Modular Chemoenzymatic Synthesis of Terpenes and their Analogues

Luke A. Johnson, Alice Dunbabin, Jennifer C. R. Benton, Robert J. Mart, and Rudolf K. Allemann*

Abstract: Non-natural terpenoids offer potential as pharmaceuticals and agrochemicals. However, their chemical syntheses are often long, complex, and not easily amenable to large-scale production. Herein, we report a modular chemoenzymatic approach to synthesize terpene analogues from diphosphory-lated precursors produced in quantitative yields. Through the addition of prenyl transferases, farnesyl diphosphates, (2E,6E)-FDP and (2Z,6Z)-FDP, were isolated in greater than 80 % yields. The synthesis of 14,15-dimethyl-FDP, 12-methyl-FDP, 12-hydroxy-FDP, homo-FDP, and 15-methyl-FDP was also achieved. These modified diphosphates were used with terpene synthases to produce the unnatural sesquiterpenoid semiochemicals (S)-14,15-dimethylgermacrene D and (S)-12-methylgermacrene D as well as dihydroartemisinic aldehyde. This approach is applicable to the synthesis of many non-natural terpenoids, offering a scalable route free from repeated chain extensions and capricious chemical phosphorylation reactions.

Terpenoid natural products are highly diverse secondary metabolites with economic importance in fields as diverse as fragrances (geraniol, linalool, citronellol),[1] flavorings (menthol),[2] pharmaceuticals (taxol, artemisinin),[3–5] biofuels (bisabolene),[6, 7] and agrochemicals (farnesenes).[8] All known terpenes are produced from two universal five-carbon precursors, dimethylallyl diphosphate (DMADP, 1) and isopentenyl diphosphate (IDP, 2), which are derived from either the mevalonate (MVA) or non-mevalonate (1-deoxy-d-xylulose-5-phosphate, DXP) pathways (Supporting Infor-mation, Scheme S1). As precursors to many useful chemicals, significant efforts have been made to upregulate production of 1 and 2 in vivo for terpene biosynthesis.[3, 9–14] While high terpene yields can be achieved, most notably for amorpha-4,11-diene by yeast fermentation for the production of the anti-malarial artemisinin,[3, 15] efficient production often requires significant metabolic engineering and fine-tuning.[16–18]

The synthesis of products beyond the natural terpenome offers the possibility of improved chemical properties and altered biological activities. Analogues of terpene synthase substrates have been pivotal to determining the mechanisms of terpene cyclizations by establishing the sequence of carbocation migrations, hydride and methyl shifts and the position of any proton eliminations.[19–22] These analogues can often act as substrates, forming modified terpene prod-ucts.[22–25] Point mutations of terpene synthases can dramat-ically alter product distributions, catalytic efficiencies, and substrate preferences allowing access to a plethora of new, non-natural terpenoid products.[26–29] For example, in addition to farnesyl diphosphate (FDP), (S)-germacrene D synthase (ScGDS) accepts 14,15-dimethyl-FDP as a substrate and cyclizes it to (S)-14,15-dimethylgermacrene D. Strikingly, while the natural product (S)-germacrene D is a repellent of aphids, (S)-14,15-dimethylgermacrene D acts as an attrac-tant.[30, 31] Similarly, 12-hydroxy-FDP is an alternative sub-strate for amorpha-4,11-diene synthase (AaADS), which forms dihydroartemisinic aldehyde, thereby reducing the subsequent number of chemical steps needed for the synthesis of artemisinin.[32]

The chemistry used to synthesise modified geranyl- (GDP, C10), farnesyl- (FDP, C15) and geranylgeranyl- (GGDP, C20) diphosphates is often unsuitable for the large-scale produc-tion required for many applications, for instance in crop protection or medicinal therapy. Analogues are commonly derived by iterative chain extensions using enolate and Wittig chemistry, often suffering from low yields and requiring repeated chromatography.[31–35] The final step is generally addition of the diphosphate group by substitution of a halo-genated allylic alcohol with tris(tetra-n-butylammonium) hydrogen diphosphate. This step requires multiple ion exchange procedures and results in highly variable yields (10–60 %).[38] To address these challenges, we designed a modular chemoenzymatic synthesis to generate terpene analogues from modified IDP and DMADP intermediates. We envisioned using promiscuous kinases to diphosphorylate prenol (3) and isoprenol (4), and then adding prenyl trans-ferases and terpene cyclases to create a complete enzymatic pathway to synthesize terpenes directly from prenols. The natural promiscuity of the kinases, prenyl transferases, and terpene cyclases constituting this truncated pathway should then allow for the use of modified prenols as substrates. Sequential use of different prenyl transferases would enable modular chemoenzymatic assembly of isoprenoids with unprecedented ease of placement of functional groups and access to synthetic space not naturally available or limited to more difficult post-cyclisation tailoring reactions (P450 oxi-dation, S-adenosyl-l-methionione-dependent methyla-tion).[39, 40]

