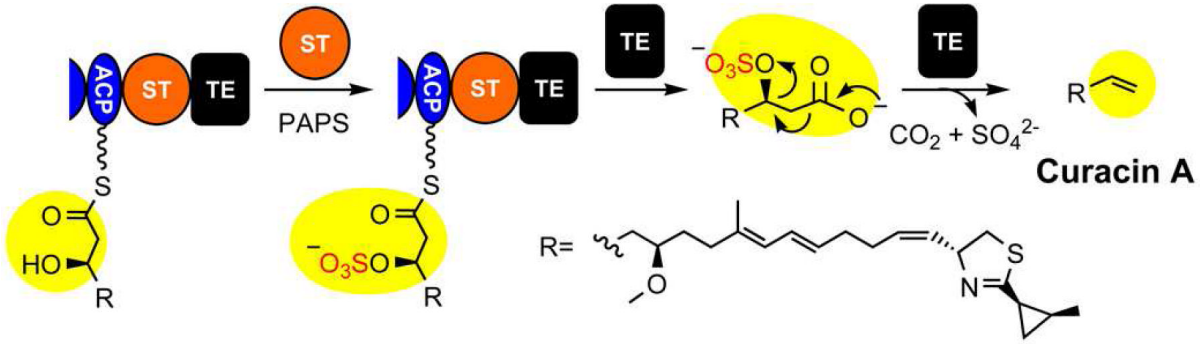
Published in final edited form as:
J Am Chem Soc. 2009 November 11; 131(44): 16033–16035. doi:10.1021/ja9071578.

Polyketide Decarboxylative Chain Termination Preceded by *O*-Sulfonation in Curacin A Biosynthesis

Liangcai Gu[†,¶], **Bo Wang**[†], **Amol Kulkarnid**[‡], **Jennifer J. Gehret**[†], **Kayla R. Lloyd**[‡], **Lena Gerwick**[§], **William H. Gerwick**[§], **Peter Wipf**[‡], **Kristina Håkansson**[†], **Janet L. Smith**[†], and **David H. Sherman**[†]

Life Sciences Institute, Department of Medicinal Chemistry, Department of Chemistry, Department of Biological Chemistry, and Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109, Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, and Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093

Abstract



Biosynthetic innovation in natural product systems is driven by the recruitment of new genes and enzymes into these complex pathways. Here, an unprecedented decarboxylative chain termination mechanism is described for the polyketide synthase of curacin A, an anticancer lead compound isolated from the marine cyanobacterium *Lyngbya majuscula*. The unusual chain termination module containing adjacent sulfotransferase (ST) and thioesterase (TE) catalytic domains embedded in CurM was biochemically characterized. The TE was proved to catalyze a hydrolytic chain release of the polyketide chain elongation intermediate. Moreover, a selective ST-mediated sulfonation of the (R)-β-hydroxyl group was found to precede TE-mediated hydrolysis, triggering a successive decarboxylative elimination and resulting in the formation of a rare terminal olefin in the final metabolite.

Curacin A (Figure 1A), a marine cyanobacterial metabolite isolated from marine cyanobacterium *Lyngbya majuscula*, is a mixed-polyketide nonribosomal-peptide natural product with potent anticancer activities.[1] The biosynthetic pathway generates a series of intermediates with increasing hydrophobicity leading to the final product.[2] Accordingly,

davidhs@umich.edu .
[†]University of Michigan
[‡]University of Pittsburgh
[§]University of California, San Diego
[¶]Current address: Department of Genetics, Harvard Medical School, 77 Ave Louis Pasteur, NRB 232, Boston, MA 02115.

Supporting Information Available: Experimental details including Chemical Synthesis, and Materials & Methods, and supplementary data. This material is available free of charge via the Internet at http://pubs.acs.org.

several types of decarboxylases have been identified to catalyze release of eleven CO_2 molecules during curacin A biosynthesis.[3-7] In addition, polar functional groups formed in the chain elongation intermediates (e.g., keto and hydroxyl groups) are eliminated by the coupled β-keto reduction/β-hydroxyl dehydration, HMG polyketide β-branching,[4,5] and cyclopropanation[5] reactions, with the only remaining hydroxyl group modified as the methyl ether derivative. Moreover, as a linear polyketide, curacin A contains an unusual hydrophobic terminal olefin instead of a typical terminal carboxyl, aldehyde, or alcohol group.[8] Annotation of the pathway revealed an unprecedented chain termination module in which a putative sulfotransferase (ST) domain is flanked by an acyl carrier protein (ACP) domain and a thioesterase (TE) domain[2]. Notably, queries of the public database have revealed a number of PKS-derived pathways deposited from genome sequencing projects that also contain homologous ACP-ST-TE genes, but none of their products have been identified. These pathways might be readily assigned to specific metabolites that contain a terminal olefin, such as kalkitoxin from *L. majuscula*,[9] and haliangicin from *Haliangium luteum*,[10] if a biosynthetic correlation between the terminal olefin and the ACP-ST-TE module is established.

