NJC

Considerations of the Biosynthesis and Molecular Diversity of Tolyporphins

Journal:	*New Journal of Chemistry*
Manuscript ID	NJ-ART-04-2021-001761.R1
Article Type:	Paper
Date Submitted by the Author:	04-Jun-2021
Complete List of Authors:	Lindsey, Jonathan; North Carolina State University, Chemistry

Considerations of the Biosynthesis and Molecular Diversity of Tolyporphins

Jonathan S. Lindsey

Department of Chemistry

North Carolina State University

Raleigh, NC 27695-8204

e-mail: jlindsey@ncsu.edu

Abstract

Tolyporphins A–R constitute fundamentally distinct members of the tetrapyrrole pigments of life family. The 18 members present diversity at multiple levels including the chromophore (dioxobacteriochlorin, oxochlorin, porphyrin); composition of the pyrroline substituents (hydroxy, acetoxy, or one of four *C*-glycosides); and stereochemical configuration of the pyrroline substituents. Eleven of the 18 tolyporphins contain at least one *C*-glycoside; each *C*-glycoside has a β,D configuration and lacks a 6′-hydroxy group: 3′,6′-dideoxygalactose (common name abequose), 2′-*O*-acetyl-3′,6′-dideoxygalactose (2′-*O*-acetylabequose), 6′-deoxygalactose (D-fucose), or 6′-deoxygulose (antiarose). Rare are such glycosides outside of tolyporphins: (2′-*O*-acetyl)abequose is reported only in the glycan polymer attached to the cell wall of two strains of Gram-negative bacteria, and antiarose is reported in one bacterial natural product and ~50 plant cardiac glycosides. Eight of the 18 tolyporphins are bis(*C*-glycosides), an exceptionally uncommon motif in natural products. The biosynthetic pathways to the family of tolyporphins remain unknown. Regardless of such diversity, each tolyporphin member shares a common pattern of perimeter methyl substituents that coheres with derivation from uroporphyrinogen III, the universal precursor in the established pathway to native tetrapyrroles. Here, transformations required to convert uroporphyrinogen III to all 18 tolyporphins are considered in the context of plausible biosynthetic pathways. Heme d_1, perhaps the closest relative (yet still a distant cousin) of tolyporphins, and for which key biosynthetic transformations remain undeciphered, provides a point of reference. Taken together, the work provides the foundation for bioinformatic searching for enzymes associated with the biosynthesis and diversification of tolyporphins.

Introduction

The discovery of tolyporphin A in 1992 added a fundamentally new tetrapyrrole structure to the venerable pigments of life family.[1] Tolyporphin A was identified in an empirical search for anti-neoplastic agents in cyanobacteria.[2] The cyanobacterial culture HT-58-2, collected in Nan Madol on the island of Pohnpei in Micronesia, remains the only known producer of tolyporphin A. Tolyporphin A presents striking features especially in contrast with other members of the pigments of life family (Fig. 1).[3,4] The features include the following: (i) the absence of alkyl substituents at two β-pyrrole positions; (ii) the presence of a *C*-glycoside as part of a gem-dialkyl motif, along with a flanking oxo group, in each pyrroline ring; and (iii) the absence of a centrally chelated metal (i.e., a free base) versus the customary metalation observed with heme (Fe), chlorophylls and bacteriochlorophylls (Mg), cobalamin (Co) and coenzyme F_{430} (Ni).

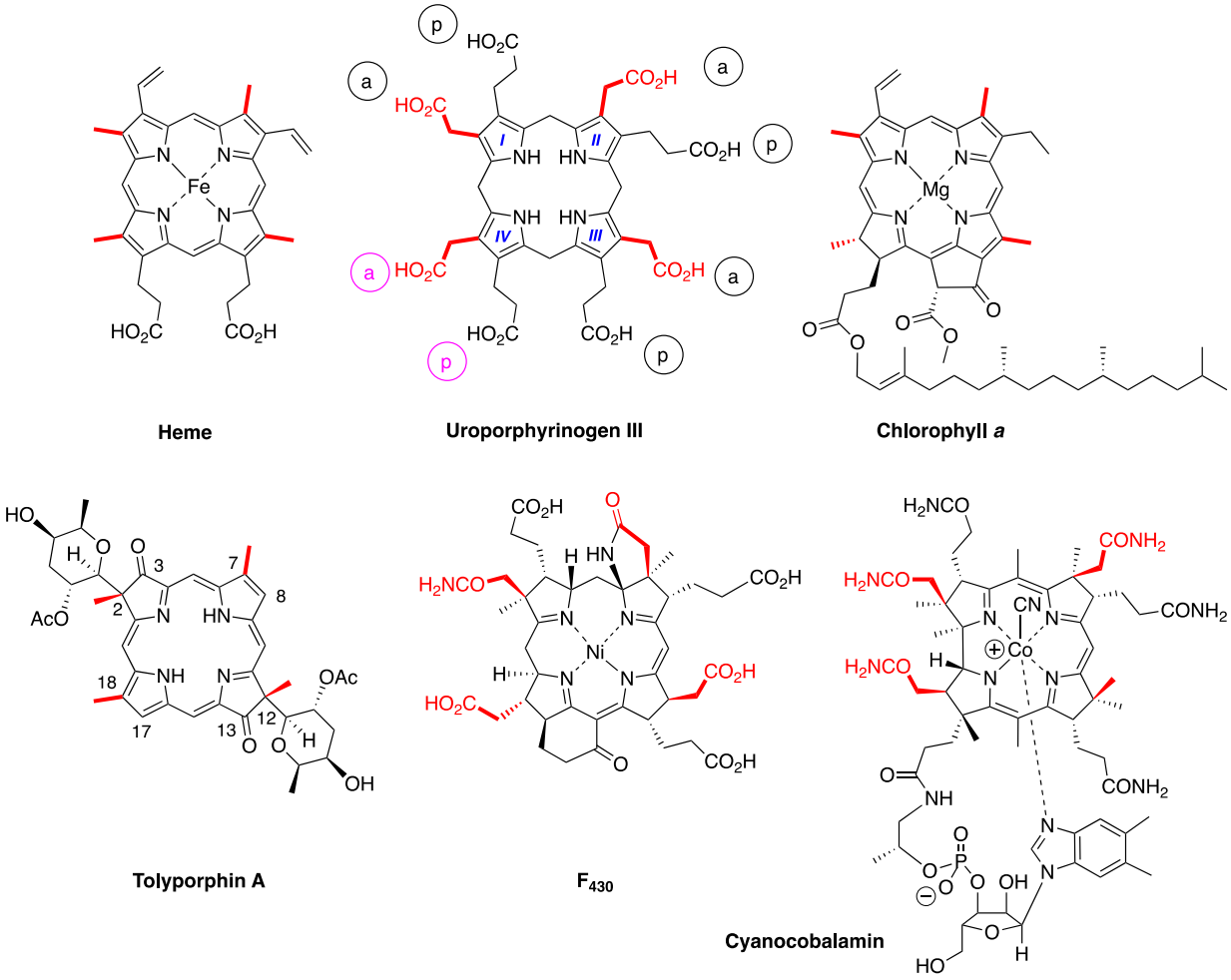


Fig. 1 Structures of tetrapyrrole macrocycles (common substituent pattern shown in red).

In ensuing years, the tolyporphin family has grown to include 18 members (A–R), all from the HT-58-2 culture,[5-7] although tolyporphin A is the dominant member (up to approximately half of the total).[8] The structures are shown in Fig. 2. The various tolyporphins have been noted in a number of reviews.[9-13] The structures vary in the nature of the chromophore (dioxobacteriochlorin, oxochlorin, or porphyrin) and the pyrroline substituent (hydroxy, acetoxy, or one of several *C*-glycosides). The presence of such variegated diversity suggests the tolyporphins are likely secondary metabolites and may play roles in defense functions of the microorganism rather than as protein-bound, biological cofactors.

