way lies in its similarity to the allylic oxidation of olefins by Hg(II)[81] and to the mercuration of $FADH^-$-related heterocyclic ring systems.[82] The second mechanism invokes nucleophilic attack of $FADH^-$ on the Cys_{140} thiolate ligand attached to Hg(II), resulting in an outer-sphere reduction of the metal to Hg(0) with concomitant formation of an intermediate flavin C-4a thiol adduct. While no direct chemical precedent exists for such an outer-sphere reduction of Hg(II), mechanisms utilizing a "bridging" sulfur ligand (analogous to the Cys_{140} thiolate) are well-known in the redox reactions between Co(III) and Cr(II).[83] Also, strong evidence has recently been obtained that a C-4a Cys_{140} adduct can readily form in MerA (Figure 4, III).[84] The third possible enzyme mechanism is that of two successive single-electron transfers from dihydroflavin to Hg(II), with Hg(I) as an intermediate, similar to that previously implicated in the reduction of Hg(II) by flavin semiquinone.[85] A fourth possibility not shown is that of a formal two-electron "hydride" transfer from $FADH^-$ to Hg(II), similar to the reduction of Hg(II) by metal hydride reagents.[33] However, this mechanism seems highly unlikely in light of our findings that 5-deaza-$FADH_2$, either free in solution[58,60] or bound within the enzyme active site,[86] cannot reduce Hg(II).

Conclusions. In this Account we have described two unique enzymes, organomercurial lyase (MerB) and mercuric ion reductase (MerA), that have evolved efficient strategies to carry out organometallic and bioinorganic chemistry on mercury species. Study of the structure and function of these enzymes has revealed the importance of key cysteinyl thiols for the ligation of RHgX and Hg(II) in such a way as to selectively lower the energy barriers for protonolytic fragmentation of C–Hg bonds (MerB) and electron transfer to thiol-coordinated Hg(II) species (MerA). Understanding of the catalytic mechanisms of these enzymes may ultimately allow purposeful redesign of their genes and encoded proteins to permit enzymatic processing of other organometallic and inorganic compounds in the environment.

We acknowledge the significant contributions to the MerA and MerB projects made in this laboratory by Drs. Barbara Fox, Karin Au, Peter Schultz, Tadhg Begley, and Alan Waltz. Additionally, many of the experiments described in the MerA section were performed in collaboration with Drs. Susan Miller, Vincent Massey, David Ballou, and Charles Williams, Jr., at the University of Michigan; current understanding of the MerA mechanism is due in large part their input and expertise. Nikolaus Schiering and Dr. Emil Pai at the Max Planck Institut für Medizinische Forschung, Heidelberg, FRG, continue efforts to solve the MerA crystal structure.

Registry No. Hg, 7439-97-6; mercuric ion reductase, 67880-93-7; organomercurial lyase, 72560-99-7.

(81) Rappoport, Z.; Sleezer, P. D.; Winstein, S.; Young, W. G. *Tetrahedron Lett.* **1965**, 3719–3728. Rappoport, Z.; Dyall, L. K.; Winstein, S.; Young, W. G. *Tetrahedron Lett.* **1970**, 3483–3488. Fieser, L. F.; Fieser, M. *Reagents for Organic Synthesis*; Wiley: New York, 1967; Vol. 1, pp 644–646.

(82) Dale, R. M. K.; Martin, E.; Livingston, D. C.; Ward, D. C. *Biochemistry* **1975**, *14*, 2447–2457. Marshall, J. L.; Booth, J. E.; Williams, J. W. *J. Biol. Chem.* **1984**, *259*, 3033–3036.

(83) Deutsch, E.; Root, M. J.; Nosco, D. L. *Adv. Inorg. Bioinorg. Mech.* **1982**, *1*, 269.

(84) Miller, S. M.; Massey, V.; Ballou, D.; Williams, C. H.; Jr.; Distefano, M. D.; Moore, M. J.; Walsh, C. T. *Biochemistry* **1990**, *29*, 2831–2841.

(85) Singh, A. N.; Gelerinter, E.; Gould, E. S. *Inorg. Chem.* **1982**, *21*, 1232.

(86) Fox, B. S. Ph.D. Thesis, Massachusetts Institute of Technology, 1983.

Mechanistic and Evolutionary Aspects of Vitamin B_{12} Biosynthesis

A. Ian Scott

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

Received February 23, 1990 (Revised Manuscript Received May 21, 1990)

Introduction

Since the discovery of vitamin B_{12} as the anti-pernicious anemia factor, chemists and biochemists alike have been fascinated not only by the complex problems inherent in the structure of B_{12}, the biochemistry mediated by its coenzyme, and the total synthesis of the corrin macrocycle but also by the challenge of unraveling the biosynthetic pathway, which is shared at the outset by nature's route to the other biologically important metallopigments—heme and siroheme (iron), chlorophyll (magnesium), and coenzyme F430 (nickel)—before diverging along the corrin (cobalt) branch.

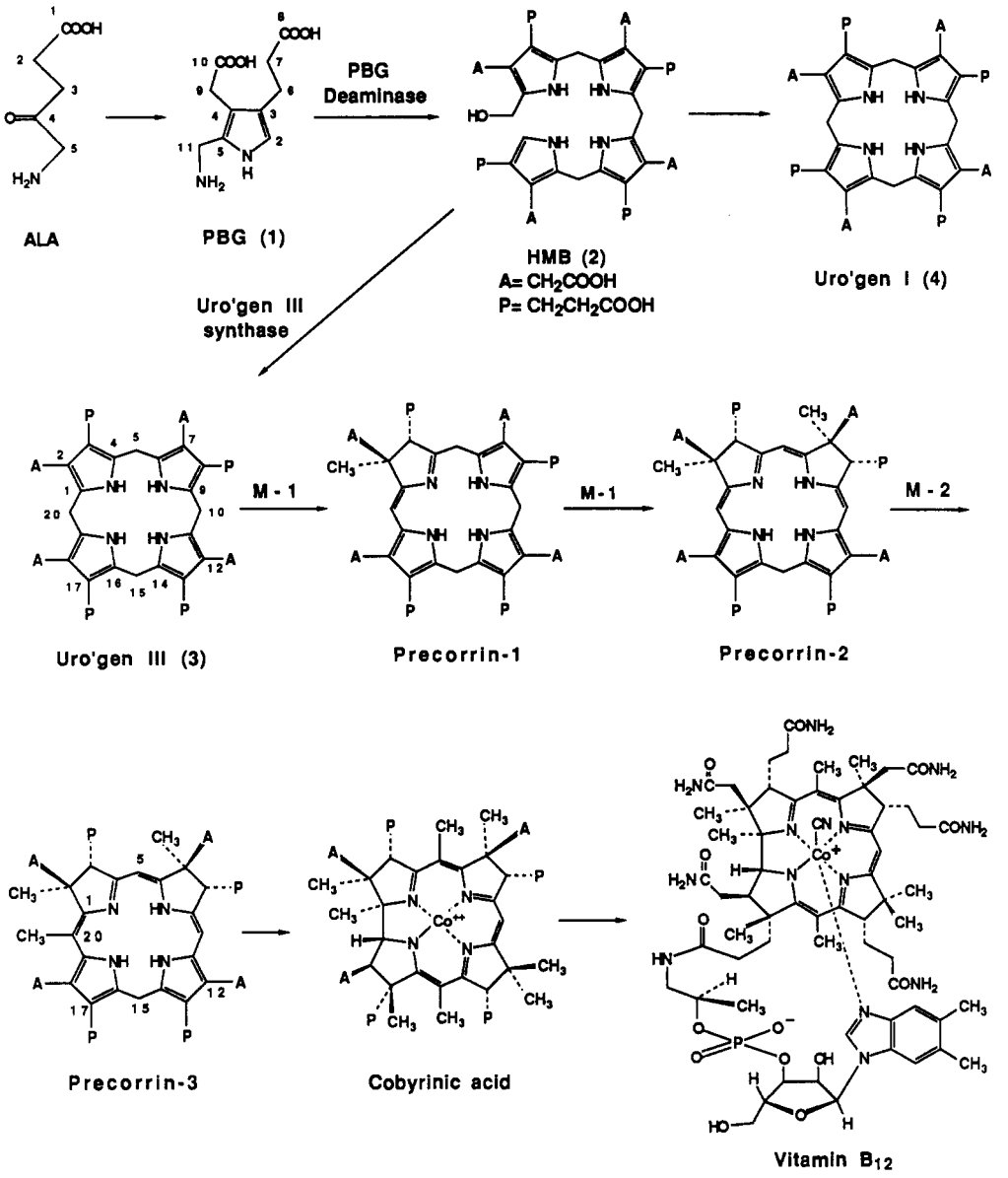
For over 20 years we and others have been studying the B_{12} pathway, the first decade of which was reviewed in earlier Accounts.[1] The results of these endeavors are summarized in Scheme I, which introduces the early intermediates 5-aminolevulinic acid (ALA) and por-

