Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Multi-domain terpenoid cyclase architecture and prospects for proximity in bifunctional catalysis

Mengbin Chen[1], Golda G Harris[1], Travis A Pemberton[1] and David W Christianson[1,2]

Crystal structures of terpenoid cyclases reveal assemblies of three basic domains designated α, β, and γ. While the biosynthesis of cyclic monoterpenes (C_{10}) and sesquiterpenes (C_{15}) most often involves enzymes with α or αβ domain architecture, the biosynthesis of cyclic diterpenes (C_{20}), sesterterpenes (C_{25}), and triterpenes (C_{30}) can involve enzymes with α, αα, βγ, or αβγ domain architecture. Indeed, some enzymes of terpenoid biosynthesis are bifunctional, with distinct active sites that catalyze sequential reactions. Interestingly, some of these enzymes oligomerize to form dimers, tetramers, and hexamers. Not only can such assemblies enable enzyme regulation by allostery, but they can also provide a modest enhancement of terpenoid product flux through proximity channeling or cluster channeling. The mixing and matching of functional terpenoid cyclase domains through tertiary and/or quaternary structure may also comprise an evolutionary strategy for facile product diversification.

Addresses
[1] Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6323, United States
[2] Radcliffe Institute for Advanced Study and Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, United States

Corresponding author: Christianson, David W (chris@sas.upenn.edu)

Current Opinion in Structural Biology 2016, **41**:27–37

This review comes from a themed issue on **Catalysis and regulation**

Edited by **David W. Christianson** and **Nigel Scrutton**

http://dx.doi.org/10.1016/j.sbi.2016.05.010

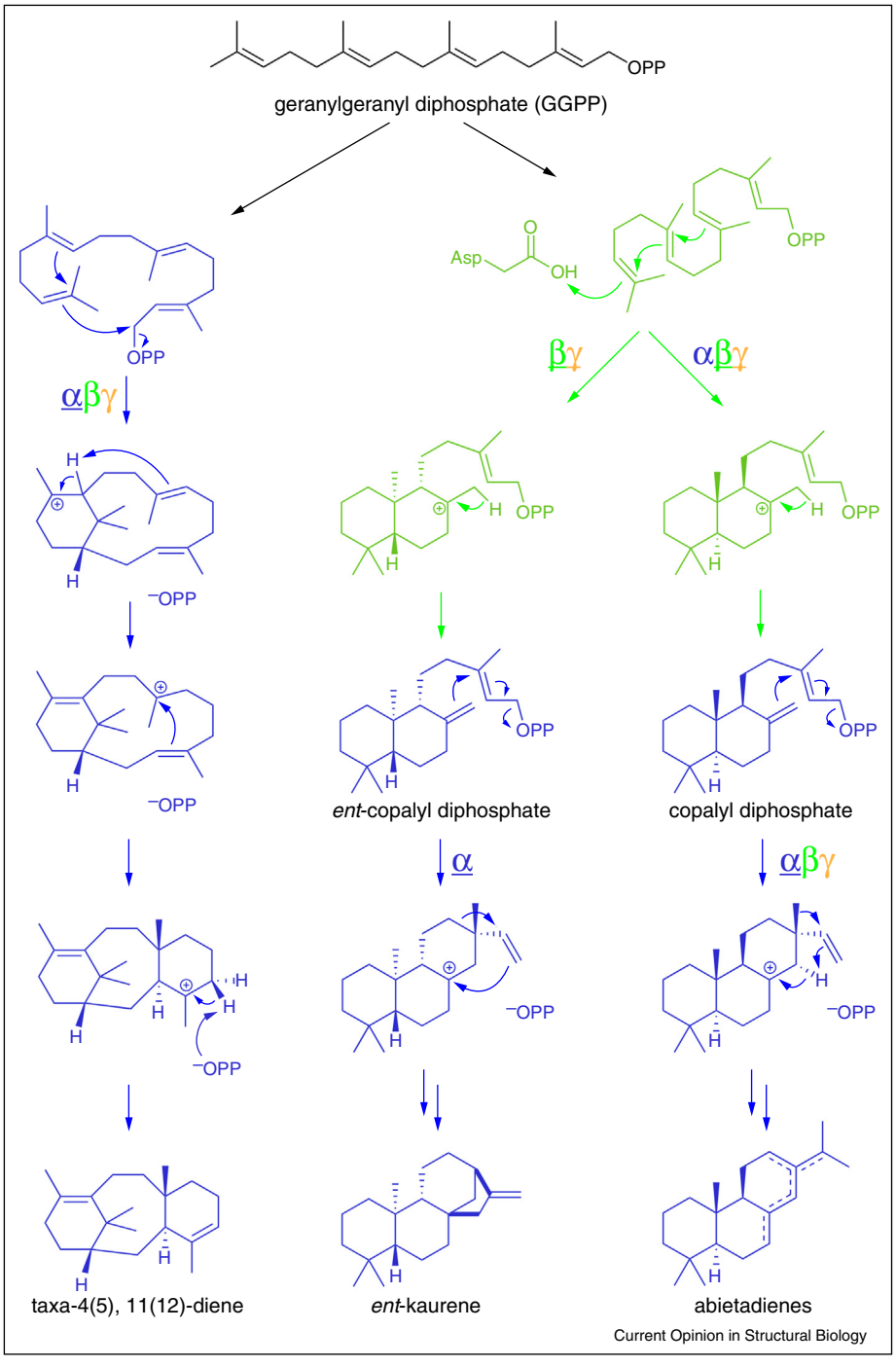
Introduction

Terpenoid cyclases catalyze the formation of thousands of complex natural products containing multiple rings and stereocenters starting from a single achiral precursor such as C_{10} geranyl diphosphate (GPP), C_{15} farnesyl diphosphate (FPP), or C_{20} geranylgeranyl diphosphate (GGPP) [1–3]. From the perspective of mechanistic organic chemistry, these reactions are the most complex in biology, in that on average at least half of the carbon atoms of the substrate undergo changes in bonding and/or hybridization during the course of a single enzyme-catalyzed reaction [4,5]. For example, consider the cyclization of GGPP catalyzed by taxadiene synthase in the first committed step of paclitaxel (Taxol) biosynthesis (Figure 1) [6–8]: a minimum of 10 carbon atoms in the C_{20} substrate are involved in multiple bond-making or bond-breaking steps. While terpenoid cyclases are not exceptionally efficient catalysts (e.g., for taxadiene synthase, $k_{cat}/K_M = 3.5 \times 10^3\ M^{-1}\ s^{-1}$) [9], these enzymes are nonetheless impressive in that the rate enhancement in comparison with the uncatalyzed rate is immeasurably large. The cyclization cascades illustrated in Figure 1 simply cannot occur in the absence of an enzyme catalyst.

A terpenoid cyclase serves a critical role as a template for catalysis. The polyisoprenoid substrate is quite flexible — with three bond rotations per 5-carbon isoprenoid unit, GGPP is more flexible than a glycine polypeptide, which only has two bond rotations per glycine. Moreover, since a polyisoprenoid contains no polar residues capable of hydrogen bond interactions, the only way to hold the substrate in a cyclization-competent conformation is for the enzyme to provide a snug-fitting template. In essence, the specific cyclization products generated by a terpenoid cyclase are encoded in the shape of the active site contour.

