HHS Public Access

Author manuscript

Chem. Author manuscript; available in PMC 2024 March 09.

Published in final edited form as:

Chem. 2023 March 09; 9(3): 698–708. doi:10.1016/j.chempr.2022.12.006.

First *trans*-eunicellane terpene synthase in bacteria

Zining Li[1], **Baofu Xu**[1], **Volga Kojasoy**[2], **Teresa Ortega**[2], **Donovon A. Adpressa**[3], **Wenbo Ning**[1], **Xiuting Wei**[1], **Jamin Liu**[1], **Dean J. Tantillo**[2], **Sandra Loesgen**[1,4], **Jeffrey D. Rudolf**[1,5,*]

[1]Department of Chemistry, University of Florida, Gainesville, FL, United States

[2]Department of Chemistry, University of California–Davis, Davis, CA, United States

[3]Loxo Oncology, Inc., Louisville, CO, United States

[4]Whitney Laboratory for Marine Biosciences, University of Florida, St. Augustine, FL, United States

[5]Lead contact

SUMMARY

Terpenoids are the largest family of natural products, but prokaryotes are vastly underrepresented in this chemical space. However, genomic*s* supports vast untapped biosynthetic potential for terpenoids in bacteria. We discovered the first *trans*-eunicellane terpene synthase (TS), AlbS from *Streptomyces albireticuli* NRRL B-1670, in nature. Mutagenesis, deuterium labeling studies, and quantum chemical calculations provided extensive support for its cyclization mechanism. In addition, parallel stereospecific labeling studies with Bnd4, a *cis*-eunicellane TS, revealed a key mechanistic distinction between these two enzymes. AlbS highlights bacteria as a valuable source of novel terpenoids, expands our understanding of the eunicellane family of natural products and the enzymes that biosynthesize them, and provides a model system to address fundamental questions about the chemistry of 6,10-bicyclic ring systems.

Graphical Abstract

*Correspondence: jrudolf@chem.ufl.edu.

AUTHOR CONTRIBUTIONS

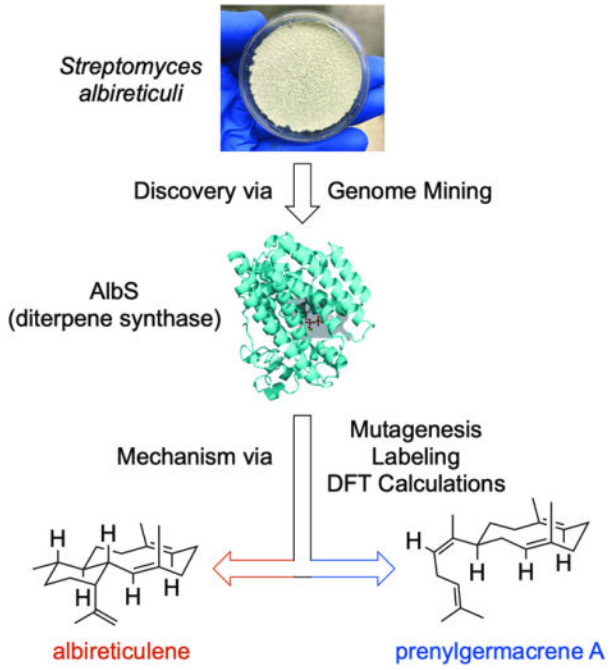
Conceptualization, J.D.R.; methodology, Z.L., B.X., D.A.A., D.J.T., S.L., and J.D.R.; investigation, Z.L., B.X., V.K., T.O., D.A.A., W.N., X.W., and J.L.; writing – original draft: Z.L. and J.D.R.; writing – review & editing, Z.L., B.X., V.K., W.N., X.W., D.J.T., S.L., and J.D.R.; visualization, Z.L. and J.D.R.; supervision, J.D.R.; funding acquisition, D.J.T., S.L., and J.D.R.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online.

DECLARATION OF INTERESTS

The authors declare no competing interests.



eTOC Blurb

Terpenoids, the largest and most structurally diverse class of natural products, are widely distributed in nature but less commonly reported in bacteria. Here, the first *trans*-eunicellane terpene synthase, AlbS, is discovered by genome mining in bacteria. Extensive evidence of mutagenesis, isotope-labeling studies, and DFT calculations reveal its cyclization mechanism. AlbS and the novel skeleton of its enzymatic product highlight bacteria as a valuable source of terpenoids and expand our understanding of the eunicellane family of natural products and their biosyntheses.

Keywords

Bacterial terpenoids; terpene synthase; enzymes; mechanism; genome mining; isotope labeling; quantum chemical calculations; diterpenoid; eunicellane

INTRODUCTION

Terpenoids, the largest and most structurally diverse class of natural products with over 80,000 known compounds, are a rich reservoir of pharmaceuticals, vitamins, flavors, fragrances, and biofuels.[1–5] While the vast majority of terpenoids have been isolated from plants, fungi, or marine organisms, relatively few are of bacterial origin.[1] However, this contradicts the extraordinary biosynthetic potential for terpenoid biosynthesis seen in the genomes of bacteria.[1,6–8] Terpene synthases (TSs), the enzymes that construct the hydrocarbon skeletons of terpenoids from prenyl diphosphates and are often responsible for the first step in their respective biosynthetic pathways, are widely distributed in bacteria.[1,9] Both canonical type I and type II TSs as well as about a dozen non-canonical TS families

are known to exist in prokaryotes.[10–12] However, most of these enzymes have not been characterized and products from their biosynthetic gene clusters (BGCs) remain enigmatic.

The eunicellanes are a unique family of diterpenoids that possess a 6/10-bicyclic skeleton (Fig. 1A). Many members of this family have diverse biological activities,[13,14] some of which have promising clinical applications such as eleutherobin and the sarcodictyins exhibiting potent antitumor activities via the paclitaxel-like mechanism of inducing tubulin polymerization and microtubule stabilization.[14–16] Most eunicellane diterpenoids have been found in soft corals;[13] there are only a few examples from plants[17,18] and most recently bacteria.[19,20] These natural products are highly oxidized and frequently possess at least one transannular ether bridge. One structural consistency in almost all eunicellanes is the presence of a *cis* ring fusion. Only six of more than 360 known eunicellanes contain a *trans* ring fusion: the plant-derived magdalenic acid and the magdalenic acid-containing diterpene dimers bisyinshanic acids,[17,18] two cytotoxic bacterial microeunicellols,[19] and the antibacterial coral-derived eunicellol A (Fig. 1A).[21]

Until recently, there were no biosynthetic studies on the eunicellane diterpenoids and no TSs known to produce their 6/10-bicyclic diterpene skeleton. In conjunction with our discovery of the antibacterial diterpenoid benditerpenoic acid from *Streptomyces* sp. (CL12–4), we identified the first TS and BGC responsible for eunicellane formation and functionalization.[20] Mechanistic studies of this TS, Bnd4, revealed that *cis*-eunicellane formation likely occurs via initial 1,10-cyclization, followed by a 1,3-hydride shift, 1,14-cyclization, and deprotonation to yield benditerpetriene (**2**) (Fig. 1B).[22] Shortly after, two independent studies led to the identification of two coral TSs that produced klysimplexin R, a *cis*-eunicellane with a *Z*-configured alkene at C-2/C-3 and a hydroxyl at C-11.[23,24] Mechanistically, klysimplexin R is plausibly formed via initial 1,14-cyclization with isomerization of the $\Delta^{2,3}$ bond (likely via geranyllinalyl diphosphate, GLPP), followed by two sequential 1,2-hydride shifts, 1,10-cyclization, and cation quench by water (Fig. 1B).[23] This cyclization cascade is likely conserved throughout the coral eunicellanes given the nature of the *cis* ring fusion, 2,3-alkene configuration, and placement of the captured water or alkene that results from final cation quench. The mechanistic differences between the bacterial and coral *cis*-eunicellane synthases are not surprising given that these TSs do not share any appreciable sequence similarities.