STs catalyze transfer of a sulfonate group from the donor, adenosine 3′-phosphate 5′-phosphosulfate (PAPS), to a hydroxyl or amine group on receptors ranging from proteins to small molecules.[11] They are involved in a broad range of biological processes, such as detoxification, hormone regulation, drug metabolism, signaling pathways, and others.[11] Typically, a polyketide synthase or non-ribosomal peptide synthetase TE catalyzes a chain termination process by releasing a full-length acyl, peptidyl, or hybrid chain from the carrier protein phosphopantetheine thiol group as a free acid, or employs an intramolecular nucleophile (e.g., hydroxyl or amine group) to generate a macrolactone or macrolactam product.[12-14] Based on the predicted full-length intermediate tethered to the CurM ACP (Figure 1A), a series of reactions including thioester hydrolysis, decarboxylation and dehydration are required to form the terminal olefin moiety. According to the canonical functions of ST and TE, the CurM ST was predicted to transfer a sulfonate group to the β-hydroxyl of the full-length intermediate, and the CurM TE was presumed to release the intermediate from the ACP as a free acid. However, how these enzymes are coordinated to render the terminal olefin formation has remained elusive. Thus, we were motivated to pursue biochemical studies to elucidate the mechanisms of this novel chain termination process.

Our bioinformatics analysis and protein solubility tests suggested that the previously sequenced *curM* TE region[2] might have been partially replaced with exogenous DNA during construction of the genomic DNA library of *L. majuscula*. Thus, we sequenced *curM* and additional parts of the 3’ flanking region (GenBank accession no. GQ412749) of the *cur* cluster in another cosmid (pLM14), from the *L. majuscula* genomic DNA library[15] and compared the data with those from a *L. majuscula* genome sequencing project (unpublished data). It revealed that the 3’ end of the deposited *cur* gene cluster[2] starting from the middle of the *curM* TE region was indeed chimeric. The revised gene cluster (Figure 1A) lacks “*curN*”, and a complete TE domain is encoded by the 3’ end of *curM*.[2] The adjacent downstream genes show high homology to tRNA 2-selenouridine synthase (ATPase) and adenylate/guanylate cyclase (Cy), and are not likely involved in curacin biosynthesis. Thus, *curM* now appears to mark the 3’-end boundary of the *cur* gene cluster, and the ST-TE di-domain presumably functions as the chain termination module.

To assess biochemically the ST-TE–mediated chain termination process, we cloned and overexpressed CurM ACP, ST and TE as soluble single-domain constructs (Figure S1A). ST was eluted as a monomer and TE as a dimer from an analytical size-exclusion column (Figure S1B). ACP was overexpressed in the apo form in order to generate ACP-linked substrates. A simplified model substrate, 3-hydroxy-5-methoxytetradecanoyl-ACP (**1**-ACP, Figure 1B), was designed to mimic the full-length chain intermediate tethered to CurM ACP (Figure 1A). The

ACP-linked substrates were generated by loading the acyl-CoA substrates to the apo ACP by using the highly flexible *Streptomyces verticillus* Svp phosphopantetheinyltransferase.[16]

