method used for precorrin-6A (**20**); thus, it was found that it is the double bond between C-18 and C-19 of **20** that is reduced. Also, the deuteride equivalent from [4-$^2H_2$]NADPH (**14a**) was found (*25*) to be transferred to C-19 of precorrin-6A; the chemistry makes sense if it is the C-18–protonated form of precorrin-6A that is reduced. At that point, attention focused on the cofactor, and by preparing [4*R*-^2H]-NADPH (**14b**) and [4*S*-^2H]NADPH (**14c**), researchers were able to show that the reductase stereospecifically transfers H_R of the cofactor **14** (*26*). Thus, these experiments, which examined all three components in the process (substrate, cofactor, and product), left no doubt that reduction is a necessary step for B_{12} synthesis in the aerobic bacterium *P. denitrificans*.



14

14a X = Y = ^2H
14b X = H Y = ^2H
14c X = ^2H Y = H



16a

Taking into account the elimination of acetic acid, comparison of the oxidation levels of precorrin-3 (**9**) with those of hydrogenobyrinic acid (**11**) shows that there is no overall redox change. Because reduc-

tion is needed at precorrin-6A (**20**), it followed that there has to be an oxidation step at some stage between precorrin-3 (**9**) and precorrin-6A (**20**). It was satisfying to demonstrate (*27*) that an equivalent reduction step is necessary for B_{12} biosynthesis in *P. shermanii*, an organism grown essentially anaerobically and useful for the early research.

The Pathway Beyond Precorrin-6B

With precorrin-6B (**21**) set firmly in place in the biosynthetic pathway (Fig. 2), we can take a forward look to hydrogenobyrinic acid (**11**) to see what still has to be done. The required steps are as follows, though not necessarily in this order: (i) decarboxylation of the 12-acetate residue; (ii) rearrangement of the 11-methyl group to C-12; and (iii) methylation at C-5 and C-15. Precorrin-6B (**21**) was tested as a potential substrate for the various remaining enzymes known to be involved in B_{12} biosynthesis. Of course, a suitable assay was needed. Success came from assaying for methyltransferase activity, and the enzyme thus identified transformed precorrin-6B (**21**) by addition of two methyl groups presumably at C-5 and C-15, as these are the only two needed. The purified enzyme, encoded by the *cobL* gene, also catalyzed the decarboxylation of the 12-acetate group (*28, 29*). The product of this enzymatic conversion was isolated and named precorrin-8x. Its standing as another new intermediate on the B_{12} pathway was established by a demonstration of its efficient enzymatic conversion into hydrogenobyrinic acid without significant change in the ratio of the two isotopes (^3H and ^{14}C) used for labeling (*30*).

Of the seven structure determinations carried out by the Paris and Cambridge teams during recent research on B_{12} intermediates, the work on precorrin-8x was the most frustrating. Its heptamethyl ester was very unstable (*30*), and even after gentle isolation the heptacarboxylic acid changed in aqueous solution to give a mixture of at least five closely related forms, resulting in complex NMR spectra. Work on B_{12} was never dull. When finally it was possible to hold on to precorrin-8x long enough for NMR studies, the structure **24** was deduced (*31*). Multiple ^{13}C labeling was carried out as earlier, but here an additional experiment, based on [2,3-$^{13}C_2$]ALA [see (**4**)] as starting material, allowed the carbon attached to C-12 of **24** and C-12 itself to be labeled with ^{13}C. Proof was provided in this way that it is a methyl group that is attached to C-12 and that C-12 is an sp^2 center. Thus, the methylene system, which is presumably formed as an intermediate (**23**) during decarboxylation (**22** → **23**), underwent tautomerization to **24** (Fig. 2). Structure **24** for precorrin-8x still lacks some stereochemical detail, but the main features are clear. All the C methyl groups for B_{12} were now attached, and only rearrangement of the 11-methyl to C-12 was needed to form hydrogenobyrinic acid (**11**).

The methyl migration from C-11 in **24** to C-12 in **11** is illustrated as a suprafacial 1,5-sigmatropic rearrangement and is catalyzed by a relatively small enzyme (relative molecular mass of 22,000) encoded by the *cobH* gene (*30*). How the enzyme achieves the methyl shift is an interesting problem for the future. This shift allows the double bonds to move into conjugation, thus setting up the characteristic corrin chromophore of hydrogenobyrinic acid.

