Genetically Engineered Synthesis of Natural Products

A. Ian Scott

J. Nat. Prod., **1994**, 57 (5), 557-573• DOI:
10.1021/np50107a001 • Publication Date (Web): 01 July 2004

Journal of Natural Products
Vol. 57, No. 5, pp. 557-573, May 1994

GENETICALLY ENGINEERED SYNTHESIS OF NATURAL PRODUCTS[1]

A. IAN SCOTT

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

ABSTRACT.—The feasibility of multi-step, one-flask total synthesis of natural products is demonstrated by cloning and overexpression of the corresponding biosynthetic enzymes from plant and microbial sources. As many as eight or nine enzymes can be recombined in vitro to reach an advanced intermediate in the vitamin B_{12} pathway.

Since the 1950s organic chemists have been exploring biosynthetic pathways to natural products, at first with radioactive tracers fed to whole plants and growing cells, then more "biochemically" with cell-free extracts, and finally with purified enzymes. By the beginning of the 1970s it became clear that techniques of nmr spectroscopy could be used to follow ^{13}C-enriched substrates through the maze of multienzyme conversions (frequently 15 or 20 steps) leading to a labeling pattern in the target from which biochemical processing could be deduced—albeit speculatively. Now another powerful technique is beginning to change the way in which many of us interested in Nature's methods of organic synthesis are approaching the problem. The new dimensions afforded by genetics and molecular biology have revolutionized our concepts of what is feasible in natural product biosynthesis. In Figure 1 the pathway from substrate A to the target

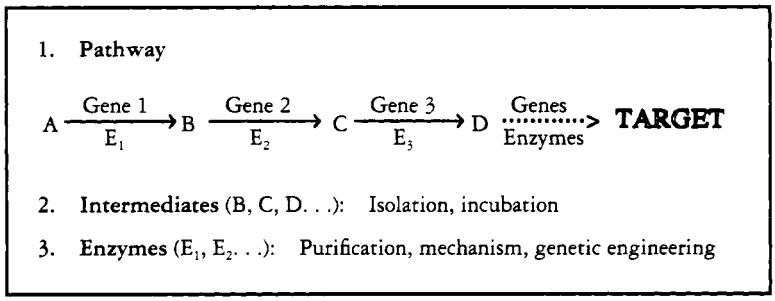


FIGURE 1. Natural Product Biosynthesis.

molecule via intermediates B, C, D, etc., now includes the gene for each biosynthetic enzyme. The appropriate gene products, i.e., the enzyme for each step, can be purified, sequenced, and the cDNA amplified in *E. coli* (or other vector) to produce amounts of enzyme (up to 1 gram per liter of cells), hitherto unimaginable. Thus, by employing the techniques shown in Figure 2 using either a cDNA library or more conveniently a series of open reading frames (ORFs) obtained by genetic complementation, mapping and sequencing, the gene products (enzymes) can be expressed "conventionally" or by PCR (Figure 3) and tested for their biosynthetic capabilities, using high-field spectroscopy to assay each biosynthetic step **as it occurs** in the nmr tube. Each enzyme can be studied separately in mechanistic detail, but when **recombined**, the full synthetic machinery can be activated, leading to multi-enzyme **total synthesis** of the target. Site-directed

[1]The 1993 Research Achievement Award address delivered at the 34th Annual Meeting of the American Society of Pharmacognosy, San Diego, California, 21 July 1993.

I. Isolate the gene encoding the enzyme.
   A. Construct a cDNA library.
   B. Make a probe (antibody or oligonucleotide) to identify the gene of interest from the cDNA library.
   C. Screen the cDNA library with the labeled probe to locate a plasmid or bacteriophage bearing the gene for the biosynthetic enzyme.
   D. Amplify and purify the plasmid or bacteriophage DNA bearing the gene and determine the nucleotide sequence of the gene.

II. Express and purify the enzyme.

FIGURE 2. Genetic Engineering of Biosynthetic Enzymes for Natural Products.

mutagenesis can then be used to modify substrate specificity thus producing new variations on the biosynthetic theme, as portrayed in Figure 4. This review illustrates the power of such an approach applied to several completely different natural product targets.

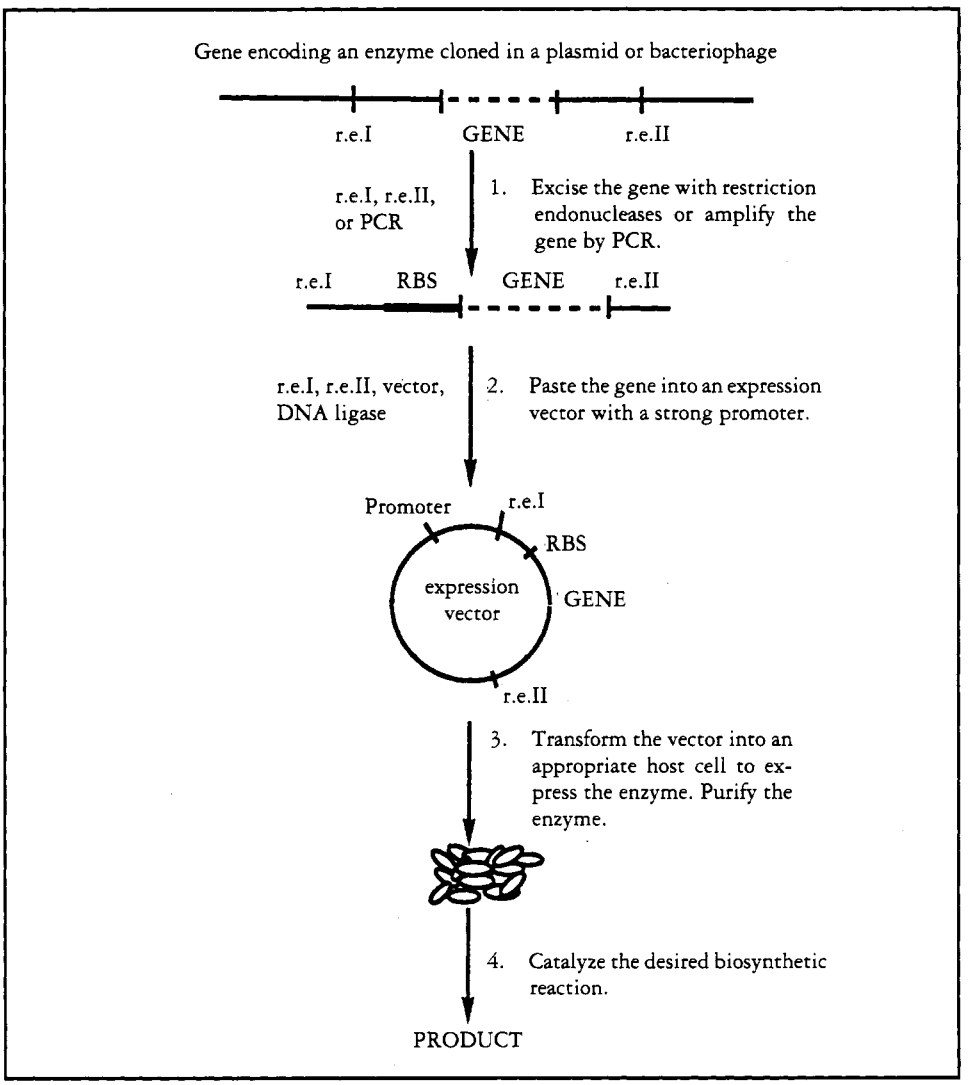


FIGURE 3. Expression of Biosynthetic Enzymes.

