Synthetic and Systems Biotechnology 9 (2024) 11–18

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Cyclization mechanism of monoterpenes catalyzed by monoterpene synthases in dipterocarpaceae

Xiaoyun Lu [a,b,1], Jie Bai [b,1], Zunzhe Tian [a,1], Congyu Li [b], Nida Ahmed [b], Xiaonan Liu [b], Jian Cheng [b], Lina Lu [b], Jing Cai [a], Huifeng Jiang [b,*], Wen Wang [a,**]

[a] School of Ecology and Environment, Northwestern Polytechnical University, Xi'an, Shanxi, 710072, China
[b] Key Laboratory of Engineering Biology for Low-Carbon Manufacturing, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China

ARTICLE INFO

Keywords:
Cyclization
Monoterpene synthases
Intermediate stabilization

ABSTRACT

Monoterpenoids are typically present in the secretory tissues of higher plants, and their biosynthesis is catalyzed by the action of monoterpene synthases (MTSs). However, the knowledge about these enzymes is restricted in a few plant species. MTSs are responsible for the complex cyclization of monoterpene precursors, resulting in the production of diverse monoterpene products. These enzymatic reactions are considered exceptionally complex in nature. Therefore, it is crucial to understand the catalytic mechanism of MTSs to elucidate their ability to produce diverse or specific monoterpenoid products. In our study, we analyzed thirteen genomes of Dipterocarpaceae and identified 38 MTSs that generate a variety of monoterpene products. By focusing on four MTSs with different product spectra and analyzing the formation mechanism of acyclic, monocyclic and bicyclic products in MTSs, we observed that even a single amino acid mutation can change the specificity and diversity of MTS products, which is due to the synergistic effect between the shape of the active cavity and the stabilization of carbon-positive intermediates that the mutation changing. Notably, residues N340, I448, and phosphoric acid groups were found to be significant contributors to the stabilization of intermediate terpinyl and pinene cations. Alterations in these residues, either directly or indirectly, can impact the synthesis of single monoterpenes or their mixtures. By revealing the role of key residues in the catalytic process and establishing the interaction model between specific residues and complex monoterpenes in MTSs, it will be possible to reasonably design and engineer different catalytic activities into existing MTSs, laying a foundation for the artificial design and industrial application of MTSs.

1. Introduction

Dipterocarpaceae plants have served as a primary source of natural monoterpenoids, which have been used for treating cardiovascular diseases including stroke and coronary heart diseases in China for over 2000 years [1–3]. However, due to over-harvesting, many species within Dipterocarpaceae family are now endangered [4]. To solve this problem, synthetic biology offers a promising alternative approach by enabling the engineering of biosynthetic pathways in microbes to synthesize monoterpenes [5–8]. Microbes have demonstrated significant capabilities to produce certain monoterpenes [9,10]. The abundance of monoterpenes' carbon skeleton can be attributed to diverse monoterpene synthases (MTSs), which convert the C10 precursor geranyl pyrophosphate (GPP) into cyclic or acyclic products [11–14]. However, only five MTSs have been identified in the Dipterocarpaceae species so far [15]. Therefore, mining more MTSs in Dipterocarpaceae species holds great potential in reconstructing the biosynthetic pathway of different monoterpenes in microbes.

Monoterpene cyclization by MTSs is one of the most complex chemical reactions in nature [16]. Over half of the carbon atoms in the substrate undergo modifications in bonding, hybridization, and stereochemistry during the multi-step catalytic cyclization process [17,18],

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author.
** Corresponding author.
E-mail addresses: jiang_hf@tib.cas.cn (H. Jiang), wwang@mail.kiz.ac.cn (W. Wang).
[1] Xiaoyun Lu, Jie Bai and Zunzhe Tian contributed equally to this work.

https://doi.org/10.1016/j.synbio.2023.11.009
Received 20 September 2023; Received in revised form 7 November 2023; Accepted 25 November 2023
Available online 4 December 2023

which will form a variety of products [19]. For instance, (+)-sabinene synthase, one of the first cloned MTSs, demonstrated a broad product spectrum *in vitro*, containing 63 % (+)-sabinene, along with 21 % terpinene, 7.0 % terpinolene, 6.5 % limonene and 2.5 % myrcene [20]. In contrast, borneol diphosphate synthases (BPPSs) from four Dipterocarpaceae plants can specifically synthesize borneol with minimal production of other monoterpene byproducts [15]. Although the catalytic process of monoterpenes has been reported in previous studies [21–26], there is still a need for further investigation into the catalytic mechanism responsible for the diversity of products catalyzed by MTSs, particularly in more species from the Dipterocarpaceae family.

In this study, we performed annotation of terpene synthases (TPS) in genomes of 13 Dipterocarpaceae species and narrowed down 38 MTSs that can generate a variety of monoterpene products by simulations and experiments. By analyzing the formation mechanism of acyclic, monocyclic and bicyclic products in monoterpene synthases, the minimum unit that affects the selectivity of monoterpene products, namely a single amino acid mutation, is predicted for the first time. We validated the relationship between residue mutations and product spectra, and further revealed that the mutations altered the synergistic effect between the shape of the active cavity and the stability of the carbocation intermediates are sufficient to relocate the reaction intermediates to select the reaction pathway. Monoterpenes are important plant natural products those are widely used as nutraceuticals and pharmaceuticals. The monoterpene synthases and mutants mined in this study will lay a foundation for the efficient production and application of monoterpenes in microorganisms.