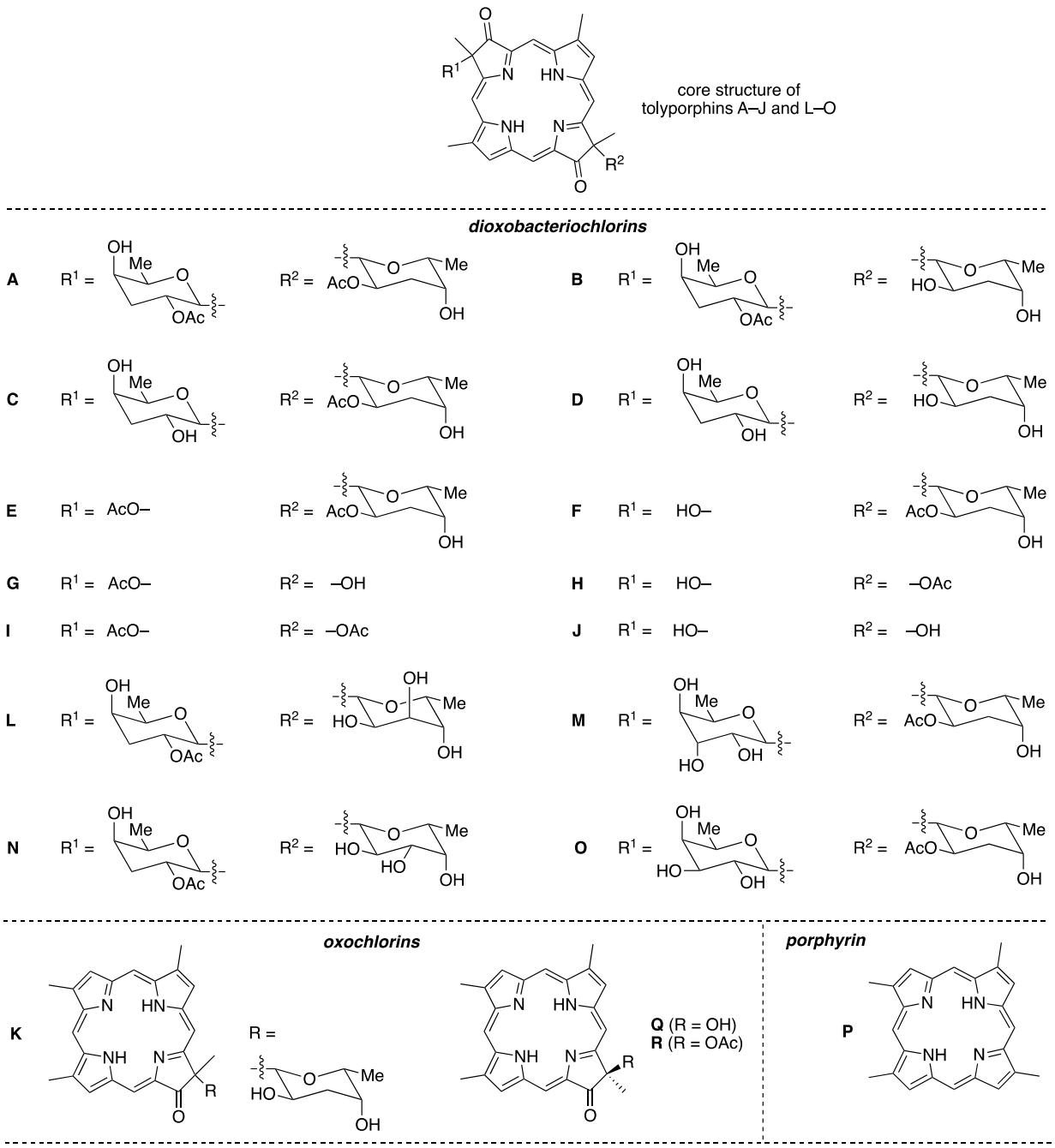


Fig. 2 Structures of tolyporphins.

While there is sizeable diversity within the tolyporphins family, one commonality concerns the pattern of methyl groups about the perimeter of the macrocycle. In tolyporphin A, the pattern of methyl groups upon circumambulating the macrocycle from rings I through IV is MeX-MeX-MeX-XMe, where X is one of the substituents at the β-pyrrolic or β-pyrrolinic position (Fig. 1).

The same pattern with one "flipped MeX" occurs for all tolyporphins A–R (Fig. 2). Inspection of heme, chlorophyll, cobalamin and F_{430} reveals the identical pattern (Fig. 1). The latter tetrapyrroles are known to derive from a common biosynthetic precursor, uroporphyrinogen III,[14] wherein the substituent pattern in circumambulating the macrocycle (from rings I to IV) is ap-ap-ap-pa with the abbreviations "a" and "p" standing for acetic acid and propionic acid, respectively. The acetic acid substituents undergo decarboxylation, yielding methyl groups, in the conversion of uroporphyrinogen III to coproporphyrinogen III. The congruent substitution pattern is strong evidence in support of derivation of the family of tolyporphins from uroporphyrinogen III.[3,4]

We obtained the HT-58-2 culture in 2015 and have carried out a number of studies, including assessing the non-axenic nature of the culture and determining[15] and annotating[16] the closed genome of the cyanobacterium (7.85 Mbp) as well as that of a dominant community (non-cyanobacterial) bacterium (3.23 Mbp).[17] Yet a chasm exists between knowing the genome sequence and knowing the enzymes that enable the biosynthetic pathway, particularly for a family of compounds with unprecedented structural features. Other groups have continued to mine the HT-58-2 culture and have found non-tetrapyrrole natural products, the tolypodiols.[18,19] Our long-term objectives are both fundamental and practical – we want to gain deeper insight into the dichotomy between the formation of hydroporphyrins with gem-dialkyl groups (e.g., cobalamin, F_{430}) and hydroporphyrins with *trans*-dialkyl groups (chlorophylls and bacteriochlorophylls), and we would like to exploit the enzymes responsible for tolyporphins in chemoenzymatic preparation of designer hydroporphyrins.

This paper considers the structures of tolyporphins in the context of possible routes of formation, not with regards to specific enzymes, but in terms of known transformations in the biosynthesis of tetrapyrrole macrocycles and in the larger body of organic chemistry. This work

provides a conceptual framework for the bioinformatics search for enzymes for the biosynthesis of tolyporphins.

Results and Discussion

Derivation of tolyporphin A from uroporphyrinogen III would require the following steps: (i) removal of all four propionic acid side chains; (ii) installation of two oxo groups; (iii) attachment of the two *C*-glycosides; and (iv) dehydrogenation of the macrocycle. For formation of the tolyporphins B–R, the same general features hold with the following twists: (a) only one oxo group is installed for the oxochlorins (tolyporphins K, Q and R), and none for the porphyrin (tolyporphin P); (b) a hydroxy or acetoxy group is attached in lieu of a glycoside (tolyporphins F–I); (c) one of four distinct glycosides may be installed; and/or (d) the installed glycoside moiety may undergo hydroxylation or dehydroxylation at the 3′-position, stereochemical inversion of the 3′-hydroxyl group, and/or acetylation at the 2′-hydroxyl group (glycosyl numbering distinct from the parent tolyporphin numbering). A further non-trivial wrinkle concerns the stereochemistry of the appended groups: the two glycosyl groups are on the same face of the macrocycle in tolyporphin A[20-22] whereas the analogous substituents are on opposite faces in tolyporphin E, and the one substituent in the oxochlorin tolyporphin R (–OH) is opposite that in the dioxobacteriochlorin tolyporphin A.[7] The corresponding stereochemistry in the other tolyporphins has not yet been unambiguously established. Here, each process is considered in turn.

Reconnaissance. The removal of a propionic acid group and introduction of an oxo group at a β-pyrrolic position, while an exquisitely rare biosynthetic transformation, occurs twice in the formation of heme d_1 (Fig. 3). Heme d_1 is a dioxo*iso*bacteriochlorin. A radical SAM enzyme[23] (NirJ) is believed to carry out this process,[24] yielding the gem-dialkyl-substituted pyrroline units of pre-heme d_1, although the mechanism[24] and even the exact intermediates (e.g., **1**–**3**)[25] remain

unclear. One proposed radical mechanism proceeds via intermediates **4** and **5** with hypothetical liberation of acrylic acid (**6**). The installation of the β-oxo group occurs in subsequent steps to form heme d_1. The removal of all four propionic acid groups as required to form tolyporphin A is unprecedented in tetrapyrrole science, as is the variable installation of *C*-glycoside, hydroxyl, or acetoxy groups in the larger family of tolyporphins (B–R). The objectives here concern the origin of diversity in the context of plausible reaction pathways, encompassing but not limited to that proposed for the NirJ-mediated process in heme d_1 biosynthesis, given how profoundly different the tolyporphins are from other members of the pigments of life family.

