Pure Appl. Chem., Vol. 80, No. 8, pp. 1791–1798, 2008.
doi:10.1351/pac200880081791

Chemical wizardry? The generation of diversity in terpenoid biosynthesis*

Rudolf K. Allemann

School of Chemistry, Cardiff University, Park Place, Cardiff, CF10 3AT, UK

Abstract: Terpene synthases catalyze the conversion of linear prenyl-diphosphates to a multitude of hydrocarbon skeletons with often high regio- and stereoselectivity. These remarkable enzymes all rely on a shared fold for activity, namely, the class I terpene cyclase fold. Recent work has illuminated the catalytic strategy used by these enzymes to catalyze the arguably most complex chemical reactions found in Nature. Terpene synthases catalyze the formation of a reactive carbocation and provide a template for the cyclization reactions while at the same time providing the necessary stability of the carbocationic reaction intermediates as well as strictly controlling water access.

Keywords: catalysis; terpene; carbocation; conformation; enzyme.

INTRODUCTION

Several tens of thousands of terpene natural products have been described to date, yet all of these derive from only a small number of linear precursors. The 15-carbon isoprenoid, farnesyl diphosphate (FPP), is the biogenetic precursor of more than 300 different sesquiterpene hydrocarbon scaffolds in plants, bacteria, and fungi [1–6]. The hydrocarbon products undergo further derivatization to lead to many thousands of sesquiterpene products such as the antimalarial artemisinin, PR- [7] and T-2 toxins [8], the male natural contraceptive gossypol [9], antibiotics such as pentalenolactone [10], or parthenolide [11], which has an anti-migraine effect (Fig. 1). Cyclization of FPP (and other isoprenyl diphosphates) is catalyzed by terpene synthases that rely on a common mainly α-helical structure known as the class I terpene cyclase fold (Fig. 2) to promote the arguably most complex chemical reactions occurring in Nature [4,5]. These enzymes promote the Mg^{2+}-dependent expulsion of the pyrophosphate group from the substrate and serve as high-fidelity templates to subtly channel conformation and stereochemistry during the cyclization reactions. They are key to the generation of the enormous diversity in structure and stereochemistry found in terpenoids. Terpene biosynthesis thus appears to be a striking example of chemical wizardry used by Nature to generate structural and stereochemical diversity.

We provide here an overview of results from a combination of chemical and molecular biological experiments to address how a specific sesquiterpene cyclase, namely, aristolochene synthase from *Penicillium roqueforti* (PR-AS) chaperones its substrate and reaction intermediates along complex reaction pathways leading to the production of aristolochene with high specificity (Fig. 3). In this cyclization cascade two rings, three chiral centers, and two double bonds are formed regio- and stereospecifically from a linear achiral substrate within the active site of the enzyme, and approximately half of the carbon atoms undergo changes in hybridization, configuration, or bonding.

---

*Paper based on a presentation at CHEM-BIO-TECH-2007, a joint meeting of the IUPAC 1st Symposium on Chemical Biotechnology (ISCB-1) and the 8th Symposium on Bioorganic Chemistry (ISBOC-8), 8–11 August 2007, Turin, Italy. Other presentations are published in this issue, pp. 1773–1882.

