Biological formation of pyrroles: Nature’s logic and enzymatic machinery

Christopher T. Walsh,* Sylvie Garneau-Tsodikova† and Annaleise R. Howard-Jones

Received (in Cambridge, UK) 2 May 2006
First published as an Advance Article on the web 15th June 2006
DOI: 10.1039/b605245m

This Highlight provides an overview of the molecular logic employed in the formation and elaboration of pyrroles during natural product biosyntheses. The pyrrole ring, either in its unadorned form or as the pyrole-2-carboxylate moiety commonly encountered in natural products, possesses many useful chemical and electronic features that are readily exploited in biological contexts. The ring itself is planar and electron rich, highly susceptible to electrophilic attack, capable of oxidation and able to participate in both π–π stacking and hydrogen-bonding interactions; the 2-carboxylate moiety provides a nucleophilic center for derivatization, hydrogen bonding and reaction with biological targets. The focus of this Highlight is primarily on the construction of pyrrole rings during secondary metabolite formation, and on their derivatization and incorporation into natural products. Overall biosynthetic strategies are discussed as well as relevant enzymology and mechanistic details.

1 Introduction

The five-membered heteroaromatic pyrrole ring, planar and electron rich, is a useful recognition element in many biological contexts—forming hydrogen bonds, coordinating metals, and...
providing stacking interactions. It is a key constituent of the bicyclic indole side chain of L-tryptophan residues in proteins and is present in a wide range of natural products (Fig. 1). Perhaps the most celebrated biological context of pyrroles is in the tetrapyrrole scaffold of heme and related porphinoid cofactors (e.g., heme b 1, vitamin B12 2, chlorophyll α 3 and factor 430 4, Fig. 1A) where iron, cobalt, magnesium or nickel atoms are chelated in the equatorial plane of the macrocycles. The linear tetrapyrrole, hydroxymethylbilane, is the common precursor to these macrocyclic pyrrole-bearing scaffolds.

There are also many natural examples where the pyrrole ring is found excised from this tetrapyrrole array and divorced from the benzenoid partner it possesses in the indole scaffold. Examples include the coumarin family of natural product antibiotics (e.g., 8

Fig. 1 Examples of pyrrole-containing natural products: (A) the tetrapyrroles; and (B) pyrrole-containing small molecules.
and 9), the pyrrolomycins (e.g., 10), pyoluteorin 11, pyrrolnitrin 12 and the prodigines (e.g., 6 and 7) (Fig. 1B). These are secondary metabolites isolated from biocontrol bacterial strains, which modulate antagonizing fungi colonizing on plant surfaces. A variety of pyrrole-containing natural products, including the pyrrolomycins (e.g., 10), pyoluteorin 11 and pyrrolnitrin 12, are halogenated during biosynthesis. Of note also are the natural products of the prodigine family (e.g., prodigiosin 6 and undecylprodigiosin 7) where three pyrrole rings are present, two directly coupled in tandem array (Fig. 1B). The indolocarbazole scaffolds of the natural products staurosporine 13 and rebeccamycin 14 (Fig. 1B) are biosynthesized from two molecules of L-tryptophan 5 by an oxidative dimerization pathway, creating the maleimide or pyrrolinone rings by way of a dicarboxypyrrole intermediate. We focus on these particular examples due to the availability of molecular information on the enzymes involved in pyrrole ring construction. This discussion may therefore serve as a blueprint for Nature’s broader strategies in pyrrolic natural product biosynthesis.

2 Biogenesis of the pyrrole ring in L-tryptophan and in tetrapyrrole cofactors

The indole ring in the proteinogenic amino acid L-tryptophan 5 is fashioned from N-(5′-phosphoribosyl)anthranilate 15 after enzymatic isomerization to the ring opened ketone 16 and cyclization to an iminium ion 17 (Fig. 2). Decarboxylation and subsequent dehydration of this iminium moiety affords the first cyclization to an iminium ion 17. Subsequent dehydration of this iminium moiety affords the first cyclization to an iminium ion 17. The hydrophobic indole side chain generally packs into the interior of proteins, forming π–π stacking interactions with other side chains or cofactors. The pyrrole moiety can also engage in hydrogen bonding and, in rare instances, may be susceptible to one electron redox transfers. In the indole nucleus, the presence of a benzene ring fused to the pyrrole modulates the electron density on the pyrrole ring and directs reaction with electrophiles to the 3-position (rather than the 2-position). The indole ring in L-tryptophan 5 acts as a key precursor to the large class of structurally complex indole alkaloid natural products, such as the indolocarbazoles and pyrroloquinolines discussed below. Such alkaloids arise via secondary transformations of L-tryptophan, thus diverting part of the metabolic flux of this amino acid away from protein synthesis and towards the molecule of interest. The structural diversity within the indole alkaloid family is astounding, as is the wide variety of producer organisms, with many plants, bacteria and other marine organisms acting as rich sources of these natural products.11,21