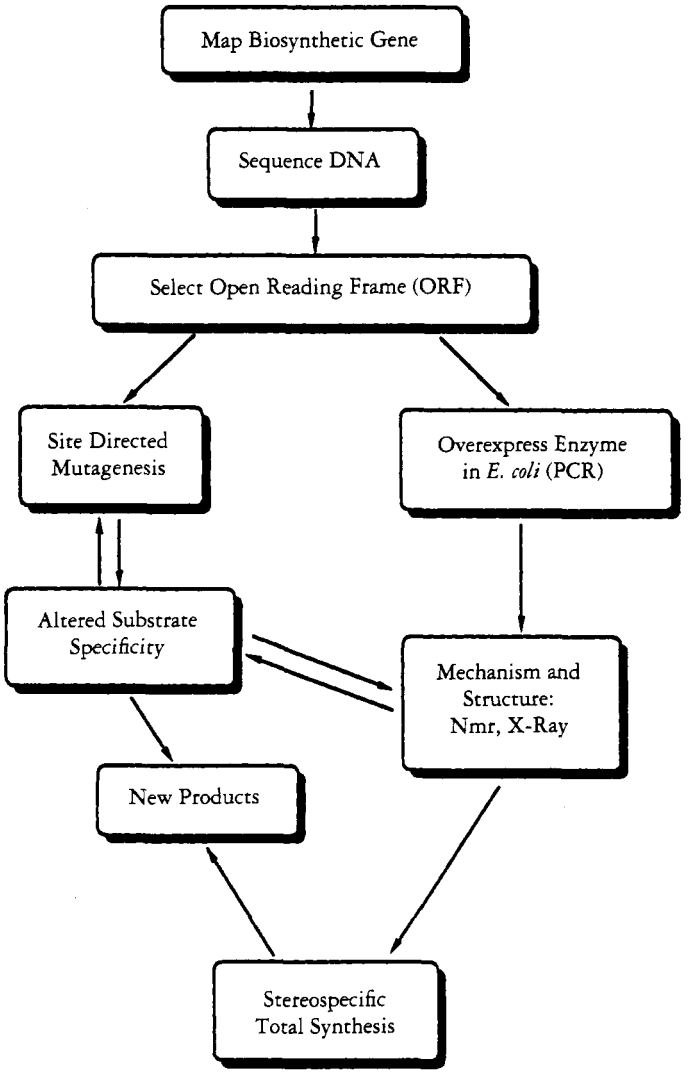


FIGURE 4. Genetic Engineering of Biosynthetic Enzymes.

MECHANISM OF ISOPENICILLIN N SYNTHASE (IPNS).—It is now firmly established that, following the enzymatic assembly of the Arnstein tripeptide, L-aminoadipyl-L-cysteinyl-D-valine [ACV; **1**] in a manner that retains the absolute configuration of the β-methyl groups (● and ▲) of the D-valine residue, cyclization of ACV [**1**] to isopenicillin N [**2**] leads directly to the penicillin nucleus (1,2) and thence to cephalosporin C [**6**] via [**3**], [**4**], and [**5**] (Figure 5).

Ever since the discovery by Demain (3) and Abraham (4) that cell-free preparations from *C. acremonium* catalyze the synthesis of the penicillin nucleus from ACV [**1**], intensive investigations have been undertaken to discover the mechanism(s) of the two bond-forming processes, viz., the removal of two hydrogens to fuse the β-lactam ring and a further 2H to generate the second, thiazolidine ring, [**1**]→[**8**]→[**2**] (Figure 6). Although recent studies with the purified enzyme IPNS (mol wt ca. 38,000) indicate that the cofactors Fe^{++}, O_2, and ascorbic acid are required in the cell-free system together with dithiothreitol (1,2,5), all attempts to characterize the putative lactam thiol [**8**]

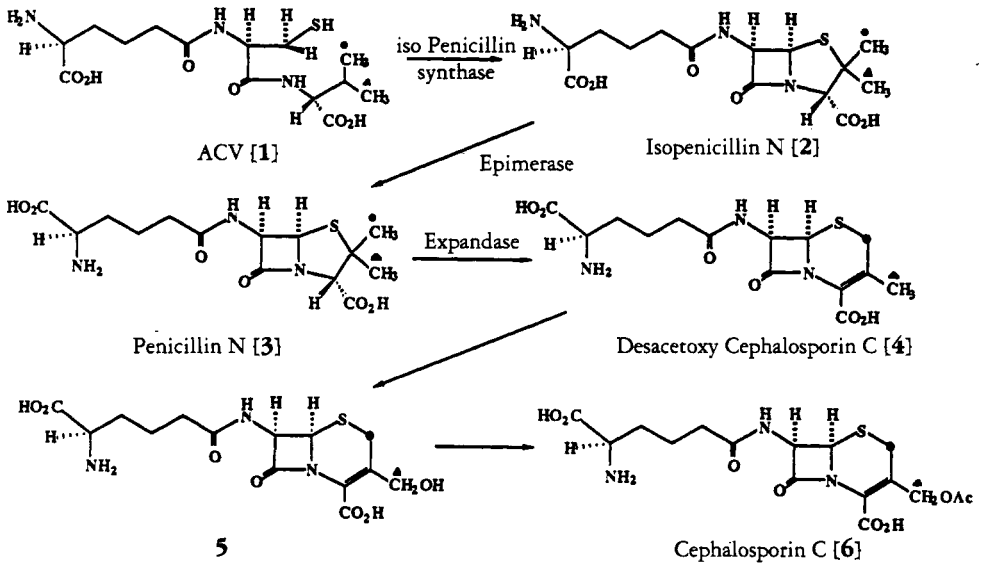


FIGURE 5. Biosynthesis of Cephalosporin C.

directly have so far failed (6–8), although kinetic isotopic effect studies strongly suggest that the reaction takes place via initial monocyclic β-lactam formation (9).

Synthetic preparation of **8** has been achieved in two independent studies (6–8) but the attempted conversion of **8** →**2** has not been realized due to the short half-life of **8** at pH 6-7, the working range of the cell-free system. Only by indirect kinetic isotope methods can it be deduced that the rate-determining step involves the proton (*) in ACV [**1**] rather than the β-valine proton (●) (1,2,9). In our laboratory the use of doubly labeled ACV, in which C-3 of Cys and N of Val were enriched with ^{13}C and ^{15}N, respectively, provided an excellent probe (10) whereby the discharge of **8** from the enzyme could have been recognized by ^{13}C-nmr spectroscopy. In the event, only the final molecule of isoPen N [**2**] exhibiting the anticipated ^{13}C-^{15}N coupling (J=4 Hz) was detectable under the experimental conditions.

Also tested was the ACV hydroxamate (11) (Figure 7, **9**) which proved to be an inhibitor of IPNS.

A similar experiment was conducted with the IPNS inhibitor ACG [**10**] synthe-

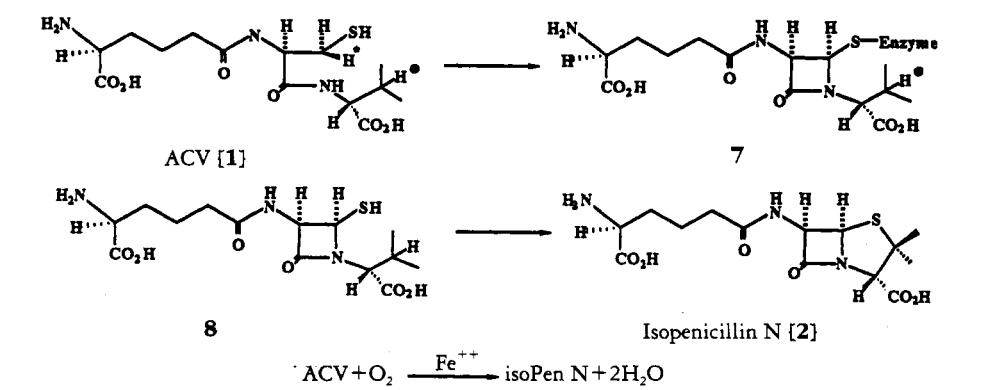
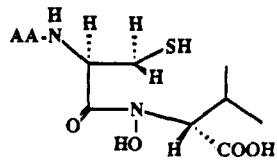


FIGURE 6. Proposed Intermediates in Pencillin Biosynthesis.