A. Ian Scott was born in Glasgow, Scotland, in 1928. After receiving B.Sc., Ph.D., and D.Sc. degrees from Glasgow University (with R. A. Raphael), he spent a postdoctoral year with M. S. Newman (Ohio State) and, following an industrial position (Imperial Chemical Industries), joined Sir Derek Barton's laboratory for a second postdoctoral period in London and Glasgow. His academic career began in 1957 at Glasgow and was followed by Professorships at U.B.C. (Vancouver), Sussex (U.K.), Yale, and finally Texas A&M University, where he has remained, apart from a brief return to Scotland in 1980–1981, since 1977. He is a Fellow of the Royal Societies of London and Edinburgh and a founding member of the Yale Sherlock Holmes Society. His research interests have been focused for the last 20 years on the biosynthesis of porphyrins and B_{12} and on enzyme mechanisms, investigated by using a combination of NMR spectroscopy, organic chemistry and, more recently, molecular biology.

(1) (a) Scott, A. I. *Acc. Chem. Res.* **1978**, *11*, 29. (b) Battersby, A. R.; McDonald, E. *Acc. Chem. Res.* **1979**, *12*, 14.

0001-4842/90/0123-0308$02.50/0 © 1990 American Chemical Society

Scheme I



phobilinogen (PBG, 1) and their enzymatic transformation to the macrocyclic template of uro'gen III (3) via the hydroxymethylbilane (2) and to the partially methylated precorrins-1–3. These represent successive C-methylation of 3 by two methyltransferases, M-1 and M-2, with *S*-adenosylmethionine (SAM). In sharp contrast to the slow but steady acquisition of the biosynthetic sequence over two decades, the last few years have witnessed an exponential increase in the rate of progress thanks to a powerful combination of molecular biology and high-field pulse-programmed NMR spectroscopy, and as a result of these fresh approaches we can now predict with confidence that most of the remaining intermediates and their associated biosynthetic enzymes will be discovered in the near future.

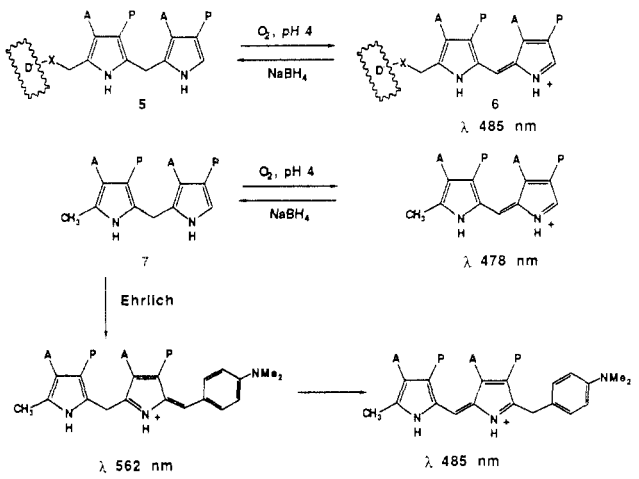
In this Account we illustrate the basis for our optimism by focusing on several recent developments associated with the "early" and "late" segments of Scheme I, including the assembly of the macrocyclic templates, the definition of the sequence of C-methylation, and finally a proposal for the overall mechanism of corrin synthesis based largely on in vitro models, which in turn casts light on the evolution of the type-III macrocycles from a "primordial" corrin template. Most of the examples are taken from our own research, but references to independent studies from the laboratories of A. R. Battersby (Cambridge), G. Müller (Stuttgart), and P. M. Jordan (London), which have provided valuable and stimulating contributions to porphyrin and corrin biosynthesis, will be found in the text.

The Enzymes of Tetrapyrrole Synthesis: PBG Deaminase and Uro'gen III Synthase

PBG deaminase (EC 4.3.1.8) catalyzes the tetramerization of PBG (1) to preuro'gen (hydroxymethylbilane, HMB; 2),[2,3] which is cyclized with rearrangement to the unsymmetrical uro'gen III (3) by uro'gen III synthase[3–5] (EC 4.2.1.75) (Scheme I). In the absence of the latter enzyme, preuro'gen (2) cyclizes to uro'gen I (4), which,

(2) Burton, G.; Fagerness, P. E.; Hosazawa, S.; Jordan, P. M.; Scott, A. I. *J. Chem. Soc., Chem. Commun.* **1979**, 202.
(3) Battersby, A. R.; Fookes, C. J. R.; Gustafson-Potter, K. E.; Matcham, G. W. J.; McDonald, E. *J. Chem. Soc., Chem. Commun.* **1979**, 1115.
(4) Jordan, P. M.; Burton, G.; Nordlov, H.; Schneider, M. M.; Pryde, L. M.; Scott, A. I. *J. Chem. Soc., Chem. Commun.* **1979**, 204.
(5) Reviews, Leeper, F. J. *Nat. Prod. Rep.* **1985**, *2*, 19, 561; **1987**, *4*, 441; **1989**, *6*, 171.

Scheme II



as discussed below, turns out to be a substrate for the methylases of the vitamin B_{12} pathway. We have used genetic engineering to construct a plasmid pBG 101 containing the *Escherichia coli hemC* gene[6,7] encoding deaminase. *E. coli* (TB1) transformed with this plasmid produces deaminase at levels greater than 200 times those of the wild strain[7] thereby allowing access to substantial quantities of enzyme for detailed study of the catalytic mechanism.