The best-known pathway for de novo biological construction of the pyrrole ring is that seen in the biosynthesis of the tetrapyrrolic macrocycles, including the hemes and vitamin B12 (Fig. 3). The first dedicated intermediate in this pathway, δ-aminolevulinic acid (ALA) 19, is biosynthesized from glycine and succinyl CoA by a PLP-dependent enzyme, ALA synthase.23,24 The enzyme ALA dehydratase (porphobilinogen synthase) (Fig. 4) then catalyzes condensation of two δ-aminolevulinic molecules by a Knoore-type cyclization reaction, creating the trialkyl-substituted pyrrole known as porphobilinogen (PBG) 20. Such alkyl-substituted pyroles are electron rich, prone to auto-oxidation and likely kept at low concentrations in the cell. Indeed, the monomeric PBG is swept through to hydroxymethylbilane 21 (otherwise known as preuroporphyrinogen) (Fig. 5) by the remarkable enzyme, porphobilinogen deaminase (hydroxymethylbilane synthase).25 This protein catalyst has a covalently tethered dipyrromethane cofactor attached to an active site residue and generates a hexapyrrole covalent enzyme intermediate from the resting form of the enzyme.26–28 This linear hexapyrrole is assembled by iterative nucleophilic attack of the growing protein-tethered poly-pyrrole chain on four porphobilinogen-derived moieties by an S_n1-type mechanism.28 At this stage, regioselective protonolytic cleavage is effected, leading to formation of the linear tetrapyrrolic hydroxymethylbilane product, with regeneration of the starting dipyrromethanyl cofac-

![Fig. 2 Biosynthesis of L-tryptophan 5 from N-5′-phosphoribosyl]anthranilate 15.14](image-url)
tor in the active site. The subsequent steps of enzymatic cyclization and maturation to the tetrapyrrolic macrocyclic scaffolds of the hemes, corrins, chlorophylls and factor 430—with their specific complexation of iron, cobalt, magnesium or nickel ions—have been exhaustively studied in the latter part of the last century.31–34

3 Dehydrogenation of L-proline to pyrrole-2-carboxyl units after covalent protein tethering

In the past decade, the sequencing of hundreds of microbial genomes has provided insights into alternative molecular logic for the generation of pyrrolic natural products. A recurring biological form of pyrrole is the pyrrole-2-carboxylate skeleton. This moiety is sometimes seen as the free acid, but more often is further transformed in biosynthetic coupling steps.

In the antibiotics clorobiocin 8 and coumermycin A, 9, pyrrole-2-carboxylate moieties are present in ester linkages to the 3′-hydroxyls of the noviosyl sugars,35 suggesting the intermediacy of an activated form of pyrrole-2-carboxylate as an acyl donor. The pyrrole moiety in pyoluteorin 11 also utilizes an activated pyrrole-2-carboxylate scaffold, assembled in an analogous manner.5 Similarly, one—but only one (ring A)—of the three pyrrole rings in the red prodigine pigments produced by a number of Streptomyces and Serratia bacterial species is derived from such a scaffold.

In producer microorganisms, it seemed likely that the readily available proteinogenic amino acid L-proline would serve as precursor to the pyrrole-2-carboxylate moiety. This pathway must involve a controlled four-electron oxidation process. Indeed, we have previously determined that bacteria are frequently able to shunt some fraction of the pool of a proteinogenic amino acid
to a nonproteinogenic one for use in a conditional metabolic pathway. In such instances, the amino acid is often sequestered in a thioester linkage as a protein-bound intermediate on a 10 kDa peptidyl carrier protein (PCP) domain (also known as a thiolation (T) domain) (Fig. 6). This is typically the first step of nonribosomal peptide synthetase assembly line machinery. For example, L-tyrosine and L-histidine loaded as thioesters on a phosphopantetheinyl arm of specific PCP domains serve as substrates for β-hydroxylation in the first steps of vancomycin and nikkomycin biosynthesis (A similar transformation is also observed during vancomycin biosynthesis). This sequestration ensures that only the loaded fraction of the amino acid pool gets chemically modified, leaving the bulk available for primary metabolism, including protein synthesis. In addition to achieving physical sequestration by covalent tethering of amino acid substrates to proteins, thioester activation of the amino acid lowers the activation barriers for downstream aminoacyl transfer chemistry.

