HOW NATURE SYNTHESIZES B_{12} WITHOUT OXYGEN. DISCOVERIES ALONG THE ANCIENT, ANAEROBIC PATHWAY†

A. I. Scott

Department of Chemistry, Texas A&M University
College Station, Texas 77843, U.S.A.

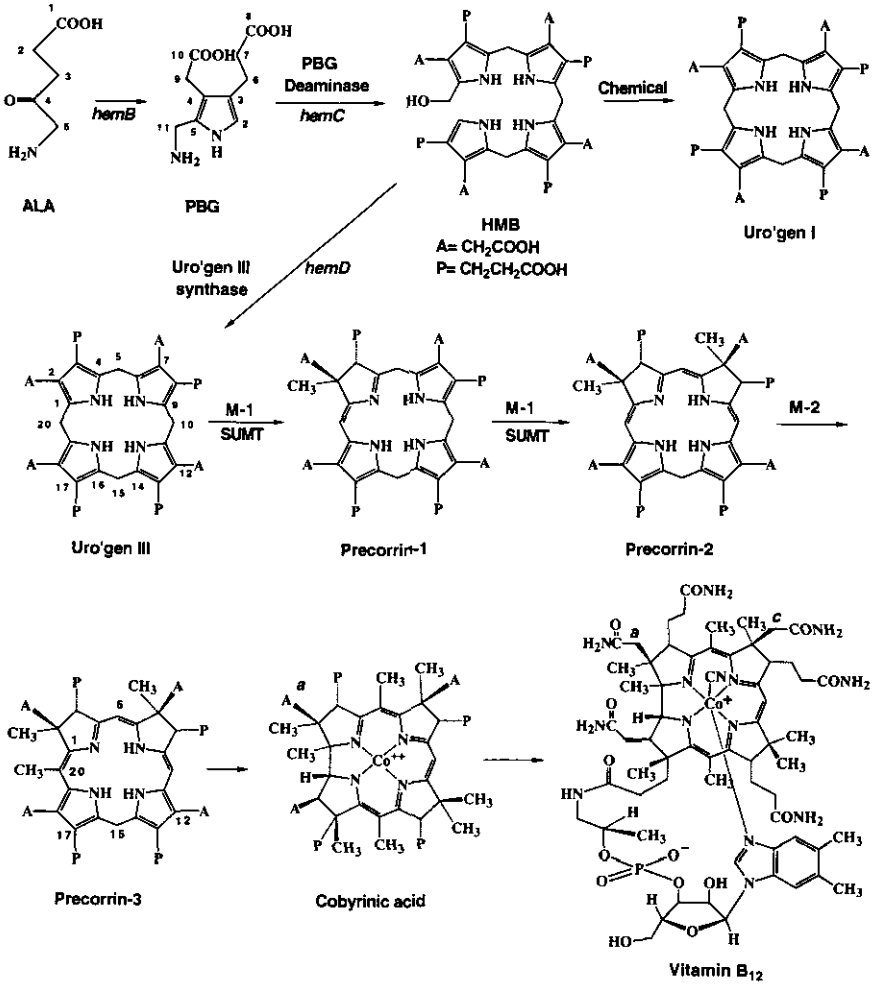
Abstract - A second pathway to vitamin B_{12} has been discovered in anaerobic and microaerophilic organisms such as *Propionibacterim shermanii*. Although running parallel to the biosynthetic route established in the aerobe, *Pseudomonas denitrificans*, the anoxic pathway which has existed for *ca.* $4x10^9$ years exhibits many differences from the aerobic sequence, including early insertion of cobalt, a ring contraction mechanism which does not use oxygen, and the extrusion of acetaldehyde (rather than acetic acid) as the two carbon fragment. The main features of both biosynthetic routes to B_{12} are compared and contrasted.

1 Introduction

Previous studies have defined the biosynthesis of vitamin B_{12} in the aerobic organism *Pseudomonas denitrificans*.[1,2] In order to set the stage for this account of the discovery of a parallel yet frequently distinct pathway to B_{12} in anaerobic or semi-anaerobic bacteria, some overlap with the aerobic pathway is not only inevitable but vital to the understanding of the main differences between the sequences which occur in the "central" region between precorrin-2 and cobinamide, and, in particular, how nature arranges ring contraction, loss of the C-2 unit and cobalt insertion *in the absence of oxygen*. Having established the connection between urogen III and B_{12} in 1972, work over the next seven years had led to the "library" of intermediates shown in Scheme 1. In spite of prodigious efforts by all of the participating research groups,[1,2] no further intermediates beyond precorrin-3 came to light between 1979 and 1990. By adding the techniques of molecular biology to our repertoire we were ready (by 1988) to move towards the goal of discovering the missing intermediates on the road to corrins, by matching each gene product to its biosynthetic function.

†Dedicated with respect to Professor Koji Nakanishi on the occasion of his 75th birthday and wishing him many more happy returns.

Scheme 1



2 The early methyl transferases

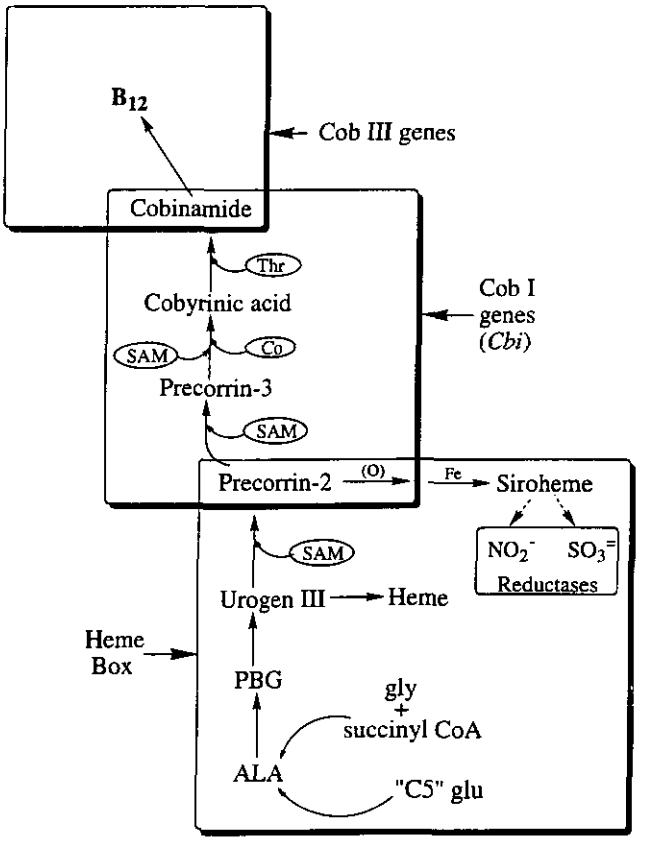
The first 3 enzymes of the pathway, overexpressed in *E. coli* (*hemB, C, D*) were combined to prepare substantial amounts of uro'gen III from ALA.[3] Next, in order to enter the chiral world of corrins, uro'gen III has to be C-methylated twice by the same enzyme, SAM-uro'gen III methyl transferase (SUMT), which also bears strong homology to *cysG* from *E.coli.*

2.1 Cys G and SUMT

SUMT was first partially purified from *P. shermanii* by G. Müller[4] and has been overexpressed in *Pseudomonas denitrificans* as CobA.[5] In *E. coli* it was found that the *CysG* gene encodes Uro'gen III methylase (M-1) as part of the synthetic pathway to siroheme, the cofactor for sulfite reductase, and overproduction (30 mg/L) was achieved by the appropriate genetic engineering.[6] Although SUMT and M-1 appear to perform the same task, it has been found that their substrate specificities differ. In fact, the enzyme CysG turned out to be multi-functional, (see "heme box", Scheme 2) catalyzing the complete

synthesis of siroheme via the NAD mediated dehydrogenation of precorrin-2, followed by insertion of Fe^{2+}, and so has been renamed siroheme synthase.[7]

