Alternative Active Site Confinement by Enforcing Substrate Pre-Organization in Cyclases

Alternatives *Confinement* in der aktiven Tasche durch forcierte Substrat Prä-Organisation in Zyklasen

Von der Fakultät 4: Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Kristina Schell

aus Nienburg/Weser

Hauptberichter: Prof. Dr. Bernhard Hauer
Mitberichter: Prof. Dr.-Ing. Ralf Takors
Vorsitzender: Prof. Dr. Arnd G. Heyer

Tag der mündlichen Prüfung: 08.03.2024

Institut für Biochemie und Technische Biochemie der Universität Stuttgart

Abteilung für Technische Biochemie

-2023-

The presented work was developed at the suggestion and under the supervision of Prof. Dr. Bernhard Hauer from August 2019 to February 2023 at the Institute of Biochemistry and Technical Biochemistry, Department of Technical Biochemistry, at the University of Stuttgart. Parts of this work were published in "Alternative Active Site Confinement in Squalene–Hopene Cyclase Enforces Substrate Pre-organization for Cyclization".[1]

Declaration of authorship

I hereby certify that the dissertation entitled "*Alternative Active Site Confinement by Enforcing Substrate Pre-Organization in Cyclases*" is entirely my own work except where otherwise indicated. Passages and ideas from other sources have been clearly indicated.

Erklärung über die Eigenständigkeit der Dissertation

Ich versichere, dass ich die vorliegende Arbeit mit dem Titel „*Alternatives Confinement in der aktiven Tasche durch forcierte Substrat Prä-Organisation in Zyklasen*" selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe; aus fremden Quellen entnommene Passagen und Gedanken sind als solche kenntlich gemacht.

Köln, 28.08.2023

Kristina Schell

Acknowledgements

» Above all, don't fear difficult moments. The best comes from them «

Rita Levi-Montalcini (1909-2012)

Table of Contents

Abbreviations

*Aac*SHC	Squalene hopene cyclase from *Alicyclobacillus acidocaldarius*
*Apa*SHC	Squalene hopene cyclase from *Acetobacter pasteurianum*
*Ath*LCY-B	lycopene β-cyclase from *Arabidopsis thaliana*
*Ath*LCY-E	lycopene ε-cyclase from *Arabidopsis thaliana*
Amp	Ampicillin
AU	Arbitrary units
Bp	Base pairs
*Can*LCY-B	Lycopene β-cyclase from *Capsicum annuum*
CD	(2-hydroxypropyl)-β-cyclodextrin
CDW	Cell dry weight
COSY	Correlated Spectroscopy
C_V	Coefficient of variation
CV	Column volume
CWW	Cell wet weight
D.f.	Degrees of freedom
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
DpnI	Restriction enzyme from *Diplococcus pneumoniae* G41
DWP	Deep well plate
E. coli	*Escherichia coli*
FPP	Farnesyl pyrophosphate
Fwd	forward
GA	Geranyl acetone
GC	Gas chromatography
HIV	Human immunodeficiency virus
HMBC	Heteronuclear Multiple Bond Correlation
HSQC	Heteronuclear Single Quantum Coherence
IEX	Ion exchange chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside

KSCN	Potassium thiocyanate
LB medium	Lysogeny broth medium
LCY	Lycopene cyclase
LCY-B	β- lycopene cyclase
LCY-E	ε- lycopene cyclase
*Mca*SHC	Squalene hopene cyclase from *Methylococcus cupsulutus*
MD	Molecular dynamics
MS	Mass spectrometry
MSA	Multiple sequence alignment
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser Enhancement Spectroscopy
PCR	Polymer chain reaction
Rev	Reverse
RMSD	Root-mean-square deviation
Rpm	Rounds per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SHC	Squalene hopene cyclase
TB medium	Terrific broth medium
TBDMSCl	*tert*-butyldimethylsilyl chloride
T-DAB	Terrific broth defined autoinduction medium
THF	Tetrahydrofuran
TKO	Triple knockout
TMS	Tetramethylsilane
Tris	Tris-(hydroxymethyl)-aminomethan
TS	Terpene synthase
WT	Wild type
*Zmo*SHC1	Squalene hopene cyclase from *Zymomonas mobilis*

Abstract

Confinement of an enzyme's active site is critical to the efficiency of chemical reactions and has been recognized as an important tool for catalysis. Confined active sites facilitate the pre-organization of substrates and intermediates to control the reaction course, protect against premature quenching and provide unique products. The catalytic center activates the substrate, and its activity can be enhanced by residues surrounding the substrate in the active site, changing the local catalyst geometry, and maintaining a constrained structural and/or electronic configuration of the catalytic center. These properties are characteristic of confinement, resulting in the generation of proximity between the substrate and the catalytic center, as well as complementary binding of the substrate into the active site. Effectively, this accelerates the reaction, controls the progress of the reaction, and positions the substrate in a productive conformation. The reaction course is selectively controlled by the stabilization of intermediates and by the interaction of electron-rich residues with electron-poor molecules and vice versa. Despite these benefits, a strongly confined active site is inherently limited to compounds that resemble the native substrate, with only small deviations tolerated. This restricts the applications for new reactivities and prevents broad substrate scopes. In this work, analysis of the enzyme structure in combination with iterative saturation mutagenesis were employed for the development of biocatalysts with alternative confinement and productive substrate pre-organization. To unlock the potential of confined Brønsted acid catalysts this approach was applied on terpene synthases. These enzymes form several carbocations as transition states and intermediates, which can be selectively converted by confinement of the active site. Exploiting the potential of terpene synthases to convert modified terpene scaffolds could provide interesting building blocks with branched isoprene/terpene motifs. In addition, the control of the reaction progression to specific products rather than a mixture of products could be targeted by stabilizing carbocation intermediates or transition states.

The present work demonstrates a structure-guided strategy to create an alternative confinement in the squalene hopene cyclase from *Alicyclobacillus acidocaldarius* (*Aac*SHC). This strategy aims to create proximity between substrate and catalytic center and complementarity between substrate and active site to obtain productive pre-organization of the substrate.