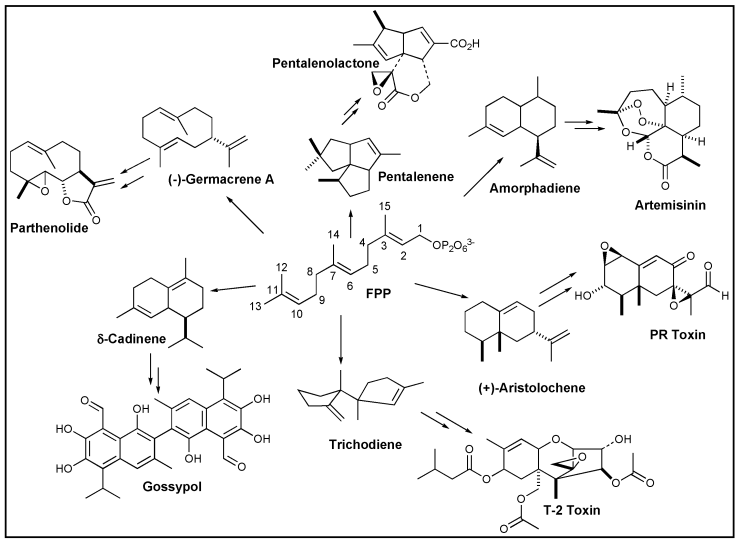

Fig. 1 Structure of FPP, the precursor of a myriad of sesquiterpenoid natural products. The action of sesquiterpene synthases converts FPP to over 300 different hydrocarbons, which can be further modified in downstream metabolic processes into complex natural products such as pentanolactone, artemisinin, PR-toxin, T-2 toxin, gossypol, and parthenolide.

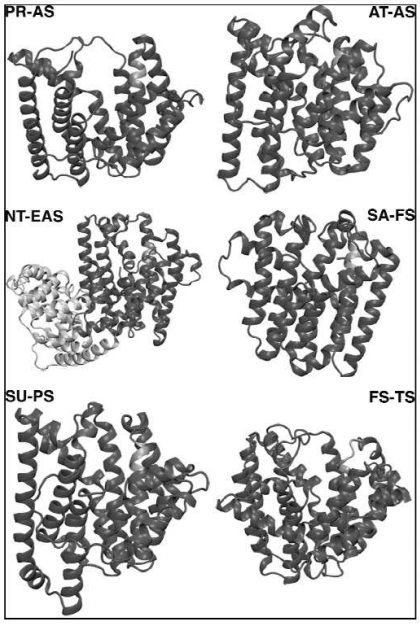

Fig. 2 Structural similarities among sesquiterpene synthases define the class I terpenoid fold (blue). Aristolochene synthases from *Penicillium roqueforti* (PR-AS) [26] and *Aspergillus terreus* (AT-AS) [20], 5-epi-aristolochen synthase from *Nicotiana tabacum* (NT-EAS) [27], *Staphoylococcus aureus* farnesyl-diphosphate synthase (SA-FS) [28], pentalenene synthases from *Streptomyces* UC5319 (SU-PS) [29] and trichodiene synthase from *Fusarium sporotrichioides* [30] are shown. The N-terminal domain of NT-EAS (gray) is of unknown function but shows structural similarity to glucoamylase [31] and endoglucanase CelD [32] and has been designated the class-II terpene synthase fold [4]. This fold is typical for plant terpene synthases.

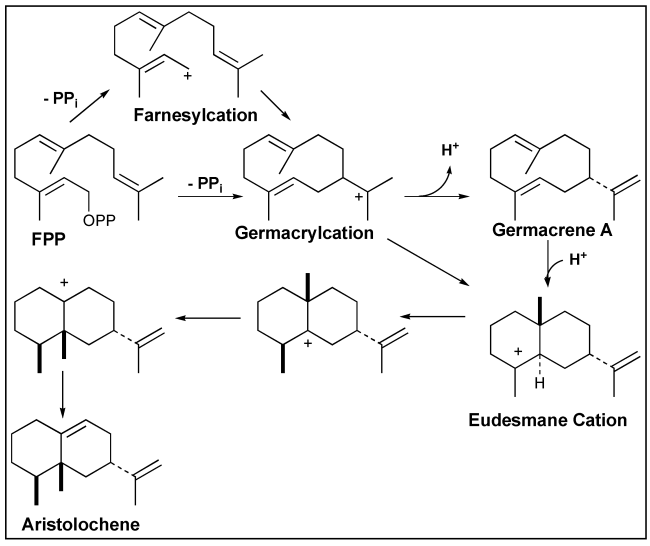

Fig. 3 Possible mechanisms for the cyclization and rearrangement of FPP to aristolochene catalyzed by aristolochene synthases (AT-AS and PR-AS) proceeding through germacryl and eudesmane cation.