Scheme 2



2.2 The Salmonella genes

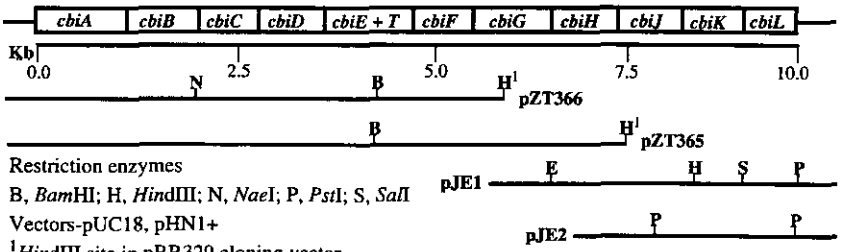
Fortunately, the heme and corrin pathways intersect in *E. coli* and *Salmonella typhimurium*. Although the former organism does not make B_{12}, the discovery by J. Roth[8] that anaerobic fermentation of *S. typhimurium* produces vitamin B_{12} allowed the vast array of genetic and cloning techniques available with this organism to be used in the search for the B_{12} pathway. Three *loci* at minutes 14, 34 and 42 have been identified by mutation and complementation studies.[9] The main gene cluster at 42 min contains all of the DNA (Cob I) necessary for the synthesis of cobinamide from precorrin-2 (Scheme 2), a process involving 6 C-methylations (at C-1, C-5, C-11, C-15, C-17 and C-20), decarboxylation (of the acetate residue at C-12), ring contraction, loss of the equivalent of acetic acid (from C-20 and its attached methyl), amidation and cobalt insertion under control of the *cbi* genes. We were able to use the 10 kb sequence data provided by J. Roth to clone and overexpress the gene products, corresponding to 12 open reading frames (ORF's) necessary to synthesize cobyrinic acid from precorrin-2 (Figure 1).

Ten of the *cbi* genes found in the *S. typhimurium cob* operon were subcloned for expression from the 4 different plasmids. The gene products of *cbiE*, *cbiF*, *cbiH* and *cbiL* were shown to be SAM binding proteins, and based on their homology with other methyltransferases, were considered to be the most likely candidates for methyltransferase activity. Meanwhile the Rhone-Poulenc group reported that the *P. denitrificans cobL* gene product[2] had two functions, methylation (of C-5 and C-15) and decarboxylation of the ring C acetate. SDS-PAGE and NH_2-terminal sequence analysis revealed that two separate gene

products in *S. typhimurium* (*cbiE* and *cbiT*) correspond to the *cobL* gene product with *cbiE* homologous to the methyltransferase region and *cbiT* homologous with the decarboxylase region, i.e. in *Salmonella* the 5,15-methylase and decarboxylation activities are separated.

From its homology with the *P. denitrificans cobI* gene product (31% identity, 71% conservation), the *cbiL* gene product was predicted to be the *S. typhimurium* precorrin-2 methyltransferase (M-2), and using this expressed protein, (or more efficiently the Cob I enzyme from *P. denitrificans*), the multi-enzyme one flask synthesis of precorrin-3 from the building block ALA was accomplished as shown in Scheme 3, both in the NMR tube and preparatively (on the fifty-milligram scale), by adding the 5 overexpressed enzymes to the substrate ALA in the presence of SAM.[10,11] The structure of precorrin-3 could then be studied in detail and revealed subtle differences in the ^{13}C-NMR spectrum which reflected the influence of the new methyl group at C-20 on the conjugated system, resulting in a preponderance of the tautomer shown, whose electronic array is prepared for the next C-methylation step.

As a guide to the anticipated order of insertion of the remaining methyl groups on the periphery of this last intermediate, precorrin-3, we recall earlier pulse-labeling experiments from three research groups (Cambridge, Paris, and Texas-Stuttgart) whereby the substrates uro'gen III and precorrins 2 and 3 (in their oxidized forms) were incubated with SAM for several hours with a cell free extract capable of synthesizing cobyrinic acid followed by a pulse of labeled [$^{13}CH_3$] SAM. By examining the different intensities of the ($^{13}CH_3$)-methyl resonances in the resultant cobester the **sequence** of methyl group insertion was found to correspond to C-methylation at C-17, followed by C-12 (now known to occur first at C-11, *vide infra*) then at C-1, C-5 and C-15.[12-14] There is thus agreement from three independent studies that C-17 is the first site of alkylation on the precorrin-3 template and this is followed by C-methylation at C-11, then at C-1, imposing certain restrictions on the type of structures expected for the missing intermediates precorrins 4, 5 and 6, corresponding to methyl insertion at C-17, C-11 and C-1 respectively.

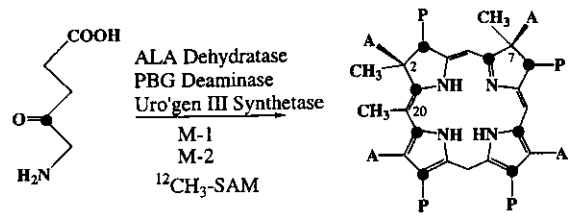


Genes	Plasmid	Construction
cbiC&D	pCAR345	pZT366 (N-B) into pHN1+
cbiE	pCAR340	ecpcr[3] product into pHN1+
cbiT&F	pCAR276	pZT366 (B-H) into pUC18
cbiT,F,G,[4]*H*	pCAR293	pZT365 (B-H) into pUC18
cbiH&J[4]	pCAR309	pJE1 (E-H) into pUC18
cbiK&L	pCAR292	pJE1 (P-P) into pUC18
cbiK	pCAR311	pJE2 (S-P) into pUC18
cobF[2]	pCAR332	ecpcr[3] product into pHN1+
cobI[2]	pCAR333	ecpcr[3] product into pHN1+

[2]*P. denitrificans* gene
[3]Expression cassette polymerase chain reaction
[4]Form insoluble inclusion bodies

Figure 1: Plasmids used to identify and express the open reading frames (ORF's) of the *cbi* genes of *S. typhimurium* encoding B_{12} synthetic enzymes.

Scheme 3



In order to begin the task of discovering the remaining precorrin intermediates, precorrin 3 was incubated in turn with each of the putative, overexpressed methyltransferases from the genes *cbiE*, *H*, *F* and *L* and SAM. A surprising result was obtained. The only C-methylation observed was that catalyzed by ORF-7 (*cbiF*) and turned out to be methylation at C-11! The new isolate (Figure 2), is a modified corphin bearing a **fourth** methyl group at C-11. Since biochemical conversion to cobyrinic acid was not demonstrated, we named the new compound **(4x)**[15] whose structural variant recalls one of the 4 possible structures proposed for Factor S_3, an unusual tetramethylated zinc complex based on the type -I (symmetrical) porphyrin structure.