9

FIGURE 7. The Proposed Hydroxamate Intermediate.

sized with ^{13}C at the Cys-3 position. This specimen is easily prepared from protected AC, which is condensed with protected glycine followed by deprotection. The tripeptide [**10**] can only form one ring, namely, the β-lactam [**12**], so it was of considerable interest to observe the steady state spectrum of the enzyme-inhibitor complex at 1 mM concentration. Because 50-mg batches of pure IPNS are required for these nmr experiments at high concentrations, a new source of the enzyme was needed. In collaboration with Dr. Y. Aharonowitz, Tel Aviv, the IPNS gene from *Streptomyces jumonjinensis* cloned into a T7 polymerase system was expressed (12) and, after solubili-

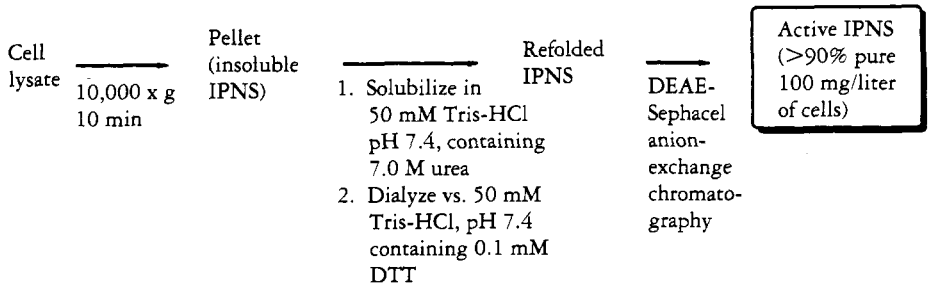


FIGURE 8. Purification of IPNS.

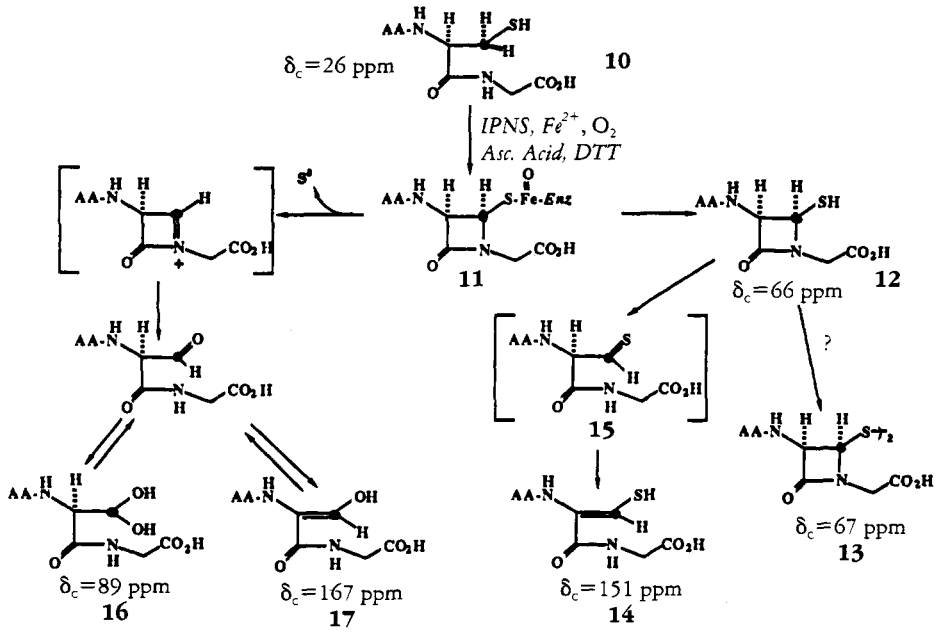


FIGURE 9. Intermediates and Products Observed on Incubation of ACG with IPNS.

zation of the resultant inclusion bodies, could be refolded to give the active enzyme as shown in Figure 8.

Incubation of ACG (2 mM) with IPNS (1 mM) at 5° in the nmr tube revealed (Figure 9) a persistent new signal at δ 66, coincident with the chemical shift of the synthetic β-lactam thiol [**12**] labeled with ^{13}C-nmr in the β-carbon of the cysteine moiety. On warming the solution to 25° the resonance shifted to δ 67, ascribed to the disulfide structure [**13**]. At the same time another new signal at δ 151 marked the formation of the enethiolate [**14**], labeled as shown (●) and formed by C-N bond cleavage to the labile thioaldehyde [**15**] and enolization. This experiment provides the first direct evidence for the monocyclic β-lactam structure [**12**] generated by IPNS. In addition, two other new resonances mark the operation of β-elimination of the cysteinyl sulfur to form the aldehyde hydrate [**16**; δ 89] and its **enolate** [**17**; δ 167], which had previously been discovered by Baldwin *et al.* (13) during studies on the mechanism of IPNS employing ACG and by isotopic substitution of ACV with deuterium (12) to suppress the second ring closure, which also led to the extrusion of sulfur.

By using "anti-freeze" solvents such as DMSO-H_2O (1:3) it has been possible to extend the catalytic turnover of IPNS with ACV from 20 min at 25° to 20 h at −15°. In this way, the observation of any enzyme-bound intermediates can be made at 1 mM concentration of enzyme-substrate complex. All of the above nmr observations are in accord with current views (1,2) on the active site of IPNS in which Fe^{II} is coordinated to two histidine residues (14) and participates in the mechanism by forming a complex with

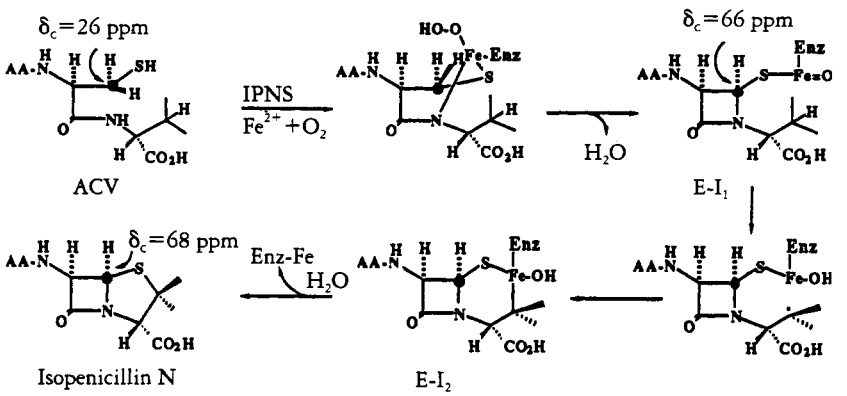


FIGURE 10a. Mechanistic Proposal for IPNS (J.E. Baldwin).

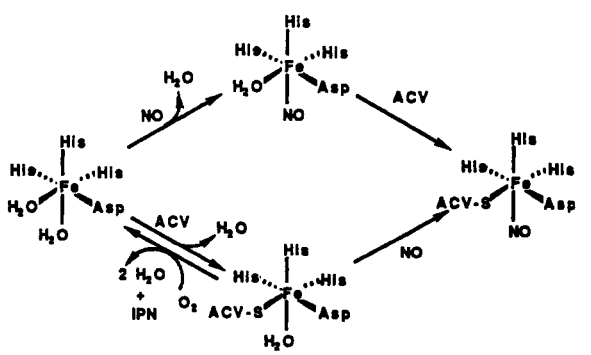


FIGURE 10b. Coordination of Iron at the Active Site of IPNS.

the cysteine thiolate of ACV (Figure 10a and 10b). This is followed by addition of O_2 and thence via E-I_1 and the iron-carbon bonded species (E-I_2) both rings are formed with the expulsion of 2 H_2O. Confirmation of these proposals for the mechanism for IPNS must await the 3-dimensional structure whose solution by X-ray diffraction is now in progress (with Prof. S.J. Remington, Oregon).

The understanding of the mechanisms of penicillin biosynthesis will be an important step in the design of novel mutants of IPNS with altered substrate specificity. The coupling of the complete set of enzymes for cephalosporin biosynthesis in vitro (Figure 5) is our eventual goal.

TRANSFER OF PLANT GENES AND EXPRESSION IN BACTERIA: SYNTHESIS OF INDOLE ALKALOIDS.—The second topic concerns our recent research on the biosynthesis of the indole alkaloids of *Catharanthus roseus*, a subject which has been actively pursued in our laboratories at Vancouver, Sussex, Yale, and Texas since 1962 (15). Again, the techniques of radio- and stable-isotope tracer feeding, cell-free systems and enzyme purification have been powerfully complemented by genetic engineering by expression of plant enzymes in heterologous systems (bacteria, yeast, insect cells). We chose as our first example the key enzyme strictosidine synthase (16) which catalyzes the Pictet-Spengler type condensation of the monoterpenoid glucoside secologanin and tryptamine (Figure 11). The product, 3α-(*S*)-strictosidine, is the progenitor of more than 1200 indole alkaloids as well as the important quinoline, camptothecin. Pioneering experiments by the Zenk group (17) have already shown that the cDNA library from *Rauwolfia serpentina* can be probed with oligonucleotide and a cDNA clone isolated and transferred to *E. coli*.