This may allow the conversion of geranylacetone (GA), farnesol and farnesylacetone analogs as substrates with modified isoprene patterns. Among different rational and semi-rational approaches only variant G600M (VD1), in which the substrate tunnel was modified, yielded starting activity. Structural analysis of VD1 led to the identification of a bottleneck in the tunnel. This resulted presumably in steric interactions and proximity by decreasing average distances between the double bond of the substrate's terminal isoprene unit and the catalytic center. Furthermore, a lower fluctuation of these distances around this mean value was observed in VD1 compared to the wild type (WT). These observations support the hypothesis of improved substrate pre-organization and confirm the creation of proximity between the substrate and the catalytic center. The development of a screening method and optimization of reaction conditions facilitated iterative saturation mutagenesis to investigate the evolvability and the potential of the approach. The positions for saturation mutagenesis targeted the shape complementarity of the active site to the GA analog dihydropseudoirone, and the finally developed variant (VD5) showed an 1174-fold increase of the total turnover number and 111-fold increase in catalytic efficiency compared to the WT. Creation of alternative active site confinement demonstrates evolvability and great potential to overcome limitations in the engineering of biocatalysts and allowed the generation of novel building blocks in preparative mg scale.

Limitations in the generation of alternative confinement were approached by using lycopene cyclases. These catalysts convert linear lycopene to carotenes under physiological conditions and were mainly studied for the conversion of pseudoionones in this work. The latter substrate could not be converted by *Aac*SHC through generation of alternative confinement. Three different lycopene cyclases were tested, of which *Can*LCY-B showed immediate activity in converting pseudoionone to a monocyclic product. *Ath*LCY-B and *Ath*LCY-E initially showed no conversion of selected terpenes. After optimizing the reaction conditions, a multiple sequence alignment (MSA) was performed to identify non-conserved positions around the catalytic acid of LCY-B. It was hypothesized that these amino acids influence the confinement and the pre-organization of the substrate. Saturation mutagenesis at the identified positions improved β-Ionone formation by 4-fold and conversion by 5% with variant V335L compared to the WT. The applicability of this MSA-based alternative confinement strategy for engineering of further lycopene cyclases was demonstrated by using saturation mutagenesis of the

Abstract

respective residues in *Ath*LCY-E. Under optimized conditions α-Ionone product formation increased 4.5-fold with the best performing variant *Ath*LCY-E S359F compared to WT *Ath*LCY-E. Application of lycopene cyclases to complement activities with challenging steric interactions demonstrated the successful expansion of the diversity for protonation reactions by biocatalysts and the successful application of an MSA-based approach to generate alternative confinement.

To further investigate the generation of alternative confinement and expand the toolbox of Brønsted acid catalysis, the acid isomerization of monoterpenes catalyzed by squalene hopene cyclases (SHCs) was investigated. To access selective product formation with monoterpenes a strategy based on cation stabilization was applied to overcome the challenging pre-organization of the cyclic C10 compounds. The focus was on aromatic residues with high electron density and residues surrounding the carbocation to direct the reaction course toward a single monoterpene product. In an initial screening, four monoterpenes were converted by *Aac*SHC, resulting in complex product mixtures. Of these, one monocyclic and one bicyclic substrate were selected for further engineering. The goal here was to increase the formation of terpinen-4-ol, a hydrated monoterpene. In addition, limonene that was not converted by *Aac*SHC was tested. Optimization of reaction conditions and semi-rational engineering to stabilize the carbocation intermediate produced improved variants in terms of selectivity and terpinen-4-ol formation. Saturation mutagenesis of hydrophobic amino acids that interact with the docked substrate and surrounding residues resulted in variants VT3 and VS2, which had the best selectivity and the highest measured terpinene-4-ol formation, respectively. VT3 converted monocyclic terpinolene with a selectivity of 64% and with a 3.4-fold increase in total turnover number (TTN) compared to the WT. The highest terpinene-4-ol formation of 219 µM and a 2-fold increase in TTN compared to the WT was measured with the bicyclic compound sabinene. Features, such as bulkier residues at position 600, found in VT3 and VS2 are likely responsible for generation of alternative confinement by the positioning of aromatic residues, which stabilize cationic intermediates along the reaction trajectory towards terpinene-4-ol formation.

Creating alternative confinement shows great potential for overcoming limitations in biocatalyst engineering. Interesting building blocks were generated and new reactivities with improved selectivity in protonation reactions have been discovered. Moreover, this strategy could be used for predicting potential hot spots in enzyme engineering

campaigns and data-driven predictions of enzyme functions to decipher the catalytic potential of enzyme scaffolds.

Zusammenfassung

Confinement der aktiven Tasche eines Enzyms ist entscheidend für die Effizienz chemischer Reaktionen und wurde als wichtiges Instrument für die Katalyse anerkannt. *Confinement* der aktiven Tasche begünstigt die Prä-Organisation von Substraten und Zwischenprodukten, um den Reaktionsverlauf zu steuern, vor vorzeitigem Reaktionsabbruch zu schützen und einzigartige Reaktionsprodukte zu gewinnen. Das katalytische Zentrum aktiviert das Substrat und dessen Aktivität kann erhöht werden, durch die das Substrat umgebenden Reste im aktiven Zentrum, Veränderungen der lokalen Katalysatorgeometrie und das Aufrechterhalten einer begrenzenden strukturellen und/oder elektronischen Konfiguration des katalytischen Zentrums. Diese Eigenschaften kennzeichnen *Confinement* und führen dazu, dass Nähe des Substrats zum katalytischen Zentrum generiert und das Substrat komplementär zu der aktiven Tasche gebunden wird. Daraus resultiert die Beschleunigung der Reaktion, die Steuerung des Reaktionsverlaufs und das Positionieren des Substrats in einer produktiven Konformation. Der Reaktionsverlauf wird selektiv gesteuert, indem Zwischenprodukte durch *Confinement* der aktiven Tasche stabilisiert und elektronenreiche Reste mit elektronenarmen Molekülen interagieren und vice versa. Trotz dieser Vorteile ist *Confinement* der aktiven Tasche jedoch von Natur aus auf Verbindungen, die dem nativen Substrat ähneln, limitiert und nur geringfügige Abweichungen werden toleriert. Dies schränkt die Anwendungsmöglichkeiten der Katalysatoren im Hinblick auf neue Reaktivitäten ein und verhindert ein breites Substratspektrum. In dieser Arbeit wurde die Analyse der Enzymstruktur in Kombination mit iterativer Sättigungsmutagenese für die Entwicklung von Biokatalysatoren mit alternativem *Confinement* und produktiver Substrat Prä-Organisation eingesetzt. Um das Potenzial von *Confinement* in einem Brønstedt-Säure-Katalysator zu erschließen, haben wir Terpensynthasen verwendet. Diese Enzyme bilden mehrere Carbokationen als Übergangszustände und Zwischenprodukte, die durch *Confinement* der aktiven Tasche selektiv umgewandelt werden können. Die Nutzung des Potenzials von Terpensynthasen zur Umwandlung modifizierter Terpengerüste könnte interessante Bausteine mit verzweigten Isopren-/Terpenmotiven liefern. Darüber hinaus könnte der Reaktionsverlauf durch die Stabilisierung von Carbokation-Zwischenstufen oder Übergangszuständen auf bestimmte Produkte gerichtet werden, statt zu einem Produktgemisch zu führen. Die vorliegende Studie demonstriert eine strukturgeleitete Strategie zur Schaffung von alternativem

