Recent discoveries in the pathways to cobalamin (coenzyme B$_{12}$) achieved through chemistry and biology[*],[**]

A. Ian Scott and Charles A. Roessner[‡]

Center for Biological NMR, Department of Chemistry, Texas A&M University, College Station, TX 77843-3255, USA

Abstract: The genetic engineering of *Escherichia coli* for the over-expression of enzymes of the aerobic and anaerobic pathways to cobalamin has resulted in the in vivo and in vitro biosynthesis of new intermediates and other products that were isolated and characterized using a combination of bioorganic chemistry and high-resolution NMR. Analyses of these products were used to deduct the functions of the enzymes that catalyze their synthesis. CobZ, another enzyme for the synthesis of precorrin-3B of the aerobic pathway, has recently been described, as has been BluB, the enzyme responsible for the oxygen-dependent biosynthesis of dimethylbenzimidazole. In the anaerobic pathway, functions have recently been experimentally confirmed for or assigned to the CbiMNOQ cobalt transport complex, CbiA (*a,c* side chain amidation), CbiD (C-1 methylation), CbiF (C-11 methylation), CbiG (lactone opening, deacylation), CbiP (*b,d,e,g* side chain amidation), and CbiT (C-15 methylation, C-12 side chain decarboxylation). The dephosphorylation of adenosylcobalamin-phosphate, catalyzed by CobC, has been proposed as the final step in the biosynthesis of adenosylcobalamin.

Keywords: cobalamin; vitamin B$_{12}$; CobZ; BluB; Cbi proteins; CobC.

INTRODUCTION

The solutions to some of the most complex natural product biosynthetic pathways have been attained only through combining chemistry with biology. One of the more successful applications of this approach has been the delineation of the aerobic and anaerobic pathways to the enzyme cofactor, adenosylcobalamin, and, hence, its cyano derivative, vitamin B$_{12}$ (Fig. 1). (For the purpose of this review, only selected portions of the anaerobic pathway are shown. The more complete aerobic and anaerobic pathways from uroporphyrinogen III to adenosylcobalamin may be seen and downloaded at <http://people.tamu.edu/~c-roessner>.) Some of the early (and current) phases of this work were performed by bacterial geneticists who isolated and sequenced clusters of genes, necessary for cobalamin biosynthesis, from the bacteria *Pseudomonas denitrificans* [1] and *Salmonella enterica* serovar typhimurium (*S. typhimurium*) [2]. It was found that cobalamin is manufactured (by bacteria and archaea) via two alternate routes that differ primarily in the timing of cobalt insertion and the early stages that lead to the contraction of the macrocycle and excision the extruded carbon molecule and its attached methyl group. The pathway in *P. denitrificans*, an aerobe, was found to incorporate molecular oxygen into the precorrin-3 macrocycle as a prerequisite to ring contraction and was consequently termed the aerobic