The natural substrate of the Escherichia coli hydroxyethylthiazole kinase (EcTHIM) is 5-(2-hydroxyethyl)-4-methylthiazole, but EcTHIM has also been reported to phosphorylate 3.[41] We speculated that by combining EcTHIM with the isopentenyl phosphate kinase from Methanocaldococcus

[*] Dr. L. A. Johnson, Dr. A. Dunbabin, J. C. R. Benton, Dr. R. J. Mart, Prof. Dr. R. K. Allemann
School of Chemistry, Cardiff University
Park Place, Cardiff CF10 3AT (UK) E-mail: allemannRK@cardiff.ac.uk

jannaschii (MjIPK) both 1 and 2 could be synthesized directly from their respective alcohols (Figure 1 a). Compared to natural biosynthetic pathways (Supporting Information, Scheme S1), the number of enzymatic steps and equivalents of energy-rich molecules, such as ATP, CTP, and NADPH, is reduced. Crucially, this more straightforward synthetic route should also allow greater substrate variation. Efficient phosphorylation of 3 and 4 by EcTHIM was found to be possible using phosphoenol pyruvate (PEP) in the presence of pyruvate kinase to recycle a catalytic quantity of ATP and to prevent significant ATP hydrolysis. After optimising the relative concentrations of PEP and substrates to minimize inhibition of EcTHIM by PEP (Supporting Information, Figure S2), this system generated quantitative yields of the corresponding prenyl phosphates as determined by ^{31}P and ^{1}H NMR and LC-MS (Figure 1 b and Supporting Information, Figures S1–S4).

Having confirmed that 3 and 4 act as substrates of EcTHIM by ^{1}H and ^{31}P NMR, we next investigated the

