

# Mutational Analysis of Linalool Dehydratase-Isomerase (LinD) Suggests Alcohol and Alkene Transformations are Catalysed Using Non-Covalent Mechanisms

Received 00th January 20xx,  
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Anibal Cuetos,<sup>a</sup> Javier Iglesias,<sup>b</sup> Hamid-Reza Danesh-Azari,<sup>a</sup> Erna Zukic,<sup>a</sup> Adam Dowle,<sup>c</sup> Sílvia Osuna<sup>b,d</sup> and Gideon Grogan<sup>\*a</sup>

The interconversion of non-activated alkenes and alcohols, catalysed by (de)hydratases, has great potential in biotechnology for the generation of fine and bulk chemicals. LinD is a cofactor-independent enzyme that catalyses the reversible (de)hydration of the tertiary alcohol (*S*)-linalool to the triene  $\beta$ -myrcene, and also its isomerization to the primary alcohol geraniol. Structure-informed mutagenesis of LinD, followed by activity studies, confirmed essential roles for residues C171, C180 and H129 in water activation for the hydration of  $\beta$ -myrcene to linalool. However, no evidence of covalent thioether intermediate was found using either X-ray crystallography, mass spectrometry, or QM/MM nudged elastic band simulations. Labelling and NMR experiments confirmed a role for residue D39 in (de)protonation of the linalool carbon C10 in the isomerization of linalool to geraniol and also the intermediacy of  $\beta$ -myrcene in this isomerization reaction. X-ray, molecular dynamics and activity studies also suggested a significant role in catalysis for a mobile methionine residue M125, which exists in substantially altered orientations in different mutant structures.

## Introduction

The addition of water to carbon-carbon double bonds to make alcohols, and the reverse reaction, the dehydration of alcohols to give alkenes, are important reactions in organic chemistry owing to the abundance of cheap petrochemical olefins and the requirement for alkenes and alcohols for industrial and synthetic chemical applications.<sup>1</sup> However, abiotic water addition or dehydration processes involve, in many cases, heavy-metal catalysts such as tungsten oxide, or multistep reactions involving toxic reagents, such as hydroboration-oxidation and oxymercuration-reduction, each of which are required owing to the poor reactivity of water as either a nucleophile or an electrophile. In addition, few of these processes permit the asymmetric hydration of alkenes, a process that would lead to desirable chiral alcohols, although bio-inspired solutions have been implemented in the design of asymmetric hydration catalysts incorporating DNA complexes<sup>2</sup> or chiral bio-polymers.<sup>3,4</sup>

In nature, the hydration of alkenes is catalysed by hydratases (E.C. 4.2.1.X), enzymes that enable the reversible (de)hydration of

alcohols to alkenes with high regio- and enantioselectivity.<sup>5-8</sup> Depending on their substrates, hydratases may be classified into two broad classes: those such as fumarase<sup>9</sup> and enoyl-CoA hydratase,<sup>10</sup> which catalyse the addition of water to an electron-deficient alkene, *via* nucleophilic Michael addition to  $\alpha,\beta$ -unsaturated carbonyl compounds, and those that catalyse an electrophilic water addition to electron-rich, isolated C=C double bonds. The asymmetric addition of water to these non-activated alkenes presents a special challenge in terms of catalysis, but the ability to hydrate these substrates asymmetrically, to give optically enriched alcohols, would have remarkable significance given the lack of options within conventional chemical catalysis. Enzymes that catalyse the hydration of such alkenes include cofactor dependent kievitone hydratase<sup>11-13</sup> and carotenoid hydratases<sup>14,15</sup> and also a promiscuous activity of phenolic acid decarboxylases, which has been applied by the group of Faber to the asymmetric hydration of hydroxystyrenes.<sup>16,17</sup> Recently, Hauer and co-workers have presented engineered mutants of a flavin-dependent fatty acid hydratase from *Elizabethkingia meningoseptica*,<sup>18</sup> which, in the presence of a decoy molecule, catalyse the asymmetric hydration of simple alkenes, such as dec-1-ene to (*S*)-2-decanol with >99% e.e.<sup>19</sup>

The cofactor-independent linalool dehydratase isomerase (LinD) was identified from the crude extract of the  $\beta$ -proteobacterium *Castellaniella defragrans* 65Phen<sup>20</sup> and catalyses the enantioselective (de)hydration and isomerization of the monoterpenoid (*S*)-linalool **2** to  $\beta$ -myrcene **1** and geraniol **3** respectively (**Figure 1A**). The wider interest in LinD has been prompted by its reported activity in the dehydration of simple

<sup>a</sup> York Structural Biology Laboratory, Department of Chemistry, University of York, YO10 5DD York, U.K.

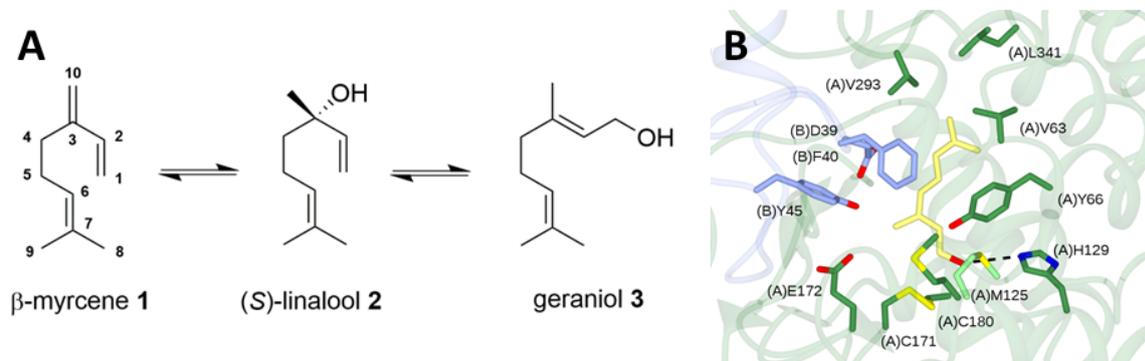
<sup>b</sup> CompBioLab group, Institut de Química Computacional i Catàlisi, Departament de Química, Carrer Maria Aurèlia Capmany 69, 17003 Girona, Spain.

<sup>c</sup> Bioscience Technology Facility, Department of Biology, University of York, YO10 5DD, York, UK.

<sup>d</sup> ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain

† Footnotes relating to the title and/or authors should appear here.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x



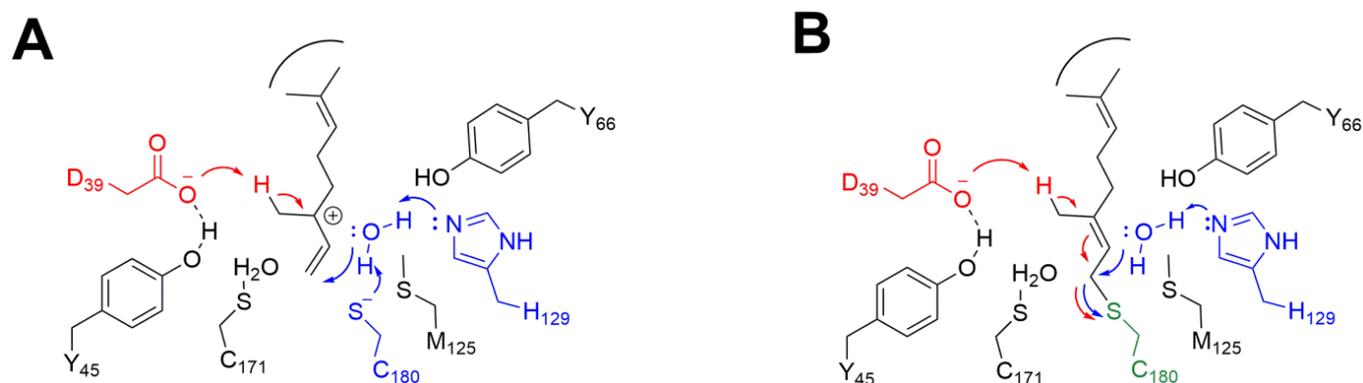
**Figure 1.** A: (De)hydration and Isomerisation reactions catalysed by LinD; B: Active site of LinD at the monomer-monomer interface (PDB 5G1U) with backbone and carbon atoms of subunits A and B in dark green and blue, respectively. Geraniol is shown in cylinder format with carbon atoms in yellow. The two positions of the side chain of M125 are shown with carbon atoms in dark green ('open') and light green ('closed' – from 5G1W).