pathway. The pathway in *S. typhimurium*, which synthesizes cobalamin only in the absence of oxygen and, thus, is called the anaerobic pathway, was found to take advantage of a cobalt ion chelated into pre-corrin-2 (or Factor II) to set the stage for ring contraction.
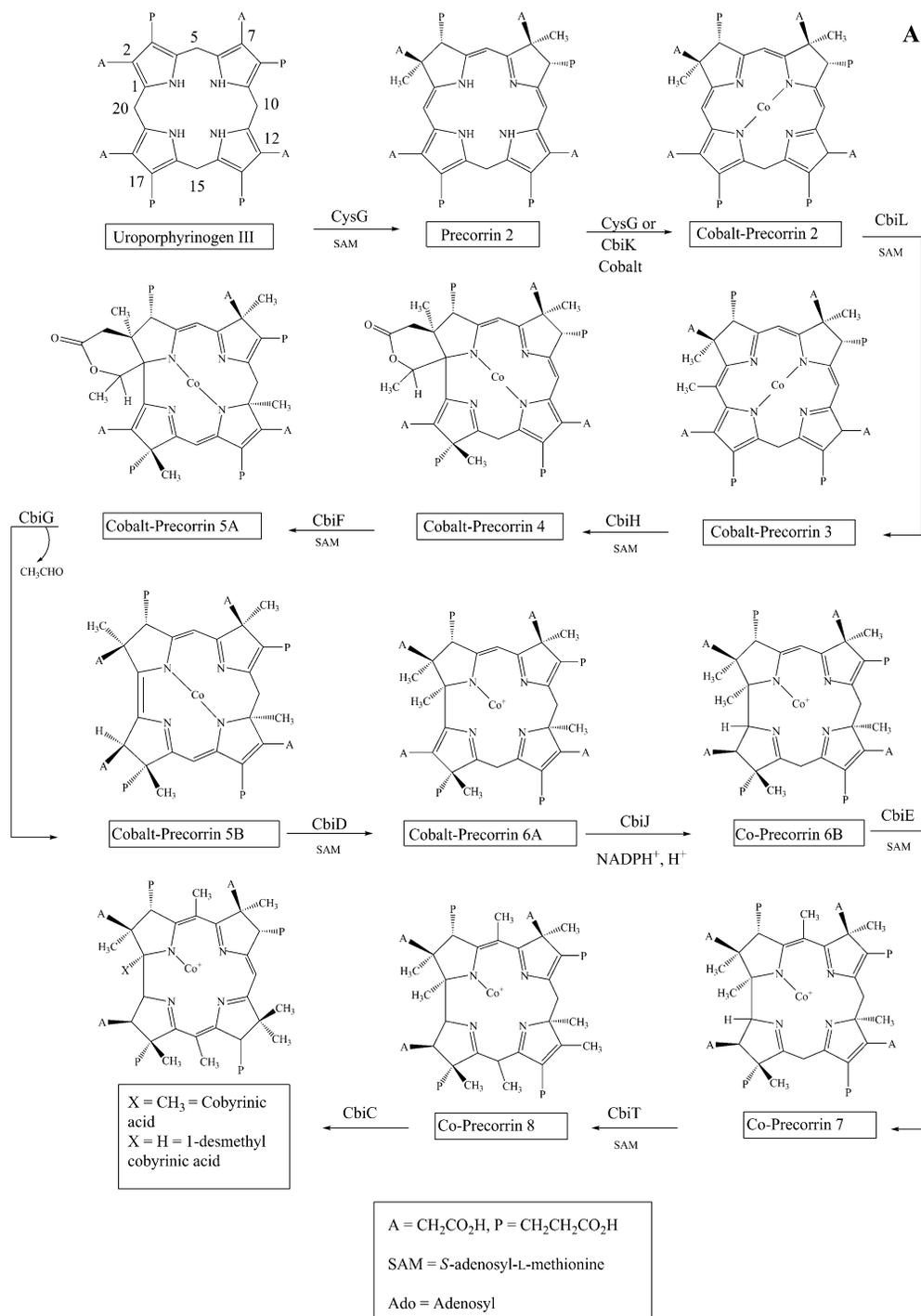

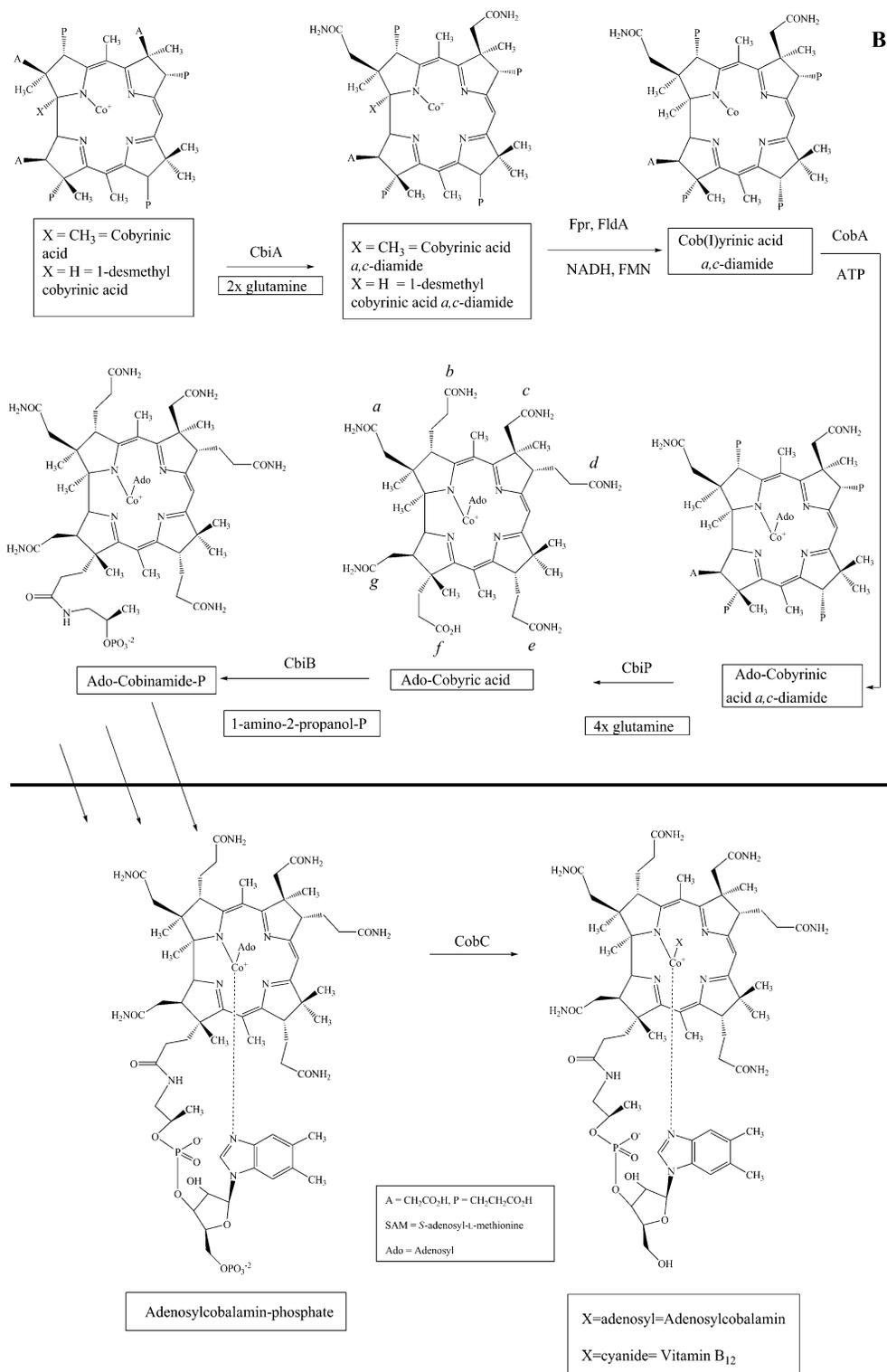**Fig. 1A** The anaerobic pathway from uroporphyrinogen III to cobyrinic acid.