2. Results

2.1. Discovery of monoterpene synthases in 13 species of dipterocarpaceae

The Dipterocarpaceae species are rich in monoterpenoids [15], and the biosynthesis of monoterpenes involves a complex rearrangement of GPP carbocations catalyzed by MTSs, resulting in a diverse collection of products (Fig. 1a). In order to identify more MTSs for synthesizing different monoterpenes, we performed annotation of terpene synthases (TPS) in genomes of thirteen Dipterocarpaceae species (*Dipterocarpus alatus*, *D. gracilis, D. zeylanicus, D.intricatus, Hopea mollissima, H. odorata*, *Parashorea chinensis, Shorea leprosula, S. roxburghii, S. henryana, Vatica xishuangbannaensis, V. odorata* and *V. rassak*) within the Dipterocarpaceae family. Initially we obtained 373 TPS genes based on local BLASTP and Hidden Markov Models analysis (Supplementary Fig. 1). These TPS genes can be clustered into five subfamilies, namely TPS-a, TPS-b, TPS-c, TPS-e/f and TPS-g, which is consistent with the annotation in *Arabidopsis thaliana* (Supplementary Fig. 1) [27]. To further refine these sequences, we focused on sequences containing the functional motif (RR(X)8W), which is crucial for the isomerization of GPP to linalyl cation and is highly conserved among all monoterpene synthases [28,29]. After filtering sequences that do not have the conserved functional motif, we identified 14 to 48 MTS candidates in the thirteen Dipterocarpaceae species, respectively (Supplementary Table 1). For enzymatic functional characterization, we randomly selected 60 candidate genes from these candidates (Supplementary Fig. 1). Notably, none of the sequences in the TPS-c and TPS-e/f subfamilies contained a motif (R(R)X8W), which is consistent with previous findings (Supplementary Fig. 1) [30,31]. Except 22 genes that did not produce targeted products, a total of 38 MTSs were identified to be able to yield monoterpenes, including 5 previously identified (+)-bornyl diphosphate synthases [15] (Fig. 1b and Supplementary Table 2). In conjunction with the outcomes from phylogenetic tree and Multiple Expectation Maximums for Motif Elicitation (MEME) motif search of these 38 protein sequences (Supplementary Fig. 2), we found that MTSs have obvious sequence and functional motif homology. Those having close phylogenetic relationships show higher similarity (Supplementary Fig. 2).

Starting with the same GPP precursor, these MTSs yielded distinct product profiles. The proteins sharing closest sequence similarity tend to produce similar monoterpenes. However, high conserved sequence homology was not the sole determinant of product profiles. We found that some proteins with high sequence similarity have very different product profiles. For instance, vod.ctg000225_np551212.959 (VOMTS1) was capable of producing acyclic linalool (compound1), while vod.ctg000225_np551212.964 (VOMTS2) could form four common

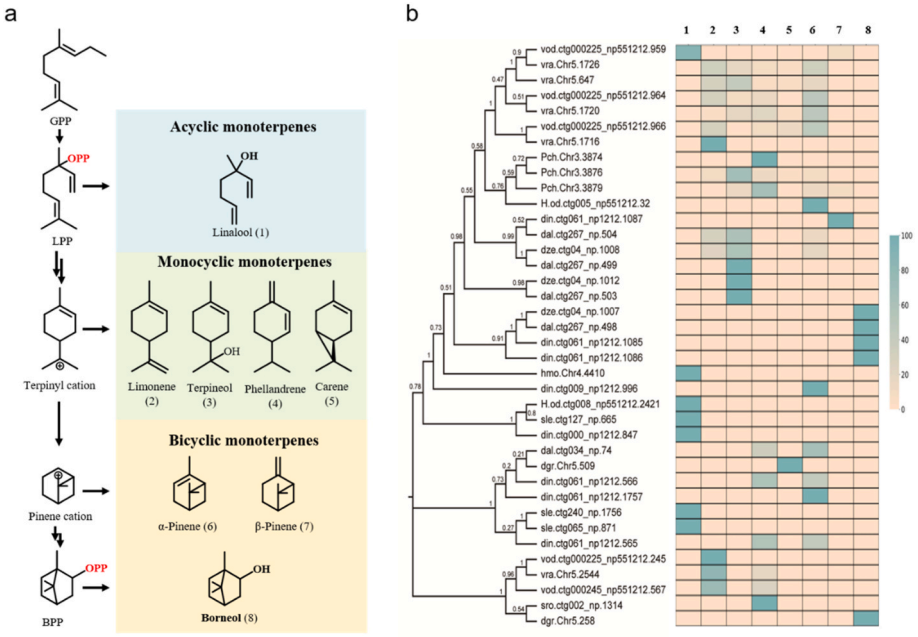

Fig. 1. | Monoterpene biosynthetic pathway and functional identification of MTS genes. a) The formation process of different monoterpenes. b) Phylogenetic tree of MTSs and its product distribution. Identification of product using GPP as a substrate and analysis by GC-MS. Heat map of the proportion of different monoterpenes in the total product of GPP catalyzed by MTSs. n = 3. Orange to green color indicates lowest to highest percentage of the product. 1, Linalool; 2, Limonene; 3, Terpineol; 4, Phellandrene; 5, Carene; 6, α-Pinene; 7, β-Pinene; 8, Borneol.

monoterpenes (α-pinene (compound 6), limonene (compound 2), phellandrene (compound 4), terpineol (compound 3)). Moreover, specific synthesis of linalool (compound 1) and borneol (compound 8) without byproducts was observed, whereas most cyclic monoterpenes (compound 2–6) tend to be produced simultaneously. These findings suggest that functional similarity is not entirely dependent on homology of sequences, indicating the existence of other factors influencing function of enzymes. However, the concrete mechanism driving these differences remains to be fully elucidated.

2.2. Broad product spectrum of monoterpene synthases

MTSs possess a conserved C-terminal α-helical fold and rely on a cluster of three Mg^{2+} ions to facilitate the ionization of the substrate and release of pyrophosphate moiety (PPi), thereby initiating a cyclization cascade involving multiple carbocation intermediates. This enzymatic process is aided by a closed active site conformation that effectively guards against premature quenching [32]. Despite the shared precursor geranyl diphosphate (GPP), different MTSs exhibit distinct product profiles. This divergence in products may be attributed to the rearrangement of carbocation facilitated by specific residues within the active site cavity, given the conserved structure of the enzyme (Fig. 2a). To investigate the catalytic mechanism of MTS specific synthesis of monoterpenoids, we focused on 4 MTSs involved in the synthesis of acyclic monoterpenes (VOMTS1), monocyclic monoterpenes (PCMTS1, pch.Chr3.3879), bicyclic monoterpenes (DZMTS3, dze.ctg04_np.1007), and various complex products (VOMTS2) (Fig. 2a). By constructing the ternary structure of these four MTSs using Alphafold2, we found that these MTSs are similar to the reported crystal structure (1N20) [33], with the three α-helices contributing to the formation of the active cavity (Supplementary Fig. 3).

