

Enzyme Catalysis

International Edition: DOI: 10.1002/anie.201509060
German Edition: DOI: 10.1002/ange.201509060Blocking Deprotonation with Retention of Aromaticity in a Plant *ent*-Copalyl Diphosphate Synthase Leads to Product Rearrangement

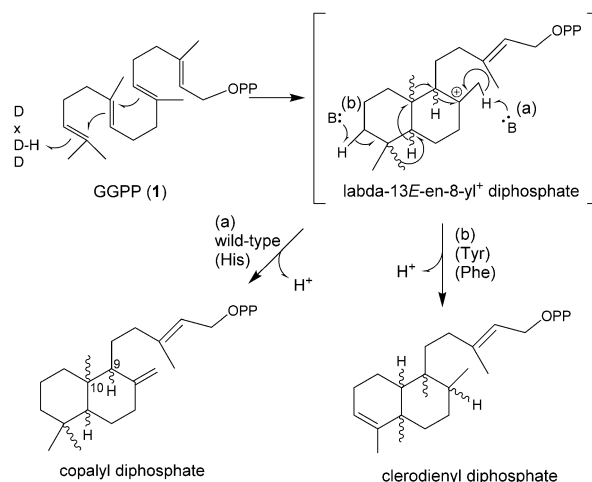
Kevin C. Potter, Jiachen Zi, Young J. Hong, Samuel Schulte, Brandi Malchow, Dean J. Tantillo, and Reuben J. Peters*

Abstract: Substitution of a histidine, comprising part of the catalytic base group in the *ent*-copalyl diphosphate synthases found in all seed plants for gibberellin phytohormone metabolism, by a larger aromatic residue leads to rearrangements. Through a series of 1,2-hydride and methyl shifts of the initially formed bicycle predominant formation of (–)-kolavenyl diphosphate is observed. Further mutational analysis and quantum chemical calculations provide mechanistic insight into the basis for this profound effect on product outcome.

Biosynthesis of the labdane-related diterpenoids, a large superfamily of over 7000 natural products, is characterized by the use of class II diterpene cyclases.^[1] The vast majority of these biocatalysts are exclusively found in plants, stemming from repeated evolutionary diversification of the *ent*-copalyl diphosphate synthase (CPS) required for gibberellin phytohormone biosynthesis, that is, by gene duplication and neofunctionalization.^[2] As implied by this evolutionary scenario, these enzymes must exhibit catalytic plasticity which enables new products to readily arise. Consistent with this hypothesis, we have shown that substitution of smaller residues for the catalytic base group can lead to the incorporation of water, thus yielding hydroxylated products.^[3] Here it is reported that replacement of the histidine in the catalytic base group with larger aromatic groups leads to even more profound alteration of product outcome, specifically rearrangement through a series of 1,2-hydride and methyl shifts. This finding, coupled with additional mutational analysis and quantum chemical calculations, clarifies the catalyzed reaction and underlying enzymatic structure–function relationships of the ancestral CPS.

Class II diterpene cyclases catalyze cycloisomerization of the general diterpenoid precursor (*E,E,E*)-geranylgeranyl diphosphate (GGPP; **1**), thus utilizing an acid/base mechanism wherein the middle aspartate of a highly conserved DxDD motif is employed as the general acid to protonate the terminal C=C bond of **1**.^[4] Sequential *anti* addition from the

two internal C=C bonds leads to the bicyclic (*trans*-decalin) intermediate, labda-13*E*-en-8-yl⁺ diphosphate. Typically this intermediate is deprotonated at C17 to yield copalyl diphosphate (CPP), which can differ in stereochemistry, for example, the (9*R*,10*R*) *ent*-CPP (**2**; for structure see Scheme 2), which is relevant to gibberellin biosynthesis.^[1] However, other product outcomes are possible. Particularly relevant here, rather than deprotonation to produce CPP, is a sequence of 1,2-hydride and methyl shifts, thus resulting in new skeleton backbones such as the fully rearranged clerodanes (Scheme 1).^[1]



Scheme 1. Relevant cyclization reactions.