alcohols such as crotyl alcohol and methylbutanediol to valuable bulk chemicals such as isoprene and butadiene respectively.<sup>21–23</sup>

In an effort to understand the mechanism of this valuable hydratase, and to rationally engineer its activity for further applications in biotechnology, the LinD crystal structure was independently solved by three groups: The Harder group has reported the structure of LinD in complex with both  $\beta$ -myrcene and geraniol (PDB code 5HSS).<sup>24</sup> Apo-structures have also been deposited by this group (5HLR),<sup>24</sup> and also by LaMattina and co-workers (5I3T). Although initially the native oligomerization state of LinD was thought to be homotetrameric,<sup>20</sup> the enzyme in all cases was shown to adopt a toroidal homopentameric structure, with each monomer displaying an  $(\alpha, \alpha)_6$  barrel fold. The active sites were revealed at the interface of two neighboring monomers. In previous work,<sup>25</sup> we have also determined the structure of LinD, both in an apo-form (5G1W) and also in complex with the isomerisation product geraniol (5G1U), although in a different conformation to that presented by Harder.<sup>24</sup> The active site of LinD (**Figure 1B**) revealed a hydrophobic binding pocket including V63, V293 and L341, involved in recognition of the geminal dimethyl group of the terpene substrate(s). Residues within reactive distance of C3 and C10 atoms that participate in the double bond to be hydrated in linalool include (B)D39 and (B)Y45, and on the other face of the ligand, (A)Y66. At the bottom of the active site, and within reactive distance of the ligand, are two cysteine residues, (A)C180 and (A)C171, a methionine (A)M125 a histidine (A)H129 and a glutamate (A)E172. M125 was shown to be substantially mobile: In structures 5HSS, 5HLR, 5I3T and 5G1U, it exists in an 'open' conformation as shown in **Figure 1B** away from the ligand, with

either a water molecule or the primary hydroxyl of geraniol (5G1U) situated between the side chains of H129 and C180; however, in the apo structure 5G1W, the side chain has a 'closed' conformation with the sulfur atom replacing the water molecule between those side chains.

The assignment of catalytic roles to any of the amino acids in the active site of LinD is complex, as any proposed mechanism must explain the function of these residues in four possible separate reactions involving three possible substrates, and also account for the (S)-enantioselectivity of the linalool (de)hydration reaction. Studies using cell-extracts by both ourselves<sup>25</sup> and the group of Harder<sup>24</sup> suggested that mutation of either cysteine residue to alanine led to ablation of any LinD activity, with reduced activity also observed for extracts containing alanine mutants of Y45, M125 and H129. The geraniol complex, in combination with a preliminary study of mutants using cell extracts, led us to propose two possible mechanistic foundations for all transformations: the first where a central carbocation intermediate resulting from linalool dehydration can be quenched through either proton abstraction from C10 (dehydration to  $\beta$ -myrcene) or hydration at C1 (isomerization to geraniol, **Figure 2A**); the second where a cysteine-ligated thioterpene intermediate, derived from attack of Cys at the terminal C1 atom, could again be either deprotonated to  $\beta$ -myrcene or hydrolyzed to give geraniol (**Figure 2B**). (A detailed diagrammatic summary of the mechanistic proposals for LinD action can be found in the Supporting Information (**Figure S1**)).



**Figure 2.** Mechanistic hypotheses for LinD action: **A:** A carbocation intermediate, resulting from linalool dehydration is quenched by deprotonation (red) by D39 at C10 to give  $\beta$ -myrcene, or hydration at C1 (blue), catalysed either by H129 or C180, to give geraniol; **B:** A covalent thioether intermediate formed by attack of C180 (green) at C1 of linalool with dehydration is followed by either deprotonation (red) by D39 at C10 to give  $\beta$ -myrcene, or hydrolysis at C1 (blue), catalysed by H129, to give geraniol.

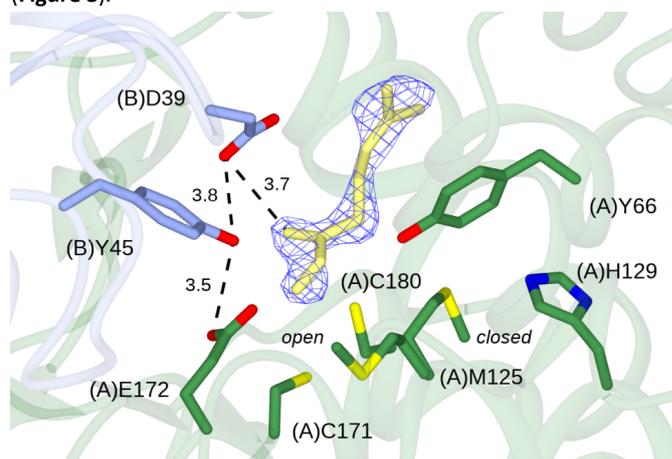
In order to investigate the nature of the mechanism of LinD, in this report we have expressed and purified the wild-type enzyme and mutants of the residues within the first catalytic shell. We also present an improved structural complex of wt-LinD with  $\beta$ -myrcene, measurements of mutant activity using GC and analysis of ligand complexes using MS and X-ray crystallography. The results, in combination with Quantum Mechanics/Molecular-Mechanics (QM/MM) reaction path calculations, strongly support a hypothesis for non-covalent mechanisms of LinD action, including an elimination pathway from geraniol to linalool *via* a  $\beta$ -myrcene intermediate. In addition, these studies suggest a role for the mobile M125 side-chain in the catalysis of linalool- $\beta$ -myrcene (de)hydration.

## Results and discussion

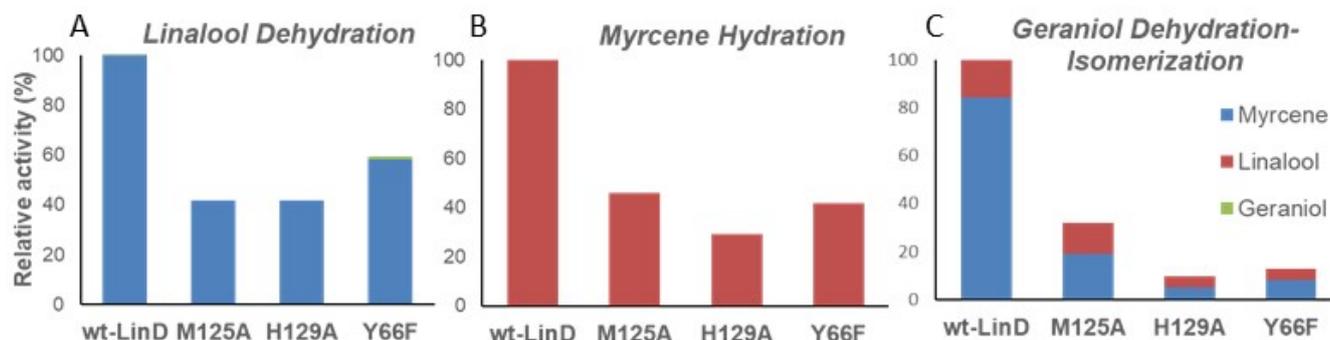
### Structure of wt-LinD in complex with $\beta$ -myrcene

As a first step in the detailed investigation of LinD activity, we sought to obtain ligand complexes of the enzyme with either  $\beta$ -myrcene or linalool using X-ray crystallography. Wt-LinD was expressed and purified as described previously (SI Section 2 describes the cloning, expression and purification of wt-LinD and mutants).<sup>25</sup> In contrast to the LinD-geraniol complex (5G1U), which was obtained by soaking LinD crystals with linalool, new crystals of LinD were obtained using co-crystallization with racemic linalool in the crystal drop (SI Section 3; Table S2). The solution of one of these structures, refined to a resolution of 2.18 Ångstroms, gave an alternate crystal form to those obtained previously, in space group  $P2_1$ , with ten monomers (comprising two pentamers) in the asymmetric unit. Data collection and refinement statistics can be found in the Supporting Information (SI Section 4; Table S3). A distinguishing feature of this structure was

greatly improved electron density for  $\beta$ -myrcene than had previously been reported.<sup>24</sup> Four molecules of  $\beta$ -myrcene of six modelled in the ten available active sites were especially well-defined, with *B*-factors as low as 55, and with an average of 65, compared to 97 for the previously reported  $\beta$ -myrcene complex.<sup>24</sup> The improved resolution and density permits the definition of the C1-C2 and C3-C10 alkene bonds within the context of active site residues (Figure 3).