Here, we report the identification and characterization of the first *trans*-eunicellane forming TS. Albireticulene synthase (AlbS), from *Streptomyces albireticuli* NRRL B-1670, was found by genome mining for TSs in bacteria. Using a series of spectroscopic and chemical techniques, we unambiguously determined that albireticulene (**1**) possesses a *trans*-eunicellane skeleton. Mutagenesis, deuterium labeling studies, and quantum chemical calculations provided extensive support for its cyclization mechanism. In addition, parallel stereospecific labeling studies with AlbS and Bnd4 revealed a mechanistic distinction between *cis* and *trans* eunicellane formation in bacteria.

RESULTS AND DISCUSSION

Genome mining for bacterial terpene synthases

Using a sequence similarity network (SSN)[25] of all bacterial protein sequences categorized as “terpene cyclase-like 2” proteins (IPR034686), we annotated the network with all characterized members (Figs. 2 and S1). At an e-value of 10^{-70}, which was chosen based on the best separation of TS subfamilies with known products, there are 43 clusters with at least one characterized TS and 102 clusters and 151 singletons that are uncharacterized. In this study, we selected a small cluster containing five homologues for functional analysis. Each of these TSs possessed the canonical type I DDxxD and NSE metal-binding motifs as well as the C-terminal WxxxxxRY motif (Fig. S2). The closest known homologue to this cluster of TSs is (−)-germacradien-4-ol synthase from *Streptomyces citricolor* (27%/38% sequence identity and similarity, respectively, over 81% coverage).[26] In addition, each of these TSs was encoded in genetic proximity to a polyprenyl synthase and at least one cytochrome P450 (Fig. 2), suggesting the genuine products of these BGCs are oxidized terpenoids.

Characterization of albireticulene synthase, a *trans*-eunicellane diterpene synthase

We cloned the gene encoding UniProt ID A0A2A2D8W5 from *Streptomyces albireticuli* NRRL B-1670, into *E. coli* for heterologous expression and protein purification (Tables S1–S3, Fig. S3). We tested A0A2A2D8W5, which previously showed C_5 and C_{10} prenylation activity with small molecule nucleophiles and was designated TS118,[27] for in vitro activity with farnesyl diphosphate (FPP) and geranylgeranyl diphosphate (GGPP). GC-MS analysis revealed that FPP was converted into five sesquiterpenes (Fig. 3AB). Analysis of the EI-MS fragmentation patterns and Kovats retention indices identified the products as β-elemene (**3**), shyobunol (**4**), elemol (**5**), *trans*-nerolidol (**6**), and *epi*-α-cadinol (**7**) (Figs. S4–S8). In the presence of GGPP, a single major product (**1**) was formed; production of **1** was also achieved in *E. coli* using our artificial GGPP production system (Fig. 3C).[28] When analyzed by GC-MS, **1** exhibited an M^+ peak at *m/z* of 272.17, which is consistent with a molecular formula of $C_{20}H_{32}$ (Fig. S9).

Compound **1** was isolated, and its structure was elucidated by 1D and 2D NMR analysis in benzene-d_6 and toluene-d_8 (SI, Tables S4 and S5, Figs. 3B and S10–S22). The HSQC spectrum displayed five methines, seven methylenes, and four methyls. HMBC correlations supported that two methines and one methylene are alkenes with Me-17, Me-19, and Me-20 connected to C-15/C-16, C-6/C-7, and C-2/C-3, respectively. The characteristic methyl doublet of Me-18 revealed it is connected to C-11. Key ^{1}H-^{1}H COSY correlations of H-1/H-10/H-11/H-12/H-13/H-14/H-1 assembled a cyclohexane ring system. Additional COSY correlations of H-8/H-9/H-10/H-1/H-2 and H-4/H-5/H-6, together with HMBC correlations between H-2 and C-3, H-4 and C-2, and Me-19 and both C-6 and C-8 revealed the existence of a 10-membered ring. IPAP-HMBC NMR[29,30] data displayed the $^3J_{(H\text{-}2,C\text{-}20)}$ and $^3J_{(H\text{-}6,C\text{-}19)}$ coupling constants to be approximately 8.9 and 8.5 Hz, respectively, indicating that both the C-2/C-3 and C-6/C-7 alkenes are *E* configured (Figs. 3D and S16–S17). Altogether, the NMR data supported that **1** possesses a 6/10-fused bicyclic eunicellane diterpene skeleton. Chemical exchange crosspeaks in the HSQC spectra at 258 K in toluene-d_8 revealed two major conformers in a ratio of 2.2:1 and suggested a flexible structure

(Fig. S12). Dynamic behavior in solution was similarly seen for the 10-membered ring in benditerpetriene (**2**).[20] Peak broadening and conformational flexibility posed a significant challenge to determine both the relative and absolute configuration of **1**.

Here, we turned to chemical modification to solve the configuration of **1**. First, we observed that **1** rapidly degraded in chloroform to two products **8** and **9**, in a ratio of 7:1 (SI, Fig. S23). Both gersemiene A (**8**) and gersemiene B (**9**), named after the structurally similar gersemiols,[21] bear a 6/6/6-tricyclic ring system (SI, Table S6, Figs. 3D and S24–S42) that is proposed to originate from protonation-initiated 2,7-cyclization and subsequent deprotonation. In **8**, a 1D TOCSY spectrum with selective excitation of H-14 exhibited magnetization transfer for H-1, H-2, H-10, and H-11. H-1, H-2, H-10, and H-11 are each axial given that their vicinal coupling constants are all ~11 Hz ($^3J_{(H\text{-}2,H\text{-}1)}$ = 11.6 Hz; $^3J_{(H\text{-}1,H\text{-}10)}$ = 11.3 Hz; $^3J_{(H\text{-}10,H\text{-}11)}$ = 11.0 Hz). H-14 was equatorial based on its small ^{3}J coupling constants to H-1 (4.6 Hz) and H-13$_{ax}$ (5.8 Hz). Me-19 displayed stronger NOE effects to H-1 and H$_{ax}$-9 than to H-2 or H-11, supporting that Me-19 is axial and *cis* to H-1 (Fig. 3D). This gave the relative configuration of **8**, which was further supported by a series of IPAP-HMBC NMR experiments (Figs. S31–S33).