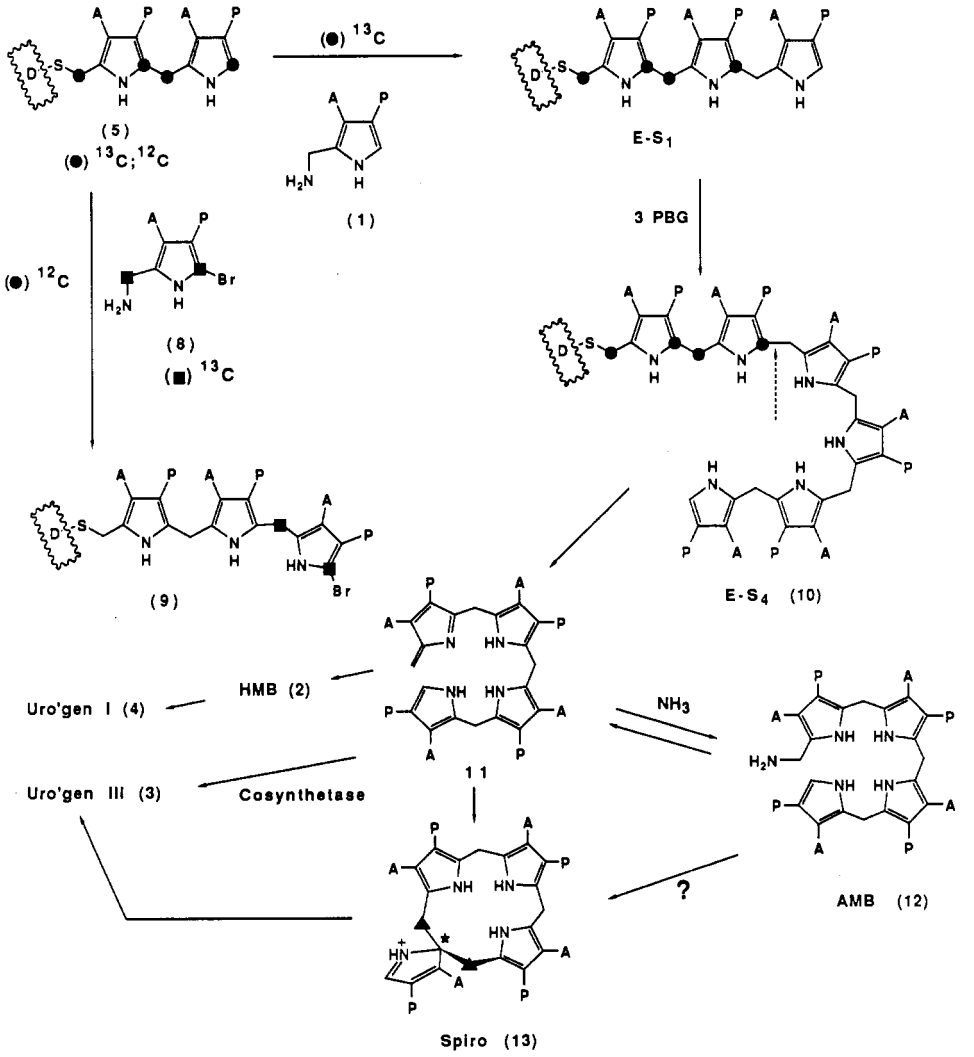
Previous work with deaminase[8-11] had established that a covalent bond is formed between substrate and enzyme, thus allowing isolation of covalent complexes containing up to three PBG units (ES_1–ES_3). Application of ^3H NMR spectroscopy to the mono-PBG adduct (ES_1) revealed a rather broad ^3H chemical shift indicative of bond formation with a cysteine thiol group at the active site.[12] However, with adequate supplies of pure enzyme available from the cloning of *hemC* it was shown that a novel cofactor, derived from PBG during the biosynthesis of deaminase, is covalently attached to one of the four cysteine residues of the enzyme in the form of a dipyrromethane which, in turn, becomes the site of attachment of the succeeding 4 mol of substrate during the catalytic cycle. Thus, at pH < 4, deaminase (**5**) rapidly develops a chromophore (λ_{max} = 485 nm) diagnostic of a pyrromethene (**6**), while reaction with Ehrlich's reagent generates a chromophore typical of a dipyrromethane (λ_{max} = 560 nm) changing to 490 nm after 5–10 min. The latter chromophoric interchange was identical with that of the Ehrlich reaction of the synthetic model pyrromethane (**7**) and can be ascribed to the isomerization shown (Scheme II) for the model system (**7**). Incubation of *E. coli* strain SASX41B/pBG101 (*hemA*$^-$ requiring ALA for growth) with [5-^{13}C]ALA afforded highly enriched enzyme for NMR studies. At pH 8, the enriched carbons of the dipyrromethane (py-CH_2-py) were clearly recognized at 24.0 ppm (py-CH_2-py), 26.7 ppm (py-CH_2-X), 118.3 ppm (α-free pyrrole), and 129.7 ppm (α-substituted pyrrole) (see **5**, Scheme III). The signals are sharpened at pH 12, and the CH_2X-resonance is shifted to δ 29.7 in the unfolded enzyme. Comparison with synthetic models reveals that a shift of 26.7 ppm is in the range expected for an α-thiomethyl pyrrole (py-CH_2-SR). Confirmation of the dipyrromethane (rather than oligo pyrromethane) came from the ^{13}C INADEQUATE spectrum taken at pH 12, which reveals the expected coupling only between py-CH_2-py (δ 24.7) and the adjacent substituted pyrrole carbon (δ 128.5 ppm). When the enriched deaminase was studied by ^1H detected heteronuclear filtered spectroscopy, only five protons (in a total of *several thousand*!) attached to ^{13}C-enriched nuclei were observed in accord with structure **5**. A specimen of deaminase was then covalently inhibited with the suicide inhibitor [2,11-$^{13}C_2$]-2-bromo-PBG (**8**, Scheme III) to give a spectrum consistent only with structure **9**. The site of covalent attachment of substrate (and inhibitor) is therefore *the free α-pyrrole carbon at the terminus of the dipyrromethane* in the native enzyme, leading to the structural and mechanistic proposal for deaminase portrayed in Scheme III.

It was then found that 2 mol of PBG is incorporated autocatalytically into the apoenzyme (isolated via overexpression in a *hemB*$^-$ strain of *E. coli* which *does not make PBG*)[13] and that the first (kinetic) encounter of PBG deaminase with substrate involves attachment of PBG (with loss of NH_3) to the α-free pyrrole position of the dipyrromethane to form the ES_1 complex[14,15] (Scheme III). The process is repeated to produce the "tetra-PBG" (ES_4) adduct (**10**). At this juncture site-specific cleavage of the *hexapyrrole* chain (at →) releases the azafulvene bilane (**11**) which either becomes the substrate of uro'gen III synthase or, in the absence of the latter enzyme, is stereospecifically hydrated[3] to HMB (**2**) at pH 12 and then cyclized chemically to uro'gen I (**4**) at pH ≤8. Independent, complementary work in 1987–1988 from two other laboratories[16,17] reached identical conclusions regarding the catalytic site. The ^{13}C labeling defines the *number* of PBG units (two) attached in a head-to-tail motif to the native enzyme at pH 8 and reveals the identity of the nucleophilic group (Cys-SH) which anchors the dipyrromethane (and hence the growing oligopyrrolic chain) to the enzyme. Site-specific mutagenesis[7] and chemical cleavage[16,17] were employed to determine that Cys-242 is the point of attachment of the cofactor. Thus, replacement[7] of cysteine with serine at residues 99 and 242 (respectively) gave fully active and inactive specimens of the enzyme, respectively.

The use of the α-carbon of a dipyrromethane as the nucleophilic group responsible for oligomerization of 4 mol of PBG (with loss of NH_3 at each successive en-

(6) Scott, A. I. *J. Heterocycl. Chem.* **1987**, *24*, S-75.
(7) Scott, A. I.; Roessner, C. A.; Stolowich, N. J.; Karuso, P.; Williams, H. J.; Grant, S. K.; Gonzalez, M. D.; Hoshino, T. *Biochemistry* **1988**, *27*, 7984.
(8) Anderson, P. M.; Desnick, R. J. *J. Biol. Chem.* **1980**, *255*, 1993.
(9) Berry, A.; Jordan, P. M.; Seehra, J. S. *FEBS Lett.* **1981**, *129*, 220.
(10) Battersby, A. R.; Fookes, C. J. R.; Hart, G.; Matcham, G. W. J.; Pandey, P. S. J. *J. Chem. Soc., Perkin Trans. 1* **1983**, 3041.
(11) Jordan, P. M.; Berry, A. *Biochem. J.* **1981**, *195*, 177.
(12) Evans, J. N. S.; Burton, G.; Fagerness, P. E.; Mackenzie, N. E.; Scott, A. I. *Biochemistry* **1986**, *25*, 905.
(13) Scott, A. I.; Roessner, C. A.; Clemens, K. R.; Stolowich, N. J.; Santander, P. J.; Gonzalez, M. D. *FEBS Lett.* **1988**, *242*, 319.
(14) Scott, A. I.; Stolowich, N. J.; Williams, H. J.; Gonzalez, M. D.; Roessner, C. A.; Grant, S. K.; Pichon, C. *J. Am. Chem. Soc.* **1988**, *110*, 5898.
(15) Jordan, P. M.; Warren, M. J.; Williams, H. J.; Stolowich, N. J.; Roessner, C. A.; Grant, S. K.; Scott, A. I. *FEBS Lett.* **1988**, *235*, 189.
(16) (a) Warren, M. J.; Jordan, P. M. *FEBS Lett.* **1987**, *225*, 87. (b) Warren, M. J.; Jordan, P. M. *Biochemistry* **1988**, *27*, 9020.
(17) (a) Hart, G. J.; Miller, A. D.; Leeper, F. J.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1987**, 1762. (b) Miller, A. D.; Hart, G. J.; Packman, L. C.; Battersby, A. R. *Biochem. J.* **1988**, *254*, 915. (c) Hart, G. J.; Miller, A. D.; Battersby, A. R. *Biochem. J.* **1988**, *252*, 909. (d) Beifus, U.; Hart, G. J.; Miller, A. D.; Battersby, A. R. *Tetrahedron Lett.* **1988**, *29*, 2591.

Scheme III



counter with an α-free pyrrole) not only is remarkable for the exquisite specificity and control involved but is, as far as we know, a process unique in the annals of enzymology in that a *substrate* is used not only once but *twice* in the genesis of the active-site cofactor! Even more remarkable is the fact that the apoenzyme is automatically transformed to the active holoenzyme by addition of two PBG units without the intervention of a second enzyme. Crystallization and X-ray diffraction studies of both the native enzyme and several of its genetically altered versions are now in progress.