**Figure 3.** Active site of wt-LinD in complex with  $\beta$ -myrcene. Backbone and side chain carbon atoms of monomers A and B are shown in dark green and blue. Electron density corresponds to the  $2F_o - F_c$  omit map at a level of  $1\sigma$  and is the map obtained prior to addition refinement of the  $\beta$ -myrcene atoms, which have been added for clarity and are shown in yellow. The side chain of M125 is observed in both the 'open' and 'closed' conformations. Distances are given in Ångstroms.



**Figure 4.** Activity of purified wt-LinD and mutants determined by GC end-point assay. % axes normalized to 100% activity for wt-LinD. **A:** % reaction mixture following incubation with racemic linalool; **B:** % reaction mixture following incubation with myrcene; **C:** % reaction mixture following incubation with geraniol. % linalool,  $\beta$ -myrcene and geraniol are represented by red, blue and green bars respectively.

The ligand conformation presents the prochiral *re*-face of the C3-C10 double bond to the side of the active site with (B)D39 and (B)D45. Another notable observation within the active sites of the complex structure was the side-chain of M125 in both its 'open' and 'closed' conformation. In four out of ten active sites this side-chain was modelled in both conformations with equal occupancy (**Figure 3**).

#### Preparation and Analysis of a Library of LinD mutants

Following our previous structural studies,<sup>25</sup> a library of site-directed mutants was designed to explore the mechanism of LinD using activity assay, crystallography and mass spectrometry. In contrast to previous studies, which employed only cell extracts of *E. coli* containing mutant enzymes, each LinD mutant was purified by nickel-affinity chromatography and gel filtration prior to activity measurements (**SI Section 2**). The mutants created were: C171A, and C180A, to investigate possible roles in water activation and/or the formation of a covalent thioether intermediate; C171S, designed in an effort to slow reaction mechanism with a view to again possibly trapping a thioether intermediate; D39A and Y45F, to examine their possible roles as acid-base catalysts in the (de)protonation of the C10 methyl group; M125A, to examine the effect of removing this mobile residue; E172A and H129A, to investigate their possible role in the activation of water for the formation of the geraniol primary alcohol; and Y66F, as Y66 was suggested to be a candidate for the catalytic acid required to protonate the linalool hydroxyl in the linalool dehydration reaction.<sup>26</sup>

Following nickel affinity chromatography, gel filtration of each mutant permitted the observation of different oligomeric forms of the LinD variants, most of which existed as a mixture of monomeric and pentameric forms (**SI Section 5; Figure S2**), which were inactive and active, respectively, for LinD reactions. Interestingly, the ratio of monomer: pentamer was very different: Mutants C180A or M125A displayed more than 90% of the pentamer, whereas Y45F and E172A contained less than 10%, and undetectable amounts, of the

pentamer respectively. Similar observations were recently recorded for a selection of LinD mutants by Engleder and co-workers<sup>27</sup>.

The activity of pentameric forms of wt-LinD and the soluble mutants, isolated from gel filtration fractions, was determined using three separate, GC-based, end-point assays for the dehydration, the hydration and the isomerization reactions, using racemic linalool,  $\beta$ -myrcene and geraniol as substrates respectively (**SI Section 6 and 7; Figure 4**).

The dehydration of racemic linalool catalysed by wt-LinD displayed excellent enantioselectivity, giving  $\beta$ -myrcene with 48% conversion within 24 h and an *E* (enantioselectivity) value of >200 for the kinetic resolution. The reversible reaction was carried out with wt-LinD using  $\beta$ -myrcene as substrate, affording enantiopure (*S*)-linalool with 24% conversion after 24 h. When geraniol was used as the starting material with wt-LinD, nearly 65% of substrate was consumed after 24 h, with 53% transformation to  $\beta$ -myrcene and 12% to enantiopure (*S*)-linalool.

Mutation of cysteine residues C171 or C180 to either alanine or, in the case of C171, also serine, resulted in complete inactivation for either the (de)hydration of linalool and  $\beta$ -myrcene or the isomerization of geraniol, with no other intermediates detected by GC, in accordance with experiments performed on LinD cell extracts by ourselves<sup>25</sup> and the Harder group<sup>24</sup>. The activities of D39A and Y45F for linalool dehydration and  $\beta$ -myrcene hydration were dramatically reduced in each case, giving the corresponding products with a conversion of <5% (data not shown). This behavior was also observed for the D39A and Y45F in the geraniol isomerization, with each displaying  $\leq$  2% conversion after 24h. While the activity of the mutant M125A was reduced compared to the wt-LinD, it was not catalytically essential for either reaction, with 20% conversion for the linalool dehydration, 11% of conversion for  $\beta$ -myrcene hydration

**Table 1.** Kinetic constants for wt LinD and LinD mutants for linalool dehydration,  $\beta$ -myrcene hydration and geraniol isomerization (See SI for experimental details).

Variant	Reaction								
	Linalool Dehydration			$\beta$ -Myrcene Hydration			Geraniol Isomerisation		
	$k_{\text{cat}}^*$ ( $\text{min}^{-1}$ )	$K_m$ (mM)	$k_{\text{cat}}/K_m$ ( $\text{min}^{-1}$ $\text{mM}^{-1}$ )	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$K_m$ (mM)	$k_{\text{cat}}/K_m$ ( $\text{min}^{-1}$ $\text{mM}^{-1}$ )	$k_{\text{cat}}$ ( $\text{min}^{-1}$ )	$K_m$ (mM)	$k_{\text{cat}}/K_m$ ( $\text{min}^{-1}$ $\text{mM}^{-1}$ )
<b>Wt-LinD</b>	4.37	0.8 $\pm$ 0.1	6.24	9.0	0.2 $\pm$ 0.1	45.4	1.66	0.5 $\pm$ 0.1	3.32
<b>M125A</b>	0.07	0.6 $\pm$ 0.1	0.12	0.27	0.2 $\pm$ 0.1	1.38	0.88	0.5 $\pm$ 0.1	1.75
<b>H129A</b>	0.15	0.9 $\pm$ 0.1	0.17	0.31	0.4 $\pm$ 0.1	0.77	0.019	0.8 $\pm$ 0.1	0.023
<b>Y66F</b>	2.50	0.9 $\pm$ 0.1	2.77	0.96	0.4 $\pm$ 0.1	2.40	0.019	1.2 $\pm$ 0.1	0.017

and 20% for geraniol-isomerization reaction, affording 12% of  $\beta$ -myrcene and 8% of (*S*)-linalool. The mutants H129A and Y66F also displayed reduced activities for the linalool dehydration reaction (20% and 28% conversion respectively) and for the LinD-catalysed hydration of  $\beta$ -myrcene (7% and 10% conversion respectively). In the case of the geraniol dehydration-isomerization reaction, their activity was reduced to just 6% by mutant H125A (3% of  $\beta$ -myrcene and 3% of (*S*)-linalool) and 8% by Y66F (5% of  $\beta$ -myrcene and 3% of (*S*)-linalool). Interestingly, although the enzymatic activity was reduced significantly for some mutants, the enantioselectivity of the reactions was not affected in any case.