Although class II diterpene cyclases have been characterized to produce rearranged labdane-related diterpenoids from bacteria (see Scheme S1 in the Supporting Information),^[5] none are known from plants, and the enzymatic basis for such rearrangement is not known.

Based on previously determined crystal structures,^[6] the catalytic base group in the CPS from *Arabidopsis thaliana* (AtCPS) was recently identified.^[3c] In this group the general base appears to be a water molecule, activated by ligation, in part by the side chains of a histidine and asparagine, which form a catalytic dyad conserved in all CPSs involved in gibberellin plant hormone biosynthesis (H263 and N322 in AtCPS). The role of the water was indicated by the fact that substitution of alanine for either residue, or both, led to addition of water to the *ent*-labda-13*E*-8-yl⁺ intermediate (**A**; for structure see Scheme 2), with formation of *ent*-labda-13*E*-en-8-ol diphosphates, thus yielding both epimers (8*α*- and 8*β*-hydroxyls).^[3c] Substitution of larger aliphatic residues for

[*] K. C. Potter, Dr. J. Zi, S. Schulte, B. Malchow, Prof. Dr. R. J. Peters Roy J. Carver Department of Biochemistry
Biophysics & Molecular Biology, Iowa State University
Ames, IA 50011 (USA)
E-mail: rjpeters@iastate.edu

Dr. Y. J. Hong, Prof. Dr. D. J. Tantillo
Department of Chemistry, University of California, Davis
Davis, CA 95616 (USA)

Supporting information and ORCID(s) from the author(s) for this article are available on the WWW under <http://dx.doi.org/10.1002/anie.201509060>.

H263 in AtCPS led to increasingly inactive enzymes, and was interpreted as suggesting that a correctly positioned base (e.g., either the imidazole side chain of His or sufficient space for more water molecules) is required for catalysis. However, it is possible that such an electron-rich group is required to initiate the reaction, but not necessarily as a general base. This hypothesis was tested by substitution of larger aromatic residues for the catalytic His. Modeling indicated that a tryptophan would be too large, thus, phenylalanine and tyrosine were substituted for H263 in AtCPS (AtCPS:H263F and AtCPS:H263Y; see the Experimental Section in the Supporting Information). Strikingly, novel enzymatic activity was observed, either in vitro or through expression in *E. coli* which was also engineered to produce **1** (Figure 1).^[7] The new

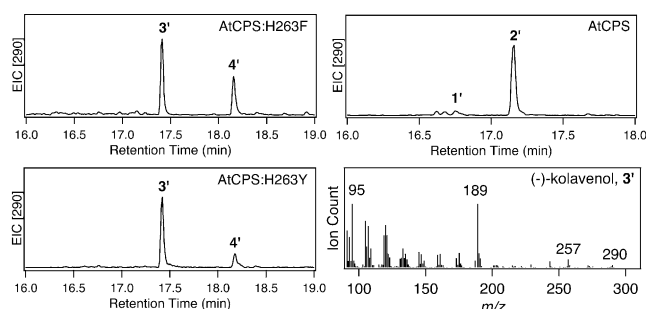
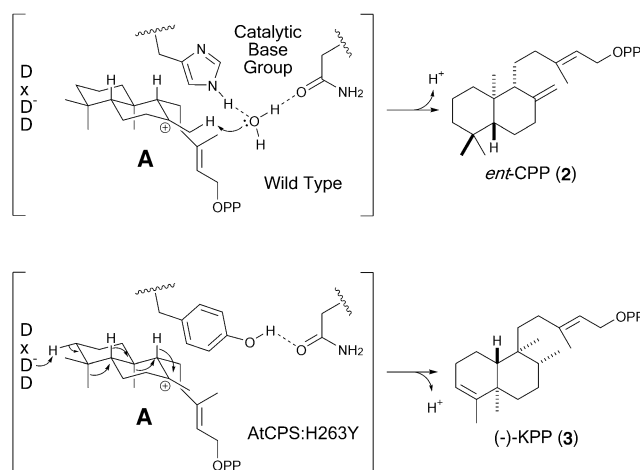