Fig. 1B Top: The anaerobic pathway from cobyrinic acid to adenosylcobinamide. Bottom: The proposed final step of cobalamin biosynthesis.

Over-expression of the cobalamin (*cob*) biosynthetic genes from *P. denitrificans* allowed the determination of the intermediates of the aerobic biosynthetic route from uroporphyrinogen III (urogen III) to adenosylcobalamin (for the most recent reviews, see [3,4]) and the concommitant assignment of the functions of the encoded enzymes based on their catalytic activity. *Cbi* genes for the biosynthesis of cobalamin by the anaerobic pathway have been over-expressed from not only *S. typhimurium* but also from *Bacillus megaterium* [5] and *Propionibacterium freudenreichii* (*shermanii*) [6] and, while the catalytic activity had been determined experimentally for a few of the gene products, until recently most of their functions had been derived on the basis of sequence similarities to enzymes of the aerobic pathway. In this paper, we review some of the recent discoveries resulting from the combination of chemistry with biology that have helped provide answers to some of the questions that remained in both pathways to adenosylcobalamin.

THE DISCOVERY OF TWO NEW FLAVIN-MONOOXYGENASES FOR THE BIOSYNTHESIS OF PRECORRIN-3B AND DIMETHYLBENZIMIDAZOLE

Molecular oxygen has been shown to play an important role in two separate portions of the aerobic pathway to cobalamin, i.e., in the biosynthesis of precorrin-3B and in the biosynthesis of 5,6-dimethylbenzimidazole (DBMI).

CobZ: An FAD-bearing enzyme for the biosynthesis of precorrin-3B in *Rhodobacter*