To identify the residues that exert significant influence on the substrate profile, we excluded residues situated at the back of the helix and those engaged in interactions with the phosphate group of GPP. This enabled us to narrow down our investigation to the 16 remaining amino acids within a 6 Å distance from the active cavity, which were identified as critical residues governing the product profile [34]. Among these residues, we found 9 residues exhibited similarities or possess shared characteristics across the analyzed set of 4 MTS enzymes, highlighting their conserved roles (Supplementary Fig. 4). Therefore, only 7 residues showed variability across the studied MTSs, indicating their potential contribution to the diversification of product profiles (Fig. 2b). In Fig. 2c, the distribution of these 7 residues within the pocket of DZMTS3 is illustrated. It is worth noting that T563 was not in direct contact with the substrate due to the presence of W319, which acts as a barrier between them. Meanwhile, the other residues mentioned were in direct contact with the substrate.

2.3. Cyclization mechanism of monoterpene synthase

To determine the key factors affecting the first cyclization step in monoterpene synthesis, we focused on residues within VOMTS1 and other MTSs. In VOMTS1, residue 320 is alanine (A), while the corresponding residues in other enzymes are asparagine (N). Through molecular dynamics simulations, we found that the A320N mutation resulted in the compression of linear pyrophosphate (LPP) by occupying space, resulting in smaller pockets that promote LPP folding (Fig. 3a, Supplementary Fig. 5). Additionally, the G323I mutation introduced larger side chains, increasing the distance between neighboring helices and expanding the pocket space, allowing the substrate to undergo conformational change (Fig. 3a, Supplementary Fig. 5). Functional verification of these mutants supported the findings from molecular dynamics simulations (Fig. 3b). Interestingly, the A320N-G323I double mutation further enhanced cyclization, completely eliminating the production of linear linalool (Fig. 3b). This suggests that the introduction of aspartic acid in the A320N mutation may facilitate the cyclization of linear LPP, but whether the active cavity shape is suitable for LPP folding cyclization is more critical.

Next, we explored the factors contributing to the formation of borneol during secondary cyclization. Based on BLAST analysis, we

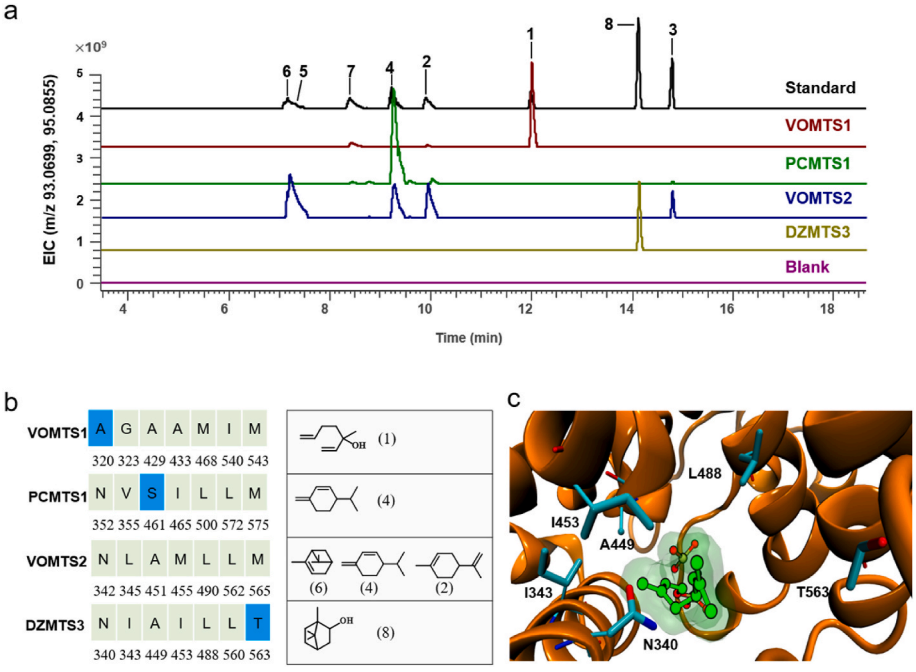

Fig. 2. | The distribution of products and key amino acids in MTSs. a) GC-MS analysis of four MTS genes (*VOMTS1*, *PCMTS1*, *VOMTS2*, and *DZMTS3*). The black line represents mixed standards of different monoterpenes at equal concentrations. The Blank represents control without MTSs. Peak 1, Linalool; Peak 2, Limonene; Peak 3, Terpineol; Peak 4, Phellandrene; Peak 5, Carene; Peak 6, α-Pinene; Peak 7, β-Pinene; Peak 8, Borneol. b) Different amino acids present in the active cavity and main product structures of four MTSs. Residues with distinct property differences are shown in blue. c) The location of key residues in the DZMTS3 active cavity. The DZMTS3 protein skeleton is shown in brown. Key residues are shown in green. The substrate GPP is shown in blue.

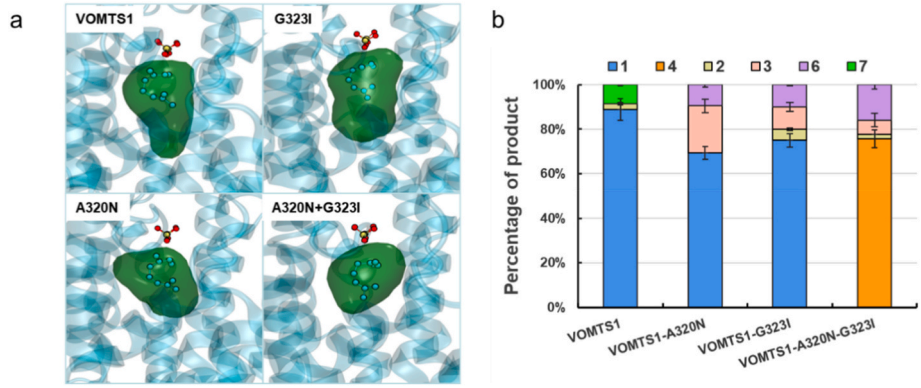

Fig. 3. | Functional characterization of VOMTS1 and its mutants. The active cavity volumes and product distribution in different VOMTS1 mutants. a) The active cavity volumes of VOMTS1 and mutants. Green represents the volumes of the active cavity. The red dots represent phosphoric acid groups. The blue dots in the active cavity represent intermediate transition states. b) Product distribution in VOMTS1 and mutants. 1, Linalool; 2, Limonene; 3, Terpineol; 4, Phellandrene; 6, α-Pinene; 7, β-Pinene. Error bars reflect mean ± s.d through three replicates.

