ChemPubSoc Europe

DOI: 10.1002/cbic.201300763

An Improved Technique for the Rapid Chemical Characterisation of Bacterial Terpene Cyclases

Jeroen S. Dickschat,* Khomaizon A. K. Pahirulzaman, Patrick Rabe, and Tim A. Klapschinski[a]

A derivative of the pET28c(+) expression vector was constructed. It contains a yeast replication system (2μ origin of replication) and a yeast selectable marker (URA3), and can be used for gene cloning in yeast by efficient homologous recombination, and for heterologous expression in *E. coli*. The vector was used for the expression and chemical characterisation of three bacterial terpene cyclases.

With over 50000 known members, terpenoids represent the largest class of natural products that can be isolated from all kingdoms of life. The chemical space that is covered by terpenoids ranges from small acyclic hydrocarbons, such as isoprene, to highly functionalised polycyclic molecules with dense stereochemical information. The biosynthesis of all terpenoids starts from just a few acyclic precursors, including prenyl diphosphate (hemiterpenes, C_5), geranyl diphosphate (GPP, monoterpenes, C_{10}), farnesyl diphosphate (FPP, sesquiterpenes, C_{15}) and geranylgeranyl diphosphate (GGPP, diterpenes, C_{20}).[1,2] The transformation of these precursors into structurally complex terpene hydrocarbons or alcohols is catalysed by terpene cyclases—some of the most sophisticated catalysts in natural product biosynthesis. In type I terpene cyclases, cyclisation is initiated by abstraction of the precursor's diphosphate by complexation to a trinuclear Mg^{2+} cluster, which, in turn, is coordinated to two highly conserved motifs: the aspartate-rich motif (DDXXD) and the NSE triad (ND(L,I,V)XSXXXE). The resulting allylic cation can undergo a multistep cascade with several intramolecular attacks of olefinic double bonds (i.e. cyclisation reactions), Wagner–Meerwein rearrangements and proton or hydride shifts. Termination mechanisms usually include deprotonation to yield a terpene hydrocarbon or quenching of the cation by attack by water to produce a terpene alcohol. The intermediate cations along the cascade cannot be observed directly but can become evident from shunt products that arise by deprotonations or attack by water, along the pathway to the main product.[3]

The products of several bacterial terpene cyclases have been chemically identified,[3–23] and structural data for a few enzymes have been obtained.[23–26] Furthermore, recent genome sequencing has revealed several hundred genes coding for terpene cyclase homologues in bacterial genomes, most in actinomycetes.[20–22] The accumulated knowledge demonstrates that bacterial terpene cyclases share a common three-dimensional architecture, including a recently discovered helix-break motif that is involved in substrate activation,[23,27] although amino acid sequence homology is low. Chemical analysis of the terpenes emitted into the headspace above bacterial agar plate cultures has shown that terpene cyclases with closely related amino acid sequences catalyse the formation of the same main product,[20,22] although it is impossible to predict the structure of a terpene product from the cyclase's amino acid sequence. This underlines the requirement for further experimental data to shed light on the functions of as yet uncharacterised terpene cyclases. As genome sequencing has now become a very fast and affordable standard technology, more candidate terpene cyclase sequences emerge every day. To keep track of these fast developments, we recently established a rapid analytical method for the chemical characterisation of bacterial terpene cyclases. This is based on the heterologous expression of terpene cyclases in *Escherichia coli* and direct sampling of volatile terpenes by a closed-loop stripping apparatus (CLSA) followed by GC/MS.[21] The bottleneck in this approach is the ligation step for cloning following restriction endonuclease digestion (often only successful after several trials). Therefore, we developed an improved technique for the rapid chemical characterisation of bacterial terpene cyclases; the technique makes use of a very efficient cloning strategy: homologous recombination in yeast. This technique and the structures of the terpene products generated by three previously uncharacterised bacterial terpene cyclases are presented here. The three enzymes were also investigated by incubation experiments with linear terpene precursors.