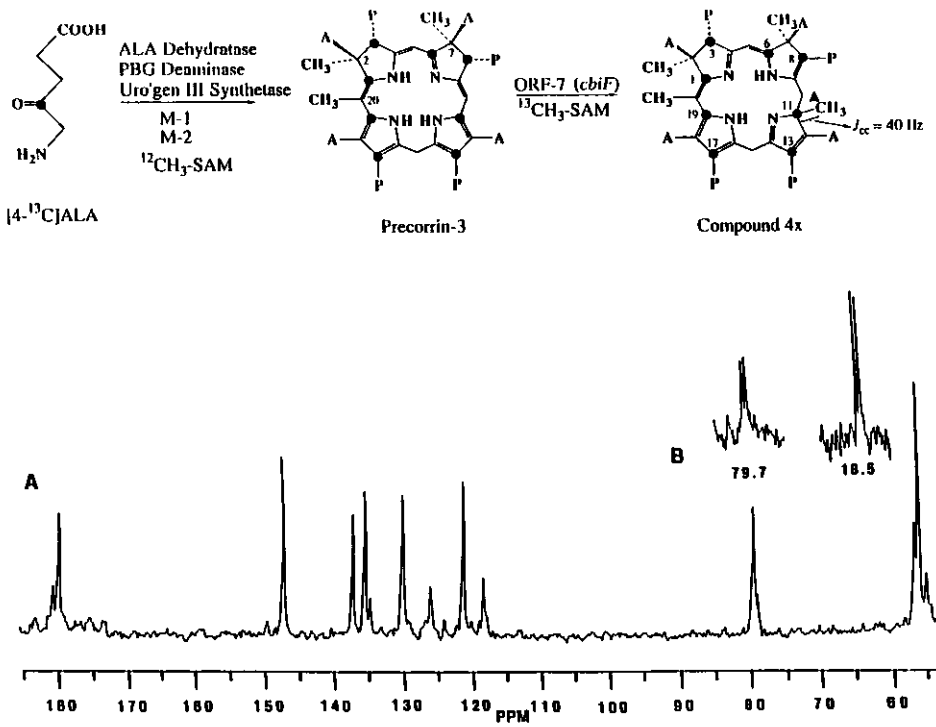
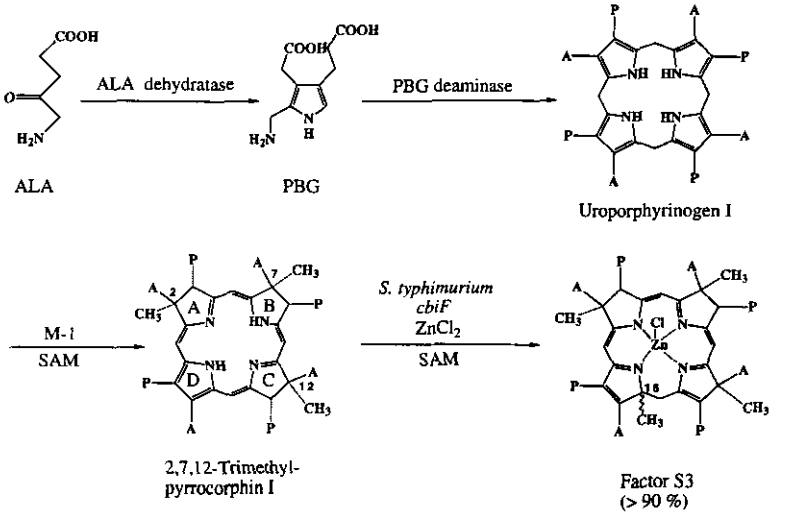


Figure 2: Multi-enzyme synthesis of compound 4x with the *cbiF* gene product and SAM. The ^{13}C-NMR spectrum of 4x (**A**) reveals 3 sp^2 (C- 8, C-13 and C-17) but only one sp^2 propionate terminus (C-3) and a signal at δ79.7 typical of sp^3 carbon (C-11) adjacent to nitrogen. The inset (**B**) shows the coupling (Jcc = 40 Hz) of the new $^{13}CH_3$ (d18.5) to the C-11 signal at δ79.7.

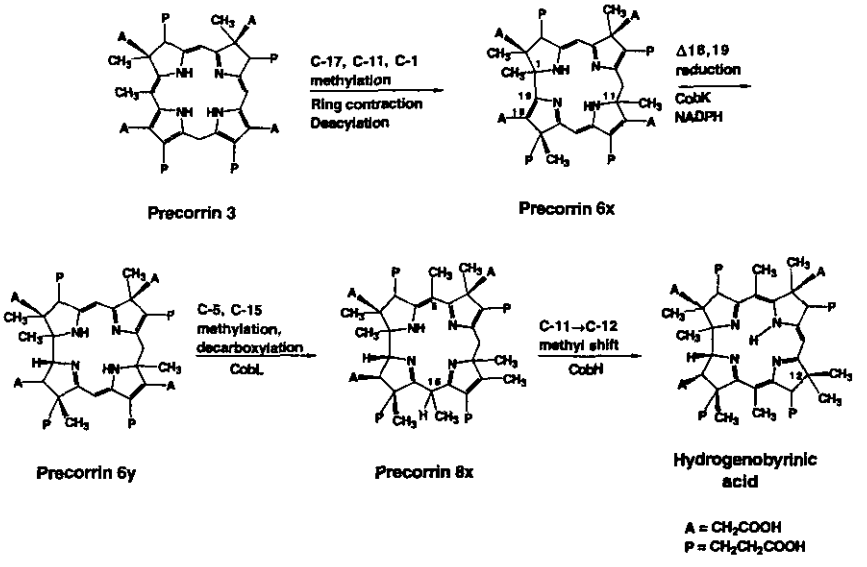
The spectroscopic data for the latter metabolite isolated from *P. shermanii* could not distinguish between methylation at C-1 or C-11 but it is now clear that the correct isomer is the one in which 11 (≡16) α methylation has taken place on the uro'gen I template. This new structural proposal for factor S_3 was

nicely confirmed by a second multi-enzyme synthesis, this time using [4-^{13}C]ALA as substrate in a one flask reaction mixture containing dehydratase, deaminase (→uro'gen I) *cysG*, *cbiF* and SAM (Scheme 4). In the absence of the last of these enzymes (*cbiF*) the product 2,7,12-trimethylpyrrocorphin accumulates, but when the fourth enzyme (*cbiF*) was added, a new signal appeared at δ 79 heralding the insertion of a fourth (α-) methyl group on the C-16 position (Scheme 4). The resultant zinc complex (as octamethyl ester) was identical in every respect with factor S_3 isolated earlier from *P. shermanii*.[16] It only remains to define the absolute stereochemistry of the new chiral center at C-16. These findings suggest a lack of substrate specificity in C-methylation by the methyl transferases and provide confirmation of the function of *cbiF* as a methyl transferase which attacks the α-position of pyrrole rings of pyrrocorphins, whether type III or type I, as well as opening the door to synthetic chemistry based on C-methylation at electron rich centers.

Scheme 4



Scheme 5



3 Some comparisons with the aerobic pathway

At this stage, in 1990, the Rhone-Poulenc scientists in collaboration with A. R. Battersby's group in Cambridge, using engineered *P. denitrificans*, discovered three late intermediates, precorrins-6x, 6y, and 8x[17-19] (Scheme 5) whose structures[20,21] were most informative about the timing of ring contraction and, most interestingly, revealed a final [1,5]-sigmatropic shift which moves the methyl at C-11 in precorrin-8x to its final resting place (C-12) in the corrin structure of hydrogenobyrinic acid (HBA).

Still hidden from view, however, was the nature of the pivotal processes whereby the porphyrinoid ring is contracted at some stage during the conversion of precorrin-3 to -6x. Ever since the discovery of the porphyrinoid-corrin connection, we had always regarded the solution of the ring contraction process as the central problem in B_{12} biosynthesis.