As an aid to our efforts to study the enzymatic mechanism of strictosidine synthase by nmr (which requires 20–40 mg of pure enzyme for each experiment), and to produce large quantities of strictosidine for use as substrate for the next enzyme in the pathway (cathenamine synthase), a glucosidase which forms cathenamine, we have cloned and

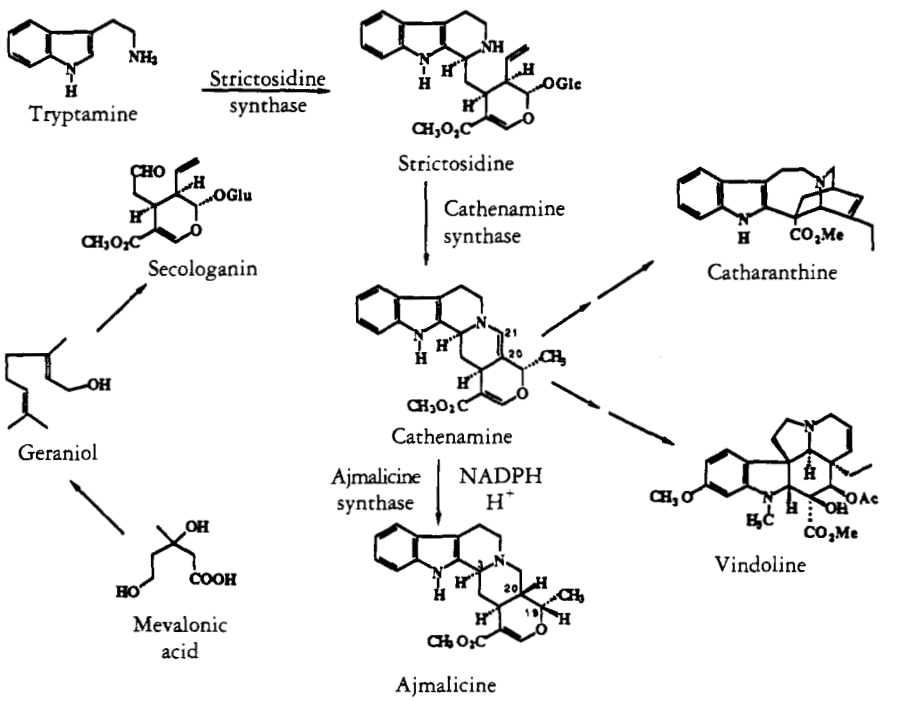


FIGURE 11. Biosynthesis of Some Indole Alkaloids.

determined the nucleotide sequence of the homologous gene from *C. roseus* (18). This gene has also been transformed into tobacco plants where it has been shown to express active strictosidine synthase (19). The overexpression of active *C. roseus* strictosidine synthase in *E. coli* was achieved (20) using the expression cassette polymerase chain reaction (ECPCR) technique (21) and led to the capability of purifying the enzyme in quantities exceeding those previously attainable. Strictosidine synthase has been shown by ultrastructural immunolocalization to be associated with plant vacuoles. To effect transport across the endoplasmic reticulum and into the vacuole, it is normally synthesized with a signal peptide which is, presumably, removed by signal peptidase during transport. Therefore, we designed our ECPCR primers so that the expressed protein would contain a methionine residue immediately followed by the amino acids of the **processed** enzyme, which were predicted from the DNA sequence (18).

When induced with IPTG, *E. coli* strain XA90(pRD1) displayed a new protein band with an M_r of about 34,000 on SDS-PAGE gels that was not seen in uninduced cells. A time-course of the expression demonstrated that collecting the cells two hours after induction provided good expression of the new protein. After being grown in the presence of IPTG for two hours, the whole cells displayed strictosidine synthase activity as demonstrated by the appearance of strictosidine by cells bearing the pRD1 but not by control cells bearing pHN1+ with no insert. A cell-free extract prepared by sonication of the cells followed by centrifugation at 12,000×g had about four times higher activity than whole cells. This enzymatic activity was in the soluble fraction even though most of the protein associated with the induced band was found in the insoluble fraction, presumably in the form of inclusion bodies. However, only a very small portion of the protein was soluble as there was not enough to be seen by SDS-PAGE and attempts to purify it required up to eight different steps and resulted in less than 1 mg of 50% pure enzyme from 8 liters of cells. The insoluble pellets were next washed with buffer and 3 M urea, solubilized in 8 M urea, dialyzed, and subjected to anion-exchange chromatography. The resulting protein was >90% pure as judged by SDS-PAGE and had a specific activity of 31 nkat/mg compared to 104 nkat/mg reported by other workers for enzyme isolated from cell cultures of *C. roseus* (17). The sequence of the first ten amino acids of the purified enzyme was found to be **Met Ser Pro Ile Leu Lys Lys Ile Phe Ile**, corresponding exactly to the sequence predicted from the nucleotide sequence. The M_r of the protein usually appeared to be about 34,000, slightly less than the 36,074 molecular weight predicted from the nucleotide sequence of the gene, suggesting that the protein had been partially hydrolyzed by proteolysis. Similar processing of the enzyme has also been reported elsewhere (17). Therefore, the lower specific activity of our recombinant enzyme (about 1/3 less than that of the enzyme isolated from *C. roseus* cell cultures) may be due to proteolysis, which would have to occur at the C-terminal since the N-terminal is unaltered. The availability of the recombinant enzyme is highly advantageous, for we now have the capability to purify as much enzyme activity from **1 liter** of a 2-h-induced culture of bacterial cells (about 2.5 g of cells) as was previously possible starting with **2 kg of *C. roseus*** cells, which require a complex growth medium and at least a week to grow. The purification procedure is also greatly simplified, requiring 2–3 days and only one chromatographic step.

The above technique is applicable to all of the enzymes necessary for the synthesis of the alkaloids of *C. roseus* and indeed is a general solution to targeted natural product synthesis of complex molecules. Desirable but "endangered plant" substances, e.g., taxol and camptothecin, can in principle be synthesized once the cDNA library and the repertoire of **sequenced** enzymes is available. We believe that this approach to plant natural products is not only viable, but offers a novel departure from, and a valuable

complement to, conventional synthetic chemistry, and may become the method of choice by the turn of this century.