Plasmid pET28c(+) is a vector for heterologous expression of genes in *E. coli*. To make this vector replicable in yeast, the region of vector pRS426 that contains the 2μ origin of replication and the URA3 yeast selectable marker were amplified by PCR and cloned by yeast homologous recombination into a pET28c(+) vector that had been linearised by restriction endonuclease digestion, (Figure 1; details in the Supporting Information). The resulting vector is replicable in yeast and in *E. coli* and can therefore be used for gene cloning in yeast by homologous recombination, followed by gene expression in *E. coli*. The vector was named pYE-Express (Yeast-*E. coli*-Expression system).

The first investigated enzyme was a terpene cyclase from *Saccharothrix espanaensis* DSM 44229 (accession number YP_007039597). This enzyme was previously suggested by us to function as a caryophyllene synthase,[22] based on the identification of caryophyllene in headspace extracts from *S. espanaensis*. Three terpene cyclases are encoded in the genome of this species. One was easily identified as geosmin synthase,

[a] Dr. J. S. Dickschat, Dr. K. A. K. Pahirulzaman, P. Rabe, T. A. Klapschinski
Institut für Organische Chemie, TU Braunschweig
Hagenring 30, 38106 Braunschweig (Germany)
E-mail: j.dickschat@tu-bs.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300763.

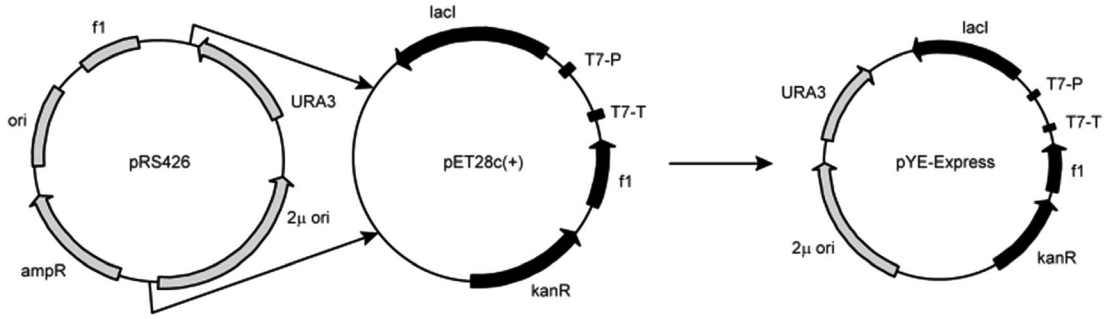


Figure 1. Construction of the expression vector pYE-Express.

and a second was 2-methylisoborneol synthase (as determined by sequence homology). As formation of caryophyllene by either of these enzymes would be difficult to understand, the involvement of the third uncharacterised enzyme in caryophyllene biosynthesis was considered the most likely. After cloning of the corresponding gene into pYE-Express by homologous recombination in yeast followed by heterologous expression in *E. coli*, CLSA headspace sampling and GC/MS analysis showed indeed the production of a single sesquiterpene hydrocarbon; this was identified as caryophyllene (**1**, Figure 2 A). Incubation of the purified enzyme with GPP, FPP and GGPP revealed that FPP was the only accepted substrate for conversion into **1** as a single product (Figure 2 B). These data unequivocally proved that the function of the third *S. espanaensis* enzyme was caryophyllene synthase, in full agreement with our previous findings.

Heterologous expression of the terpene cyclase WP_010314578 from *Saccharopolyspora spinosa* NRRL 18395 in *E. coli* resulted in headspace extracts containing (*E*)-β-farnesene (**2**) as the main product, as well as (*E*,*E*)-α-farnesene (**3**), nerolidol (**4**), germacrene A (**5**, observed as its thermal Cope rearrangement product β-elemene, **5***), hedycaryol (**6**), germacrene D (**7**), bicyclogermacrene (**8**), allo-hedycaryol (**9**), viridiflorol (**10**) and avermitilol (**11**, Figure 3 A). All these compounds are biosynthetically related; a conceivable formation by one terpene cyclase via common cationic intermediates is shown in Scheme 1. The pathway starts with the ionisation of FPP by diphosphate abstraction to yield the farnesyl cation **A**. Deprotonation or attack by water gives rise to **2**, **3** and **4**. Cyclisation to the (*E*,*E*)-germacradienyl cation **B** paves the way to **5** by loss of a proton and to **6** by capture with water. A 1,3-hydride shift in **B** to produce the allyl cation **C** and deprotonation yields **7**; the alternative deprotonation-initiated cyclisation to **8** followed by attack by water under ring opening gives rise to **9**.