Confinement in der Squalen-Hopen Zyklase aus *Alicyclobacillus acidocaldarius* (*Aac*SHC). Diese Strategie zielt darauf ab Nähe zwischen Substrat und katalytischem Zentrum und Komplementarität zwischen Substrat und aktiver Tasche herzustellen, um eine produktive Prä-Organisation zu erhalten. Dies könnte die Umsetzung von Geranylaceton-(GA), Farnesol- und Farnesylaceton-Analoga als Substrate mit modifizierten Isoprenmustern ermöglichen. Von verschiedenen rationalen und semi-rationalen Struktur-basierten Ansätzen führte nur die G600M-Variante (VD1), die den Substrattunnel beeinflusst, zu Startaktivität. Die ausführliche Strukturanalyse der VD1 Variante führte zur Identifizierung einer Verengung im Substrattunnel. Diese führte wahrscheinlich zu sterischen Wechselwirkungen mit den Substraten und Nähe zum katalytischen Zentrum durch Verringerung des durchschnittlichen Abstandes zu der Doppelbindung der terminalen Isopreneinheit des Substrats. Des Weiteren wurde eine geringere Fluktuation dieser Abstände um diesen Mittelwert in VD1 im Vergleich zum Wildtyp (WT) beobachtet. Diese Beobachtungen unterstützen die Hypothese der verbesserten Prä-Organisation des Substrates und bestätigen die Schaffung von Nähe des Substrats zum katalytischen Zentrum. Die Entwicklung einer Screening-Methode und die Optimierung der Reaktionsbedingungen ermöglichten die Durchführung der iterativen Sättigungsmutagenese, um die Evolvierbarkeit und das Potenzial dieses Ansatzes zu untersuchen. Die Positionen für die Sättigungsmutagenese zielten auf die Formkomplementarität der aktiven Tasche zum GA-Analogon Dihydropseudoiron ab und die entwickelte finale Variante (VD5) zeigte eine 1174-fache Steigerung des Gesamtumsatzes und eine 111-fache Steigerung der katalytischen Effizienz im Vergleich zum Wildtyp. Dieser Ansatz zur Schaffung von alternativem *Confinement* der aktiven Tasche zeigt Evolvierbarkeit und großes Potenzial zur Überwindung von Limitationen bei der Entwicklung von Biokatalysatoren und ermöglichte die Herstellung neuartiger Bausteine im präparativen mg-Maßstab.

Die Anwendbarkeit dieser Methode wurde des Weiteren mit Lycopin Zyklasen validiert. Diese alternativen Katalysatoren setzen unter physiologischen Bedingungen lineares Lycopin zu Carotenen um und wurden in dieser Arbeit vor allem auf die Umsetzung von Pseudoiononen untersucht. Für dieses Substrat war die Erzeugung von alternativem *Confinement* zur Aktivitätssteigerung mit *Aac*SHC limitiert. Drei verschiedene Lycopin Zyklasen wurden getestet von denen *Can*LCY-B sofort Aktivität für die Umsetzung von Pseudoionon zu monozyklisiertem Produkt zeigte. *Ath*LCY-B und *Ath*LCY-E zeigten in

ersten Versuchen keine Umsetzung ausgewählter Terpene. Nach Optimierung der Reaktionsbedingungen wurde ein *Multiple-Sequence-Alignment* (MSA) durchgeführt, um nicht konservierte Positionen, die die katalytische Säure von LCY-B umgeben, zu identifizieren. Es wurde die Hypothese aufgestellt, dass diese Aminosäuren den Einschluss und damit die Prä-Organisation des Substrats beeinflussen. Sättigungsmutagenese an den identifizierten Positionen verbesserte die β-Ionon Bildung um das 4-fache und dem Umsatz um 5 % mit der Variante *Can*LCY-B V335L im Vergleich zum WT. Die Anwendbarkeit dieser MSA-basierten alternativen *Confinement*-Strategie für das Engineering von weiteren Lycopin Zyklasen wurde durch Sättigungsmutagenese der entsprechenden Reste in *Ath*LCY-E demonstriert. Unter optimierten Bedingungen war die Bildung des α-Ionon-Produkts bei der besten Variante *Ath*LCY-E S359F, im Vergleich zum WT, um das 4,5-fache erhöht. Die Anwendung von Lycopin Zyklasen zur Komplementierung der Aktivitäten mit schwierigen sterischen Wechselwirkungen zeigte die erfolgreiche Erweiterung der Vielfalt für Protonierungsreaktionen durch Biokatalysatoren und die erfolgreiche Anwendung des MSA-basierten Ansatzes zur Erzeugung von alternativem *Confinement*.