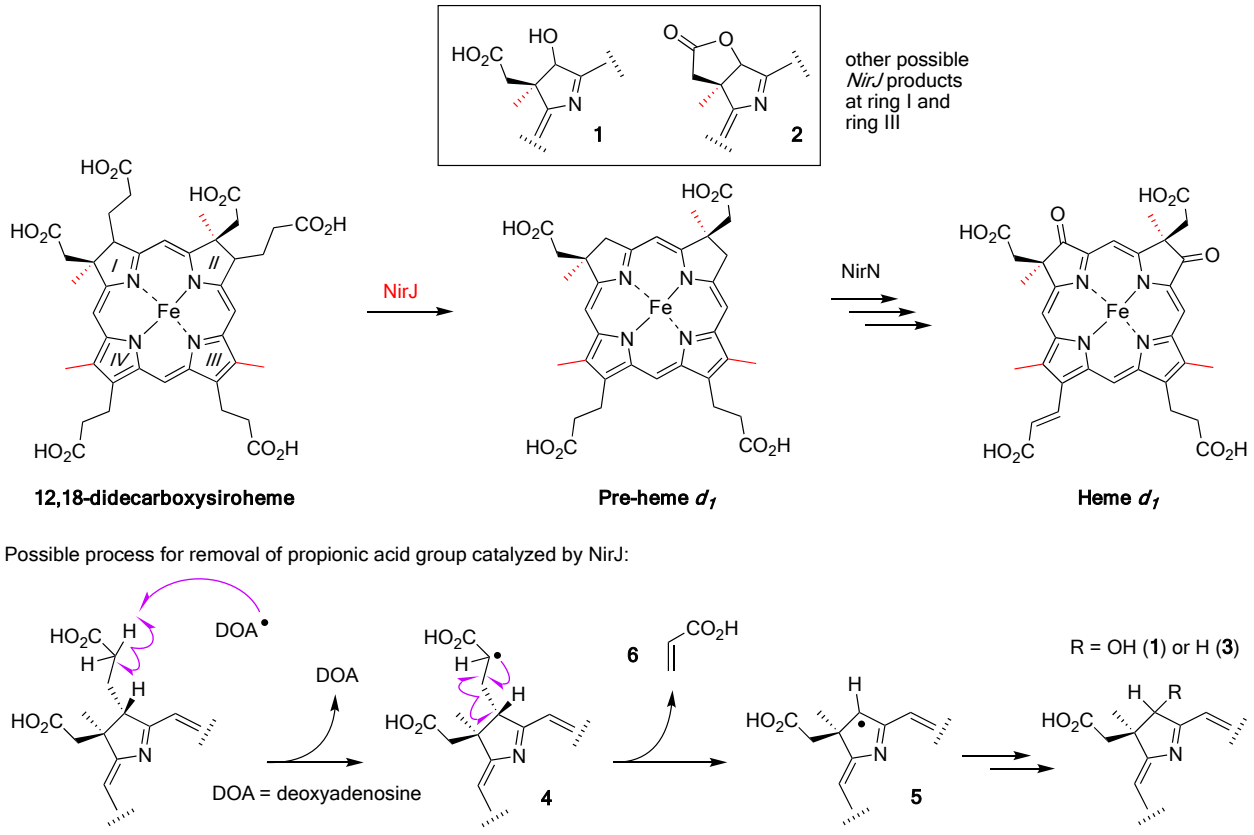


Fig. 3 Late stages in the biosynthesis of heme d_1 (top) and possible mechanism[24] for NirJ-mediated removal of a propionic acid group (bottom).

Pyrroline formation from a porphyrinogen. One pathway to form the pyrroline motif of tolyporphin A begins with coproporphyrinogen III (Fig. 4, path 1). Alkylation of the 2-position (ring I) with a glycosylating agent (such as a glycosyl phosphoester[26,27] as indicated by G–O–Pi) installs the *C*-glycoside (**7**), simultaneously creating the gem-dialkyl motif and introducing unsaturation between the pyrrole α-carbon and the adjacent meso-carbon. The process mirrors that for electrophilic substitution of pyrrole (see Appendix) where the glycosylating agent serves as an electrophile (E^+). Subsequent reaction at the 3-position with an oxygen electrophile XOZ (e.g., dioxygen or other activated oxygen entity such as in a cytochrome P_{450}) affords the pyrroline bearing both the propionic acid group and an oxygen species (**8**). Base abstraction of an α-proton of propionic acid and β-elimination in a retro-Michael sense liberates acrylic acid (**6**). Enamine-like displacement of the nucleofugal group attached to oxygen then affords the macrocycle containing a 3-oxo group (**9**). A radical-SAM mediated process (via NirJ)[24,25] is equally plausible (not shown).

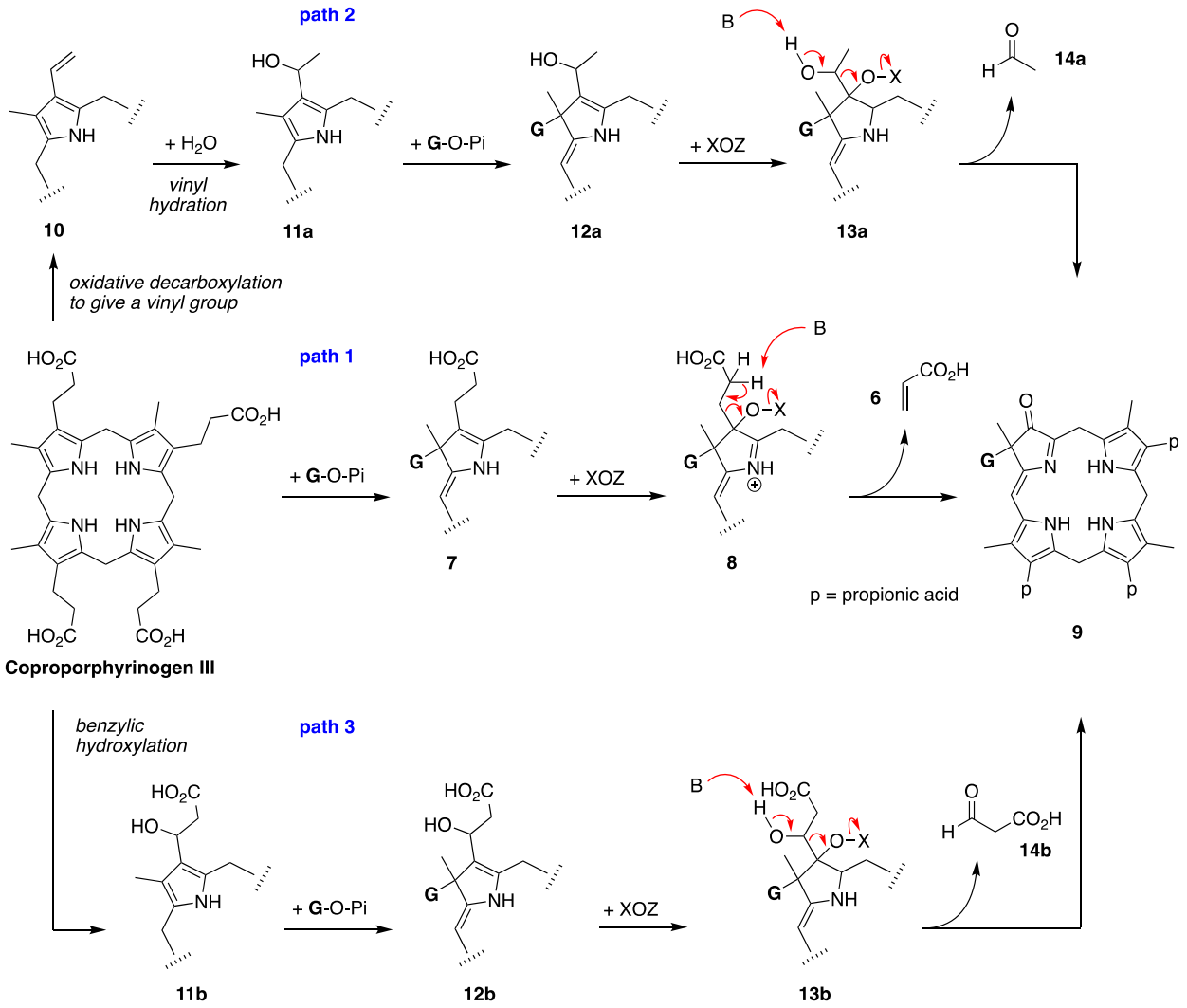


Fig. 4 Formation of the 2-glycosyl-3-oxopyrroline motif (**9**) accompanied by 3-dealkylation (G = glycosyl unit; X and Z are leaving groups; B is a base; Pi is phosphate).

A second pathway (Fig. 4, path 2) begins with the 3-vinyl group (**10**), which is produced by oxidative decarboxylation of the propionic acid substituent, an established pathway for the conversion of coproporphyrinogen III to protoporphyrinogen IX.[28] Hydration of the vinyl group in a Markovnikov sense affords the α-hydroxyethyl substituent (**11a**), a process known in tetrapyrrole biosynthesis albeit of intact chlorins, not porphyrinogens (*vide infra*). *C*-glycosylation occurs with a glycosyl–phosphoester at the pyrrolic β-position (**12a**). Enamine attack on an

electrophilic oxygen species (XOZ) results in oxidation at the adjacent pyrrolic β-position (**13a**). Base abstraction of the hydroxyl proton and cleavage in a retro-aldol sense liberates acetaldehyde (**14a**) along with the nucleofugal group attached to oxygen to afford the 3-oxo group (**9**).