found the critical residue T563 in DZMTS3 (Fig. 2b), while other MTSs had methionine (M) at the corresponding site. Molecular dynamics simulation revealed that the T563M mutation resulted in enlarged reactive cavity structure (Fig. 4a). T563 directly affected the position of the indole ring of W319 in DZMTS3, resulting in a 1 Å increase in the distance between its backbone carbon-α (CA) and the oxygen backbone of I448. The enlargement of the reactive cavity structure was not conducive to secondary cyclization and resulted in an increased production of monocyclic monoterpenes. Experimental results further confirmed that the substitution of T563M in DZMTS3 led to the reduction of borneol and increased the production of the terpineol byproduct, while the corresponding M543T mutation in VOMTS2 improved its ability to synthesize borneol (Fig. 4c). Similar observations were found in PCMTS1, although simultaneous mutation of S461A was required. We hypothesize that A449 of DZMTS3 (corresponding to S461 in PCMTS1) functions by influencing I448, a highly conserved residue that affects cyclization in MTSs.

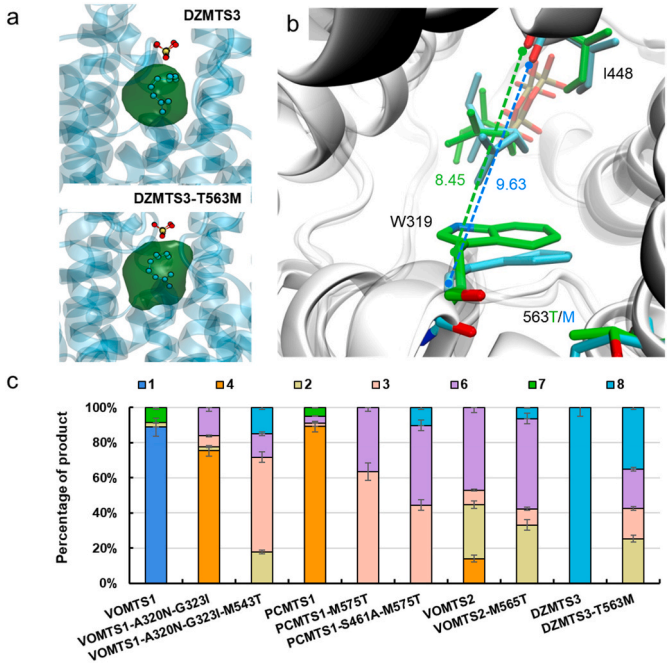

Fig. 4. | The key residues affecting the secondary cyclization of borneol. a) Effect of T563 mutation on the structure of DZMTS3 active cavity. Green represents the volume of the active cavity. The red dots represent phosphoric acid groups. The blue dots in the active cavity represent intermediate transition states. b) Effect of T563 mutation on W319 location in the active cavity. Green represents DZMTS3. Blue represents DZMTS3-T563M mutation. c) The product distribution of four MTS caused by mutation T563. 1, Linalool; 2, Limonene; 3, Terpineol; 4, Phellandrene; 6, α-Pinene; 7, β-Pinene; 8, Borneol. Error bars reflect mean ± s.d through three replicates.

In contrast to PCMTS1-M575T, the PCMTS1-M575T-S461A variant demonstrates a notable reduction in Terpineol yield along with a simultaneous increase in Borneol production. This observation suggests that the S461A mutation plays an effective role in prolonging the intermediates' stability. Additionally, molecular dynamics simulations have shed light on a significant hydrogen bond interaction forming between S461 and I460 in PCMTS1-M575T (Supplementary Fig. 6). This interaction is likely to weaken the stabilizing influence of I460. Consequently, we put forth a hypothesis that A449 in DZMTS3 (corresponding to S461 in PCMTS1) may exert its influence by affecting I448 (corresponding to I460 in PCMTS1), which is a highly conserved residue known to have a substantial impact on the cyclization process in MTSs [33]. Based on the findings of prior research [33–35], it has been established that the residue I448 exhibits a high degree of conservation across all MTSs. This conservation serves as a compelling rationale for refraining from introducing mutations in this particular residue.

2.4. Steady state of the monoterpene cyclization intermediate

Although the folding process of borneol in enzymes has been recapitulated based on co-crystal structures of substrate or product analogs [11,33], the mechanism by which the enzymatic environment promotes the steady state of the intermediate remains unclear. Here, we adopted QM/MM approach [36] to investigate the cyclization process by scanning the bond lengths between reactive atoms. After performing molecular dynamics simulations, we obtained a stable snapshot of LPP as a phosphorylated substrate serving as the starting point for remodeling of the cyclization pathway (Fig. 5a). To ensure an accurate representation the electrostatic environment in the pocket, the QM region included residues W319, N340, I448 and A449, in addition to the LPP, which have been shown to play a critical role for borneol cyclization [34,35].