At this juncture, to the sequences of the genes and the necessary plasmids responsible for corrin synthesis from the anaerobic pathway in *Salmonella typhimurium* had been added the corresponding DNA sequences from the aerobic B_{12} producer *P. denitrificans*,[23] thus allowing the cloning and overexpression in *E. coli* of the two sets of gene products corresponding to the open reading frames shown in Figure 3.[15] We were thus in the most unusual situation in having all of the biochemical machinery necessary for the synthesis of corrins in hand without knowing the structures of the intermediates between precorrin-3 and 6x (in the aerobic series) (Scheme 5) or **any** of the intermediates between precorrin-3 and cobyrinic acid in the anaerobic *S. typhimurium*. Parallel studies on both sets of gene products were therefore initiated. For the anaerobic pathway, our experience with cell free extracts of *P. shermanii* served as a guide, for we had already shown (with G. Müller)[24] that cobalt is inserted early, in this pathway, at the stage of precorrin-2, whereas in the aerobe *P. denitrificans*, cobalt is not inserted until **after** the formation of the complete corrin, hydrogenobyrinic acid (Scheme 5). So it was already clear that profound differences existed between the aerobic and anaerobic pathways, but due to the complications arising from the necessity of learning to work entirely with cobalt complexes (and the attendant problems of Co^{II} NMR spectroscopy), we concentrated first on the metal-free route to corrins, i.e. the aerobic pathway.

3.1 The steps between precorrin-3 and precorrin-6x in the aerobic *P. denitrificans*

We had earlier discovered the function of CbiF (i.e. C-11 methylation, *vide supra*) but this enzyme (from *Salmonella*) had been presented with precorrin-3 so had actually handled a close relative, rather than the natural substrate, to make compound 4x (Figure 2). In other words, a methyl group should already be at C-17 before C-11 methylation takes place. To make matters worse, in spite of many months of intensive search, no enzyme activity of any of the remaining *cob* genes could be demonstrated. It then became clear that an essential cofactor was missing from our incubations. It turned out that our transatlantic competitors in the B_{12} gene hunt, the research group of Rhone-Poulenc Rorer had assumed (reasonably) that completely anaerobic conditions were necessary for enzymatic transformation through the processes leading from precorrin-3 to 6x and beyond. When factor IV (oxidized precorrin-4) was discovered by the French workers[25] it was immediately apparent that the ring contraction process must be **oxidative** in nature. We therefore revisited our collection of gene products and found that the enzyme CobG, which had always been tested *in the absence of air*, served as an excellent catalyst, (and only in the presence of O_2 and NADH) not for the ring contraction as anticipated, but for the insertion of one atom of dioxygen into the 20 position of precorrin-3, concomitant with participation of the ring A acetate in γ-lactone formation terminating at C-1 (scheme 6) to yield precorrin-3x.[26] It was then possible to go forward[1]

from 3x to the ring-contracted precorrin-4, a process which is formally seen as an acyloin reaction, triggered by methylation at C-17, and catalyzed by CobJ, thus discovering Nature's wonderful spring-loaded device for oxidative ring-contraction (Scheme 6). As described in the previous lectures, quite independently the Rhone-Poulenc group discovered and characterized precorrin-3x, which they named precorrin-3B.[27] We found that O_2 is required as part of the catalytic machinery and that one atom of $^{18}O_2$ is incorporated at C-20 in the product, precorrin 3x (Scheme 6), ultimately finding its way into the carbonyl of the methyl ketone in the ring contracted precorrin-4, as shown by the isotopic shift of ^{18}O on the ^{13}C resonances of C-20 in precorrin-3x and 4. The next step, between precorrin-4 and 6x could then be shown to involve C-11 methylation, mediated this time by CobM (≡CbiF) in the presence of SAM, to reach the last of the missing intermediates, precorrin-5[28] (Scheme 7).

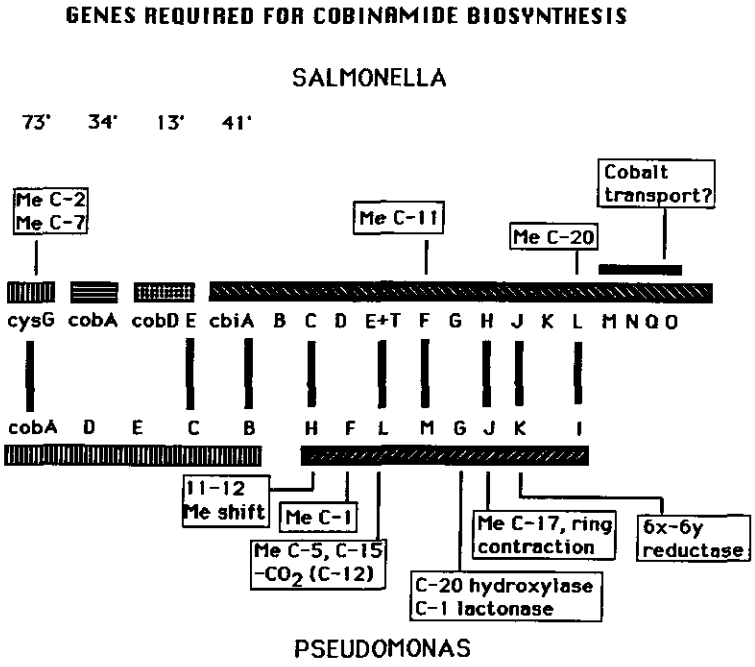
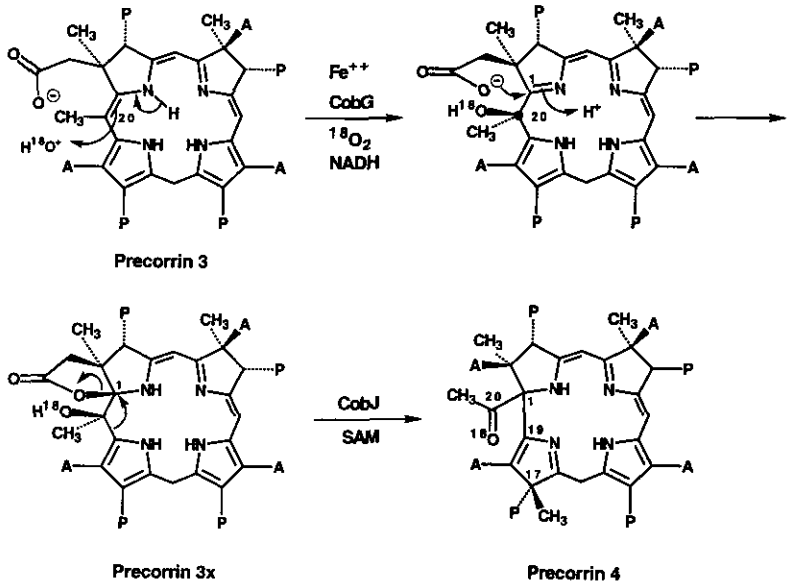


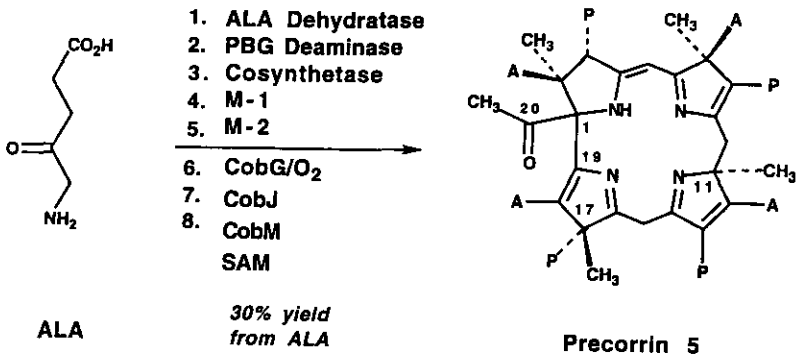
Figure 3: The location and functions of some of the known genes for B_{12} biosynthesis in *Salmonella typhimurium* and *Pseudomonas denitrificans*. In *S. typhimurium* the genes map at 14' (*cobD*, *cobE*; addition of aminopropanol), 34' (*cobA*; adenosylation), 41' (*cbiA-R*; cobinamide biosynthesis), and 73' (*cysG*; uro'gen III methyltransferase). Homologies between the *S. typhimurium* and *P. denitrificans* gene products are shown. See the text for a discussion of their functions.

