How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 20308–20312
International Edition: doi.org/10.1002/anie.202106718
German Edition: doi.org/10.1002/ange.202106718

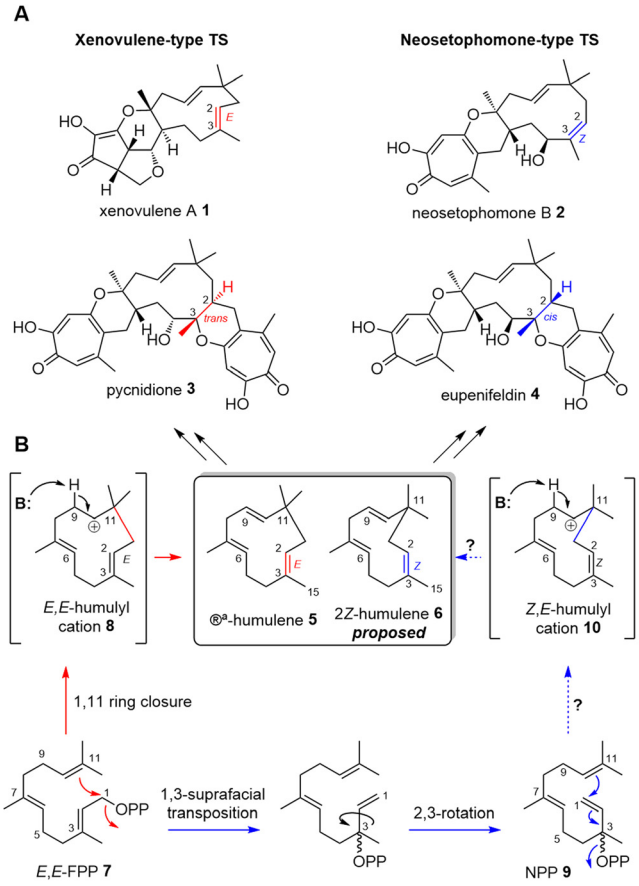
Biosynthesis

Understanding and Engineering the Stereoselectivity of Humulene Synthase

Carsten Schotte, Peer Lukat, Adrian Deuschmann, Wulf Blankenfeldt, and Russell J. Cox**

Abstract: *The non-canonical terpene cyclase AsR6 is responsible for the formation of 2E,6E,9E-humulene during the biosynthesis of the tropolone sesquiterpenoid (TS) xenovulene A. The structures of unliganded AsR6 and of AsR6 in complex with an in crystallo cyclized reaction product and thiolodiphosphate reveal a new farnesyl diphosphate binding motif that comprises a unique binuclear Mg^{2+}-cluster and an essential K289 residue that is conserved in all humulene synthases involved in TS formation. Structure-based site-directed mutagenesis of AsR6 and its homologue EupR3 identify a single residue, L285/M261, that controls the production of either 2E,6E,9E- or 2Z,6E,9E-humulene. A possible mechanism for the observed stereoselectivity was investigated using different isoprenoid precursors and results demonstrate that M261 has gatekeeping control over product formation.*

Tropolone sesquiterpenoids (TS) are privileged meroterpenoid natural products of hybrid terpene/polyketide origin. Representative examples (e.g. xenovulene A **1**, neosetophomone B **2**, pycnidione **3**, eupenifeldin **4**; Scheme 1 A) possess potent psychoactive, antitumor and antiarthritic properties.[1–5] All TS feature a core 11-membered macrocycle that is derived from the sesquiterpene humulene.[6,7] The biosynthesis of TS proceeds via a core intermolecular hetero Diels–Alder reaction that connects humulene with the polyketide-derived tropolone nucleus observed in **1**–**4** (Electronic Supplementary Information [ESI] Figure S1–S2).[6] TS can further be divided into two major subfamilies.[6] Xenovulene-type TS are derived from 2*E*,6*E*,9*E*-humulene **5** and the 2*E*-alkene configuration in **5** is mirrored in the final pathway product (2*E*-alkene in **1**; *trans*-fusion at the C-2/C-3 ring junction in **3**; Scheme 1 A,B).[8] Neosetophomone-type TS display a 2*Z*-alkene configuration in monosubstituted pathway products (e.g. **2**) and *cis* fusion at the C-2/C-3 ring junction in disubstituted TS such as **4** (Scheme 1 A,B). 2*Z*,6*E*,9*E*-humulene **6** was proposed as the precursor of **2** and **4**, to explain the difference in stereochemistry, but there is only limited evidence for the existence of **6** as a biosynthetic intermediate.[6]



Scheme 1. **A**, representative examples of tropolone sesquiterpenoids; **B**, proposed humulene formation during the biosynthesis of **1**–**4**; dashed arrows = putative steps in the formation of **6**; B = general base.