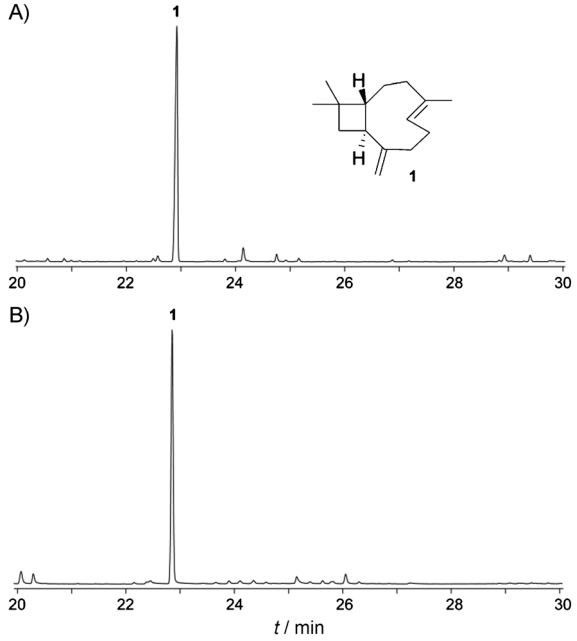


Figure 2. Total ion chromatogram of A) headspace extract from *E. coli* upon heterologous expression of the caryophyllene synthase from *S. espanaensis*, and B) extract from incubation of recombinant caryophyllene synthase with FPP.

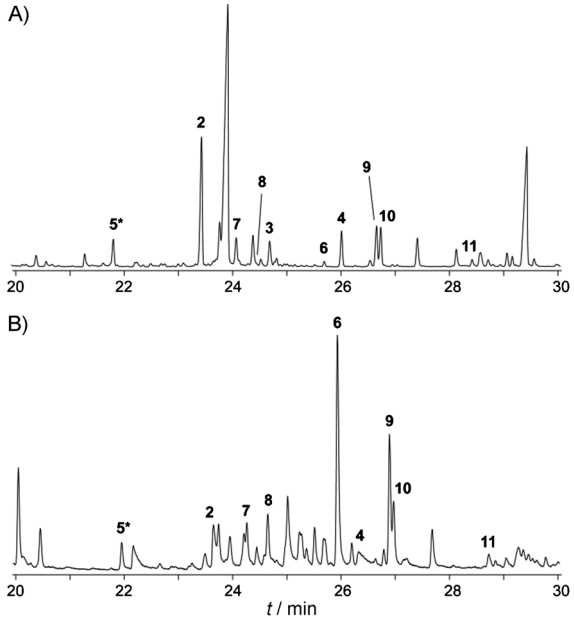
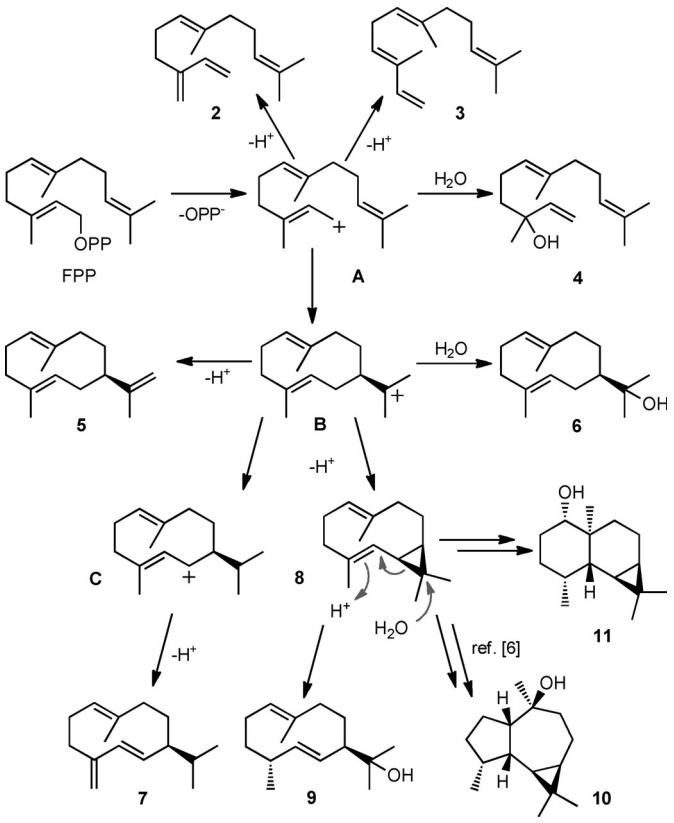