A third pathway begins with benzylic hydroxylation of the propionic acid group (Fig. 4, path 3) to give intermediate **11b**. The process thereafter proceeds through intermediates **12b** and **13b**, resembling that in path 2, albeit with expulsion of 3-oxopropionic acid (**14b**, malonaldehydic acid) to install the 3-oxo group flanked by the gem-dialkyl substituent (**9**). The path is not precedented in tetrapyrrole biosynthesis, but benzylic hydroxylation is an established chemical and metabolic process.

Repeat of either of the oxo-pyrroline-forming processes at the distal pyrrole ring (ring III) installs the second *C*-glycoside at the 12-position and liberates an additional acrylic acid or acetaldehyde unit (not shown). Alternatively, the presence of an electrophilic oxygen species in lieu of the glycosylating agent would result in installation of a hydroxy group at the 2- or 12-positions.

Dealkylation without pyrroline formation. A striking structural feature of all tolyporphins (except P) is the presence of at least one pyrroline motif. Yet the presence of 2–4 open β-pyrrole positions across the family of tolyporphins A–R is equally unusual. The putative transformation from coproporphyrinogen III requires removal of 2–4 propionic acid groups and construction of 2–0 pyrroline groups. Pathways similar to those in Fig. 5 are readily extended to strip the macrocycle of propionic acid groups.

When the electrophile is a proton instead of a glycosyl moiety, β-pyrrole protonation sets up a reaction channel for dealkylation without concomitant formation of the geminal substituent. For example, protonation of the hydrated vinyl product (**11a**) affords intermediate **15a**, which eliminates acetaldehyde (**14a**) (Fig. 5, top). Similarly, protonation of the product of benzylic

hydroxylation (**11b**) affords intermediate **15b**, which eliminates malonaldehydic acid (**14b**) (Fig. 5, top). In both cases, dealkylation proceeds without *C*-glycosylation or hydroxylation to afford the macrocycle descended from uroporphyrinogen III that has an open (i.e., unsubstituted) β-pyrrole position (**17**). Alternatively, protonation of the β-pyrrole position bearing the propionic acid group in coproporphyrinogen III affords **16**, setting up a process where the propionic acid substituent can be lost as acrylic acid (**6**), again without *C*-glycosylation or hydroxylation (Fig. 5, bottom). Repetition of any of the dealkylation processes at the distal pyrrole ring (ring IV) liberates an additional molecule of acrylic acid, acetaldehyde, or malonaldehydic acid (not shown). While the particular pathway of course remains to be elucidated, the ability to remove a propionic acid group with retention of the pyrrole unit (i.e., without formation of an oxopyrroline) is essential for biosynthesis of all tolyporphins: there are two such pyrroles in the dioxobacteriochlorins tolyporphins A–J,L–O; three in the oxochlorins tolyporphins K, Q and R; and four such pyrroles in the porphyrin tolyporphin P.

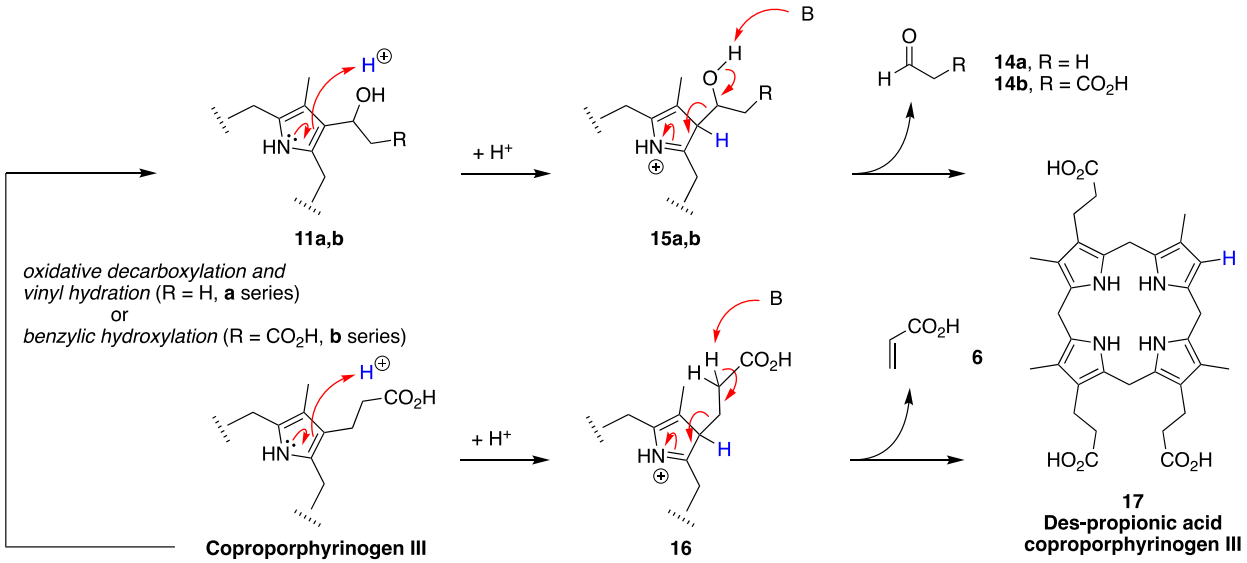
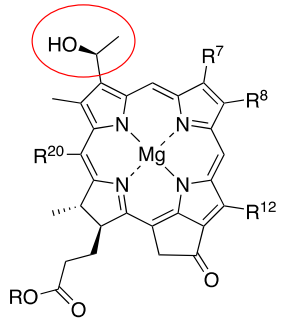


Fig. 5 Dealkylation processes following β-protonation.

Precedence for hydrated vinyl groups. Several routes above propose the intermediacy of an α-hydroxyethyl group (**11a–13a**, **15a**). Hydration of a vinyl group to form an α-hydroxyethyl substituent is a known biosynthetic process (as well as an established synthetic process). Bacteriochlorophylls *c–e*, the "*Chlorobium* chlorophylls" of green sulfur bacteria, each contain an α-hydroxyethyl group at the 3-position (Fig. 6). Hydration occurs in the penultimate biosynthetic step in the path to bacteriochlorophyll *d*, with the final step entailing attachment of the farnesyl chain.[28] The hydrated product, bacteriochlorophyllide *d*, is the biosynthetic precursor to bacteriochlorophylls *c* and *e*.



R = farnesyl
R^8 = Et, Pr, *iso*-Bu, neopentyl
R^{12} = Me or Et

Bacteriochlorophyll *c* : $R^7 = CH_3$, $R^{20} = CH_3$
Bacteriochlorophyll *d* : $R^7 = CH_3$, $R^{20} = H$
Bacteriochlorophyll *e* : $R^7 = CHO$, $R^{20} = CH_3$

Fig. 6 Natural tetrapyrroles with an α-hydroxyethyl substituent.

Pyrroline formation following oxidation. The presence of unsubstituted β-pyrrole positions opens a site for oxidation, which in turn opens a distinct pathway for formation of the β-oxopyrroline motif as shown in Fig. 7. Thus, removal of a propionic acid group (as in Fig. 5) is followed by formation of the porphyrin (**18a,b**). Subsequent arene-like hydroxylation of the β-pyrrole position affords an enol (**19a,b**), which engenders alkylation at the adjacent β-pyrrole position. In so doing, the β-pyrroline sp^3-hybridized center is created adjacent to the oxo motif

(**20a,b**). This route lacks the concision of the processes shown in Fig. 4 and Fig. 5, where removal of the alkyl group (propionic acid, α-hydroxypropionic acid, or α-hydroxyethyl) is accompanied by β-oxidation. However, the fact that alkyl groups can be removed without concomitant oxidation prompts consideration of a route where formation of the oxo motif follows dealkylation. Note that this route is displayed for use with a porphyrin (**18**), not a porphyrinogen (e.g., **17**). A porphyrinogen contains four isolated pyrrole units, which upon β-hydroxylation would not likely support enol alkylation owing to loss of pyrrole aromaticity.