Sesquiterpenes such as **5** are biosynthesized by class I terpene cyclases from the universal precursor 2*E*,6*E*-farnesyl diphosphate **7** (FPP).[9] All known class I terpene cyclases share a common α-helical fold and two conserved aspartate-rich motifs (**DDxxD** and **NSE/DTE**; bold indicates metal

[*] C. Schotte, A. Deuschmann, Prof. Dr. R. J. Cox
Institute for Organic Chemistry and BMWZ, Leibniz Universität Hannover
Schneiderberg 38, 30167 Hannover (Germany)
E-mail: carsten.schotte@oci.uni-hannover.de
russell.cox@oci.uni-hannover.de
Dr. P. Lukat, Prof. Dr. W. Blankenfeldt
Structure and Function of Proteins, Helmholtz Centre for Infection Research
Inhoffenstr. 7, 38124 Braunschweig (Germany)
Prof. Dr. W. Blankenfeldt
Institute for Biochemistry, Biotechnology and Bioinformatics, Technische Universität Braunschweig
Spielmannstr. 7, 38106 Braunschweig (Germany)
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202106718.

binding) that coordinate a trinuclear Mg^{2+}-cluster responsible for binding **7** and triggering the loss of diphosphate (PP_i).[9] Cleavage of PP_i generates a reactive allylic cation that can undergo diverse intramolecular cyclizations, rearrangements and final proton loss.[9]

From a mechanistic viewpoint 2*E*-humulene synthases belong to the simplest cyclases. 1,11-cyclization of FPP **7** yields the *E,E*-humulyl cation **8** and deprotonation then affords **5** without further rearrangement (Scheme 1B). The occurrence of *Z*-configured alkenes in terpenes (e.g. **6**) requires a formal isomerization of the 2*E*-alkene in FPP **7**, that is assumed to proceed via a 1,3-suprafacial transposition of the PP_i in **7** to give *cisoid* nerolidyl diphosphate **9b** (NPP) after rotation of the C-2/C-3 bond (Scheme 1B).[10–12] Cleavage of PP_i and 1-11-cyclization would yield the *Z,E*-humulyl cation **10** and deprotonation could then afford **6**. However, the structural factors that control the transposition from **7** to **9** are poorly understood.

Two cryptic terpene cyclases (AsR6, PycR6) have recently been identified as responsible for 2*E*-humulene **5** formation in TS natural products.[8,13] They bear no significant sequence identity to any known terpene synthase and the typical metal-binding motifs involved in FPP **7** coordination are not identifiable.[8] Homologous enzymes (EupE; EupfG and EupR3) have been linked to the production of neosetophomone B **2** and eupenifeldin **4**, but the terpene product of these enzymes was not directly identified.[6,13,14]

Here, the *E. coli* codon-optimized genes *asR6* and *eupR3* were expressed in *E. coli* BL21 as polyhistidine-tagged enzymes. Incubation of AsR6 with FPP **7** gives, as previously reported,[7] 2*E*-humulene **5** (*m/z* 204). Production of **5** was confirmed by gas chromatography mass spectrometry (GCMS) analysis (Figure 1A; ESI Figure S3).[8] Identical in vitro assays with EupR3 afforded a single product with the same molecular mass (*m/z* 204) as **5**, but a different retention time (Figure 1B). A preparative-scale biotransformation of FPP **7** with EupR3 allowed analysis of the terpene product by nuclear magnetic resonance (NMR). Structure elucidation confirmed the product as the expected 2*Z*-humulene **6** (ESI Figure S4–S23; Table S1).[6] The different configurations of the C-2/C-3 alkene in **5** and **6** were confirmed by NOESY NMR. Key nOe correlations of H_3-15 to H-2 in **6**, but from H_3-15 to H_2-1 in **5**, together with other correlations, confirm the expected alkene configuration (ESI Figure S24; Table S1 + S2).