INTERMEDIACY OF GERMACRENE A

Based on mechanistic studies with labeled substrates and through the analysis of the reaction products obtained from PR-AS catalysis, a chemical mechanism for the Mg^{2+}-dependent conversion of FPP to (+)-aristolochene was proposed, in which the C10–C11 double bond displaces the diphosphate leaving group in a S_N2-like reaction to generate germacryl cation followed by proton loss from C12 to generate germacrene A (Fig. 3). This uncharged intermediate was postulated to undergo protonation of the C6–C7 double bond and a further cyclization to form the bicyclic eudesmane cation. Successive 1,2 hydride shift and methyl migration followed by loss of H_{Si} on C8 results in the generation of (+)-aristolochene. The intermediacy of germacrene A is tentatively supported by the observations that germacrene A is a product of PR-AS catalysis [12] and that several mutants of PR-AS displayed increased production of germacrene A (Tyr 334, Tyr 92) [12,13]. On the other hand, no germacrene A was observed during catalysis by AS from a different fungal source, *Aspergillus terreus*. In addition, germacrene A does not appear to act as a substrate of AS and the active site acid required to generate eudesmane cation has so far been elusive [12,14,15]. These observations suggest that germacrene A might be an off-pathway product rather than an intermediate in the predominant channel leading to the formation of aristolochene [14].

Rather than analyzing the effects of structural changes of the enzyme on the product distribution, substrate analogs can be used to study the reaction mechanism. Fluorinated substrates have proved useful in the elucidation of mechanistic details of terpenoid biosynthesis since fluorine substituents do not greatly affect the binding affinities while at the same time exerting a strong influence on the electronic environment at the site of replacement. 2-Fluorofarnesyl-diphosphate was found to act as a substrate of PR-AS [16]. Due to the destabilizing inductive effect of the fluorine substituents on cations located on the β-carbon, the intermediate 2F-eudesmane cation is destabilized sufficiently so that 2F-FPP is converted exclusively to 2-fluorogermacrene A.

The substrate analog 12,13-difluoro-FPP was found to be a potent reversible competitive inhibitor of PR-AS with a K_I of 0.8 ± 0.2 μM, suggesting that the initial cyclization of FPP to germacryl cation occurs along a reaction pathway in which FPP ionization is accompanied by electrophilic attack of C1 by the C10, C11 π-bond with inversion of configuration at C1 (Figs. 3 and 4) [17]. On the other hand,

experiments with site-specific mutants in which Phe 112 or Tyr 92 were replaced with alanine leading to the production of (*E*)-β- and (*E*,*E*)-α-farnesene and some β-bisabolene in the case of AS-Y92A, had been interpreted to suggest that the cyclization of FPP might proceed in a stepwise fashion through farnesyl cation (Fig. 3) [18,19]. The aromatic residues Phe 92 and Phe 112 (together with other active site residues such as Phe 178 and Trp 334) exert a stereoelectronic effect on the reaction through their steric bulk by aligning the π-orbital of the C10, C11 double bond with the breaking C1–O bond of the diphosphate group (Fig. 4). When the size of these side chains is reduced in the mutants enzymes, the immediate quenching of the developing positive charge on C1 may be prevented, leading to production of the linear products through deprotonation. This interpretation is strongly supported by the observation that 12,13-difluoro-FPP was a substrate of the mutant enzyme AS-F112A [17].