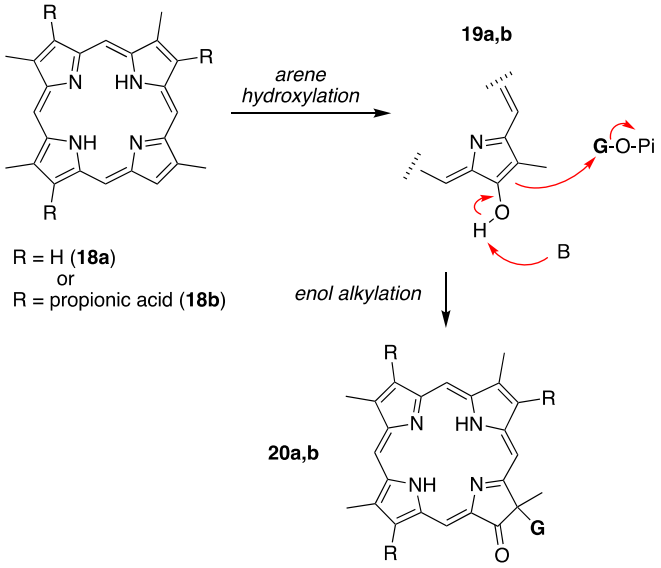


Fig. 7 Arene hydroxylation followed by enol alkylation.

Macrocycle diversity and dehydrogenation. The issue of how molecular diversity might arise in the tolyporphins family is considered next. The installation of a hydroxyl or *C*-glycosyl group at the 2- and 12-positions is a matter of the presence of the glycosyl or hydroxyl electrophile. Thus, the presence of an electrophilic hydroxyl entity (HO–X) instead of a glycosylating agent (G–O–Pi) in Fig. 4 and Fig. 7 would install an –OH group instead of a *C*-glycosyl group, respectively. The question of the number of such groups depends on the extent of processing to form β-oxopyrroline groups. With installation of two β-oxopyrroline groups, the resulting

macrocycle is a 5,15-dihydro-3,13-dioxobacteriochlorin (**21**). With installation of one or no β-oxopyrroline groups, the resulting macrocycle is a tetrahydro-13-oxochlorin (**23**) or a hexahydroporphyrin (**25**) (i.e., a porphyrinogen), respectively. The dehydrogenation of meso-positions is well known, as in the conversion of a porphyrinogen to a porphyrin.[14,28] Here, the final step following tailoring of peripheral substituents entails dehydrogenative removal of $2e^-/2H^+$ (**21** → **22**), $4e^-/4H^+$ (**23** → **24**), or $6e^-/6H^+$ (**25** → **26**) to form the corresponding dioxobacteriochlorin, oxochlorin, or porphyrin macrocycle (Fig. 8).

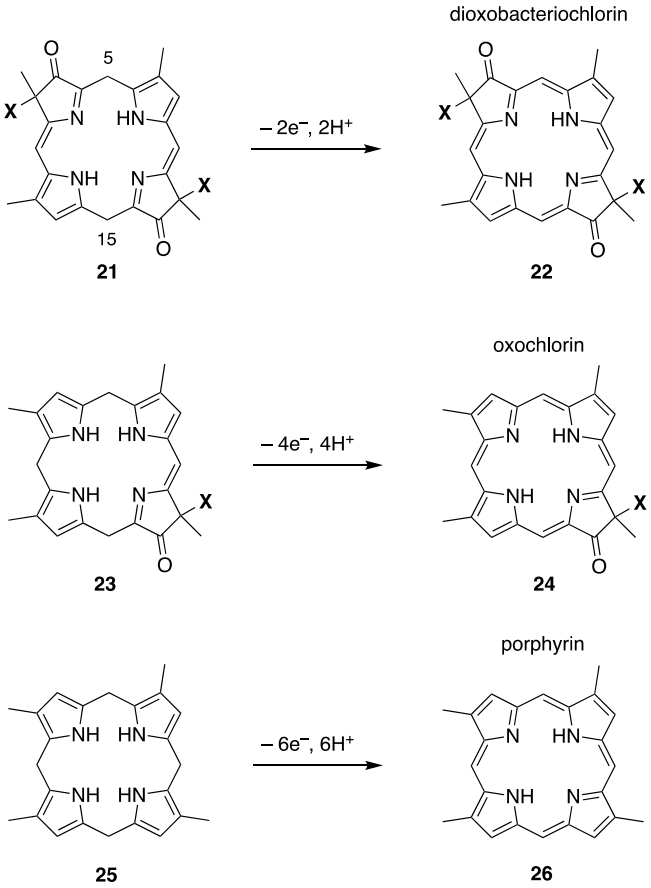


Fig. 8 Dehydrogenation in the final step of formation of tolyporphins.

Glycosyl diversity. Formation of *C*-glycosides, while less common than *O*- or *N*- or *S*-glycosides, is well established in natural products chemistry[13] albeit unprecedented in tetrapyrrole science outside of the examples provided by tolyporphins. Inspection of the sugar structures in

Fig. 9 shows that each has the D-configuration, is a pyranose, contains the 4′-axial hydroxy group characteristic of galactose, and lacks a 6′-hydroxy group. Variation accrues at the 3′-position (equatorial OH, axial OH, or no OH) and the 2′-position (OH or *O*-acetyl). The specific name for 3′,6′-dideoxy-D-galactose is abequose (**27**); 2′-*O*-acetyl-3′,6′-dideoxy-D-galactose is 2′-*O*-acetylabequose (**28**); 6′-deoxy-D-galactose is D-fucose (**29**); and the 6′-deoxy sugar with axial hydroxyl groups at both the 3′- and 4′-positions is 6′-deoxygulose, also known as antiarose (**30**), where gulose is a 3′-epimer of galactose. Such glycosyl units are found elsewhere in natural products:

- 6′-Deoxysugars are found in diverse bacteria.[29] The 6′-deoxy sugar D-fucose is found in a sizeable number of bacterial natural products of diverse composition,[13] as well as in polysaccharides of plant[30] and bacterial origin.[31-33]
- Antiarose as a glycosyl motif appears to be exceptionally rare among *bacterial* natural products, given only a single example (an *N*-glycoside of the bis-indole tjipanazole) in a comprehensive survey of several thousand glycosylated bacterial natural products.[13] In contrast, some 50 cardiac *O*-antiarosides are known from the bark of plants.[34,35]
- 3′,6′-Dideoxysugars such as abequose[36] are rarely found in nature except as a constituent of the glycan polymer attached to the cell wall of certain Gram-negative bacteria (*Salmonella enterica*, *Yersinia pseudotuberculosis*).[37,38] The 2′-*O*-acetylabequose in such polysaccharides also is known and engenders distinct serotypes.[39,40] The acetylation can be incomplete due to partial acetylation of the abequose unit,[41] or partial deacetylation[42] of the 2′-*O*-acetylabequose unit. The introduction of *O*-acetyl groups is a biological strategy for late-stage diversification of intact glycosides.[43] The presence of abequose and 2′-*O*-

acetylabequose units in tolyporphins appears to be the only natural examples known outside of the aforementioned cell-surface glycan polymers.

Beyond the issue of the diverse sugars, eight tolyporphins (A–D, L–O) are unusual in bearing two *C*-glycosides. Bis(*C*-glycosylated) compounds are known in plants and microbes, but are exceptionally uncommon,[44] no doubt a reflection of the limited probability of the presence of appropriate substitution sites and twice occurrence of an event that alone is rare, namely *C*-glycosylation. The distinct pyrroline substituents in each tolyporphin member (A–R) are listed in Table 1.

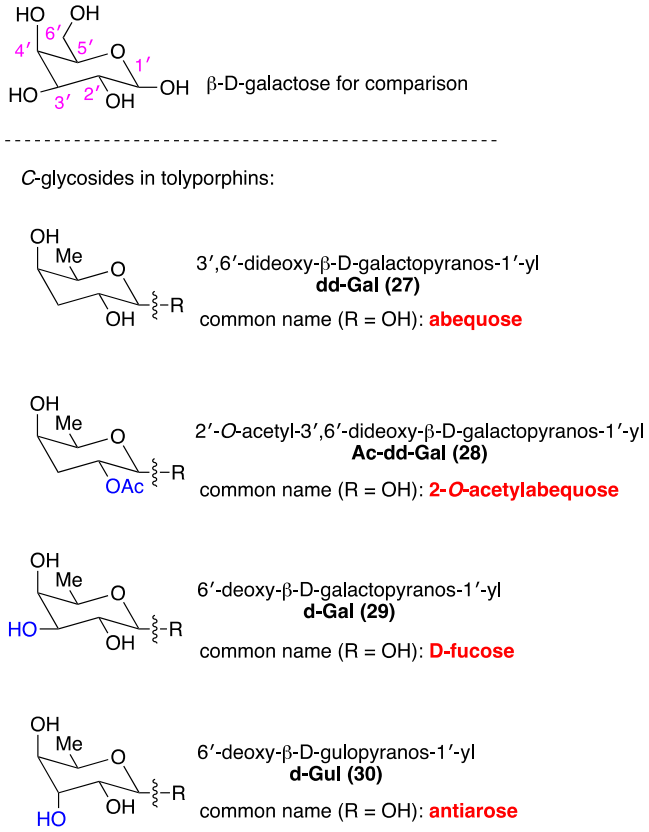


Fig. 9 Glycoside diversity in tolyporphins.