Um die Erzeugung von alternativem *Confinement* weiter zu untersuchen und den Werkzeugkasten der Brønsted-Säure-Katalyse zu erweitern, wurde die saure Isomerisierung von Monoterpenen, die durch Squalen-Hopen-Zyklasen (SHCs) katalysiert wird, untersucht. Um die selektive Produktbildung durch saure Isomerisierung von Monoterpenen zu erreichen, wurde eine Strategie angewendet, die auf der Stabilisierung von Kationen basiert, da die Prä-Organisation der zyklischen C10-Verbindungen schwer zu erreichen ist. Insbesondere konzentrierten wir uns auf aromatische Reste mit hoher Elektronendichte und Reste, die die Carbokationen umgeben, um den Reaktionsverlauf zu einem einzigen Monoterpen zu lenken. In einem ersten Screening wurden vier Monoterpene durch *Aac*SHC umgewandelt, wobei komplexe Produktmischungen entstanden. Davon wurden ein monozyklisches und ein bizyklisches Substrat für das weitere Engineering ausgewählt. Ziel war es dabei die Produktion von Terpinen-4-ol, einem hydratisierten Monoterpen, zu erhöhen. Zusätzlich wurde Limonene getestet, das nicht von *Aac*SHC umgesetzt wurde. Durch Optimierung der Reaktionsbedingungen und semi-rationales Engineering zur Stabilisierung des Carbokation-Intermediates, wurden verbesserte Varianten in Bezug auf Selektivität und Terpinen-4-ol-Bildung generiert. Sättigungsmutagenese von hydrophoben Aminosäuren,

die mit dem gedockten Substrat und den umgebenden Resten interagierten, führten zu den Varianten VT3 und VS2, die die beste Selektivität bzw. die höchste gemessene Terpinen-4-ol-Bildung aufwiesen. VT3 setzte monozyklisches Terpinolen mit einer Selektivität von 64 % und einer 3,4-fach erhöhten TTN im Vergleich zum WT um. Mit dem bizyklischem Sabinen wurde die höchste Terpinen-4-ol-Bildung von 219 µM und eine 2-fach erhöhte TTN im Vergleich zum WT gemessen. Strukturelle Merkmale wie voluminösere Reste an Position 600, die in VT3 und VS2 gefunden wurden, sind wahrscheinlich für die Positionierung von aromatischen Resten verantwortlich, die kationische Zwischenprodukte entlang des Reaktionsverlaufs stabilisieren.

Dieser Ansatz zur Schaffung alternativen *Confinements* des aktiven Zentrums birgt ein großes Potenzial für die Überwindung von Limitierungen bei der Entwicklung von Biokatalysatoren. Es wurden interessante Bausteine erzeugt und neue Reaktivitäten mit verbesserter Selektivität bei Protonierungsreaktionen entdeckt. Darüber hinaus könnte diese Strategie zur Vorhersage potenzieller Hotspots in Enzym-Engineering-Kampagnen und zur datengesteuerten Vorhersage von Enzymfunktionen verwendet werden, um das katalytische Potenzial von Enzymen zu entschlüsseln.

1. Introduction

1.1. Principles of catalysis and confinement

Catalysis is one of the key principles in chemistry. In the modern chemical industry catalysis is pivotal in the efficient use of natural resources and minimization of waste production for sustainable applications. It is used in medical, food, fragrance, solar cell, diagnostic and polymer industry to provide access to new chemical reactions and to accelerate existing reactions, making them profitable.[2]

In particular, reducing the overall activation energy by a regenerating catalyst to accelerate the chemical transformation opens novel reaction pathways in homogeneous, heterogeneous or biocatalysis.[3,4] The activation energy (E_a) is associated with the gas constant R and the absolute temperature T. The Arrhenius equation describes the velocity constant associated with E_a, R, T and a frequency factor (A) (Equation 1).

$$k = Ae^{-\frac{Ea}{RT}} \tag{1}$$

The action and principles of catalysts to accelerate the reactions are complex.[4,5] High activation energies in chemical reactions are often due to bond breaking. The formation of new bonds stabilizes the system on the product side. Catalysts form transient bonds with the reactant via covalent and noncovalent interactions. These lower the energy of the transition state by partially compensating for the energy required to break the old bond until the formation of the new bond ensues and the products are formed.[5] Notably the product is usually not the thermodynamic endpoint, but the catalyst selectively forms an intermediate which determines the end of catalysis.[5,6] This gives rise to reaction pathways that are selective for specific products and reactions that would not have been observed in spontaneous reactions without catalysts despite competing pathways with multiple products.[7–9] Inspired by biocatalysts and metal catalysts, establishment and application of catalysts for enantioselective reactions in organocatalysis was recently awarded with a Nobel Prize.[10] It was shown that organocatalysts modified by molecular engineering could be applied for a multitude of new type reactions and used for iminium catalysis, enamine catalysis and H-bonding catalysis.[6,11,12]

The catalysts activate the reactant in the active site and the coverage of the active site is dependent on e.g. the adsorption of the substrate or the binding of the substrate in heterogenous catalysts.[13] The optimum binding energy in the active site to perform

catalysis but circumvent product inhibition depends on, e.g., structural and/or electronic configuration of the active site and requires balance between constrictive binding and void space.[13] Shape selective catalysts, e.g., small molecule catalysts, enzymes and nanoporous catalysts select the substrate from a wide range of compounds, with multifold higher fidelity observed in enzymatic catalysis compared to homogenous and heterogeneous catalysis.[14–16] The active site activates the reactant, but the matrix of the active site outside of the first shell is of equal importance as it often enhances the activity of the ground state of the catalyst attacking groups. This is enabled by flexible alterations of the local catalyst geometry and holding a constrained structural and/or electronic configuration of the active site.[17] For example, electron density distribution changes in the catalyst favor the formation of new bonds by holding the molecules in proximity and specific relative orientation to the catalytic site.[3,18] These flexible adaptations of a catalysts structure along the reaction path are essential for lowering the activation energy.

Recently, the concept of confinement received greater attention in the context of small molecule organocatalysis and was identified as key feature of all types of selective catalysis.[3] Confinement describes the shaping of a catalyst's active site in a well-defined three-dimensional spatial arrangement and leads to restriction of small molecules motion in the active site to essentially influence the reactivity and selectivity of the catalyst.[3] Confined active sites control the conformation of substrates and transition states precisely and stabilize reactive chemical species to facilitate reaction rate acceleration and direct the reaction progress and termination.[18–20] Moreover, confined active sites shield substrate and intermediates in a buried cavity to protect against premature quenching by solvents and reduce byproduct formation.[8,9] In fact, a large and unusual selectivity compared to classical reactions in bulk is observed.[8,9] Two major features in a confined active site are fundamentally involved in the positive effects for catalysts, proximity and substrate specific secondary interactions in the active site.[21] The term proximity has also been used in the context of proximity channeling between active sites in previous studies. In this work, however, only the proximity of the substrate to the catalytic center is referred to when the term proximity is used. Proximity of the reactant to the active site results in reaction initiation and acceleration, while secondary substrate shape complementary interactions stabilize intermediates, transition states and affect substrate binding and substrate selectivity.[21,22] The ideal confined catalyst shapes the

active site cavity exclusively complementary to the transition state of the desired transformation by limiting the degrees of freedom of the interacting reactants.[3,23] It was shown that shape complementarity and precisely placed catalytic groups can be successfully harnessed to design and evolve artificial enzymes for Kemp elimination and Diels-Alder reaction. In these reactions chemical rate accelerations of 6×10^8-fold and 100-fold improved catalytic efficiency were afforded, respectively.[23,24] Proximity and secondary substrate specific interactions are based on geometrical constraints (shape and size selectivity), entropy effects, adsorption and desorption.[25] Moreover, they contribute in productive substrate binding in the active site to enforce the pre-organization of the substrate in a higher energy conformation and allow the correct reaction propagation to unique, selective products.[25–29] In particular, basic atomic properties change in a confined catalyst framework and as an effect increase the excitation energy, lower polarizability, influence π-π interactions and lead to electrostatic stabilization.[3,30]