Figure 1. Effect of H263Y and H263F mutations on AtCPS product outcome. Wild-type also shown for comparative purposes. Extracted ion chromatographs (m/z 290) from GC-MS analyses of the dephosphorylated products (numbering as in text, with prime' notation used to indicate that these are dephosphorylated derivatives). Also presented is the mass spectrum for the novel peak (**3'**, kolavenol).

compound, following dephosphorylation and analysis by gas chromatography with mass spectral detection (GC-MS), exhibited a mass spectrum similar to that of the rearranged clerodane produced by a bacterial terpenedienyl diphosphate synthase.^[5a] However, these compounds exhibited different retention times, thus demonstrating some (stereo)-chemical difference between them (see Figure S1). Therefore, approximately 1.5 mg of the dephosphorylated compound was purified and its structure investigated by NMR spectroscopy. This structure was found to correspond to a neoclerodane,^[8] and comparison to previously reported spectral data and optical rotation demonstrated that this was (–)-kolavenol (see Table S1 and Figures S2–S4).^[9] Therefore, these mutants, AtCPS:H263F and, particularly, AtCPS:H263Y, produce (–)-kolavenyl diphosphate [(–)-KPP; **3**; Scheme 2]. Thus, they not only catalyze bicyclization of **1**, but also promote a subsequent series of 1,2-hydride and methyl shifts, thereby leading to full rearrangement and a novel product outcome. Consistent with the good yields observed from the bacterial metabolic engineering system (Figure 1), kinetic analysis demonstrates that both mutants exhibit only relatively small decreases in catalytic efficiency relative to wild-type AtCPS (ca. 10-fold), with 30-fold decreases in k_{cat} offset by about twofold increases in substrate affinity, as well as approximately fivefold decreases in substrate inhibition (see Table S2).



Scheme 2. Reactions catalyzed by wild-type and the H263Y mutant of AtCPS.

Formation of **3** requires removal of a proton from the same carbon atom that is protonated to initiate the bicyclization reaction. This mechanism suggests that the same proton used to initiate bicyclization might be removed after the subsequent rearrangement. To test this mechanistic hypothesis, labelling studies were performed with in vitro enzymatic assays in D_2O with wild-type and the H263Y mutant enzymes. Upon overnight incubation, dephosphorylation of the products, and analysis by GC-MS, it was found that (–)-kolavenol-derived from **3**, produced by the mutant enzyme in D_2O , retained a molecular weight of 290 Da, that is, deuterium is not incorporated. However, *ent*-copalol derived from **2**, formed by the wild-type enzyme in D_2O , exhibited a molecular weight of 291 Da, that is, as expected, deuterium is incorporated (see Figure S5). This result demonstrates that formation of **3** proceeds by removal of the same proton originally added to initiate (bi)cyclization. Although this result does not rule out other possible bases (e.g., another water molecule), it seems likely that the middle aspartate of the DxDD motif (D379 in AtCPS) acts as the catalytic acid, as well as the catalytic base to remove the proton in formation of **3**.

These results provide insights into the overall reaction mechanism of class II diterpene cyclases. Substitution of His263 with aromatics still promotes protonation and bicyclization, which presumably results from not only enforcement of the folding of **1** into the correct conformation, but also interactions of the aromatic π electrons with the resulting intermediate **A** (Scheme 2).^[10] As previously described,^[3c] substitution of H263 with larger aliphatic residues (i.e., valine and isoleucine) simply resulted in either decreased yields of the same *ent*-labda-13*E*-en-8*α*-ol diphosphate (**4**) primary product as formed by AtCPS:H263A (i.e., with AtCPS:H263V), or essentially complete loss of any product (i.e., with AtCPS:H263I). Although this may be due in part to improper folding (of the protein and/or substrate), no rearranged compounds were detected from reactions with any of these mutants. The ability of larger aromatic side chains, at this position (residue 263 in AtCPS), to promote the production of **3** presumably arises from sterically blocked

access to the carbocation center of **A** (e.g., by a water molecule for production of **4**). Notably, the AtCPS:H263Y mutant leads to greater production of **3** (77 %), at the expense of **4**, than does the AtCPS:H263F mutant (55 % of **3**). This difference may be due to the potential ability of the tyrosine hydroxy group to form a hydrogen bond with N322, thus generating a more stable enzyme and/or, possibly even more importantly, increased effectiveness in blocking the binding of water at this position in the active site (see Scheme 2 and Figure S6).