With the soluble enzymes and model substrates in hand, we first sought to test whether the CurM TE exhibits canonical hydrolysis activity to cleave the thioester bond (Figure 1B). (3*S*)-**1**-ACP and (3*R*)-**1**-ACP were prepared as the TE substrates to test its stereoselectivity for the β-hydroxyl group. The reactions were analyzed by reverse-phase HPLC, and the separated ACP fractions were examined by Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) and infrared multiphoton dissociation (IRMPD)[17] techniques. In addition, the chain-release products were detected by LC-MS and confirmed by co-injection with authentic standards. Both of the acyl groups were found to be hydrolyzed from (3*S*)-**1**-ACP and (3*R*)-**1**-ACP with low efficiency (Figure 2A ii, 2B iii, 2C ii and Figure S2A). However, TE-catalyzed hydrolysis of (3*S*)-**1**-ACP was ~5-fold faster than with (3*R*)-**1**-ACP (Figure S2A). In addition, steady-state kinetic analysis using the CoA-linked substrates confirmed that (3*S*)-**1** is slightly preferred for TE hydrolysis (Figure S2B). However, based on bioinformatic analysis of PKS ketoreductase (KR) sequences,[18] the CurM KR was predicted to reduce the β-keto to an (*R*)-β-hydroxyl group in the full-length chain intermediate (see Materials and Methods in Supporting Information). The resolution of this paradox is described below.

Based on the known mechanism of ST enzyme function, CurM ST was predicted to bind to PAPS, and transfer a sulfonate moiety to the β-hydroxyl group of the intermediate tethered to or released from CurM ACP (Figure 1B). First, we sought to identify the ST substrate by testing the sulfonation of **1** and **1**-ACP. The sulfonation of **1** was analyzed by LC-MS, and that of **1**-ACP was examined by HPLC and FTICR-MS. No substrate conversion was observed for **1** in the presence of ST and PAPS. In contrast, complete conversion of (3*R*)-**1**-ACP was determined by HPLC (Figure 2A iii), and the product with an 80-dalton mass addition was confirmed to be (3*R*)-**2**-ACP by FTICR-MS (Figure 2B ii). Notably, for (3*R*)-**1**-ACP substrate, ST-catalyzed sulfonation was ~1000-fold faster than TE hydrolysis. Likewise, the stereoselectivity of ST for the β-hydroxyl group was determined by comparing the sulfonation efficiency of (3*S*)-**1**-ACP and (3*R*)-**1**-ACP. Under the same reaction conditions, ST sulfonation with (3*R*)-**1**-ACP was ~2.5-fold faster than with (3*S*)-**1**-ACP (Figure S3). Thus, CurM ST stereoselectivity is consistent with the predicted formation of a (*R*)-β-hydroxyl group by CurM KR. The catalytic efficiency and substrate preference of ST and TE revealed in these experiments provide convincing evidence that ST sulfonation precedes the TE hydrolysis step.

Next, we sought to examine the TE-catalyzed hydrolysis after the sulfonation by coupling the two reactions in one pot. When (3*R*)-**1**-ACP was treated with both ST and TE, complete release of the acyl chain from CurM ACP was observed by HPLC (Figure 2A *iv*), and confirmed by FTICR-MS (Figure 2B *iii*) and IRMPD. We found under the same reaction conditions that the TE-catalyzed hydrolysis reaction with sulfonated product (3*R*)-**2**-ACP was ~800-fold faster than with (3*R*)-**1**-ACP. Moreover, TE-catalyzed hydrolysis of (3*R*)-**2**-ACP is ~120-fold faster than (3*S*)-**2**-ACP (Figure S4). These results further corroborated the timing and stereoselectivity of the ST sulfonation and TE hydrolysis steps.

Finally, we assessed whether a proposed sulfonated intermediate (3*R*)-**2** (Figure 1B) is released following hydrolysis of (3*R*)-**2**-ACP, or whether formation of **3** (Figure 1B) occurs directly by a coordinated decarboxylative elimination process. First, the ST-TE coupled reactions with (3*R*)-**1**-ACP were analyzed by LC-MS and a single product was confirmed to be (3*R*)-**2** by co-injection with the corresponding authentic standard (Figure 2C *iii*). Due to the expected volatility of **3**, the contents of the coupled reaction were extracted with hexane and analyzed by GC-MS. A second product was thus detected and confirmed to be **3** by co-injection with the corresponding authentic standard (Figure 2D *iv*). We found that **2** was relatively stable in the reaction mixture, and was not spontaneously converted to **3** (Figure 2C *iii*). To determine

which enzyme catalyzes the decarboxylative elimination of **2**, authentic (i.e. synthetic) (3*R*)-**2** was separately treated with CurM ST or TE, and a small amount of **3** was observed only for the TE reaction (estimated ~2% conversion in 1 h). Under the same conditions, the TE reaction with (3*R*)-**2**-ACP leading to **3** was more efficient than with (3*R*)-**2** (estimated ~20% conversion in 5 min), suggesting that the decarboxylative elimination is likely coupled with hydrolysis in the TE active site. The significant level of (3*R*)-**2** as a product of the TE reaction from model substrate (3*R*)-**1**-ACP suggests that some interactions between the native substrate and the TE active site might be essential for efficient coupling of the hydrolysis and decarboxylative elimination reactions.