Next, inspired by the nonenzymatic conversion of **1** into **8** and **9**, we installed an oxygen atom at C-6 by oxidizing **1** with *m*-chloroperoxybenzoic acid (*m*CPBA). Epoxidation of the $\Delta^{6,7}$ bond in **1** led to **10**, NMR analysis of which led to the elucidation of the expected 6/6/6-tricyclic framework with a hydroxyl at C-6 (SI, Table S7, Figs. 3D and S43–S50). A strong NOESY correlation between H-2 and H-6 supported that the hydroxyl at C-6 was equatorial and Mosher ester analysis[31,32] unequivocally determined C-6 to be *S* configured (SI, Table S8, Figs. 3D and S51–S60). Together, the absolute configuration of **10** was assigned as 1*S*,2*R*,6*S*,7*S*,10*R*,11*S*,14*R*. Considering the relative configuration of **1** and the absolute configuration of **10**, the absolute configuration of **1** was thus unambiguously determined to be 1*R*,2*E*,6*E*,10*R*,11*S*,14*R*, a *trans*-eunicellane skeleton (Fig. 3B). Compound **1**, which was named albireticulene, is the C-1-epimer of **2**.

Mutagenesis reveals 10-membered ring formation occurs first in the AlbS mechanism

With AlbS and Bnd4 forming eunicellane epimers, we were interested in how these two TSs were different and how AlbS controls stereospecific cyclization of GGPP. AlbS and Bnd4 share 24%/41% identity and similarity, respectively, over 75% coverage (Fig. S2). In the absence of a protein structure, we utilized tFold to create a model of AlbS[33] and aligned it with our previous model of Bnd4 docked with GGPP (Fig. 4A).[22] As was expected for a type I di-TS, the D^{121}DxxD and N^{251}xxxSxxxE metal-binding motifs and the R^{342}Y diphosphate sensor are near the diphosphate moiety. Lining the active site of AlbS are seven aromatic amino acids, F94, H114, Y178, Y214, F329, and W336 and Y343 from the WxxxxxRY motif. There are only minor differences between the AlbS and Bnd4 models, namely the side chains of F94, H114, Y178, and F329 in AlbS are W67, T87, F162, and M309 in Bnd4, respectively (Fig. 4A).

Based on the AlbS model and the structural comparison with Bnd4, we conducted single-point mutation studies to determine which residues are essential for eunicellane formation

(Fig. 4A). Several mutations, including those made to D121, F94, H114, W336, and Y343 either created insoluble protein or completely abolished diterpene cyclization activity, suggesting essential roles in Mg^{2+} binding (D121) or protein solubility and catalytically competent active site (F94, H114, W336, Y343) for these residues (Figs. 4B and S61). Mutation of E185, which was modeled near the diphosphate moiety and is essential for catalysis in selinadiene synthase,[34] only moderately affected activity. An aromatic residue at Y178 is important for initial cyclization as Y178A only produced geranylgeraniol (GGOH, **13**), but Y178F fully restored cyclization activity. Similarly, hydrophobicity at residue 329 is important for folding and catalysis as F329A was insoluble but F329M retained 79% relative activity.

Finally, Y214 is implicated in playing an important role in eunicellane formation. When the homologous residue in Bnd4, Y197, was mutated to Ala, only GGOH was formed.[28] When Y214 in AlbS was mutated to Ala, Trp, His, Ile, or Leu, solubility was greatly affected. The Y214F mutant, however, was soluble and retained a minor level (4%) of **1** formation and produced a major new peak, **14** (Fig. 3C). Isolation and NMR analysis of **14** led us to identify it as a 10-membered monocyclic system with a *Z*-configured $\Delta^{11,12}$ bond (SI, Table S9, Figs. 3B and S62–S68). The ^{1}H, HSQC, and HMBC NMR spectra led us to identify at least two conformers of **14** present at 25 °C (Fig. S69).[35–38] Compound **14**, which we named prenylgermacrene A, is an isomer of both prenylgermacrene B and the *E*-configured eunicene A (Fig. S70).[39,40] When $AlbS^{Y214F}$ was incubated with FPP, both **3** and **6** were detected with **6** being the major product (Fig. 3A). The formation of **14**, evidently via the deprotonation of a monocyclic cationic intermediate, indicates that AlbS initially catalyzes 1,10-cyclization in the pathway to eunicellanes and that Y214 plays a role in stabilizing the monocyclic cationic intermediate and/or guiding the intermediate on the pathway to the 6,10-bicycle.

Isotope-labeling and quantum chemical calculations support the cyclization mechanism and differentiate *trans*- and *cis*-eunicellane terpene synthases

To elucidate the cyclization mechanism by which the 6/10-bicyclic eunicellane skeleton is formed by AlbS, we next performed stable isotope labeling experiments.[41] In comparison with unlabeled GGPP, 1,1-$^{2}H_2$-GGPP and chiral 1*R*-^{2}H-GGPP (Fig. S71) were incubated with AlbS. The deuterated enzymatic products were isolated, purified, and spectroscopically characterized using GC-MS and NMR. The M^+ peaks for the singly labeled and doubly labeled products matched the expected *m/z* values of 273 and 274, respectively (Fig. S9). Using the ^{1}H and HSQC spectra to visualize the protons at C-18 (δ_H = 0.98 ppm), C-1 (δ_H = 2.07 ppm), and C-11 (δ_H = 1.09 ppm), it was clear that the 1*R*-^{2}H in GGPP migrated from C-1 to C-11 in **1** while the 1*S*-^{2}H was retained on C-1 (Figs. 5A and S72). The lack of any observed scrambling of deuterium in the singly labeled reaction, along with the retention of the 2,3-*E*-alkene, suggests that the initial cyclization reaction may follow an S_N2-like mechanism (although an S_N1-like process where the enzyme prevents stereochemical scrambling is also possible). This labeling study also supports that a direct 1,3-hydride shift from C-1 to C-11, rather than two sequential 1,2-hydride shifts,[23] follows the initial 1,10-cyclization (Fig. 5B). The resulting cationic charge at C-1 would then be attacked by the C-14/C-15 π-bond (1,14-cyclization), which would be followed by deprotonation to

yield **1**. The production of sesquiterpenes **3**, **4**, **5**, and **7** by AlbS additionally supports an initial 10-membered cyclization as they all originate from the monocyclic germacradienyl cation (Fig. S73). In addition, both **4** and **7** follow cyclization with a 1,3-hydride shift, as seen in eunicellane cyclization. The proposed mechanism is different to the one proposed to be catalyzed by the coral eunicellane synthases but similar to the proposed mechanism of Bnd4 (Fig. 1B),[13,22,23] except for the stereocenter difference at C-1. To investigate this difference, we also incubated Bnd4 with 1*R*-^{2}H-GGPP. The product was found to be 1-^{2}H-**2**, the expected product of Bnd4 with the configuration at C-1 opposite to that seen in AlbS (Figs. 5A, 5C, and S74–S76). Overall, the data support that the active sites of AlbS and Bnd4 provide GGPP with similar molecular templates to form the same monocyclic intermediate, yet different enough to favor conformers disposed for 1,3-hydride transfer of the pro-*R* in AlbS and the pro-*S* in Bnd4.

