Nucleophilic Water Capture or Proton Loss: Single Amino Acid Switch Converts δ-Cadinene Synthase into Germacradien-4-ol Synthase

Marianna Loizzi, Veronica González, David J. Miller, and Rudolf K. Allemann

δ-Cadinene synthase is a sesquiterpene cyclase that utilises the universal achiral precursor farnesyl diphosphate (FDP) to generate predominantly the bicyclic sesquiterpene δ-cadinene and about 2% germacradien-4-ol, which is also generated from FDP by the cyclase germacradien-4-ol synthase. Herein, the mechanism by which sesquiterpene synthases discriminate between deprotonation and reaction with a nucleophilic water molecule was investigated by site-directed mutagenesis of δ-cadinene synthase. If W279 in δ-cadinene synthase was replaced with various smaller amino acids, the ratio of alcohol versus hydrocarbon product was directly proportional to the van der Waals volume of the amino acid side chain. DCS-W279A is a catalytically efficient germacradien-4-ol synthase ($k_{cat}/K_m = 1.4 \times 10^3 \text{ M}^{-1} \text{s}^{-1}$) that produces predominantly germacradien-4-ol in addition to 11% δ-cadinene. Water capture is not achieved through strategic positioning of a water molecule in the active site, but through a coordinated series of loop movements that allow bulk water access to the final carbocation in the active site prior to product release.

Introduction

Present in all kingdoms of life, terpene synthases catalyse highly complex biosynthetic reactions, in which achiral, linear isoprenyl diphosphates are converted into complex often cyclic or polycyclic structures.[1,2] In most cases, these carboxatonic reaction cascades are characterised by high stereo- and regioselectivity and involve changes in the hybridisation of up to half of the carbon atoms. The final carbocation can either lose a proton to generate a hydrocarbon product or react with water to produce a terpene alcohol.[1] Subsequent biosynthetic reaction steps convert the products generated by terpene synthases into tens of thousands of terpenoids that act, among other things, as pigments; phytoalexins; semiochemicals; or in medicine, provided the original work is properly cited.

Motifs (DDXXD and NSE/DTE) on opposite sides of the active site.[2] They provide a three-dimensional template to bind the carbocation interactions and general acid–base mechanisms of the catalysed reactions has revealed common structural features and distinct phases of activity.[1–5] Class I terpene synthases share a predominantly α-helical fold with an active site lined with mostly hydrophobic and aromatic amino acid residues; they contain two conserved Mg$^{2+}$-binding motifs (DDXXD and NSE/DTE) on opposite sides of the active site.[2] They provide a three-dimensional template to bind the carbocation intermediates along distinct reaction paths. Class I terpene synthases initiate the chemical reaction by catalysing the cleavage of the carbon–oxygen bond of the substrate to generate a tightly bound diphosphate (PP)–carbocation pair.[2,7] Single-crystal X-ray structures of several sesquiterpene synthases complexes with (E,E)-farnesyl diphosphate (FDP, 1), several analogues of 1, Mg$^{2+}$ and PP, together with molecular dynamics simulations, have provided strong support that loop movements and conformational changes are required to form the closed form of the enzyme, in which substrate 1 is in a reaction-ready conformation.[2,8] After conformational rearrangements of enzyme and substrate necessary to form the Michaelis complex, the chemical reaction occurs, with major contributions from carbocation stabilisation by the π electrons of aromatic amino acid residues, PP, carbocation interactions and general acid–base catalysis by PP, and/or the enzyme.[1–12] The closed conformation of sesquiterpene synthases that generates hydrocarbon products also prevents access of bulk solvent, which avoids quenching of the reactive carbocationic reaction intermediates by water.[1–3,6–11]

In contrast to the wealth of knowledge available on the mechanistic details of sesquiterpene synthase catalysed reactions, in which the last cationic intermediate is deprotonated by PP, little information is available for terpene alcohol synthases, such as epicubenol,[13] hedycaryol,[14] avermitil[15] epicedrol[16] and germacradien-4-ol synthases[17] (GdoIS), which generate their products through the reaction of the final carbocation with water. The mechanism by which sesquiterpene synthases discriminate between deprotonation or water cap-
ture has not been explored in detail. In enzymes that generate alcohols, deprotonation of the final cation must be prevented, and capture of the final carbocationic intermediate must be tightly controlled to prevent quenching of early cationic intermediates, whereas synthases that generate terpene alcohols must prevent deprotonation of the final carbocationic intermediate. Clearly, terpene synthases have evolved to control water access and reactivity. Tightly bound water molecules can be found in the active sites of terpene synthases in their closed conformation, even for synthases that do not generate alcohol products.\cite{2,7,9} This finding might suggest that these sequestered water molecules could be responsible for nucleophilic capture of the final carbocationic species.\cite{2,14} However, a previously published investigation into the structure and mechanism of GdolS revealed that the reaction of the final carbocation most likely depended on specific loop movements of the enzyme that allowed bulk water to access the active site.\cite{18} GdolS must prevent deprotonation from C6 in intermediate 6, so as not to produce δ-cadinene (7; Scheme 1), which in contrast is the pathway for catalysis by δ-cadinene synthase (DCS).\cite{21}