In summary, the ST-TE–mediated chain termination represents a remarkable example of metabolic innovation via recombination and co-evolution of multifunctional enzymes in biosynthetic pathways. Insertion of a functional ST domain into the CurM PKS chain termination module has resulted in a unique process for polyketide chain release. Specifically, these catalytic events transform the β-hydroxyl of the penultimate chain elongation intermediate into a β-sulfate, an excellent leaving group that is positioned chemically to facilitate decarboxylative elimination in the presence of the terminal carboxylate following TE-mediated hydrolysis of the acyl-thioester (Scheme 1A). A closely related strategy of terminal olefin formation occurs in the mevalonate pathway for isoprenoid biosynthesis.[19] Likewise, a mevalonate-5-diphosphate decarboxylase (MDD) catalyzes a decarboxylative elimination reaction by first converting a β-hydroxyl group into a phosphate leaving group (Scheme 2B). In contrast, introduction of a terminal olefin as the final step in tautomycetin biosynthesis, a polyketide metabolite isolated from *Streptomyces* sp. CK4412, is more likely generated by the action of a discrete decarboxylase via yet another unique mechanism.[20] Given the substrate tolerance of the ST-TE module, it likely represents a useful new metabolic engineering tool for deliberate incorporation of terminal olefins into high value natural products and other long chain hydrocarbons.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Sergei Chernyak and Kathleen Noon for mass spectrometry assistance. This work was supported by NIH grants GM076477 and the Hans W. Vahlteich Professorship (to D.H.S.), Rackham Predoctoral Fellowship (to L.G.), and NSF career award CHE-05-47699 (to K.H.).

References

1. Verdier-Pinard P, Lai J, Yoo HD, Yu J, Marquez B, Nagle DG, Nambu M, White JD, Falck JR, Gerwick WH, Day BW, Hamel E. Mol. Pharm 1998;53:62–76.
2. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH. J. Nat. Prod 2004;67:1356–1367. [PubMed: 15332855]
3. Gu L, Geders TW, Wang B, Gerwick WH, Hakansson K, Smith JL, Sherman DH. Science 2007;318:970–974. [PubMed: 17991863]
4. Gu L, Jia J, Liu H, Hakansson K, Gerwick WH, Sherman DH. J. Am. Chem. Soc 2006;128:9014–9015. [PubMed: 16834357]
5. Gu L, Wang B, Kulkarni A, Geders TW, Grindberg RV, Gerwick L, Hakansson K, Wipf P, Smith JL, Gerwick WH, Sherman DH. Nature 2009;459:731–735. [PubMed: 19494914]
6. Calderone CT, Iwig DF, Dorrestein PC, Kelleher NL, Walsh CT. Chem. Biol 2007;14:835–846. [PubMed: 17656320]
7. Calderone CT, Kowtoniuk WE, Kelleher NL, Walsh CT, Dorrestein PC. Proc. Natl. Acad. Sci. U.S.A 2006;103:8977–8982. [PubMed: 16757561]

8. Walsh CT. Acc. Chem. Res 2008;41:4–10. [PubMed: 17506516]
9. LePage KT, Goeger D, Yokokawa F, Asano T, Shioiri T, Gerwick WH, Murray TF. Toxicol. Lett 2005;158:133–139. [PubMed: 16039402]
10. Fudou R, Iizuka T, Sato S, Ando T, Shimba N, Yamanaka S. J. Antibiot 2001;54:153–156. [PubMed: 11302488]
11. Chapman E, Best MD, Hanson SR, Wong CH. Angew. Chem. Int. Edit 2004;43:3526–3548.
12. Walsh CT. Science 2004;303:1805–1810. [PubMed: 15031493]
13. Fischbach MA, Walsh CT. Chem. Rev 2006;106:3468–3496. [PubMed: 16895337]
14. Khosla C, Tang Y, Chen A, Schnarr NA, Cane DE. Annu. Rev. Biochem 2007;76:11.1–11.27.
15. Chang Z, Flatt P, Gerwick WH, Nguyen VA, Willis CL, Sherman DH. Gene 2002;296:235–247. [PubMed: 12383521]
16. Sanchez C, Du L, Edwards DJ, Toney MD, Shen B. Chem. Biol 2001;8:725–738. [PubMed: 11451672]
17. Dorrestein PC, Bumpus SB, Calderone CT, Garneau-Tsodikova S, Aron ZD, Straight PD, Kolter R, Walsh CT, Kelleher NL. Biochemistry 2006;45:12756–12766. [PubMed: 17042494]
18. Caffrey P. Chem. Biol 2005;12:1060–1062. [PubMed: 16242648]
19. Bonanno JB, Edo C, Eswar N, Pieper U, Romanowski MJ, Ilyin V, Gerchman SE, Kycia H, Studier FW, Sali A, Burley SK. Proc. Natl. Acad. Sci. U.S.A 2001;98:12896–12901. [PubMed: 11698677]
20. Choi SS, Hur YA, Sherman DH, Kim ES. Microbiol.-SGM 2007;153:1095–1102.