Next, we obtained the crystal structure of AsR6 in its open conformation (2.0 Å, PDB: 7OC5, Table S7) and in complex with the synthetic substrate analogue farnesyl-thiolodiphosphate (FSPP, 2.0 Å, PDB: 7OC4, Figure 2A,B, Table S7).[15,16] AsR6 crystallized as a homodimer, which was confirmed by analysis with PDBePISA[17] and is in accordance with multi angle light scattering data ([MALS] ESI Figure S25). Each polypeptide chain folds into a single, purely α-helical domain consisting of eleven helices (HL) with HL-3 to HL-10 forming a barrel-like core with a central hydrophobic cavity. The overall protein fold of AsR6 is similar to other Type I sesquiterpene cyclases. DALI analysis[18] identified the bacterial selinadiene synthase (PDB: 4OKM) as the closest structural relative (12% sequence identity, RMSD 4.1 Å,

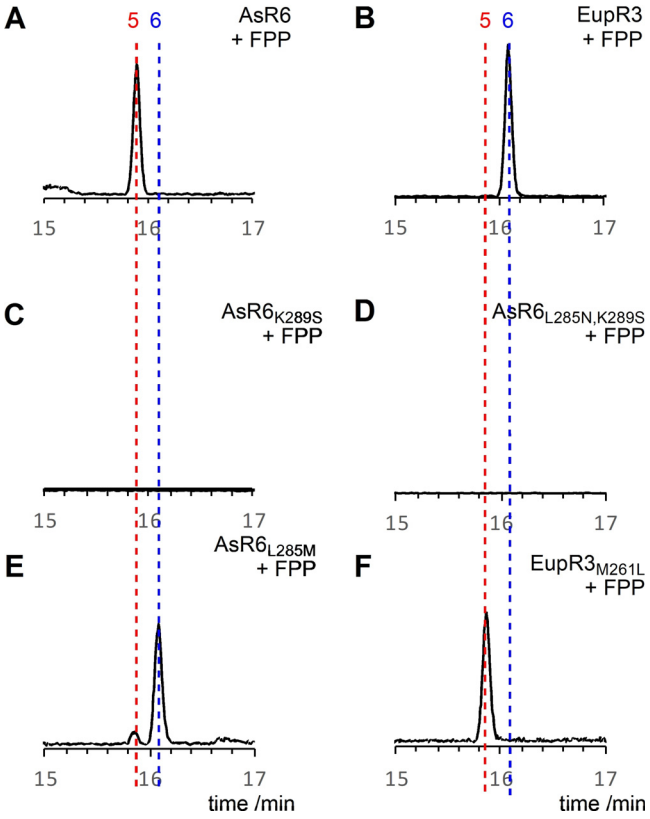


Figure 1. Total ion chromatograms of extracts from assays of wildtype enzymes and mutants with FPP **7**: **A,** incubation of AsR6 with **7**; **B,** incubation of EupR3 with **7**; **C,** incubation of $AsR6_{K289S}$ with **7**; **D,** incubation of $AsR6_{L285N\_K289S}$ with **7**; **E,** incubation of $AsR6_{L285M}$ with **7**; **F,** incubation of $EupR3_{M261L}$ with **7**.

ESI Table S3 and Figure S26).[18,19] To our surprise FSPP did not crystallize intact. Analysis of the electron density identified a cleaved thiolodiphosphate (SPP_i) moiety and a ring-like structure within the active site that best matches the reaction product 2*E*-humulene **5**. The exact conformation of the ring-like structure could not be determined unambiguously and the displayed structure is in best agreement with the observed electron density (ESI Figure S45).

The SPP_i is coordinated by R240, K289, R350, Y351 and R354, whereas W138 forms an aromatic wall of the reaction chamber and may be involved in stabilizing cationic intermediates (Figure 2C). R350 and Y351 form a conserved RY-dimer that is widely encountered in class I terpene cyclases.[20] Intriguingly, despite the overall structural similarity to canonical terpene cyclases, analysis of the electron density map identifies only two coordinated Mg_{A+B}^{2+}-ions (Figure 2C). Canonical terpene cyclases typically harbour a trinuclear Mg^{2+} cluster.[21] Both Mg^{2+}-ions in AsR6 are coordinated by the carboxylate of D164 (situated on an aspartate-rich motif on HL-4).