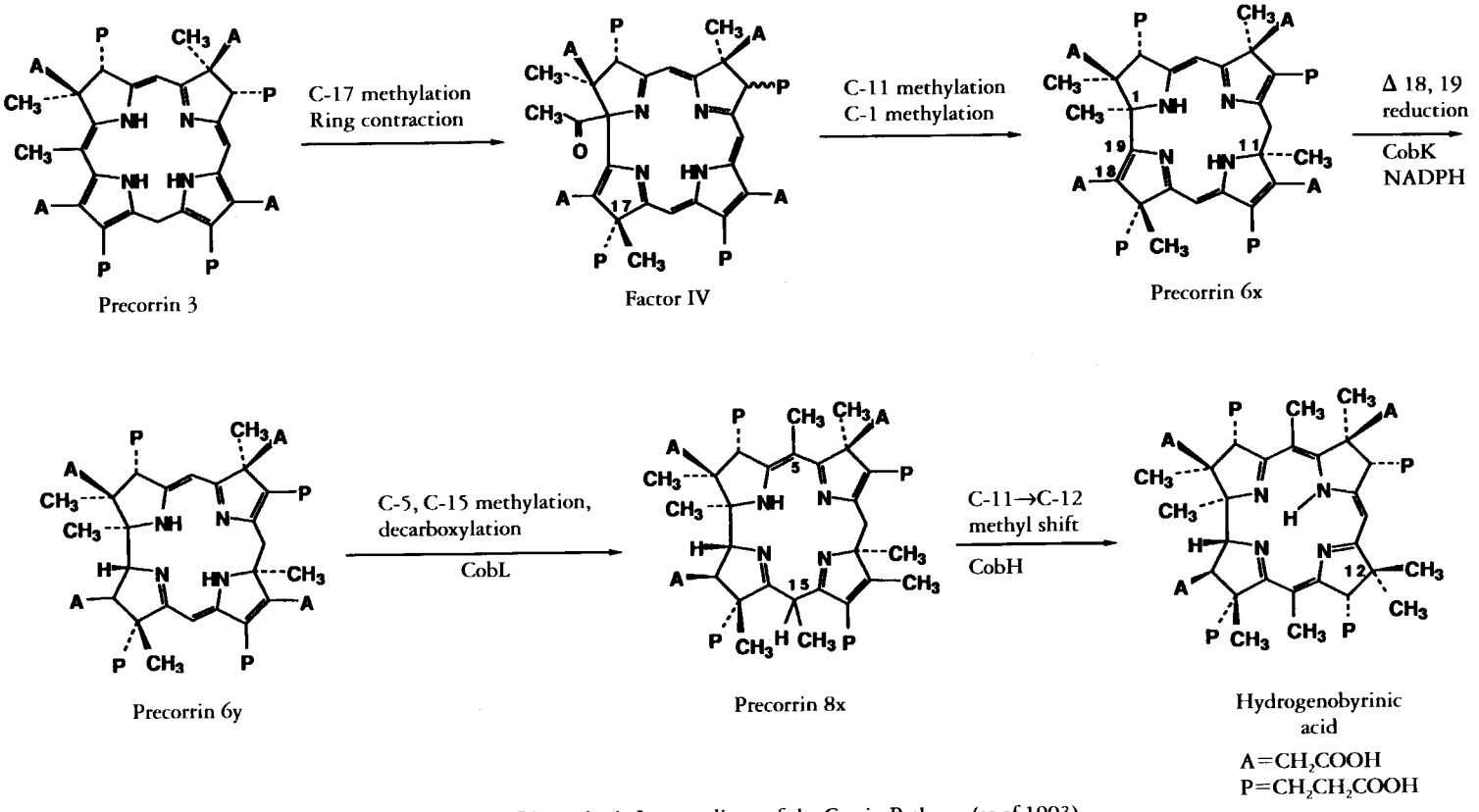
MULTI ENZYME SYNTHESIS OF NATURAL PRODUCTS EXEMPLIFIED.—During the course of our studies on vitamin B_{12} biosynthesis, a pathway involving over 20 enzymes, we have been fortunate not only in expressing the genes encoding almost all of the desired enzymes of the corrin operon, but, to our surprise, it was found that their combination in sets of as many as eight at a time provides a synthetic machine *par excellence* for one flask preparation of several advanced intermediates endowed both with complex functionality and a wealth of stereochemical detail. From the many experiments performed recently, the syntheses of precorrin-3x (22,23) and precorrin-5 (24) are illustrations of the power of genetically engineered synthesis.

The pathway to corrins (25–28) in the aerobic organism *Pseudomonas denitrificans* involves the conversion of precorrin-3 to hydrogenobyrinic acid via the isolated intermediates corresponding to sequential oxidative ring contraction and insertion of *C*-methyl groups from *S*-adenosyl methionine (SAM) at C-17β (→Factor IV) (29), two further methylations at C-11,C-1 (→precorrin 6x) (30,31), reduction at C-18, 19 (precorrin 6y) (32), methylation at C-5, C-15 with decarboxylation of the C-12 acetate side-chain (precorrin 8x) (33) and, finally (1,2,6) sigmatropic shift of the C-11 methyl to C-12 to afford the cobalt-free corrin (34) (Scheme 1).

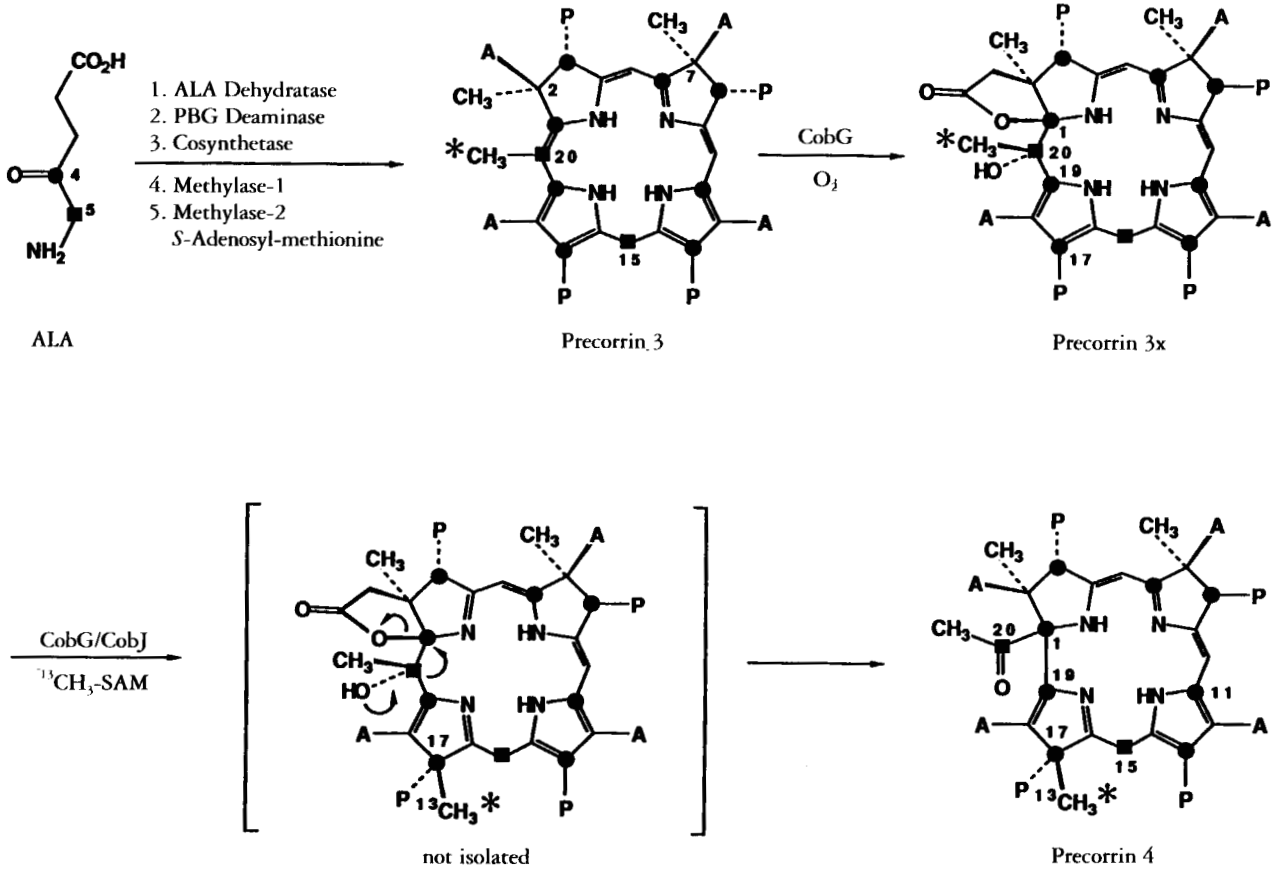
Because most of the above steps have been defined through the use of mutants rather than single enzymes, the fascinating problem of the mechanism of the ring contraction step has remained enigmatic, and is further complicated by the discovery of a parallel, *anaerobic* pathway in *Propionibacterium shermanii*, which features early introduction of cobalt into the dipyrrocorphin, precorrin-3 (35).

Using ^{13}C-nmr spectroscopy as a probe for the activity of each of the overexpressed enzymes of the B_{12} pathway in *P. denitrificans*, we have tested the individual enzymes in turn with precorrin-3 as substrate and find that CobG (36,37) serves as an O_2-dependent enzyme whose role is to install oxygen-derived functionality at C-20, thus preparing the macrocycle for ring contraction. Remarkably, the resultant spring-loaded mechanism **does not operate until after the fourth *C*-methylation has occurred at C-17**, an event that is mediated by CobJ (37) a SAM-requiring enzyme, thereby defining both the ring contraction and C-17 methylation sequences.