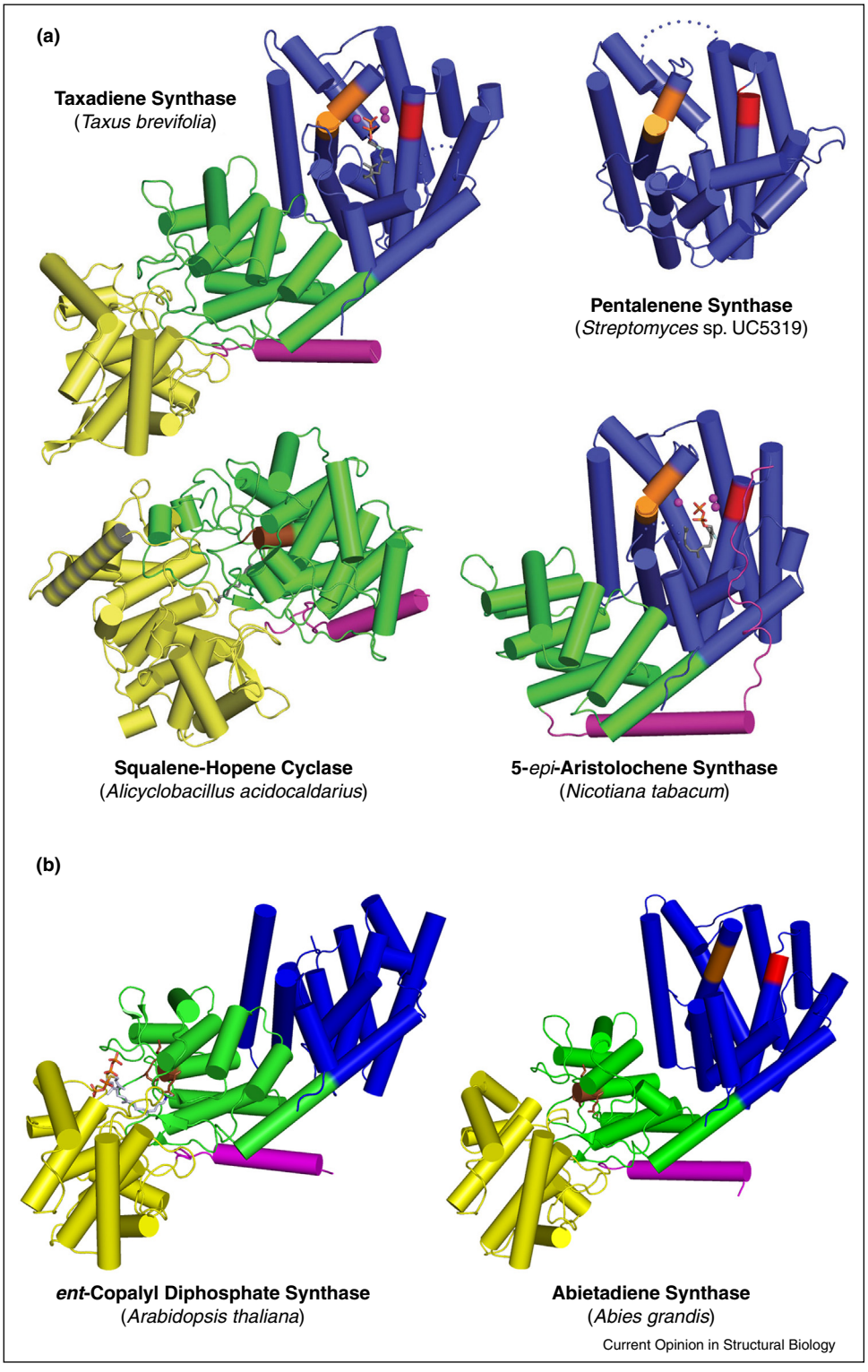
A well-defined template for catalysis is a common feature of both class I and class II terpenoid cyclases. The first crystal structures of class I and class II terpenoid cyclases were reported nearly 20 years ago and revealed distinctive protein folds for each class of enzyme (Figure 2a) [10–12]. For example, the α-helical fold observed for the class I bacterial cyclase pentalenene synthase (designated the α domain fold [13]) is similar to that of 5-*epi*-aristolochene synthase, although the latter enzyme contains an additional domain so as to adopt overall αβ domain architecture. In contrast, the βγ architecture of the class II cyclase squalene-hopene cyclase likely arose from gene duplication and fusion due to structural similarities between the two domains (note that α, β, and γ domains each adopt α-helical folds) [14]. The β-domain of βγ squalene-hopene cyclase is homologous to the β-domain of αβ 5-*epi*-aristolochene synthase [14], suggesting the possibility of linked αβγ domain architectures [13,15]. The first X-ray crystal structure of a terpenoid cyclase with αβγ domain architecture was that of taxadiene synthase, a diterpene cyclase (Figure 2a) [16••]. Similar αβγ domain architecture was subsequently observed in the diterpene cyclases

Figure 1



GGPP cyclization reactions catalyzed by diterpene synthases. Class I cyclization reactions are catalyzed by α domains and class II cyclization reactions are catalyzed at the βγ domain interface. Reactions can occur exclusively in a single active site of an αβγ cyclase, as exemplified by the taxadiene synthase reaction. Alternatively, tandem class II/class I reactions can occur in a bifunctional αβγ cyclase, as exemplified by abietadiene synthase, or in separate α and βγ cyclases (that may nonetheless associate), as for some *ent*-kaurene synthases. Domain architecture for each enzyme is indicated (α, β, γ) and catalytic domains for each reaction are underlined; class I and class II cyclization reactions are blue and green, respectively.

Figure 2



Domain architecture in terpenoid cyclases. **(a)** The α, β, and γ domains are blue, green, and yellow, respectively. The active site of a class I cyclase resides in an α domain flanked by signature metal-binding motifs DDXXD and (N,D)DXX(S,T)XXXE (red and orange, respectively). Class I cyclases are found as single α domains, αβ assemblies, and αβγ assemblies as exemplified by pentalenene synthase (PDB 1PS1), 5-*epi*-aristolochene synthase (PDB 1LZ9), and taxadiene synthase (PDB 3P5R), respectively. The active site of a class II cyclase resides at the interface of β and γ domains as exemplified by squalene-hopene cyclase (PDB 1SQC), where the central aspartic acid in the motif DXDD (brown) serves as a general acid. Class II cyclases are found with βγ and αβγ domain assemblies. The role of N-termini (purple) in class I plant cyclases is to help

ent-copalyl diphosphate synthase [17] and abietadiene synthase (Figure 2b) [18••], as well as the sesquiterpene cyclase α-bisabolene synthase [19].

In addition to their differences in tertiary structure, class I and class II terpenoid cyclases also differ by the chemical strategy employed for initial carbocation formation. For class I cyclases, an induced-fit conformational change triggered by the binding of three metal ions (typically Mg^{2+}, sometimes Mn^{2+}, and rarely Co^{2+}) and substrate leads to active site closure and substrate ionization to yield an allylic cation plus inorganic pyrophosphate (PP_i) [20]. Metal ions are liganded by two characteristic amino acid sequence motifs: the ''aspartate rich'' sequence **D**DXX**D** and the ''NSE'' sequence (**N,D**)DXX(**S,T**)XXX**E** (boldface indicates metal ligands). A class II cyclase utilizes the central aspartic acid in the sequence DXDD as a general acid to protonate the terminal C=C bond of the polyisoprenoid to generate a tertiary carbocation [21,22]. In each cyclase class, one of the remaining π bonds of the substrate is held close to the incipient carbocation with the proper orientation for cyclization by the active site template.