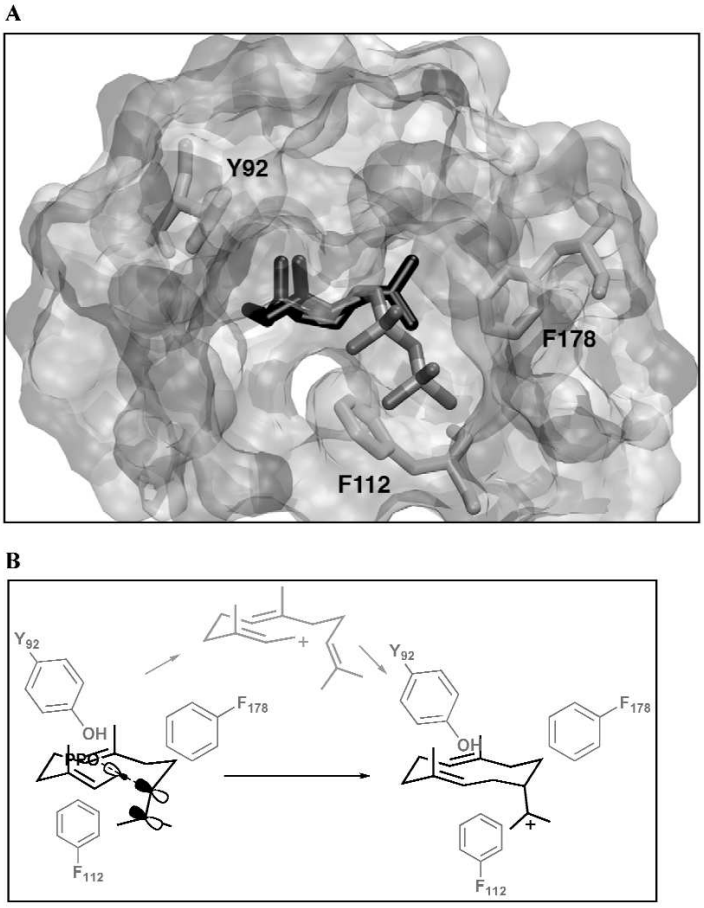

Fig. 4 AS serves as a template for the folding of its substrate into a reactive conformation. (A) Some of the bulky residues in the active site of AS that shape the active site contour and mutagenesis of which alters the product distribution as a consequence of altered substrate conformation. (B) Stereoelectronic control of the formation of germacryl cation by AS; proper alignment of the 10,11-double bond with the C–O bond of the leaving group leads to the formation of germacryl cation in a S_N2-like reaction and avoids the formation of farnesyl cation.

ACTIVE SITE ACID

The results described above are in agreement with the suggestion that germacrene A is an on-path reaction intermediate during AS catalysis. However, if germacrene A is indeed formed through deprotonation of germacryl cation, then an active site acid is required to reionize germacrene A. Several active site acids have been proposed, but there is a lack of experimental evidence to support their involvement in the reaction. Molecular modeling of the enzyme-complexed germacrene A suggested that Tyr 92 could serve as the general acid to protonate the C6–C7 double bond and subsequently as the general base for the deprotonation from C8, but ASY92F produced appreciable quantities of aristolochene, ruling out Tyr 92 as the sole catalytically obligatory general acid [12,15]. Attempts to support a shuttle mechanism in which a proton was delivered from the surface of the protein to the active site by way of a hydrogen-bonding network involving Lys 206, Asp 200, Arg 203, and Tyr 92 [12] led to largely inconclusive results (Taylor and Allemann, unpublished). Alternative proposals such as an unprecedented active site oxonium [15] ion or the pyrophosphate itself [20] also suffer from lack of experimental evidence to support them.

Density functional theory calculations in the gas phase on the cyclization of FPP along with molecular docking and modeling studies in the active site of PR-AS revealed an alternative, thermodynamically more economical reaction pathway, in which protonation of the C6–C7 double bond is accomplished by intramolecular proton transfer rather than through the involvement of a general acid (Fig. 3), thereby avoiding the quenching of the positive charge in the high-energy germacryl cation [14]. The small amount of germacrene A produced by the wild-type enzyme is the result of proton loss from C12 in germacryl cation to generate the isopropylidene group. To the best of our knowledge, this intramolecular proton-transfer mechanism is in agreement with all available experimental data.

STABILIZING CARBOCATIONIC INTERMEDIATES

The positive transition states leading to the formation of germacryl and eudesmane cation and the cations themselves must be stabilized efficiently within the "mild" environment provided by the active sites of enzymes, irrespective of whether eudesmane cation is formed through intra- or intermolecular proton transfer. Negatively charged amino acids are largely ruled out since their interaction with the carbocationic intermediates might easily lead to alkylation and suicide inhibition. The analysis of the composition of the active sites of terpene cyclases indicated that these enzymes provide an unreactive hydrophobic environments that mimics aprotic organic solvents. Additionally, carbocationic intermediates appear to be stabilized substantially by interaction with the π-systems of aromatic residues such as phenylalanine, tyrosine, and tryptophane. Such cation-π or charge-quadrupole interactions can contribute significant stabilization [21] and have been found to promote both the formation of germacryl and eudesmane cations.