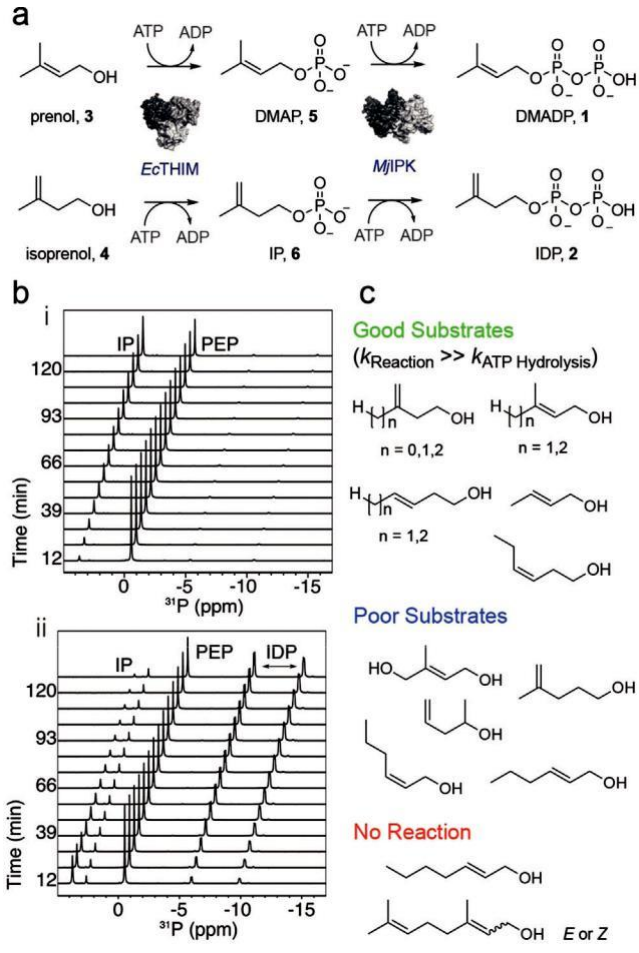
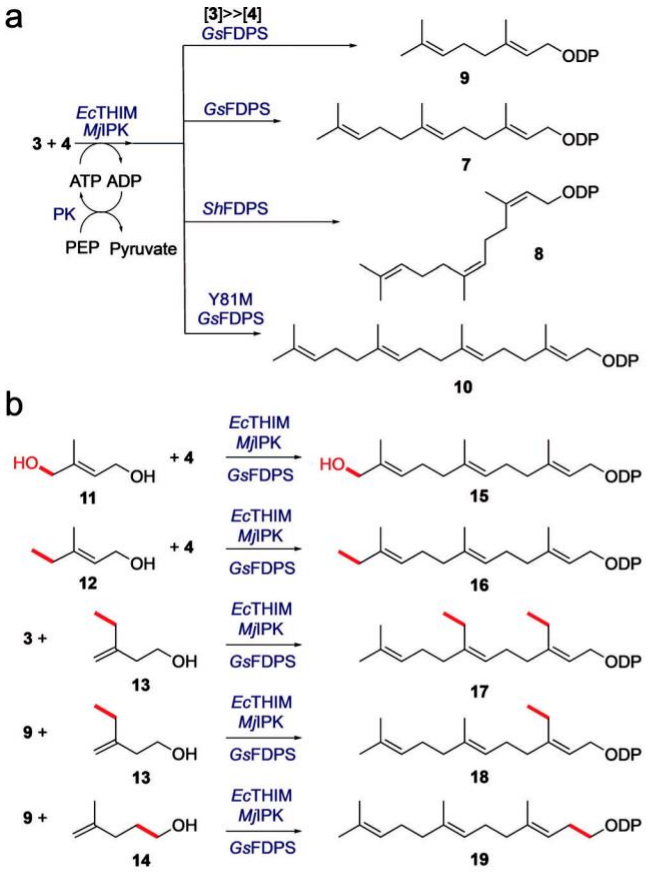


Figure 1. a) Stepwise ATP-dependent monophosphorylation of prenol (3) and isoprenol (4) by EcTHIM (PDB: 1EKK) and diphosphorylation of 5 and 6 by MjIPK (PDB: 3K4Y) to form dimethylallyl diphosphate (DMADP, 1) and isopentenyl diphosphate (IDP, 2). b) ^{31}P NMR time-course following i) monophosphorylation of 4 to 6 (d 3.7) by EcTHIM and ii) phosphorylation of 6 to 2 (d 5.9, 10.0) by MjIPK. Phosphoe-nol pyruvate (PEP, d 0.5) is used to recycle catalytic quantities of ATP. c) Analogues of 3 and 4 tested with EcTHIM.

substrate scope of EcTHIM. A variety of substrates with structural similarity to 3 and 4 featuring methylated, demethylated, and hydroxylated substrates, and increased chain lengths were tested (Figure 1 c and Supporting Information, Figures S5–S7). For substrates where the phosphorylation rate was significantly higher than the background rate of ATP hydrolysis, kinetic parameters were determined by a coupled assay with lactate dehydrogenase and spectroscopically monitoring consumption of NADH (Supporting Information, Figure S8 and Table S1). EcTHIM was found to turn over 4 with an approximately 10-fold greater k_{cat} than was measured for 3. Exploring the substrate scope established a maximum accepted chain length of seven carbons (Figure 1 c and Supporting Information, Figure S8 and Table S1). Those substrate affinities that could be measured were all compa-rable, suggesting that EcTHIM s promiscuity arises from weak non-specific contacts with its substrates. All products from reactions with EcTHIM were accepted as substrates by MjIPK; the resulting diphosphates were characterised by ^{1}H and ^{31}P NMR and LC-MS (Supporting Information, Figur-es S6 and S7 and Table S2).[42]

Synthesis of (2E,6E)-FDP (7) was achieved using a 1:2 ratio of 3 to 4, EcTHIM, MjIPK, and the (2E,6E)-FDP synthase from Geobacillus stearothermophilus (GsFDPS, Scheme 1 a). In the presence of magnesium (5 mm), a neces-sary cofactor for the enzymes (Supporting Information, Figures S9 and S10), 7 precipitated from solution. Purification of 7 was achieved by repeatedly washing the precipitate with water to remove any impurities before addition of excess inorganic diphosphate to coordinate the magnesium, disrupt-ing the FDP-Mg^{2+} complex. Once soluble, 7 was purified using a C18 reverse phase column to give an isolated yield of 83 %.