### Kinetics

In order to derive more information on the catalytic role of residues, the mutation of which resulted in measurable residual activity (M125, H129 and Y66), kinetic constants (Table 1; SI Section 8; Figure S4, S5, S6) for these and also wt-LinD were determined using GC analysis of samples from a homogenous system using 10% (v/v) DMSO as a co-solvent. Wt-LinD exhibited a  $K_m$  value of 0.8 mM for linalool dehydration, with a  $k_{\text{cat}}$  of 4.37  $\text{min}^{-1}$ . For the hydration of  $\beta$ -myrcene a lower  $K_m$  value (0.2 mM) was observed, with a higher  $k_{\text{cat}}$  (9.0  $\text{min}^{-1}$ ). For geraniol isomerization the values for the  $K_m$  and  $k_{\text{cat}}$  were 0.5 mM and 1.66  $\text{min}^{-1}$  respectively. The wt-LinD-catalysed hydration of  $\beta$ -myrcene displayed the highest  $k_{\text{cat}}$  and lowest  $K_m$  of all reactions therefore. Although  $K_m$  values observed for linalool dehydration and geraniol isomerization by wt-LinD broadly agreed with those reported by Brodkorb and co-workers,<sup>20</sup> our results suggest that linalool dehydration proceeds 2.6 times faster than geraniol isomerization, in contrast to their previous results where

geraniol isomerization was reported to be four times faster than the linalool dehydration.<sup>26</sup>

$K_m$  values for M125A were similar to the wt-LinD-catalysed processes, however,  $k_{\text{cat}}$  values observed for linalool dehydration and  $\beta$ -myrcene hydration were significantly reduced, to 0.07 mU and 0.27  $\text{min}^{-1}$  respectively. Geraniol isomerization activity by M125A was reduced only from 1.66  $\text{min}^{-1}$  to 0.88  $\text{min}^{-1}$ , suggesting that M125 plays a less crucial role in the catalysis of this process.

The  $K_m$  values for H129A were similar, or slightly higher, than those of wt-LinD for all reactions.  $k_{\text{cat}}$  values were substantially lower in all cases, with rates of linalool dehydration,  $\beta$ -myrcene hydration and the geraniol isomerization reaction reduced by 29-, 29- and 87-fold respectively. Finally, mutant Y66F exhibited a slightly higher  $K_m$  value for linalool dehydration (0.9 mM) and  $\beta$ -myrcene hydration (0.4 mM) than for wt-LinD. The  $K_m$  value for geraniol isomerization reaction (1.2 mM) was two-fold higher than for wt-LinD. Despite these changes in  $K_m$  the activity of Y66F remained relatively high for linalool dehydration and  $\beta$ -myrcene hydration, at 60% and 11% of wt activity respectively, although geraniol isomerization activity was proportionately much lower, reduced 87-fold compared to wt-LinD. In summary, a consideration of GC-based endpoint assays and the kinetics results points to more significant roles for M125 in (de)hydration reactions over geraniol isomerisation and for Y66 in the isomerisation over (de)hydration processes, but H129 appears to make a significant contribution to both.

### Crystal Structures of LinD Mutants

We next aimed to obtain crystallographic evidence for a possible thioterpene intermediate (**Figure 2B**) from a slow LinD mutant, wherein the mechanism might be interrupted through mutation of a residue involved in the second half of the (de)hydration reaction. This strategy had previously been employed successfully to obtain evidence of a cysteine-ligated covalent intermediate in maleate isomerase.<sup>28</sup> Hence, mutants C171A, C171S, C180A and M125A were crystallized in the presence of linalool as for the wt-LinD  $\beta$ -myrcene complex described above. Data collection and refinement statistics are given in the Supporting Information. Structures were obtained at resolutions of 2.52, 2.58, 1.45 and 1.99 Å respectively. However, in no case was density obtained consistent with a covalent intermediate bound through either Cys residue; nor was other significant ligand density observed within any active site. However, C171S and C171A featured M125 in the 'open' and 'closed' conformations respectively. In the C180A mutant, M125 is observed in both conformations and the higher resolution reveals this movement to extend to include the peptide bond to I124. This higher resolution also reveals two distinct conformations of the loop comprising E172, P173 and D174 at the monomer-monomer interface (**SI Section 9; Figure S7**). This movement displaces E172 from its position close to myrcene, as observed in the wt-LinD structure and breaks an H-bonding interaction with the side chain phenolic group of Y45.

### Mass Spectrometry studies of LinD, mutants and complexes

The possibility of a thioterpene intermediate was also investigated using mass spectrometry analysis of wt-LinD and mutants C171S, C171A, C180A, M125A, H129A, D39A, and Y45F, each of which had been incubated with linalool and digested overnight at 37°C following digestion with chymotrypsin. Resulting peptides were analyzed by LC-MS/MS using a Thermo Orbitrap Fusion mass spectrometer. Crucially, peptides A167GIVCEPDNY and V178QCNSVAY containing residues at positions 171 and 180, where linalool modification would be most likely, were confidently identified in their unmodified forms in all samples, so no additions to any cysteine residue were identified. Details of the MS analysis are given in Section 10 of the Supporting Information.

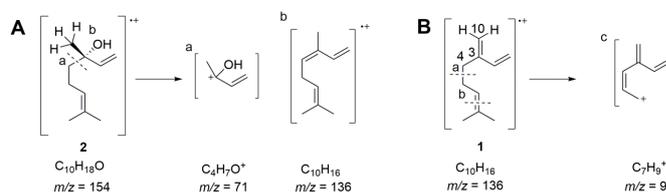
### LinD catalyses linalool-myrcene (de)hydration and isomerization to geraniol through non-covalent chemistry

A discussion of the results requires that hypotheses are presented for the two reversible reactions catalysed by LinD; the (de)hydration of linalool-myrcene, and the isomerization of linalool-geraniol. The discussion is simplified by a failure to find evidence of a covalent intermediate in LinD catalysis in these studies. While the involvement of such an intermediate cannot be absolutely ruled out, the absence of evidence resulting from extensive X-ray crystallography and mass spectrometry experiments prompts us to exclude the involvement of a covalent intermediate from further discussion. We initially propose therefore that LinD catalysis proceeds *via* non-covalent mechanisms, perhaps involving the intermediacy and stabilization of a central carbocation intermediate,

the subsequent transformation of which can yield either (de)hydration or isomerization products.

### (De)hydration

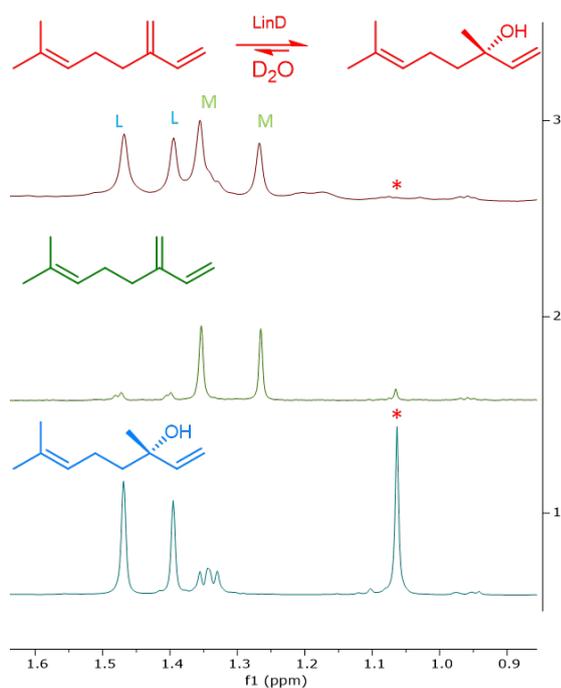
The position of  $\beta$ -myrcene in the structure presented, in conjunction with the activity studies on LinD mutants, permits assignment of roles to active-site residues in the hydration of this substrate. In agreement with QM/MM studies on LinD by Liu and co-workers,<sup>26</sup> we propose that the side chain of D39, on the *si*-face of the C3-C10 double bond of  $\beta$ -myrcene, is well placed to donate a proton to C10, to generate a carbocation, as it is positioned within approximately 3.5-4.0 Å of C10 in the  $\beta$ -myrcene complex structure, and its D39A mutant displays barely measurable residual activity for linalool and  $\beta$ -myrcene (de)hydration. The carboxylate of D39 interacts with the side-chain phenol of Tyr45, the mutation of which to phenylalanine also results in a mutant of greatly reduced activity. The (de)protonation at C10 was investigated using mass spectrometric analysis of the (de)hydration reactions in deuterated water (**SI Section 11**). Using  $\beta$ -myrcene as a substrate, 24% conversion to (*S*)-linalool was observed after 24 h incubation. The mass spectrum of non-labelled (*S*)-linalool yields, among others, major peaks at  $m/z = 71$  (resulting from bond cleavage between C3 and C4; 'a' in **Figure 5**) and  $m/z = 136$  (following dehydration; 'b' in **Figure 5**). MS analysis of the linalool peak after LinD-catalysed  $\beta$ -myrcene hydration (**SI Section 12; Figure S9**) revealed mass increases in the relevant fragments of +3 Da compared to the linalool standard, from  $m/z = 71$  to 74 and  $m/z = 136$  to 139 after 30 min and 12 h incubation periods.