Table 1 Pyrroline substituents in tolyporphins

Substituent[a]	Tolyporphin where the

Abbreviation	Common name	substituent is found
Ac-dd-Gal	2′-*O*-acetylabequose	A,[a] B, C, E, F, L, M, N, O
dd-Gal	abequose	B, C, D,[a] K
d-Gal	D-fucose	N, O
d-Gul	antiarose	L, M
AcO	acetoxy	E, G, H, I,[a] R
HO	hydroxyl	F, G, H, J,[a] Q

[a]The substituent appears twice in the tolyporphin.

Three distinct models for diversification of the pyrroline substituents are displayed in Fig. 10. The six types of products are displayed in blue in each model, where "pre-T" stands for a precursor to the tolyporphin. The three models are presented as distinct processes, although combinations thereof are possible.

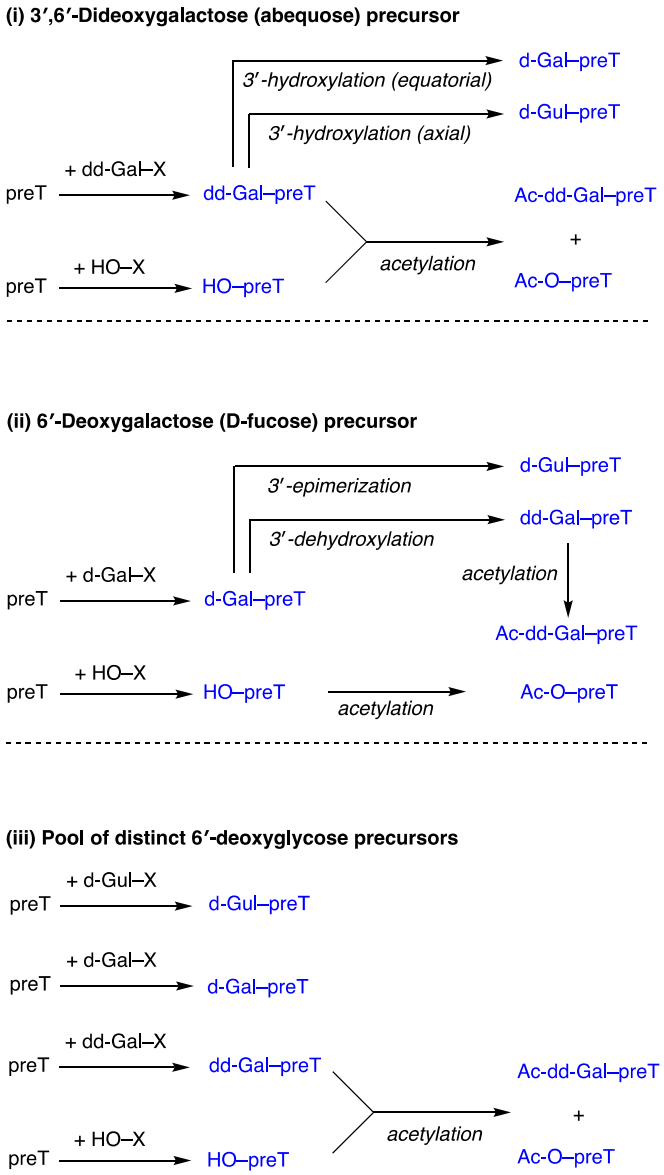


Fig. 10 Models for the origin of molecular diversity of pyrroline substituents.

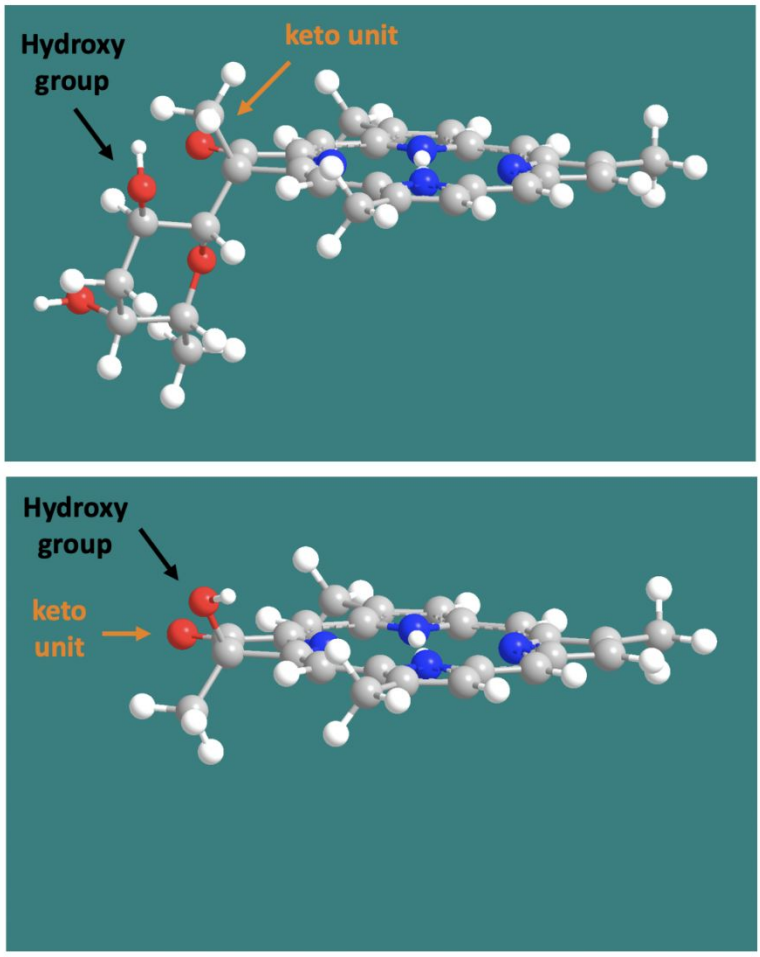
In model (i), 3′,6′-dideoxygalactose (abequose) is installed via dd-Gal–X at the nascent tolyporphin 3-position. In parallel, a hydroxy group also can be installed via HO–X. Both of the resulting products undergo acetylation. The acetylation occurs at the 2′-position of the glycosyl moiety or at the pyrroline hydroxy group, both of which are in a similar location with respect to the tetrapyrrole macrocycle (*vide infra*). In this manner, four distinct substituents are created. The two remaining types require hydroxylation at the 3′-position of the 3′,6′-dideoxygalactosyl unit,

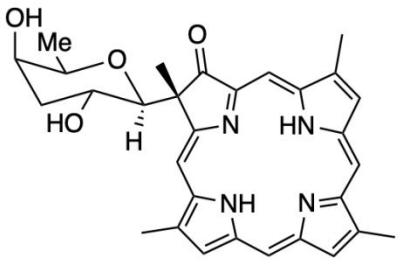
either in an equatorial manner (affording the 6′-deoxygalactose, D-fucose) or an axial manner (affording the 3′-epimer of galactose, the 6′-deoxygulose, antiarose).

In model (ii), the process begins with 6′-deoxygalactose, which upon installation undergoes 3′-epimerization to give the 6′-deoxygulosyl unit (antiarose), or 3′-dehydroxylation (3′-hydro-3′-dehydroxylation to be exact) followed by optional acetylation to give the 3′,6′-dideoxygalactosyl substituents abequose and 2′-*O*-acetylabequose. Installation of a hydroxy group proceeds as in model (i) followed optionally by acetylation.

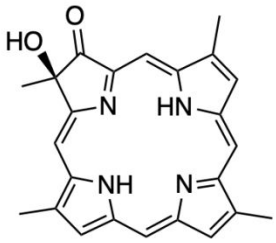
In model (iii), a pool of pre-formed glycosyl units is attached, after which acetylation occurs on the 3′,6′-dideoxygalactosyl (abequose) unit. This process entails limited enzymatic processing of the glycosyl motif following attachment to the tetrapyrrole ligand, but requires an extant, full complement of reactive sugar moieties. Installation of a hydroxy group again proceeds as in model (i) followed optionally by acetylation.