Uro'gen III Synthase: The Ring-D Switch

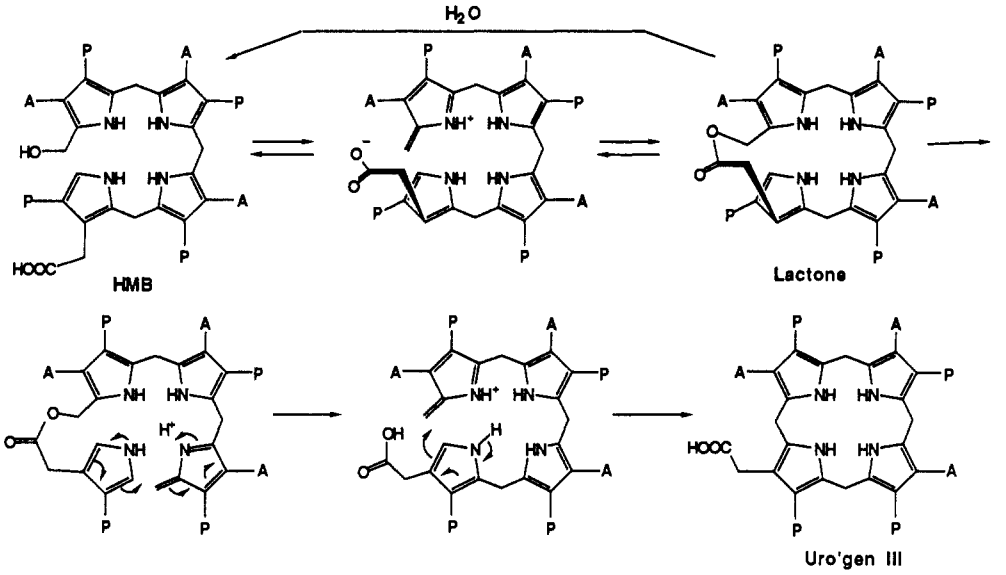
We now turn briefly to the *rearranging* enzyme uro'gen III synthase. The early ideas of Bogorad[18] involving the aminomethylbilane (AMB; **12**) as substrate were helpful in finally tracking down the elusive species which, after synthesis by PBG deaminase, becomes the substrate for uro'gen III synthase. This was shown to be HMB (**2**), and at this stage (1978) the lack of activity of AMB (**12**) as substrate relegated the latter bilane to an interesting artifact produced quantitatively by addition of ammonia to deaminase incubations (Scheme III). However, with the acquisition of substantial quantities of pure uro'gen III synthase obtained by cloning the genes *hemC* and *hemD together*[19] and overexpression in *E. coli*, the substrate specificity of the synthase (often called cosynthetase) has been reinvestigated. Ever since its conception[20] by Mathewson and Corwin in 1961, the spiro compound (**13**) (Scheme III) has been a favorite construct with organic chemists, since both its genesis through α-pyrrolic reactivity and its fragmentation–recombination rationalize the intramolecular formation of uro'gen III from the linear bilane, preuro'gen (HMB; **2**). A careful search for the spiro compound (**13**) was conducted at subzero temperatures in cryosolvent (−24 °C; ethylene glycol/buffer) by using various ^{13}C isotopomers of HMB as substrate.[19] Although the synthase reaction could be slowed down to 20 h (rather than 20 s), no signals corresponding to the quaternary carbon ((*); δ ~80 ppm) or to the α-pyrrolic methylene groups (▲; δ 35–40 ppm) could be observed. During these studies, however, it was found that totally synthetic AMB (**12**) (thus free of deaminase which can catalyze interconversion of AMB into HMB, **2**) served as a slow but productive substrate for uro'gen III synthase *at high concentrations* (1 mmol in enzyme and substrate concentrations) and that if care is not taken to remove the ammonia liberated from PBG by the action of deaminase, not only is the enzymatic

(18) Radmer, R.; Bogorad, L. *Biochemistry* **1972**, *11*, 904.

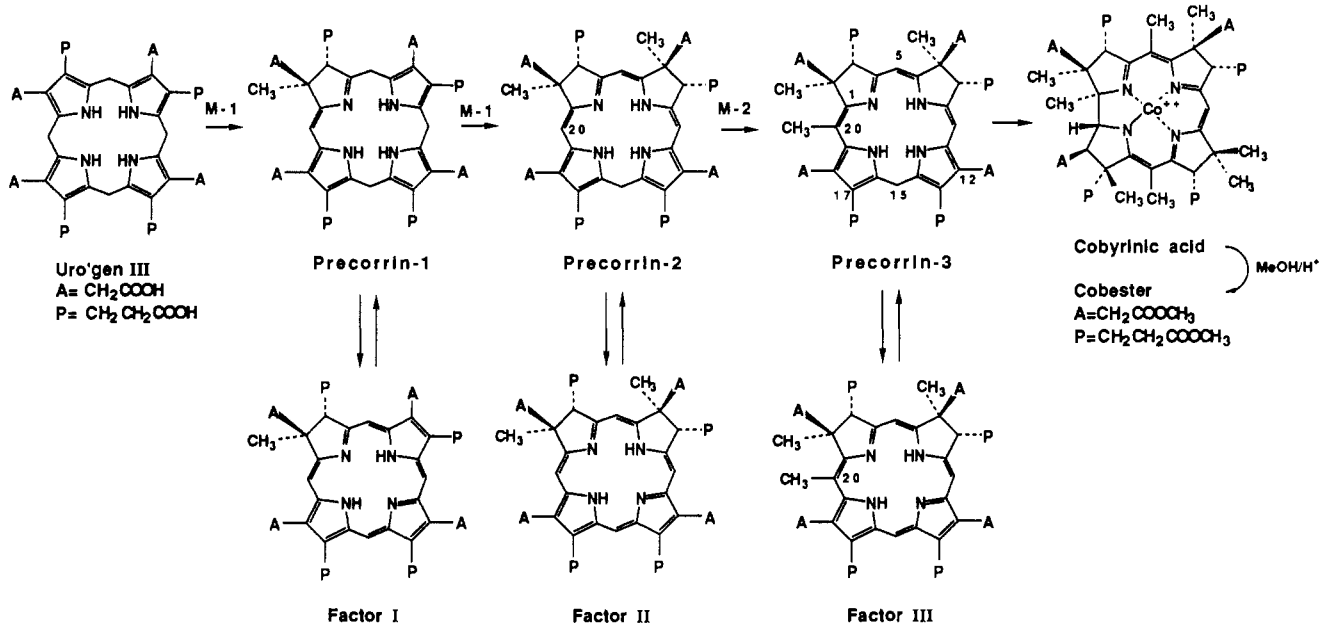
(19) Scott, A. I.; Roessner, C. A.; Karuso, P.; Stolowich, N. J.; Atshaves, B.; Xue, T., manuscript in preparation. Scott, A. I. *17th IUPAC Symp. Nat. Prod.*, New Delhi, Feb 1990; *Pure Appl. Chem.* **1990**, *62*, 1269.

(20) Mathewson, J. H.; Corwin, A. H. *J. Am. Chem. Soc.* **1961**, *83*, 135.

Scheme IV



Scheme V



formation of AMB observed[21] but, in the presence of uro'gen III synthase, the product is again uro'gen III. Regardless of which version of the bilane system (HMB, AMB) is involved, the lack of observation of any intervening species free of enzyme strongly suggests that the intermediate is enzyme-bound and therefore difficult to detect. Indirect evidence for the "spiro" mechanism was adduced when it was found that one of the enantiomers of the corresponding synthetic ring D *lactam* served as an inhibitor of uro' gen III synthase.[22a] However, a novel alternative to the "spiro intermediate" (**13**) hypothesis, still consistent with the known facts, has been proposed[19] using a self-assembly concept involving *lactone* formation as portrayed in Scheme IV. We have suggested that generation of the azafulvene (**11**) (as before) is followed not by carbon–carbon bond formation (→**13**) but by closure of the macrocycle via the macrolide (Scheme IV). The regiospecificity of this step is reflected by the failure[5] of synthetic bilanes lacking the acetic acid side chain at position 17 in ring D to undergo enzyme-catalyzed rearrangement to the type-III system, together with the observation that a "switched" bilane carrying a propionate at position 17 (and an acetate at 18) does indeed give a certain amount of (rearranged) uro'gen I (but also uro'gen III) enzymatically,[5,22b] indicating that lactone formation could also be achieved (but less efficiently) by a propionate group at the 17-position. The subsequent chemistry is quite similar to the fragmentation–recombination postulated for the spiro system except that a "twisted macrolide" becomes the pivotal intermediate species and the sequence proceeds as shown in Scheme IV. Although this novel hypothesis is rather difficult to test experimentally, specific chemical traps for the lactones combined with the use of ^{18}O labeling are being used

(21) Rosé, S.; Frydman, R. B.; de los Santos, C.; Sburlati, A.; Valasinas, A.; Frydman, B. *Biochemistry* **1988**, *27*, 4871.