![Scheme 1. Catalytic mechanisms of the DCS (pathway a) and GdolS (pathway b) catalysed conversions of 1 to 7 and germacradien-4-ol (8).](image)

In selina-4(15),7(11)-diene synthase from Streptomyces pristinaespiralis, a G1/2 helix break motif, combined with a diphosphate sensor–linker–effector motif that is conserved throughout bacterial sesquiterpene synthases, has been proposed to play a major role in substrate binding and active-site closure.\cite{21} This "kink" in the G-helix has also been noted as a potentially important catalytic feature in human squalene synthase\cite{22} and hedycaryl synthase.\cite{23} Interestingly, DCS was converted into a GdolS through saturation mutagenesis by Keasling and colleagues.\cite{21} A re-evaluation of that work reveals that the amino acid residues that generate this change of function are located in the G1/2 helix break motif (see below).

DCS from Gossypium arboreum produces 7 in the first committed step of the biosynthetic pathway to the phytoalexin gossypol (Scheme 1).\cite{21} Despite little sequence identity outside the conserved Mg\textsuperscript{2+}–binding motifs, DCS and GdolS share the typical structure of class I terpene synthases in the catalytic domain,\cite{9,18} many aspects of their active-site compositions and the respective catalytic reaction pathways from 1 to 7 or 8 are shared (Scheme 1). The co-crystal structure of DCS, Mg\textsuperscript{2+} and the substrate analogue (2Z,6E)-2-farnesyl diphosphate (2F-FDP) revealed an unusual Mg\textsuperscript{2+}–binding motif, in which the NSE/DTE motif is replaced by a second DDXXD motif.\cite{21} Interestingly, depending upon the substrate used, DCS appeared to follow a 1,6; 1,10 or 1,11 ring closure;\cite{21} this suggested that these pathways were energetically similar and that DCS might have some inherent promiscuity, despite its high fidelity when acting on 1.\cite{21} Only the N-terminal tail of the N-terminal β-domain of DCS is involved in catalysis and plays an important role in protecting the active site from water. Truncated proteins that lack the first 8 and 20 amino acids of the β-domain produce increasing amounts of 8.\cite{21}

The double-mutant protein DCS-N403P/L405H was shown to convert 1 into 8 (93 %) and an additional unidentified cyclic sesquiterpene alcohol.\cite{21} However, the catalytic activity of DCS-N403P/L405H is severely compromised relative to that of the wild-type (WT) enzyme.

In DCS-N403P/L405H, the active site is exposed to solvent through a potential water channel created by alteration of the G-helix residues.\cite{21} Analysis of the X-ray structure of DCS reveals that the aromatic residues W279 and Y410 are closer to the isoprenyl chain of the substrate and on the opposite side of the active-site contour relative to N403 and L405. W279 is on the C helix of DCS and within 7 Å of Y410, just below the G2 helix and towards the bottom of the active-site cleft (Figure 1). These two residues are ideally placed not only to stabilise carbocationic intermediates during the formation of 7, but also to form hydrophobic interactions that may help to control the active-site conformation of 1 and mediate active-site closure and opening. Hence, to test the hypothesis that alteration of W279 can disrupt hydrophobic interactions with Y410 and 1 and allow increased water access to the active site, the contribution of W279 to catalysis was explored by site-directed mutagenesis. Herein, we show that single amino acid changes can convert DCS into GdolS that produce up to 90 % 8 with high catalytic efficiency. The results suggest that W279 plays a key role in shielding the active site of DCS from solvent.

**Results and Discussion**

DCS-His\textsubscript{5} was produced in Escherichia coli and its catalytic properties determined. Steady-state kinetic experiments with radiolabelled [1\textsuperscript{4}H]\textsuperscript{3} 9,\textsuperscript{10} revealed a turnover number, \( k_{\text{cat}} \), of \( 1.26 \times 10^{-1} \text{ s}^{-1} \); a Michaelis constant, \( K_m \), of \( 0.58 \mu \text{M} \) (Table 1) and a catalytic efficiency, \( k_{\text{cat}}/K_m \), of \( (2.17 \pm 0.4) \times 10^{-4} \text{ s}^{-1} \mu \text{M}^{-1} \).
which was identical to the value previously measured for WT-DCS with no His-tag ($k_{cat}/K_m = (3.1 \pm 0.2) \times 10^{-3} \, \text{s}^{-1} \, \text{mM}^{-1}$).