**Figure 5.** Fragmentation of A  $\beta$ -myrcene and B (*S*)-linalool in MS analysis.

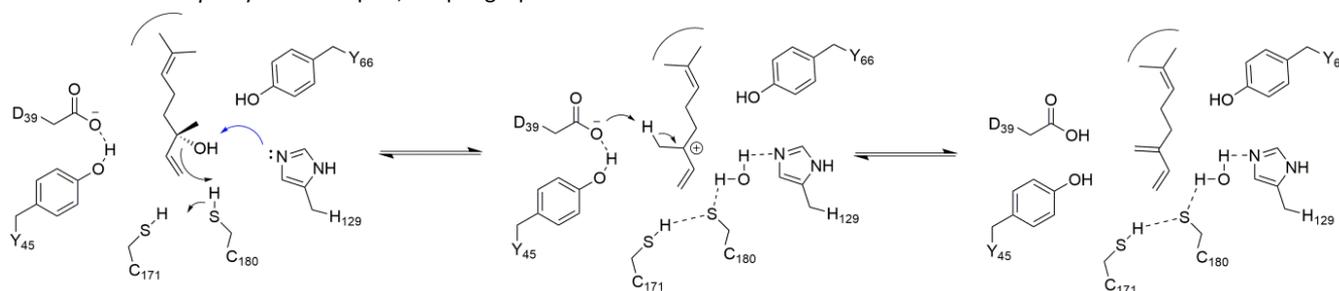
Moreover, the mass spectrum of the residual  $\beta$ -myrcene substrate (**SI Section 12; Figure S10**) displayed mass increases of +1 and +2 for the characteristic fragment  $m/z = 93$  ('c' in **Figure 5**), in a ratio of 1.0:0.5:1.0 after 30 min, and predominantly +2 after 12 h. In addition, monitoring the reaction by <sup>1</sup>H-NMR also revealed the absence of evolution of the C10 methyl peak of (*S*)-linalool (singlet expected at 1.21 ppm) following hydration of  $\beta$ -myrcene in deuterated water (**Figure 6; SI Section S13; Figure S13**). Each of these experiments is indicative of rapid and reversible proton transfer at C10 through the course of the (de)hydration reaction.

We<sup>25</sup> and others<sup>24,26</sup> have suggested that, following protonation at C10, a carbocation would be generated at C3. The hydration from one face of this intermediate exclusively would be necessary to generate the (*S*)-enantiomer of linalool. The only candidate water molecule within the  $\beta$ -myrcene complex structure is the one that is



**Figure 6.** Comparison of  $^1\text{H}$ -NMR spectra of linalool (blue) and  $\beta$ -myrcene (green) standards and LinD-catalysed  $\beta$ -myrcene hydration following 30 min incubation in deuterated water (red; L = linalool; M =  $\beta$ -myrcene; asterisk indicates absence of C10 methyl signal).

positioned between the side chains of H129 and C180. In a preliminary mechanistic hypothesis, Harder has also proposed these residues as the most significant for interaction with the hydrating water molecule.<sup>24</sup> Mutation of C180 and H129 to alanine gave, respectively, an inactive variant and one with 24-fold and 29-fold reduced velocity than wt-LinD for the linalool dehydration and the  $\beta$ -myrcene hydration reactions respectively (Table 1). The presence of this water molecule requires that the side-chain of M125 is in the 'open' conformation. However, as positioned in the  $\beta$ -myrcene complex, this water is on the *re*-face of the C3-C10 double bond, and would give the (*R*)-enantiomer of linalool as a product. We propose therefore, that the intermediate would have to be rotated to assume a conformation approximately  $160^\circ$  on its longitudinal axis relative to that observed in the  $\beta$ -myrcene complex, adopting a position that is

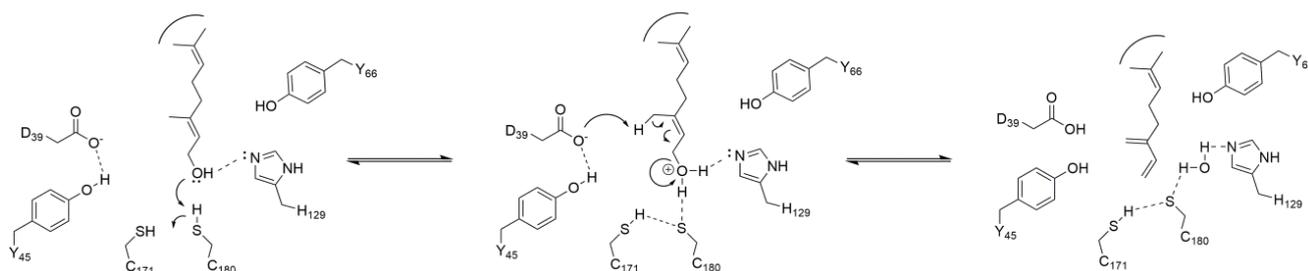


**Figure 7.** Proposed mechanism for (de)hydration reaction of linalool catalysed by LinD. H129 and C180 (and C171) catalyse the removal of water from the chiral carbon C3 of (*S*)-linalool to yield a carbocation that is protonated by D39, assisted by Y45.

more similar to the  $\beta$ -myrcene modelled in structure 5HSS.<sup>24</sup> This would present the *si*-face of the intermediate to the water molecule, which quenches the carbocation to give the (*S*)-linalool product (Figure 7). This mechanism is at variance with that proposed by Liu and co-workers,<sup>26</sup> in which Y66 activates a water molecule ('W14') for the hydration of myrcene. Mutation of Y66 to phenylalanine has comparatively little effect on LinD (de)hydration activity (Table 1), and no crystal structure of LinD at any resolution features a water molecule, interacting with both C180 and Y66, in the position proposed in that model.

#### Isomerisation

For the geraniol-linalool isomerization reaction, the results of both mutation and structural studies should be reconciled with the previously obtained structure of LinD in complex with geraniol (5G1U). The complex with geraniol suggests that the hydroxyl on carbon-1 occupies a position in the active site between the side chains of H129 and C180, the proposed location of the water molecule needed for the hydration of  $\beta$ -myrcene, and which is occupied in some structures by the side chain sulfur atom of M125. The geraniol complex suggests therefore that the cavity between C180 and H129 hosts and activates the water for addition/elimination to the C1-C2 double bond of  $\beta$ -myrcene in the isomerization reaction. In our hands, mutation of H129 to alanine has an even more marked effect on geraniol isomerization than on linalool dehydration with an 83-fold reduction in velocity compared to wt-LinD, consistent with this hypothesis. However, it should be noted that structure 5HSS, geraniol was modelled to occupy an alternate conformation, in which the 1-hydroxyl occupies a position at the 'bottom' of the active site, interacting with C180, Y45 and also E172. In that structure the electron density is less clear than in 5G1U, and the average *B*-factors of the geraniol ligand are 90-100, compared to 55-65 for 5G1U. The *in silico* model,<sup>26</sup> based on the 5HSS geraniol conformation, also suggests that the geraniol 1-hydroxyl is derived from a water molecule activated by E172. Our mutation of E172 gave inactive E172A, although analytical gel filtration studies suggested that this mutant does not oligomerise, and inactivity might thus be attributed to a failure to form an active site. In addition, the conformation of geraniol in 5HSS does not explain the large reduction in activity in geraniol isomerization when H129 is mutated to alanine.



**Figure 8.** Proposed mechanism for dehydration reaction of geraniol to  $\beta$ -myrcene leading to (*S*)-linalool catalysed by LinD. H129 and C180 (with C171) catalyse the removal of water from C1 of geraniol in an elimination reaction that yields the intermediate  $\beta$ -myrcene, which can be converted to (*S*)-linalool as shown in **Figure 7**.