Filling the Last Remaining Gap

At this stage, the structures were known of all the intermediates on the biosynthetic pathway to hydrogenobyrinic acid (en route to vitamin B_{12}), from ALA (**4**) through to precorrin-3 (**9**). There was a gap, however, before precorrin-6A (**20**) could be reached. But the various precursors on the pathway running forward from precorrin-6A were known and structurally characterized all the way through to hydrogenobyrinic acid (**11**). Moreover, much was known about the genes and enzymes required to convert **11** into **13** and to add the adenosyl residue and the nucleotide loop, which are needed to complete the construction of the vitamin B_{12} coenzyme (*16*).

The biosynthetic pathway from uropor-

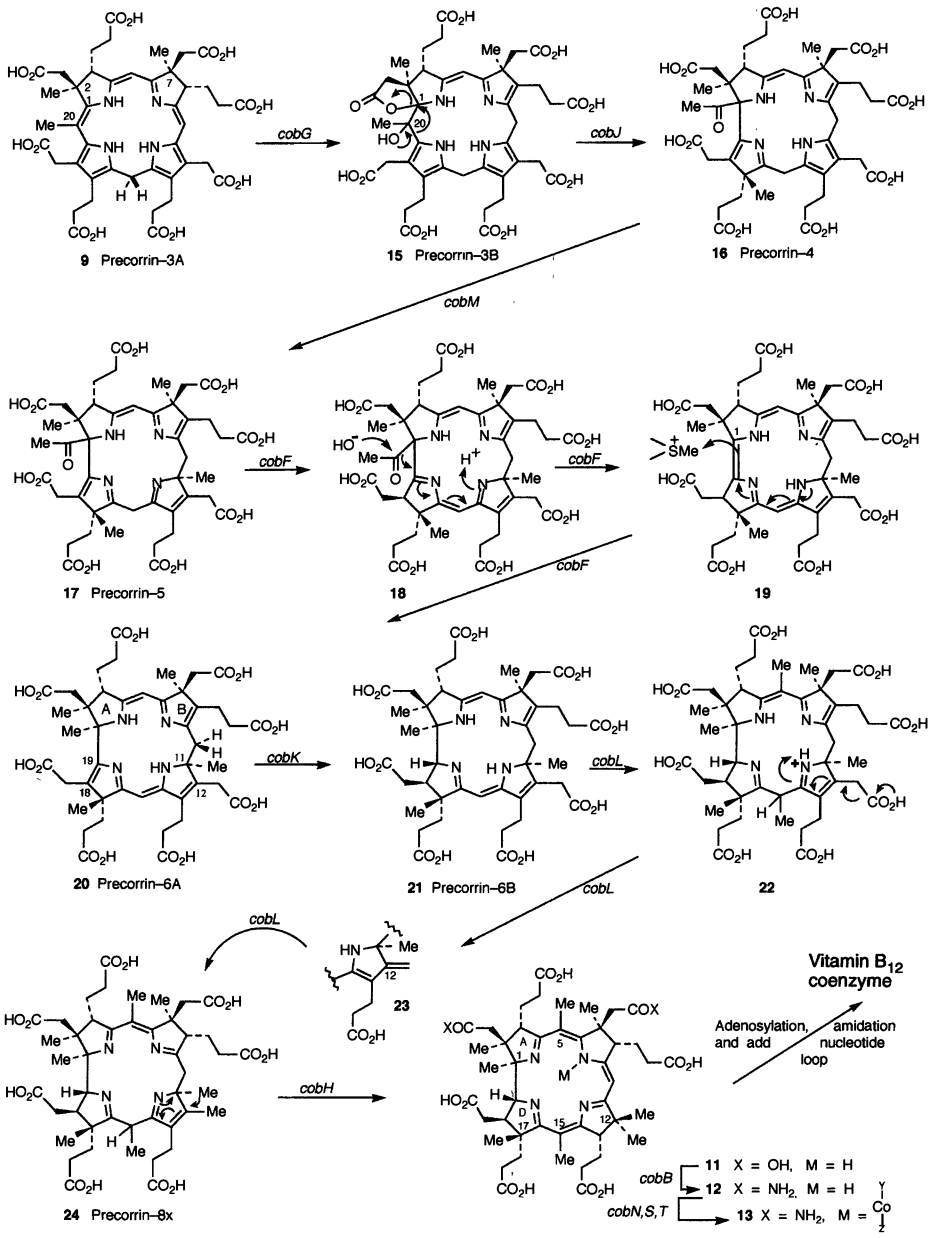


Fig. 2. B_{12} biosynthetic pathway from precorrin-3A through to the corrin macrocycle.