Analysis of the pentane-extractable products arising from incubations with 1 by GC-MS showed that, in addition to $m$-aminobenzyl alcohol relative to that of DCS-His$_6$ (7 and 8 in approximately 2:1 ratio). These results show that the relative amounts of 7 and 8 are dependent on the volume of the side chain of residue 279. The values of $K_m$ for DCS-His$_6$-W279M and DCS-His$_6$-W279L were slightly increased, relative to that of the DCS-His$_6$ (Table 1). Smaller residues that were hydrophilic, rather than hydrophobic, were tested to examine the possibility that hydrophilic residues might form a repulsive interaction with Y410, leading to a poorly defined active site that compromised the catalytic activity. Alternatively, an increase in the mobility of the G1/G2 helix break motif might generate larger quantities of 8. Consequently, W279 was replaced with glutamine, glutamate and aspartate. When incubated with 1, DCS-His$_6$-W279Q, DCS-His$_6$-W279E and DCS-His$_6$-W279D generated decreasing amounts of 7 and an increasing proportion of 8, with the ratio of the two products showing a near-linear relationship between the van der Waals volume of the amino acid and alcohol production (Figure 2 and Table 1). This provides powerful evidence that the van der Waals volume of the C-loop residue 279 is of central importance for product distribution in DCS catalysis.

These results establish an essential role for Trp279 in DCS catalysis. In the WT enzyme, residue 279 prevents water access to 5, which is also stabilised through cation–π interactions with the indole ring. Trp279, hence, facilitates ring closure to 6 and proton loss to generate 7. Replacement of Trp279 with tyrosine had only a modest effect on the outcome, with a 7% increase in 8. However, if Trp279 was replaced with smaller, non-aromatic residues the quantity of 8 relative to 7 increased in a manner that depended on the van der Waals volume of the residue (Figure 2), irrespective of the hydrophobicity of the amino acid in position 279. Replacement with Gln, Asp or Glu only significantly affected the product distribution and not the catalytic efficiency of the enzymes, thus suggesting that changes to the size of residue 279 might open a channel that

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}$ [$\times 10^{-3} , \text{s}^{-1}$]</th>
<th>$K_m$ [$\mu$M]</th>
<th>$\Delta$ (%)</th>
<th>$%$</th>
<th>$k_{cat}/K_m$ [$\times 10^{-3} , \text{mM}^{-1} , \text{s}^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT-DCS</td>
<td>1.00 $\pm$ 0.4</td>
<td>3.20 $\pm$ 0.02</td>
<td>98</td>
<td>2</td>
<td>3.1 $\pm$ 0.2</td>
</tr>
<tr>
<td>DCS-His$_6$</td>
<td>1.27 $\pm$ 0.005</td>
<td>0.85 $\pm$ 0.11</td>
<td>90</td>
<td>10</td>
<td>2.17 $\pm$ 0.4</td>
</tr>
<tr>
<td>DCS-W279E</td>
<td>0.59 $\pm$ 0.01</td>
<td>1.45 $\pm$ 0.09</td>
<td>50</td>
<td>50</td>
<td>0.41 $\pm$ 0.04</td>
</tr>
<tr>
<td>DCS-W279Q</td>
<td>3.00 $\pm$ 0.02</td>
<td>8.00 $\pm$ 4.00</td>
<td>60</td>
<td>40</td>
<td>0.25 $\pm$ 0.13</td>
</tr>
<tr>
<td>DCS-W279D</td>
<td>0.67 $\pm$ 0.01</td>
<td>9.24 $\pm$ 3.00</td>
<td>40</td>
<td>60</td>
<td>0.07 $\pm$ 0.02</td>
</tr>
<tr>
<td>DCS-W279L</td>
<td>3.80 $\pm$ 0.09</td>
<td>9.45 $\pm$ 1.10</td>
<td>68</td>
<td>32</td>
<td>0.40 $\pm$ 0.05</td>
</tr>
<tr>
<td>DCS-W279M</td>
<td>1.09 $\pm$ 0.04</td>
<td>1.90 $\pm$ 0.20</td>
<td>65</td>
<td>35</td>
<td>0.57 $\pm$ 0.06</td>
</tr>
<tr>
<td>DCS-W279A</td>
<td>3.12 $\pm$ 0.01</td>
<td>2.23 $\pm$ 0.51</td>
<td>11</td>
<td>89</td>
<td>1.40 $\pm$ 0.32</td>
</tr>
<tr>
<td>DCS-W279Y</td>
<td>0.73 $\pm$ 0.01</td>
<td>2.48 $\pm$ 0.70</td>
<td>83</td>
<td>17</td>
<td>0.17 $\pm$ 0.01</td>
</tr>
</tbody>
</table>