Replacing GsFDPS with (2Z,6Z)-FDP synthase from Solanum habrochaites, allowed ready access to (2Z,6Z)-FDP (8) in 92 % yield (Supporting Information, Figure S11), avoiding the cumbersome synthesis, and losses during diphos-phorylation of isomerically pure (2Z,6Z)-farnesol. When 1 was present in large excess over 2, GDP (9) was formed using GsFDPS (Supporting Information, Figure S12). Unlike FDPs, 9 is soluble in the presence of magnesium and was separated from FDP by filtration. Using the Y81M variant of GsFDPS, allowed GGDP (10) to be generated from a 1:3 ratio of 1 and 2 (Supporting Information, Figure S11).[43] When methylated or hydroxylated analogues of 2 were combined with 1 in the presence of GsFDPS, 15, 16, and 17 were produced (Scheme 1 b and Supporting Information, Figures S13–S15). As with FDP, the products precipitated from the reaction mixture in the presence of magnesium allowing their straightforward isolation and purification. To access terpene precursors derived from both 2 and analogues thereof without forming statistical mixtures of products, 9 was mixed with methyl-IDP in the presence of GsFDPS to synthesise 15-methyl-FDP (18, Supporting Information, Fig-ure S16). Similarly, when homoisoprenol (14) was diphos-phorylated and added to 9, homo-FDP (19) was produced (Supporting Information, Figure S17), although in this case the reaction progressed extremely slowly, most likely because of suboptimal substrate alignment within GsFDPS.



Scheme 1. a) Pathway for the synthesis of GDP (9), (2E,6E)-FDP (7), (2Z,6Z)-FDP (8), and GGDP (10). b) FDP analogues 12-hydroxy-FDP (15), 12-methyl-FDP (16), 14,15-dimethyl-FDP (17), 15-methyl-FDP (18), homo-FDP (19), enzymatically synthesized from prenol (3), isoprenol (4), 4-hydroxyprenol (11), methylprenol (12), methylisoprenol (13), and homoisoprenol (14).