The new isolate, precorrin-5, undergoes facile tautomerism, and isomerizes on esterification as seen from the NMR data of its octamethyl ester, revealing a close similarity in methylation pattern with that of factor S_3 in that a fourth methyl group is inserted into the a-pyrrolic position and remains in place until the penultimate step (Scheme 8). Thus it appears likely that the genes *cobM* and *cbiF* (which show considerable sequence homology) catalyze the same reaction in their respective pathways although we recall that *cbiF* "mis-methylated" the *wrong* substrate precorrin-3, in forming 4x.

Scheme 6



Scheme 7



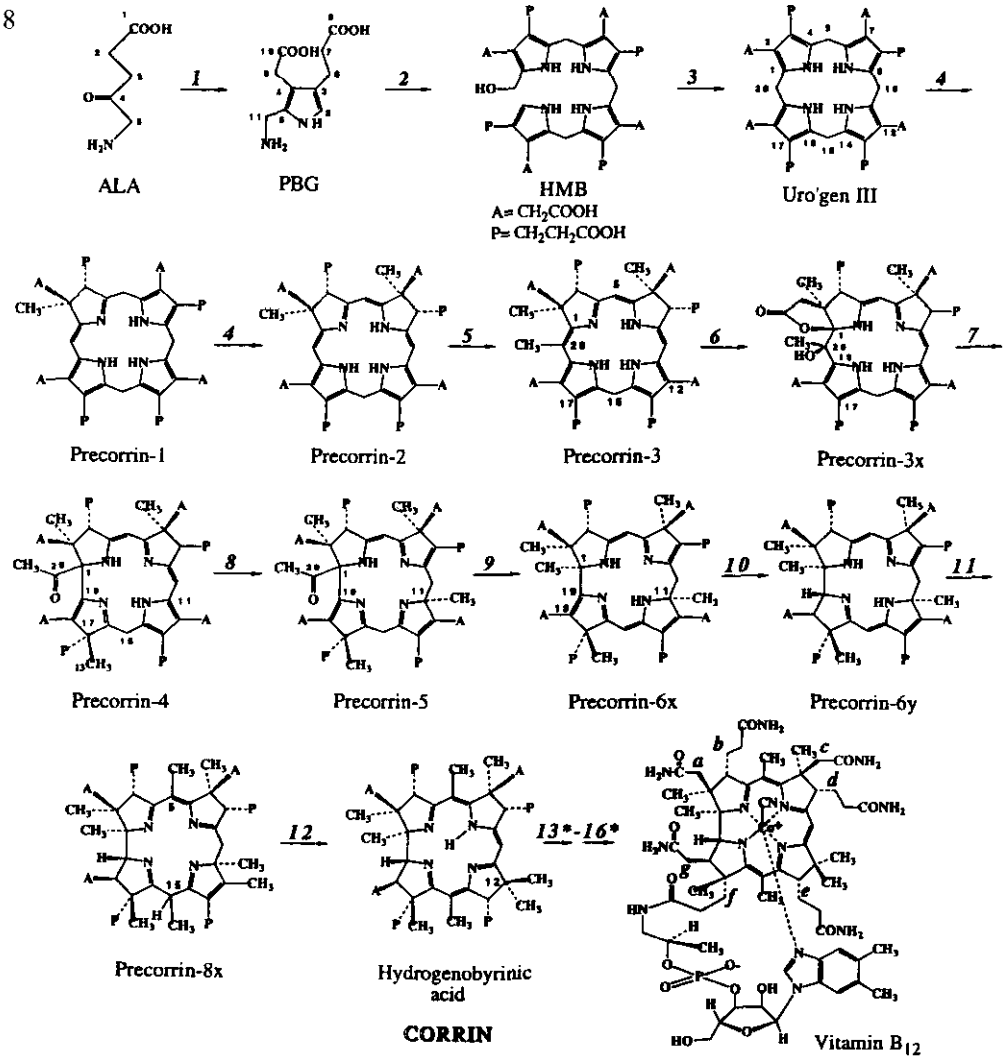
Finally CobF was found to be yet another bifunctional enzyme, which catalyzes both the deacylation and the methylation at C-1 of precorrin-5 in the presence of SAM. Esterification of the product provided a sample with NMR and FAB - MS data identical with those of precorrin-6x octamethyl ester.[29] This biosynthetic correlation (Scheme 8) confirms the structure and defines the absolute stereochemistry (except at C-1) of precorrin-5. It is important to stress that, in all of these precorrins, tautomerism plays a major role in their intrinsic chemistry. In other words, the structures elucidated by NMR after isolation may correspond to the most stable tautomer which, as was found for precorrin-8x, **may not be the true biosynthetic intermediate**.

3. 2 Multi-enzyme synthesis of corrins using the *Cob* genes

With all of the gene products in hand as single enzymes[30] we were ready to test the most advanced version of our multi-enzyme synthesis by combining all 12 gene products to synthesize corrin. Although optimistic about the outcome, we were concerned that previous attempts to reconstruct the bio-conversion of precorrin-3 to precorrin-6x, using four purified enzymes (CobG,J,M,F) proceeded in low,

poorly reproducible yields.[31] This is in sharp contrast to the incubation of precorrin-3 with a cell free extract of the engineered strain of *P. denitrificans*, afforded precorrin-6x in 20-40% yield using conditions which contained enough oxygen to catalyze the CobG reaction. In spite of the possibility that unfavorable substrate/product concentrations could have inhibited several of the enzyme catalyzed steps, to our great satisfaction, the overexpressed enzymes in combination behaved towards substrates and cofactors *in vitro* just as they do in the living cell. It was an exciting moment to see the final corrin target, HBA, being reached before our eyes in 20% overall yield, based on the 5-carbon starting material, ALA, when all 12 enzymes were added to a single flask containing ^{13}C-ALA, SAM, NADPH (in buffer, pH7) under aerobic conditions for 12h at 30 ℃ (steps 1-12; Scheme 8). The resultant ^{13}C-enriched HBA showed the expected set of 8 signals in the ^{13}C-NMR spectrum (corresponding to incorporation of 8 units of [4-^{13}C]-ALA and, for final characterization, was converted by chemical insertion of cobalt to cobyrinic acid showing complete identity (NMR, CD, MS) with the natural product. The latter serves as a relay to B_{12} since it has already been converted in *ca.* 40% yield to the vitamin. Thus the chemo-enzymatic synthesis of vitamin B_{12} was completed in two phases - a 17-step, 12-enzyme synthesis of the relay HBA[30] and a 4 step chemical conversion of the latter to cyanocobalamin, already in place thanks to the work of Müller[32] and Eschenmoser.[33]

Scheme 8



4 Towards the Elucidation of the Anaerobic Pathway.

4.1 Cloning the *P. shermanii* genes

Using the genes from *P. shermanii* together with the complete set of B_{12} biosynthetic genes already cloned from *Salmonella* we are now beginning to solve the mysteries of the anaerobic pathway, whose intermediates, from precorrin-3 onwards, are cobalt complexes and whose structures and mechanisms of formation can be predicted to open up a whole new sequence of events, perhaps even more surprising than the aerobic pathway discussed above. The current status of the sequencing of plasmids isolated from the *P. shermanii* gene bank are shown in Figure 4 and compared with the known ORF's from *Salmonella* and *P. denitrificans*. It is already clear that there are several "orphan" genes e.g. *CbiD* in *Salmonella* whose sequences correspond to the new *P. shermanii* ORF's. It is also of interest that *CobG* (aerobic) has no counterpart in *Salmonella*. The next task is to find the ring contraction mechanism in the anaerobic series and current work is concentrated on this part of the pathway, as follows.