Structure comparison with known fungal terpene cyclases reveals identical aspartates for example, in trichodiene synthase[22] and aristolochene synthase,[23] that coordinate the equivalent Mg_A^{2+} and Mg_B^{2+} ions, respectively (ESI Figure S27 + 28). However, whereas in canonical terpene cyclases the third Mg_C^{2+}-ion is coordinated by a second metal

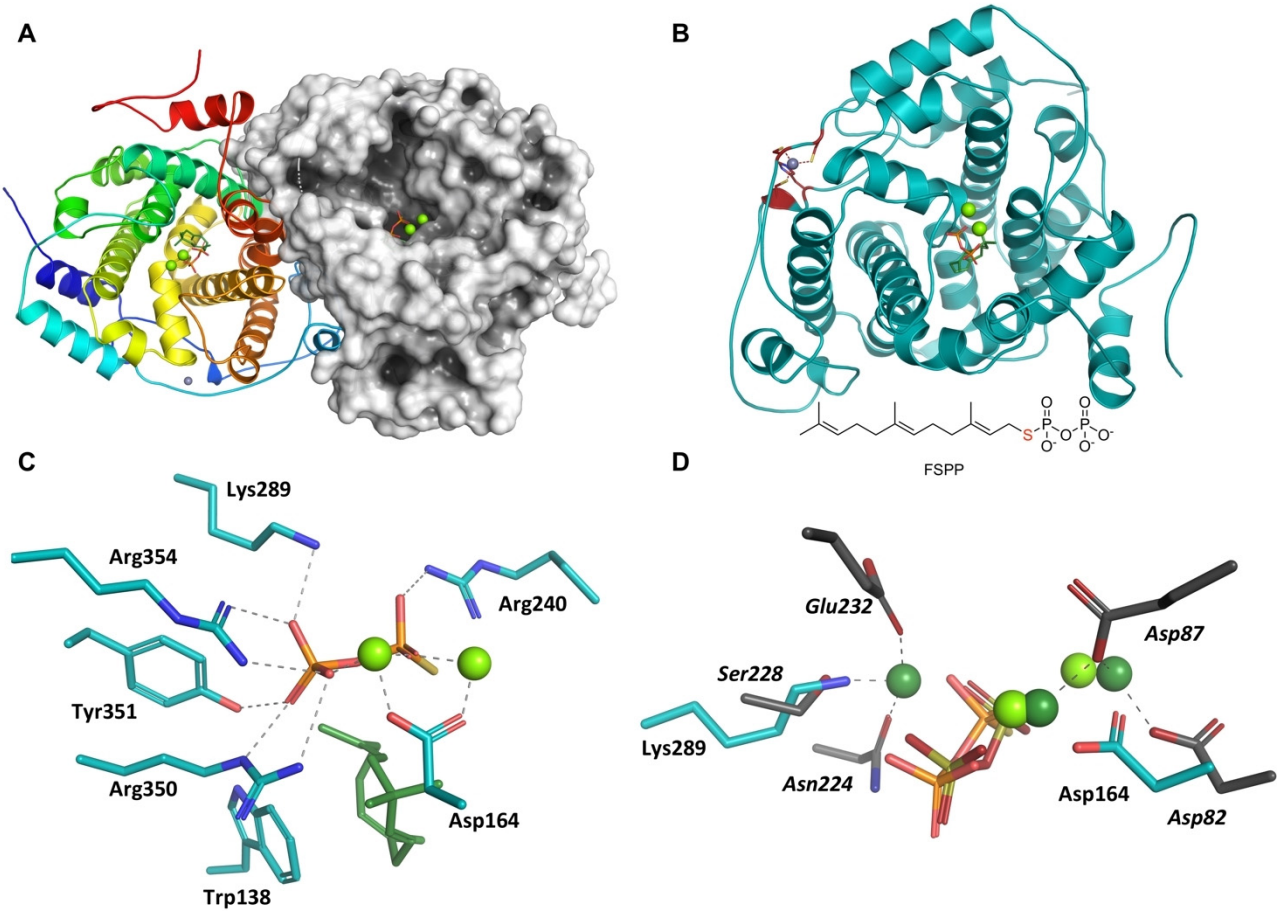


Figure 2. Crystal structure of AsR6: **A**, crystal structure of the AsR6 dimer; chain A shown as cartoon coloured from *N*- (blue) to *C*-terminus (red); chain B shown in surface representation; ligands are represented as sticks/spheres; **B**, crystal structure of AsR6 in complex with *in crystallo* cyclized reaction product and thiolodiphosphate; Zn^{2+}-binding site highlighted in red; **C**, AsR6 Mg^{2+}-binding site showing bound humulene (green); **D**, structural alignment of the diphosphate binding site in AsR6 (Mg^{2+} = light green; residues = cyan) and selinadiene synthase (PDB: 4okm, Mg^{2+} = dark green; residues = grey, italic labels).

binding motif ($\mathbf{N}_{224}\mathbf{S}_{228}\mathbf{E}_{232}$ in selinadiene-synthase; Figure 2D),[19] in AsR6 the Mg_C^{2+}-ion is replaced by the ε-ammonium of K289 that occupies the corresponding position in the active site (Figure 2D). K289 is located on the opposite side of the active site on HL-7 and coordinates the diphosphate via its lysyl side chain.