The role of the Asn from the catalytic dyad (N322 in AtCPS) in the observed production of **3** was probed by additional mutagenesis. Substitution of leucine alone (i.e., AtCPS:N322L) led to only production of unrearranged products, predominantly the hydroxylated **4**, along with small amounts of the epimeric *ent*-labda-13*E*-en-8 β -ol diphosphate (**5**). However, in the context of aromatic substitutions for H263, significant amounts of **3** were observed as well, particularly with AtCPS:H263F/N322L (75 % versus the 33 % of **3** observed with AtCPS:H263Y/N322L). Significant loss of yield in the bacterial metabolic engineering system was observed with these two double mutants (Figure 2). While this loss could be restored to some extent by alanine

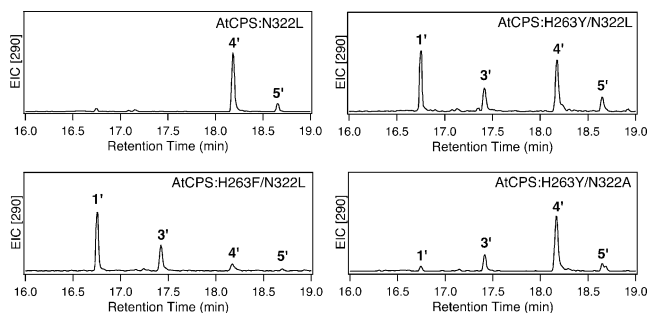


Figure 2. Effect of N322L mutation, in the absence or presence of H263Y or F mutations, on AtCPS product outcome. H263Y/N322A double mutant also is shown for comparative purposes. Extracted ion chromatographs (m/z 290) from GC-MS analyses of the dephosphorylated products (numbering as in text, with prime' notation used to indicate that these are dephosphorylated derivatives).

substitution instead (e.g., AtCPS:H263Y/N322A), it also reduced the production of **3** relative to **4**. Intriguingly, these results differ from those observed for substitution of H263, and seems to support the hypothesis that either an aromatic side chain or one sufficiently small enough to allow water to bind in its place is required at this position to initiate bicyclization. And there is only a minimal (if any) role for the asparagine residue from the catalytic base group, apart from serving as an activating ligand for the water, which normally serves as the general base.

Structural studies and quantum chemical analysis of the interactions of the π faces of aromatic groups with carbocations indicates that C–H $\cdots\pi$ interactions may play important roles in terpene synthase catalyzed reactions.^[10,11] However, the ability of an aromatic group in the catalytic base group to stabilize **A** does not immediately suggest how this group might also promote initiation of the reaction. Accordingly,

this effect was further probed by quantum chemical calculations (see the Supporting Information for details), the application of which to terpene cyclization reactions has been validated by experimental tests of mechanistic predictions.^[12] Here three model systems were examined: the geranylgeranyl-15-yl⁺ diphosphate formed by initial protonation (the fully optimized conformer most suited for cyclization was used; the tail leading to the OPP group was removed to simplify this series of calculations), a carbocation in which the two C–C bonds, which form during bicyclization, were constrained to 2.50 Å (to approximate partially formed bonds, although these likely do not form synchronously in the enzymatic reaction; Figure 3, left), and **A**. A theozyme-based

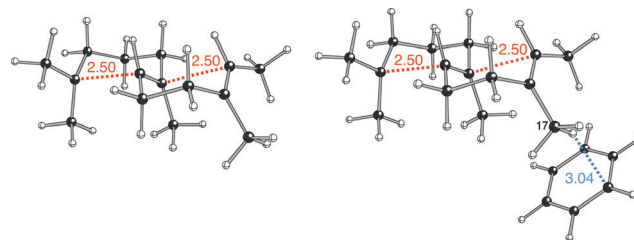


Figure 3. The free constrained carbocation (left) and the constrained carbocation in the presence of benzene (right). Constrained distances [Å] are shown in red. The closest H $\cdots$ C_{benzene} distance is shown in blue.