1.1.2. Types of catalysis

The investigation of principles of catalysis and especially the concept of confinement is important to understand the influence of catalysts on activity and selectivity. Enzymes are naturally occurring and diverse confined catalysts according to the lock-and-key theory proposed by Emil Fisher in 1894 and the extension to the induced fit and the conformational selection theory.[31,32] The latter theories describe the mechanism of ligand binding coupled to conformational changes in enzymes. In one of the most complex one-step enzymatic reactions hopene or hopanol with five ring structures, 13 covalent bonds, and nine stereo centers are formed from the linear substrate squalene by SHC (Figure 1).[33] This assumes that the ligand is properly bound, and the conformational changes are accurately defined in the enzyme. Several reactions with remarkable activity and selectivity were shown to be catalyzed by enzymes and studies aimed to mimic these biocatalysts and demonstrated the application in selective cross-coupling, cycloaddition and isomerization reactions and the major influence of acid/base catalysis in heterogeneous- and homogenous confined catalysts.[34–37]

Similar to enzyme active sites, zeolites possess a well-defined pore structure which enables the catalyst to separate molecules by size and trap them in the active site.[38] In heterogeneous catalysis zeolites demonstrated the ability to catalyze a variety of transformations e.g. Beckmann rearrangements, Brønsted acid transformations and electrophilic aromatic substitutions.[39] In addition, the starting material lactide for

synthesis of renewable and biodegradable plastics was synthesized by a H-beta zeolite.[34] In the reaction without the confined catalyst with simple Brønsted acid catalysts the unselective oligomerization product was formed.[34] The confined reaction site of a Sn-beta zeolite was shown to catalyze the chemoselective oxidation of several ketones including the reaction of dihydrocarvone to the 7-membered lactone with a selectivity above 98% when combined with a Lewis acid (Figure 1).[40]

In homogenous catalysis the second coordination sphere results in a (chiral) pocket. This is a prerequisite for a catalyst. The confined active site exhibits additional features such as shape- and size-selectivity, prevention of side reactions, substrate pre-organization, and stabilization of reactive intermediates.[3,41] Multiple supramolecular catalysts including cyclodextrins and capsule like structures have been developed for a wide scope of reactions and supramolecular catalysts were recognized as artificial enzymes.[42,43] In this context the $[Ga_4L_6]^{12-}$ cage was designed with an anionic and hydrophobic pocket to preferentially incorporate organic and cationic molecules.[44] With these catalysts reaction rates could be accelerated by up to 1000-fold for aza-Cope rearrangements. The improvement was attributed to the effective pre-organization of the substrate and efficient stabilization of the cation.[26] In a carbonyl ene cyclization of citronellal with challenging stereochemistry and side reactivity the substrate was converted with 97% selectivity to isopulegol by the $[Ga_4L_6]^{12-}$ cage with good diastereoselectivity and differed significantly in the product outcome observed in the reaction in aqueous KH_2PO_4 (Figure 1).[45] In addition, a self-assembled octahedral cage was shown to catalyze the selective epoxidation of the diterpene tail protruding from the supramolecular cavity.[18]

1. Introduction

The confinement of the active site can also play a crucial role in increasing the overall reactivity of a given process by stabilizing a more reactive conformation of the chosen substrate. Zhang *et al.* demonstrated the intramolecular hydroarylation of simple olefins by a homogenous small molecule catalyst IDPi (Figure 1).[46] The high acidity of the IDPi catalysts and the pre-organization of the substrate are key requisites for the protonation. Gem-dimethyl substitution of substrates on the aliphatic chain bearing a nucleophilic functional group is known to enormously increase the rate of cyclization in any intramolecular cyclization.[36]
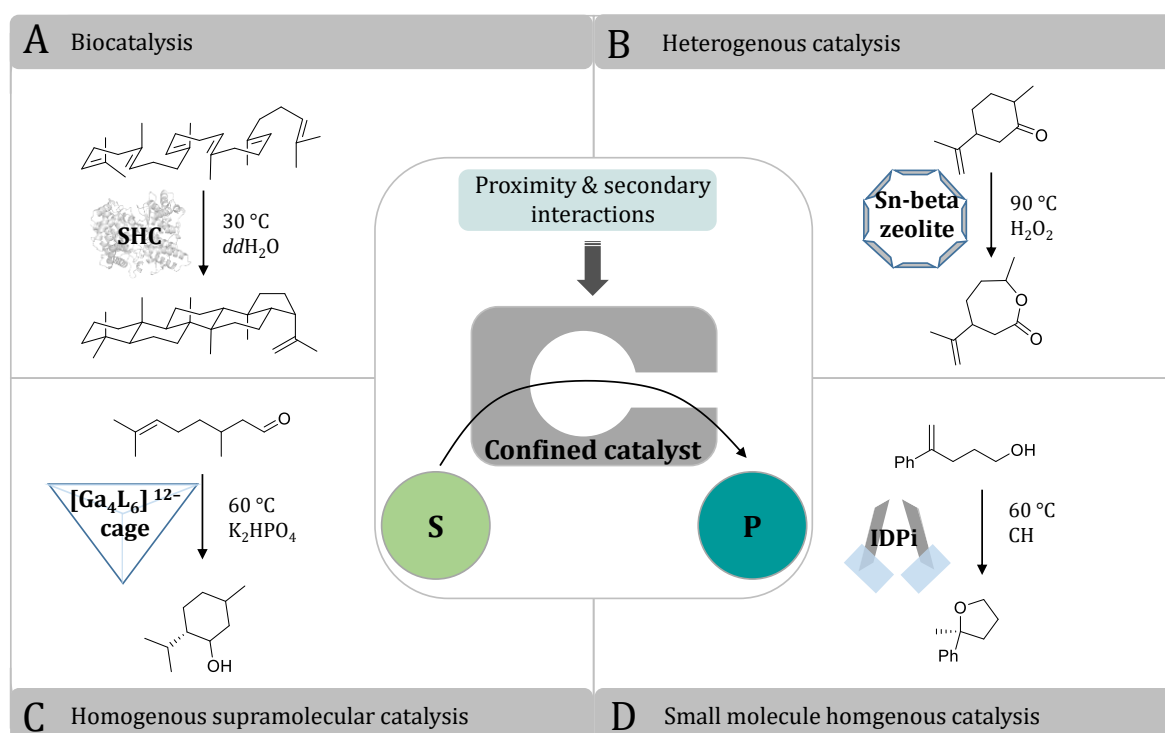
Figure 1. Confined catalysis via proximity and secondary interactions. (A) Polycyclization of squalene by SHC as example for selective reactions in biocatalysis.[85] (B) Heterogeneous catalysis of dihydrocarvone to the 7-membered lactone via Sn-beta zeolite.[40] Homogenous (C) Prins reaction with citronellal by a [Ga$_4$L$_6$]$^{12-}$ cage and (D) protonation of simple olefins and subsequent intramolecular cyclization catalyzed by IDPi.[36,45]