Following a relaxed scan of the C3–O and C1–C7 bonds in LPP, followed by unconstrained complete relaxation to find the lowest energy state, the first cyclization step yielded the lowest-energy intermediate 1 (Terpinyl cation) (Fig. 5b). To identify the residues that stabilized this conformation, an IGMH [37] analysis was performed on the cluster containing residues and phosphate groups surrounding the intermediates. The electron density at different critical points were 0.0253, 0.0143 and 0.0102, respectively. The results revealed that the phosphate group (point 1, 0.0253) possessed the strongest interaction with the intermediates, followed by carbonyl oxygen of the I448 main

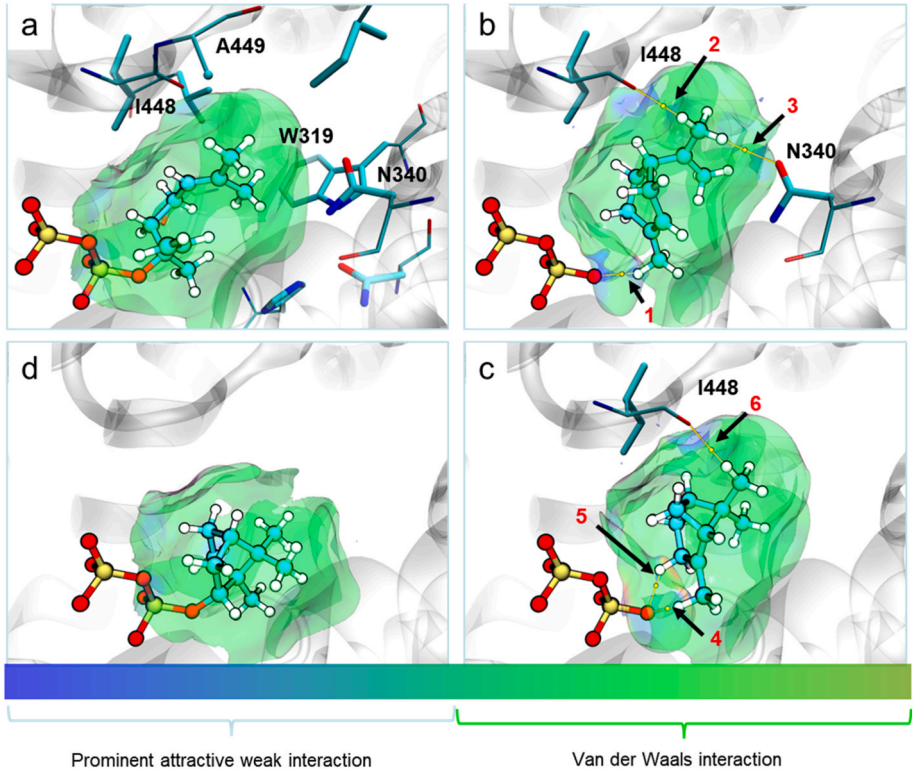

Fig. 5. | The conformational cyclization process analysis of borneol in IGMH using QM/MM scanning. The green color represents regions influenced by van der Waals forces, whereas the blue color represents regions with strong electrostatic attractions. The balls in the figure represent the maximum electron density points. a) A stable snapshot of the phosphorylated substrate, LPP. b) Single ring intermediates involving the oxygen atoms of phosphoric acid (1), I448 (2), and N340 (3), which contribute to the formation of stable carbon-positive intermediates. c) Double ring intermediates that are stabilized by phosphoric acid (4, 5) and I448 (6). d) The final phosphorylated end product undergoes hydrolysis to produce borneol.

chain (point 2, 0.0143), while the N340 side chain showed the weakest interaction (point 3, 0.0102) (Fig. 5b). Further IRI [38] cluster analysis revealed significant interactions between the atoms of the intermediates themselves, particularly with marked attraction between the hydrogens of the C9 and C2 intermediates (point 7, 0.0259) (Supplementary Fig. 7). These findings indicate that the intermediates possess a self-stabilizing mechanism in addition to stabilization mediated by enzymatic environment.

For the second positive carbon attack, the C2–C8 bond is formed, instead of direct cyclization to borneol diphosphate by terpinyl cations, resulting in the generation of intermediates 2 (pinene cation). Based on IGMH analysis, the phosphate groups exhibited significant attraction to the hydrogens of C4 and C5 (point 4, 0.0453; point 5, 0.0376). This stable state prevents the deprotonation of C3 or C5, which would lead to the production of α-pinene or β-pinene. Additionally, the carbonyl oxygen within the L448 backbone plays an important role in stabilizing the pinene cation (point 6, 0.0119). As cyclization is a continuous process, both N340 and I448 contribute to the correct stabilization of the terpinyl and pinene cations, which are essential for borneol synthesis. The positional changes of N340, I448, and phosphate groups are pivotal for the stability of intermediates. Consequently, any mutations affecting the positions of N340, I448, and phosphate groups may indirectly lead to the instability of intermediates.

3. Discussion

MTS enzymes catalyze cascade reaction involving multiple carbocation intermediates, which play a crucial role in defining the reaction coordinate of catalysis within a non-polar active cavity and provide a theoretical basis for the ability to make multiple products [17]. Our research provides further evidence that the extent of MTS product diversity may be determined to a large degree by the three-dimensional profile of the active cavity defined by the overall structure of the enzyme, rather than by the precise residues of the direct reaction alone. Guided by the three-dimensional structure of the enzyme-active cavity, subsequent reactions of these carbocation intermediates lead to the formation of diverse carbon skeletons. This phenomenon has been observed in other terpene synthases [39–42]. Yoshikuni et al., for instance, demonstrated that the combination of multiple mutations in γ-humulene synthase can result in a narrowed product specificity whereas the enzyme would otherwise produce multiple products [40]. Molecular simulations indicated that these mutations altered the shape and volume of the active cavity, leading to the repositioning of reaction intermediates (Fig. 4). In addition, Greenhagen et al. achieved interconversion of TEAS (*Nicotiana tabacum* 5-epiaristolochene synthase) and HPS (*Hyoscyamus muticus* premnaspirodiene synthase) activities using a set of 7 amino acid residues based on a contact mapping strategy [43]. While previous reports have shed light on how clusters of amino acid residues contribute to reaction cascades of terpene synthases [40,43], the precise structural changes required to modify or control product specificity in MTSs remain unresolved.