phyrinogen III (**6**) to hydrogenobyrinic acid (**11**) in *P. denitrificans* (*13, 15*) requires the enzymes encoded by nine *cob* genes (*cobA, F, G, H, I, J, K, L, M*). The results already outlined established the functions of three of the enzymes, encoded by *cobK, cobL*, and *cobH*. Also, the enzymes encoded by *cobA* and *cobI* were shown to act in the part of the pathway that goes from **6** to **9**. Though a full discussion of these enzymes is outside the scope of this article, it is nevertheless essential to know that *cobA* encodes the methyltransferase that sets in place the C methyl groups at C-2 and C-7 of precorrin-2 (**8**) (*32, 33*), whereas *cobI* encodes the enzyme responsible for the methylation at C-20, which generates precorrin-3 (**9**) (*34*). The four genes, *cobG, cobF, cobJ*, and *cobM*, must therefore code for the enzymes needed for conversion of precorrin-3 (**9**) into precorrin-6A (**20**). Also, the enzymes from *cobF, cobJ*, and *cobM* were known to be methyltransferases because of substantial sequence homology between these genes and the *cobA* and *cobI* genes (*15*). Thus, the three C methylations needed (with other chemistry as well) to convert precorrin-3 (**9**) into precorrin-6A (**20**) were well matched with the three methyltransferases available. Now, the remaining puzzle was this: What does the protein encoded by *cobG* do and where and in what order do the methylases from *cobF, cobJ*, and *cobM* carry out their task? The relatively easy part of setting these last few puzzle pieces in place depended entirely on having the rest of the picture established.

With only these four genes to study, a logical approach could be used that involved both gene deletion and gene overexpression. The first step forward came from a strain of *P. denitrificans* in which the *cobM* gene had been deleted. This strain afforded a protein extract that converted precorrin-3 into a new intermediate that proved to be tetramethylated and hence was named precorrin-4. This was isolated in pure form as a reasonably stable didehydroderivative called factor IV (*35*). Factor IV was reduced back to precorrin-4 when it was incubated with the complete enzyme system from *P. denitrificans* together with NADH, and precorrin-4 was further transformed into precorrin-6x (**20**), a known precursor of vitamin B_{12} (*35*). This firmly set precorrin-4 on the biosynthetic pathway to the vitamin. Once again, multiple ^{13}C labeling and NMR revealed the structure of factor IV as **16a**, and the structure of precorrin-4 to be **16** or possibly a tautomer (*35*).

The structure of precorrin-4 held more surprises as well. Ring contraction had already occurred at this early stage, and the location of the acetyl group was unexpected. The chemistry for ring contraction had pointed to the concomitant formation of an acetyl group (*36*) that could later be extruded as acetic acid, the established elimination product. But everyone's speculations before this work on precorrin-4 placed the acetyl group at C-19 rather than at C-1. Its location at C-1 was confirmed (*37*) by synthesizing precorrin-3 (**9**) with ^{13}C labels at both C-1 and C-20, followed by enzymic conversion into precorrin-4 (**16**). NMR analysis of the derived Factor IV **16a** demonstrated that the two ^{13}C labels were still directly bonded. Comparison of the structures of precorrin-3 (**9**) and precorrin-4 (**16**) shows that, to reach precorrin-4, both C-17 methylation and oxidation (adding one oxygen atom overall) had occurred. Things were locking neatly into place because an oxidative step must precede the formation of precorrin-6x (**20**).

That the fourth methyl group to be attached to the macrocycle appears at C-17 (*35, 37*) in precorrin-4 (**16**) was in full agreement with early pulse-labeling studies (*9*). The later pulse-labeling work showed that the 12α-methyl group of cobyrinic acid (**10**) (*10, 11*) and of hydrogenobyrinic acid (**11**) (*38*) is the fifth methyl to be added, though it is placed initially at C-11 (*19, 20*) [see precorrin-6A (**20**)]. As this fifth methylation does not occur when *cobM* is deleted, it follows that *cobM* encodes the 11-methyltransferase of *P. denitrificans*.