To validate this highly unusual diphosphate binding motif, we mutated K289. Incubation of FPP **7** with K289S and analysis by GCMS showed that formation of 2*E*-humulene **5** was abolished, consistent with the proposed crucial role of K289 in pyrophosphate coordination (Figure 1C). We then probed whether the typically observed NSE-motif could be (re)introduced into AsR6, to reconstitute the canonical diphosphate binding motif. Structural alignment of AsR6 with trichodiene and aristolochene synthases reveals that the double mutant $AsR6_{L285N\_K289S}$ would fulfil this requirement, as the complementary E298 already occupies the appropriate position in the active site (ESI Figure S29). However, analysis of the $AsR6_{L285N\_K289S}$ double mutant revealed that enzyme activity was not restored upon incubation with FPP **7** (Figure 1D).

To complete the structural characterization of AsR6 we report coordination of a Zn^{2+} ion by C43, C85, C87 and H41. This binding motif is conserved in the AsR6-like humulene synthases (ESI Figure S30), but absent in canonical class I terpene cyclases and might be involved in protein stability but was not further investigated as it is located away from the active site (Figure 2B).

Despite sharing ≈50% sequence identity, AsR6 and EupR3 each give rise to a single humulene isomer (**5** or **6**) in the presence of FPP **7** and exhibit complete control of the stereochemical outcome. To identify structural features that drive the different product configurations we identified all residues in the active site of AsR6 that are within 6 Å of the bound FSPP/humulene substrate (ESI Figure S31). Mapping of these residues to a global sequence alignment of the humulene synthases involved in formation of **5** (AsR6, PycR6) and **6** (EupR3, EupE) showed that all active site residues are highly conserved in both types of humulene synthases. Only one residue is consistently different in the two types of cyclases: humulene cyclases producing **5** have a conserved L285 residue that is replaced by a methionine in cyclases giving **6** (ESI Figure S32).

L285 extends into the active site cavity of AsR6 with one of its C_δ atoms in close proximity to the C-2/C-3 alkene (4.5 Å and 3.9 Å; Figure 3A). We reasoned that in the absence of

other conserved changes residues at this position are likely involved in controlling product stereoselectivity.

Site-directed mutagenesis afforded the *cross-convergent mutants* $AsR6_{L285M}$ and $EupR3_{M261L}$.[24] Incubation of the $AsR6_{L285M}$ mutant with FPP **7** and analysis by GCMS identified production of minor 2*E*-humulene **5** (12%; based on peak integration) and major 2*Z*-humulene **6** (>85%; Figure 1E). Analysis of the $EupR3_{M261L}$ mutant validated the role of this residue.

Here, production of the native terpene product **6** is abolished and $EupR3_{M261L}$ exclusively produces **5** instead (Figure 1F). Notably, the introduced M261L mutation leads to a complete switch in the observed product stereochemistry and indicates that the presence of methionine at this position is a prerequisite for **6** formation. Further site-directed mutagenesis experiments were directed to investigate the influence of smaller amino acids on this position (L285A and L285V) but did not affect product formation (ESI Figure S33).