Current research aims at expanding and generalizing the application of catalysts for available substrate classes through the rational development of more active, but highly selective catalyst motifs.[47,48] Versatile bio- and organocatalysts with an outstanding range of transformations are Brønsted acid catalysts.[49,50] In particular, chiral Brønsted acid catalysis contributes significantly to the formation of chiral scaffolds of elaborate bioactive molecules and is a rapidly growing area with great potential when combined with confined catalysts such as in BINOL-derived phosphoric acid catalysts.[50]

1.2. Brønsted acid catalysis

In Brønsted acid catalysis the protonation of a functional group or formation of a electrophilic transition state enhances the reactivity towards nucleophiles thereby enabling unique versatility and a diversity of complex scaffolds.[50] The origin of this reaction diversity is the efficient activation of a large number of structures, which depend on the acidity of the catalyst and the alkalinity of the substrate.[50] Initially Brønsted acid catalysis was limited to hydrolysis reactions and the formation of esters and acetals, and asymmetric catalysis was scarcely explored.[51] However, in the past decades BINOL-derived phosphoric acids, which are esters of phosphoric acid with chiral backbones, were characterized and it was reported that a protonated substrate can be closely associated with its anion in a Mannich-type reaction. This can confer enantioselectivity to the reaction.[52,53] Since, Brønsted acid catalysts attracted much attention and asymmetric transformations of imines, carbonyls, alcohols, epoxides and unfunctionalized aliphatic substrates like alkenes and alkynes have been realized.[54–56] Nowadays Brønsted acid catalysis demonstrates versatility for a wide range of reactions and extensive possibilities for formation of new C-X and C-C bonds.[50,53,55] Subsequently, asymmetric/chiral Brønsted acid catalysis has emerged as one of the most prominent subfields of organocatalysis.

1.2.1. Chiral Brønsted acid catalysts

The recognition of the crucial relationship between the configuration of a chiral molecule and its biological activity initiated a tremendous interest in the efficient and atom economic enantioselective catalysis to perform asymmetric reactions. In the field of asymmetric Brønsted acid catalysis two types of asymmetric catalysts exist with regard to the reaction mechanism: A) general and B) specific Brønsted acid catalysts. General Brønsted acid catalysts act as hydrogen bond donors to activate compounds and stabilize the transition states (Figure 2).[57–60] Besides general Brønsted acid catalysis, small molecules possessing hydrogen bond donors catalyze an array of C–C and C–X bond-forming reactions.[59,61] Specific Brønsted acid catalysts activate the substrates by protonation. The reaction mechanism is linked to the acidity of the catalysts, which is weaker in general compared to specific catalysts (Figure 2).[57,60,62]

Fundamental research in the field of Brønsted acid catalysis elaborated the use of neutral compounds such as thiourea and diols (e.g. TADDOL and squaramide derivatives) in

general acid catalysis for Strecker reactions, hetero-Diels-Alder reactions and conjugate addition reactions, respectively.[56,63] Overall, several hydrogen bonding catalysts and reactions were reported for activation of electrophiles.[17,59] Nevertheless, the general acid catalysts are limited for broad applications, because of their weak acidity.[57,60] Only with the application of strong acids within specific acid catalysts the substrate scope could be extended thus since 2004, BINOL-derived phosphoric acids were used for Mannich reactions.[52,53] These catalysts form a chiral cavity due to their axial chirality.[52,53] The size of the cavity can be designed for different substrates by modification of the 3,3'-substituents. In the TRIP catalyst reported by List *et al.* placing bulky groups near the active site as 3,3'-substituents allowed the asymmetric transfer hydrogenation of imines.[64] Moreover, the success of the BINOL derived catalysts prompted the development of catalysts with alternative backbones to modify geometrical parameters of the active site. Selected examples are the bis-BINOL derived phosphoric acids and the Biphenol-derived phosphoric acids catalyzing the transfer hydrogenations of quinolines or the asymmetric C-H functionalization via an internal redox process, respectively.[65,66] In addition to chiral phosphoric acids, chiral dicarboxylic acids, chiral disulfonic acids, chiral sulfonimides and *N*-triflylphosphoramide have emerged as stronger Brønsted acids, and their applicability for synthetic synthesis has become widely accepted.[55,67–69] Multiple asymmetric reactions including transfer hydrogenations, reductive aminations, Friedel-Crafts reactions, Michael reactions, Prins reactions, Nazarov cyclizations and Diels-Alder reactions could be performed with these catalysts.[50,64,70–72] For the alkylation with weak basicity, the BINOL backbone imidodiphosphate with N-triflyl groups with high acidity and confined active site was designed.[46,73] The moderately strong or strong general and specific Brønsted acid catalysts entail the application for asymmetric catalysis with tunable reactivity and selectivity for a myriad of reactions.