The enzyme encoded by *cobG* afforded yet another surprise. It was found that this enzyme alone transforms precorrin-3 (**9**) and carries out an oxidative step without methylation (*39, 40*). The new biosynthetic intermediate carries the same number of methyl groups as the starting material, previously called precorrin-3; this name thus had to be modified (*8*) to precorrin-3A. The oxidation product is precorrin-3B. Biotransformation of precorrin-3B into precorrin-4 (**16**) proved that it really is an intermediate in the biosynthesis of B_{12} (*40*). Pure samples of precorrin-3B were prepared in multiply labeled forms that allowed the structure for precorrin-3B (*41*) to be deduced (**15**) (*39, 40*). The conclusion was that the oxidative step does not cause the ring contraction; rather, it lays the foundation for contraction to be catalyzed by the next enzyme, the C-17 methyltransferase encoded by *cobJ*, which generates precorrin-4 (**16**). Remarkably, the purified oxidative enzyme does not use heme as a cofactor but has properties supporting the presence of iron-sulfur clusters (*40*). Some of the ^{13}C NMR signals corresponding to the published structure of precorrin-3B (**16**) have been observed (*42*) directly from the species produced by incubating ^{13}C-labeled precorrin-3A (**9**) and the *cobG* enzyme. This experiment was run in the presence of air, which was an important change because it allowed the enzyme, obtained by overexpression in *Escherichia coli*, to carry out more than one cycle (*43*). The sum of all the above evidence showed that the *cobG* enzyme is the initial oxidase and that the enzyme from *cobJ* methylates C-17 with ring contraction, whereas the *cobM* protein methylates C-11. This leaves the methyltransferase encoded by *cobF* to do the remaining necessary methylation at C-1.

When this C-1 methylation was blocked by omitting the protein encoded by the *cobF* gene from the cocktail of enzymes, a pentamethylated intermediate was produced, which is precorrin-5. It could be biosynthesized either from precorrin-3A, by using the *cobG, cobJ*, and *cobM* enzymes (*40*), or from precorrin-4, when just the *cobM* enzyme was needed (*44*). Multiply ^{13}C-labeled forms of precorrin-5 were generated by the latter approach to allow the structure **17** to be deduced by NMR (*44*). There are indications that the C-1 acetyl group, which must be enzymatically removed to allow C-1 methylation, is intrinsically labile. This acetyl group can be slowly lost nonenzymatically (*40, 44*), and the product of this loss has been isolated as its didehydro derivative, factor V (*40*).

The Genes and Their Enzymes, Including Cobalt Insertion

I have concentrated here on the biosynthesis of hydrogenobyrinic acid (**11**) from uroporphyrinogen III (**6**) in *P. denitrificans* because it is in this part of the pathway that all the excitement of a rapidly unraveling puzzle has been generated. In Table 1, I have also included information on four genes and their enzymes from another part of the B_{12} pathway. The enzyme encoded by *cobB* converts hydrogenobyrinic acid **11**

Table 1. The genes and enzymes for corrin biosynthesis in *P. denitrificans*.

Gene	Enzyme function
cobA	Methylation at C-2 and C-7 of uroporphyrinogen III (**6**)
cobI	Methylation at C-20 of precorrin-2 (**8**)
cobG	Oxidation of precorrin-3A (**9**)
cobJ	Ring contraction and C-17 methylation of precorrin-3B (**15**)
cobM	C-11 methylation of precorrin-4 (**16**)
cobF	C-1 methylation (and deacylation?) of precorrin-5 (**17**)
cobK	Reduction of precorrin-6x (**20**)
cobL	Methylation at C-5 and C-15 of precorrin-6y (**21**) and decarboxylation of C-12 acetate
cobH	Rearrangement of precorrin-8x (**24**) to hydrogenobyrinic acid (**11**)
cobB	*a,c*-Amidation of hydrogenobyrinic acid to give **12**
cobN	
cobS	Cobaltochelatase to give **13**
cobT	

into its *a,c*-diamide (**12**) (*45*) ready for cobalt insertion that is catalyzed by an enzyme complex encoded by three genes, *cobN*, *cobS*, and *cobT* (*46*). This latter step yields cobyrinic acid *a,c*-diamide (**13**). It was cobyrinic acid (**10**) that was used in early studies of B_{12} biosynthesis. The pathways in *P. denitrificans* and *P. shermanii* merge at this stage; the earlier parts of the two pathways differ at least in the timing of cobalt insertion (*47*, *48*). The various transformations that have been surveyed here evidently take place in *P. shermanii* on cobalt-containing macrocycles (*49*).