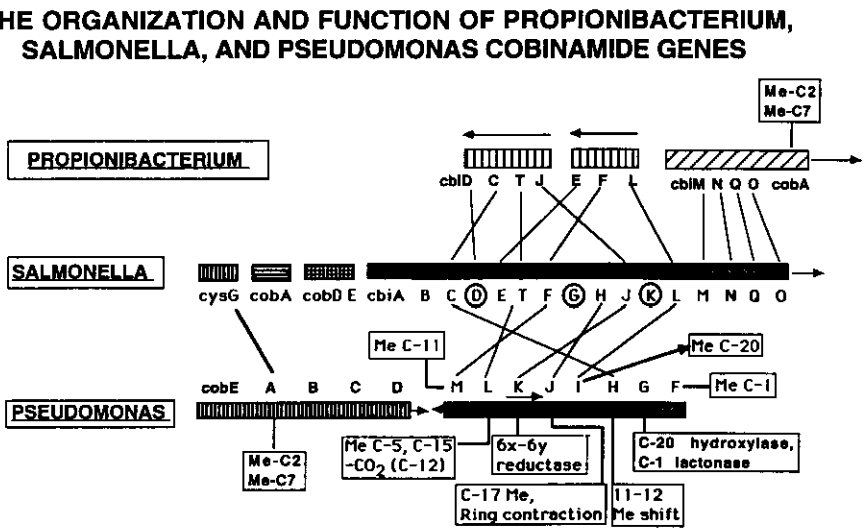


Figure 4: Organization of the aerobic (*P. denitrificans*) and anaerobic (*P. shermanii*, *S. typhimurium*) genes.

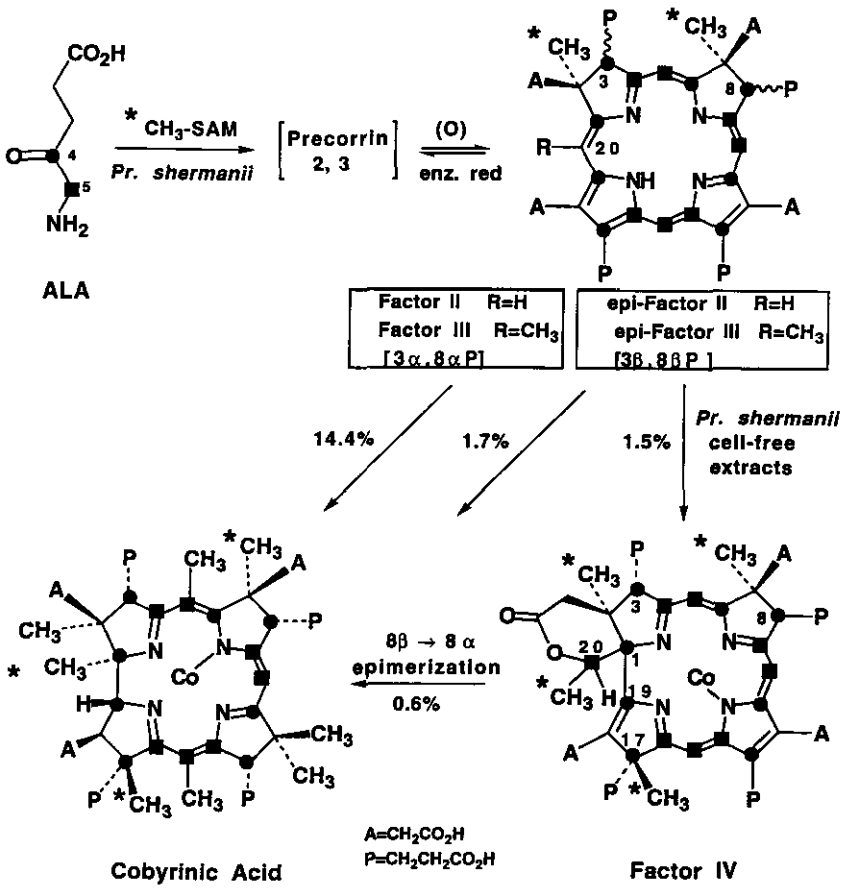
4.2 Oxygen exchange

It was possible to trace both the source and the fate of the oxygen functionalities involved at or near the site of ring contraction in both pathways. Whereas the C-20 hydroxyl in precorrin-3x is derived from one atom of dioxygen shown by ^{18}O labeling (Scheme 8) and is eventually lost in the acetic acid excised from C-1 of precorrin-5, the oxygen at C-1 in precorrin-3x stems from the original ringA-acetate carboxyl and is retained in this carboxyl group throughout all of the subsequent rearrangements and processes culminating in the structure of hydrogenobyrinic acid.[34,35] In contrast, it was found (with M. Kajiwara)[36] that in *P. shermanii* a regiospecific exchange (with water from the medium) of a unique carbonyl function, again the

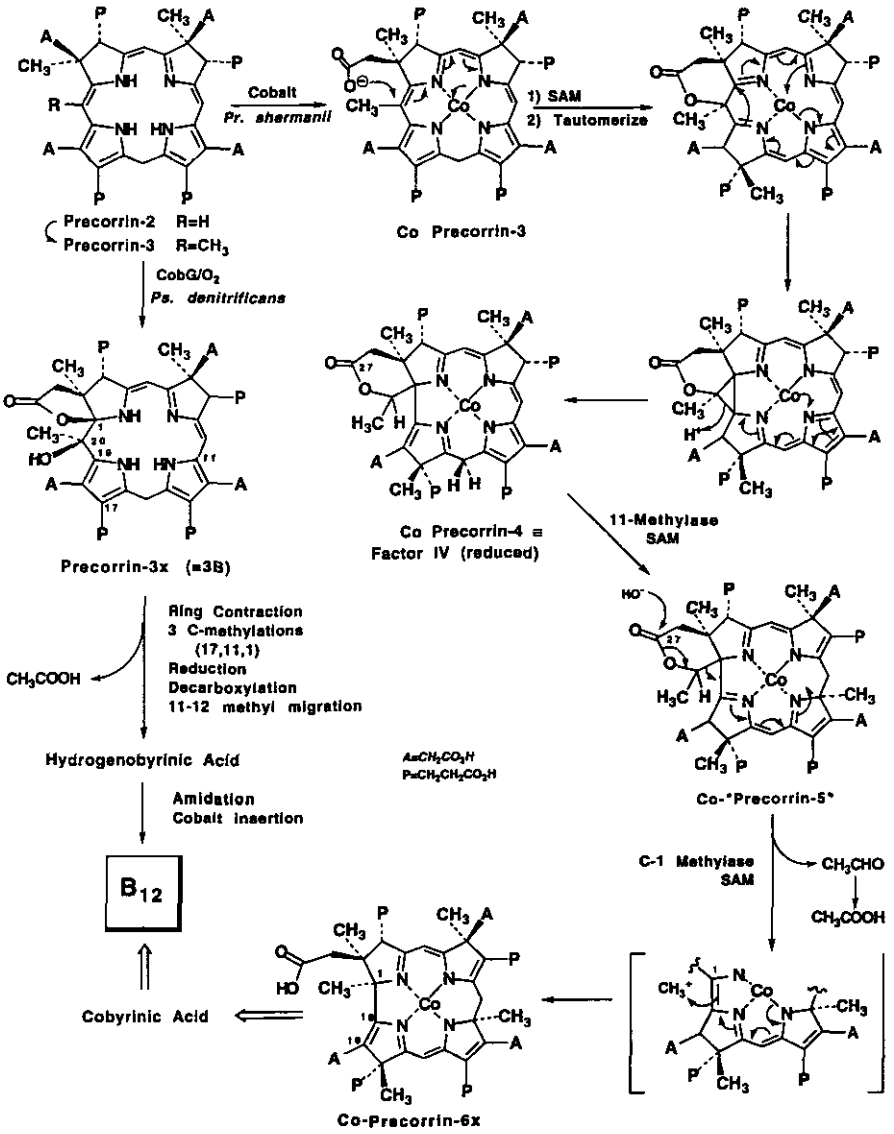
ring A-acetamido group (C-27) takes place during formation of B_{12}, some time after precorrin-3, but before cobyrinic acid (see ***a*** in vitamin B_{12}, scheme 1) which has already lost one half of its -C=^{18}O label at this ***a*** position.[37] Also to be rationalized in the transfer of ^{18}O label from this same C-27 CO_2H to the acetic acid liberated during B_{12} synthesis in *P. shermanii*.[38] Thus there is a subtle distinction between the role of the acetate side chain in the aerobic pathway compared with the parallel anaerobic pathway in *P. shermanii*, for the latter organism cannot utilize dioxygen as part of the mechanism and cobalt is inserted at an early stage, into precorrin-2. It is of considerable evolutionary interest that, while the aerobes such as *P. denitrificans* carry out the synthesis of the complete corrin, HBA and its a,c-diamide **before** metal insertion, i.e. are driven from porphyrinoids to ring contracted corrin *via* functionalization by O_2, the anaerobic pathway could employ the 2-electron valency changes of cobalt ion to mediate similar, but non-identical, steps most probably using the internal oxygen of the ring -A acetate function as the source of C-20 "hydroxylation." The pathways in fact diverge at precorrin-2 and do not intersect again until the a,c-diamide of cobyrinic acid is reached.[1,2]