More than two dozen crystal structures of class I and class II terpenoid cyclases have been determined to date, and these structures provide a foundation for understanding complex chemical mechanisms of isoprenoid cyclization [23]. While the simplest terpenoid cyclase structure is that of a single α domain class I enzyme, such as the bacterial sesquiterpene cyclase pentalenene synthase [10] or the diterpene cyclase CotB [24•], αβ, βγ, and αβγ domain architectures of increasing complexity are observed in both class I and class II plant diterpene cyclases (Figure 2); sometimes, αβγ cyclases are bifunctional enzymes in which tandem cyclization reactions are catalyzed by the class II and class I active sites. Intriguingly, αα domain architectures with $(\alpha\alpha)_n$ assemblies (n = 1, 2, and 6) are also observed for terpenoid synthases, some of which catalyze tandem biosynthetic reactions. The combination of tandem bifunctionality, variable domain architecture, and variable quaternary structure hints of a potential for assembly line catalysis or even a terpenoid biosynthetic metabolon. In the remainder of this review, we outline some recent examples of bifunctional catalysis resulting from the evolution of specific domain architecture in terpenoid cyclases.

Bifunctional αβγ assemblies

Biosynthesis of the labdane diterpene *ent*-kaurene in the bacterium *Bradyrhizobium japonicum* requires two separate class II and class I cyclases, *ent*-copalyl diphosphate synthase and *ent*-kaurene synthase, respectively [15]; the recently-reported crystal structure of *ent*-kaurene synthase exhibits characteristic α domain architecture [25•]. Substrate GGPP initially undergoes cyclization by the class II cyclase to form *ent*-copalyl diphosphate, which is cyclized to form *ent*-kaurene by the class I cyclase (Figure 1). In the plant *Marah macrocarpus*, these enzymes form a complex (presumably assembled from αβγ monomers) believed to preferentially channel intermediate *ent*-copalyl diphosphate from the active site of the class II cyclase to the class I cyclase [26]. In the fungus *Phaeosphaeria* sp. L487, a bifunctional *ent*-copalyl diphosphate synthase/*ent*-kaurene synthase results from the fusion of the class I and class II cyclases in a single αβγ polypeptide chain [27]. Although the herb *Salvia miltiorrhiza* Bunge has separate *ent*-copalyl diphosphate synthase and *ent*-kaurene synthase enzymes, the engineering of a fusion construct yields a bifunctional αβ–αβγ enzyme that exhibits ~4-fold enhanced product flux [28]. Thus, association or fusion of terpenoid cyclase domains catalyzing successive reactions may confer a modest catalytic advantage. Even in the absence of a direct channel between enzyme active sites, such ''proximity channeling'' can enhance product flux [29]. Conceivably, proximity channeling could also be augmented by electrostatic channeling [30] of anionic isoprenoid diphosphate intermediates if the protein surface were significantly cationic.

Single αβγ domain assemblies are also formed in plant diterpene synthases such as taxadiene synthase from *Taxus brevifolia* (Pacific yew), *ent*-copalyl diphosphate synthase from *Arabidopsis thaliana* (mouse-ear cress), and abietadiene synthase from *Abies grandis* (grand fir) (Figure 2) [16••,17,18••]. However, in taxadiene synthase, only the class I active site in the α domain is functional, and in *ent*-copalyl diphosphate synthase, only the class II active site at the βγ domain interface is functional. Strikingly, both active sites are functional in abietadiene synthase (Figure 2b): the class II active site catalyzes the cyclization of GGPP to form *ent*-copalyl diphosphate, which in turn is further cyclized to form abietadiene in the class I active site (Figure 1). In this system, there is no evidence for direct channeling; the transfer of intermediate product *ent*-copalyl diphosphate from the class II active site to the class I active site occurs by diffusive transfer [31,32].

Bifunctional αα and $(\alpha\alpha)_6$ assemblies

In recent years, bifunctional terpenoid cyclases with αα domain architectures have been identified, and these enzymes encompass various combinations of isoprenoid chain elongation, cyclization, and fragmentation chemistry. For example, geosmin is a potent odorant molecule

(Figure 2 Legend Continued) cap the active site in the α domain, as shown for 5-*epi*-aristolochene synthase. Reprinted with permission from Ref. [16••]. **(b)** *ent*-Copalyl diphosphate synthase is a class II terpenoid cyclase with αβγ domain architecture; the active site at the βγ interface is indicated by a bound substrate analogue, and the α domain is vestigial (PDB 3PYA). Abietadiene synthase is a bifunctional class I/class II terpenoid cyclase with fully functional α and βγ domains (PDB 3S9V). Protein structures color-coded as in (a).

that signals *Streptomyces* contamination in public water supplies [33]. The N-terminal α domain of geosmin synthase catalyzes the cyclization of FPP to form germacradienol, which subsequently undergoes a retro-Prins fragmentation in the C-terminal α domain to generate the 12-carbon product geosmin and the 3-carbon product acetone [34,35]. Recently, the crystal structure of the N-terminal α domain and a homology model of the C-terminal α domain were used to fit the *ab initio* molecular envelope calculated from small-angle X-ray scattering (SAXS) data [36•]. These structural studies revealed a monomeric architecture in which the active site of one domain was not oriented directly toward the active site of the second domain. However, the physical proximity of the two active sites of geosmin synthase may provide a modest catalytic advantage through proximity channeling [29]. This has been demonstrated in protein engineering experiments with other sesquiterpene synthases. For example, the fusion of the FPP synthase α domain and the 5-*epi*-aristolochene synthase αβ domains into tandem ααβ or αβα domain protein constructs yields approximately two-fold increases in product flux [37].

A naturally occurring example of an isoprenoid chain elongation domain fused with a cyclase domain that adopts hexameric quaternary structure $(\alpha\alpha)_6$ is found in fusicoccadiene synthase from *Phomopsis amygdali* (PaFS). Fusicoccadiene is a precursor of the diterpene glucoside fusicoccin A, a phytotoxin that also exhibits promising antitumor properties [38,39], and PaFS was the first bifunctional diterpene synthase with αα domain architecture to be discovered [40]. A second bifunctional synthase was later discovered in the same organism [41], and more recent reports describe additional bifunctional diterpene and sesterterpene synthases with presumed αα domain architecture [42–45]. The C-terminal α domain of PaFS catalyzes the condensation of dimethylallyl diphosphate (DMAPP) and three isopentenyl diphosphate (IPP) molecules to generate GGPP, and the N-terminal α domain catalyzes the cyclization of GGPP to form fusicoccadiene (Figure 3a) [40,46]. Each α domain of the αα assembly is a metal-dependent class I terpenoid synthase, with conserved metal-binding motifs that coordinate to three divalent metal ions required for substrate activation and ionization to form the initial substrate carbocation [47••]. The GGPP synthase and GGPP cyclase domains are related by 19% amino acid sequence identity, suggesting gene duplication, fusion, and divergence from a primordial terpenoid synthase.