Viewing the Complete Picture

Now we can view the entire pathway of B_{12} synthesis from beginning to end, with the chemistry for each transformation at the critical stages preparing the ground for the next one—or perhaps more accurately, with the available intrinsic reactivity at a given stage predetermining the nature of the next step (*36*) as the biosynthetic pathway of B_{12} evolved.

Figure 1 shows how uroporphyrinogen III (**6**), the parent of the pigments of life, is channeled toward vitamin B_{12} by methylation at C-2 to produce precorrin-1 (**7**), the first intermediate, with further methylation at C-7 to generate precorrin-2 (**8**). The *cobI* enzyme then places a third methyl group at C-20, producing precorrin-3A (**9**). The biosynthetic steps downstream from precorrin-3A in *P. denitrificans* have been my main theme here (Fig. 2) (*50*). The *cobG* enzyme catalyzes the oxidative formation of precorrin-3B (**15**) from precorrin-3A (**9**). This enzyme does not bring about the ring contraction but instead simply sets the stage for this step, which is carried out by the *cobJ* enzyme together with methylation at C-17. The contraction process is illustrated as a pinacol-type rearrangement (*36*) on the γ-lactone, but it is possible that this step actually involves the δ-lactone (to C-20). These γ- and δ-lactones probably interconvert readily. The precorrin-4 (**16**) so formed then undergoes methylation at C-11 leading to precorrin-5 (**17**), which is converted into precorrin-6A (**20**). Satisfying mechanistic interpretations can be given of the deacetylation and C methylation processes required for this last transformation and are illustrated in Fig. 2 by structures **18** and **19**. Only a simple prototropic shift is needed for **17** to yield the more conjugated system **18**, which is ideally constituted to allow hydrolytic extrusion of the C-1 acetyl group by providing the necessary electron sink (*51*). Moreover, its product **19** is an extended enamine, thus providing reactivity for methylation at C-1 (*52*). In this way is generated precorrin-6A (**20**), the intermediate that set rolling the surge of experimentation leading to the discovery of B_{12}'s synthetic pathway.

It is at this point that the oxidation level is adjusted by NADPH-dependent enzymic reduction of the C-18–protonated form of precorrin-6A (**20**) (hydride delivery to C-19) to give precorrin-6B (**21**). Next comes the transformation by an enzyme that displays two activities, transmethylation and decarboxylation, perhaps as a result of ancestral gene fusion. Methylation is carried out at C-5 and C-15, setting in place the last of the eight required methyl groups, and it is the 12-acetate that is decarboxylated. The product is precorrin-8x (**24**). The order of the methylations and decarboxylation is not known, but a reasonable sequence is shown as methylation to give **22** and then forward to **23** followed by **24**. Before 1990, those who studied B_{12} had been puzzled about how the C-12 acetate was specifically decarboxylated. Because we know now that prior methylation occurs at C-11, a reasonable mechanism for decarboxylation is available. Precorrin-8x (**24**) is isomeric with hydrogenobyrinic acid (**11**); it carries the correct number of both methyl groups and double bonds. All that is needed is the 1,5-methyl shift from C-11 to C-12—rather like flicking a switch—to allow the characteristic conjugated system of the corrin macrocycle of hydrogenobyrinic acid to be generated. Amidation of the *a* and *c* acetic side chains then provides the substrate **12** for the complex cobaltochelatase, which slips in the cobalt ion to form cobyrinic acid *a,c*-diamide (**13**). More steps of amidation and others (outside the scope of this article) are needed to add the adenosyl residue and the nucleotide loop to generate the vitamin B_{12} coenzyme. As all of these steps are now understood, the elucidation of vitamin B_{12} biosynthesis is complete, marking the end of an epoch that started 39 years ago, when the structure of B_{12} was first solved (*53*) and chemists first wondered how this vitamin was built by living organisms.