1.2.2. Brønsted acid catalysis in Enzymes

The versatility of Brønsted acid catalysis for a wide range of reactions makes it significant for biosynthetic pathways.[74,75] By forming hydrogen bonds enzymes act as general Brønsted acid catalysts, enhancing the basicity of the substrate and stabilizing transition states during the formation/breaking of bonds to decrease the free energy of the activated complex and accelerate the reaction.[76–78] Similar to TADDOL, hydrolytic enzymes perform general Brønsted acid catalysis in e.g. hydrolases, lipases, proteases and esterases.[79] Acting as a general Brønsted acid catalyst, the serine protease uses an oxyanion hole to

form hydrogen bonds to position and activate the carbonyl group for the hydrolysis of the amide bonds of the substrate (Figure 2).[80] The carbonyl is activated via a serine associated with an aspartate and a histidine in a catalytic triad, forming a tetrahedral intermediate upon nucleophilic attack by the serine residue (Figure 2).[80] The amine then leaves the active site and a water molecule is activated to hydrolyze the acyl-enzyme intermediate.[80] Most often, the oxyanion hole is composed of amides or positively charged residues to stabilize the negative charge of the intermediate and allow positioning of the substrate when it is otherwise sterically hindered.[13] Another example for general Brønsted acid catalysis by enzymes is the hydrolysis of amide bonds catalyzed by Human immunodeficiency virus (HIV) protease. Here, a *syn*-oriented hydroxyl group of the aspartate activates the substrate and hydrogen bonding stabilizes the tetrahedral intermediate (Figure 2).[81] This general Brønsted acid catalysis is well recognized in enzymes, however due to the low acidity of proton donors in water only limited examples for specific catalysis exist for enzymes.[76] In 2014 Hammer *et al.* postulated specific Brønsted acid catalysis in squalene hopene cyclases, which demonstrated high acidity to enable olefin protonation in a confined microenvironment with high selectivity, avoiding potential side reactions.[82] These enzymes' catalytic acid is an active site aspartate coordinated by a specific amino acid network (Figure 2).[82] This aspartates' hydroxyl group is presumably *anti*-orientated and increases the acidity by 10^4-fold compared to *syn*-protons commonly observed in general Brønsted acid catalysts.[83,84]
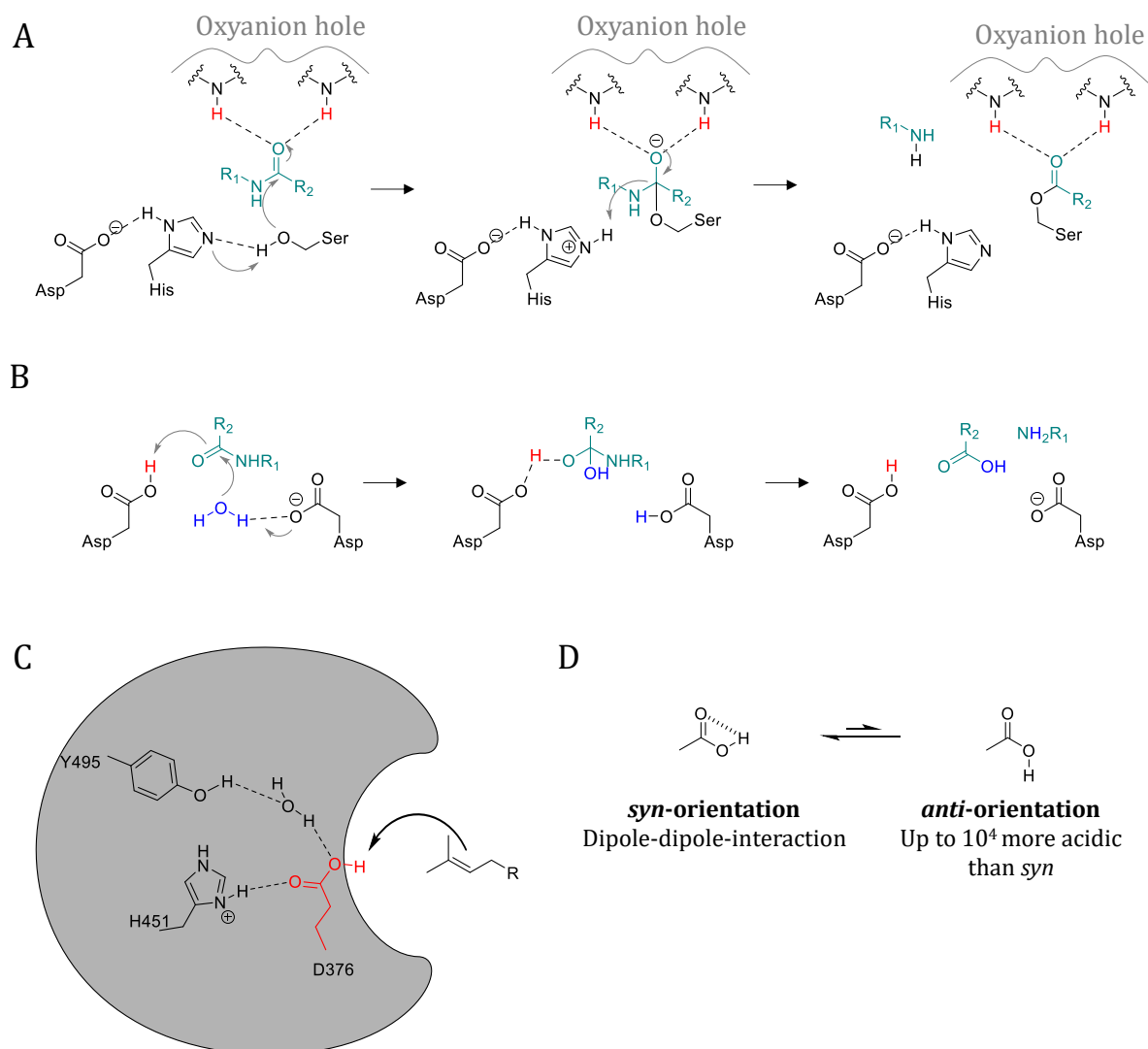
Figure 2. Chiral Brønsted acid catalysis with general and specific acid catalysts. (A) Hydrolysis of amide bonds via a two-step mechanism with a covalent enzyme-substrate complex. Two hydrogens of the protein backbone form the oxyanion hole, which positions and activates the carbonyl group of the substrate (teal) as a general Brønsted acid.[80] (B) Hydrolysis of an amide bond by HIV protease.[81] The substrate (teal) binds in the active site and a coordinated water molecule (blue) is activated for nucleophilic attack. The tetrahedral intermediate is stabilized by hydrogen bonding through an aspartic acid in *syn*-orientation of the general Brønsted acid. (C) Protonation of C-C bonds by *anti*-oriented D376 (red) in *Aac*SHC mediated by a hydrogen bond interaction with Y495 and H451 and (D) physiologically relevant *syn*- and *anti*-orientations of the hydroxyl group of carbonic acids in general and specific Brønsted acid catalysts, respectively.[83–85]

Overall, Brønsted acid catalysis in chemical or biological reactions is efficient for the activation of multiple structures. However, after the activation of the substrates, especially by specific Brønsted acid catalysts, the control of the formed carbocation and the course of the reaction is essential for chirality. In this context, investigation on the main properties of carbocations and their stabilization are essential.