approach^[10,13] was then used to examine the effect of complexation by an aromatic group, modelled as a benzene. The position of the benzene molecule was fully optimized in all three systems (see Figure S7). This optimization was initiated with the benzene near to the C17 methyl group of **A**, with one aromatic carbon atom held in line with the C8–C17 bond of **A** at a distance of 4.05 Å as in the case of the protonated GGPP (to prevent the benzene from wandering towards the distal carbocation center, something which would not be possible in the enzyme). These calculations (carried out with several different theoretical methods) indicate that **A** can be selectively stabilized (versus protonated GGPP) by 3–5 kcal mol^{−1}, the carbocation with partial C–C bonds can be selectively stabilized by 1–3 kcal mol^{−1}, and the basicity of the π bond, in GGPP, which is protonated increases by about 1 kcal mol^{−1} (see Figure S7). While these results are preliminary, in the sense that alternative orientations of the components of the substrate–theozyme complex have not yet been examined, and the remainder of the enzyme active site is not included, they nevertheless show that the presence of an aromatic group in the catalytic base group can promote cyclization in both a thermodynamic and kinetic sense.

Formation of the fully rearranged **3** almost certainly reflects the intrinsic reactivity of **A**. Consistent with previous studies demonstrating that addition of a C=C to a carbocation to form a cyclohexane is quite exothermic,^[14] the quantum chemical calculations reported herein indicate that bicyclization is highly exothermic (by > 30 kcal mol^{−1}). Accordingly, the formation of **A** is expected to be irreversible, but its *trans*-decalin geometry also permits a series of antiparallel 1,2-hydride and methyl migrations via tertiary carbocations (Scheme 1).

These rearrangements were probed by quantum chemical calculations. The relevant intermediates begin with **A**, which undergoes 1,2-hydride migration to form the corresponding 9-yl⁺ **B**, with subsequent 1,2-methyl migration forming the *ent*-halima-13E-en-10-yl⁺ **C**. Another 1,2-hydride migration forms the corresponding 5-yl⁺ **D**, and subsequent 1,2-methyl migration forms the fully rearranged *ent*-clerodan-13E-en-4-yl⁺ (**E**). Notably, quantum chemical calculations predict that the barriers for the first H/Me/H 1,2-shifts are 2–7 kcal mol^{−1}, with only small differences in energy for the corresponding intermediates (i.e., **A**–**D**), while there is a substantial energetic barrier of nearly 15 kcal mol^{−1} for the final Me shift, and formation of **E** is highly endothermic, approximately +12 kcal mol^{−1} relative to that of **D** (Figure 4; see also

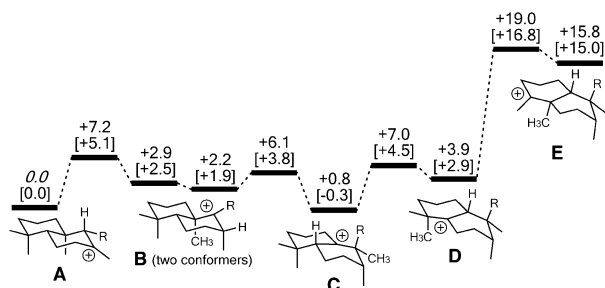


Figure 4. Energies of carbocation minima and transition-state structures involved in hydride and methyl shifts en route to **3**. Relative energies (B3LYP/6-31 + G(d,p) and mPW1PW91/6-31 + G(d,p))//B3LYP/6-31 + G(d,p) [in brackets] are shown in kcal mol^{−1}; in this model, R = (CH₂)₂C(Me)=CHCH₂OPO₃H₂.

Figure S8). Intriguingly, these results suggest that **A**–**D** exist in equilibrium, with formation of the fully rearranged **3** critically dependent on the absence of any other suitably located general base in the AtCPS active site other than D379, the residue which acted as the general base to initiate the overall reaction by protonation, and which presumably quenches **E** by removal of that same proton.