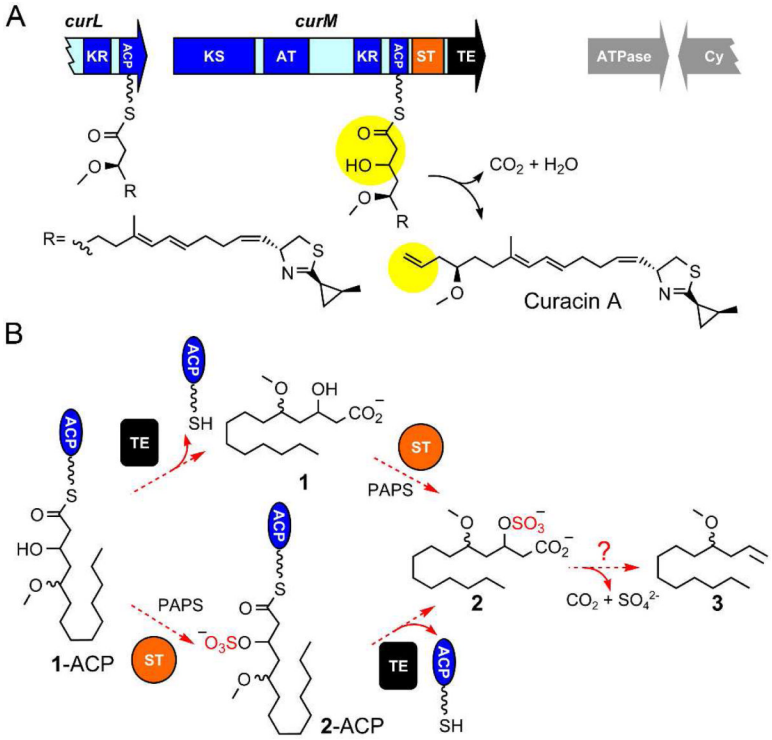


Figure 1.

Chain termination module of the PKS in curacin A biosynthesis. (A) Annotation of the 3' end of the *cur* gene cluster. ACP, acyl carrier protein; AT, acyltransferase; ATPase, putative tRNA 2-selenouridine synthase; Cy, putative adenylate/guanylate cyclase; KR, ketoreductase; KS, ketosynthase; ST, sulfotransferase; TE, thioesterase. (B) Proposed mechanisms with the model substrates for the decarboxylative chain termination leading to product with a terminal olefin.