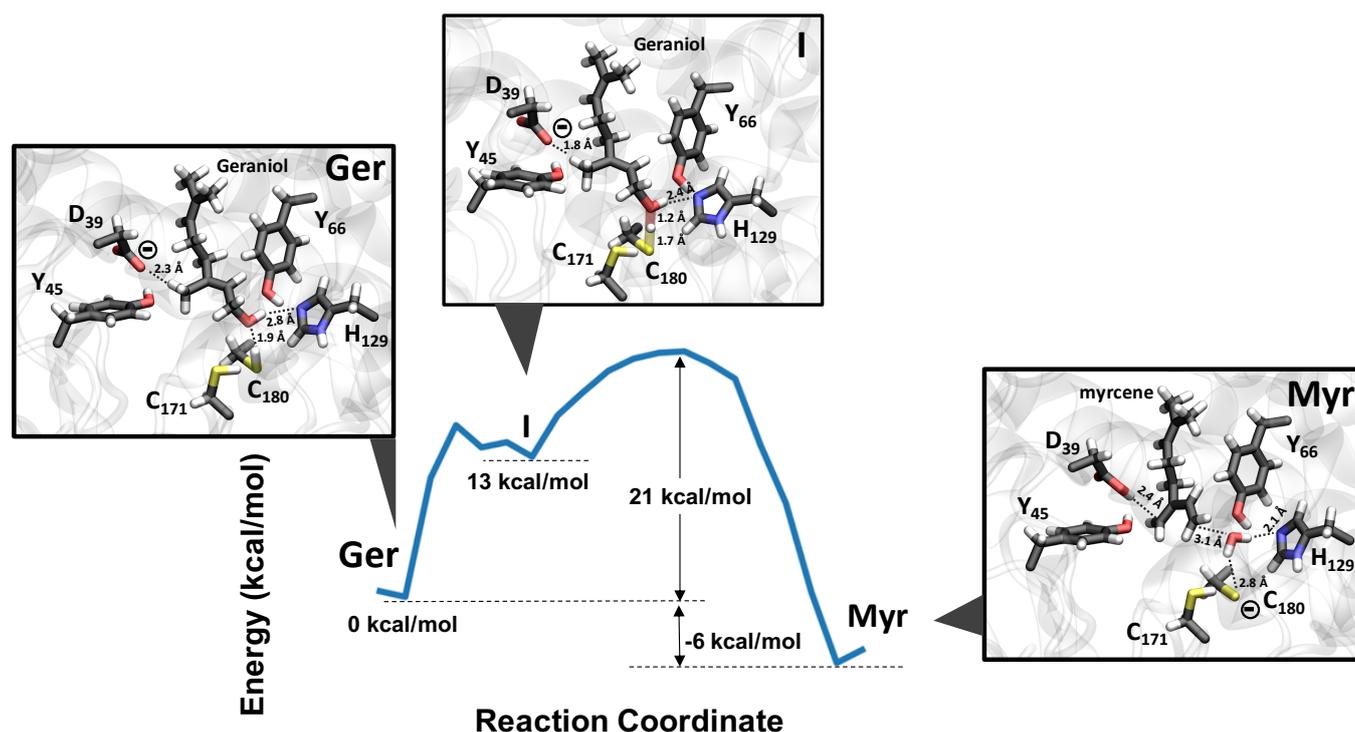
The modelling studies by Liu and co-workers have also suggested that the dehydration of geraniol leads to the central intermediate carbocation at C3, from which either protonation at C10 or hydration at C3 leads to  $\beta$ -myrcene or (*S*)-linalool respectively.<sup>26</sup> When geraniol is used as the substrate for wt-LinD-catalysed isomerization in deuterated water, after 30 min, geraniol was converted into 50%  $\beta$ -myrcene and 28% (*S*)-linalool (**SI Section 12; Figure S12**), with 80% and 20% of these products after 12 h. The *m/z* for the characteristic  $\beta$ -myrcene fragment after 12 h was again observed to be 95 Da, +2 Da higher than that of the standard fragment. The *m/z* for (*S*)-linalool fragments were increased by +3 Da, to 74 and 139 (**Figure S12**). The accumulation of deuterium at C10 in (*S*)-linalool suggests that conversion of geraniol to linalool *via* the carbocation at C3 is less likely, as exchange of H for D at C10 would not be expected *via* this pathway. Hence, we propose that 1,4-elimination of geraniol, catalysed first *via* protonation of the C1 hydroxyl by C180, would first occur to give  $\beta$ -myrcene, as this would then implicate H/D exchange, catalysed by D39, as part of the mechanism. This would also explain the almost complete inactivation of LinD for geraniol isomerization when D39 and Y45 are mutated to alanine (*vide supra*).  $\beta$ -myrcene would then undergo hydration to (*S*)-linalool as described previously to complete the geraniol-linalool conversion (**Figure 8**).

To further validate the proposed elimination mechanism, we performed QM/MM simulations, coupled to a nudged elastic band (NEB) approach, to investigate with atomistic details the geraniol to  $\beta$ -myrcene conversion, and to estimate the associated potential energy landscape along the reaction coordinate (**SI Section 14**). The binary complex LinD-geraniol was considered as the starting point of the simulations, equilibrated by a combination of force field based and QM/MM geometry optimizations. A detailed description of the reaction mechanism can be obtained by following the reaction path on the reconstructed energy landscape (**Figure 9**). For the elimination mechanism, the reactants state, *i.e.* the complex between LinD and geraniol (**Ger, Figure 9**), is characterized by a

hydrogen bond between the substrate hydroxyl group and enzyme residues H129, and C180 in agreement with the reduced activity of both alanine mutants. Besides, D39 is close to the C10 atom from geraniol and, therefore, well suited to donate/abstract a proton from this carbon atom. This group-state interaction is in agreement with previous structural and mechanistic hypotheses but also with those presented in this work. The transformation of geraniol to myrcene starts with the efficient transfer of a hydrogen atom from the C180 thiol group to the geraniol hydroxyl group, thus forming the protonated alcohol (**Figure 9, I**). This protonated alcohol is stabilized by the His129 - N $\epsilon$ ... OH<sub>2</sub> - geraniol hydrogen bond, which not only stabilizes the positive charge being developed at the hydroxyl group, but also properly positions the ligand molecule to favor this proton transfer. The protonated alcohol species is completely formed at intermediate I, which is 13 kcal mol<sup>-1</sup> higher in energy than **Ger**. Finally, water removal followed by proton abstraction at C10 by D39 in a 1,4 elimination process yields  $\beta$ -myrcene **Myr**, which is approximately 6 kcal mol<sup>-1</sup> more stable than **Ger** (**Figure S15**). The estimated energy barrier for the geraniol to  $\beta$ -myrcene conversion is approximately 21 kcal mol<sup>-1</sup>, which is in reasonable agreement with previous calculations by Liu and co-workers.<sup>26</sup> The energy difference between geraniol and  $\beta$ -myrcene of approximately 6 kcal mol<sup>-1</sup> is also in agreement with the product distribution observed experimentally. It is also worth mentioning that the computed energy barrier for the formation of a covalent intermediate is 39 kcal mol<sup>-1</sup> (**Figure S16**), therefore, excluding this reaction path for the conversion for geraniol to  $\beta$ -myrcene. In conclusion, the proposed computational methodology has shown that the lowest energy reaction path from reactants to products takes place *via* the 1,4-elimination mechanism as postulated experimentally.