The ability of tyrosine substitution for the catalytic histidine to promote rearrangement was further probed by quantum chemical calculations, again using a theozyme-based approach. Here, the tyrosine was modelled as phenol, with hydrogen bonding to a formamide model of the asparagine, and the hydroxy group positioned near the C8 methyl group which would be deprotonated in formation of the usual product **2** (Figure 5). In this arrangement, these functional groups remove the barrier to rearrangement, and provide selective stabilization for **B**, such that the initial 1,2-hydride shift becomes exothermic (by 4–5 kcal mol^{−1}). Again, these results are preliminary in the sense that alternative orientations of the components of the substrate–theozyme complex have not yet been examined, but nevertheless clearly show that the hydride shift can be promoted by the functional groups present in the mutant.

Although the structurally and mechanistically analogous triterpene cyclases also often catalyze rearrangement reactions similar to those observed here, the underlying mecha-

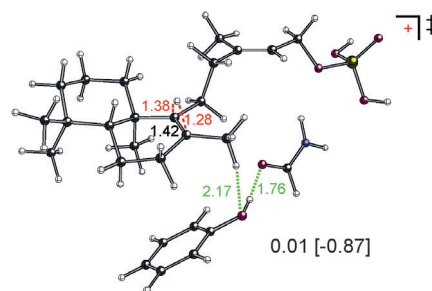


Figure 5. The barrier for the 1,2-hydride shift which initiates rearrangement of *ent*-labda-13E-8-yl⁺ (**A**) is removed in the presence of a Tyr/Asn model (computed energies of the transition state, corresponding to the TS between **A** and **B1**) in Figure 4 (see also Figure S8), relative to **A** in kcal mol^{−1} with B3LYP/6-31 + G(d,p) and mPW1PW91/6-31 + G(d,p)//B3LYP/6-31 + G(d,p) [within brackets] are shown, as are distances in Å; see the Supporting Information for details).

nism remains unknown for these enzymes as well. It has been suggested that triterpene rearrangement may depend on organization of a π -electron gradient and/or positioning of the catalytic base.^[15] Our results suggest that such rearrangement largely relies on the latter, that is, positioning of the catalytic base, since replacement of the histidine, in the catalytic base group in AtCPS, with a larger aromatic residue, particularly tyrosine, leads to predominant formation of the fully rearranged **3**. Significantly, the results reported here provide even further insight into such catalysis. Specifically, bicyclization is enabled by the ability of such aromatic substitution to enforce correct folding of **1** within the active site, yet block deprotonation, while also providing some selective stabilization of the initial bicyclic intermediate **A** and other structures along the reaction coordinate for protonation/bicyclization through carbocation– π interactions. The exothermic nature of bicyclization of **1** to **A** then leads, in the absence of a suitable base, to a series of 1,2-hydride and methyl shifts to the fully rearranged intermediate **E**, wherein the carbocation is proximal to the middle Asp of the DxDD motif, the residue which acted as the catalytic acid, and which seems likely to be the only remaining general base.

Finally, it should be noted that the AtCPS:H263Y mutant appears to represent the first identified (–)-kolavenyl diphosphate synthase, thus offering potential access to labdane-related diterpenoids requiring such rearrangement. For example, a number of diterpenoids derived from **3** are of pharmaceutical interest, such as salvinorin A, a known hallucinogen that, along with salvinicins A and B, are potential treatments for Alzheimer's disease. The first two operate as κ -opioid agonists, while the latter is a μ -opioid antagonist. In addition, barbatins A–C showed significant cytotoxic effects against several cancer lines.^[16,17] Indeed, even (–)-kolavenol itself has been shown to exhibit antitumor activity.^[18] Finally, the ability of this single-residue change (requiring change of only a single nucleotide) to lead to such a dramatic change in product outcome highlights the plasticity of class II diterpene cyclases, and presumably underlies the observed diversification of LRD natural products.

Acknowledgements

We would like to thank the W. M. Keck Metabolomics Research Laboratory for the generous use of their analytical equipment and facility. This work was supported by a grant from the NIH (GM076324 to R.J.P.) and NSF (CHE-0957416 and CHE-030089 [supercomputing resources via XSEDE] to D.J.T.).