To assess the energetic viability of our proposed mechanism and gain additional insight into the differences in stereochemical control by AlbS and Bnd4, we performed computational analyses with density functional theory calculations, mPW1PW91/6–31+G(d,p)//B3LYP/6–31+G(d,p)[42–48]. Calculated relative free energies of cationic intermediates and transition state structures indicated that the path shown in Fig. 5B is reasonable (Fig. 5D). After $\mathbf{A}^{+}$ is formed, the 1,3-hydride shift requires 11.1 kcal mol^{-1} and forms the allylic cation $\mathbf{B1}^{+}$ in a relatively deep well, 11.6 kcal mol^{-1} lower than the monocyclic intermediate $\mathbf{A}^{+}$ and 8.5 kcal mol^{-1} lower than the subsequent transition state for 1,14-cyclization ([**B**-**C**]‡). We propose that a conformational change from $\mathbf{B1}^{+}$ to $\mathbf{B2}^{+}$, which requires 2.8 kcal mol^{-1}, is necessary to allow the subsequent ring closure to $\mathbf{C}^{+}$. The *trans*-eunicellane cation $\mathbf{C}^{+}$ is not strongly energetically favored as it is higher in energy than $\mathbf{B1}^{+}$ (by 3.9 kcal mol^{-1}) or $\mathbf{B2}^{+}$ (by 1.1 kcal mol^{-1}). This contrasts with the calculated energies for the formation of *cis*-eunicellane by Bnd4, which followed a facile downhill pathway where the energies of $\mathbf{B}^{+}$ and the transition state for 1,14-cyclization were essentially equivalent.[22] Therefore, AlbS must prevent premature cation quench of the $\mathbf{B}^{+}$ intermediates and provide an energetically favorable deprotonation to form the *trans*-eunicellane skeleton. This scenario is reminiscent of that proposed previously for taxadiene synthase.[50]

Conclusion

Plants, fungi, and very recently marine animals are often the main sources of terpenoid natural product discovery and biosynthetic studies. Consequently, bacteria continue to be overshadowed, even though they are well regarded as producers of structurally unique and biologically active natural products. Using genome mining, we discovered a bacterial TS that constructs the rare *trans*-eunicellane skeleton, the first enzyme of its kind found in nature. AlbS highlights bacteria as a valuable source of novel terpenoids, expands our understanding of the eunicellane family of natural products, and provided a model system to address fundamental questions about the chemistry of 6/10-bicyclic ring systems.

Given our recent discovery and characterization of Bnd4,[20,22] we were in a unique position to compare the mechanisms of the two known bacterial eunicellane synthases, AlbS and Bnd4. In this study, we provided extensive support for the cyclization mechanism of AlbS and revealed atomic-level details of the stereospecific control of eunicellane

skeleton formation. Protein model-guided mutation studies and evaluation of enzyme products supported initial 1,10-cyclization followed by a 1,3-hydride transfer. Deuterium labeling studies confirmed the direct 1,3-hydride transfer from C-1 to C-11 using the pro-*R* hydrogen on GGPP. Quantum chemical calculations corroborated the cationic cascade and implicated a conformational change prior to the second cyclization step. Differences in primary sequence, active site residues, the stereospecificity of the 1,3-hydride shift, and the calculated energies and geometries of structures on the mechanistic pathway provide insight into why AlbS and Bnd4 selectively form eunicellane epimers. This study adds to a growing knowledge base of sequence-structure-function relationships that allow TSs to exquisitely control their carbocation-based cyclization reactions. Understanding these general principles will help the community engineer TSs to create tools for chemoenzymatic synthesis or synthetic biology.

The discovery and characterization of TSs such as AlbS provide access to synthetically challenging backbones, tools for synthetic biology and biocatalyst development, and optimism that many novel terpene skeletons and new natural products are awaiting to be found in bacteria via genome mining. Albireticulene is expected to be a precursor of a more tailored, complex natural product or starting point to a family of natural products. This hypothesis raises questions such as: (i) What is the genuine natural product(s) of this BGC? (ii) What types of biological activity or ecological function does the biosynthetic product have? and (iii) What types of enzymes can functionalize eunicellane skeletons?

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact—Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Jeffrey D. Rudolf (jrudolf@chem.ufl.edu).

Materials availability—Plasmids generated in this study will be made available on request. All other materials are commercially available or can be easily prepared as indicated.

Data and code availability—All data are included in the main text or supplemental information.

Present address—Baofu Xu: Shandong Laboratory of Yantai Drug Discovery, Bohai Rim Advanced Research Institute for Drug Discovery, Yantai, Shandong, China. State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China

Jamin Liu: New York University Tandon School of Engineering, Brooklyn, NY, United States

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was funded in part by NIH Grants R00 GM124461 and R35 GM142574 (to J.D.R.), NSF Grants CHE-1856416 and CHE-030089 (XSEDE) (to D.J.T.), and NSF Grant CHE-2020110 (to S.L.). We acknowledge the University of Florida Mass Spectrometry Research and Education Center (MSREC), which is supported by the NIH (S10 OD021758-01A1), and Jodie Johnson for GC-MS support. We thank Dr. Annika Jagels for HRESIMS analysis. We acknowledge the University of Florida Center for Nuclear Magnetic Resonance Spectroscopy and the University of Florida McKnight Brain Institute at the National High Magnetic Field Laboratory's Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) Facility, which is supported by the US NSF Cooperative Agreement No. DMR-1644779 and the State of Florida, for NMR support. We wish to thank James Rocca and Ion Ghiviriga for excellent NMR support.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