Inspection of the X-ray structure of PR-AS suggested that the π-system of Trp 334 could interact favorably with the positive charge at C3 of eudesmane cation [13]. This is an ideal strategy for the stabilization of high-energy carbocationic intermediates since it prevents the quenching of the positive charge. In agreement with this proposal, replacement of Trp 334 with Phe did not significantly change the distribution of terpenoids produced in incubations with FPP. However, the enzymes PR-ASW334V produced predominantly germacrene A in addition to ~5 % aristolochene, while PR-ASW334L produced exclusively germacrene A. It is noteworthy that the catalytic activity of PR-ASW334L and PR-ASW334V were dramatically reduced, indicating that Trp 334 also plays a role in the catalytic steps before the formation of eudesmane cation (vide infra).

The conversion of germacyl cation to aristolochene is also facilitated by the bulky aromatic side chain of Phe 178, a residue that is ideally placed to stabilize the developing positive charge on C2/C3 of eudesmane cation [18,22]. However, it appears that it is the large size of Phe 178 that catalyzes the 1,2 hydride shift from C2 in eudesmane cation rather than its aromaticity that promotes the formation

of aristolochene during AS catalysis. The aromaticity of Phe 178, on the other hand, is involved in the efficient formation of germacryl cation from FPP. One of the other key residues for the efficient formation of germacryl cation is Phe 112. Its replacement with alanine [18] or cysteine (Forcat and Allemann, unpublished) led to reduced catalytic activity and the production of small amounts of germacrene A (and in the case of PR-AS-F112C, 5 % aristolochene) in addition to the linear (*E*,*E*)-α- and (*E*)-β-farnesene, indicating that residue 112 contributes significant stabilization to the transition state leading to germacryl cation. Interestingly, matrix-assisted laser desorption ionization with time-of-flight mass spectrometry (MALDI-TOFMS) analysis of PR-AS-F112C after incubation with substrate revealed an increase in the mass of the protein by 204 units, suggesting the formation of a covalent bond between PR-AS-F112C and a reaction intermediate (Fig. 5) (Forcat and Allemann, unpublished). Identification of the trapped intermediate is currently under way.

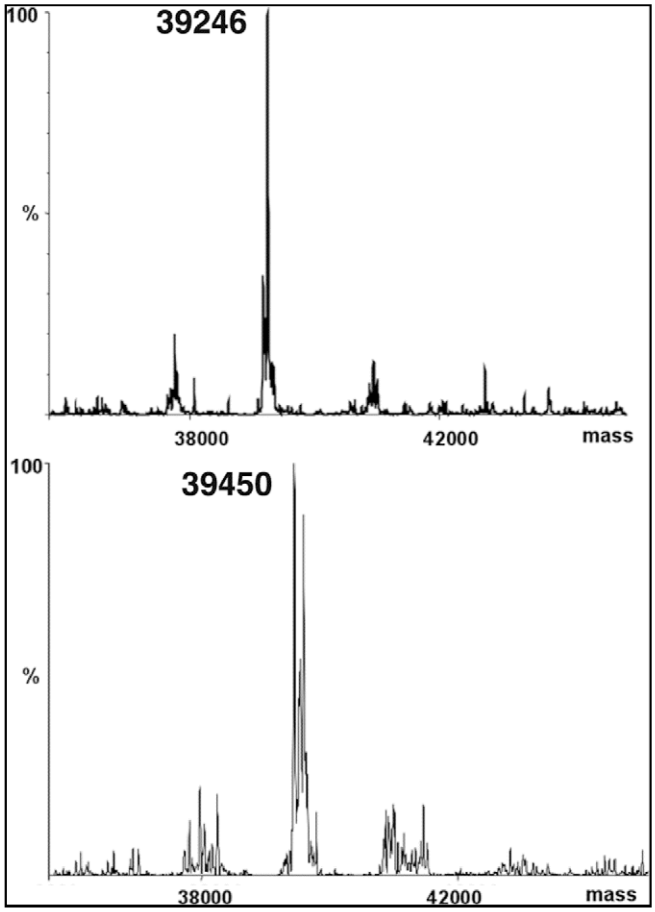

Fig. 5 Mass spectrum of AS-F112C before (top) and after incubation with substrate. The increased mass suggests the modification of the enzyme with a sesquiterpenoid.