#### Role of M125 and Y66

The electron density observed in our previous structures 5H1U (geraniol complex) and 5H1W (*apo*-structure) clearly show that the

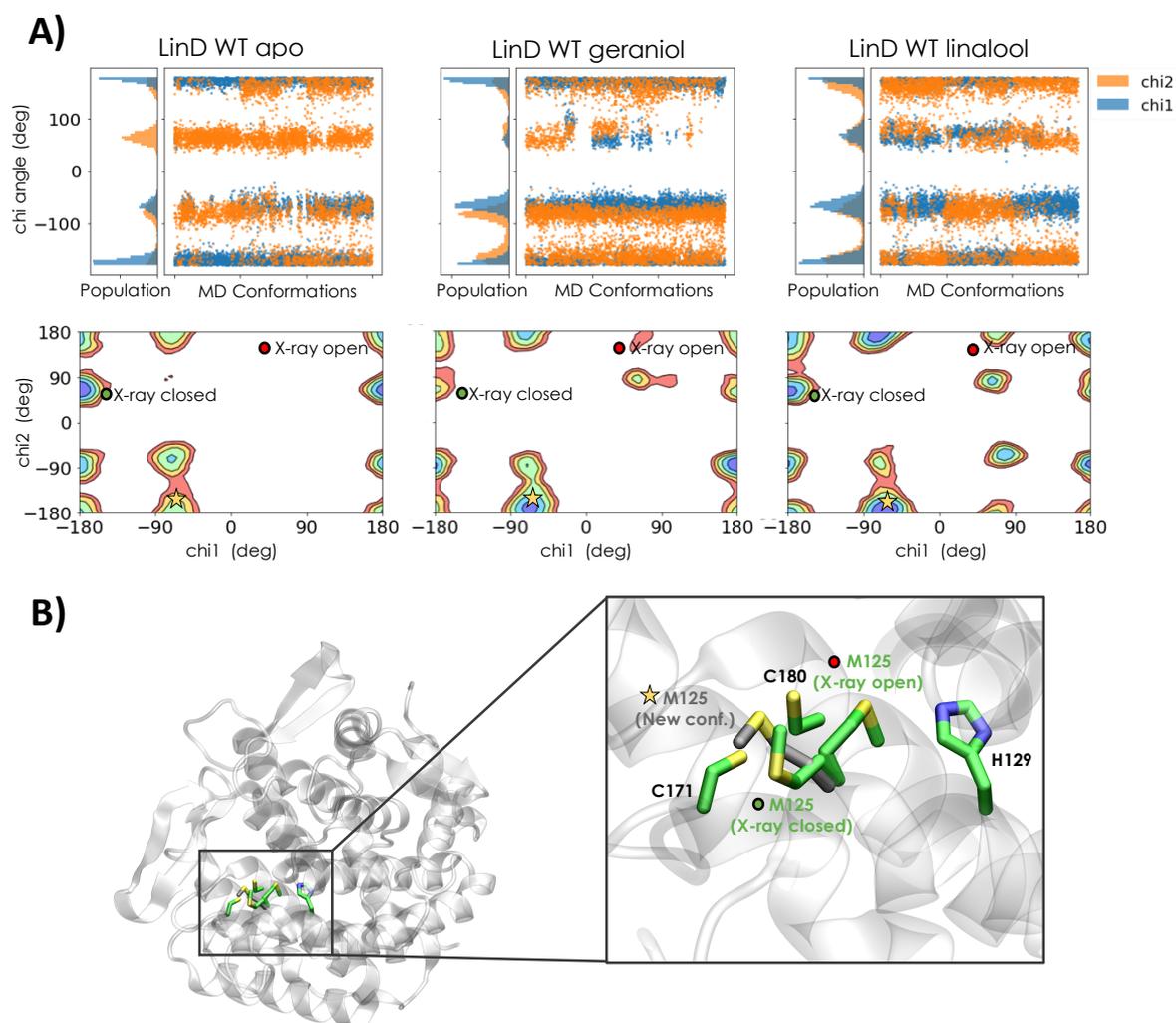


**Figure 9.** QM/MM-neb simulations for the geraniol (**Ger**) to  $\beta$ -myrcene (**Myr**) conversion. The associated potential-energy landscape along the reaction coordinate is shown, and the relative stabilities and activation energies are expressed in kcal mol<sup>-1</sup>. The most relevant distances at the **Ger**, **I**, and **Myr** minima are shown in Å.

side chain of M125 is able to adopt two conformations, which we have termed ‘open’ and ‘closed’, in the active site. Indeed, the newly acquired  $\beta$ -myrcene complex in this report displays mixed conformations of M125 in some active sites, and C180A and C171A structures feature the ‘open’ and ‘closed’ conformations exclusively. Considered together, the structural, MD simulation and mutation data suggest that the side-chain of M125 must vacate a water site between C180 and H129 in order for the enzyme to accommodate a water molecule, or the C1-hydroxyl of geraniol, for catalytic turnover. One role of M125 may therefore be to act as a switch, enabling water/geraniol binding when the enzyme is in the presence of sufficient amounts of substrate. However, the  $K_m$  values displayed by M125A for linalool,  $\beta$ -myrcene, or geraniol were not substantially affected by mutation. The catalytic activity of M125A, by contrast, was reduced for all the reactions, but substantially so for the linalool (de)hydration reactions, suggesting a more significant role for M125 in the transformation of these substrates.

In an attempt to answer why M125 displays two different conformations and specific interactions within the LinD active site, we modelled the LinD enzyme in the *apo*- and geraniol/linalool bound states. Three independent classical molecular dynamics (MD) simulations of 100 ns were run for each homopentameric system (*i.e.*

300 ns in the *apo*, and geraniol/linalool bound states). Monitoring of M125 ‘open’/‘closed’ conformations was performed by measuring side chain *chi1* and *chi2* torsional angles in all MD trajectories, discarding the first 20 ns of simulation time (see full computational details in SI). We first analyzed whether the ‘open’/‘closed’ M125 conformation could have an impact in substrate binding by measuring substrate accession channels in ‘open’ and ‘closed’ X-ray structures and representative MD snapshots in the *apo* state. As shown in **Figure S14**, ‘open’/‘closed’ M125 orientations do not affect the accession channel bottleneck radius (BR) either in X-ray structures (BR = 1.9 Å) or in the MD simulations (BR of  $1.21 \pm 0.13$  Å and  $1.22 \pm 0.14$  Å for ‘open’ and ‘closed’ structures, respectively). Conformational population analysis was subsequently performed from the MD simulations to monitor the different stable conformations explored in each system. ‘Closed’ conformations of M125 are mainly found for LinD in the *apo* state (*chi1* values of ca. -170 deg), whereas both ‘open’ and ‘closed’ conformations are visited in the ligand-bound state (value of 30 deg and -170 deg, respectively). This is in agreement with the requirement to accommodate a water molecule in the *apo* state or geraniol hydroxyl group. Interestingly, a new intermediate conformation of M125 was characterized in the MD simulations,



**Figure 10.** Conformational population analysis of M125 in apo and substrate-bound LinD systems: A: Histogram and conformational population analysis reconstructed from MD simulations in the apo, geraniol bound or linalool bound states. M125 X-ray structures are shown with a green or red circle, whereas the newly characterized conformation is shown with a yellow star. B: Detailed atomistic representation of the different conformations observed for M125.

which is particularly stable in the LinD ligand-bound state (**Figure 10**). This new conformation is characterized by a  $\chi_1$  angle of ca. -70 deg, which lies between ‘open’ and ‘closed’ conformations, thus allowing the presence of a water molecule or the ligand hydroxyl group. This new conformation also presents a rather different  $\chi_2$  angle (-175 deg) as the methionine sulfur atom points towards the ligand active site region, having therefore potential implications in catalysis as discussed below.

Methionine residues have been suggested to have an antioxidant role,<sup>29</sup> in which they serve to protect essential and sensitive cysteine residues, such as those encountered in LinD, from oxidative stress, however the requirement for DTT in *in vitro* reactions catalysed by

LinD suggests that M125 does not have an antioxidant role adequate to support the activity of the enzyme in the absence of another agent. Methionine residues have, however, also been implicated in catalytic roles through the stabilization of transition states, as observed for phosphite dehydrogenase,<sup>30</sup> and indeed for terpenoid carbocations in the prenyltransferases UbiA and NphB.<sup>31</sup> We hypothesize that the increased dependence upon M125 for catalytic activity in the linalool (de)hydration reaction may be because it assists in carbocation stability in those reactions, but not in the conversion of geraniol to  $\beta$ -myrcene, for which a carbocation-independent route is proposed. To shed some light on the catalytic role of M125, QM/MM geometry optimizations were additionally performed. We modelled the LinD – linalool complex with M125 in a

conformation taken from the MD simulations that presents the thiol group pointing towards the ligand hydroxyl group (*vide supra*). A TS-like structure for the isomerization to linalool was enforced by elongating the C3–hydroxyl group bond to a distance of approximately 3 Å. Analysis of the geometry optimized structure (Figure S17) revealed that the M125 sidechain indeed forms a hydrogen bond with the hydroxide anion being formed, thus stabilizing the formation of the carbocation transition state of the reaction and having an impact on the energetics of the reaction. Altogether, the combination of experiments, structural analysis and molecular modelling techniques demonstrate the importance of M125 in LinD catalysis.