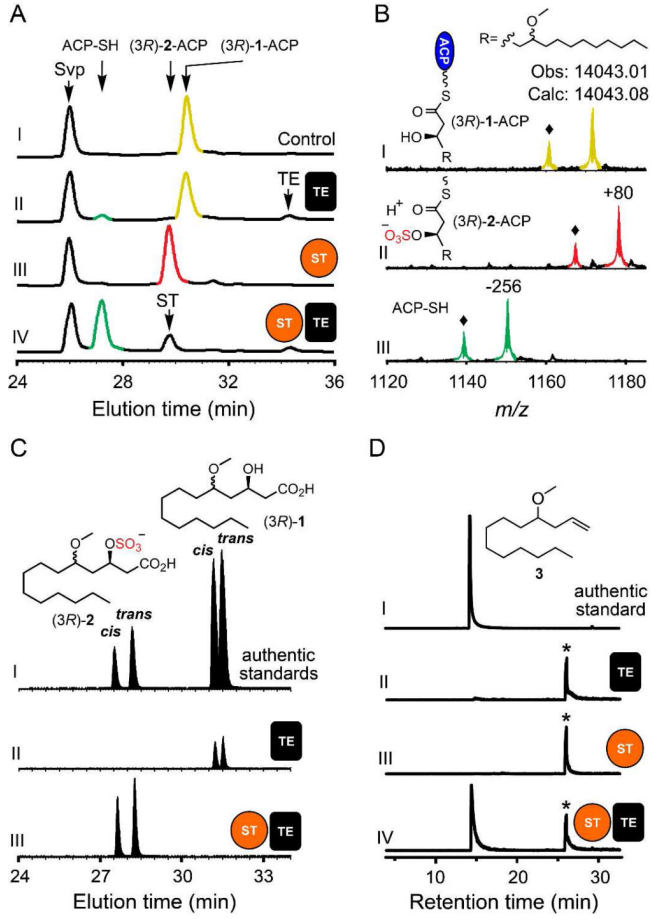
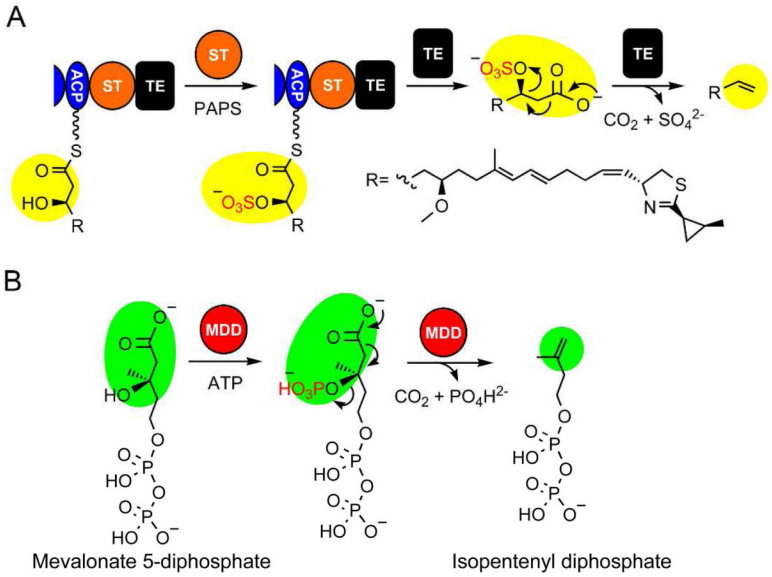


Figure 2.

Analysis of ST and TE reactions with (3*R*)-**1**-ACP substrate. (A) HPLC analysis of ST and TE reactions. ~300 μM (3*R*)-**1**-ACP was added with 2 μM TE, ST or both, and 2 mM PAPS in 50 mM Tris-HCl buffer (pH 7.0), and incubated at room temperature for 30 min. (B) Partial FTICR mass spectra (+12 charge state) for the ACP-linked substrate and products. The HPLC peaks of ACP species were collected for FTICR-MS analysis, and only one MS spectrum is shown for each ACP species. Both of the ACP species with and without an N-terminal methionine (indicated by black diamond) were detected. The ACP species with the 80-dalton mass addition corresponds to the protonated form of (3*R*)-**2**-ACP in gas-phase. (C) LC-MS chromatograms (273.2 and 353.1 mass range) showing products of (3*R*)-**1**-ACP treated with TE, or both TE and ST. ~300 μM (3*R*)-**1**-ACP was treated with 2 μM TE, or both TE and ST in 50 mM Tris-HCl buffer (pH 7.0) for 1 h or 5 min, respectively. The enantiomers of (3*R*)-**1** and (3*R*)-**2** were able to be separated and assigned as *cis* and *trans*. (D) GC-MS chromatograms of ST and TE reactions. ~300 μM (3*R*)-**1**-ACP was added to 2 μM TE, ST or both, and 2 mM PAPS in 50 mM Tris-HCl buffer (pH 7.0), and incubated at room temperature for 30 min. The chromatograms were recorded at 60-420 atomic mass units, and the Y axis is ion abundance. Y axis scaling of the authentic standard was adjusted for comparison with the samples. Asterisks denote unidentified species from hexane extracts of the reaction mixture.



Scheme 1.

Decarboxylative elimination mechanisms and terminal olefin formation in (A) the curacin A pathway and (B) the mevalonate pathway.