(22) (a) Stark, W. M.; Baker, M. G.; Raithly, P. R.; Leeper, F. J.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1985**, 1294. (b) Battersby, A. R.; Fookes, J. R.; Matcham, G. W. J.; McDonald, E.; Hollenstein, R. *J. Chem. Soc., Perkin Trans. 1* **1983**, 3031.

to confirm or refute the possible role of such macrocyclic lactones as the basis for the mechanism of uro'gen III synthase.

Temporal Resolution of the Methylation Sequence

The bioconversion of uro'gen III to cobyrinic acid is summarized in Scheme V, where it is postulated that the *oxidation level of the various intermediates is maintained at the same level as that of uro'gen III.* The lability of the reduced isobacteriochlorins, the fact that they are normally isolated in the oxidized form, and the requirement for chemical reduction of factor I (but not factors II and III) to the corresponding precorrins, before incorporation into corrin in cell-free systems, lent credence to this idea, but rigorous proof for the in vivo oxidation state of factors I and II has only recently been obtained as discussed below.[23,24] In order to distinguish between the oxidation levels of the *isolated* chlorins, isobacteriochlorins, corphins, etc. and to avoid possible confusion in using the term "factor" (which has other connotations in B_{12} biochemistry, e.g., intrinsic factor/factor III), the term precorrin-**n** has been suggested[24] for the actual structures of the biosynthesized intermediates after uro'gen III where **n** denotes the number of SAM-derived methyl groups. Thus, although the names factors I–III for the *isolated* species will probably survive (for historical reasons), all of the true intermediates probably have the same oxidation as uro'gen III, viz., tetrahydro factor I (precorrin-1), dihydro factor II (precorrin-2), and dihydro factor III (precorrin-3) (Scheme V).

In spite of intensive search, no new intermediates containing four or more methyl groups (up to a possible total of eight) have been isolated, but the biochemical conversion of factor III to cobyrinic acid must involve the following steps (see Scheme V), not necessarily in the order indicated: (1) the successive addition of five methyls derived from *S*-adenosylmethionine (SAM) to *reduced* factor III (precorrin-3); (2) the contraction of the permethylated macrocycle to corrin; (3) the extruction of C-20 and its attached methyl group leading to the isolation of acetic acid;[25-28] (4) decarboxylation of the acetate side chain in ring C (C-12); (5) insertion of Co^{3+} after adjustment of oxidation level from Co^{2+}. In order to justify the continuation of the search for such intermediates whose inherent lability to oxygen is predictable, ^{13}C pulse labeling methods were applied to the cell-free system from *Propionibacterium shermanii* which converts uro'gen III (3) to precorrin-2 and thence to cobyrinic acid, a technique used previously in biochemistry to detect the flux of radiolabels through the intermediates of biosynthetic pathways.

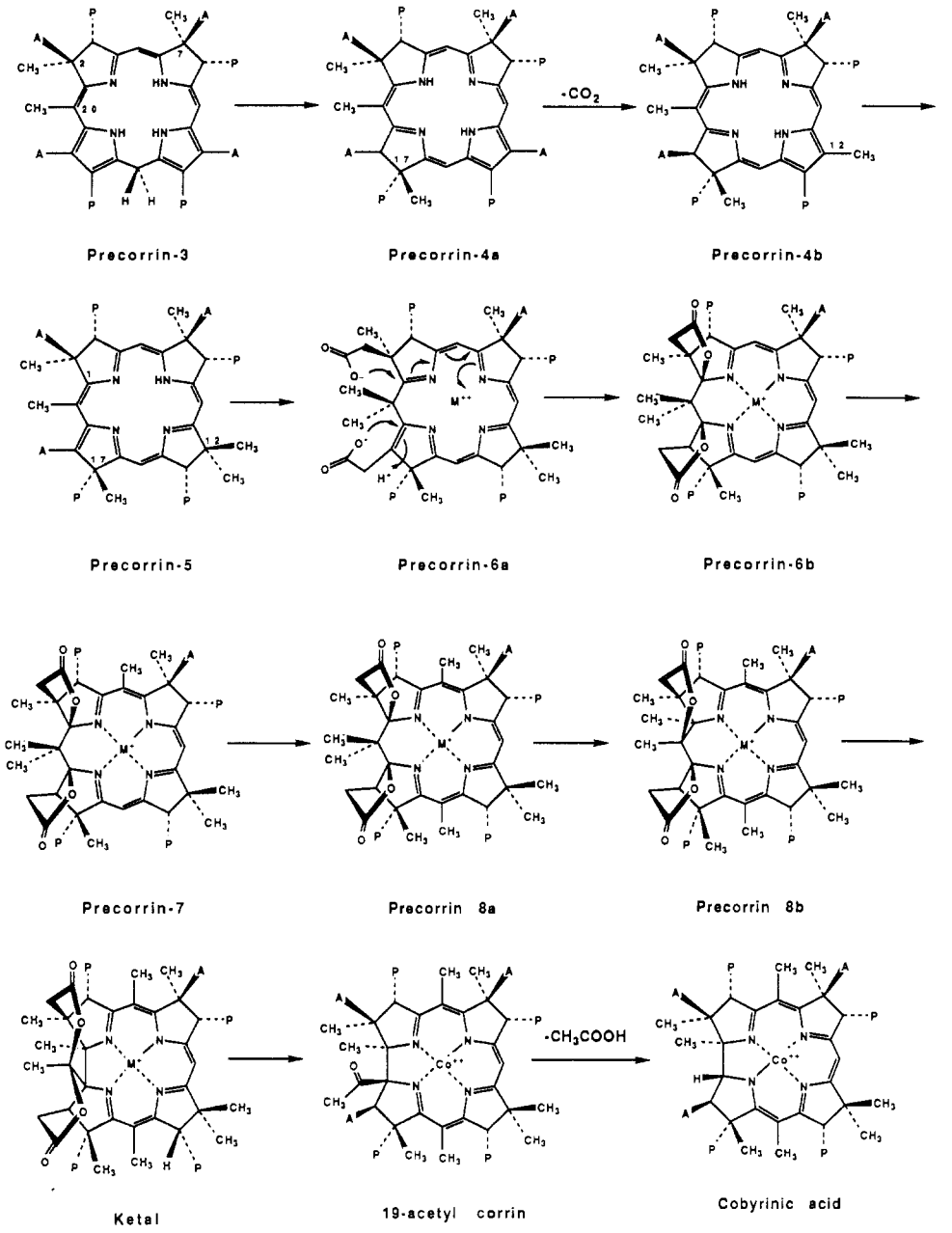
The Pulse Experiments

The methylation sequence can be resolved temporally provided that *enzyme-free* intermediates accumulate in sufficient pool sizes to affect the intensities of the resultant methyl signals in the ^{13}C NMR spectrum of the target molecule, cobyrinic acid, when the cell-free system is challenged with a pulse of $^{12}CH_3$-SAM followed by a second pulse of $^{13}CH_3$-SAM (or vice versa) at carefully chosen intervals in the total incubation time. By this approach it is possible to "read" the biochemical history of the methylation sequence as reflected in the dilution (or enhancement in the reverse experiment) of $^{13}CH_3$ label in the methionine-derived methyl groups of cobyrinic acid after conversion to cobester, whose ^{13}C NMR spectrum has been assigned.[29,30] Firstly, precorrin-2 is accumulated in whole cells containing an excess of SAM in the absence of Co^{2+}. The cells are then disrupted, and Co^{2+} is added immediately, followed by a pulse of $^{13}CH_3$-SAM (90 atom %) after 4 h. After a further 1.5 h, cobyrinic acid is isolated as cobester. The ^{13}C NMR spectrum of this specimen defines the complete methylation sequence, beginning from precorrin-2, as C-20 > C-17 > C-12α > C-1 > C-5 > C-15, with a differentiation of 25% (±5%) in the relative signal intensities for the C_5 and C_{15} methyl groups[31,32] further confirmed by hetero-filtered ^1H spectroscopy of the ^{13}C-enriched sample. Methylation at C-20 of precorrin-2 to give precorrin-3 is not recorded in the spectrum of cobester since C-20 is lost on the way to cobyrinic acid, together with the attached methyl group, in the form of acetic acid.[25-28] The sequence C-17 > C-12α > C-1 has been found in *Clostridium tetanomorphum*,[24] and further differentiation between C-5 and C-15 insertion suggests the order C-15 > C-5 in this organism for the last two methylations, i.e., opposite from the *P. shermanii* sequence.