[a] Percentage of total products.
Figure 2. Histogram of the percentage distributions of products 7 (●) and 8 (■) generated by DCS-His₆ and its mutants versus the van der Waals volume²² of residue 279 (●).

allowed water access to the active site of DCS.²¹ The size of the channel appears to depend on the size of residue 279, so that small residues allow for the generation of larger amounts of 8. Inspection of the X-ray crystal structure of DCS²⁰ reveals the G1/G2 helix break motif first identified by Pandita et al. (Figure 1).²⁰ This motif was missed in an earlier report in which a homology model for DCS based on epi-aristolochene synthase was used.²² N403 and L405 sit at either end of the helix-break motif, which suggests that the G1/G2 helix break motif is important for essential loop movements of terpene synthases, including those found in plants.¹⁹,²¹ Aside from W279, N403 and L405, the active site of DCS is highly robust to site-directed mutagenesis, in that changes to G276, I130, T407, C408, G409, L413, E455 and M523—all of which are located in or around the active site—do not lead to the generation of products other than 7.²⁶ This is in stark contrast to many bacterial and fungal terpene synthases for which changes to the active-site composition often lead to alternative products.³⁸,²⁷–³⁸ Plant terpene synthases appear to possess a robust architecture that, in most cases, ensures product fidelity. There are, however, a few hot spots where mutation alters the reactivity dramatically and, in the case of DCS, these are found in regions that control the movements of loops involved in closure of the active site prior to formation of the Michaelis complex.³⁸,⁹,¹⁹,²¹ Specifically, in addition to the G1/G2 helix mutants reported previously,²¹ Trp279 mediates a loop movement to ensure that proton loss occurs in the final step of DCS catalysis, as opposed to allowing water to capture 5 to generate 8. Because W279 is located directly across the active site from Y410, these two residues may form a favourable hydrophobic contact with the substrate as it folds into its reactive conformation with the active site. The aromatic nature of W279 does not appear to determine product distribution because replacement with even negatively charged aspartate or glutamate, or neutral alkyl residues, such as Leu or Ala, do not alter the reaction products; they only increase the proportion of water captured in direct proportion to the van der Waals volume of the side chain. Subsequently, the diphosphate group may act as a general base for the final deprotonation step (Figure 3). If W279 is replaced with a smaller, non-aromatic residue this process may be perturbed; C1 and C6 of 1 (Scheme 1 and Figure 3) are too far apart to facilitate the final ring closure and extra space in the active site opens a pore,²¹ whereby water can enter the active site and 5 is quenched (Figure 3). It is also notable that alteration of the C terminus through the addition of a hexahistidine tag led to the production of significant quantities of 8; this is consistent with a precise series of loop movements that effect closure of the active site of DCS. As mentioned above, water molecules have been observed in the crystal structures of several terpene synthases, in both open and closed conformations.²² These water molecules, however, do not take part in reactions and simply cushion the substrate in the active site.²¹ It is perhaps surprising that water molecules remain tightly bound in an unreactive state, even in mutant enzymes in which the active site has been altered. In the case of DCS, we have never observed nucleophilic capture of the bicyclic cadinonyl cation (6; Scheme 1).