The nmr assays for the activities of CobG and J were developed as follows. First, precorrin-3 was prepared in two ^{13}C-isotopomeric versions, A, from [^{13}C-4]-(●)-, and B, from [^{13}C-5]-(■)-5-aminolevulinic acid, using the multi-enzyme synthesis described earlier (38). The reaction of isotopomer A (●) with CobG in presence of O_2 and NADH resulted in a spectrum almost identical with that of the substrate except for the disappearance of the signal (●) for C-1 at δ 146 ppm and the appearance of a peak (●) at δ 106 ppm corresponding to sp^3 geometry and a new environment at C-1. When isotopomer B (■) served as substrate, the pattern of nmr signals remained constant except for the resonance for C-20 (at δ 103 ppm), which was replaced by a signal at δ 78 ppm corresponding to oxygen insertion and resultant sp^3 hybridization at this center. In a separate experiment using precorrin-3 enriched at the C-20 methyl, the CH_3 signal (*) at C-20 (δ 17.6 ppm) in precorrin-3 underwent a downfield shift to δ 25 during incubation with CobG. The above spectral changes are in accord with the addition of oxygen at both positions 1 and 20 and together with the observation by ir spectroscopy of a γ-lactone (1799 cm^{-1}) lead to the structural proposal (Scheme 2) for the new product, precorrin-3x, whose formation can be rationalized by epoxidation at C-1, 20 by the O_2-dependent CobG enzyme followed by participation of the ring A carboxylate in a



SCHEME 1. Biosynthetic Intermediates of the Corrin Pathway (as of 1993).



ALA | Precorrin 3 | Precorrin 3x

not isolated | Precorrin 4

SCHEME 2. Recently Discovered Intermediates and Mechanism of Ring Contraction.

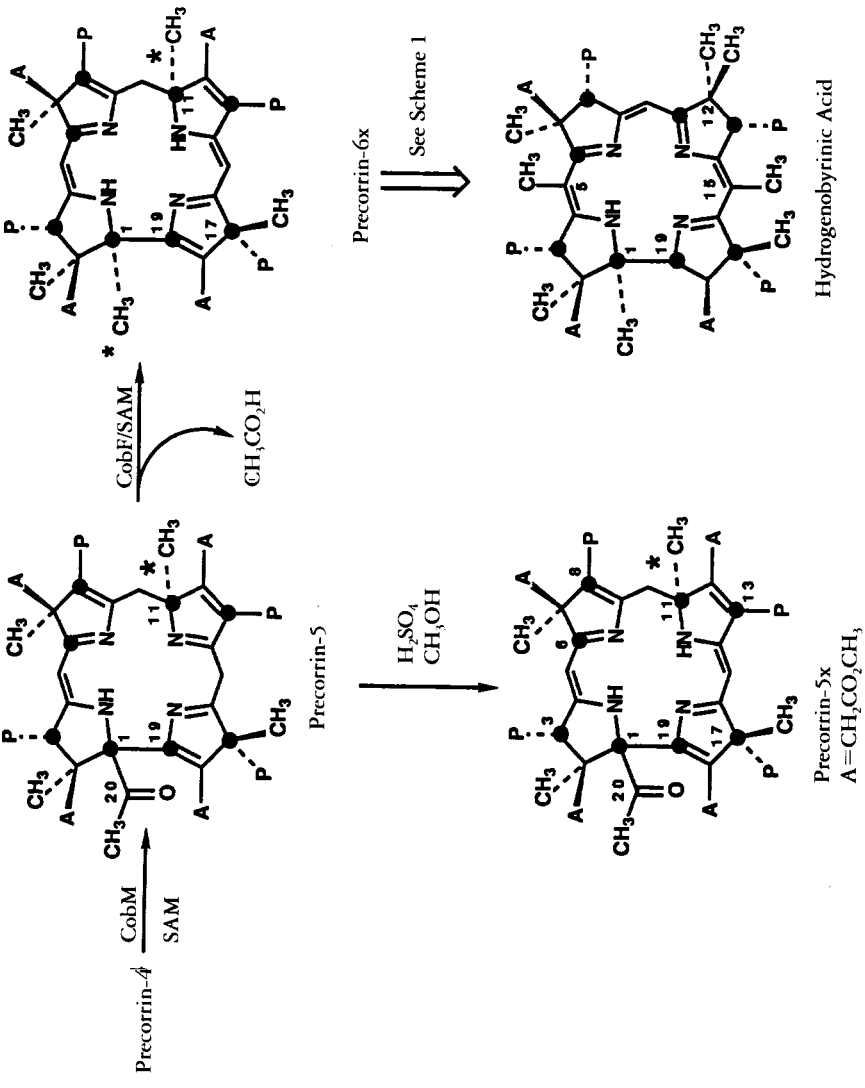
lactonization-ring opening sequence. In the absence of O_2 no reaction of precorrin-3 with CobG was observed.

Next, addition of SAM and, in turn, each of the remaining putative methyl transferases in the *P. denitrificans* repertoire (CobF, J, M) to precorrin-3x labeled from [4-^{13}C]ALA (●) resulted in a new signal for C-17 (δ 66) corresponding to *C*-methylation at this center, **only when CobJ was present**. Confirmation that C-17 is indeed the site of the fourth methyl insertion came from double-labeled incubation of isotopomer A (●) and $^{13}CH_3$-SAM (*) in the presence of **both** CobG and J. In this experiment the new CH_3 signal (*) appearing at δ 22.5 was coupled (J=37 Hz) to the C-17 resonance (●; δ 66 ppm) (Scheme 2). Most significantly, the ^{13}C-nmr spectrum of the new intermediate also displays a new pair of coupled carbons, C-1, C-19 (δ 82, 142, J=52 Hz), showing that ring contraction occurs during incubation with CobJ. When the two-enzyme incubation was repeated using isotopomer B (■) the signal for C-20 in the product appeared at δ 210 ppm indicating that ring contraction is accompanied by the genesis of a new methyl ketone function pendant from C-1, by a process which corresponds formally to the pinacol type rearrangement (39) illustrated in Scheme 2.[2] Analysis of the full set of spectral data (nmr; fabms; Ft-ir) leaves no doubt that the new isolate is precorrin-4. Aerial oxidation to factor IV (29) completed the proof of structure and absolute stereochemistry (except at position C-1) for precorrin-4, the long-sought tetramethylated intermediate of corrin biosynthesis.