With this concept in mind, inspection of the gene clusters for pyoluteorin, undecylprodigiosin, clorobiocin, and coumermycin revealed conserved sets of three open reading frames (orfs) whose predicted functions seemed to validate this logic for L-proline metabolism (Fig. 7). The identified orfs were: (1) a free-standing 50 kDa adenylation (A) domain specific for L-proline activation as the L-prolyl-AMP, (2) a free standing 10 kDa PCP domain that, once primed with phosphopantetheine, may be loaded with L-prolyl-AMP to yield the tethered L-prolyl-S-PCP protein, and (3) a predicted 50 kDa flavoprotein dehydrogenase (DH), which is presumably capable of catalyzing oxidation to the pyrrolyl-2-carboxyl-S-PCP product (Fig. 6). In the pyoluteorin system, the prodigiosin system, and the coumermycin, clorobiocin, and pyoluteorin systems, this logic could be proven by in vitro studies using purified proteins. Thus it was possible to detect the L-prolyl-S-PCP substrate and, after action of the flavoprotein desaturase, the pyrrolyl-2-carboxyl-S-PCP product, or the cleaved protein-free small molecule equivalents. In the case of clorobiocin, it was shown that the net four-electron oxidation of L-proline to pyrrolyl-2-carboxylate proceeds in two discrete two-electron steps, via a dehydroprolyl intermediate tethered to the PCP domain, as established by stopped flow mass spectrometry. A [13C, 15N]-depleted 10 kDa PCP domain was used in this work to narrow the isotopic mass envelope for direct detection of the starting L-prolyl-S-PCP, the [M-2] dehydro intermediate, and the final [M-4] pyrrolyl-2-carboxyl-S-PCP product (Fig. 8). We conjecture two consecutive C–N desaturation steps by the flavoprotein, with isomerization out of conjugation in between the first and second oxidation steps (Fig. 6). The 3,4-dehydro-L-prolyl-S-PCP protein was in fact a kinetically competent intermediate en route to the pyrrolyl-2-carboxyl-S-PCP product.

This pathway of L-proline activation—involving covalent tethering and then desaturation to the heteroaromatic pyrrolyl-2-thioester—confers several advantages. The nascent pyrrolyl product is conjugated with an electron-withdrawing substituent (the 2-carboxythioester moiety), thus reducing the electron rich nature of the pyrrole and stabilizing it against auto-oxidation and adventitious capture by electrophiles. Furthermore, if desaturation of L-proline to the pyrrole-2-carboxylate occurred on the free amino acid, the carboxylate acidity may then be too low for ready activation by an ATP-dependent ligase to allow subsequent formation of the protein-bound thioester. Thus, the timing of L-proline activation prior to desaturation to the heteroaromatic pyrrole ring is unlikely to be accidental with regard to control of
reactivity. That the nascent pyrrolyl-2-carboxyl group is activated as a thioester is key to its ability to function as an acyl donor in the subsequent suite of reactions, both for carbon–carbon bond formation (e.g., in pyoluteorin 11 biosynthesis) and for ester formation (e.g., in coumermycin A₁, 9 formation), as will be noted below.

It is likely that this biosynthetic strategy is common across the entire family of mono-pyrrolic natural products, and investigations into related metabolic pathways should therefore be relatively straightforward. The identification of analogous orfs should be possible using cloning methods, with DNA primers designed against conserved sequences in one or more of the three key orfs (Fig. 7).

The pyrrole-2-carboxylate nucleus is chemically useful in that it possesses an electron-rich pyrrole core, capable of reacting with electrophiles, whilst the electron-deficient carbonyl of its thioester is susceptible to nucleophilic attack. Both of these modes of reactivity are observed in the elaboration of PCP-tethered pyrrole-2-carboxylates to their ultimate natural products.