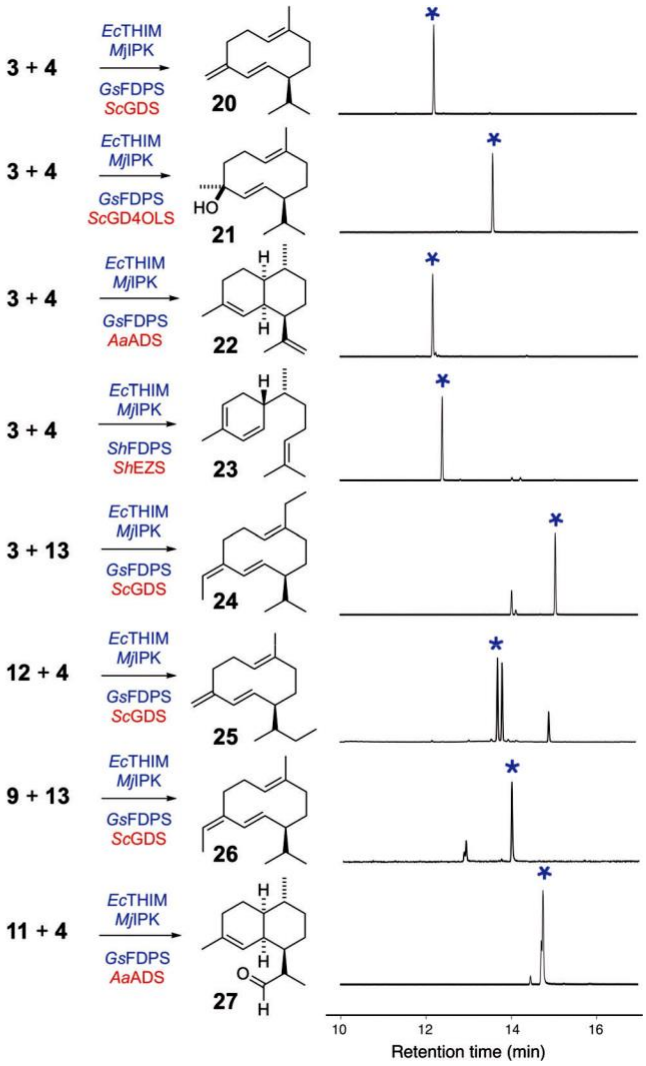


Figure 2. Synthesis of natural and unnatural sesquiterpenes from prenols. GC-traces of products formed under AaADS, ScGDS, ScGD4OLS, and ShEZS catalysis in combination with EcTHIM, MjIPK, an appropriate FDPS, pyruvate kinase, phosphoenol pyruvate, and catalytic ATP. Terpene products (*) (S)-germacrene D (20), ( )-germacradiene-4-ol (21), amorpha-4,11-diene (22), 7-epizingiberene (23), (S)-14,15-dimethylgermacrene D (24), (S)-12-methylgermacrene D (25), (S)-15-methylgermacrene D (26), and dihydroartemisinic aldehyde (27) were produced.

Finally, terpene synthases were added to reaction mix-tures to produce high-value sesquiterpene products and unnatural analogues (Figure 2). To avoid precipitation of FDP or its analogues by the presence of magnesium, 2-hydroxypropyl-b-cyclodextrin equimolar to the FDP product was added (Supporting Information, Figure S9). The hydro-phobic interior of the cyclodextrin acts as a host, preventing guest FDP molecules from interacting with each other and forming larger scale complexes (Supporting Information, Figure S18). GC-MS analyses of the pentane extractable products obtained from phosphorylation/cyclization reactions starting from 3 and 4 catalyzed by ScGDS, AaADS, 7-epizingiberene synthase from Solanum habrochaites (ShEZS), and ( )-germacradiene-4-ol synthase from Strep-tomyces citricolor (ScGD4OL) showed that the respective natural products had been generated (20–23) (Figure 2, S19– 30).[30, 32, 44] Furthermore, using methylated prenol ana-logues 12 and 13 directly yielded the desirable unnatural semiochemicals (S)-14,15-dimethylgermacrene D (24), (S)-12-methylgermacrene D (25), and (S)-15-methylgermacr-ene D (26) (Figure 2). Similarly, using 4-hydroxyprenol (11) and 4 to form 15 followed by addition of AaADS yielded dihydroartemisinic aldehyde (27). Meanwhile a number of unknown sesquiterpene products were formed using methy-lated prenols and ScGD4OL or AaADS. GC-MS results were consistent with unnatural methylated terpenoids (Supporting Information, Figures S25, S26, S29, and S30).

In summary, we report a novel, efficient modular chemo-enzymatic synthesis of terpenes and their analogues by directly phosphorylating five-carbon precursors. Coupling phosphorylation with ATP recycling allows quantitative yields to be achieved for the two phosphorylation steps. Combining these reactions with prenyl transferases results in excellent yields (greater than 80 %) of natural farnesyl diphosphates, key intermediates in the synthesis of sesquiter-penoid natural products. As 3 and 4 precursors are signifi-cantly cheaper than farnesol, this methodology offers a com-

petitive alternative synthesis of FDP; given the difficulty of synthesising (2Z,6Z)-farnesyl diphosphate, this work repre-sents the first feasible route to its synthesis. Purification of the products is straightforward due to the formation of insoluble magnesium complexes. Adding prenyl transferases and ter-pene synthases to the reaction mixture allows synthesis of natural and non-natural terpenes, highlighting the strength of this modular approach. Finally, the ease of addition in different combinations of prenol analogues to the synthesis, allows unprecedented access to modified linear terpene precursors. Importantly, this capacity to synthesize unnatural FDPs is a key advantage of this method over whole-cell procedures, where competition with natural pathways would be expected to result in challenges with toxicity. We anticipate that this new approach will allow modified terpene products to be produced on scales suitable for medicinal and agro-chemical use.