4.3 Timing of cobalt insertion

Scheme 9



Scheme 10



The isolation of factors I-III from *P. shermanii* is only possible in the absence of cobalt. The circumstantial evidence was therefore strong that cobalt is inserted into the macrocycle at an early stage. In collaboration with G. Müller[24] we were able to show that cobalt factor II was efficiently converted to cobyrinic acid in *P. shermanii* and that cobalt factor III (which could be isolated from cell free extracts) also served as a precursor. With this new knowledge and independent evidence from Cambridge that cobalt is inserted early,[39] the reason for the failure of the *Salmonella* gene products to process the cobalt free substrate precorrin-3 becomes clear. Now we had to begin our "gene hunt" all over again, but this time using cobalt precorrin-2, 3 *etc.* as substrates. Logically, we therefore searched for the cobalt chelatase of the anaerobic pathway. It will be recalled that in the aerobic route, cobalt is inserted late, into hydrogenobyrinic acid a, c-diamide by a specific enzyme complex (*Cob NST*) requiring ATP, which

inserts the cobalt as Co^{II} i.e. without valency change.[27] Since no homologies for *Cob NST* could be found in the *S. typhimurium* gene bank another impasse had been reached. However, remembering the ferrochelatase activity of the versatile CysG we tested the catalytic ability of this multifunctional enzyme to insert cobalt, using factor II as substrate and found a greatly enhanced rate ($\sim 10^3$) over the control reaction.[7] Confirmation of the role of cysG as a cobalt inserting enzyme has now come from genetic experiments by Roth,[40] who showed that 66 cysG auxotrophs of *Salmonella* do not make B_{12} or siroheme.

4.4 A new intermediate from *P. shermanii* - at last!

After many years of searching in cell-free extracts of *P. shermanii*, a novel B_{12} intermediate, Factor IV, which is a cobalt tetradehydrocorrin with 4 methyl groups has just been discovered[41] and its structure (Scheme 9) is indeed a surprise, for not only has ring contraction taken place but the ring A acetate has formed a unique δ-lactone terminating at C-20 which also bears a proton. This means that, unlike C-20 in precorrins 4 and 5, which is present as a ketone (*vide supra*), the same carbon in Factor IV is at the oxidation level of an *alcohol* and must formally lose *acetaldehyde* rather than acetic acid when the 2 carbon fragment is extruded at some later stage. By trapping with dimedone in the incubation, we could show that the 2-carbon fragment is indeed excised as acetaldehyde, which is otherwise oxidized to acetic acid[43,44] in the cell free extract. This oxidation takes place with great facility and we could follow the transformation of acetaldehyde to a mixture of ethanol and acetic acid in *P. shermanii* extracts.

Since ring contraction in the anaerobic *P. shermanii* cannot use O_2, we have suggested[41,42] in Scheme 10 that the process is initiated by attack of the ring A carboxylate on C-20 with concomitant formation of a cyclopropane intermediate. The structure of Factor IV also rationalizes the unique exchange of ^{18}O label at the ring A acetate carbonyl during anaerobic biosynthesis[36,37] and the transfer of ^{18}O label to the acetic acid isolated during B_{12} biosynthesis in *P. shermanii*.[38] A plausible biosynthetic route from Factor IV to Co precorrin-6x is shown in scheme 10. It is clear that the ancient, anaerobic route to B_{12} was able to orchestrate the ring contraction using an internal rather than an external source of oxygen functionality, whose installation at C-20 is necessary in both pathways but whose origin and subsequent fate is quite distinct in aerobic *vs* anaerobic metabolism. That Nature should have discovered, and still uses, two independent routes to such a complex structure can only be regarded as fantastic.

5 Epilogue

We regard the successful conclusion of the synthesis of B_{12} not so much as the end of a chapter in biosynthetic exploration but rather as the beginning of a whole new era in genetically engineered synthesis of complex natural products, regardless of their origin, since it is now possible to prepare and express cDNA or genomic libraries from any species of plant or organism.

ACKNOWLEDGMENTS

It is a pleasure to pay tribute to the contributions of an enthusiastic team of graduate students, postdoctoral fellows and technical assistants whose names are mentioned in the references and the reviews.[1,10,12] Over the years we have enjoyed continuing and fruitful collaborations with Professors G. Müller (Stuttgart), M. Kajiwara (Tokyo), J. Roth (Utah) and P. M. Jordan (Southampton). Last, but by no means least, my profound debt to the senior members of the group, Drs. C. A. Roessner (genetics/enzymology), N. J. Stolowich (enzymology/NMR), P. J. Santander (Chemistry), and H. J.

Williams (NMR) is gratefully acknowledged. We thank the National Institute of Health, Robert A. Welch Foundation, and the Texas Advanced Research Program for generous financial support.