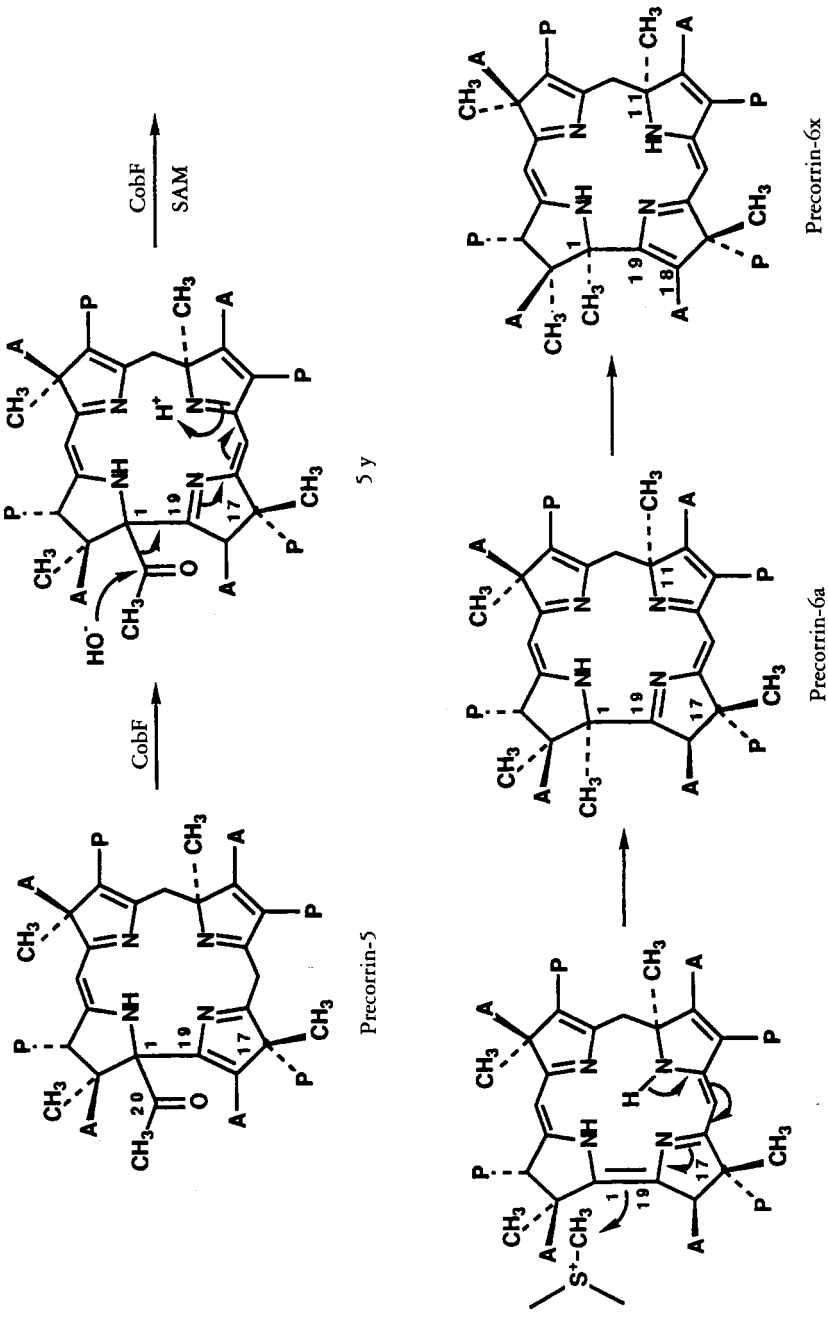
Prior to the discovery of precorrin-4 we had predicted (40) that CobM catalyzes insertion of a C-methyl group at C-11 based on studies with model substrates. Using the true substrate, precorrin-4, biosynthetically labeled from [$^{13}C_4$-]-5-aminolevulinic acid (ALA) as shown (●) (Scheme 3), the ^{13}C-nmr spectrum of the octacarboxylic acid isolated from the incubation mixture containing precorrin-4 CobM, and $^{13}CH_3$*-SAM revealed, in addition to C_{1-19} coupling, a second coupled pair corresponding to a new CH_3* group at δ 23.7 (d, J=37 Hz) and an sp^3 signal for C-11 (δ 76, d, J=37 Hz). Proof that the C-1 acetyl function was still intact came from (a) the observation of a resonance for C-1 at δ 82 (d, J=52 Hz) and (b) the retention of signals (δ 211 ppm) for the carbonyl at C-20 and the sp^3 center at C-15 when a second isotopomer of precorrin-4, labeled with ^{13}C at C-15 and C-20 from [^{13}C-5]-ALA, was used as substrate. The new isolate is thus precorrin-5. When the gene product CobF (40) was included in the incubation, the signal for C-1 (δ 82) in the spectrum of precorrin-5 disappeared and was replaced by a resonance at δ 73 (dd) coupled to the signal for C-19 at δ 154 and to a new *C*-methyl group (*) at δ 29.5 (J=36 Hz). Esterification of the free acid provided a sample with nmr and fabms identical with those of precorrin-6x [octamethyl ester] (31). This biosynthetic correlation confirms the structure and defines the absolute stereochemistry (except at C-1) of precorrin-5 as shown in Scheme 3.

Thus it appears that deacylation of precorrin-5 is catalyzed by the CobF enzyme to provide the short-lived intermediate precorrin-**5y** which then undergoes C-1 methylation. A mechanistic rationale shown in Scheme 4 features the bis-imino chromophore in 5y extended from C-19 as an electron sink created by reaction with CobF. This structural motif in which C-11 methylation insulates rings C and D electronically from rings A/B not only facilitates deacetylation, but allows the return of electron density to C-1 followed by electrophilic C-methylation mediated by CobF/SAM. The process is completed by prototropic shift from C_{18} in the kinetic product precorrin-6a (not isolated)

[2]For simplicity, only two of the eight enriched carbons derived from [^{13}C-5]-ALA are shown in Scheme 2 (■) since there is virtually no change in chemical shift of the remaining labeled carbons from this isotopomer.



SCHEME 3. Precorrin-5, a New Intermediate on the B_{12} Pathway.



SCHEME 4. Mechanisms of Deacylation and C-1 Methylation.

to give precorrin-6x. That the facile removal of acetic acid from C-1 is also a **chemical** process was demonstrated with precorrin-5x, which loses acetic acid spontaneously (albeit slowly) at pH <7.

When the entire repertoire of the eight enzymes[3] (35) necessary for the nine-step synthesis of precorrin-5 was incubated with the substrates ALA and SAM, the genetically engineered synthesis of precorrin-5 was achieved in 30% overall yield based on ALA[4] (36). This striking example of the in vitro reconstitution of a complex biosynthetic pathway shown in Figure 12 serves to illustrate both the power and generality of multi-enzyme synthesis (41,42) featuring many different types of C-C, C-N, and C-O bond formation and rearrangements. Most importantly, the removal of the rigid, metabolic

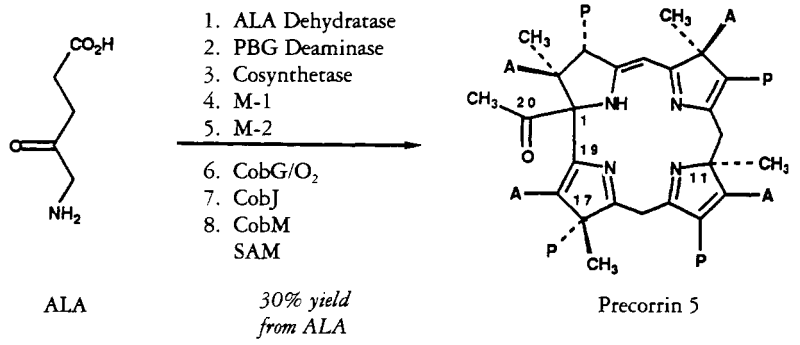


FIGURE 12. Multi-enzyme Synthesis of Precorrin-5.

networking of the living cell, by using only those biosynthetic enzymes necessary to synthesize the target, ensures that each intermediate is dedicated to a predetermined pathway rather than being diverted into the main arteries of primary metabolism in vivo. We would like to suggest that with the demonstration of this nine-step genetically engineered process, the multi-enzyme synthesis of complex natural products has come of age (43,45).

The structures of precorrins-3x, -4, and now precorrin-5, complete our knowledge of the steps between precorrins-3 and -6x and consequently of all of the intermediates (and enzymes that interconnect them) between ALA and hydrogenobyrinic acid in the aerobe *P. denitrificans*. It remains for us now to reconstitute the entire pathway in vitro and then to uncover the details of the parallel, but non-identical, synthetic route in anaerobic bacteria, e.g., *Salmonella typhimurium* (46) and *Propionibacterium shermanii* (35) where, in the latter organism, the intermediates are known to be cobalt complexes beyond the stage of precorrin-3 and where the redox chemistry necessary for ring contraction must be independent of O_2 (27).

In summary, what was once a pipe dream, namely the observation of natural products and their intermediates being formed in a test tube by the sequential action of as many as 5 or 10 enzymes is now a reality, thanks to the powerful combination of

[3]In biosynthetic order these are ALA dehydratase, PBG deaminase, uro'gen III synthase, uro'gen III methylase (CobA; 2 steps), precorrin-3 synthase (CobI and CbiL), CobG, CobJ, CobM.