### 4 Enzymatic halogenation of pyrrolic natural products

Dozens of pyrrolic natural products are halogenated, including via mono- and dihalogenation at C₄ and C₅ (as seen in pyrrolnitrin 12, pyoluteorin 11 and pyrrolomycin B 10, Fig. 1), trihalogenation at C₃, C₄ and C₅, and tetrahalogenation at C₂, C₃, C₄ and C₅. In the case of pyoluteorin 11, we have established that 4,5-dichlorination occurs while the heterocycle is still tethered to a PCP domain (PltL) as the protein-bound thioester, but after four-electron oxidative desaturation of the l-prolyl ring to its pyrrolyl equivalent. In vitro, the halogen atoms are installed enzymatically by a dedicated halogenase (PltA) using chloride ion, dioxygen and reduced FAD to generate a “Cl⁺” equivalent in the active site that effects iterative regioselective halogenation (Fig. 9). This oxidative halogenation chemistry for delivering “X⁺” equivalents to the electron rich pyrrole is probably general for all of the halogenated pyrrolic and indolic natural products. The halogenated pyrrolyl thioesters thus generated are used as acyl donors for subsequent enzyme-catalyzed steps in their respective biosynthetic pathways.

4-Bromo- and 4,5-dibromopyrrolyl-2-carboxylate scaffolds are common structural elements in marine natural products. These arise by analogous oxidative halogenation strategies to those previously described, however the enzymes generating HOBr equivalents are more commonly haloperoxidases, using hydrogen peroxide rather than dioxygen as co-substrate. These haloperoxidases contain either vanadate or heme iron in the active sites to bind peroxide and generate reactive hypohalite equivalents.

### 5 Enzymatic transfer of pyrrolyl-2-carboxyl to co-substrate nucleophiles

The protein-tethered pyrrolyl-2-carboxyl units generated by oxidative desaturation of l-prolyl-S-protein moieties can serve as
Fig. 7 Comparison of the gene clusters for pyoluteorin 11, undecylprodigiosin 7, clorobiocin 8 and coumermycin A 9. Adenylation (A) domains are indicated in green; peptidyl carrier protein (PCP) domains in red and dehydrogenase (DH) domains in blue.

electrophilic acyl donors in subsequent steps of enzymatic natural product assembly (Fig. 10). One example of such a transamidification step to an oxygen nucleophile is seen in the final stages of clorobiocin 8 and coumermycin A 9 formation (Fig. 10B).58,59 The putative thioesterase CloN7 (or CouN7)—acting in concert with the acyl carrier protein (ACP) CloN1 and the acyltransferase CloN2—catalyzes nucleophilic attack of the 3′-hydroxyl of the clorobiocin 8 (or coumermycin A 9) noviosyl sugar on the pyrrolyl-2-carboxyl-S-ACP domain (Fig. 11A).58,59 The roles of purified CouN1 and CouN7, from the coumermycin A 9 system, were recently confirmed in vitro.60 These Clo/Cou enzyme systems thus effect regiospecific acylation of the coumarin scaffold in the final stages of antibiotic biosynthesis. C-Methylation of the pyrrole ring at C5 by a putative methylcobalamin-containing radical-generating enzyme (CloN6/CouN6)61 may occur prior to the action of CloN7/CouN7 or may complete the biosynthesis of clorobiocin 8 and coumermycin A 9. X-Ray analysis of clorobiocin 8 in complex with its bacterial target protein, the DNA gyrase B subunit (GyrB) indicates that the pyrrolyl moiety is a key pharmacophore for binding of the antibiotic in the ATP binding site of GyrB.62 In so doing, the 2-acyl group is converted to a ketone functionality that persists on further chain elongation. On final cyclization and aromatization, this carbonyl is present as the ketone moiety of the released pyoluteorin 11. It is most likely that this is the general mechanism for formation of equivalent pyrrolyl-β-keto moieties that become embedded in natural product scaffolds. This carbon–carbon bond-forming step represents an interface between NRPS and PKS modules in the assembly of hybrid NRP/PK skeletons. Such NRPS/PKS interfaces are of interest in understanding the functional and structural interplay between these different classes of enzymes.