REFERENCES

1. Rudolf JD, Alsup TA, Xu B, and Li Z (2021). Bacterial terpenome. Nat. Prod. Rep 38, 905–980. 10.1039/d0np00066c. [PubMed: 33169126]
2. Avalos M, Garbeva P, Vader L, Wezel G.P. van, Dickschat JS, and Ulanova D (2022). Biosynthesis, evolution and ecology of microbial terpenoids. Nat. Prod. Rep 39, 249–272. 10.1039/D1NP00047K. [PubMed: 34612321]
3. Lange BM (2015). The evolution of plant secretory structures and emergence of terpenoid chemical diversity. Annu. Rev. Plant Biol 66, 139–159. 10.1146/annurevarplant-043014-114639. [PubMed: 25621517]
4. Huang M, Lu J-J, Huang M-Q, Bao J-L, Chen X-P, and Wang Y-T (2012). Terpenoids: natural products for cancer therapy. Expert Opin. Investig. Drugs 21, 1801–1818. 10.1517/13543784.2012.727395.
5. Mewalal R, Rai DK, Kainer D, Chen F, Külheim C, Peter GF, and Tuskan GA (2017). Plant-derived terpenes: A feedstock for specialty biofuels. Trends Biotechnol. 35, 227–240. 10.1016/j.tibtech.2016.08.003. [PubMed: 27622303]
6. Cane DE, and Ikeda H (2012). Exploration and mining of the bacterial terpenome. Acc. Chem. Res 45, 463–472. 10.1021/ar200198d. [PubMed: 22039990]
7. Kalkreuter E, Pan G, Cepeda AJ, and Shen B (2020). Targeting bacterial genomes for natural product discovery. Trends Pharmacol. Sci 41, 13–26. 10.1016/j.tips.2019.11.002. [PubMed: 31822352]
8. Kautsar SA, van der Hooft JJJ, de Ridder D, and Medema MH (2021). BiG-SLiCE: A highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters. GigaScience 10, giaa154. 10.1093/gigascience/giaa154. [PubMed: 33438731]
9. Yamada Y, Kuzuyama T, Komatsu M, Shin-ya K, Omura S, Cane DE, and Ikeda H (2015). Terpene synthases are widely distributed in bacteria. Proc. of the Natl. Acad. Sci. U. S. A 112, 857–862. 10.1073/pnas.1422108112.
10. Dickschat JS (2016). Bacterial terpene cyclases. Nat. Prod. Rep 33, 87–110. 10.1039/c5np00102a. [PubMed: 26563452]
11. Dickschat JS (2019). Bacterial diterpene biosynthesis. Angew. Chem., Int. Ed 58, 2–15. 10.1002/anie.201905312.
12. Rudolf JD, and Chang CY (2020). Terpene synthases in disguise: Enzymology, structure, and opportunities of non-canonical terpene synthases. Nat. Prod. Rep 37, 425–463. 10.1039/c9np00051h. [PubMed: 31650156]
13. Li G, Dickschat JS, and Guo YW (2020). Diving into the world of marine 2,11-cyclized cembranoids: A summary of new compounds and their biological activities. Nat. Prod. Rep 37, 1367–1383. 10.1039/d0np00016g. [PubMed: 32458945]
14. Gross H, and König GM (2006). Terpenoids from marine organisms: Unique structures and their pharmacological potential. Phytochem. Rev 5, 115–141. 10.1007/s11101-005-5464-3.

15. Lindel T, Jensen PR, Fenical W, Long BH, Casazza AM, Carboni J, and Fairchild CR (1997). Eleutherobin, a new cytotoxin that mimics paclitaxel (Taxol) by stabilizing microtubules. J. Am. Chem. Soc 119, 8744–8745. 10.1021/ja9717828.
16. Long BH, Carboni JM, Wasserman AJ, Cornell LA, Casazza AM, Jensen PR, Lindel T, Fenical W, and Fairchild CR (1998). Eleutherobin, a novel cytotoxic agent that induces tubulin polymerization, is similar to paclitaxel (Taxol). Cancer Res. 58, 1111–1115. [PubMed: 9515790]
17. Pinto AC, Pizzolatti MG, De Epifanio RA, Frankmölle W, and Fenical W (1997). The isolation of novel diterpenoids, including a C40 bis-diterpenoid, from the Brazilian plant Vellozia magdalenae (Velloziaceae). Tetrahedron 53, 2005–2012. 10.1016/S0040-4020(96)01183-0.
18. Zhang BY, Wang H, Luo XD, Du ZZ, Shen JW, Wu HF, and Zhang XF (2012). Bisyinshanic acids A and B, two novel diterpene dimers from the roots of Euphorbia yinshanica. Helv. Chim. Acta 95, 1672–1679. 10.1002/hlca.201200092.
19. Ma LF, Chen MJ, Liang DE, Shi LM, Ying YM, Shan WG, Li GQ, and Zhan ZJ (2020). Streptomyces albogriseolus SY67903 produces eunicellin diterpenoids structurally similar to terpenes of the gorgonian Muricella sibogae, the Bacterial Source. J. Nat. Prod 83, 1641–1645. 10.1021/acs.jnatprod.0c00147. [PubMed: 32367724]
20. Zhu C, Xu B, Adpressa DA, Rudolf JD, and Loesgen S (2021). Discovery and biosynthesis of a structurally dynamic antibacterial diterpenoid. Angew. Chem., Int. Ed 60, 14163–14170. 10.1002/anie.202102453.
21. Angulo-Preckler C, Genta-Jouve G, Mahajan N, de la Cruz M, de Pedro N, Reyes F, Iken K, Avila C, and Thomas OP (2016). Gersemiols A–C and eunicellol A, diterpenoids from the arctic soft coral Gersemia fruticosa. J. Nat. Prod 79, 1132–1136. 10.1021/acs.jnatprod.6b00040. [PubMed: 26894524]
22. Xu B, Tantillo DJ, and Rudolf JD (2021). Mechanistic insights into the formation of the 6,10-bicyclic eunicellane skeleton by the bacterial diterpene synthase Bnd4. Angew. Chem. Int. Ed 60, 23159–23163. 10.1002/ANIE.202109641.
23. Scesa PD, Lin Z, and Schmidt EW (2022). Ancient defensive terpene biosynthetic gene clusters in the soft corals. Nat. Chem. Biol 18, 659–663. 10.1038/s41589-022-01027-1. [PubMed: 35606556]
24. Burkhardt I, de Rond T, Chen PY-T, and Moore BS (2022). Ancient plant-like terpene biosynthesis in corals. Nat. Chem. Biol 18, 664–669. 10.1038/s41589-022-01026-2. [PubMed: 35606558]
25. Zallot R, Oberg NO, and Gerlt JA (2018). 'Democratized' genomic enzymology web tools for functional assignment. Curr. Opin. Chem. Biol 47, 77–85. 10.1016/j.cbpa.2018.09.009. [PubMed: 30268904]
26. Nakano C, Kudo F, Eguchi T, and Ohnishi Y (2011). Genome mining reveals two novel bacterial sesquiterpene cyclases: (−)-germacradien-4-ol and (−)-epi-α-bisabolol synthases from Streptomyces citricolor. ChemBioChem 12, 2271–2275. 10.1002/cbic.201100418. [PubMed: 23106076]
27. Xu B, Li Z, Alsup TA, Ehrenberger MA, and Rudolf JD (2021). Bacterial diterpene synthases prenylate small molecules. ACS Catalysis 11, 5906–5915. 10.1021/acscatal.1c01113. [PubMed: 34796043]
28. Xu B, Ning W, Wei X, and Rudolf JD (2022). Mutation of the eunicellane synthase Bnd4 alters its product profile and expands its prenylation ability. Org. Biomol. Chem 20, 8833–8837. 10.1039/D2OB01931K. [PubMed: 36321628]
29. Saurí J, Nolis P, and Parella T (2020). How to measure long-range proton-carbon coupling constants from 1H-selective HSQMBC experiments. Magn. Reson. Chem 58, 363–375. 10.1002/mrc.4928. [PubMed: 32239575]
30. Saurí J, and Parella T (2013). On the interference of J(HH) modulation in HSQMBC-IPAP and HMBC-IPAP experiments. Magn. Reson. Chem 51, 509–516. 10.1002/mrc.3977. [PubMed: 23780917]
31. Seco JM, Quiñoá E, and Riguera R (2004). The assignment of absolute configuration by NMR. Chem. Rev 104, 17–118. 10.1021/cr000665j.
32. Hoye TR, Jeffrey CS, and Shao F (2007). Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons. Nat. Protoc 2, 2451–2458. 10.1038/nprot.2007.354. [PubMed: 17947986]