Experimental Section

Introduction of C-terminal hexahistidine tag into DCS: The gene encoding WT-DCS was available in a pET21d vector from previous work.²⁶,²⁸ A single nucleotide deletion was required to bring the His₆ coding sequence of pET21d in frame with the DCS coding sequence. A Quickchange site-directed mutagenesis kit was used to introduce the desired deletion, according to the manufacturer’s instructions. PCR primers were as follows: 5'-GAACC AATTG CACTT GAGGA TCCGA ATTC-3' and 5'-GAATT CGGAT CCTCA AGTGC AATTG GTTC-3'. Plasmids were transformed into E. coli XL1 Blue and then purified from overnight cultures (lysogeny broth (LB) medium (10 mL) containing ampicillin (100 µg/mL)) by using the miniprep kit, as described by the manufacturer. Deletion was confirmed by DNA sequencing.
Expression and purification: DCS-His<sub>6</sub> was produced in *E. coli* BL21(DE3) cells that harboured the cDNA for DCS-His<sub>6</sub> under control of the T7 promoter. *E. coli* BL21(DE3) cells were gently defrosted on ice before plasmid (1 μL; 60 ng μL<sup>−1</sup>) was added to the cell suspension. The resulting mixture was stored on ice (30 min), heat-shocked in a water bath (42 °C, 30–35 s) and then returned to the ice (2 min). LB medium (1 mL) was added and the solution was incubated for 1 h at 37 °C with shaking (150 rpm). The cells were harvested by centrifugation (1 min, 3300 × g), resuspended in a minimum amount of LB medium and spread on an agar plate containing ampicillin (100 μg mL<sup>−1</sup>). The plate was then incubated overnight at 37 °C. A single colony from the agar plate harbouring the transformed cells was used to inoculate LB medium (100 mL) containing ampicillin (100 μg mL<sup>−1</sup>) and the culture was incubated at 37 °C with shaking (150 rpm) overnight. The overnight culture (10 mL) was transferred to each of 6 × 500 mL of LB medium containing the same concentration of ampicillin as before. Cells were incubated at 37 °C with shaking (150 rpm). When an optical density (OD<sub>600</sub>) of 0.6 was reached, isopropyl β-D-thiogalactopyranoside (IPTG) was added (0.5 mM final concentration) and the cultures were incubated for 24 h with shaking (250 rpm), at 20 °C. Cells were harvested by centrifugation at 5 °C (4200 g, 10 min). The supernatant solution was discarded and the pellets were stored at −20 °C.

Pellets were allowed to thaw at 5 °C and resuspended in cell lysis buffer (50 mM; 20 mM Tris-Base, 5 mM β-mercaptoethanol (βME), pH 8) and stirred gently for 1 h at 0 °C. Cells were then disrupted by sonication at 5 °C (40% amplitude for 3 min with 5 s on/10 s off cycles) and the resulting suspension was centrifuged at 5 °C (17 000 g, 30 min). SDS-PAGE analysis showed that protein was in the soluble fraction and the pellets were discarded. The supernatant solution was then loaded onto a 2 cm Amintranz nitrilotriacetic acid (NTA) Ni<sup>2+</sup> column (Expedeon, Over, UK) and eluted under gravity-controlled drip flow. After 40 min, the column was washed with four column volumes (CV) of binding buffer (Tris·HCl 100 mM, βME 5 mM, NaCl 500 mM, imidazole 5 mM, pH 8). The column was then washed with a gradient of imidazole (from 5 to 300 mM, 20 CV) in binding buffer. DCS-His<sub>6</sub> eluted in the range 60–100 mM imidazole; column fractions were analysed by SDS-PAGE. The fractions containing pure protein corresponding to a molecular weight of 64 000 (DCS-His<sub>6</sub>) were pooled, dialysed overnight (10 mM Tris-Base, 5 mM βME, pH 7.5; molecular weight cutoff (MWCO) 30000) and then concentrated to a final volume of about 5 mL (AMICON)
system, YM 30). The solution was aliquoted and stored at 0°C. The concentration of protein was estimated by using the method of Bradford.[24]

Site-directed mutagenesis of recombinant DCS-His and mutagenic primers is described in the Supporting Information. The expression and purification of mutant DCS-His enzymes was identical to that described for the WT.

Analytical incubations of DCS-His and mutants with 1: Compound 1 (25 μL, 10 mM) was added to assay buffer (250 μL; 20 mM Tris, 5 mM βME, 10 mM MgCl₂ at pH 7.5) followed by addition of enzyme (100 μL, 40 μM). The aqueous solution was overlaid with HPLC-grade pentane (0.5 mL) and the resulting mixture was incubated with gentle agitation (18–24 h) at 25°C. The incubations were repeated without enzyme as negative controls. The pentane extracts were then analysed by GC-MS as described in the Supporting Information.

Steady-state kinetics of DCS-His and mutants: Kinetic assays were performed according to the standard, linear range, microassay procedure previously developed for DCS (see the Supporting Information).[9,10]

Acknowledgements

We thank Dr Robert Jenkins and Simon Waller for assistance with GC-MS experiments. This work was supported by the UK’s Biotechnology and Biological Sciences Research Council (BBSRC) through grants (BB/H01683X/1, BB/M022463/1, BB/N012526/1). We thank Cardiff University for the provision of a postgraduate studentship to M.L.

Keywords: cations  •  enzymes  •  natural products  •  reaction mechanisms  •  site-directed mutagenesis