Inspection of the structures of the pyrrole-imidazole class of alkaloids (e.g., oroidin 26, Fig. 12A) suggests operation of the same strategy, diverting L-proline to pyrrolyl-2-carboxyl-S-carrier proteins for enzyme-catalyzed capture by amine nucleophiles (Fig. 10D). Biosynthesis of such mono- or dibromo pyrrole-imidazole alkaloids may be initiated by transfer of the halogenated pyrrolyl-2-carboxyl-S-PCPs to the α-amino group of L-arginine or L-homoarginine. Subsequent oxidation and decarboxylation, and cyclization of the guanidine side chain, may then afford the natural product, oroidin 26 (Fig. 12A).64 Oroidin itself or its
monobromo-analogue may in turn be a precursor to more complex bromopyrroles such as the marine natural products, axinellamines A–D (e.g., 27, Fig. 12A), sceptrin 28 and ageliferin 29. In an analogous manner, capture of the halogenated pyrrolyl-2-carboxyl-S-PCP by the ε-amino group of L-lysine leads to additional variants of the pyrrole-amide scaffold.

Further variations of the pyrrolyl-2-carboxylate building block occur in the pyrrolosesquiterpenes such as the glaciapyrroles (e.g., 31) the pyrrolomonoterpenes such as pyrrolostatin 25 and the signaling molecule ryanodine 30 (Fig. 12). Pyrrolostatin 25 is likely formed via the pyrrolyl-2-carboxyl-S-PCP route converging with a C₁₀ Δ²-isoprenoid, geranyl pyrophosphate, by nucleophilic attack of C₄ of the pyrrole ring on C₁ of geranyl pyrophosphate (Fig. 12B). Similarly, nucleophilic attack on farnesyl pyrophosphate by a pyrrole-2-carboxylate moiety is likely during the biosynthesis of the glaciapyrroles (e.g., 31), with further (or
concomitant) decarboxylation of the pyrrole-2-carboxyl species and subsequent oxidative modifications.

Biosynthesis of the pyroloquinoline quinone (PQQ) cofactor 32 (Fig. 12C), on the other hand, is initiated by the dimerization of L-tyrosine and L-glutamate, followed by extensive oxidation. The structurally related makaluvic acid pyrroloquinolines (e.g., makaluvamine M 33, Fig. 12C) probably arise by a similar oxidative strategy, perhaps starting instead from L-tryptophan 5. A large family of prenylated L-tryptophan-based natural products have been isolated, the biosyntheses of which often involve extensive oxidation as well as prenylation and further tailoring modifications. Among many examples of such L-tryptophan-derived alkaloids that could be cited are the fischerindoles and welwitindolinones (e.g., welwitindolinone A 34), the stephacidins (e.g., stephacidin B 35)72–74 and the chartellines (e.g., chartelline C 36) (Fig. 12C).75

6 Prodigiosin biogenesis: three routes to the three pyrrole rings

*Streptomyces* and *Serratia* bacterial species are known to produce red pigments that have antibiotic activity. These compounds, the prodiginines (e.g., 6 and 7, Fig. 1B) have a wide array of biological activities—including anticancer, antimicrobial and immunosuppressive activities76–80—and an intriguing arrangement of three pyrrole rings, two of them directly coupled. Sequencing of gene clusters from multiple prodiginine-producing bacteria has allowed for bioinformatic prediction, genetic evaluation and direct biochemical studies of the biosynthetic routes to each of the three pyrrole rings.6,7

The pyrrole A-ring is made from L-proline by exactly the logic and enzymatic machinery noted above (Fig. 6). Subsequently, a ketosynthase (KS) domain-tethered pyrrolyl-2-carboxythioester is transferred to a malonyl-ACP domain on PigH 30 by decarboxylative Claisen condensation (Fig. 13A). The two contiguous ACP domains of PigH are required for condensation to occur. At this point, the amino acid L-serine 37 provides the remaining two carbons for the pyrrole B-ring by undergoing the same kind of PLP-mediated decarboxylation71 used by δ-aminolevulinic acid synthase at the start of tetrapyrrole assembly.23,24

As noted in Fig. 13A, the masked carbanion resulting from PLP-mediated loss of the L-serine carboxylate can attack the pyrrolyl-b-ketoacyl chain giving carbon–carbon bond formation. Enzymatic hydrolysis of the PLP-imine linkage liberates the free amino group that then serves as an intramolecular nucleophile for formation of the second pyrrole ring. In vitro characterization of the enzymes PigA (L-prolyl DH domain), PigG, PigH, PigI (L-prolyl A domain) and PigJ has recently been reported, demonstrating reconstitution of the biosynthetic pathway to the prodigiosin A- and B-rings.45

The free bi-pyrrolic alcohol 38, and the aldehyde 39 formed by subsequent enzymatic oxidation and methylation, are known intermediates in the prodigiosin pathway.7 Unlike the linear poly-pyrroles involved in heme assembly, these two pyrroles are directly tethered via their C2 carbons, without an intervening methylene bridge, reflecting a different chemical logic inherent to their assembly.