The role of residue Y66 in catalysis is more cryptic. Mutation of this residue to phenylalanine has a modest effect on linalool (de)hydration. This may be expected as in no structure is the phenolic hydroxyl of Y66 in sufficiently close proximity to a water molecule to suggest a role in water activation/removal in those processes. A role in carbocation stabilization in the (de)hydration reactions through  $\pi$ -cation interactions is plausible, as these should be retained in the conservative mutation of this side-chain to phenylalanine, and a role in terpene carbocation stabilization has been established for tyrosine in terpene synthases previously.<sup>32</sup> However, activity is more markedly reduced for Y66F in geraniol conversion initially to  $\beta$ -myrcene, which we propose to be predominantly a carbocation-independent process. Determination of the precise role of Y66 in LinD catalysis therefore requires further investigation.

## Conclusions

Although LinD catalyses what at first sight appears to be a fairly simple chemical reaction, it is clear that a complex cascade of reaction steps, assisted by a constellation of active site residues, is required to achieve the necessary reaction specificity and stereoselectivity in response to the binding of different substrates. Our combined experimental and computational data suggest that the interconversion of linalool and  $\beta$ -myrcene is likely achieved using exclusively acid-base chemistry *via* a C3 carbocation intermediate. However labelling experiments and QM/MM-neb simulations suggest that the isomerization reaction of geraniol to linalool is partially decoupled from the carbocation route and proceeds through the intermediate generation of  $\beta$ -myrcene *via* an addition-elimination process to or from geraniol. These observations provide further information for the rational engineering of LinD for applications in the production of both flavor-fragrance compounds but also bulk alkenes for the production of plastics precursors.

## Acknowledgements

A.C. was funded by grant BB/P005578/1 from the BBSRC. We thank Dr Johan P. Turkenburg and Mr Sam Hart for assistance with X-ray data collection and the Diamond Light Source for access to beamlines I02, I03 and I04-1 under proposal number mx-9948. The York Centre

of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1). This study was also supported in part by the European Research Council Horizon 2020 research and innovation program (ERC-2015-StG-679001, S.O.), Spanish MINECO (project PGC2018-102192-B-I00, S.O.), Generalitat de Catalunya for the emerging group CompBioLab (2017 SGR-1707); Juan de la Cierva-Incorporación fellowship IJCI-2017-34129, J. I.F), and Marie Curie EnzVolNet fellowship (H2020-MSCA-IF-2016-753045, J. I. F).

## Conflicts of interest

There are no conflicts to declare.

## Notes and references

- 1 J. Clayden, N. Greeves and S. Warren. Organic Chemistry. Oxford University Press, 2nd ed. (2012).
- 2 A.J. Boersma, D. Coquière, D. Geerdink, F. Rosati, B.L. Feringa and G. Roelfes, *Nat. Chem.*, 2010, **2**, 991.
- 3 L. Xue, B. Jia, L. Tang, X.F. Ji, M.Y. Huang and Y.Y. Jiang, *Polym. Advan. Technol.* 2004, **15**, 346.
- 4 S. Wang, Z. Zhang, C. Chi, G. Wu, J. Ren, Z. Wang, M. Huang and Y. Jiang, *React. Funct. Polym.*, 2008, **68**, 424.
- 5 J. Jin and U. Hanefeld, *Chem. Commun.*, 2011, **47** 2502.
- 6 V. Resch and U. Hanefeld, *Catal. Sci. Technol.*, 2015, **5**, 1385.
- 7 M. Engleder and H. Pichler, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 5841.
- 8 R.M. Demming, M-P. Fischer, J. Schmid and B. Hauer, *Curr. Opin. Chem. Biol.*, 2018, **43**, 43.
- 9 Z. Chi, Z.-P. Wang, G.-Y. Wang, I. Khan and Z.-M. Chi, *Crit. Rev. Biotechnol.*, 2014, **36**, 99.
- 10 G. Agnihotri and H.-W. Liu, *Bioorg. Med. Chem.*, 2003, **11**, 9.
- 11 P.J. Kuhn and D.A. Smith, *Physiol. Plant Pathol.*, 1979, **14**, 179.
- 12 D. Li, K.R. Chung, D.A. Smith and C.L. Scharl, *Mol. Plant. Microbe*, 1995, **8**, 388.
- 13 M. Engleder, M. Horvat, A. Emmerstorfer-Augustin, T. Wriessnegger, S. Gabriel, G. Strohmeier, H. Weber, M. Müller, I. Kaluzna, D. Mink, M. Schürmann and H. Pichler, *PLOS ONE*, 2018, **13**, e0192653.
- 14 A. Hiseni, L.G. Otten and I.W.C.E. Arends, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 1275.
- 15 Z. Sun, S. Shen, C. Wang, H. Wang, Y. Hu, J. Jiao, T. Ma, B. Tian, and Y. Hua, *Microbiology*, 2009, **155**, 2775–2783.
- 16 C. Wuensch, J. Gross, G. Steinkellner, K. Gruber, S.M. Glueck, and K. Faber, *Angew. Chem. Int. Ed.*, 2013, **52**, 2293.
- 17 S.E. Payer, H. Pollak, S.M. Glueck and K. Faber, *ACS Catal.*, 2018, **8**, 2438.
- 18 M. Engleder, T. Pavkov-Keller, A. Emmerstorfer, A. Hromic, S. Schrempf, G. Steinkellner, T. Wriessnegger, E. Leitner, G. A. Strohmeier, I. Kaluzna, D. Mink, M. Schürmann, S. Wallner, P. Macheroux, K. Gruber and H. Pichler, *ChemBioChem*, 2015, **16**, 1730.
- 19 R.M. Demming, S.C. Hammer, B.M. Nestl, S. Gergel, S. Fademrecht, J. Pleiss and B. Hauer, *Angew. Chem. Int. Ed.*, 2019, **58**, 173.
- 20 D. Brodkorb, M. Gottschall, R. Marmulla, F. Lüddecke and J. Harder, *J. Biol. Chem.*, 2010, **285**, 30436.
- 21 P. Marliere, 2013. EP Grant 2890793.
- 22 A. Botes and A. Conradie, 2015, US Patent Grant 9862973.

- 23 P. Marliere, M. Delcourt, S. Mazaleyrat and R.B. Le, 2016, US Patent Application 2016186161.
- 24 S. Weidensweber, R. Marmulla, U. Ermier and J. Harder, *FEBS Lett.*, 2016, **590**, 1375.
- 25 B.M. Nestl, C. Geinitz, S. Popa, S. Rizek, R.J. Haselbeck, R. Stephen, M.A. Noble, M.P. Fischer, E.C. Ralph, H.T. Hau, H. Man, M. Omar, J.P. Turkenburg, S. Van Dien, S.J. Culler, G. Grogan and B. Hauer, *Nat. Chem. Biol.*, 2017, **13**, 275.
- 26 B. Ling, X. Wang, H. Su, R. Liu and Y. Liu, *Phys. Chem. Chem. Phys.*, 2018, **20**, 17342.
- 27 M. Engleder, M. Müller, I. Kaluzna, D. Mink, M. Schurmann, H. Pichler and A. Emmerstorfer-Augustin, *Molecules*, 2019, **24**, 2092.
- 28 F. Fisch, C.M. Fleites, M. Delenne, N. Baudendistel, B. Hauer, J.P. Turkenburg, S. Hart, N.C. Bruce and G. Grogan, *J. Am. Chem. Soc.*, 2010, **132**, 11455.
- 29 R.L. Levine, J. Moskovitza and E.R. Stadtman, *IUBMB Life*, 2000, **50**, 301.
- 30 K.E. Ranaghan, J.E. Hung, G.J. Bartlett, T.J. Mooibroek, J.N. Harvey, D.N. Woolfson, W.A. van der Donk and A.J. Mulholland, *Chem. Sci.*, 2014, **5**, 2191.
- 31 R. Mutendam, 2015, PhD thesis, University of Groningen.
- 32 D.W. Christianson, *Chem. Rev.*, 2017, **117**, 11570.