REFERENCES AND NOTES

1. A. R. Battersby and F. J. Leeper, *Chem. Rev.* **90**, 1261 (1990).
2. A. R. Battersby and E. McDonald in *B_{12}*, D. Dolphin, Ed. (Wiley, New York, 1982), vol. 1, pp. 107–144; F. J. Leeper, *Nat. Prod. Rep.* **6**, 171 (1989).
3. A. R. Battersby, *Acc. Chem. Res.* **19**, 147 (1986).
4. A. I. Scott, *ibid.* **23**, 308 (1990); *Angew. Chem. Int. Ed. Engl.* **32**, 1223 (1993).
5. Intermediates that precede the formation of the corrin macrocycle are called precorrins. The attached number corresponds to the number of C methyl groups introduced from SAM into uroporphyrinogen III to form that intermediate. Letters are also added to the number to allow discrimination of different intermediates carrying the same number of SAM-derived methyl groups (*8*).
6. L. Mombelli, C. Nussbaumer, H. Weber, G. Müller, D. Arigoni, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 11 (1981).
7. A. R. Battersby, M. J. Bushell, C. Jones, N. G. Lewis, A. Pfenninger, *ibid.*, p. 13.
8. H. C. Uzar, A. R. Battersby, T. A. Carpenter, F. J. Leeper, *J. Chem. Soc. Perkin Trans. I* **1987**, 1689 (1987).
9. H. C. Uzar and A. R. Battersby, *J. Chem. Soc. Chem. Commun.* **1982**, 1204 (1982).
10. ———, *ibid.* **1985**, 585 (1985).
11. A. I. Scott *et al.*, *Bioorg. Chem.* **12**, 356 (1984).
12. I have concentrated here on the essential information needed to appreciate the transformation in knowledge achieved recently. In fact, many other questions about B_{12} biosynthesis were answered during the first phase of research (*2–4*).
13. B. Cameron, K. Briggs, S. Pridmore, G. Brefort, J. Crouzet, *J. Bacteriol.* **171**, 547 (1989).
14. J. Crouzet *et al.*, *ibid.* **172**, 5968 (1990).
15. J. Crouzet *et al.*, *ibid.*, p. 5980.
16. J. Crouzet *et al.*, *ibid.* **173**, 6074 (1991).
17. For the genetics of B_{12} production in *Salmonella typhimurium*, see J. R. Roth, J. G. Lawrence, M. Rubenfield, S. K. Higgins, G. M. Church, *J. Bacteriol.* **175**, 3303 (1993); in *Bacillus megaterium*, see (*54*).
18. D. Thibaut, L. Debussche, F. Blanche, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8795 (1990).
19. D. Thibaut, F. Blanche, L. Debussche, F. J. Leeper, A. R. Battersby, *ibid.*, p. 8800.
20. F. Blanche *et al.*, *J. Chem. Soc. Chem. Commun.* **1992**, 138 (1992).
21. Ready conversion of one double bond tautomer into another can occur in the macrocyclic B_{12} intermediates. Thus, precorrin-6A (**20**) may have a double bond at C-9 to N rather than at C-8 or C-9 as illustrated; these close tautomers are included when one refers to precorrin-6A. This also holds true for all the other structures discussed here.
22. For a related case, see W. G. Whittingham *et al.*, *J. Chem. Soc. Chem. Commun.* **1989**, 1116 (1989), and B. Müller *et al.*, *ibid.*, p. 1119.
23. F. Blanche *et al.*, *J. Bacteriol.* **174**, 1036 (1992).
24. D. Thibaut *et al.*, *J. Chem. Soc. Chem. Commun.* **1992**, 139 (1992).
25. G. W. Weaver *et al.*, *ibid.* **1991**, 976 (1991).
26. F. Kiuchi *et al.*, *ibid.* **1992**, 306 (1992).
27. K. Ichinose, F. J. Leeper, A. R. Battersby, *ibid.* **1993**, 515 (1993).
28. F. Blanche *et al.*, *J. Bacteriol.* **174**, 1050 (1992).
29. A plausible explanation of the ability of one enzyme to catalyze two such different reactions as methyl transfer and decarboxylation is that *cobL* could have arisen by fusion of two ancestral genes, each initially coding for an enzyme having one activity.
30. D. Thibaut *et al.*, *J. Bacteriol.* **174**, 1043 (1992).
31. D. Thibaut *et al.*, *J. Chem. Soc. Chem. Commun.* **1992**, 982 (1992).
32. F. Blanche, L. Debussche, D. Thibaut, J. Crouzet, B. Cameron, *J. Bacteriol.* **171**, 4222 (1989).
33. There has been considerable interest in the equivalent enzyme from other sources [M. Warren, in *The Biosynthesis of the Tetrapyrrole Pigments* (Ciba Foundation Symposium 180, Wiley, New York, 1994), p. 26; A. I. Scott, *ibid.*, p. 285].
34. D. Thibaut *et al.*, *J. Bacteriol.* **172**, 6245 (1990).
35. D. Thibaut *et al.*, *J. Chem. Soc. Chem. Commun.* **1993**, 513 (1993).
36. A. Eschenmoser, *Angew. Chem. Int. Ed. Engl.* **27**, 6 (1988).
37. A. I. D. Alanine *et al.*, *J. Chem. Soc. Chem. Commun.* **1994**, 103 (1994).
38. F. Blanche *et al.*, *Angew. Chem. Int. Ed. Engl.* **29**, 884 (1990).
39. L. Debussche *et al.*, *J. Chem. Soc. Chem. Commun.* **1993**, 1100 (1993).
40. L. Debussche *et al.*, *J. Bacteriol.* **175**, 7430 (1993).
41. The possibility must be borne in mind that precorrin-3B (**15**) is initially generated as the γ-lactone (to C-20) but that, on keeping after release from the enzyme, it rearranges to the δ-lactone (to C-1).
42. A. I. Scott *et al.*, *FEBS Lett.* **331**, 105 (1993).
43. Forward transformation of the enzymic product to some established B_{12} intermediate was not carried out in this case, but this necessary step with precorrin-3B (**15**) had already been made (*39*).
44 C. Min *et al.*, *J. Am. Chem. Soc.* **115**, 10380 (1993).