Figure 3. Total ion chromatogram of A) headspace extract from *E. coli* upon heterologous expression of the hedycaryol synthase from *S. spinosa*, and B) extract from incubation of recombinant hedycaryol synthase with FPP. Peak numbers refer to compound numbers (see Scheme 1).



Scheme 1. Suggested biosynthetic pathway to all products of the hedycaryol synthase from *S. spinosa*. For chiral terpenes, relative configurations are shown.

This mechanism is similar to that for the formation of (1(10)*E*,5*E*)-germacradien-11-ol as an intermediate in geosmin biosynthesis, where isolepidozene (a stereoisomer of **8**) reacts analogously with water under cyclopropane ring-opening.[28] The hydrocarbon **8** is also the precursor to both **10** and **11** by a known mechanism.[6] None of the terpenes detected here were observed in a *S. spinosa* headspace extract (Figure S1 in the Supporting Information), thus suggesting that the respective terpene cyclase is not expressed in the natural host under laboratory culture conditions. Incubation of the purified enzyme resulted in the formation of the same products (Figure 3B), except that alcohols **6** and **9** were detected in significantly larger amounts. These alcohols might be under-represented in the headspace extracts because of their much lower volatility in comparison to sesquiterpene hydrocarbons. Another reason for the low amount of **6** in the headspace extracts might be its known instability, which could result in nonspecific degradation during collection on charcoal traps. The *S. spinosa* enzyme is best described as a "hedycaryol synthase", as **6** was the main product of in vitro incubation.

Interestingly, incubation of the recombinant avermitilol synthase from *Streptomyces avermitilis* with FPP was reported to yield mainly **11**, with minor amounts of **5**, **10** and germacrene B,[6] but none of the other terpenes that were observed here. Although similar products were observed (albeit in different proportions), the *S. spinosa* hedycaryol synthase and the avermitilol synthase from *S. avermitilis* share only 27% amino acid identity (by comparison, there is 33% residue identity between avermitilol and pentalenene synthases from *S. avermitilis*).

Expression of the terpene cyclase YP_003342315 from *Streptosporangium roseum* DSM 43021 in *E. coli* resulted in the production of sesquiterpenes (Figure 4A) and monoterpenes (Figure 4B). The main compound in the sesquiterpene fraction of the volatiles was identified as δ-cadinene (**13**); minor products were α-muurolene (**14**), T-muurolol (**15**), α-copaene (**16**), β-copaene (**17**), cadina-1,4-diene (**18**), *trans*-cadina-1(6),4-diene (**19**), 1-*epi*-cubenol (**20**), cubebol (**21**), *epi*-cubebol (**22**), β-cubebene (**23**), α-cubebene (**24**), zonarene (**25**), bicyclosesquiphellandrene (**26**) and cadina-3,5-diene (**27**). All these compounds were also observed in headspace extracts from *S. roseum* (Figure S2). Incubation of the recombinant enzyme with FPP followed by liquid–liquid extraction yielded essentially the same products as detected in the headspace extracts; relative amounts differed, as much larger quantities of sesquiterpene alcohols were found (Figure 4C). The main product was alcohol **22**, and thus the enzyme is best described as an "*epi*-cubebol synthase". In addition, the sesquiterpene alcohol germacradien-4-ol (**12**) was observed in the in vitro incubation of the *epi*-cubebol synthase with FPP, but not in the headspace ex-