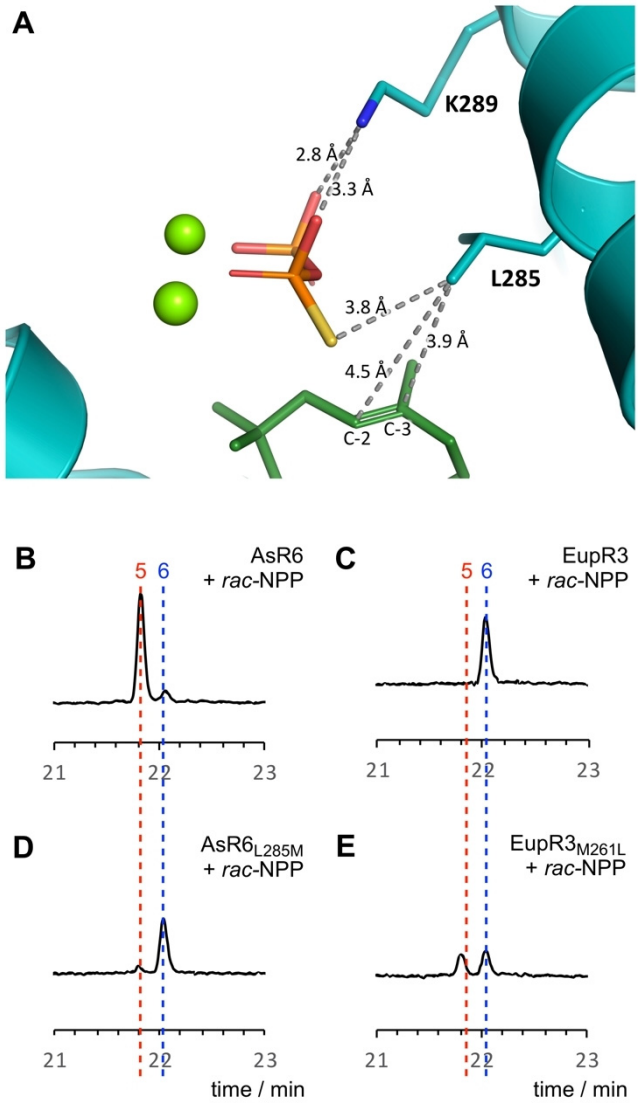


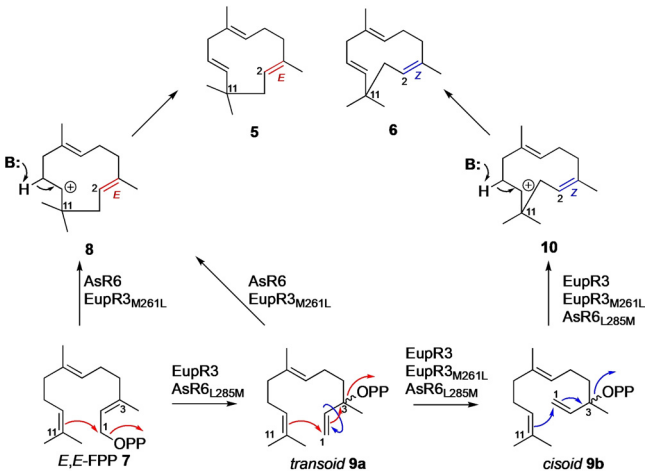
Figure 3. Investigations into the stereoselectivity of humulene formation: **A**, position of key residue L285 in the active site; **B**, total ion chromatogram of extracts from incubation of AsR6 with *rac*-NPP **9**; **C**, incubation of EupR3 with **9**; **D**, incubation of $AsR6_{L285M}$ with **9**; **E**, incubation of $EupR3_{M261L}$ with **9**.

As formation of the 2*Z* alkene requires a formal isomerization of the 2*E* alkene in FPP **7**, presumably via the well-established initial formation of NPP **9** (Scheme 1B), we reasoned that M261 could be involved in the initial transformation of **7** to **9**. Alternatively, the residue could be involved in controlling the possible carbocation trajectories that lead to the *E,E*- or *Z,E*-humulyl cations **8** and **10** respectively. Cane et al. reported similar effects of an individual amino acid residue on the stereochemical outcome for amorpha-4,11-diene synthase (ESI Figure S34).[25] The single point mutant T296V abolished production of amorphadiene from FPP **7** and production was only restored upon incubation with (3*R*)-NPP **9**.[25]

We probed the in vitro activity of wild-type enzymes and mutants with synthetic (±) NPP **9**. Incubation of **9** with AsR6 affords the native terpene product **5** as the major product (Figure 3B). Minor formation of the *Z*-isomer **6** confirms that **9** is indeed a precursor of *Z*-humulene **6**—but in AsR6 this trajectory is less favorable. Incubation of EupR3 with **9** exclusively affords 2*Z*-humulene **6** (Figure 3C). For AsR6 the *transoid* conformer of the tertiary allylic isomer NPP, **9a**, is the likely true substrate (Scheme 2). EupR3 likely catalyzes the conversion of **9a** into the *cisoid* conformer **9b** and then formation of **6** (Scheme 2). GCMS analysis of the cross-convergent mutants revealed that $AsR6_{L285M}$ produces 2*Z*-humulene **6** and only traces of **5** were observed (Figure 3D).