The process of *O*-acetylation appears in each of the models shown in Fig. 10. The structural model for *O*-acetylation of the 2′-hydroxyglycosyl groups as envisaged for tolyporphin A is shown in Fig. 11. Only one fully decorated *C*-glycosyl-oxo-pyrroline motif is shown for clarity. The molecular model of an analogous tolyporphin bearing a hydroxy group at the pyrroline β-position also is displayed; this structure is tolyporphin R. The 2′-hydroxy group of the glycosyl moiety (nascent tolyporphin A) and the hydroxy group at the pyrroline position (tolyporphin R) occupy similar positions with respect to the tetrapyrrole macrocycle. The proposition here is that a promiscuous acetyltransferase could acetylate both hydroxy groups, and if incomplete, would resemble the *O*-acetylation of polysaccharides, including abequose at the 2′-position, in the outer membrane of Gram-negative bacteria.[41] In this manner, partial *O*-acetylation can give rise to a considerable extent of the observed diversity of tolyporphins.





Tolyporphin A analogue



Tolyporphin R

Fig. 11 Molecular models showing hydroxy groups prior to *O*-acetylation; the keto group is indicated for visual orientation. Top panel: oxopyrroline and unacetylated *C*-glycosyl units are shown for a simplified analogue of tolyporphin A. Bottom panel: tolyporphin R.

Tolyporphins interrelated by *O*-acetylation of a *C*-glycosyl unit (abequose) include the following: D → B,C → A. Tolyporphins interrelated by *O*-acetylation of a hydroxyl group include the following: F → E; J → G,H → I; Q → R. (Such relationships assume congruent stereochemistry of the corresponding groups, which has not yet been established.) Thus, if tolyporphins A, E, I and R are regarded as fully *O*-acetylated, then eight additional tolyporphins with incomplete or no *O*-acetylation reside in the same respective structural lineage. Tolyporphins L and M are isomers and contain one 2′-*O*-acetylabequose and one antiarose, whereas tolyporphins

N and O are isomers and contain one 2′-*O*-acetylabequose and one D-fucose. As abequose is the only sugar in the tolyporphins family that is present as an *O*-acetyl entity, tolyporphins L–O are regarded as fully *O*-acetylated. In other words, biosynthesis of four core tolyporphins accompanied by incomplete *O*-acetylation affords 2/3 of the tolyporphins: set A–D; set E,F; set G–I; and set Q,R. The remaining 1/3 of the tolyporphins are K, P, isomers L,M, and isomers N,O.

Other routes for glycosyl diversification can be envisaged as well, such as direct installation of the 2′-*O*-acetyl-3′,6′-dideoxygalactose unit and the acetoxy group, rather than *O*-acetylation after installation of the hydroxyl species. If so, the diversity observed could stem from partial deacetylation via the interrelationships outlined above. Also, 6′-dehydroxylation could be achieved globally following installation of the 6′-hydroxy glycosyl motifs, but the absence to date of any tolyporphins with "normal" 6′-hydroxyglycosyl motifs, yet the presence otherwise of considerable molecular diversity, would tend to argue against such a process. Each of the proposed routes likely originates from enzymatic promiscuity in the following manner: (i) an enzyme installs a 3′-hydroxy group in an equatorial or axial position; (ii) an enzyme causes epimerization or dehydroxylation of the 3′-hydroxy group; (iii) an enzyme installs diverse sugars on the macrocycle. Enzymes of natural products biosynthesis are known to be promiscuous (including *C*-glycosyltransferases[45]), affording a source of molecular diversity.[46] An alternative view is that there exists a constellation of enzymes, each with high specificity, operating together to give the observed collection of tolyporphins.

Outlook

Tolyporphins are members of a new class of tetrapyrrole macrocycles. The observed molecular diversity suggests secondary metabolites, with roles as yet undefined. The 18 tolyporphins known to date present three distinct chromophores (dioxobacteriochlorin, oxochlorin,

and porphyrin), six distinct substituents in the pyrrole ring (hydroxy, acetoxy, and four *C*-glycosides), and variable stereochemistry of the latter substituents among the three compounds where definitive knowledge is available. As natural products, selected members of the tolyporphins family present the rare cases of (1) bis(*C*-glycosylation), (2) incorporation of abequose (otherwise found in the glycan polymer of certain Gram-negative bacteria), and (3) incorporation of antiarose (chiefly found in cardiac antiarosides derived from plants). Studies of the biosynthesis of heme d_1, which exhibits a subset of the molecular structural issues found in the tolyporphins family, have occupied talented scientific teams for some years.[23-25] Delineating the biosynthetic pathway(s) to tolyporphins also is expected to be a long road even with access to substantially complete genomic data.[15-17] The possibility must not be given short shrift that tolyporphins could arise from pathways completely unrelated to those that flow through the otherwise universal and fecund uroporphyrinogen III, but if so the congruence of substituent patterns with all other natural tetrapyrrole macrocycles would stand as a remarkably beguiling biomolecular signature. Regardless, consideration of known biosynthetic pathways to tetrapyrroles, enzymatic processes in natural products chemistry, and organic chemistry fundamentals affords a conceptual framework for the origin of the repertoire of tolyporphins. The present work may enable a focused search for enzymes for the biosynthesis of tolyporphins, and in so doing facilitate gaining a deeper understanding of this novel class of tetrapyrrole macrocycles.

Appendix

Pyrrole (**A1**) is a π-excessive heterocycle,[47] given the presence of six π-electrons shared over a five-atom framework, with electrophilic aromatic substitution (S_EAr) a common reaction. The standard S_EAr process entails electrophilic attack to give a tetrahedral center (**A2**) followed

by ipso-proton loss; the carbon atom at the site of substitution undergoes conversion from the initial sp^2 of pyrrole to sp^3 of the intermediate followed by regeneration of the unsaturated, sp^2 hybridization (**A3**) (Fig. A1, top). Porphyrinogens (**A4**, and analogous tetrahydroporphyrins and dihydroporphyrins) have another reaction pathway available, however, that begins in the same manner as S_EAr with electrophile addition (**A5**), but entails loss of the proton on the adjacent meso-methylene rather than loss of the ipso proton (Fig. A1, bottom). This process remains an example of electrophilic substitution, but the site of electrophile addition becomes saturated (sp^3 hybridization) and the adjacent meso carbon becomes unsaturated (**A6**). In this context, the reaction profile of the pyrrole more resembles that of alkylation of an enamine. The "cryptic enamine" in pyrrole is illustrated in Fig. A1.

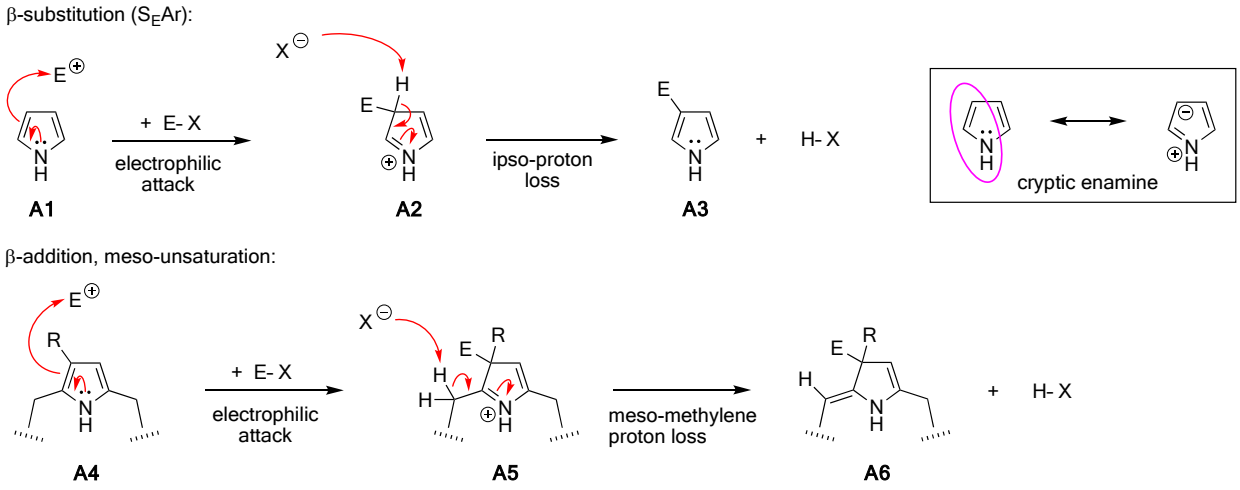


Fig. A1 Standard S_EAr of pyrrole (top), altered process in a β-substituted porphyrinogen (bottom), and pyrrole as a cryptic enamine (upper insert).

Author Information

Corresponding Author

E-mail: jlindsey@ncsu.edu. Tel: 919-515-6406.