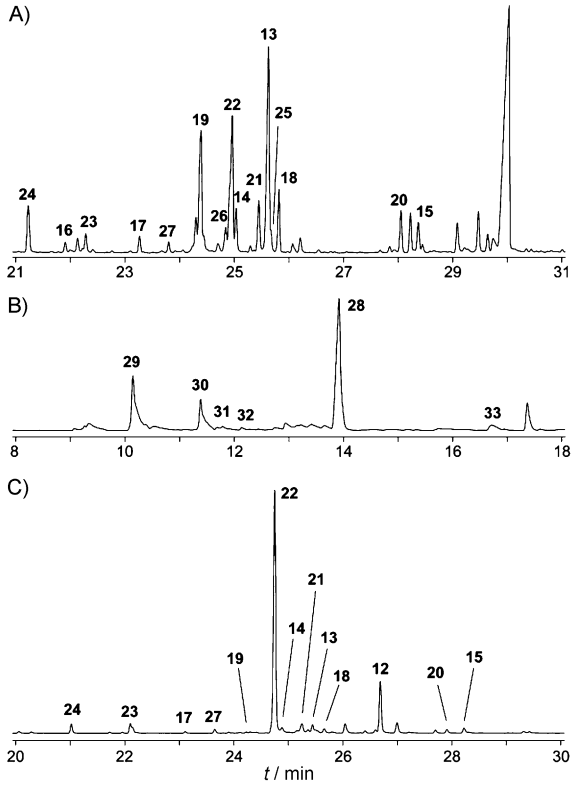
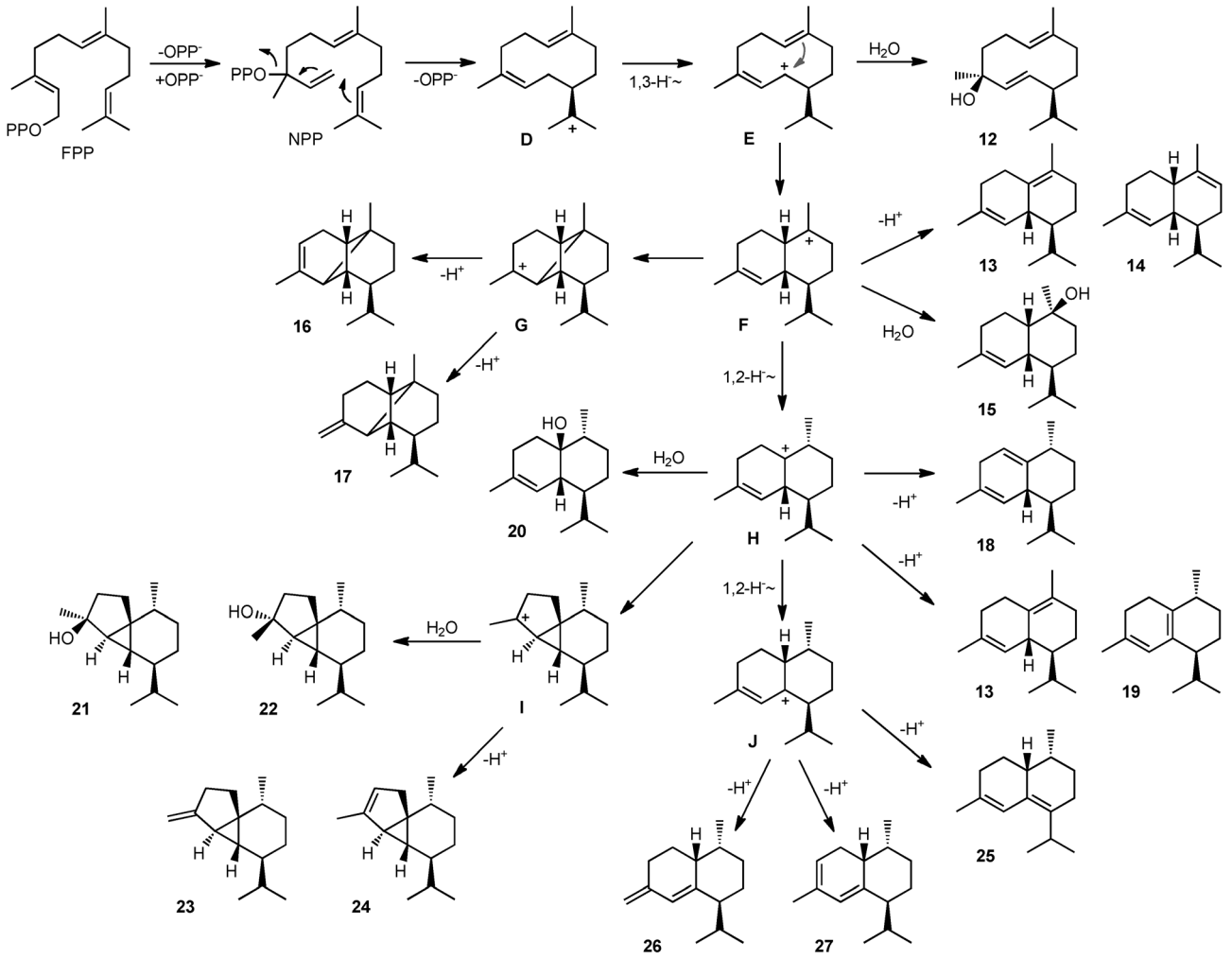