Studies with site-specific mutants of PaFS in which the GGPP synthase domain or the GGPP cyclase domain have been inactivated reveal an approximately two-fold enhancement of product flux, so the covalent linkage of the two α domains confers a modest catalytic advantage through proximity channeling [29] in the absence of a direct channel between active sites. Since PaFS is a hexamer [47••], this catalytic advantage may be enhanced through an agglomeration effect [48••] herein designated "cluster channeling" (Figure 3b). Such quaternary structural assemblies ensure that multiple active sites for each enzyme are close to each other even in the absence of a direct channel between them.

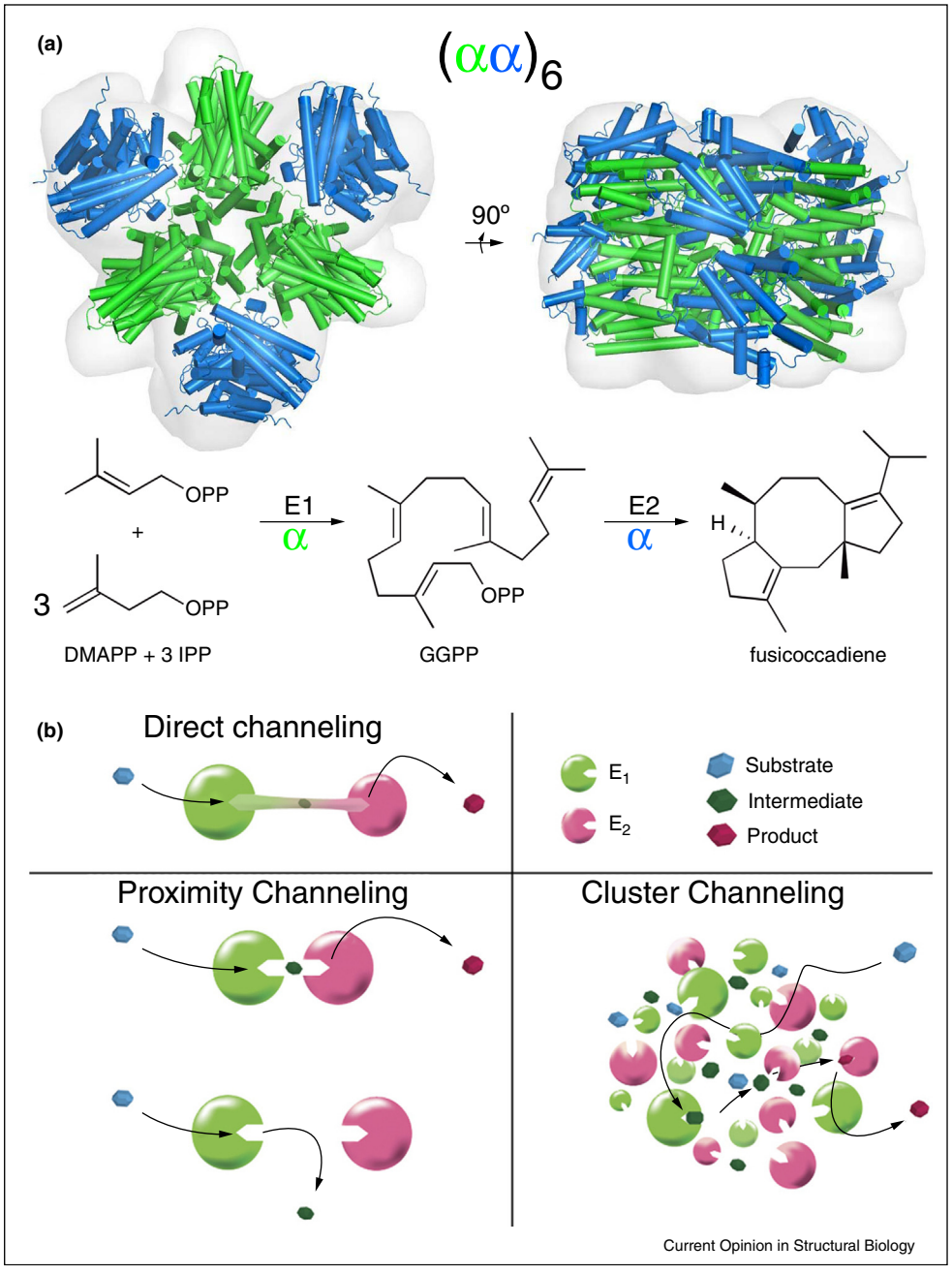
X-ray crystal structures of the individual α domains of PaFS have been determined, and these structures were then used to fit the *ab initio* molecular envelope of the PaFS hexamer generated from SAXS data (Figure 3a) [47••]. The hexameric quaternary structure of the full-length protein appears to be driven by the hexameric quaternary structure of the GGPP synthase domain, although slight structural adjustments were necessary to fit three additional GGPP cyclase dimers into the SAXS envelope. Because PaFS is a hexameric assembly, the active site of a single GGPP synthase domain is close to more than one GGPP cyclase domain. This ensures that the intermediate product GGPP has a higher probability of diffusing into a cyclase active site. Thus, αα domain assembly in PaFS is a paradigm example of proximity and cluster channeling in terpenoid biosynthesis.

Insight gained from the PaFS structure informs our understanding of structure–function relationships in related bifunctional sesterterpene synthases that generate even larger cyclic terpenoids. For example, ophiobolin F synthase from *Aspergillus oryzae* (AcOS) is a sesterterpene cyclase that generates a 25-carbon product sharing a 5-8-5 ring structure identical to that of fusicoccadiene [42]. Pairwise comparison of residues comprising the active site contours of each cyclase reveals that most residues are isosteric or strictly conserved; however, W225 and V228 in PaFS are replaced by smaller residues L217 and A220 in AcOS, respectively. These amino acid substitutions in the AcOS active site serve to accommodate the longer C_{25} sesterterpene substrate [47••].