Here, we observed that even a single amino acid substitution could yield sufficient changes to alter the spectrum of major reaction products across all four different MTSs, which themselves produce different types of monoterpenes (Fig. 4c). To gain deeper insights, we employed molecular dynamic simulations to analyze the different mutant structures within MTSs. The results show that the initial cyclization event requires sufficient space for folding the linear LPP, which is further confirmed by *in vitro* mutation experiments. In VOMTS1, the mutation A320N facilitates the conversion of enzyme activity by reshaping the reaction chamber and collaboratively stabilizing the terpinyl cation intermediate. This results in the production of diverse monocyclic monoterpenes

instead of acyclic monoterpenes. Meanwhile, in DZMTS3, T563M mediates product diversity by modifying the three-dimensional profile of the enzyme reaction cavity. In addition, our investigation revealed that the second cyclization event requires a shorter distance from the alternate axis of the active cavity compared to the initial cyclization in PCMTS1. In summary, these mutations mediate product diversity by inducing changes in the active site shape and volume that are sufficient to reposition reactive intermediates for alternative reaction pathways.

Finally, previous reports have indicated that functional remodeling of terpene synthases based on sequences is challenging due to the lack of a clear sequence-function correlation [40,44]. Establishing functional interaction models of specific amino acids, thereby indicating the role of key residues in the catalytic process, holds the promise to predict how alternative or novel catalytic activities can be rationally designed into specific MTSs. Overall, our findings shed novel light on the evolutionary interplay between MTS sequence diversity and complex chemical products. This mechanism holds the potential for rationalizing the design and engineering of different catalytic activities into existing MTS [45]. In a long run, the monoterpene synthases and mutants excavated in this study are expected to achieve direct synthesis of specific proportions of monoterpenes in microorganisms, and have a strong industrial application prospect.

4. Materials and methods

4.1. Mining of monoterpene synthases (MTS) in dipterocarpaceae

From our previous study [15], 13 Dipterocarpaceae genomes were used in this investigation. We obtained the 13 Dipterocarpoideae protein sequences from the China National GeneBank Database (CNGBdb, https://db.cngb.org/) under accession no. CNP0002104. To identify candidate Dipterocarpaceae MTS proteins, we performed search against two Pfam domains: PF01397 and PF03936 using Hidden Markov Model (HMM) analysis in HMMER v.3.3(http://www.hmmer.org/). Protein sequences with significant hits (e-value <10-3) were retrieved as potential Dipterocarpaceae MTS candidates. To verify these sequences, we conducted BLASTp analysis using queries from Arabidopsis thaliana with a threshold identity> 30 ($p < 10\text{-}5$).

To analyze the conserved motifs of candidate MTS genes in Dipterocarpaceae proteins, were used MEME software (http://meme-suite.org/) with default parameters. Multiple sequences alignment was performed using MTS proteins from *A. thaliana* and Dipterocarpaceae with ClustalX2.1 software (www.clustal.org/). A phylogenetic tree was constructed using the Maximum-Likelihood (ML) method with 100 bootstrap replicates in MEGA version 7.0 [46]. The resulting tree was visualized and annotated by Figtree version 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

4.2. Functional verification of MTSs

Following the previously described methods [15], heterologous expression and functional determination of MTSs were performed. Candidate MTS genes were constructed into the pET-28a vector between *Nde* I and *Xho* I restriction sites, and all proteins were heterologously expressed in *E.coli* BL21 (DE3) strain. The target proteins were purified by nickel affinity chromatography and their concentrations were determined using BCA Protein Assay Reagent Kit (Pierce, USA). The enzyme reaction system comprised 50 mM Tris-HCl pH 7.2, 10 mM $MgCl_2$, 5 mM DTT, 1 mM PMSF, 3 mM GPP. The catalytic reaction was initiated by adding 1 mg/mL purified MTS protein, and the reaction mixtures were incubated at 30 °C for 15 min. The reactions were terminated at 80 °C for 5 min and subsequently quenched on ice. To promote dephosphorylation of the product, 3 μL of calf intestinal alkaline phosphatase (CIAP) (TaKaRa, Japan) was added and the mixture was incubated for another 1 h at 37 °C. The monoterpene products were extracted by ethane and analyzed by GC-MS.

The GC-MS analysis was carried out on an Agilent 7890A gas chromatograph system coupled with quadrupole time-of-flight (Q-TOF) mass spectrometer and an inert electron ionization (EI) ion source (Agilent, USA). An ADB-5MS capillary column coated with 5 % diphenyl cross-linked with 95 % dimethylpolysiloxane (30 m × 250 μm inner diameter, 0.25 μm film thickness; J&W Scientific, USA) was used. The oven temperature program consisted of 50 °C (2 min), followed by an increase to 180 °C (5 min) at a rate of 5 °C min^{-1}, and then a ramp of 10 °C min^{-1} to 230 °C. The total run time was 38 min. The injection, transfer line, and ion source temperatures were set at 250, 290, and 230 °C, respectively. Agilent MassHunter 10.0 software with NIST2020 library was used for raw peak extraction, peak alignment, deconvolution analysis, peak identification and integration of the peak areas. Before each experiment, perflubuterine was used as the internal standard, and the response intensity of the instrument was corrected to the same level. In addition, the standard mixture of different concentrations was used as external standard for absolute quantification.

5. Computational method

The system was initially set up using Alphafold [47] for structure prediction. The protein was placed in a water box (OPC [48]) with a boundary that was kept at least 10 Å away from the protein, to ensure neutralize total charges, Na^+ ions were added to the system. The ff19SB [49] force field was used to describe the protein in all simulations. The system was then subjected to molecular dynamics simulations. Firstly, all atoms except the protein backbone underwent minimization. Later, the system was heated from 0 to 300 K in 50 ps under the NVT ensemble with a time step of 2 fs. Following the heating, the system density was equilibrated at 1.0 atm and a temperature of 300 K under the NPT ensemble. The protein backbone was fixed during heating and density equilibrium steps. Subsequently the system underwent further equilibration with a gradient restraint applied to the protein backbone for 100 ns, followed by 100 ns of MD simulations without any restraints. Finally, QM/MM calculations were performed using the QM/MM method of Orca [50], with the active region including all residues within 10 Å from the QM region. The hybrid UB3LYP-D4 [51] and all-electron def2-SVP basis set were used to treat the QM region, and geometry optimizations were performed using the L-BFGS algorithm and the semiempirical quantum mechanical methods GFNn-xTB [52]. The QM/MM boundary was treated using hydrogen link atoms and the electronic embedding scheme.