It is now apparent that several discrete methyl transferases are involved in the biosynthesis of cobyrinic acid from uro'gen III, since enzyme-free intermediates must accumulate in order to dilute the ^{13}C label. A rationale for these events is given in Scheme VI, which takes the following facts into account: (a) the methionine-derived methyl group at C-20 of precorrin-3 does not migrate to C-1 and is expelled together with C-20 from a late intermediate (as yet unknown) in the form of acetic acid; (b) neither 5,15-norcorrinoids[33] nor descobaltocobyrinic acid[34] are biochemical precursors of cobyrinic acid; (c) regiospecific loss of ^{18}O from [1-^{13}C,1-$^{18}O_2$]-5-aminolevulinic acid derived cyanocobalamin from the ring-A acetate occurs,[35] in accord with the Eschenmoser concept of lactone formation (see below) as portrayed in Scheme VI, where precorrin-5

(23) Battersby, A. R.; Frobel, K.; Hammerschmidt, F.; Jones, C. *J. Chem. Soc., Chem. Commun.* **1982**, 455–457.
(24) (a) Uzar, H. C.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1982**, 1204; (b) *Ibid.* **1985**, 585. (c) Uzar, H. C.; Battersby, A. R.; Carpenter, T. A.; Leeper, F. J. *J. Chem. Soc., Perkin Trans. 1* **1987**, 1689–1696 and references cited therein.
(25) Nussbaumer, C.; Imfeld, M.; Wörner, G.; Müller, G.; Arigoni, D. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 9–10.
(26) Battersby, A. R.; Bushnell, M. J.; Jones, C.; Lewis, N. G.; Pfenninger, A. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 13–15.
(27) Mombelli, L.; Nussbaumer, C.; Weber, H.; Müller, G.; Arigoni, D. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 11–12.
(28) Müller, G.; Gneuss, K. D.; Kriemler, H.-P.; Irwin, A. J.; Scott, A. I. *Tetrahedron (Suppl).* **1981**, *37*, 81–90.
(29) Ernst, L. *Liebigs Ann. Chem.* **1981**, 376–386.
(30) Battersby, A. R.; Edington, C.; Fookes, C. J. R.; Hook, J. M. *J. Chem. Soc., Perkin Trans. 1* **1982**, 2265–2277.
(31) Scott, A. I.; Mackenzie, N. E.; Santander, P. J.; Fagerness, P.; Müller, G.; Schneider, E.; Sedlmeier, R.; Wörner, G. *Bioorg. Chem.* **1984**, *13*, 356–362.
(32) Scott, A. I.; Williams, H. J.; Stolowich, N. J.; Karuso, P.; Gonzalez, M. D.; Müller, G.; Hlineny, K.; Savvidis, E.; Schneider, E.; Traub-Eberhard, U.; Wirth, G. *J. Am. Chem. Soc.* **1989**, *111*, 1897–1900.
(33) Nussbaumer, C.; Arigoni, D., work described in C. Nussbaumer's Dissertation No. 7623, E.T.H., Zurich, 1984.
(34) Podschun, T. E.; Müller, G. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 46–47.
(35) Kurumaya, K.; Okasaki, T.; Kajiwara, M. *Chem. Pharm. Bull.* **1989**, *37*, 1151.

Scheme VI

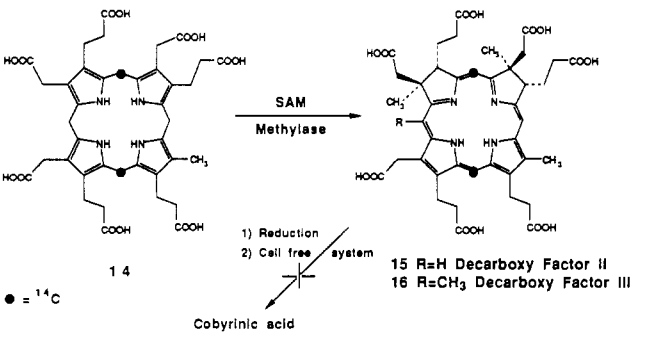


is methylated at C-20 followed by C-20 → C-1 migration and lactonization.

Timing of the Decarboxylation Step in Ring C

It has also proved possible to define the point in the biosynthetic sequence where ring-C decarboxylation occurs, using synthetic 5,15-$^{14}C_2$-labeled 12-decarboxylated uro'gen III (**14**) as a substrate for the nonspecific methylases of *P. shermanii* to prepare the 12-methyl analogues **15** and **16** of factors II and III respectively (Scheme VII). It was found[32] that these ring-C decarboxylated analogues are *not* substrates for the enzymes of corrin biosynthesis, leading to the conclusions that (a) in normal biosynthesis, precorrin-**3** is *not* the intermediate that is decarboxylated and (b) decarboxylation occurs at some stage after the fourth methylation (at C-17) and, by mechanistic analogy, before the fifth methylation at C-12. Hence, two pyrrocorphin intermediates, viz., precorrins-**4a** and -**4b**, should intervene between precorrins-**3** and -**5**, i.e., precorrin-**3** is C-methylated at C-17 to give precorrin-**4a** followed by

Scheme VII



decarboxylation (→**4b**) and subsequent C-methylations at C-12α, C-1, C-5, and C-15, in that order as suggested in Scheme VI. Although this sequence differs in the penultimate stages from that reported elsewhere,[24c] consensus over the timing of decarboxylation in ring C has been reached since identical conclusions have been reported using *Pseudomonas denitrificans*.[36]

The decarboxylation mechanism most plausibly involves the addition of a proton to a ring-C exomethylene species. The stereospecificity (retention) of this step has been demonstrated by Kuhn-Roth oxidation of the chirally labeled C-12β methyl group of cobester following incorporation of ALA, chiral (^{2}H, ^{3}H) on the methylene group, which becomes incorporated into the C-12 acetate side chain.[37]

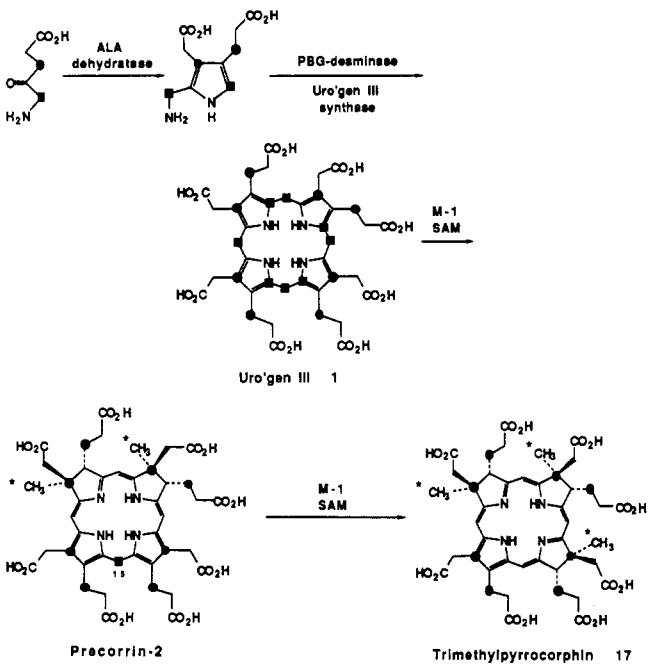
Although factors I–III, when adjusted to the requisite oxidation level (i.e., precorrins-**1**–**3**), all serve as excellent substrates for conversion to cobyrinic acid in cell-free systems from *P. shermanii* and *C. tetanomorphum*, direct evidence for their sequential interconversion was obtained only recently. In the presence of $^{13}CH_3$-SAM and the complete cell-free system containing a reducing system (factor II → precorrin-**2**), a specimen of factor II is converted to factor III containing a single ^{13}C-enriched methyl at C-20 (δ 19.2 ppm). The demonstration of this conversion removes any doubt that separate pathways could exist for the biosynthesis of vitamin B_{12} from precorrins-**2** and -**3**, respectively.[32]

We now return to the construction of a working hypothesis for corrin biosynthesis. The formulations precorrins-**6b**, -**7**, -**8a**, and -**8b** take into account the idea of lactone formation using rings A and D acetate side chains. The migration C-20 → C-1 could be acid or metal-ion catalyzed and the resultant C-20 carbonium ion quenched with external hydroxide or by the internal equivalent from the carboxylate anion of the C-2 acetate in ring A (precorrin-**8a** → precorrin-**8b**) as suggested by the results of labeling with ^{18}O discussed above.[35] In any event, the resultant dihydrocorphinol–bislactone precorrin-**8b** is poised to undergo the biochemical counterpart of Eschenmoser ring contraction[38] to the 19-acetylcorrin. Before this happens we suggest that the final methyl groups are added at C-5 (precorrin-**7**) and then C-15 (precorrin-**8a**) to take account of the nonincorporation of the 5,15-norcorrinoids.[33] The resultant precorrin-**8b** (most probably with cobalt in place) then contracts to 19-acetylcorrin which, by loss of acetic acid, leads to cobyrinic acid, a process that is accompanied by a kinetic deuterium isotope effect at C-19.[41,42] Cobalt insertion must precede methylation at C-5 and C-15 since hydrogenocobyrinic acid[39] does not insert cobalt enzymatically.[34] The valency change Co^{2+} → Co^{3+} during or after cobalt insertion has so far received no explanation.