$(\beta\alpha)_2$ and α_n oligomers

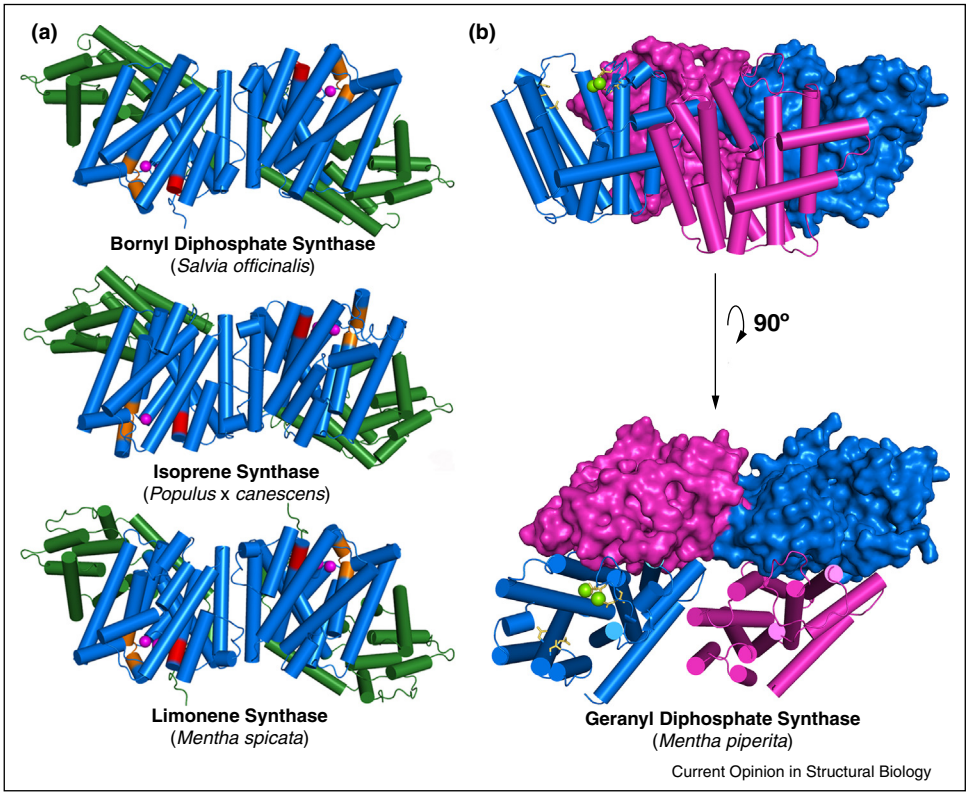
Homodimeric quaternary structure has often been observed for crystalline α or αβ domain terpenoid synthases since the first structure determination of a terpenoid synthase, farnesyl diphosphate synthase [49]. Parallel [49] or antiparallel [50] αα subunit association modes can utilize different regions of the α domain surface [51]. Recurring motifs in dimerization of class I terpene synthases with αβ folds forming βααβ assemblies are evident. For example, the monoterpene cyclase limonene synthase is an αβ monomer in solution as determined by gel filtration chromatography [52], but it forms a dimer in the crystal [53] with quaternary structure identical to that of bornyl diphosphate synthase [54] and isoprene synthase [55] (Figure 4a). Thus, there is potential for weak homodimeric interactions between α subunits, even when not detectable in solution. Perhaps this signals the potential for the assembly of terpenoid biosynthetic

Figure 3



Fusicoccadiene synthase (PaFS). **(a)** The X-ray crystal structures of the individual C-terminal α domain (green, GGPP synthase) and the N-terminal α domain (blue, GGPP cyclase) are fit into the *ab initio* SAXS molecular envelope of the fusicoccadiene hexamer to illustrate a plausible model for the bifunctional enzyme cluster. Adapted with permission from Ref. [47••]. Copyright 2016 American Chemical Society. **(b)** Direct channeling occurs between enzyme active sites E1 and E2 through a protein tunnel that prevents the intermediate from diffusing into solution. Proximity channeling occurs when enzyme active sites E1 and E2 are positioned sufficiently close so that the intermediate is processed by E2 before it can diffuse into solution. If E1 and E2 are not appropriately positioned, the intermediate will diffuse into solution. Cluster channeling occurs in a multienzyme or multisubunit cluster such that once the intermediate is generated by enzyme active site E1, even though the probability of the intermediate being processed by any particular E2 active site might be low, the probability that the intermediate will be utilized by one of the several E2 enzymes in the cluster is high.
Source: Adapted from Ref. [48••]; reprinted by permission from Macmillan Publishers Ltd, copyright 2014.

Figure 4



Subunit assembly in terpenoid synthase oligomers. **(a)** Crystal structures of bornyl diphosphate synthase (PDB 2N20), isoprene synthase (PDB 3N0G), and limonene synthase (PDB 2ONG) reveal identical dimeric quaternary structure. Although limonene synthase exists as a monomer in solution, weak intermonomer interactions are nonetheless sufficient to guide a recurring mode of subunit association in the crystal. **(b)** The crystal structure of geranyl diphosphate synthase from *M. piperita* reveals $\alpha_2\delta_2$ architecture (PDB 3KRA). Since the regulatory δ subunit derives from the α subunit, the quaternary structure approximates α_4 tetramerization.

metabolons if heterodimers are similarly capable of forming. Future studies are of course required to assess this possibility.

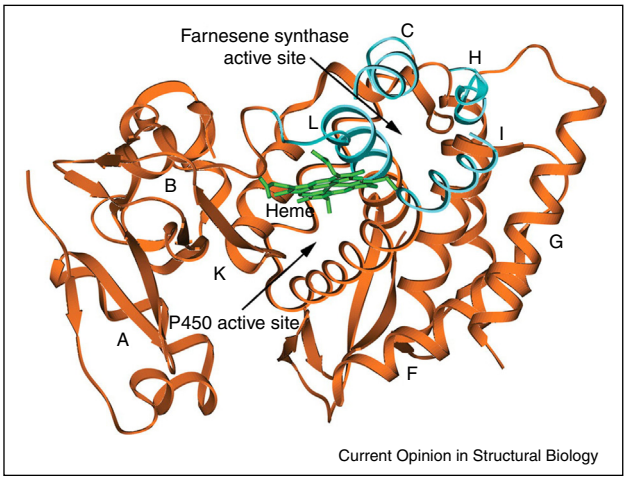
In addition to dimeric α_2 quaternary structures, α_4, and α_6 quaternary structures are also observed in terpenoid synthases. For example, the recently determined crystal structure of the C-terminal GGPP synthase domain of PaFS [47••] reveals a hexameric α_6 quaternary structure similar to that first observed for human GGPP synthase [56]. An interesting example of tetrameric α_4 quaternary structure is found in GPP synthase from *Mentha piperita*, which is actually a heterotetramer comprised of two large α subunits and two small α′ subunits [57]. The large subunit is the catalytic α domain subunit and contains two canonical DDXXD metal-binding motifs. The small subunit lacks metal-binding motifs and adopts a truncated α domain structure — it is accordingly designated as a δ subunit [58•], but structural homology suggests that δ derives from α so δ = α′. One α and one α′ subunit form a tight dimer, and a heterotetramer is formed as a dimer of dimers with overall $\alpha_2\alpha'_2$ stoichiometry (Figure 4b). Individually, both subunits are catalytically inactive; however, when the heterotetramer is formed, the α subunit generates GPP from DMAPP and IPP. The δ subunit acts as a regulatory subunit, stabilizing the α subunit with the proper active site architecture for catalysis. Thus, the association of multiple α domains through tertiary and/or quaternary structure not only enables channeling by proximity or clustering to enhance catalysis, but it can also enable the regulation of enzyme activity.

Bifunctional moonlighting synthases

Moonlighting enzymes are those that, in addition to serving a primary catalytic function, also serve a secondary catalytic function that can be quite distinct from their primary catalytic function [59]. Moonlighting enzymes thus contain features of two separate active sites inserted or overlaid on a single protein fold. To date, two examples of terpene synthase activity moonlighting on cytochrome P450 enzymes have been reported.