Precorrin-3B is derived from precorrin-3A via the introduction of oxygen into the macrocycle (Fig. 2). Most organisms with the aerobic pathway use the iron-sulfur protein, CobG, for the synthesis of precorrin-3B [7]. However, some species, e.g., *Rhodobacter* sp., that apparently have the aerobic pathway (as demonstrated by the accumulation of metal-free corrinoids in their cytoplasm and by having the cobaltochelatase complex, CobNST, as does *P. denitrificans*), do not have a gene for CobG among their cobalamin biosynthetic genes. This conundrum was solved when it was found that a flavoprotein, called CobZ, could serve as an isofunctional substitute for CobG in a strain of *E. coli* that had been genetically engineered to synthesize the cobalamin precursor, hydrogenobyrinic acid (HBA) [8]. While CobG has a single Fe–S center, CobZ was shown to have not only two Fe–S centers but also a heme group and an FAD cofactor. Even though the mechanism of action for CobZ (Fig. 2) is postulated to differ from that of CobG, the product of the two enzymes is most likely the same (precorrin-3B) since one enzyme may substitute for the other in the engineered biosynthesis of HBA.
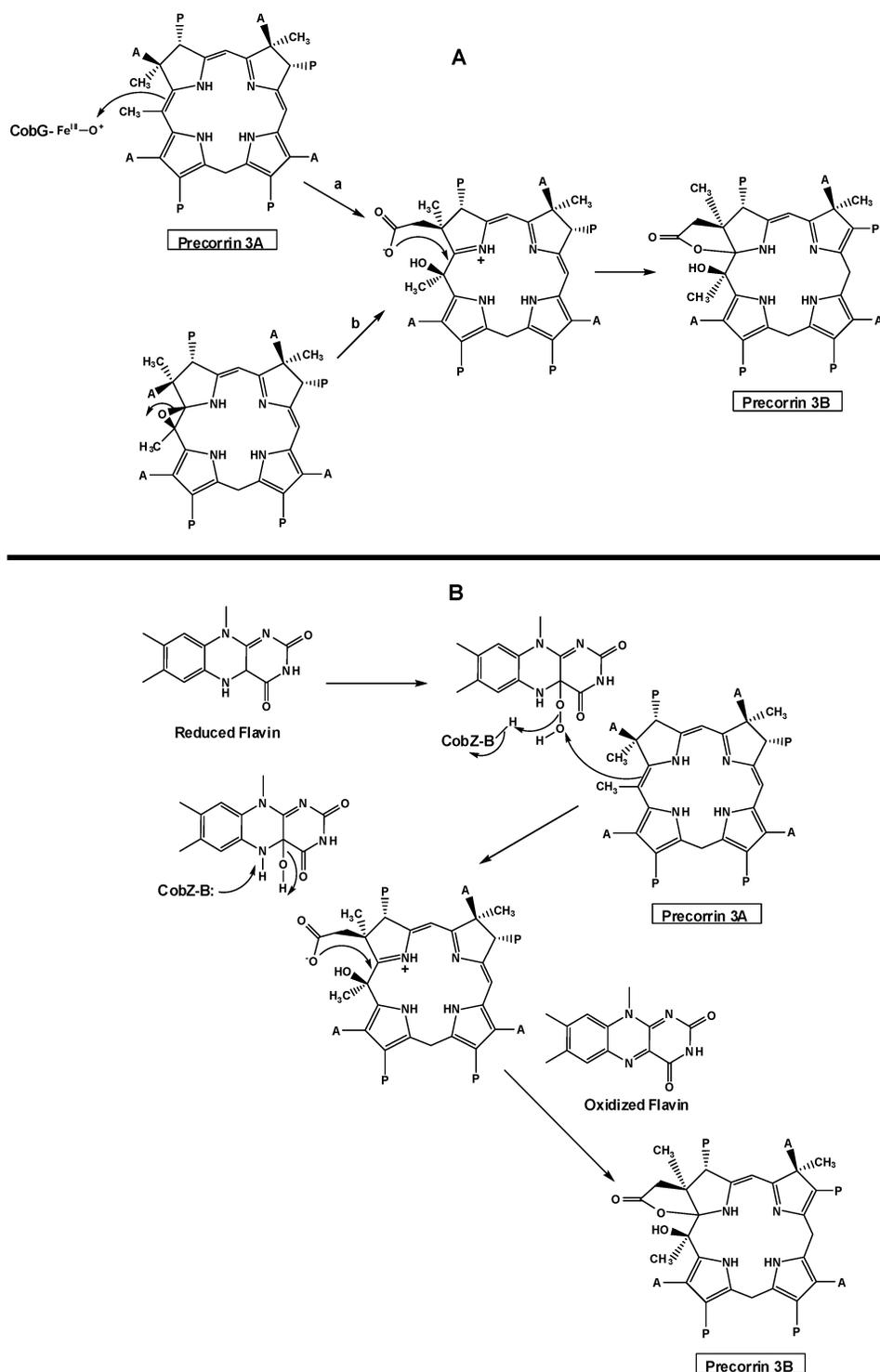
Fig. 2 Proposed mechanisms of action for CobG and CobZ. (A) Proposed mechanisms for CobG which involve either the direct insertion of oxygen from an Fe^{III}–O$^+$ species ($\equiv$OH$^+$) or, alternatively, formation of a 1,20-epoxide. (B) Proposed mechanism for CobZ involving a hydroperoxide intermediate.