Funding

This work was supported by the National Key R&D Program of China (2020YFA0908000), the National Natural Science Foundation of China (31901015), Science and Technology Partnership Program, Ministry of Science and Technology of China (KY202001017).

Availability of data and materials

All data generated or analyzed during this study are included in this research article.

Ethics approval and consent to participate

Not applicable.

CRediT authorship contribution statement

Xiaoyun Lu: Conceptualization, Methodology, Validation, Data curation, Formal analysis, Writing – original draft. **Jie Bai:** Data curation, Formal analysis, Writing – review & editing. **Zunzhe Tian:** Data curation, Formal analysis, Writing – review & editing. **Congyu Li:** Validation, Data curation, Formal analysis. **Nida Ahmed:** Writing –

review & editing. **Xiaonan Liu:** Investigation. **Jian Cheng:** Investigation. **Lina Lu:** Validation. **Jing Cai:** Writing – review & editing. **Huifeng Jiang:** Conceptualization, Writing – review & editing. **Wen Wang:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgments

We thank Yonghong Yao from Tianjin Institute of Industrial Biotechnology for his help in GC-MS.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2023.11.009.

References

[1] Vasconcelos RM, Leite FC, Leite JA, Rodrigues Mascarenhas S, Rodrigues LC, Piuvezam MR. Synthesis, acute toxicity and anti-inflammatory effect of bornyl salicylate, a salicylic acid derivative. Immunopharmacol Immunotoxicol 2012;34(6):1028–38.
[2] Sokolova AS, Yarovaya OI, Shtro AA, Borisova MS, Morozova EA, Tolstikova TG, Zarubaev VV, Salakhutdinov NF. Synthesis and biological activity of heterocyclic borneol derivatives. Chem Heterocycl Compd (N Y, NY, U S) 2017;53(371–377).
[3] Bhatia SP, Letizia CS, Api AM. Fragrance material review on borneol. Food Chem Toxicol : Int J Publ British Ind Biol Res Associat 2008;46(Suppl 11):S77–80.
[4] Ying Zha LG, Lin Bai, Shu Xuechun, Pang Yuxin, Chen Songbi. Research progress on borneolum. Chin J Tropic Agric 2019;39(259):100–7.
[5] Wang X, Wu J, Chen J, Xiao L, Zhang Y, Wang F, Li X. Efficient biosynthesis of R-(-)-Linalool through adjusting the expression strategy and increasing gpp supply in Escherichia coli. J Agric Food Chem 2020;68(31):8381–90.
[6] Liu W, Xu X, Zhang R, Cheng T, Cao Y, Li X, Guo J, Liu H, Xian M. Engineering Escherichia coli for high-yield geraniol production with biotransformation of geranyl acetate to geraniol under fed-batch culture. Biotechnol Biofuels 2016;9:58.
[7] Rolf J, Julsing MK, Rosenthal K, Lütz S. A gram-scale limonene production process with engineered Escherichia coli. Molecules 2020;25(8).
[8] Zhang H, Liu Q, Cao Y, Feng X, Zheng Y, Zou H, Liu H, Yang J, Xian M. Microbial production of sabinene–a new terpene-based precursor of advanced biofuel. Microb Cell Factories 2014;13:20.
[9] Ma R, Su P, Ma Q, Guo J, Chen S, Jin B, Zhang H, Tang J, Zhou T, Xiao C, Cui G, Huang L. Identification of (-)-Bornyl diphosphate synthase from blumea balsamifera and its application for (-)-Borneol biosynthesis in Saccharomyces cerevisiae. Synthetic Syst Biotechnol 2022;7(1):490–7.
[10] Willrodt C, David C, Cornelissen S, Bühler B, Julsing MK, Schmid A. Engineering the productivity of recombinant Escherichia coli for limonene formation from glycerol in minimal media. Biotechnol J 2014;9(8):1000–12.
[11] O'Brien TE, Bertolani SJ, Zhang Y, Siegel JB, Tantillo DJ. Predicting productive binding modes for substrates and carbocation intermediates in terpene synthases-bornyl diphosphate synthase as a representative case. ACS Catal 2018;8(4):3322–30.
[12] Davis EM, Croteau R. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. In: Leeper FJ, Vederas JC, editors. Biosynthesis: aromatic polyketides, isoprenoids, alkaloids. Springer Berlin Heidelberg; 2000. p. 53–95.
[13] Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol 2006;9(3):297–304.
[14] Christianson DW. Structural biology and chemistry of the terpenoid cyclases. Chem Rev 2006;106(8):3412–42.
[15] Tian Z, Zeng P, Lu X, Zhou T, Han Y, Peng Y, Xiao Y, Zhou B, Liu X, Zhang Y, Yu Y, Li Q, Zong H, Zhang F, Jiang H, He J, Cai J. Thirteen Dipterocarpoideae genomes provide insights into their evolution and borneol biosynthesis. Plant Commun 2022;3(6):100464.
[16] Christianson DW. Structural and chemical biology of terpenoid cyclases. Chem Rev 2017;117(17):11570–648.
[17] Zhang L, Xiao WH, Wang Y, Yao MD, Jiang GZ, Zeng BX, Zhang RS, Yuan YJ. Chassis and key enzymes engineering for monoterpenes production. Biotechnol Adv 2017;35(8):1022–31.
[18] Kim H, Srividya N, Lange I, Huchala EW, Ginovska B, Lange BM, Raugei S. Determinants of selectivity for the Formation of monocyclic and bicyclic products in monoterpene synthases. ACS Catal 2022;12(12):7453–69.
[19] Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 2009;70(15–16):1621–37.
[20] Wise ML, Savage TJ, Katahira E, Croteau R. Monoterpene synthases from common sage (salvia officinalis). Cdna isolation, characterization, and functional expression of (+)-Sabinene synthase, 1,8-cineole synthase, and (+)-Bornyl diphosphate synthase. J Biol Chem 1998;273(24):14891–9.
[21] Leferink NGH, Escorcia AM, Ouwersloot BR, Johanissen LO, Hay S, van der Kamp MW, Scrutton NS. Molecular determinants of carbocation cyclisation in bacterial monoterpene synthases. Chembiochem : Eur J Chem Biol 2022;23(5):e202100688.