33. Xiao Q, Wang L, Supekar S, Shen T, Liu H, Ye F, Huang J, Fan H, Wei Z, and Zhang C (2020). Structure of human steroid 5α-reductase 2 with the anti-androgen drug finasteride. Nat. Commun 11, 5430. 10.1038/s41467-020-19249-z. [PubMed: 33110062]
34. Baer P, Rabe P, Fischer K, Citron CA, Klapschinski TA, Groll M, and Dickschat JS (2014). Induced-fit mechanism in class I terpene cyclases. Angew. Chem., Int. Ed 53, 7652–7656. 10.1002/anie.201403648.
35. Govindam SVS, Yoshioka Y, Kanamoto A, Fujiwara T, Okamoto T, and Ojika M (2012). Cyclolobatriene, a novel prenylated germacrene diterpene, from the soft coral *Lobophytum pauciflorum*. Bioorg. Med. Chem 20, 687–692. 10.1016/j.bmc.2011.12.012. [PubMed: 22209732]
36. Faraldos JA, Wu S, Chappell J, and Coates RM (2007). Conformational analysis of (+)-germacrene A by variable-temperature NMR and NOE spectroscopy. Tetrahedron 63, 7733–7742. 10.1016/j.tet.2007.04.037. [PubMed: 20617157]
37. Terada Y, and Yamamura S (1979). Stereochemical studies on germacrenes: an application of molecular mechanics calculations. Tetrahedron Letters 20, 3303–3306. 10.1016/S0040-4039(01)95391-2.
38. Xu H, Lackus ND, Köllner TG, and Dickschat JS (2022). Isotopic labeling experiments solve the hedycaryol problem. Org. Lett 24, 587–591. 10.1021/acs.orglett.1c04021. [PubMed: 34985289]
39. Yamada Y, Arima S, Nagamitsu T, Johmoto K, Uekusa H, Eguchi T, Shin-ya K, Cane DE, and Ikeda H (2015). Novel terpenes generated by heterologous expression of bacterial terpene synthase genes in an engineered Streptomyces host. J. Antibiot 68, 385–394. 10.1038/ja.2014.171.
40. Saleh MB, Kerr RG, Saleh MB, and Kerr RG (2010). Identification of the cyclase product and its first oxidation product in the biosynthesis of fuscol and fuscosides. Aust. J. Chem 63, 901–906. 10.1071/CH10057.
41. Dickschat JS (2017). Modern aspects of isotopic labellings in terpene biosynthesis. Eur. J. Org. Chem 2017, 4872–4882. 10.1002/ejoc.201700482.
42. Lee C, Yang W, and Parr RG (1988). Development of the Colle-Salvetti correlationenergy formula into a functional of the electron density. Phys. Rev. B 37, 785–789. 10.1103/PhysRevB.37.785.
43. Adamo C, and Barone V (1998). Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. J. Chem. Phys 108, 664–675. 10.1063/1.475428.
44. Becke AD (1993). Density-functional thermochemistry. III. The role of exact exchange. J. Chem. Phys 98, 5648–5652. 10.1063/1.464913.
45. Matsuda SPT, Wilson WK, and Xiong Q (2006). Mechanistic insights into triterpene synthesis from quantum mechanical calculations. Detection of systematic errors in B3LYP cyclization energies. Org. Biomol. Chem 4, 530–543. 10.1039/B513599K. [PubMed: 16446812]
46. Tantillo DJ (2011). Biosynthesis via carbocations: Theoretical studies on terpene formation. Nat. Prod. Rep 28, 1035–1053. 10.1039/C1NP00006C. [PubMed: 21541432]
47. Tantillo DJ (2017). Importance of inherent substrate reactivity in enzyme-promoted carbocation cyclization/rearrangements. Angew. Chem. Int. Ed 56, 10040–10045. 10.1002/anie.201702363.
48. Tantillo DJ (2020). Exploring terpenoid biosynthesis with quantum chemical computations. In Comprehensive Natural Products III, Liu H.-w. and Begley T, eds. (Elsevier), pp. 644–653.
49. A data set collection of computational results is available in the IoChem-BD repository and can be accessed via 10.19061/iochem-bd-6-166.
50. Hong YJ, and Tantillo DJ (2011). The taxadiene-forming carbocation cascade. J. Am. Chem. Soc 133, 18249–18256. 10.1021/ja2055929. [PubMed: 21988104]

The bigger picture

Natural products have a long history of playing important roles in pharmaceutical, agricultural, and commercial applications. One of the current challenges is how to continue to discover new natural products. Using the genomes of organisms provides a route to select genes that are likely to contribute to the biosynthesis of new natural products. Our approach led to identifying an enzyme that produces an organic framework that is rare in nature and will lead to the identification of its genuine natural product. The goals of this research are to discover new natural products that are therapeutically relevant as drugs or drug leads, identify and develop biocatalytic tools for chemical synthesis or synthetic biology applications, and predict enzyme function from protein sequence alone.

Highlights

- First *trans*-eunicellane synthase found in nature
- Sesquiterpene and diterpene enzyme products structurally elucidated
- Mutagenesis, labeling studies and DFT calculations support cyclization mechanism
- Mechanistic comparison with a *cis*-eunicellane synthase