Keywords: biosynthesis · cyclizations · natural products · reaction mechanisms · rearrangements

How to cite: *Angew. Chem. Int. Ed.* **2016**, *55*, 634–638
Angew. Chem. **2016**, *128*, 644–648

-
- [1] R. J. Peters, *Nat. Prod. Rep.* **2010**, *27*, 1521–1530.
[2] J. Zi, S. Mafu, R. J. Peters, *Annu. Rev. Plant Biol.* **2014**, *65*, 259–286.
[3] a) J. Criswell, K. Potter, F. Shephard, M. B. Beale, R. J. Peters, *Org. Lett.* **2012**, *14*, 5828–5831; b) S. Mafu, K. C. Potter, M. L. Hillwig, S. Schulte, J. Criswell, R. J. Peters, *Chem. Commun.* **2015**, *51*, 13485–13487; c) K. Potter, J. Criswell, R. J. Peters, *Angew. Chem. Int. Ed.* **2014**, *53*, 7198–7202; *Angew. Chem.* **2014**, *126*, 7326–7330.
[4] S. Prisic, J. Xu, R. M. Coates, R. J. Peters, *ChemBioChem* **2007**, *8*, 869–874.
[5] a) Y. Hamano, Y. Kuzuyama, N. Itoh, K. Furihata, H. Seto, T. Dai, *J. Biol. Chem.* **2002**, *277*, 37098–37104; b) C. Nakano, T. Okamura, T. Sato, T. Dai, T. Hoshino, *Chem. Commun.* **2005**, 1016–1018.
[6] a) M. Köksal, K. Potter, R. J. Peters, D. W. Christianson, *Biochim. Biophys. Acta Gen. Subj.* **2014**, *1840*, 184–190; b) M. Köksal, H. Hu, R. M. Coates, R. J. Peters, D. W. Christianson, *Nat. Chem. Biol.* **2011**, *7*, 431–433.
[7] A. Cyr, P. R. Wilderman, M. Determan, R. J. Peters, *J. Am. Chem. Soc.* **2007**, *129*, 6684–6685.
[8] A. Lozama, T. E. Prisinzano, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5490–5495.
[9] a) R. Misra, R. C. Pandey, S. Dev, *Tetrahedron* **1979**, *35*, 985–987; b) T. Hubert, D. Wiemer, *Phytochemistry* **1985**, *24*, 1197–1198; c) H. Monti, N. Tiliacos, R. Faure, *Phytochemistry* **1999**, *51*, 1013–1015.
[10] Y. J. Hong, D. J. Tantillo, *Chem. Sci.* **2013**, *4*, 2512–2518.
[11] D. W. Christianson, *Chem. Rev.* **2006**, *106*, 3412–3442.
[12] a) L. Zu, M. Xu, M. W. Lodewyk, D. E. Cane, R. J. Peters, D. J. Tantillo, *J. Am. Chem. Soc.* **2012**, *134*, 11369–11371; b) D. J. Tantillo, *Nat. Prod. Rep.* **2011**, *28*, 1035–1053.
[13] Y. J. Hong, D. J. Tantillo, *J. Org. Chem.* **2007**, *72*, 8877–8881.
[14] C. Jenson, W. L. Jorgensen, *J. Am. Chem. Soc.* **1997**, *119*, 10846–10854.
[15] K. U. Wendt, *Angew. Chem. Int. Ed.* **2005**, *44*, 3966–3971; *Angew. Chem.* **2005**, *117*, 4032–4037.
[16] O. Shiota, K. Nagamatsu, S. Sekita, *J. Nat. Prod.* **2006**, *69*, 1782–1786.
[17] S. J. Dai, J. Y. Tao, K. Liu, Y. T. Jiang, L. Shen, *Phytochemistry* **2006**, *67*, 1326–1330.
[18] A. Ohsaki, Y. Lu, S. Ito, H. Edatsugi, D. Iwata, Y. Komoda, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2889–2892.
-

Received: September 28, 2015

Published online: November 25, 2015