Figure 4. Total ion chromatogram of a headspace extract from *E. coli* upon heterologous expression of the *epi*-cubebol synthase from *S. roseum*. A) Sesquiterpenes and B) monoterpenes (peak numbers refer to compound numbers in Schemes 2 and 3); chromatogram scales are identical. C) Total ion chromatogram of an extract from incubation of purified *epi*-cubebol synthase with FPP.



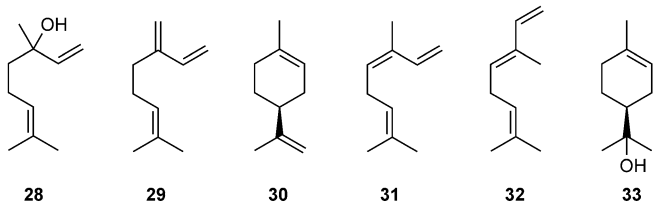
Scheme 2. Suggested biosynthetic pathway to all sesquiterpene products of the *epi*-cubebol synthase from *S. roseum*. For chiral terpenes, relative configurations are shown.

tracts obtained from *S. roseum* or from *E. coli* upon expression of the *epi*-cubebol synthase. Its formation was strongly pH dependent, with considerably higher production under slightly basic conditions (pH 8.5) than with neutral (pH 7.0) or slightly acidic conditions (pH 6.0, Figure S2).

Biosynthetic pathways to all these sesquiterpenes via common cationic intermediates are suggested in Scheme 2. These start with isomerisation of FPP to nerolidyl diphosphate (NPP), followed by cyclisation to the (*E*,*Z*)-germacradienyl cation (**D**), and a 1,3-hydride shift to cation **E**, which can be attacked by water to yield **12**. The observed strong pH dependency of this reaction is likely a result of the higher nucleophilicity of water at basic pH (pH 8.5) than at neutral or acidic pH (pH 6.0), similarly to the pH effect that we recently described for the hedycaryol synthase from *Kitasatospora setae*.[23] The fact that cultures of *S. roseum* do not emit **12** suggests that the intracellular reaction conditions under which the *epi*-cubebol synthase acts are also slightly acidic. Cyclisation of **E** to cation **F** provides the last common intermediate for all other sesquiterpenes. Hydrocarbons **13** and **14** and alcohol **15** arise from **F** by alternative deprotonations or attack by water. A second cyclisation of **F** to **G** gives rise to **16** and **17** by proton losses, whereas 1,2-hydride migration to **H** paves the way to **13**, **18** and **19** by deprotonations, or to **20** by capture with water. Cation **H** can undergo cyclisation to **I**, the direct precursor to **21**–**24** (main product, **22**). Finally, another 1,2-hydride shift in **H** results in **J**, from which hydrocarbons **25**–**27** can be formed by proton losses.