SUBSTRATE CONFORMATION, ENZYME TEMPLATING, AND FURTHER EVOLUTION OF TERPENE SYNTHASES

Clearly, AS (and other terpenoid cyclases) provides an environment in which the positive charges of transition states and intermediates are suitably stabilized and protected from solvent. A key feature of catalysis is that AS provides a template that guides the folding of the flexible linear farnesyl pyrophos-

phate into a reactive conformation that ensures optimal orbital overlap (vide supra) to induce the cyclization cascade in a chemically optimal environment. These enzymes that catalyze what are arguably the most complex chemical reaction found in Nature, may do little in terms of active catalysis after the initial ionization than to provide an "active" template that nurtures the intrinsic chemical reactivity of the folded isoprenyl-chain.

Single amino acid changes in the active site of AS were sufficient to dramatically change the course of the cyclization reaction. In addition to the examples discussed above, a particularly striking case was found when Tyr 92 was replaced with smaller residues such as tyrosine, phenylalanine, leucine, valine, cysteine, or alanine (Deligeorgopulou and Allemann, unpublished) [19,23]. PR-AS-Y92A produced more than 80 % of the linear farnesenes in addition to small amounts of α-selinene, β-selinene and selina-4,11-diene, but no germacrene A or aristolochene, while for the wild-type enzyme only cyclic hydrocarbons were observed. Interestingly, a linear relationship was observed between the van der Waals volume of residues 92 and the amount cyclic products (Deligeorgopoulou, Taylor, Calvert and Allemann, unpublished). In this context it is interesting to note that the ASs from *P. roqueforti* and from *A. terreus* have evolved in a divergent fashion (they show 61 % sequence identity) but conserved their active site contours [20]. Hence, they provide nearly identical active site templates within a shared protein fold for the binding of FPP and its conversion to aristolochene. The enzyme that produces the aristolochene epimer 5-epi-aristolochene provides a significantly different active site template for the substrate.

The plasticity of the terpene cyclases appears to provide a framework for the combinatorial production of many natural terpenoids though subtle alterations in the composition of the active site during evolution. Greenhagen et al. have recently shown that active site template can be altered within the class I terpenoid fold by altering a small number of residues in and around the active site, thereby converting one terpene synthase into another and preserving high catalytic activity [24]. Recently, we have shown that analogs of FPP containing phenyl substituents in place of methyl groups act as potent competitive inhibitors of PR-AS (K_I values ranging from 0.8 to 1.2 μM) suggesting that the active sites of terpene synthases are sufficiently flexible to accommodate even substrate analogs with large substituents [25]. Molecular modeling and preliminary X-ray crystallographic studies suggest that only small alterations of the conformation of the backbones of the inhibitors are sufficient to accommodate the phenylfarnesyl-diphosphates. The reorganization of the active site configuration did, however, lead to loss of catalytic activity. Modification of the active site residues may generate enzymes capable of catalytic turnover of the phenyl-substituted FPP molecules. Such an approach may offer a general strategy for the production of novel, unnatural "terpenoids" both in vitro and in vivo.

CONCLUDING REMARKS

Extensive studies of the AS-catalyzed conversion of FPP to aristolochene has illuminated the catalytic strategy used by this enzyme and related terpene synthases. Nature clearly does not depend on chemical wizardry to catalyze these most complex chemical transformations, but simply provides us with an impressive lesson in combinatorial chemistry that obeys the chemical rules that scientists have deciphered over many decades—a lesson that should enable further evolution of terpene synthases both in vitro and in Nature to generate novel "unnatural" terpenoids with enormous potential for chemistry, biology, and medicine.

ACKNOWLEDGMENTS

This work was supported financially by the Royal Society, Biotechnology and Biological Sciences Research Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC), and Cardiff University.