Acknowledgements

We thank Thomas Williams for assistance with the LC-MS analysis (EP/L027240/1), Gwawr Davies and Florence Huynh for help preparing 4-hydroxyprenol and geranyl diphosphate. This work was supported by the U.K. Biotechnology and Biological Research Council (BBSRC) through Grants BB/ M006158/1, BB/P009980/1, BB/M022463/1, BB/N012526/ 1 and BB/R019681/1 to RKA.

Conflict of interest

The authors declare no conflict of interest.

Keywords: biocatalysis · chemoenzymatic synthesis · natural products · terpenoids

[1] W. Schwab, R. Davidovich-Rikanati, E. Lewinsohn, Plant J. 2008, 54, 712 – 732.
[2] H. S. Toogood, A. N. Cheallaigh, S. Tait, D. J. Mansell, A. Jervis, A. Lygidakis, L. Humphreys, E. Takano, J. M. Gardiner, N. S. Scrutton, ACS Synth. Biol. 2015, 4, 1112 – 1123.
[3] V. J. J. Martin, D. J. Piteral, S. T. Withers, J. D. Newman, J. D. Keasling, Nat. Biotechnol. 2003, 21, 796 – 802.
[4] C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, et al., Nature 2013, 496, 528 – 532.
[5] P. K. Ajikumar, W.-H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, Science 2010, 330, 70 – 74.
[6] R. M. Phelan, O. N. Sekurova, J. D. Keasling, S. B. Zotchev, ACS Synth. Biol. 2015, 4, 393 – 399.
[7] P. P. Peralta-Yahya, J. D. Keasling, Biotechnol. J. 2010, 5, 147 – 162.
[8] T. J. A. Bruce, G. I. Aradottir, L. E. Smart, J. L. Martin, J. C. Caulfield, A. Doherty, C. A. Sparks, C. M. Woodcock, M. A. Birkett, J. A. Napier, et al., Sci. Rep. 2015, 5, 11183.
[9] A. O. Chatzivasileiou, V. Ward, S. M. B. Edgar, G. Stephano-poulos, Proc. Natl. Acad. Sci. USA 2019, 116, 506 – 511.
[10] S. Takahashi, Y. Yeo, B. T. Greenhagen, T. McMullin, L. Song, J. Maurina-Brunker, R. Rosson, J. P. Noel, J. Chappell, Biotechnol. Bioeng. 2007, 97, 170 – 181.
[11] S. W. Kim, J. D. Keasling, Biotechnol. Bioeng. 2001, 72, 408 – 415.
[12] C. B. Eiben, T. de Rond, C. Bloszies, J. Gin, J. Chiniquy, E. E. K. Baidoo, C. J. Petzold, N. J. Hillson, O. Fiehn, J. D. Keasling, ACS Synth. Biol. 2019, 8, 2238 – 2247.
[13] S. Lund, R. Hall, G. J. Williams, ACS Synth. Biol. 2019, 8, 232 – 238.
[14] G. Bian, Y. Han, A. Hou, Y. Yuan, X. Liu, Z. Deng, T. Liu, Metab. Eng. 2017, 42, 1 – 8.
[15] P. J. Westfall, D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, et al., Proc. Natl. Acad. Sci. USA 2012, 109, E111 – E118.
[16] S. K. Kim, G. H. Han, W. Seong, H. Kim, S.-W. Kim, D.-H. Lee, S.-G. Lee, Metab. Eng. 2016, 38, 228 – 240.
[17] J. Alonso-Gutierrez, E.-M. Kim, T. S. Batth, N. Cho, Q. Hu, L. J. G. Chan, C. J. Petzold, N. J. Hillson, P. D. Adams, J. D. Keasling, et al., Metab. Eng. 2015, 28, 123 – 133.
[18] X. Wang, W. Liu, C. Xin, Y. Zheng, Y. Cheng, S. Sun, R. Li, X. G. Zhu, S. Y. Dai, P. M. Rentzepis, et al., Proc. Natl. Acad. Sci. USA 2016, 113, 14225 – 14230.