The monoterpene fraction detected upon expression of the *S. roseum epi*-cubebol synthase in *E. coli* included linalool (**28**), along with the minor compounds myrcene (**29**), limonene (**30**), (*Z*)-β-ocimene (**31**), (*E*)-β-ocimene (**32**) and α-terpineol (**33**, Figure 4B and Scheme 3). Accordingly, incubation of the purified enzyme with GPP resulted in the formation of three major compounds, **28**–**30**; the trace compounds **31**–**33** were not observed in this experiment (Figure S2). Only **30** was detected in headspace extracts from *S. roseum* (Figure S3).

In summary, three bacterial terpene cyclases have been chemically characterised in this work. For this, vector pYE-Express was constructed for gene cloning by homologous recombination in yeast and subsequent heterologous expression in *E. coli*. By using this technique, one enzyme from *S. espanaensis*



Scheme 3. Monoterpenes made by the *epi*-cubebol synthase from *S. roseum*.

was identified as caryophyllene synthase by incubation of the recombinant enzyme with FPP, thus confirming our previous suggestion for the function of this enzyme.[22] One enzyme from *S. spinosa* yielded complex mixtures of terpenes, with hedycaryol as the main product after incubation with FPP, whereas the more volatile terpene hydrocarbon (*E*)-β-farnesene was the main product found in CLSA headspace extracts after heterologous expression in *E. coli*. The third enzyme, a terpene cyclase from *S. roseum*, produced mixtures of various sesqui- and monoterpenes upon heterologous expression in *E. coli*; δ-cadinene and linalool were the main compounds, whereas incubation experiments with FPP and GPP yielded mainly *epi*-cubebol and linalool. These data demonstrate that if complex mixtures of terpenes are formed, application of the CLSA headspace technique during heterologous expression of the enzyme might lead to biased results because of the different volatilities of the products. Another interesting observation was the strong pH dependency of the formation of germacradien-4-ol by the *epi*-cubebol synthase from *S. roseum*, whereas all other products were not influenced by the pH over the range pH 6–8.5.

The information on the chemical nature of the products from bacterial terpene cyclases obtained from this and all previous work is summarised in a phylogenetic tree (Figure S4). As can be seen, several bacterial terpene cyclases have already been chemically characterised, but there is still room for future investigations into these interesting enzymes.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft with a Heisenberg Fellowship (DI1536/4-1) and the grant "Duftstoffe aus Actinomyceten" (DI1536/2-1, to J.S.D.), and by the Beilstein Institut zur Förderung der Chemischen Wissenschaften with a PhD scholarship (to P.R.). We thank Stefan Schulz (Braunschweig) for support of our group.

Keywords: actinomycetes · biosynthesis · enzymes · GC/MS · heterologous expression · terpene cyclases