The Methyltransferases

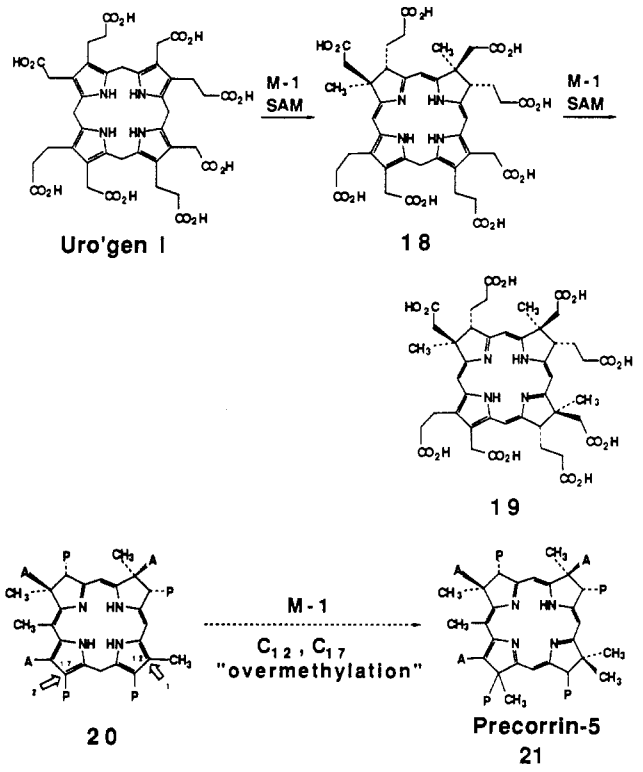
The first of the methylase enzymes catalyzes the sequential formation of factors I and II and has been named *S*-adenosylmethionine uro'gen III methyl transferase (SUMT). SUMT was first partially purified from *P. shermanii* by G. Müller[43] and recently has been overexpressed in *P. denitrificans*.[44] 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 was achieved by the appropriate genetic engineering.[45] Although SUMT and M-1 appear to perform the same task, it has been found that their substrate specificities differ. Thus, it has been possible to study in detail the reaction catalyzed by M-1 directly using NMR spectroscopy and to provide rigorous proof that the structure of precorrin-**2** is that of the dipyrrocorphin tautomer of dihydro factor II. Uro'gen III (enriched from [5-^{13}C]ALA at the positions shown in Scheme VIII) was incubated with M-1 and [$^{13}CH_3$]SAM. The resultant spectrum of the precorrin-**2** revealed an sp^3-enriched carbon at C-15, thereby locating the reduced center (shown as ■). By using a different set of ^{13}C labels (● from [3-^{13}C]ALA) and [$^{13}CH_3$]SAM), the sp^2 carbons at C-12 and C-18 were located as well as the sp^3 centers coupled to the pendant ^{13}C-methyl groups at (*) C-2 and C-7. This result confirms an earlier NMR analysis[23] of precorrin-**2** isolated by careful anaerobic purification of the methyl ester and shows that no further tautomerism takes place during the latter procedure. The two sets of experiments mutually reinforce the postulate that precorrins-**1**, -**2**, and -**3** all exist as hexahydroporphinoids, and recent labeling experiments[46] have provided good evidence that precorrin-**1** is discharged from the methylating enzyme (SUMT) as the species with the structure shown in Scheme V (or a tautomer thereof).

Scheme VIII



(36) Blanche, F.; Handa, S.; Thibaut, D.; Gibson, C. L.; Leeper, F. J.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1988**, 1117–1119.
(37) Battersby, A. R.; Deutscher, K. R.; Martinoni, B. *J. Chem. Soc., Chem. Commun.* **1983**, 698–700.
(38) Rasettti, V.; Pfaltz, A.; Kratky, C.; Eschenmoser, A. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 16–19.
(39) Dresov, B.; Ernst, L.; Grotjahn, L.; Koppenhagen, V. B. *Angew. Chem., Int. Ed. Engl.* **1981**, *20*, 1048–1049.
(40) Scott, A. I.; Georgopapadakou, N. E.; Irwin, A. J., unpublished.
(41) Scott, A. I.; Kajiwara, M.; Santander, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6616–6618.
(42) Battersby, A. R.; Edington, C.; Fookes, C. I. R. *J. Chem. Soc., Chem. Commun.* **1984**, 527–530.
(43) Müller (Müller, G. In *Vitamin B_{12}*; Zagalak, B., Friedrich, W., Eds.; de Gruyter: New York, 1979; pp 279–291) describes a methylase from *P. shermanii* capable of using uro'gen I as substrate. The product was not characterized.
(44) Cameron, B.; Briggs, K.; Pridmore, S.; Brefort, G.; Crouzet, J. *J. Bacteriol.* **1989**, *171*, 547–557. Blanche, F.; Debussche, D.; Thibaut, D.; Crouzet, J.; Cameron, B. *Ibid.* **1989**, *171*, 4222.
(45) Warren, M. J.; Roessner, C. A.; Santander, P. J.; Scott, A. I. *Biochem. J.* **1990**, *265*, 725–729.
(46) Brunt, R. D.; Leeper, F. J.; Orgurina, I.; Battersby, A. R. *J. Chem. Soc., Chem. Commun.* **1989**, 428.

Scheme IX



However, prolonged incubation (2 h) of uro'gen III with M-1 provided a surprising result: the UV and NMR changed dramatically from that of precorrin-2 (a dipyrrocorphin) to that of the chromophore of a pyrrocorphin, hitherto known only as a synthetic tautomer of hexahydroporphyrin. At first sight, this event seemed to signal a further tautomerism of a dipyrrocorphin to a pyrrocorphin catalyzed by the enzyme, but when [$^{13}CH_3$]SAM was added to the incubation, it was found that a *third* methyl group signal appeared in the 19–21 ppm region of the NMR spectrum. When uro'gen III was provided with the ^{13}C labels (●) (as shown in Scheme VIII) three pairs of doublets appeared in the sp^3 region (δ 50–55 ppm) of the pyrrocorphin product. The necessary pulse labeling experiments together with appropriate FAB-MS data finally led to the structural proposal (**17**)[47,48] for the novel trimethylpyrrocorphin produced by "overmethylation" of the normal substrate, uro'gen III, in the presence of a high concentration of enzyme. Thus M-1 has been recruited to insert a ring-C methyl and synthesizes the long-sought "natural" chromophore corresponding to that of the postulated precorrin-**4a**, although in this case the regiospecificity is altered from ring D to ring C. This lack of specificity on the part of M-1 was further exploited to synthesize a range of "unnatural" isobacteriochlorins and pyrrocorphins based on isomers of uro'gen III. Thus, uro'gen I produces three methylated products corresponding to precorrin-1, precorrin-2 (**18**), and the type-I pyrrocorphin (**19**) (Scheme IX). These compounds are reminiscent of a series of tetramethyl type-I corphinoids, factors S_1–S_4 isolated from *P. shermanii*,[49,50] which occur as their zinc complexes. Uro'gens II and IV can also serve as substrates for M-1, remarkably producing isobacteriochlorins in both cases.[48b] When uro'gen I was incubated with SUMT,[51] isolation of factor II of the type-I (sirohydrochlorin I) family revealed a lack of specificity for this methyltransferase also, although in the latter studies no pyrrocorphins were observed. This may reflect a control mechanism in the *P. denitrificans* enzyme (SUMT) which does not "overmethylate" precorrin-2 as is found for the *E. coli* M-1 whose physiological function is to manufacture sirohydrochlorin for sulfite reductase necessary for cysteine synthesis. The fact that *E. coli* does not synthesize B_{12} could reflect an evolutionary process in which the C-methylation machinery has been retained, but is only required to insert the C-2 and C-7 methyl groups.