Albaflavenone synthase is a cytochrome P450 monooxygenase from *Streptomyces coelicolor* that converts the sesquiterpene *epi*-isozizaene to the antibiotic albaflavenone [60]. This enzyme also exhibits terpene synthase

Figure 5



Moonlighting terpenoid biosynthetic functions. The crystal structure of albaflavenone synthase reveals a cytochrome P450 active site that catalyzes the monooxygenase chemistry required for the conversion of the sesquiterpene *epi*-isozizaene to an epimeric mixture of albaflavenols, and then to albaflavenone. Additionally, the crystal structure reveals a separate class I-like terpenoid synthase active site bearing the characteristic metal-binding motifs that catalyzes the conversion of farnesyl diphosphate into farnesene.
Source: Reprinted with permission from Ref. [60]. Copyright 2009, American Society for Biochemistry and Molecular Biology.

activity, and uses FPP as a substrate to produce a mixture of farnesene isomers. The amino acid sequence of albaflavenone synthase reveals DDXXD and DTE motifs typical of metal-dependent class I terpene cyclases, and structural comparison with aristolochene synthase reveals conservation of elements of the α domain of a class I terpenoid synthase in a separate active site distinct from the P450 active site (Figure 5). Intriguingly, the pH optima for the two different reactions (pH 7.0–8.2 for monooxygenase activity, and pH 5.5–6.5 for terpene synthase activity) indicate that alternate enzyme activities depend on pH-dependent conformational changes — terpene synthase activity is enhanced when the monooxygenase activity is diminished. While albaflavenone synthase provides an intriguing example of how nature can combine terpenoid synthase catalysis with redox catalysis in a single protein scaffold, it should be noted that the farnesene synthase reaction is not part of the albaflavenone biosynthetic pathway.

A second example of a moonlighting terpenoid synthase is found in a cytochrome P450 from *Penicillium aethiopicum*, VrtK [61]. This enzyme catalyzes the initial oxidation of a pendant geranyl moiety on an anhydrotetracycline-like scaffold to form an allylic cation that ultimately cyclizes to generate a spirobicyclic ring system, converting previridicatumtoxin to viridicatumtoxin. These examples demonstrate that prospects for bifunctional catalysis in terpenoid biosynthesis are not limited to isoprenoid diphosphate generation and cyclization, but also oxidation steps that ultimately enhance terpenoid product chemodiversity.

Concluding remarks

X-ray crystal structures of terpenoid cyclases determined in recent years have revealed increasingly complex assemblies of three basic domain types designated α, β, and γ, and a model for the evolution of domain fusion and function accounts for the appearance of bifunctional αβγ diterpene cyclases such as abietadiene synthase [13,58•]. Intriguingly, the assembly of multiple functional domains in a single polypeptide chain can provide a modest enhancement of product flux through proximity channeling [29]. Although proximity enhancement *in vitro* is generally modest, such enhancement could be augmented *in vivo* in the crowded environment of the cell. Indeed, colocalization of enzymes that catalyze successive steps in a biosynthetic pathway using modular synthetic scaffolds is a proven approach for increasing product flux in metabolic engineering experiments [62].

The assembly of multidomain terpenoid cyclases into dimeric, tetrameric, or hexameric quaternary structures may further facilitate bifunctional catalysis and biosynthetic flux through cluster channeling [48••]. The clustering of enzymes that catalyze sequential biosynthetic reactions enhances the probability of proximity channeling from one enzyme active site to another, and the diterpene cyclase fusicoccadiene synthase forms a hexameric assembly that serves as a paradigm for cluster channeling in terpenoid biosynthesis [47••].

The mixing and matching of functional terpenoid cyclase domains through tertiary and/or quaternary structure may comprise a route toward the enhancement or regulation of product flux, but conceivably it may also comprise a route toward the enhancement of product chemodiversity. The structural and functional relationship between the bifunctional diterpene synthase PaFS and the sesterterpene synthase AcOS demonstrates that only a few amino acid substitutions in the active site are necessary to switch cyclization specificity from C_{20} to C_{25} substrates. The α fold, originally designated the ''terpenoid synthase fold'' [10], is particularly versatile in biological systems, also appearing as a polytopic integral membrane protein that catalyzes the prenylation reaction of ubiquinone biosynthesis [63•].

The structural diversity of terpenoid cyclases directly translates into product chemodiversity. That successive reactions of terpenoid biosynthesis can be catalyzed by domain assembly through tertiary or quaternary structure interactions, or by moonlighting functions superimposed on a common protein scaffold, is perhaps suggestive of simple terpenoid biosynthetic metabolons or assembly

lines. Whether true metabolons can form in the cell, presuming that multiple enzymes are in the same compartment, is yet to be demonstrated. Nevertheless, we have shown here that the assembly of multiple domains for bifunctional catalysis through tertiary or quaternary interactions observed in X-ray crystal structures can provide a basic catalytic advantage *in vitro*. Future studies will determine the effects of this catalytic advantage for terpenoid biosynthesis *in vivo*.

Conflict of interest

None declared.

Acknowledgements

We thank Prof. Reuben Peters for his insightful comments on the manuscript, and we thank the National Institutes of Health for grant GM56838 to DWC in support of research on terpenoid biosynthesis. DWC additionally thanks the Radcliffe Institute for Advanced Study for the Elizabeth S. and Richard M. Cashin Fellowship. GGH was supported by an NIH Structural Biology and Molecular Biophysics Training Grant.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Cane DE: **Enzymatic formation of sesquiterpenes**. *Chem Rev* 1990, **90**:1089-1103.

2. Davis EM, Croteau R: **Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes**. *Top Curr Chem* 2000, **209**:53-95.

3. Zi J, Mafu S, Peters RJ: **To gibberellins and beyond! Surveying the evolution of (di)terpenoid metabolism**. *Annu Rev Plant Biol* 2014, **65**:259-286.

4. Christianson DW: **Structural biology and chemistry of the terpenoid cyclases**. *Chem Rev* 2006, **106**:3412-3442.

5. Christianson DW: **Unearthing the roots of the terpenome**. *Curr Opin Chem Biol* 2008, **12**:141-150.

6. Lin X, Hezari M, Koepp AE, Floss HG, Croteau R: **Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of Taxol biosynthesis in Pacific yew**. *Biochemistry* 1996, **35**:2968-2977.

7. Williams DC, Carroll BJ, Jin Q, Rithner CD, Lenger SR, Floss HG, Coates RM, Williams RM, Croteau R: **Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the taxadiene precursor of Taxol catalyzed by recombinant taxadiene synthase**. *Chem Biol* 2000, **7**:969-977.

8. Jin Y, Williams DC, Croteau R, Coates RM: **Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl diphosphate to 7-fluoroverticillenes**. *J Am Chem Soc* 2005, **127**:7834-7842.

9. Hezari M, Lewis NG, Croteau R: **Purification and characterization of taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of Taxol biosynthesis**. *Arch Biochem Biophys* 1995, **322**:437-444.

10. Lesburg CA, Zhai G, Cane DE, Christianson DW: **Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology**. *Science* 1997, **277**:1820-1824.

11. Starks CM, Back K, Chappell J, Noel JP: **Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase**. *Science* 1997, **277**:1815-1820.

12. Wendt KU, Poralla K, Schulz GE: **Structure and function of a squalene cyclase**. *Science* 1997, **277**:1811-1815.

13. Cao R, Zhang Y, Mann FM, Huang C, Mukkamala D, Hudock MP, Mead ME, Prisic S, Wang K, Lin F-Y *et al.*: **Diterpene cyclases and the nature of the isoprene fold**. *Proteins: Struct Funct Bioinf* 2010, **78**:2417-2432.