[4]The overall yield of 30% based on isolation of the pure ester, 5x, does not include optimization of the last two steps catalyzed by CobJ and M. Earlier work had shown that, using the first five enzymes, precorrin-3 is synthesized from ALA in 79% overall yield (ca. 95% conversion for each step) while the conversion of precorrin 3→3x (CobG) is ca. 80%. The CobJ, M sequence is therefore estimated to be operating at ca. 70% efficiency.

organic chemistry, nmr spectroscopy, and molecular biology. In other words we have changed our role as spectators of natural product biosynthesis to that of imposing control on the genetic machinery to execute a programmed synthesis of the target molecule. These are indeed exciting times for bioorganic chemistry and for those of us who are privileged to be able to stand on the shoulders of the pioneers who gave us superb tools, such as high-resolution nmr and cloning techniques, without which none of the experiments described above could even have been contemplated, let alone designed and executed.

ACKNOWLEDGMENTS

The work described in this lecture could not have been possible without the perseverance and skill of my younger colleagues whose names will be found in the references, and generous financial support from the N.I.H. and the Robert A. Welch Foundation. A special word of thanks is also due to Sandra Estrada (microbiology), Marc Holderman (enzymology), and Yanding Gao (nmr electronics) for excellent technical assistance.

LITERATURE CITED

1. J.E. Baldwin, in: "Recent Advances in the Chemistry of β-Lactam Antibiotics." Ed. by A.G. Brown and S.M. Roberts. Royal Society of Chemistry Special Publication No. 52, London, 1985, pp. 62–85, and references therein.
2. J.E. Baldwin and E. Abraham, *Nat. Prod. Rep.*, **5**, 129 (1988), and references therein.
3. T. Koroni, S. Herchen, J.E. Baldwin, M. Yoshida, N.A. Hunt, and A.L. Demain, *Biochem. J.*, **184**, 427 (1979).
4. J. O'Sullivan, R.C. Bleaney, J.A. Huddleston, and E.P. Abraham, *Biochem. J.*, **184**, 421 (1979).
5. R.L. White, E.M. John, J.E. Baldwin, and E.P. Abraham, *Biochem. J.*, **203**, 791 (1982).
6. E.P. Abraham, R.M. Adlington, J.E. Baldwin, M.J. Crimmin, L.D. Field, G.S. Jayatilake, R.L. White, and J.J. Usher, *Tetrahedron*, **40**, 1907 (1984).
7. S.K. Chung, R. Shankaranarayan, and A.I. Scott, *Tetrahedron Lett.*, 2941 (1983).
8. A.I. Scott, R. Shankaranarayan, and S.K. Chung, *Heterocycles*, **30**, 909 (1990).
9. J.E. Baldwin, R.M. Adlington, J.E. Moroney, L.B. Field, and H.-H. Ting, *J. Chem. Soc., Chem. Commun.*, 984 (1984).
10. R.L. Baxter, C.J. McGregor, G.A. Thomson, and A.I. Scott, *J. Chem. Soc., Perkin Trans. I*, 369 (1985).
11. R.L. Baxter, G.A. Thomson, and A.I. Scott, *J. Chem. Soc., Chem. Commun.*, 32 (1984).
12. O. Landman, D. Shiffman, Y. Av-Gay, Y. Aharonowitz, and G. Cohen, *FEMS Microbiol. Lett.*, **84**, 239 (1991).
13. J.E. Baldwin, M. Bradley, R.M. Adlington, W.J. Norris, and N.J. Turner, *Tetrahedron*, **47**, 457 (1991).
14. L.-J. Ming, L. Que, A. Kriauciunas, C.A. Frolik, and V.J. Chen, *Inorg. Chem.*, **29**, 1111 (1990).
15. A.I. Scott, *Acc. Chem. Res.*, **3**, 151 (1970), and references therein.
16. H. Mizukami, H. Nordlov, S.L. Lee, and A.I. Scott, *Biochemistry*, **18**, 3760 (1979).
17. T.M. Kutchan, H. Dittrich, D. Bracher, and M.H. Zenk, *Tetrahedron*, **31**, 5945 (1991), and references therein.
18. T.D. McKnight, C.A. Roessner, R. Devagupta, A.I. Scott, and C.L. Nessler, *Nucl. Acids Res.*, **18**, 4939 (1990).
19. T.D. McKnight, D.R. Bergey, R.J. Burnett, and C.L. Nessler, *Planta*, **185**, 148 (1991).
20. C.A. Roessner, R. Devagupta, M. Hasan, H.J. Williams, and A.I. Scott, *Protein Express. Purific.*, **3**, 295 (1992).
21. K.D. MacFerrin, M. Terrauera, S.L. Schreiber, and G. Verdine, *Proc. Natl. Acad. Sci. USA*, **87**, 1937 (1990).
22. A.I. Scott, C.A. Roessner, N.J. Stolowich, J.B. Spencer, C. Min and S.-I. Ozaki, *FEBS Lett.*, **331**, 105 (1993).
23. J.B. Spencer, N.J. Stolowich, C.A. Roessner, C. Min, and A.I. Scott, *J. Am. Chem. Soc.*, **115**, 11610 (1993).
24. C. Min, B.P. Atshaves, C.A. Roessner, N.J. Stolowich, J.B. Spencer, and A.I. Scott, *J. Am. Chem. Soc.*, **115**, 10380 (1993).
25. A.I. Scott, *Acc. Chem. Res.*, **23**, 308 (1990).
26. A.R. Battersby, *Acc. Chem. Res.*, **26**, 15 (1993).
27. A.I. Scott, *Angew. Chem., Int. Ed. Engl.*, **32**, 1223 (1993).

28. F.J. Leeper, *Nat. Prod. Rep.*, **6**, 171 (1989).
29. D. Thibaut, L. Debussche, D. Frechet, F. Herman, M. Vuilhorgne, and F. Blanche, *J. Chem. Soc., Chem. Commun.*, 513 (1993).
30. D. Thibaut, L. Debussche, and F. Blanche, *Proc. Natl. Acad. Sci. USA*, **87**, 8795 (1990).
31. D. Thibaut, F. Blanche, L. Debussche, F.J. Leeper, and A.R. Battersby, *Proc. Natl. Acad. Sci. USA*, **87**, 8800 (1990).
32. D. Thibaut, F. Kiuchi, L. Debussche, F.J. Leeper, F. Blanche, and A.R. Battersby, *J. Chem. Soc., Chem. Commun.*, 139 (1992).
33. D. Thibaut, F. Kiuchi, L. Debussche, F. Blanche, M. Kodera, F.J. Leeper, and A.R. Battersby *J. Chem. Soc., Chem. Commun.*, 982 (1992).
34. D. Thibaut, M. Couder, A. Famechon, L. Debussche, B. Cameron, J. Crouzet, and F. Blanche, *J. Bacteriol.*, **174**, 1043 (1992).
35. 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. Thibaug, *J. Am. Chem. Soc.*, **113**, 9891 (1991).
36. J. Crouzet, S. Levy-Schill, B. Cameron, L. Cauchois, S. Rigault, M.C. Rouyez, F. Blanche, L. Debussche, and D. Thibaut, *J. Bacteriol.*, **173**, 6074 (1991).
37. L. Debussche, D. Thibaut, M. Danzer, F. Debu, D. Fréchet, F. Herman, F. Blanche, and M. Vuilhorgne, *J. Chem. Soc., Chem. Commun.*, 1100 (1993).
38. M.J. Warren, C.A. Roessner, S.-I. Ozaki, N.J. Stolowich, P.J. Santander, and A.I. Scott, *Biochemistry*, **31**, 603 (1992).
39. U.B. Grüning, G. Holze, A. Gossauer, and L. Ernst, *Helv. Chim. Acta*, **68**, 1771 (1985).
40. 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.*, **301**, 73 (1992).
41. A.I. Scott, *Tetrahedron*, **48**, 2559 (1992), and references therein.
42. A.I. Scott, *Pure Appl. Chem.*, **65**, 1299 (1993), and references therein.
43. A.I. Scott, *Science*, **257**, 1481 (1992).
44. A.I. Scott, *Science News*, **142**, 148 (1992).
45. A.I. Scott, *Chem. Br.*, **29**, 687 (1993).
46. J.R. Roth, J.G. Lawrence, M. Rubenfield, S. Kieffer-Higgins, and G.M. Church, *J. Bacteriol.*, **175**, 3303 (1993).

Received 17 February 1994