5,6-Dimethylbenzimidazole derived from FMNH$_2$ by the action of BluB

The discovery that BluB converts FMNH$_2$ to DMBI (Fig. 3) has been heralded as the solution to the last unknown step in adenosylcobalamin biosynthesis. The *bluB* gene was first found in a cluster of *Rhodobacter capsulatus* genes encoding enzymes used in the latter stages of cobalamin biosynthesis and was shown to be required for the conversion of cobinamide to cobalamin [9]. Genetic evidence that BluB was involved in DMBI synthesis first appeared in a patent [10] using a mutant of *R. capsulatus* and was later confirmed with mutants of *Sinorhizobium meliloti* [11]. Two groups have now provided biochemical evidence that BluB is sufficient for the conversion of FMNH$_2$ to DMBI [12,13]. Production of DMBI is facilitated by FMN reductase and NAD(P)H to supply FMNH$_2$ for BluB, and the reaction requires molecular oxygen. At least three mechanisms for BluB, all involving a hydroperoxide intermediate (Fig. 3), have been proposed [13,14].
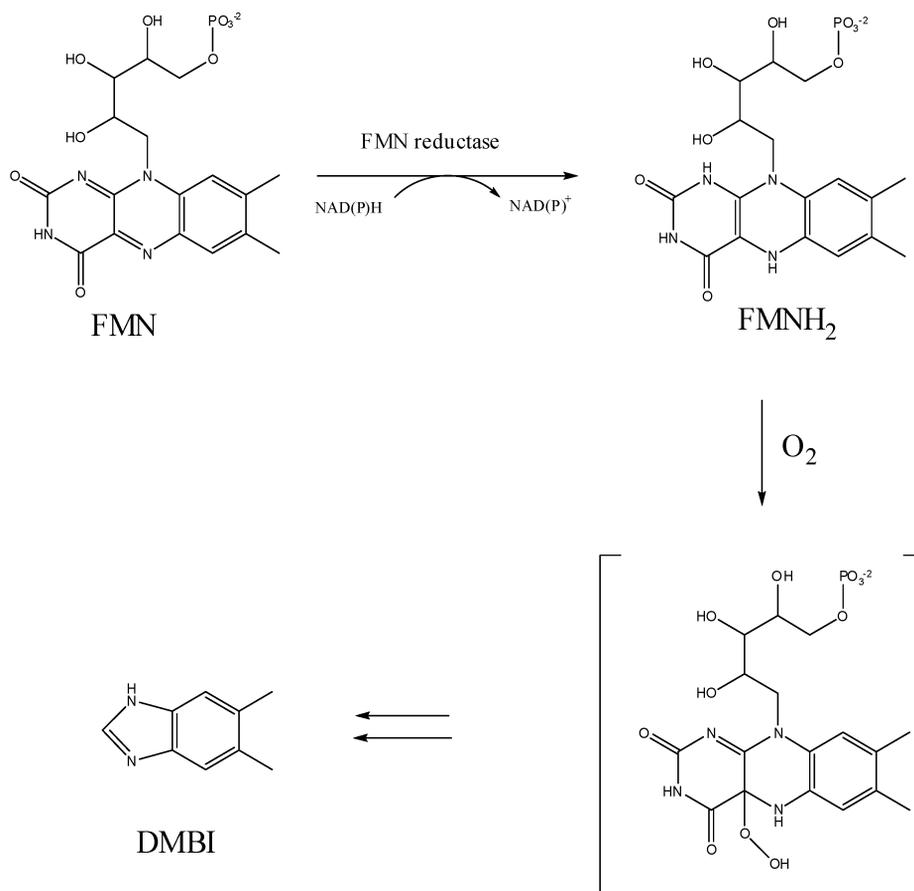

Fig. 3 The biosynthesis of DMBI from FMNH$_2$.

RECENT EXPERIMENTAL EVIDENCE FOR THE ACTIVITIES OF THE *S. TYPHIMURIUM* Cbi PROTEINS OF THE ANAEROBIC PATHWAY TO COBALAMIN

A single large *cob* operon containing most of the genes for the anaerobic biosynthesis of cobalamin in *S. typhimurium* was discovered in 1993 [2]. In this operon were 4 genes (*cbiM,N,O,Q*) that were proposed to encode a cobalt transport system and 13 other *cbi* genes that encode enzymes necessary for

the conversion of precorrin-2 to adenosylcobinamide-phosphate. The functions of CbiA,B,C,D,E,T,F,G,H,J,K,L,P for this part of the anaerobic pathway (Fig. 1) were assigned in large part based on similarity to counterparts in *P. denitrificans* which had been more extensively characterized [15–20]. Until recently, however, experimental evidence existed for the functions of only 4 (CbiB,H,K,L) of the Cbi proteins. Biochemical evidence had revealed that CbiK is a chelatase that inserts cobalt into precorrin-2 [21]. In some organisms, CbiK is replaced by either CysG or CbiX as the cobaltochelatase. In addition, there may be some variation in the oxidation state of the di- and trimethylated intermediates shown in Fig. 1A since Factor II and Factor III, the oxidized forms of precorrin-2 and precorrin-3, can also be used as substrates. CbiL has been shown to be the C-20 methyltransferase for the conversion of cobalt-precorrin-2 to cobalt-precorrin-3 [22], and CbiH catalyzes methylation at C-17 of cobalt-precorrin-3, resulting in formation of the ring-contracted, lactonized cobalt-precorrin-4 [23]. In addition, genetic evidence [24] had suggested that CbiB is involved in coupling of the aminopropanol-phosphate group to adenosyl-cobyric acid to form adenosylcobinamide-phosphate, but functions for the remaining 9 Cbi enzymes between cobalt-precorrin-4 and adenosyl-cobinamide-phosphate have proved elusive. Especially intriguing are two enzymes, CbiD and CbiG, whose activities could not be predicted because of lack of similarity to any other proteins of known function. Recent results now provide experimental evidence for the functions of CbiA, CbiD, CbiF, CbiG, CbiP, CbiT, and CbiMNOQ and reveal the structures of two new intermediates in the pathway, cobalt-precorrins-5A and -5B.