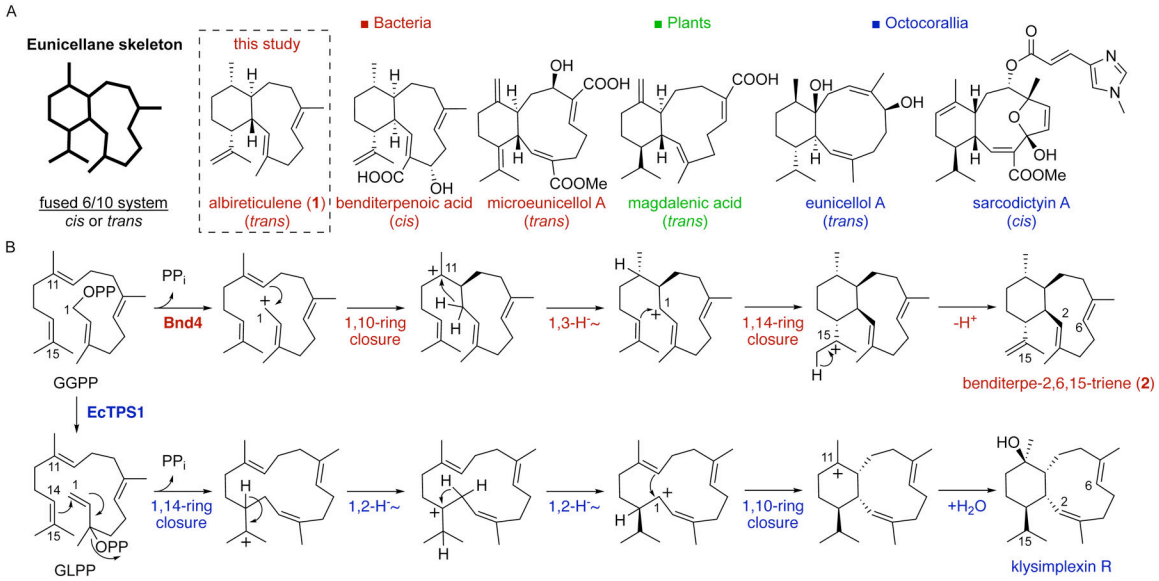


Figure 1. Eunicellane diterpenoid natural products and biosynthesis

(A) Selected members of eunicellane diterpenoids. The 6/10-bicyclic eunicellane skeleton possesses either *cis* or *trans* ring configuration and is found in coral, bacteria, and plants. The *trans* ring fusion is rare in known natural products.

(B) Recent studies support mechanistic differences between two *cis*-eunicellane di-TSs, Bnd4 from *Streptomyces* sp. (CL12–4) and EcTPS1 from *Erythropodium caribaeorum*. Prior to this study, no TSs that form the *trans*-eunicellane skeleton were known.

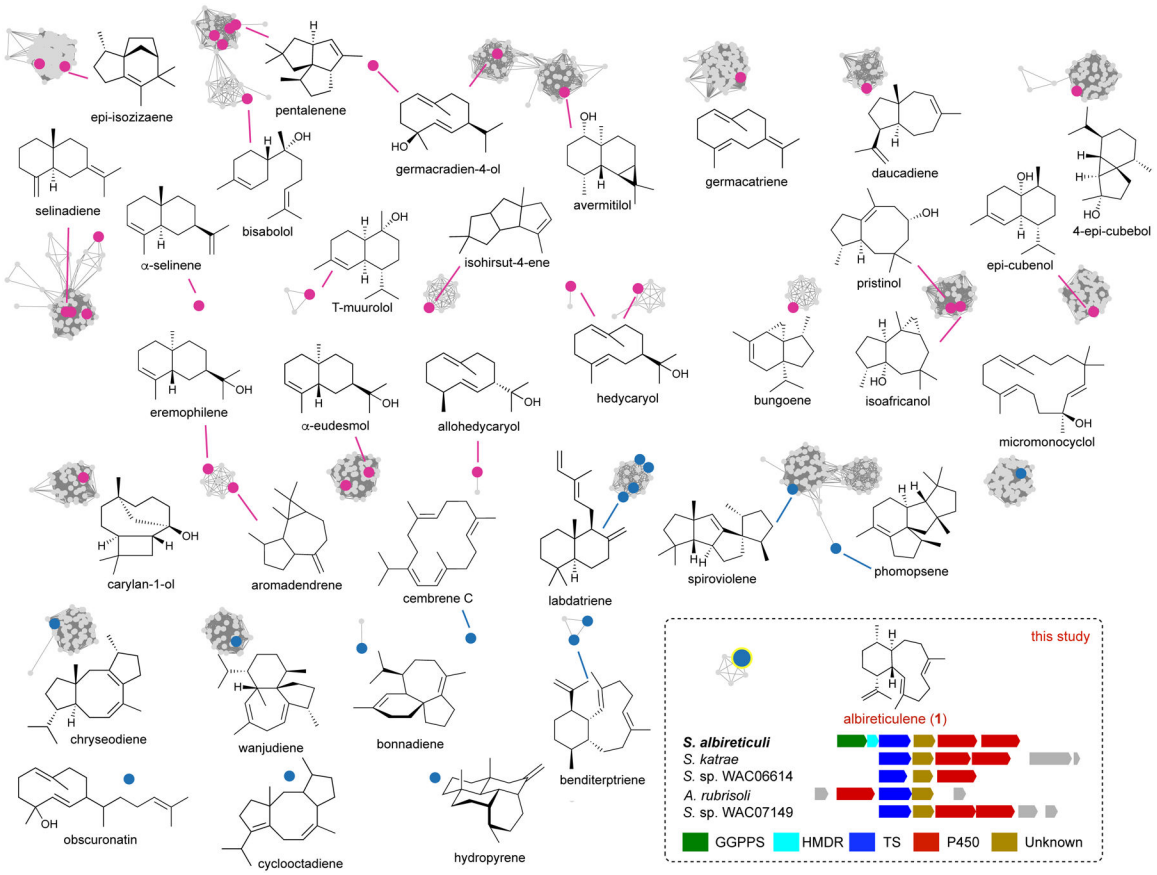


Figure 2. Sequence analysis highlights functional diversity of bacterial terpene synthases

Sequence similarity network of TSs (IPR034686) from bacteria at an e-value threshold of 10^{-70}. Functionally characterized TSs are colored (magenta, C15; blue, C20; blue with yellow highlight, AlbS) and their major products are shown. For clarity, not all characterized TSs are shown and not all products are shown for all clusters. In this enzyme family, there are 101 clusters of uncharacterized subfamilies and 151 uncharacterized singletons (not shown, see also Fig. S1). (inset) The AlbS subfamily of TSs and putative BGCs of AlbS and its homologues (see also Fig. S2). Genes are colored according to proposed function; green, GGPP synthase; cyan, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase; red, cytochrome P450; brown, conserved protein of unknown function; gray, unrelated.

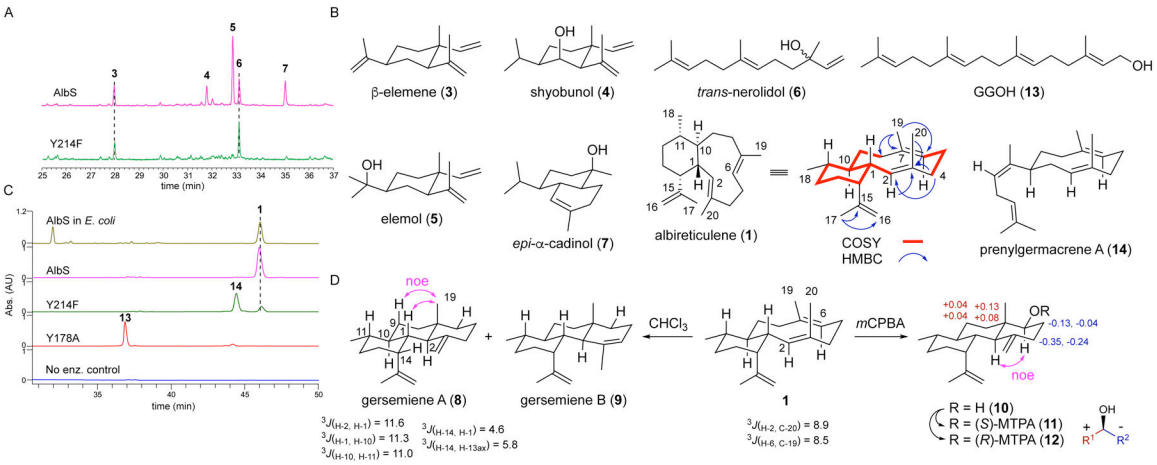


Figure 3. Enzymatic activity of AlbS and structurally determined sesquiterpene and diterpene products

(A) GC-MS analysis of AlbS and AlbS^{Y214F} reactions with FPP. The y-axis is the relative abundance of total ions.