[19] J. A. Faraldos, D. J. Miller, V. GonzQlez, Z. Yoosuf-Aly, O. Casc+n, A. Li, R. K. Allemann, J. Am. Chem. Soc. 2012, 134, 5900 – 5908.
[20] H. Sato, K. Teramoto, Y. Masumoto, N. Tezuka, K. Sakai, S. Ueda, Y. Totsuka, T. Shinada, M. Nishiyama, C. Wang, et al., Sci. Rep. 2016, 5, 18471.
[21] M. Chen, N. Al-lami, M. Janvier, E. L. D'Antonio, J. A. Faraldos, D. E. Cane, R. K. Allemann, D. W. Christianson, Biochemistry 2013, 52, 5441 – 5453.
[22] D. J. Miller, F. Yu, R. K. Allemann, ChemBioChem 2007, 8, 1819 – 1825.
[23] F. Huynh, D. J. Grundy, R. L. Jenkins, D. J. Miller, R. K. Allemann, ChemBioChem 2018, 19, 1834 – 1838.
[24] M. Demiray, D. J. Miller, R. K. Allemann, Beilstein J. Org. Chem. 2019, 15, 2184 – 2190.
[25] C. Oberhauser, V. Harms, K. Seidel, B. Schrçder, K. Ekramza-deh, S. Beutel, S. Winkler, L. Lauterbach, J. S. Dickschat, A. Kirschning, Angew. Chem. Int. Ed. 2018, 57, 11802 – 11806; Angew. Chem. 2018, 130, 11976 – 11980.
[26] R. Li, W. K. W. Chou, J. A. Himmelberger, K. M. Litwin, G. G. Harris, D. E. Cane, D. W. Christianson, Biochemistry 2014, 53, 1155 – 1168.
[27] M. Loizzi, V. GonzQlez, D. J. Miller, R. K. Allemann, Chem-BioChem 2018, 19, 100 – 105.
[28] B. T. Greenhagen, P. E. O Maille, J. P. Noel, J. Chappell, Proc. Natl. Acad. Sci. USA 2006, 103, 9826 – 9831.
[29] C. Gçrner, I. H uslein, P. Schrepfer, W. Eisenreich, T. Br4ck, ChemCatChem 2013, 5, 3289 – 3298.
[30] S. Touchet, K. Chamberlain, C. M. Woodcock, D. J. Miller, M. A. Birkett, J. A. Pickett, R. K. Allemann, Chem. Commun. 2015, 51, 7550 – 7553.
[31] "Olfactory Ligands": R. K. Allemann, J. A. Pickett, D. J. Miller, M. A. Birkett, WO2016110671A1, 2016.
[32] X. Tang, M. Demiray, T. Wirth, R. K. Allemann, Bioorg. Med. Chem. 2018, 26, 1314 – 1319.
[33] D. S. Rawat, R. A. Gibbs, Org. Lett. 2002, 4, 3027 – 3030.
[34] I. J. Fairlamb, J. M. Dickinson, M. Pegg, Tetrahedron Lett. 2001, 42, 2205 – 2208.
[35] X. Liu, G. D. Prestwich, J. Am. Chem. Soc. 2002, 124, 20 – 21.
[36] Y. Shao, J. T. Eummer, R. A. Gibbs, Org. Lett. 1999, 1, 627 – 630.
[37] F. W. Sum, L. Weiler, Tetrahedron 1981, 37, 303 – 317.
[38] V. J. Davisson, A. B. Woodside, T. R. Neal, K. E. Stremler, M. Muehlbacher, C. D. Poulter, J. Org. Chem. 1986, 51, 4768 – 4779.
[39] R. Bernhardt, J. Biotechnol. 2006, 124, 128 – 145.
[40] D. K. Liscombe, G. V. Louie, J. P. Noel, Nat. Prod. Rep. 2012, 29, 1238.
[41] P.-H. Wang, A. N. Khusnutdinova, F. Luo, J. Xiao, K. Nemr, R.
[44] D. J. Grundy, M. Chen, V. GonzQlez, S. Leoni, D. J. Miller, D. W. Christianson, R. K. Allemann, Biochemistry 2016, 55, 2112 – 2121.

Flick, G. Brown, R. Mahadevan, E. A. Edwards, A. F. Yakunin, Cell Chem. Biol. 2018, 25, 560 – 570.
[42] S. Lund, T. Courtney, G. J. Williams, ChemBioChem 2019, 20, 2217 – 2221.
[43] S. I. Ohnumaf, K. Narita, T. Nakazawa, C. Ishida, Y. Takeuchi, C. Ohto, T. Nishinof, J. Biol. Chem. 1996, 271, 30748 – 30754.