New evidence that CbiMNOQ is a cobalt transport system

Recent experimental evidence has confirmed that the CbiMNOQ proteins compose a metal transport system that prefers cobalt over nickel [25]. CbiMNO or CbiMN alone provided transport at about 25 % efficiency of the full complement of CbiMNOQ, suggesting that components of endogenous ABC-type transporters may substitute for CbiO and CbiQ.

CbiF methylates cobalt-precorrin-4 to form cobalt-precorrin-5A

As early as 1992 [22], CbiF (and thus CobM of the aerobic pathway) was predicted to be the C-11 methyltransferase necessary for the transformation of cobalt-precorrin-4 to cobalt-precorrin-5 because of its ability to mismethylate precorrin-3 at C-11. However, this mismethylation activity of CbiF and the extraordinary sensitivity of the CbiF methylation products to oxidation [26] greatly complicated the isolation of intermediates subsequent to cobalt-precorrin-4. Cobalt-precorrin 5A, the product resulting from the CbiF-catalyzed methylation of cobalt-precorrin-4 at C-11 (Fig. 1A), was recently synthesized, purified, and characterized only when oxygen was carefully eliminated from in vitro incubations containing cobalt, precorrin-3, CbiH, and CbiF, and during the isolation procedures [27].

CbiG opens the lactone ring and deacylates cobalt-precorrin-5A to afford cobalt-precorrin-5B

Unlike the equivalent precorrin-5 of the aerobic pathway, cobalt-precorrin-5A still carries the δ-lactone that formed as a consequence of C-17 methylation and ring contraction. The deacylated product, cobalt-precorrin-5B (Fig. 1A), was observed [27] only when CbiG was included in the incubation mixture described above for the biosynthesis of cobalt-precorrin-5A. Thus, the function of CbiG can now be assigned as catalyzing both the opening of the lactone ring and extrusion of the two-carbon fragment (deacylation), derived from C-20 and its associated methyl group. The acyl group has been shown to be eliminated as acetaldehyde [28].

     The observation that CobE of the aerobic pathway (whose function is unknown) shows some similarity to the carboxyl terminal of CbiG suggests that it, too, may be involved in opening of the lactone

ring or deacylation. CobE was not required for the in vitro biosynthesis of precorrin-6A [19], but the deacylation of precorrin-5 was the least efficient step in the multienzyme synthesis of precorrin-6A. It may be that CobE facilitates this process.

CbiD is necessary for C-1 methylation

In the aerobic pathway, deacylation is concomitant with C-1 methylation, catalyzed by CobF. The anaerobic pathway has no methyltransferase similar to CobF, but evidence has now been provided [29] that CbiD is required for C-1 methylation and, thus, may be a non-orthologous methyltransferase that takes the place of CobF in the anaerobic pathway. In this work, a strain of *E. coli* was genetically engineered to contain the 12 *S. typhimurium* genes believed to be required for the biosynthesis of cobyric acid (all of the enzyme-encoding *cbi* genes except *cbiB*). This strain accumulated cobyrinic acid *a,c*-diamide (Fig. 1B). However, a mutant of the strain constructed to lack only the *cbiD* gene accumulated a similar product that was still protonated, rather than methylated, at the C-1 position (1-desmethyl-cobyrinic acid *a,c*-diamide, Fig. 1B). Even though the cobalt-precorrin-6A intermediate depicted in Fig. 1A has yet to be isolated, and there is some mystery as to why the presence of the two amidating enzymes, CbiA and CbiP, was necessary for C-1 methylation, the engineered system has provided the first solid evidence that CbiD is required. CbiD has a potential SAM binding site so is probably the actual methyltransferase, but it could also work in conjunction with one of the other methyltransferases, e.g., CbiF. The structure of CbiD from *Archaeoglobus fulgidus* has been determined, but it provided no clues to its activity (unpublished information available from the Protein Data Bank Web site: <http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=1SR8>).

CbiT is both a C-15 methyltransferase and a decarboxylase

In the aerobic pathway, methylation at C-5 and C-15 and decarboxylation of the C-12 acetate side chain of precorrin-6B are catalyzed by a single enzyme, CobL. In *S. typhimurium*, however, CobL is split into two separate enzymes, CbiE and CbiT. Because of its similarity to other methyltransferases of the B_{12} pathway, it has long been assumed that CbiE is the methyltransferase that catalyzes the addition of the two methyl groups to cobalt-precorrin-6B, and that CbiT then decarboxylates the C-12 acetate side chain to afford cobalt-precorrin 8. However, the structure of CbiT [30] and its similarity to methyltransferases not of the B_{12} pathway have led to the suggestion that it, too, might be a methyltransferase. The first experimental evidence that CbiT alone can catalyze both methylation at C-15 and decarboxylation of the C-12 acetate side chain, even in the absence of C-1 methylation, has now been provided [26]. The products synthesized from cobalt-precorrin-3 in the presence of CbiF, CbiG, CbiH, and CbiT were either methylated at C-15 or methylated at C-15 *and* decarboxylated, but never decarboxylated without C-15 methylation, suggesting that CbiT first catalyzes C-15 methylation followed by decarboxylation. In the absence of CbiT, neither methylation at C-15 nor decarboxylation were observed. If CbiT is the C-15 methyltransferase, then CbiE remains as the C-5 methyltransferase. Fig. 1A depicts the CbiE-catalyzed methylation at C-5 as occuring before the action of CbiT, but this order has yet to be confirmed.