REFERENCES

1. Review: A. I. Scott, *Tetrahedron,* **1994**, *50*, 13313.
2. Review: F. Blanche, B. Cameron, J. Crouzet, L. Debussche, D. Thibaut, M. Vuilhorgne, F. J. Leeper, and A. R. Battersby, *Angew. Chem., Int. Ed. Engl.*, **1995**, *34*, 384.
3. M. D. Gonzalez and A. I. Scott, *Bioorg. & Med. Chem. Lett.*, **1994**, *4*, 743.
4. G. Müller in Vitamin B_{12}, B. Zagalak and W. Friedrich Eds. de Guyter, New York, **1979**, p. 279.
5. B. Cameron, K. Briggs, S. Pridmore, G. Brefort, and J. Crouzet, *J. Bacteriol.*, **1989**, *171*, 547.
6. M. J. Warren, C. A. Roessner, P. J. Santander, and A. I. Scott, *Biochem. J.*, **1990**, *265*, 725.
7. J. B. Spencer, N. J. Stolowich, C. A. Roessner, and A. I. Scott, *FEBS Lett.*, **1993**, *335*, 57.
8. R. M. Jeter and J. R. Roth, *J. Bacteriol.*, **1987**, *169*, 3187.
9. J. R. Roth, M. J. G. Lawrence, M. Rubenfield, M. Kieffer-Higgins, and G. M. Church, *J. Bacteriol.*, **1993**, *175*, 3303.
10. Review: A. I. Scott, *Angew. Chem., Int. Edn. Engl.*, **1993**, *32*, 1223.
11. M. J. Warren, C. A. Roessner, S. Ozaki, N. J. Stolowich, P. J. Santander, and A. I. Scott, *Biochemistry*, **1992**, *31*, 603.
12. A. I. Scott, *Acc. Chem. Res.*, **1990**, *23*, 308.
13. H. C. Uzar, A. R. Battersby, T. A. Carpenter., F. J. Leeper, *J. Chem. Soc., Perkin. Trans. I*, **1987**, 1689.
14. F. Blanche, D. Thibaut, D. Frechet, M. Vuilhorgne, J. Crouzet, B. Cameron, G. Müller, K. Hlineny, U. Traub-Eberhard, and M. Zboron, *Angew. Chem., Int. Ed. Engl.*, **1990**, *29*, 884.
15. C. A. Roessner, M. J. Warren, P. J. Santander, B. P. Atshaves, S-I Ozaki, N. J. Stolowich, K. Iida, and A. I. Scott, *FEBS Lett.*, **1992**, *301*, 73.
16. S-I Ozaki, C. A. Roessner, N. J. Stolowich, B. P. Atshaves, R. Hertle, G. Müller, and A. I. Scott, *J. Am. Chem. Soc.*, **1993**, *115*, 7935.
17. D. Thibaut, L. Debussche, and F. Blanche, *Proc. Natl. Acad. Sci.*, USA **1990**, *87*, 8795.
18. F. Blanche, D. Thibaut, A. Famechon, L. Debussche, B. Cameron, and J. Crouzet, *J. Bacteriol.*, **1992**, *174*, 1036.
19. F. Blanche, A. Famechon, D. Thibaut, L. Debussche, B. Cameron, and J. Crouzet, *J. Bacteriol.*, **1992**, *174*, 1050.
20. D. Thibaut, F. Blanche, L. Debussche, F. J. Leeper, and A. R. Battersby, *Proc. Natl. Acad. Sci., USA*, **1990**, *87*, 8800.
21. D. Thibaut, F. Kiuchi, L. Debussche, F. J. Leeper, F. Blanche, and A. R. Battersby, *J. Chem. Soc., Chem. Comm.*, **1992**, 139.
22. D. Thibaut, F. Kiuchi, L. Debussche, F. Blanche, M. Kodera, F. J. Leeper, and A. R. Battersby, *J. Chem. Soc., Chem. Commun.*, **1992**, 982.
23. J. Crouzet, B. Cameron, L. Cauchois, F. Blanche, S. Rigault, M-C Royez, D. Thibaut, and L. Debussche, *J. Bacteriol.*, **1990**, *172*, 5980.
24. G. Müller, F. Zipfel, K. Hlineny, E. Savvidis, R. Hertle, U. Traub-Eberhard, A. I. Scott, H. J. Williams, N. J. Stolowich, P. J. Santander, M. J. Warren, F. Blanche, and D. Thibaut, *J. Am. Chem. Soc.*, **1991**, *113*, 9893.
25. D. Thibaut, L. Debussche, D. Fréchet, F. Herman, M. Vuilhorgne, and F. Blanche, *J. Chem. Soc., Chem. Comm.*, **1993**, 513.

26. A. I. Scott, C. A. Roessner, N. J. Stolowich, J. B. Spencer, C. Min, and S-I Ozaki, *FEBS Lett.*, **1993**, *331*, 105.
27. L. Debussche, D. Thibaut, M. Danzer, F. Debu, D. Fréchet, F. Herman, F. Blanche, and M. Vuilhorgne, *J. Chem. Soc., Chem. Comm.*, **1993**, 1100.
28. C. Min. B. P. Atshaves, C. A. Roessner, N. J. Stolowich, J. B. Spencer, and A. I. Scott, *J. Am. Chem. Soc.*, **1993**, *115*, 10380.
29. C. A. Roessner, J. B. Spencer, S-I Ozaki, C. Min, B. P. Atshaves, P. Nayar, N. Anousis, N. J. Stolowich, M. T. Holderman, and A. I. Scott, *Protein Expression and Purification*, **1995**, *6*, 155.
30. C. A. Roessner, J. B. Spencer, N. J. Stolowich, J. Wang, G. P. Nayar, P. J. Santander, C. Pichon, C. Min, M. Holderman, and A. I. Scott, *Chemistry and Biology*, **1994**, *1*, 119.
31. L. Debussche, D. Thibaut, B. Cameron, J. Crouzet, F. Blanche, *J. Bacteriol.*, **1993**, *175*, 7430.
32. T. Podschun and G. Müller, *Angew. Chem., Int. Ed. Engl.*, **1985**, *24*, 46.
33. A. Eschenmoser, *Angew. Chem., Int. Ed. Engl.*, **1988**, *27*, 5.
34. N. J. Stolowich, J. Wang, J. B. Spencer, P. J. Santander, C. A. Roessner, and A. I. Scott, *J. Am. Chem. Soc.*, **1996**, *118*, 1657.
35. Y. Li, A. I. D. Alanine, R. A. Vishwakarma, S. Balachandran, F. J. Leeper, and A. R. Battersby, *J. Chem. Soc., Chem. Comm.*, **1994**, 2507.
36. A. I. Scott, N. J. Stolowich, B. P. Atshaves, P. Karuso, M. J. Warren, H. J. Williams, M. Kajiwara, K. Kurumaya, and T. Okazaki, *J. Am. Chem. Soc.*, **1991**, *113*, 9893.
37. R. A. Vishwakarma, S. Balachandran, A. I. D. Alanine, N. P. J. Stanford, F. Kiuchi, F. J. Leeper, and A. R. Battersby, *J. Chem. Soc., Perkin Trans. I*, **1993**, 2893.
38. D. Arigoni, CIBA Foundation Symposium, 180. The Biosynthesis of the Tetrapyrrole Pigments. Wiley New York, **1994**, pp. 280 - 281.
39. S. Balachandran, R. A. Vishwakarma, S. M. Monaghan, A. Prelle, N. P. J. Stanford, F. J. Leeper, and A. R. Battersby, *J. Chem. Soc., Perkin Trans. I*, **1994**, 487.
40. B. S. Goldman and J. R. Roth, *J. Bacteriol.*, **1993**, *175*, 1457.
41. A. I. Scott, N. J. Stolowich, J. Wang, O. Gawatz, E. Fridrich, and G. Müller, *Proc. Natl. Acad. Sci.*, **1996**, *93*, 14316.
42. J. Wang, N. J. Stolowich, P. J. Santander, J. H. Park, and A. I. Scott, *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 14320.
43. L. Mombelli, C. Nussbaumer, H. Weber, G. Müller, and D. Arigoni, *Proc. Nat. Acad. Sci., (USA)* **1981**, *78*, 11.
44. A. R. Battersby, M. J. Bushell, C. Jones, N. G. Lewis, and A. Pfenninger, *Proc. Nat. Acad. Sci., (USA)* **1981**, *78*, 13.

Received, 6th May, 1997