(B) Enzymatic products identified in this study. See also Figs. S4–S22 and S62–S69. Only key 2D NMR correlations are shown for **1**. See also Table S4.

(C) HPLC-UV analysis of AlbS and mutant reactions with GGPP; **1** was also produced in vivo in *albS*-expressing *E. coli*. Absorbance was detected at 210 nm. See also Fig. S61.

(D) Structural elucidation of **1**, **8**, **9**, and **10** using NMR spectroscopy, chemical degradation, and Mosher's analysis. See also Figs. S23–S60.

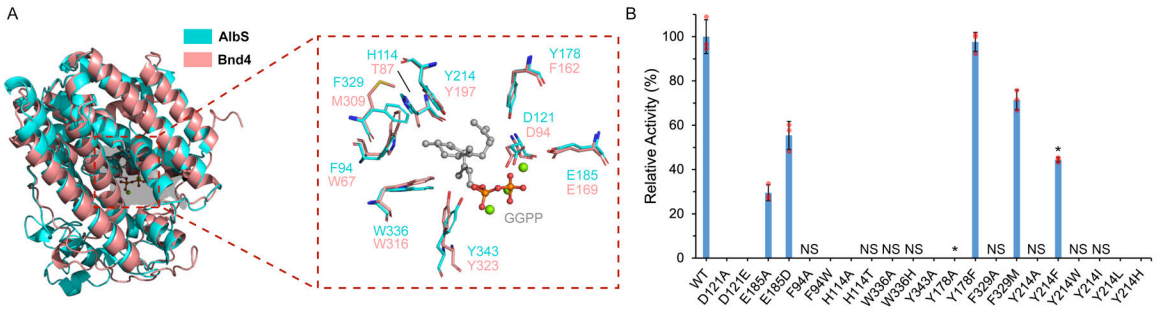


Figure 4. Structural and mutational analysis of the AlbS active site

(A) Structural models of AlbS and Bnd4 displaying key active site residues (sticks) with a docked model of GGPP (gray ball and stick). The three Mg^{2+} ions (green spheres) from selinadiene synthase (PDB ID 4OKZ)[34] were overlaid to show their approximate positions. (B) Relative cyclization activities of AlbS and mutants forming **1**. Values are the mean of three independent experiments, which are shown as overlaid orange circles, with error bars representing standard deviations. The absence of a bar denotes no activity detected; NS denotes not soluble. Mutants marked with asterisks produced other products (see also Figs. 3 and S61).

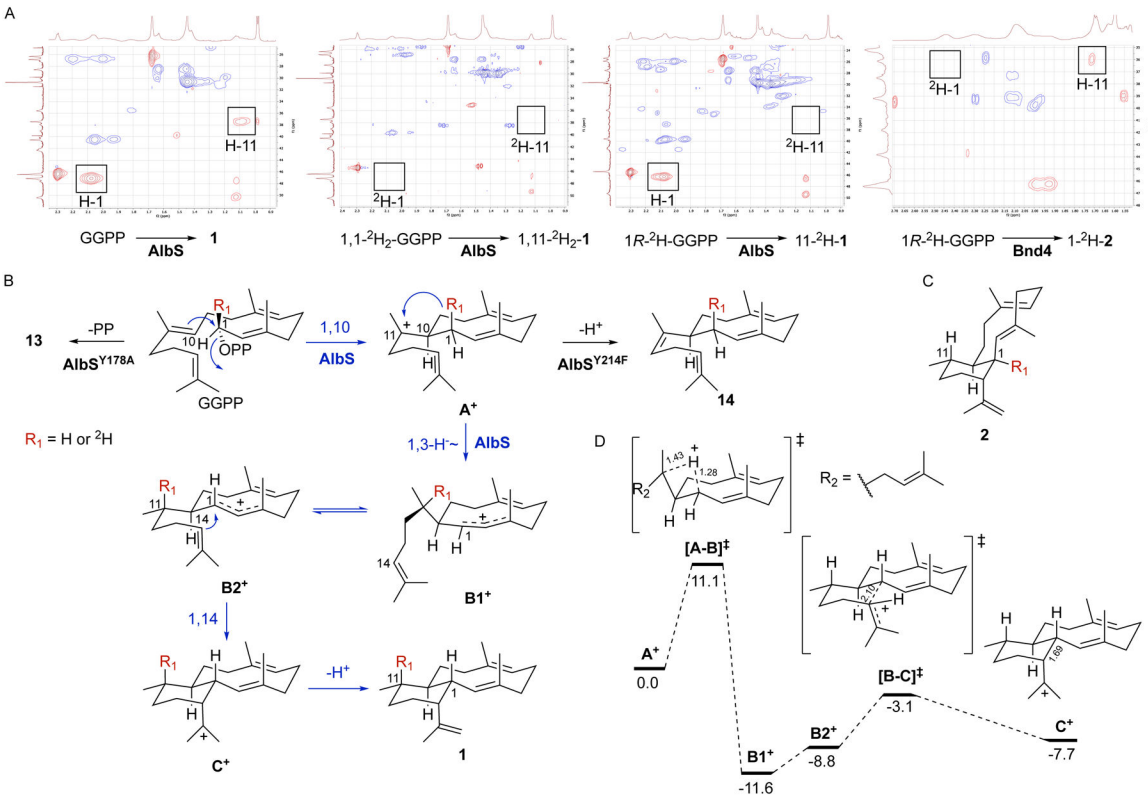


Figure 5. Mechanistic proposal and deuterium labeling support for the cyclization of GGPP into 1

(A) ^{1}H-^{13}C HSQC spectra of **1**, 1,11-$^{2}H_2$-**1**, 11-^{2}H-**1**, and 1-^{2}H-**2**, respectively. See also Figs. S9, S72, and S74–S76.

(B) Proposed mechanism for the formation of the *trans*-6,10-bicyclic eunicellane skeleton by AlbS (blue pathway). AlbS mutants yielding shunt products **13** and **14** are also shown (black pathways). The red R group depicts the location of the 1*R*-^{2}H-labeling experiment.

(C) Structure of 1-^{2}H-**2** from the 1*R*-^{2}H-labeling experiment with Bnd4 supporting a different hydride transfer in the AlbS and Bnd4 mechanisms.

(D) Relative free energies of intermediates and transition state structure in kcal mol^{-1}, calculated at mPW1PW91/6–31+G(d,p)//B3LYP/6–31+G(d,p) level of theory.[42–48] Bond distances (in Å) for key steps are listed beside the bond. The conformations depicted here are qualitative; see computed structures for actual conformations.[49]