The third pyrrole ring in the prodigiosin 6 scaffold is thought to be assembled from a fatty acid synthase-derived starting

---

**Fig. 10** Schematic of elaboration of the pyrrolyl-2-carboxyl scaffold and its incorporation into natural products by (A) hydrolysis; (B) nucleophilic attack of an alcohol; (C) attack by a carbon nucleophile; and (D) amide formation.
Fig. 11  Incorporation of the pyrrolyl moiety of pyrrolyl-2-carboxyl-S-PCP into natural products via (A) attack of noviosyl 3-hydroxyl leading to clorobiocin 8; and (B) carbanion attack during pyoluteorin 11 biosynthesis.

material, using the thiamine pyrophosphate (TPP)-dependent PigD and another variation of the PLP logic already discussed (Fig. 13B). Addition of a pyruvate-derived two-carbon fragment to 2-octenal 40 (derived from the fatty acid pathway) yields a long chain γ-ketoaldehyde moiety 41 (Fig. 13B). Action of the putative PLP-dependent aminotransferase, PigE, with an (as yet unidentified) amino acid acting as the amine donor, results in the PLP-imine form of the γ-aminoketone. Upon imine hydrolysis, the amino ketone cyclizes, and subsequent oxidation by PigB yields the mono-pyrrolic C-ring precursor 42 (Fig. 13B). The
mechanism of this PigB-mediated aromatization step has not yet been fully elucidated, but is postulated to follow an FAD-dependent oxidation pathway.

Interestingly, in the closely related undecylprodigiosin 7 system, the analogous mono-pyrrole ring is thought to be assembled by a PKS system, with PLP-mediated decarboxylation of glycine providing the final atoms of the nascent five-membered ring. In this case, carbon–carbon bond formation occurs between the α-carbon of a decarboxylated, PLP-tethered glycine and a β-ketoacyl-S-ACP species. Thus, in contrast to the prodigiosin system, one of the carbons of the undecylprodigiosin pyrrole ring is derived from glycine.

Once the alkyl-substituted pyrrole 42 is generated, this electron rich heterocycle could easily react with the bi-pyrrole aldehyde 39 (Fig. 13B) to produce the prodigiosin 6 (or undecylprodigiosin 7) backbone. Subsequent tailoring reactions then lead to the observed suite of prodigamine variants.

7 Indolocarbazole biosynthesis: oxidation and dimerization of L-tryptophan generates chromopyrrolic acid

The natural products staurosporine 13 and rebeccamycin 14 (Fig. 14) are structurally related six-ring indolocarbazole systems derived from oxidative coupling and metabolism of two molecules of L-tryptophan. The aglycones differ in two ways: (a) by chlorination at C87 of the indole moiety of each L-tryptophan; and (b) by four electrons in the oxidation state of one carbon in the five-membered ring, resulting in the rebeccamycin maleimide functionality and a four electron-reduced pyrrolinone in staurosporine. Since the naked indolocarbazole scaffolds possess very poor water solubility, the last steps in the biosynthetic pathway constitute N-glycosylation steps. In rebeccamycin 14, D-glucose is coupled to one of the indole nitrogens via an N-glycosidic linkage; in the case of staurosporine 13, an aminodeoxy-L-hexose
Fig. 13  Biosynthesis of prodigiosin 6, outlining construction of (A) the B-ring and (B) the C-ring. Biosynthesis of the A-ring proceeds as outlined in Fig. 6. The L-serine transferase domain is denoted as SerT.

(l-ristosamine) forms a glycosidic bond to one indole nitrogen, and also becomes oxidatively linked to the other indole nitrogen via its C5 position. Rebeccamycin 14 is a potent inhibitor of mammalian DNA topoisomerase I and has been investigated as an antitumor agent, and staurosporine 13 is a nanomolar inhibitor of many serine/threonine protein kinases. The planar aglycone cores are clearly important for interactions with their biological targets.