The sites of C-methylation in both the type-I and -III series are also reminiscent of the biomimetic C-methylation of the hexahydroporphyrins discovered by Eschenmoser,[52] and the regiospecificity is in accord with the principles adumbrated[53] for the stabilizing effect of a vinylogous ketimine system. In principle the methylases of the B_{12} pathway, which synthesize both natural and unnatural pyrrocorphins and corphins, can be harnessed to prepare several of the missing intermediates of the biosynthetic pathway, e.g., **20** → **21**. It is of note that their instability toward oxygen rationalizes our inability to isolate any new intermediates under aerobic conditions (>5 ppm O_2).

The possibility was also examined that a complete corrinoid structure based on a type-I template could be prepared by using uro'gen I, or reduced sirohydrochlorin I, as substrates with the cell-free system capable of converting uro'gen III to cobyrinic acid. However, no type-I cobyrinic acid was produced, suggesting that nature, although capable of inserting at least four methyl groups into uro'gen I by C-alkylation, is unable to effect the key step of ring contraction using the type-I pattern of acetate and propionate side chains. This result, although negative, is in accord with a suggestion[53] concerning the requirement for two adjacent acetate side chains in precorrinoids, one of which (at C-2) can participate in lactone formation to the C-20 meso position (see Scheme VI, precorrin-**6b**), a postulate supported by experiments[35] with ^{18}O-labeled precursors, while the second acetate function (at C-18) is used as an auxiliary to control the necessary activation of C-15 for C-methylation. The second methyltransferase which introduces the C-20 methyl into the substrate precorrin-2 (→precorrin-3) has been described recently.[36]

Evolutionary Aspects and Further Outlook

Genetic mapping of the loci of the B_{12}-synthesizing enzymes has been reported for *P. denitrificans*.[44] This complements a most interesting study on the genetics

(47) Warren, M. J.; Stolowich, N. J.; Santander, P. J.; Roessner, C. A.; Sowa, B. A.; Scott, A. I. *FEBS Lett.* **1990**, *261*, 76–80.
(48) (a) Scott, A. I.; Warren, M. J.; Roessner, C. A.; Stolowich, N. J.; Santander, P. J. *J. Chem. Soc., Chem. Commun.* **1990**, 593–595. (b) Warren, M. J.; Gonzalez, M. D.; Williams, H. J.; Stolowich, N. J.; Scott, A. I. *J. Am. Chem. Soc.* **1990**, *112*, 5343–5345.
(49) Müller, G.; Schmiedl, J.; Schneider, E.; Sedlmeier, R.; Worner, G.; Scott, A. I.; Williams, H. J.; Santander, P. J.; Stolowich, N. J.; Fagerness, P. E.; Mackenzie, N. E. *J. Am. Chem. Soc.* **1986**, *108*, 7875–7877.
(50) Müller, G.; Schmiedl, J.; Savidis, L.; Wirth, G.; Scott, A. I.; Santander, P. J.; Williams, H. J.; Stolowich, N. J.; Kriemler, H.-P. *J. Am. Chem. Soc.* **1987**, *109*, 6902–6904.
(51) Scott, A. I.; Williams, H. J.; Stolowich, N. J.; Karuso, P.; Gonzalez, M. D.; Blanche, F.; Thibaut, D.; Müller, G.; Wörner, G. *J. Chem. Soc., Chem. Commun.* **1989**, 522.
(52) Leumann, C.; Hilpert, K.; Schreiber, J.; Eschenmoser, A. *J. Chem. Soc., Chem. Commun.* **1983**, 1404–1407.
(53) Eschenmoser, A. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 5.

of *Salmonella typhimurium*, which cannot make B_{12} when grown aerobically.[54] A mutant requiring methionine, cobinamide, or cyanocobalamin when grown anaerobically produces B_{12} de novo thus leading to the isolation of other mutants blocked in B_{12} synthesis including one that cannot make factor II required for siroheme production. All of the cobalamin mutations lie close together on the chromosome, and a cluster of several methyltransferases maps at 42 min. Thus rapid progress can be expected in the isolation of the remaining biosynthetic enzymes and hence the long-sought intermediates.

Until quite recently it had been assumed that the Shemin pathway (glycine–succinate) to ALA was ubiquitous in bacterial production of porphyrins and corrins. However, it is now clear that in many archaebacteria (e.g., *Methanobacterium thermoautotrophicum*[55] and *Clostridium thermoaceticum*[56,57]) the C_5 pathway from glutamate is followed. Phylogenetically the C_5 route is conserved in higher plants, and it appears from recent work[58] that *hemA* of *E. coli* (and perhaps of *S. typhimurium*) encodes one of the enzymes for the glutamate → ALA conversion, i.e., the C_5 pathway is much more common than had been realized. In *C. thermoaceticum* it has been shown[57] that the B_{12} produced by this thermophilic archaebacterium is synthesized from ALA formed from glutamate.

Eschenmoser has speculated that corrinoids resembling B_{12} could have arisen by prebiotic polymerization of α-amino nitriles and has developed an impressive array of chemical models[53] to support this hypothesis, including ring contraction of porphyrinoids to acetylcorrins, deacetylation, and the C-methylation chemistry discussed earlier, which provide working hypotheses for the corresponding biochemical sequences. A primitive form of corrin stabilized by hydrogen,[53] rather than by methyl substitution, may indeed have existed $>4 \times 10^9$ years ago, before the origin of life[59] or the genetic code,[60] and could have formed the original "imprint" necessary for the evolution of enzymes which later mediated the insertion of methyl groups to provide a more robust form of B_{12}. Since B_{12} is found in primitive anaerobes and requires no oxidative process in its biogenesis (unlike the routes to heme and chlorophyll which are oxidative), an approximate dating of B_{12} synthesis would be $2.7–3.5 \times 10^9$ years, i.e., after DNA but before oxygen-requiring metabolism.[61]

If B_{12} indeed were the first natural substance requiring uro'gen III as a precursor, the question arises, "Why type III?" Since the chemical synthesis of the uro'gen mixture from PBG under acidic conditions leads to the statistical ratio of uro'gens (I, 12.5%; II, 25%; III, 50%; IV, 12.5%) containing a preponderance of uro'gen III, natural selection of the most abundant isomer could be the simple answer. It has been suggested[53] that the unique juxtaposition of *two adjoining* acetate side chains in the type-III isomer (which does not obtain in the symmetrical uro'gen I) may be responsible for a self-assembly mechanism requiring these functions to hold the molecular scaffolding in place via lactone and ketal formation as portrayed in Scheme VI. The logical conclusion of the concept is that the *oxidized* pigments (heme and chlorophyll) which came much later to the earth's surface in an aerobic atmosphere are "type III" *only because of the preregistration of the corrin structure* and the necessary evolution of the enzymes encoded by *hemA–D*.

With the genetic mapping of B_{12} biosynthesis now under way, it should at last be possible to discover the remaining intermediates beyond precorrin-**3**, together with the enzymes that mediate the methyl transfers, decarboxylation, ring contraction, deacetylation, and cobalt insertion. The powerful combination of molecular biology and NMR spectroscopy has been essential in solving the problems in B_{12} synthesis posed by the assembly and intermediacy of uro'gen III and the subsequent C-methylations leading to precorrin-**3**. It is anticipated that these techniques will again be vital to the solution of those enigmas still to be unraveled in the fascinating saga of B_{12} biosynthesis.

The work described in this Account has been carried out by an enthusiastic group of young colleagues whose names are mentioned in the references. Financial support over the last 20 years has been generously provided by the National Institutes of Health, the National Science Foundation, and the Robert A. Welch Foundation. It is a special pleasure to thank Professors Gerhard Müller (Stuttgart) and Peter Jordan (London) for their continued stimulating collaboration.

Registry No. Vitamin B_{12}, 68-19-9.

(54) Jeter, R. M.; Olivera, B. M.; Roth, J. R. *J. Bacteriol.* **1984**, *159*, 206–213.
(55) Zeikus, J. G. *Adv. Microb. Physiol.* **1983**, *24*, 215–299.
(56) Stern, J. R.; Bambers, G. *Biochemistry* **1966**, *5*, 1113–1118.
(57) Oh-hama, T.; Stolowich, N. J.; Scott, A. I. *FEBS Lett.* **1988**, *228*, 89–93.
(58) Jian-Ming, L.; Russell, C. S.; Coslow, S. D. *Gene* **1989**, *82*, 209–217.
(59) Decker, K.; Jungermann, K.; Thauer, R. K. *Angew. Chem., Int. Ed. Engl.* **1970**, *9*, 153–162.
(60) Eigen, M.; Lindemann, B. F.; Tietze, M.; Winkler-Oswatitsch, R.; Dress, A.; von Haeseler, A. *Science* **1989**, *244*, 673–679.
(61) Benner, S. A.; Ellington, A. D.; Taver, A. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 7054.