CbiA is the *a* and *c* side chain amidase

CbiA from *S. typhimurium* has been over-expressed and its mechanism of action studied in detail [31] providing conclusive evidence that it is the *a,c*-amidase. In addition, it has been reported [29] that a genetically engineered strain of *E. coli* containing 10 *cbi* genes (all *cbi* genes except *cbiA*, *cbiB*, and *cbiP*) accumulated 1-desmethyl-cobyrinic acid. Addition of the *cbiA* gene to this strain resulted in the accumulation of a bisamidated product (1-desmethyl-cobyrinic acid *a,c*-diamide, Fig. 1B) providing further proof that CbiA is responsible for amidation of the two side chains.

CbiP amidates the *e*, *d*, *b*, and *g* side chains to afford adenosylcobyric acid

As predicted from its similarity to CobQ of *P. denitrificans*, CbiP from *S. typhimurium* has now been proven to be adenosylcobyric acid synthase [32]. The addition of the four amide groups to adenosyl-cobyrinic acid *a,c*-diamide is not random but rather has been shown to proceed counterclockwise beginning with the *e*-side chain (*e*, *d*, *b*, *g*, Fig. 1B).

What is the final step of cobalamin biosynthesis?

It has long been believed that the final step of adenosylcobalamin biosynthesis is the replacement of the GDP group of adenosyl-GDP-cobinamide with α-ribazole, catalyzed by CobS (anaerobic pathway) or CobV (aerobic pathway). Recent evidence provided by mutants of *S. typhimurium*, however, suggests that the α-ribazole moiety is added as α-ribazole-phosphate and that the final step of cobalamin biosynthesis is removal of the phosphate group from adenosylcobalamine-phosphate by CobC [33] (Fig. 1B).

CONCLUDING REMARKS

Experimental evidence is now in hand for the functions of CobZ, BluB, CobC, and almost all of the 17 Cbi proteins. Remaining challenges to the solution of the anaerobic pathway to cobalamin include solving the oxygen-independent synthesis of dimethylbenzimidazole, confirmation of the predicted activities of CbiC, CbiE, and CbiJ (Fig. 1A) and the biosynthesis and confirmation of the stuctures of cobalt-precorrins-6A, -6B, -7, and -8. Components necessary for the reduction of Co(II) to Co(I) and the attachment of the adenosyl group (Fig. 1B) have recently been described [34,35], but the mechanism for the attachment of the lower ligand could bear further scrutiny.

Perhaps the theme of this special issue of *Pure and Applied Chemistry* was aptly described by Prof. Ian Scott when he said [36], "We are living in exciting times, since the barrier between chemistry and biology has by now almost disappeared, and from my perspective as a bioorganic chemist, the possibilities presented by combining genetic engineering with organic chemistry are almost limitless." Ian's highly successful odyssey of almost 40 years to solve the mystery of the biosynthesis of the B₁₂ cofactor and other natural products is now over but his legacy will continue in those whom he guided and who follow in his footsteps.

ACKNOWLEDGMENTS

REFERENCES

1. B. Cameron, K. Briggs, S. Pridmore, G. Brefort, J. Crouzet. *J. Bacteriol.* **171**, 547 (1989).
2. J. R. Roth, G. J. Lawrence, M. Rubenfield, S. Kieffer-Higgens, G. M. Church. *J. Bacteriol.* **175**, 3303 (1993).
3. M. J. Warren, E. Raux, H. L. Schubert, J. C. Escalante-Semerena. *Nat. Prod. Rep.* **19**, 390 (2002).
4. A. I. Scott, C. A. Roessner, P. J. Santander. In *The Porphyrin Handbook*, Vol. 12, K. M. Kadish, K. M. Smith, R. Guilard (Eds.), pp. 211–229, Academic Press, San Diego (2003).
5. E. Raux, A. Lanois, M. J. Warren, A. Rambach, C. Thermes. *Biochem. J.* **335**, 159 (1998).
6. C. A. Roessner, K. Huang, M. J. Warren, E. Raux, A. I. Scott. *Microbiology* **148**, 1845 (2002).