Finally, incubation of **9** with $EupR3_{M261L}$ reconstituted the formerly abolished formation of **6**. However, intriguingly the experiment afforded a ≈1:1 mixture of **5** and **6** (Figure 3E). Formation of both **5** and **6** suggests that the methionine M261 is not the sole controlling factor during the transformation of FPP into NPP. The exact catalytic functions of this residue (control of the allylic isomerization of FPP into NPP vs. controlling different reaction trajectories) are currently under further investigation.

We also probed whether only one of the two NPP enantiomers is converted by AsR6. The (3*R*)-enantiomer of NPP **9** is usually observed to be the intermediate during



Scheme 2. Putative biosynthetic pathways leading to **5** and **6** from the different isoprenoid precursors; B = general base.

cyclisation reactions.[26] E.g., *Z*-γ-bisabolene synthase (BbS) selectively converts (3*R*)-NPP to *Z*-γ-bisabolene, while incubation with the (3*S*)-enantiomer affords the acyclic *E*-β-farnesene.[12] First, (±)-**9** was incubated with AsR6. Then BbS was added in order to convert any residual NPP to either *Z*-γ-bisabolene or *E*-β-farnesene. GCMS analysis revealed formation of 2*E*-humulene and no production of *Z*-γ-bisabolene when compared to a control reaction lacking AsR6, suggesting that AsR6 converts both (3*R*)-NPP and (3*S*)-NPP to 2*E*-humulene **5** (ESI Figure S351&2). However, more detailed studies are required to confirm this unusual conversion of both enantiomers of (±)-**9**.

To validate the conversion of (3*S*)-NPP by humulene synthases we synthesized the (3*S*)-NPP enantiomer (see ESI for details). As expected, incubation of (3*S*)-NPP with AsR6, EupR3 and $AsR6_{L285M}$ as well as $EupR3_{M261L}$ affords identical product distributions as observed for incubation with *rac*-NPP (ESI Figure S36).

In summary, we report the first crystal structure of a non-canonical humulene synthase that shows no significant sequence homology to other known class I terpene cyclases and humulene synthases. A novel diphosphate binding motif was identified and validated for class I terpene cyclases that comprises a binuclear magnesium cluster and a strictly conserved lysine K289. Binculear magnesium clusters have been identified in very few other enzymes (e.g. UbiA-type prenyltransferases)[27] but these show no significant homology to the fungal AsR6-like humulene synthases reported here (Supporting Figure S52), and do not feature the replacement of a magnesium ion by the ε-ammonium of a lysine residue. AsR6 can utilize FPP **7** and apparently both enantiomers of NPP **9** as substrates. A single amino acid residue drives the stereochemical outcome of 2*E* vs. 2*Z*-humulene formation. Our discoveries thus broaden the reservoir of fungal non-canonical terpene cyclases and pave the way for further engineering of new-to-nature tropolone sesquiterpenoids.

Acknowledgements

CS thanks Leibniz University for funding. DFG is thanked for funding (INST 187/686-1). Prof. Jeroen Dickschat is thanked for the gift of plasmid pYE-BbS and Vanessa Harms is thanked for the gift of farnesyl pyrophosphate. Luca Codutti and Georg Krüger are thanked for MALS measurements. Ute Widow is thanked for technical assistance. We acknowledge DESY (Hamburg, Germany), a member of the Helmholtz Association HGF, for the provision of experimental facilities. Parts of this research were carried out at beamline P11 at the PETRA III storage ring and we would like to thank the beamline staff for assistance during data collection. Beamtime was allocated for proposal Xh-20010031. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: biosynthesis · enzyme engineering · humulene · meroterpenoid · tropolone sesquiterpenoid