14. Wendt KU, Schulz GE: **Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes**. *Structure* 1998, **6**:127-133.

15. Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ: **Gibberellin biosynthesis in bacteria: separate *ent*-copalyl diphosphate and *ent*-kaurene synthases in *Bradyrhizobium japonicum***. *FEBS Lett* 2009, **583**:475-480.

16. •• Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW: **Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis**. *Nature* 2011, **469**:116-120.
This paper reports the X-ray crystal structure of taxadiene synthase from the Pacific yew, which was the first crystal structure of a diterpene cyclase and the first structure of a class I $\alpha\beta\gamma$ terpenoid cyclase.

17. Köksal M, Hu H, Coates RM, Peters RJ, Christianson DW: **Structure and mechanism of the diterpene cyclase *ent*-copalyl diphosphate synthase**. *Nat Chem Biol* 2011, **7**:431-433.

18. •• Zhou K, Gao Y, Hoy JA, Mann FM, Honzatko RB, Peters RJ: **Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis***. *J Biol Chem* 2012, **287**:6840-6850.
This paper reports the X-ray crystal structure of abietadiene synthase from grand fir, which provided the first view of a bifunctional diterpene cyclase with an $\alpha\beta\gamma$ fold.

19. McAndrew RP, Peralta-Yahya PP, DeGiovanni A, Pereira JH, Hadi MZ, Keasling JD, Adams PD: **Structure of a three-domain sesquiterpene synthase: a prospective target for advanced biofuels production**. *Structure* 2011, **19**:1876-1884.

20. Aaron JA, Christianson DW: **Trinuclear metal clusters in catalysis by terpenoid synthases**. *Pure Appl Chem* 2010, **82**:1585-1597.

21. Prisic S, Xu J, Coates RM, Peters RJ: **Probing the role of the DXDD motif in class II diterpene cyclases**. *ChemBioChem* 2007, **8**:869-874.

22. Wendt KU, Schulz GE, Corey EJ, Liu DR: **Enzyme mechanisms for polycyclic triterpene formation**. *Angew Chem Int Ed* 2000, **39**:2812-2833.

23. Gao Y, Honzatko RB, Peters RJ: **Terpenoid synthase structures: a so far incomplete view of complex catalysis**. *Nat Prod Rep* 2012, **29**:1153-1175.

24. • Janke R, Görner C, Hirte M, Brück T, Loll B: **The first structure of a bacterial diterpene cyclase: CotB2**. *Acta Cryst* 2014, **D70**:1528-1537.
This paper reports the crystal structure of a single-α domain diterpene cyclase that generates a tricyclic 5-8-5 ring product. The single-domain structure of this bacterial diterpene cyclase contrasts with the $\alpha\beta\gamma$ domain structure of diterpene cyclases such as taxadiene synthase.

25. • Liu W, Feng X, Zheng Y, Huang C-H, Nakano C, Hoshino T, Bogue S, Ko T-P, Chen C-C, Cui Y *et al.*: **Structure, function and inhibition of *ent*-kaurene synthase from *Bradyrhizobium japonicum***. *Sci Rep* 2014, **4**:6214.
This paper reports the crystal structure of a single-α domain bacterial diterpene cyclase that catalyzes the cyclization of *ent*-copalyl diphosphate to form *ent*-kaurene. This enzyme works in tandem with bacterial *ent*-copalyl diphosphate synthase, a class II cyclase with presumed $\beta\gamma$ domain architecture, which generates *ent*-copalyl diphosphate from GGPP.

26. Duncan JD, West CA: **Properties of kaurene synthetase from *Marah macrocarpus* endosperm: evidence for the participation of separate but interacting enzymes**. *Plant Physiol* 1981, **68**:1128-1134.

27. Kawaide H, Sassa T, Kamiya Y: **Functional analysis of the two interacting cyclase domains in *ent*-kaurene synthase from the fungus *Phaeosphaeria* sp. L487 and a comparison with cyclases from higher plants**. *J Biol Chem* 2000, **275**:2276-2280.

28. Su P, Tong Y, Cheng Q, Hu Y, Zhang M, Yang J, Teng Z, Gao W, Huang L: **Functional characterization of *ent*-copalyl diphosphate synthase, kaurene synthase and kaurene**

oxidase in the *Salvia miltiorrhiza* gibberellin biosynthetic pathway. *Sci Rep* 2016, **6**:23057.

29. Bauler P, Huber G, Leyh T, McCammon JA: **Channeling by proximity: the catalytic advantages of active site colocalization using Brownian dynamics**. *J Phys Chem Lett* 2010, **1**:1332-1335.

30. Miles EW, Rhee S, Davies DR: **The molecular basis of substrate channeling**. *J Biol Chem* 1999, **274**:12193-12196.

31. Peters RJ, Flory JE, Jetter R, Ravn MM, Lee H-J, Coates RM, Croteau RB: **Abietadiene synthase from grand fir (*Abies grandis*): characterization and mechanism of action of the "pseudomature" recombinant enzyme**. *Biochemistry* 2000, **39**:15592-15602.

32. Peters RJ, Ravn MM, Coates RM, Croteau RB: **Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites**. *J Am Chem Soc* 2001, **123**:8974-8978.

33. Jardine CG, Gibson N, Hrudey SE: **Detection of odour and health risk perception of drinking water**. *Water Sci Technol* 1999, **40**:91-98.

34. Jiang J, He X, Cane DE: **Biosynthesis of the earthy odorant geosmin by a bifunctional *Streptomyces coelicolor* enzyme**. *Nat Chem Biol* 2007, **3**:711-715.

35. Jiang J, Cane DE: **Geosmin biosynthesis. Mechanism of the fragmentation-rearrangement in the conversion of germacradienol to geosmin**. *J Am Chem Soc* 2008, **130**:428-429.

36. • Harris GG, Lombardi PM, Pemberton TA, Matsui T, Weiss TM, Cole KE, Köksal M, Murphy FV IV, Vedula LS, Chou WKW *et al.*: **Structural studies of geosmin synthase, a bifunctional sesquiterpene synthase with αα domain architecture that catalyzes a unique cyclization-fragmentation reaction sequence**. *Biochemistry* 2015, **54**:7142-7155.

This paper reports the structure of bacterial geosmin synthase. The crystal structure of the N-terminal α domain that catalyzes the cyclization reaction forming germacradienol, and a homology model of the C-terminal α domain that catalyzes the subsequent cyclization and retro-Prins fragmentation reaction yielding geosmin and acetone, were used to fit the *ab initio* molecular envelope of the full-length enzyme calculated from SAXS data to yield a view of αα domain assembly.

37. Brodelius M, Lundgren A, Mercke P, Brodelius PE: **Fusion of farnesyldiphosphate synthase and epi-aristolochene synthase, a sesquiterpene cyclase involved in capsidiol biosynthesis in *Nicotiana tabacum***. *Eur J Biochem* 2002, **269**:3570-3577.

38. de Vries-van Leeuwen IJ, Kortekaas-Thijssen C, Nzigou Mandouckou JA, Kas S, Evidente A, de Boer AH: **Fusicoccin-A selectively induces apoptosis in tumor cells after interferon-α priming**. *Cancer Lett* 2010, **293**:198-206.