Acknowledgment

This work was supported by the NSF (CHE-1760839). Mr. Khiem Chau Nguyen performed the molecular modeling experiments.

Notes: The author declares no competing financial interests.

References

1 M. R. Prinsep, F. R. Caplan, R. E. Moore, G. M. L. Patterson and C. D. Smith, *J. Am. Chem. Soc.,* 1992, **114**, 385–387.

2 G. M. L. Patterson, C. L. Baldwin, C. M. Bolis, F. R. Caplan, H. Karuso, L. K. Larsen, I. A. Levine, R. E. Moore, C. S. Nelson, K. D. Tschappat, G. D. Tuang, E. Furusawa, S. Furusawa, T. R. Norton and R. B. Raybourne, *J. Phycol.,* 1991, **27**, 530–536.

3 D. Hood, D. M. Niedzwiedzki, R. Zhang, Y. Zhang, J. Dai, E. S. Miller, D. F. Bocian, P. G. Williams, J. S. Lindsey and D. Holten, *Photochem. Photobiol.,* 2017, **93**, 1204–1215.

4 C. Brückner, *Photochem. Photobiol.,* 2017, **93**, 1320–1325.

5 M. R. Prinsep, G. M. L. Patterson, L. K. Larsen and C. D. Smith, *Tetrahedron,* 1995, **51**, 10523–10530.

6 M. R. Prinsep, G. M. L. Patterson, L. K. Larsen and C. D. Smith, *J. Nat. Prod.,* 1998, **61**, 1133–1136.

7 J. R. Gurr, J. Dai, C. S. Philbin, H. T. Sartain, T. J. O'Donnell, W. Y. Yoshida, A. L. Rheingold and P. G. Williams, *J. Org. Chem*., 2020, **85**, 318–326.

8 Y. Zhang, R. Zhang, R.-A. Hughes, J. Dai, J. R. Gurr, P. G. Williams, E. S. Miller and J. S. Lindsey, *Phytochem. Anal.,* 2018, **29**, 205–216.

9 J. A. S. Cavaleiro, J. P. C. Tomé and M. A. F. Faustino, *Top. Heterocycl. Chem.,* 2007, **7**, 179–248.

10 J. Puddick and M. R. Prinsep, *Chem. New Zealand,* 2008, April, 68–71.

11 K. Tidgewell, B. R. Clark and W. H. Gerwick, In *Comprehensive Natural Products II Chemistry and Biology*, L. Mander and H.-W. Lui, Eds.; 2010, Vol. 2, Elsevier: Oxford, pp 141–188.

12 R. K. Singh, S. P. Tiwari, A. K. Rai and T. M. Mohapatra, *J. Antibiotics,* 2011, **64**, 401–412.

13 S. I. Elshahawi, K. A. Shaaban, M. K. Kharel and J. S. Thorson, *Chem. Soc. Rev.,* 2015, **44**, 7591–7697.

14 D. Mauzerall, *J. Am. Chem. Soc.,* 1960, **82**, 2601–2605.

15 R.-A. Hughes, Y. Zhang, R. Zhang, P. G. Williams, J. S. Lindsey and E. S. Miller, *Appl. Environ. Microbiol.,* 2017, **83**, e01068-17.

16 X. Jin, E. S. Miller and J. S. Lindsey, *Life,* 2021, **11**, 356.

17 R.-A. Hughes, X. Jin, Y. Zhang, R. Zhang, S. Tran, P. G. Williams, J. S. Lindsey and E. S. Miller, *Microbiol.,* 2018, **164**,1229–1239.

18 M. R. Prinsep, R. A. Thomson, M. L. West and B. L. Wylie, *J. Nat. Prod.,* 1996, **59**, 786–788.

19 J. R. Gurr, T. J. O'Donnell, Y. Luo, W. Y. Yoshida, M. L. Hall, A. M. S. Mayer, R. Sun and P. G. Williams, *J. Nat. Prod.,* 2020, **83**, 1691–1695.

20 T. G. Minehan and Y. Kishi, *Angew. Chem. Int. Ed.,* 1999, **38**, 923–925.

21 T. G. Minehan, L. Cook-Blumberg, Y. Kishi, M. R. Prinsep and R. E. Moore, *Angew. Chem. Int. Ed.,* 1999, **38**, 926–928.

22 W. Wang and Y. Kishi, *Org. Lett.,* 1999, **1**, 1129–1132.

23 B. Li and J. Bridwell-Rabb, *Biochemistry,* 2019, **58**, 85–93.

24 L. Boss, R. Oehme, S. Billig, C. Birkemeyer and G. Layer, *FEBS J*., 2017, **284**, 4314–4327.

25 T. Klünemann, M. Nimtz, L. Jänsch, G. Layer and W. Blankenfeldt, *FEBS J.,* 2021, **288**, 244–261.

26 A. Gutmann and B. Nidetzky, *Pure Appl. Chem.,* 2013, **85**, 1865–1877.

27 G. Tegl and B. Nidetzky, *Biochem. Soc. Trans*., 2020, **48**, 1583–1598.

28 D. A. Bryant, C. N. Hunter and M. J. Warren, *J. Biol. Chem*., 2020, **295**, 6888–6925.

29 M. Graber, A. Morin, F. Duchiron and P. F. Monsan, *Enzyme Microb. Technol.,* 1988, **10**, 198–206.

30 B.-Y. Fan, Y. He, Y. Lu, M. Yang, Q. Zhu, G.-T. Chen and J.-L. Li, *J. Nat. Prod.,* 2019, **82**, 1593–1598.

31 M. Pac, I. Komaniecka, K. Zamlynska, A. Turska-Szewczuk and A. Choma, *Carbohydr. Res.,* 2015, **413**, 37–40.

32 A. V. Perepelov, S. N. Senchenkova, N. A. Kalinchuk, A. S. Shashkov and Yu. A. Knirel, *Russ. Chem. Bull., Int. Ed*., 2018, **67**, 1931–1933.

33 I. Speciale, F. Di Lorenzo, V. Gargiulo, G. Erbs, M.-A. Newman, A. Molinaro and C. De Castro, *Angew. Chem. Int. Ed*., 2020, **59**, 6368–6374.

34 L.-S. Shi, S.-C. Kuo, H.-D. Sun, S. L. Morris-Natschke, K.-H. Lee and T.-S. Wu, *Bioorg. Med. Chem.,* 2014, **22**, 1889–1898.

35 Y. Yan, J. Yang, L. Wang, D. Xu, Z. Yu, X. Guo, G. P. Horsman, S. Lin, M. Tao and S.-X. Huang, *Chem. Sci.,* 2020, **11**, 3959–3964.

36 X. M. He and H.-W. Liu, *Annu. Rev. Biochem*., 2002, **71**, 701–754.

37 M. Hobbs and P. R. Reeves, *Biochim. Biophys. Acta,* 1995, **1245**, 273–277.

38 F. Micoli, N. Ravenscroft, P. Cescutti, G. Stefanetti, S. Londero, S. Rondini and C. A. MacLennan, *Carbohydr. Res.,* 2014, **385**, 1–8.

39 J. M. Slauch, M. J. Mahan, P. Michetti, M. R. Neutra and J. J. Mekalanos, *Infect. Immun.,* 1995, **63**, 437–441.

40 J. Ståhle and G. Widmalm, *Trends Glycosci. Glycotechnol.,* 2019, **31**, E159–E171.

41 C. De Castro, R. Lanzetta, S. Leone, M. Parrilli and A. Molinaro, *Carbohydr. Res.,* 2013, **370**, 9–12.

42 K. Ilg, G. Zandomeneghi, G. Rugarabamu, B. H. Meier and M. Aebi, *Carbohydr. Res.,* 2013, **382**, 58–64.

43 G. Samuel and P. Reeves, *Carbohydr. Res.,* 2003, **338**, 2503–2519.

44 D. Chen, S. Fan, R. Chen, K. Xie, S. Yin, L. Sun, J. Liu, L. Yang, J. Kong, Z. Yang and J. Dai, *ACS Catal*., 2018, **8**, 4917–4927.

45 J.-B. He, P. Zhao, Z.-M. Hu, S. Liu, Y. Kuang, M. Zhang, B. Li, C.-H. Yun, X. Qiao and M. Ye, *Angew. Chem. Int. Ed*., 2019, **58**, 11513–11520.

46 R. D. Firn and C. G. Jones, *Nat. Prod. Rep.,* 2003, **20**, 382–391.

47 A. H. Jackson, In *Pyrroles. Part One. The Synthesis and the Physical and Chemical Aspects of the Pyrrole Ring;* R. A. Jones, Ed., John Wiley & Sons, Inc.: New York, 1990, Vol. 1, pp 295−303.