[22] Major DT, Freud Y, Weitman M. Catalytic control in terpenoid cyclases: multiscale modeling of thermodynamic, kinetic, and dynamic effects. Curr Opin Chem Biol 2014;21:25–33.
[23] Baer P, Rabe P, Fischer K, Citron CA, Klapschinski TA, Groll M, Dickschat JS. Induced-fit mechanism in class I terpene cyclases. Angew Chem 2014;53(29):7652–6.
[24] Hong YJ, Tantillo DJ. Quantum chemical dissection of the classic terpinyl/pinyl/bornyl/camphyl cation conundrum-the role of pyrophosphate in manipulating pathways to monoterpenes. Org Biomol Chem 2010;8(20):4589–600.
[25] Peters RJ, Croteau RB. Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthases. Arch Biochem Biophys 2003;417(2):203–11.
[26] Schwab W, Williams DC, Davis EM, Croteau R. Mechanism of monoterpene cyclization: stereochemical aspects of the transformation of noncyclizable substrate analogs by recombinant (-)-Limonene synthase, (+)-Bornyl diphosphate synthase, and (-)-Pinene synthase. Arch Biochem Biophys 2001;392(1):123–36.
[27] Tholl D, Lee S. Terpene specialized metabolism in Arabidopsis thaliana. Arabidopsis Book 2011;9:e0143.
[28] Wang S, Liang H, Wang H, Li L, Xu Y, Liu Y, Liu M, Wei J, Ma T, Le C, Yang J, He C, Liu J, Zhao J, Zhao Y, Lisby M, Sahu SK, Liu H. The chromosome-scale genomes of dipterocarpus turbinatus and Hopea hainanensis (Dipterocarpaceae) provide insights into fragrant oleoresin biosynthesis and hardwood formation. Plant Biotechnol J 2022;20(3):538–53.
[29] Hurd MC, Kwon M, Ro DK. Functional identification of a lippia dulcis bornyl diphosphate synthase that contains a duplicated, inhibitory arginine-rich motif. Biochem Biophys Res Commun 2017;490(3):963–8.
[30] Jiang SY, Jin J, Sarojam R, Ramachandran S. A comprehensive survey on the terpene synthase gene family provides new insight into its evolutionary patterns. Genome Biol Evolut 2019;11(8):2078–98.
[31] Yu Z, Zhao C, Zhang G, Teixeira da Silva JA, Duan J. Genome-wide identification and expression profile of tps gene family in dendrobium officinale and the role of Dotps10 in linalool biosynthesis. Int J Mol Sci 2020;21(15).
[32] Major DT. Electrostatic control of chemistry in terpene cyclases. ACS Catal 2017;7(8):5461–5.
[33] Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. Proc Natl Acad Sci USA 2002;99(24):15375–80.
[34] Leferink NGH, Ranaghan KE, Karuppiah V, Currin A, van der Kamp MW, Mulholland AJ, Scrutton NS. Experiment and simulation reveal how mutations in functional plasticity regions guide plant monoterpene synthase product outcome. ACS Catal 2019;8(5):3780–91.
[35] Yao J, Chen F, Guo H. Qm/mm free energy simulations of the reaction catalysed by (4s)-Limonene synthase involving linalyl diphosphate (lpp) substrate. Mol Simulat 2018;44:1–10.
[36] Senn HM, Thiel W. Qm/mm methods for biomolecular systems. Angew Chem 2009;48(7):1198–229.
[37] Lu T, Chen Q. Independent gradient model based on hirshfeld partition: a new method for visual study of interactions in chemical systems. J Comput Chem 2022;43(8):539–55.
[38] Lu T, Chen Q. Interaction region indicator: a simple real space function clearly revealing both chemical bonds and weak interactions. Chem Mater 2021;1(5):231–9.
[39] Köllner TG, Schnee C, Gershenzon J, Degenhardt J. The variability of sesquiterpenes emitted from two zea mays cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. Plant Cell 2004;16(5):1115–31.
[40] Yoshikuni Y, Ferrin TE, Keasling JD. Designed divergent evolution of enzyme function. Nature 2006;440(7087):1078–82.
[41] Lodeiro S, Segura MJ, Stahl M, Schulz-Gasch T, Matsuda SP. Oxidosqualene cyclase second-sphere residues profoundly influence the product profile. Chembiochem : Eur J Chem Biol 2004;5(11):1581–5.
[42] Hyatt DC, Croteau R. Mutational analysis of a monoterpene synthase reaction: altered catalysis through directed mutagenesis of (-)-Pinene synthase from abies grandis. Arch Biochem Biophys 2005;439(2):222–33.
[43] Greenhagen BT, O'Maille PE, Noel JP, Chappell J. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. Proc Natl Acad Sci USA 2006;103(26):9826–31.
[44] Leferink NGH, Scrutton NS. Predictive engineering of class I terpene synthases using experimental and computational approaches. Chembiochem : Eur J Chem Biol 2022;23(5):e202100484.
[45] Dwyer MA, Looger LL, Hellinga HW. Computational design of a biologically active enzyme. Science (New York, NY) 2004;304(5679):1967–71.
[46] Kumar S, Stecher G, Tamura K. Mega7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33(7):1870–4.
[47] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021;596(7873):583–9.
[48] Izadi S, Anandakrishnan R, Onufriev AV. Building water models: a different approach. J Phys Chem Lett 2014;5(21):3863–71.

[49] Tian C, Kasavajhala K, Belfon KA, Raguette L, Huang H, Migues AN, et al. ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. J Chem Theory Comput 2019;16(1):528–52.
[50] Neese F. The ORCA program system. Wiley Interdiscip Rev Comput Mol Sci 2012;2 (1):73–8.
[51] Caldeweyher E, Ehlert S, Hansen A, Neugebauer H, Spicher S, Bannwarth C, et al. A generally applicable atomic-charge dependent London dispersion correction. J Chem Phys 2019;150(15):154122.
[52] Bannwarth C, Caldeweyher E, Ehlert S, Hansen A, Pracht P, Seibert J, et al. Extended tight-binding quantum chemistry methods. Wiley Interdiscip Rev Comput Mol Sci 2021;11(2):e1493.