39. Bury M, Andolfi A, Rogister B, Cimmino A, Mégalizzi V, Mathieu V, Feron O, Evidente A, Kiss R: **Fusicoccin A, a phytotoxic carbotricyclic diterpene glucoside of fungal origin, reduces proliferation and invasion of glioblastoma cells by targeting multiple tyrosine kinases**. *Transl Oncol* 2013, **6**:112-123.

40. Toyomasu T, Tsukahara M, Kaneko A, Niida R, Mitsuhashi W, Dairi T, Kato N, Sassa T: **Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi**. *Proc Natl Acad Sci USA* 2007, **104**:3084-3088.

41. Toyomasu T, Niida R, Kenmoku H, Kanno Y, Miura S, Nakano C, Shiono Y, Mitsuhashi W, Toshima H, Oikawa H, Hoshino T, Dairi T, Kato N, Sassa T: **Identification of diterpene biosynthetic gene clusters and functional analysis of labdane-related diterpene cyclases in *Phomopsis amygdali***. *Biosci Biotechnol Biochem* 2008, **72**:1038-1047.

42. Chiba R, Minami A, Gomi K, Oikawa H: **Identification of ophiobolin F synthase by a genome mining approach: a sesterterpene synthase from *Aspergillus clavatus***. *Org Lett* 2013, **15**:594-597.

43. Matsuda Y, Mitsuhashi T, Quan Z, Abe I: **Molecular basis for stellatic acid biosynthesis: a genome mining approach for discovery of sesterterpene synthases**. *Org Lett* 2015, **17**:4644-4647.

44. Qin B, Matsuda Y, Mori T, Okada M, Quan Z, Mitsuhashi T, Wakimoto T, Abe I: **An unusual chimeric diterpene synthase from *Emericella variecolor* and its functional conversion into a sesterterpene synthase by domain swapping**. *Angew Chem Int Ed* 2016, **55**:1658-1661.

45. Matsuda Y, Mitsuhashi T, Lee S, Hoshino M, Mori T, Okada M, Zhang H, Hayashi F, Fujita M, Abe I: **Astellifadiene: structure determination by NMR spectroscopy and crystalline sponge method, and elucidation of its biosynthesis**. *Angew Chem Int Ed* 2016, **128**:5879-5882.

46. Toyomasu T, Tsukahara M, Kenmoku H, Anada M, Nitta H, Ohkanda J, Mitsuhashi W, Sassa T, Kato N: **Transannular proton transfer in the cyclization of geranylgeranyl diphosphate to fusicoccadiene, a biosynthetic intermediate of fusicoccins**. *Org Lett* 2009, **11**:3044-3047.

47. •• Chen M, Chou WKW, Toyomasu T, Cane DE, Christianson DW: **Structure and function of fusicoccadiene synthase, a hexameric bifunctional diterpene synthase**. *ACS Chem Biol* 2016, **11**:889-899.

This paper reports the X-ray crystal structure determinations of the individual GGPP synthase and GGPP cyclase domains of fusicoccadiene synthase, which were then used to fit the *ab initio* molecular envelope of the intact $(\alpha\alpha)_6$ hexamer.

48. •• Castellana M, Wilson MZ, Xu Y, Joshi P, Cristea IM, Rabinowitz JD, Gitai Z, Wingreen NS: **Enzyme clustering accelerates processing of intermediates through metabolic channeling**. *Nat Biotechnol* 2014, **32**:1011-1018.

This paper describes an elegant mathematical model demonstrating that enzyme clustering is comparable to direct channeling for the enhancement of product flux in a two-enzyme system. The calculations are validated in studies of carbamoyl phosphate synthesis and utilization through a metabolic branch point in *Escherichia coli*.

49. Tarshis LC, Yan M, Poulter CD, Sacchettini JC: **Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution**. *Biochemistry* 1994, **33**:10871-10877.

50. Grundy DJ, Chen M, González V, Leoni S, Miller DJ, Christianson DW, Allemann RK: **Mechanism of germacradien-4-ol synthase-controlled water capture**. *Biochemistry* 2016, **55**:2112-2121.

51. Köksal M, Chou WKW, Cane DE, Christianson DW: **Structure of 2-methylisoborneol synthase from *Streptomyces coelicolor* and implications for the cyclization of a noncanonical *C*-methylated monoterpenoid substrate**. *Biochemistry* 2012, **51**:3011-3020.

52. Alonso WR, Rajaonarivony JIM, Gershenzon J, Croteau R: **Purification of 4*S*-limonene synthase, a monoterpene cyclase from the glandular trichomes of peppermint (*Mentha* × *piperita*) and spearmint (*Mentha spicata*)**. *J Biol Chem* 1992, **267**:7582-7587.

53. Hyatt DC, Youn B, Zhao Y, Santhamma B, Coates RM, Croteau RB, Kang CH: **Structure of limonene synthase, a simple model for terpenoid cyclase catalysis**. *Proc Natl Acad Sci USA* 2007, **104**:5360-5365.

54. Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW: **Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase**. *Proc Natl Acad Sci USA* 2002, **99**:15375-15380.

55. Köksal M, Zimmer I, Schnitzler JP, Christianson DW: **Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission**. *J Mol Biol* 2010, **402**:363-373.

56. Kavanagh KL, Dunford JE, Bunkoczi G, Russell RGG, Oppermann U: **The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding**. *J Biol Chem* 2006, **281**:22004-22012.

57. Chang TH, Hsieh FL, Ko TP, Teng KH, Liang PH, Wang AHJ: **Structure of a heterotetrameric geranyl pyrophosphate synthase from mint (*Mentha piperita*) reveals intersubunit regulation**. *Plant Cell* 2010, **22**:454-467.

58. Oldfield E, Lin FY: **Terpene biosynthesis: modularity rules**. • *Angew Chem Int Ed* 2012, **51**:1124-1137.
This review provides a valuable summary of structure–function relationships for the various protein domains found in enzymes of terpenoid biosynthesis.

59. Copley SD: **An evolutionary perspective on protein moonlighting**. *Biochem Soc Trans* 2014, **42**:1684-1691.

60. Zhao B, Lei L, Vassylyev DG, Lin X, Cane DE, Kelly SL, Yuan H, Lamb DC, Waterman MR: **Crystal structure of albaflavenone monooxygenase containing a moonlighting terpene synthase active site**. *J Biol Chem* 2009, **284**:36711-36719.

61. Chooi YH, Hong YJ, Cacho RA, Tantillo DJ, Tang Y: **A cytochrome P450 serves as an unexpected terpene cyclase during fungal meroterpenoid biosynthesis**. *J Am Chem Soc* 2013, **135**:16805-16808.

62. Lee H, DeLoache WC, Dueber JE: **Spatial organization of enzymes for metabolic engineering**. *Metab Eng* 2012, **14**:242-251.

63. Cheng W, Li W: **Structural insights into ubiquinone biosynthesis in membranes**. *Science* 2014, **343**:878-881. •
This paper reports the structure determination of UbiA, an integral membrane protein that catalyzes the electrophilic aromatic ring substitution that prenylates *para*-hydroxybenzoic acid in the biosynthesis of ubiquinone A (coenzyme Q). UbiA adopts the class I terpenoid synthase α fold; in contrast with cytosolic terpenoid synthases, UbiA is adapted for solubilization in the membrane to enable access to isoprenoid diphosphate substrates containing 6–10 isoprenoid units similarly solubilized in the membrane.