The rebeccamycin 14 and staurosporine 13 biosynthetic pathways are short and efficient with only four or five enzymes required to convert L-tryptophan 5 to the complete aglycone scaffolds, 45 and 46. The rebeccamycin pathway begins with an L-tryptophan-7-halogenase, RebH52,53 (absent in the staurosporine pathway); subsequent oxidation of L-tryptophan (or 7-Cl-L-tryptophan) to the carboxylic acid imine 43 is catalyzed by the amino acid oxidase RebO (or StaO). Oxidative dimerization by the heme oxidase (RebD or StaD) leads to the formation of a nascent pyrrolic scaffold in the five ring intermediate, chromopyrrolic acid (CPA) 44 (Fig. 14). This dicarboxypyrrole unit is formed as the joining element between the two L-tryptophan-derived skeletons. A plausible dimerization route involves attack by the enamine form of one molecule of 43 on the imine form of a second, followed by a net two-electron oxidative carbon–carbon bond-forming process to create the pyrrole ring.

To complete the formation of the planar fused six-ring indolocarbazole scaffolds 45 and 46, the oxidation state of the pyrrole ring is altered and aryl–aryl coupling effected. The enzymes responsible for this transformation are StaP/RebP, heme proteins of the cytochrome P450 family, and StaC/RebC, putative FAD monooxygenases. Acting in concert, these enzymes mediate double oxidative decarboxylation and aryl–aryl coupling of 44 to yield the sixth ring and complete the indolocarbazole scaffolds. In the process, the pyrrole ring is oxygenated once to make the staurosporine nucleus 45 and twice to yield the rebeccamycin aglycone 46. Whilst StaP (and presumably RebP) is competent to form both oxidation states of the aglycone scaffold—as well as 7-hydroxy-K252c, an aglycone that is an intermediate in oxidation state—the presence of StaC or RebC accelerates the reaction and directs the oxidative flux towards either the staurosporine- or rebeccamycin-related aglycone form.
The transient formation of the pyrrole moiety that occurs during biosynthesis of rebeccamycin 14 and staurosporine 13 is also implied in the pathways to related secondary metabolites. The biosynthetic cluster for the indolocarbazole violacein 47 (Fig. 14) contains enzymes with very high homology to enzymes of the rebeccamycin and staurosporine pathways. In particular, the only close homologue to the unusual heme oxidases RebD and StaD is the enzyme VioB from the violacein cluster.95 This enzyme is postulated to perform an analogous oxidative coupling reaction of two L-tryptophan-derived monomers with an unusual net [1,2]-shift of one of the indole rings. Interestingly, the same indolocarbazole core of violacein 47 is elaborated further in the unique pigments of the chromoviridans family (e.g., 48, Fig. 14). The central dipyromethane unit (similar in connectivity to the tetrapyrole macrocycles) is presumably formed by attack of two violacein precursors on N7,N10-methyleneetrahydrofolate (or a related species), with the central carbon unit being derived from l-serine.96 It therefore seems apparent that the pyrrole intermediate of CPA 44 is relatively versatile, being shuttled towards different oxidation states and substitution patterns depending on the constitutive parts of the biosynthetic machinery.

8 Conclusions

The electron rich, planar pyrrole ring system is a useful heterocycle for a myriad of chemical and structural roles in biological systems. In this short Highlight, we have not focused on the well established roles of metal cation coordination in the macrocyclic tetrapyrole pigments of life,97–99 nor on the reading of DNA sequences by electron donation and hydrogen bonding of poly-pyrrolic molecules such as distamycin 49 and netropsin 50, but there is extensive literature available on those topics.100–102 Rather, we have noted several distinct enzymatic routes and strategies that Nature takes to make both mono-pyrrolic and poly-pyrrolic scaffolds.

Although the knowledge base gleaned from bacterial genomics, bioinformatics and biochemical studies is extending to encompass much of the biosynthetic logic for pyrrole assembly, there are still many aspects that are unknown. Among many examples that could be cited are (1) possible biosynthetic routes to the 2,4-dicarboxypyrrrole linker in coumermycin A1,9; and (2) the pathway to the 4-aminopyrrole-2-carboxylates in distamycin family members (e.g., 49 and 50, Fig. 15). The current state of knowledge on the biosynthesis of other mono- and poly-pyrrolic small molecules should constitute a good basis for elucidation.
of this mechanistic detail. Knowledge of biosynthetic pathway logic and enzymatic machinery will be a starting point for the programming and reprogramming of enzymatic assembly lines, enabling the installation of pyrroles and related heterocycles into novel natural product variants.

9 References