REFERENCES

1. J. S. Glasby. *Encyclopedia of Terpenoids*, John Wiley, Chichester (1982).
2. D. E. Cane. *Acc. Chem. Res.* **18**, 220 (1985).
3. D. E. Cane. *Chem. Rev.* **90**, 1089 (1990).
4. K. U. Wendt, G. E. Schulz. *Structure* **6**, 127 (1998).
5. D. W. Christianson. *Chem. Rev.* **106**, 3412 (2006).
6. R. Croteau, D. E. Cane. *Meth. Enzymol.* **110**, 383 (1985).
7. R. H. Proctor, T. M. Hohn. *J. Biol. Chem.* **268**, 4543 (1993).
8. D. E. Cane, S. Swanson, P. P. N. Murthy. *J. Am. Chem. Soc.* **103**, 2136 (1981).
9. E. M. Coutinho. *Contraception* **65**, 259 (2002).
10. C. N. Tetzlaff, Z. You, D. E. Cane, S. Takamatsu, S. Omura, H. Ikeda. *Biochemistry* **45**, 6179 (2006).
11. M. J. Hewlett, M. J. Begley, W. A. Groenewegen, S. Heptinstall, D. W. Knight, J. May, U. Salan, D. Toplis. *J. Chem., Perkin Trans. 1* 1979 (1996).
12. M. J. Calvert, P. R. Ashton, R. K. Allemann. *J. Am. Chem. Soc.* **124**, 11636 (2002).
13. A. Deligeorgopoulou, S. E. Taylor, S. Forcat, R. K. Allemann. *Chem. Commun.* 2162 (2003).
14. R. K. Allemann, N. J. Young, S. Ma, D. G. Truhlar, J. Gao. *J. Am. Chem. Soc.* **129**, 13008 (2007).
15. B. Felicetti, D. E. Cane. *J. Am. Chem. Soc.* **126**, 7212 (2004).
16. D. J. Miller, F. Yu, R. K. Allemann. *ChemBioChem* **8**, 1819 (2007).
17. F. Yu, D. J. Miller, R. K. Allemann. *Chem. Commun.* 4155 (2007).
18. S. Forcat, R. K. Allemann. *Org. Biomol. Chem.* **4**, 2563 (2006).
19. A. Deligeorgopoulou, R. K. Allemann. *Biochemistry* **42**, 7741 (2003).
20. E. Y. Shishova, L. Di Costanzo, D. E. Cane, D. W. Christianson. *Biochemistry* **46**, 1941 (2007).
21. C. Jenson, W. L. Jorgensen. *J. Am. Chem. Soc.* **119**, 10846 (1997).
22. S. Forcat, R. K. Allemann. *Chem. Commun.* 2094 (2004).
23. M. J. Calvert, S. E. Taylor, R. K. Allemann. *Chem. Commun.* 2384 (2002).
24. B. T. Greenhagen, P. E. O'Maille, J. P. Noel, J. Chappell. *Proc. Natl. Acad. Sci. USA* **103**, 9826 (2006).
25. D. J. Miller, F. Yu, N. J. Young, R. K. Allemann. *Org. Biomol. Chem.* **5**, 3287 (2007).
26. J. M. Caruthers, I. Kang, M. J. Rynkiewicz, D. E. Cane, D. W. Christianson. *J. Biol. Chem.* **275**, 25533 (2000).
27. C. M. Starks, K. Back, J. Chappell, J. P. Noel. *Science* **277**, 1815 (1997).
28. D. J. Hosfield, Y. M. Zhang, D. R. Dougan, A. Broun, L. W. Tari, R. V. Swanson, J. Finn. *J. Biol. Chem.* **279**, 8526 (2004).
29. C. A. Lesburg, G. Zhai, D. E. Cane, D. W. Christianson. *Science* **277**, 1820 (1997).
30. M. J. Rynkiewicz, D. E. Cane, D. W. Christianson. *Proc. Natl. Acad. Sci. USA* **98**, 13543 (2001).
31. A. E. Aleshin, C. Hoffman, L. M. Firsov, R. B. Honzatko. *J. Mol. Biol.* **238**, 575 (1994).
32. M. Juy, A. G. Amit, P. M. Alzari, R. J. Poljak, M. Claeyssens, P. Beguin, J. P. Aubert. *Nature* **357**, 89 (1992).