7. A. I. Scott, C. A. Roessner, N. J. Stolowich, J. B. Spencer, C. Min, S. I. Ozaki. *FEBS Lett.* **331**, 105 (1993).
8. H. M. McGoldrick, C. A. Roessner, E. Raux, A. D. Lawrence, K. J. McLean, A. W. Munro, S. Santabarbara, S. E. J. Rigby, P. Heathcote, A. I. Scott, M. J. Warren. *J. Biol. Chem.* **280**, 1086 (2005).
9. M. Pollich, G. Klug. *J. Bacteriol.* **177**, 4481 (1995).
10. R. Blanche, B. Cameron, J. Crouzet, L. Debussche, D. Thibaut, E. Remy. Rhone-Poulenc Rorer, S. A. U.S. Patent 6,156,545, filed 2 Mar. 1999, issued 5 Dec. 2000.
11. G. R. O. Campbell, M. E. Taga, K. Mistry, J. Lloret, P. J. Anderson, J. R. Roth, G. C. Walker. *Proc. Natl. Acad. Sci. USA* **103**, 4634 (2006).
12. M. J. Gray, J. C. Escalante-Semerena. *Proc. Natl. Acad. Sci. USA* **104**, 2921 (2007).
13. M. E. Taga, N. A. Larsen, A. R. Howard-Jones, C. T. Walsh, G. C. Walker. *Nature* **446**, 449 (2007).
14. S. E. Ealick, T. P. Begley. *Nature* **446**, 387 (2007).
15. F. Blanche, M. Couder, L. Debussche, D. Thibaut, B. Cameron, J. Crouzet. *J. Bacteriol.* **173**, 6046 (1991).
16. L. Debussche, D. Thibaut, B. Cameron, J. Crouzet, F. Blanche. *J. Bacteriol.* **172**, 6239 (1990).
17. L. Debussche, D. Thibaut, B. Cameron, J. Crouzet, F. Blanche. *J. Bacteriol.* **175**, 7430 (1993).
18. C. Min, B. A. Atshaves, C. A. Roessner, N. J. Stolowich, J. B. Spencer, A. I. Scott. *J. Am. Chem. Soc.* **115**, 10380 (1993).
19. C. A. Roessner, J. B. Spencer, N. J. Stolowich, P. J. Santander, C. Pichon, C. Min, P. Nayar, S. Ozaki, B. A. Atshaves, N. Anousis, M. T. Holderman, A. I. Scott. *Chem. Biol.* **1**, 119 (1994).
20. A. I. Scott, C. A. Roessner, N. J. Stolowich, J. B. Spencer, C. Min, S.-I. Ozaki. *FEBS Lett.* **331**, 105 (1993).
21. E. Raux, C. Thermes, P. Heathcote, A. Rambach, M. J. Warren. *J. Bacteriol.* **179**, 3202 (1997).
22. C. A. Roessner, M. J. Warren, P. J. Santander, B. P. Atshaves, S.-I. Ozaki, N. J. Stolowich, K. Iida, A. I. Scott. *FEBS Lett.* **301**, 78 (1992).
23. P. J. Santander, C. A. Roessner, N. J. Stolowich, M. T. Holderman, A. I. Scott. *Chem. Biol.* **4**, 659 (1997).
24. E. Raux, A. Lanois, M. Levillayer, M. J. Warren, E. Brody, A. Rambach, C. Thermes. *J. Bacteriol.* **178**, 753 (1996).
25. D. A. Rodionov, P. Hebbeln, M. S. Gelfand, T. Eitenger. *J. Bacteriol.* **188**, 317 (2006).
26. P. J. Santander, Y. Kajiwara, H. J. Williams, A. I. Scott. *Bioorg. Med. Chem.* **14**, 724 (2005).
27. Y. Kajiwara, P. J. Santander, C. A. Roessner, L. M. Pérez, A. I. Scott. *J. Am. Chem. Soc.* **128**, 9971 (2006).
28. J. Wang, N. J. Stolowich, P. J. Santander, H. J. Park, A. I. Scott. *Proc. Natl. Acad. Sci. USA* **93**, 14320 (1996).
29. C. A. Roessner, H. J. Williams, A. I. Scott. *J. Biol. Chem.* **280**, 16748 (2005).
30. J. P. Keller, P. M. Smith, J. Benach, D. Christendat, G. T. de Titta, J. F. Hunt. *Structure* **10**, 1475 (2002).
31. V. Fresquet, L. Williams, F. M. Raushel. *Biochemistry* **43**, 10619 (2004).
32. L. Williams, V. Fresquet, P. J. Santander, F. M. Raushel. *J. Am. Chem. Soc. (Communication)* **129**, 294 (2007).
33. C. L. Zayas, J. C. Escalante-Semerena. *J. Bacteriol.* **189**, 2210 (2007).
34. M. V. Fonseca, J. C. Escalante-Semerena. *J. Biol. Chem.* **276**, 32101 (2001).
35. N. R. Buan, J. C. Escalante-Semerena. *J. Biol. Chem.* **280**, 40948 (2005).
36. A. I. Scott. *J. Org. Chem.* **68**, 2529 (2003).