[1] A. M. Ainsworth, M. I. Chicarelli-Robinson, B. R. Copp, U. Fauth, P. J. Hylands, J. A. Holloway, M. Latif, G. B. O'Beirne, N. Porter, D. V. Renno, et al., *J. Antibiot.* **1995**, *48*, 568 – 573.
[2] T. El-Elimat, H. A. Raja, S. Ayers, S. J. Kurina, J. E. Burdette, Z. Mattes, R. Sabatelle, J. W. Bacon, A. H. Colby, M. W. Grinstaff, et al., *Org. Lett.* **2019**, *21*, 529 – 534.
[3] C.-J. Hsiao, S.-H. Hsiao, W.-L. Chen, J.-H. Guh, G. Hsiao, Y.-J. Chan, T.-H. Lee, C.-L. Chung, *Chem.-Biol. Interact.* **2012**, *197*, 23 – 30.
[4] C. Y. Bemis, C. N. Ungarean, A. S. Shved, C. S. Jamieson, T. Hwang, K. S. Lee, K. N. Houk, D. Sarlah, *J. Am. Chem. Soc.* **2021**, *143*, 6006 – 6017.
[5] Z. Y. Al Subeh, N. Q. Chu, J. T. Korunes-Miller, L. L. Tsai, T. N. Graf, Y. P. Hung, C. J. Pearce, M. W. Grinstaff, A. H. Colby, Y. L. Colson, et al., *J. Controlled Release* **2021**, *331*, 260 – 269.
[6] Q. Chen, J. Gao, C. Jamieson, J. Liu, M. Ohashi, J. Bai, D. Yan, B. Liu, Y. Che, Y. Wang, et al., *J. Am. Chem. Soc.* **2019**, *141*, 14052 – 14056.
[7] F. Yu, S. Okamto, K. Nakasone, K. Adachi, S. Matsuda, H. Harada, N. Misawa, R. Utsumi, *Planta* **2008**, *227*, 1291 – 1299.
[8] R. Schor, C. Schotte, D. Wibberg, J. Kalinowski, R. J. Cox, *Nat. Commun.* **2018**, *9*, 1963.
[9] J. D. Rudolf, C. Y. Chang, *Nat. Prod. Rep.* **2020**, *37*, 425 – 463.
[10] D. Arigoni, *Pure Appl. Chem.* **1975**, *41*, 219 – 245.
[11] D. E. Cane, R. Iyengar, M. S. Shiao, *J. Am. Chem. Soc.* **1981**, *103*, 914 – 931.
[12] J. Rinkel, J. S. Dickschat, *Beilstein J. Org. Chem.* **2019**, *15*, 789 – 794.
[13] C. Schotte, L. Li, D. Wibberg, J. Kalinowski, R. J. Cox, *Angew. Chem. Int. Ed.* **2020**, *59*, 23870 – 23878; *Angew. Chem.* **2020**, *132*, 24079 – 24087.
[14] Y. Zhai, Y. Li, J. Zhang, Y. Zhang, F. Ren, X. Zhang, G. Liu, X. Liu, Y. Che, *Fungal Genet. Biol.* **2019**, *129*, 7 – 15.
[15] R. M. Phan, C. D. Poulter, *J. Org. Chem.* **2001**, *66*, 6705 – 6710.
[16] G. Ramamoorthy, R. M. Phan, C. D. Poulter, *J. Org. Chem.* **2016**, *81*, 5093 – 5100.
[17] E. Krissinel, K. Henrick, *J. Mol. Biol.* **2007**, *372*, 774 – 797.
[18] L. Holm, *Protein Sci.* **2020**, *29*, 128 – 140.
[19] P. Baer, P. Rabe, K. Fischer, C. A. Citron, T. A. Klapschinski, M. Groll, J. S. Dickschat, *Angew. Chem. Int. Ed.* **2014**, *53*, 7652 – 7656; *Angew. Chem.* **2014**, *126*, 7783 – 7787.
[20] P. Rabe, T. Schmitz, J. S. Dickschat, *Beilstein J. Org. Chem.* **2016**, *12*, 1839 – 1850.
[21] D. W. Christianson, *Chem. Rev.* **2017**, *117*, 11570 – 11648.
[22] M. J. Rynkiewicz, D. E. Cane, D. W. Christianson, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13543 – 13548.
[23] 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.
[24] T. G. Köllner, J. Gershenzon, J. Degenhardt, *Phytochemistry* **2009**, *70*, 1139 – 1145.
[25] Z. Li, R. Gao, Q. Hao, H. Zhao, L. Cheng, F. He, L. Liu, X. Liu, W. K. W. Chou, H. Zhu, et al., *Biochemistry* **2016**, *55*, 6599 – 6604.
[26] M. B. Quin, C. M. Flynn, C. Schmidt-Dannert, *Nat. Prod. Rep.* **2014**, *31*, 1449 – 1473.
[27] W. Cheng, W. Li, *Science* **2014**, *343*, 878 – 881.

Manuscript received: May 19, 2021
Revised manuscript received: June 21, 2021
Accepted manuscript online: June 28, 2021
Version of record online: August 11, 2021