[1] J. S. Dickschat, *Nat. Prod. Rep.* **2011**, *28*, 1917 – 1936.
[2] N. L. Brock, J. S. Dickschat in *Natural Products: Phytochemistry, Botany and Metabolism of Alkaloids, Phenolics and Terpenes* (Eds.: K. G. Ramawat, J. M. Mérillon), Springer, Berlin, **2013**, pp. 2693 – 2732.
[3] N. L. Brock, S. R. Ravella, S. Schulz, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2013**, *52*, 2100 – 2104; *Angew. Chem.* **2013**, *125*, 2154 – 2158.
[4] D. E. Cane, J.-K. Sohng, C. R. Lamberson, S. M. Rudnicki, Z. Wu, M. D. Lloyd, J. S. Oliver, B. R. Hubbard, *Biochemistry* **1994**, *33*, 5846 – 5857.
[5] X. Lin, R. Hopson, D. E. Cane, *J. Am. Chem. Soc.* **2006**, *128*, 6022 – 6023.
[6] W. K. W. Chou, I. Fanizza, T. Uchiyama, M. Komatsu, H. Ikeda, D. E. Cane, *J. Am. Chem. Soc.* **2010**, *132*, 8850 – 8851.
[7] Y. Hu, W. K. W. Chou, R. Hopson, D. E. Cane, *Chem. Biol.* **2011**, *18*, 32 – 37.
[8] C. Nakano, H.-K. Kim, Y. Ohnishi, *ChemBioChem* **2011**, *12*, 1988 – 1991.
[9] C. Nakano, H.-K. Kim, Y. Ohnishi, *ChemBioChem* **2011**, *12*, 2403 – 2407.
[10] C. Nakano, S. Horinouchi, Y. Ohnishi, *J. Biol. Chem.* **2011**, *286*, 27980 – 27987.
[11] C. Nakano, T. Tezuka, S. Horinouchi, Y. Ohnishi, *J. Antibiot.* **2012**, *65*, 551 – 558.
[12] C. Nakano, F. Kudo, T. Eguchi, Y. Ohnishi, *ChemBioChem* **2011**, *12*, 2271 – 2275.
[13] S. A. Agger, F. Lopez-Gallego, T. R. Hoye, C. Schmidt-Dannert, *J. Bacteriol.* **2008**, *190*, 6084 – 6096.
[14] D. E. Cane, R. M. Watt, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1547 – 1551.
[15] D. E. Cane, X. He, Y. Kobayashi, S. Ômura, H. Ikeda, *J. Antibiot.* **2006**, *59*, 471 – 479.
[16] S. Giglio, J. Jiang, C. P. Saint, D. E. Cane, P. T. Monis, *Environ. Sci. Technol.* **2008**, *42*, 8027 – 8032.
[17] C.-M. Wang, D. E. Cane, *J. Am. Chem. Soc.* **2008**, *130*, 8908 – 8909.
[18] M. Komatsu, M. Tsuda, S. Ômura, H. Oikawa, H. Ikeda, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7422 – 7427.
[19] S. Giglio, W. K. W. Chou, H. Ikeda, D. E. Cane, P. T. Monis, *Environ. Sci. Technol.* **2011**, *45*, 992 – 998.
[20] C. A. Citron, J. Gleitzmann, G. Laurenzano, R. Pukall, J. S. Dickschat, *ChemBioChem* **2012**, *13*, 202 – 214.
[21] P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2013**, *52*, 1810 – 1812; *Angew. Chem.* **2013**, *125*, 1855 – 1857.
[22] P. Rabe, C. A. Citron, J. S. Dickschat, *ChemBioChem* **2013**, *14*, 2345 – 2354.
[23] P. Baer, P. Rabe, C. A. Citron, C. C. de Oliveira Mann, N. Kaufmann, M. Groll, J. S. Dickschat, *ChemBioChem* **2014**, *15*, 213 – 216.
[24] C. A. Lesburg, G. Zhai, D. E. Cane, D. W. Christianson, *Science* **1997**, *277*, 1820 – 1824.
[25] J. A. Aaron, X. Lin, D. E. Cane, D. W. Christianson, *Biochemistry* **2010**, *49*, 1787 – 1797.
[26] M. Köksal, W. K. W. Chou, D. E. Cane, D. W. Christianson, *Biochemistry* **2012**, *51*, 3011 – 3020.
[27] E. Oldfield, F.-Y. Lin, *Angew. Chem. Int. Ed.* **2012**, *51*, 1124 – 1137; *Angew. Chem.* **2012**, *124*, 1150 – 1163.
[28] J. Jiang, X. He, D. E. Cane, *Nat. Chem. Biol.* **2007**, *3*, 711 – 715.

Received: December 5, 2013
Published online on February 26, 2014