45. L. Debussche, D. Thibaut, B. Cameron, J. Crouzet, F. Blanche, *J. Bacteriol.* **172**, 6239 (1990).
46. L. Debussche *et al.*, *ibid.* **174**, 7445 (1992).
47. G. Müller *et al.*, *J. Am. Chem. Soc.* **113**, 9893 (1991).
48. S. Balachandran *et al.*, *J. Chem. Soc. Perkin Trans. I.* **1994**, 487 (1994).
49. F. Blanche *et al.*, *Angew. Chem. Int. Ed. Engl.* **32**, 1651 (1993).
50. For details on the labeling methods used to elucidate the structures of the new intermediates in Fig. 2, see A. R. Battersby, *Acc. Chem. Res.* **26**, 15 (1993).
51. D. Arigoni, in *The Biosynthesis of the Tetrapyrrole Pigments* (Ciba Foundation Symposium 180, Wiley, New York, 1994), p. 280.
52. A. R. Battersby, *ibid.*, p. 267.
53. D. C. Hodgkin, A. W. Johnson, A. R. Todd, *Chem. Soc. Spec. Publ.* **3**, 109 (1955); D. C. Hodgkin *et al.*, *Nature* **178**, 64 (1956).
54. J. B. Wolf and R. N. Brey, *J. Bacteriol.* **166**, 51 (1986); R. N. Brey, C. D. B. Banner, J. B. Wolf, *ibid.* **167**, 623 (1986).

55. I wish especially to thank F. Blanche, B. Cameron, J. Crouzet, L. Debussche, D. Thibaut, and their colleagues at Rhône-Poulenc-Rorer, Paris, for the marvellous collaboration we in Cambridge have enjoyed with them. I am also most grateful to my senior colleagues F. J. Leeper, N. P. J. Stamford, and C. Abell for their invaluable help and counsel and to all other members of our Cambridge group named in the references for their many important contributions. Finally, I thank Zeneca, Hoffmann–La Roche, Roche Products, the Leverhulme Trust, and SERC for financial support.