1.3. Carbocation chemistry

Carbocations are generated from CH acidic compounds or alkyl halides and can be classified as carbenium ions, which have been defined as the classical trivalent, sp^2 hybridized carbocations and non-classical carbonium ions. The non-classical carbonium ions are species with tetra-, penta-, or even higher coordination carbon centers, which can have three-center bonding.[86] Generally, carbocations occur in nucleophilic substitution (S$_N$1) and elimination (E1) reactions, additions of electrophiles to double and triple bonds, electrophilic aromatic substitutions, rearrangements and polymerizations.[87–89] The observed versatility of reactions is caused by the simplicity of the protonation, rearrangement reactions and carbocations incorporating an unusually wide range of structures and bonding types including multiple bonding, aromaticity, strained rings, and nonclassical ions.[88–90]

Carbocations are very reactive intermediates and their stabilization is mandatory in order to facilitate the myriad of reactions observed.[91] The carbocation stabilization controls the product selectivity to prevent quenching and to lower the activation energy to accelerate the reaction.[92] In general, stability decreases from tertiary to primary carbocations, but is dependent on the substituents and can be stabilized by hyperconjugation and resonance (Figure 3).[92] Therefore, increasing the number of alkyl substituents on a carbocation leads to an increase in stability (Figure 3). Hyperconjugation stabilizes carbocations by electron delocalization via parallel overlap of the carbocation p-orbital with hybridized orbitals participating in σ-bonds (usually C-C or C-H bonds). Resonance stabilization does not involve an σ-orbital component, but only the p-orbital of the carbocation and the p-orbitals of the π-bonds in a conjugated π-system. Here the positive charge can be distributed over multiple carbons (Figure 3). Hence π-bonds allow the carbocation p-orbital to be part of a conjugated π-system, thereby

Figure 3. The three modes of carbocation stabilization via hyperconjugation and resonance stabilization.

stabilizing the carbocation.[89,93] In addition, carbocations can be stabilized by adjacent atoms bearing lone pairs by donation of electrons to the electron-poor carbon (Figure 3).

Physiologically relevant and representative carbocations are the allylic carbocations, tertiary carbocations and the secondary carbocations. Tertiary carbocations form the largest and most diverse group of carbocations stabilized by hyperconjugation with its adjacent C–C bonds and cation–π interactions with the nearby π-bonds as well as allylic carbocations.[94] Of the less stable secondary carbocations, only a few examples have been validated by computational calculations of minima on potential energy surfaces. They were documented as homoallylic carbocation or are stabilized by strong hyperconjugative interactions, e.g., in the pathway to pentalenene.[95] In addition, some hybrids of tertiary and secondary carbocations were observed, the so-called "proton sandwich".[94,95] The above carbocations represent minima on potential energy surfaces, however, transition state structures are not uncommon intermediates in natural carbocation rearrangements and can occur via asynchronous, concerted processes.[94,96]

Carbocations react with the nucleophiles in the following and the selectivity in the course of this reaction depends on the stabilization of the carbocation and methods used to control the selectivity of the nucleophilic attack. Here two different methods are differentiated i) chiral directing groups with an adjacent stereocenter blocking either the *re* or the *si* face of the carbocation which lead to geometric constraints in ring-forming reactions and ii) association of a chiral counterion with the carbocation. In this context, stereoselective delivery of a proton to an alkene or other group became useful in the development of new types of chiral Brønsted acids.[71,90,97,98] When considering the unstable/reactive transition states during the reaction and the nucleophilic attack, the environment of the compound is important for reactivity. In catalyst development, the confined active site or interactions with aromatic constituents were found to have a significant impact on reactivity and selectivity in carbocationic reactions.[92,99] A moiety at the active site can interact with the appropriate orbitals of a carbocation, leading to large geometric distortions that push and pull the carbocation along the continuum of possible structures and leads to stabilization of certain transition states for the reaction course.[92]

There is a great variety of carbocation structures and rearrangement reactions. Their stabilization is important for selective product formation. In natural biosynthesis from one compound several products can potentially be formed, however the described

principles for carbocation stabilization result in selective catalysis towards specific products. This is especially the case in the biosynthesis of terpenes where a variety of delocalization pathways with carbocations are possible. The pathway and the catalyst that facilitate this selectivity are described in the next chapters.

1.4. Terpene biosynthesis

Carbocations are important intermediates in the biosynthesis of terpenes and steroids. In the biochemical synthesis, they are steered towards the formation of various terpene natural products.[100,101] Terpene and terpenoid natural products with diverse and complex molecular architectures are produced in nature from simple precursors and their properties and functions are as diverse as their structure. They display well-known pharmaceutical properties, including anti-cancer, anti-viral, anti-bacterial and anti-inflammatory activity.[102–105] In fungi, plants, insects, amoebae and bacteria, terpenes mediate antagonistic and beneficial interactions for defense mechanisms and intraspecies communication.[103,105] In industry, terpenes find application in food, health, pharma and fragrance industry.[106–109] The common feature of terpenes is their isoprene backbone consisting of isoprene units with the formula $(C_5H_8)_n$ with naturally occurring modifications. In the first step of terpene biosynthesis, geranyl pyrophosphate is formed by the coupling of isopentenyl pyrophosphate and its regioisomer dimethylallyl pyrophosphate in the mevalonate pathway found in eukaryotes or the deoxyxylulose 5-phosphate pathway found in plants and bacteria.[110,111] Condensation with additional isopentenyl pyrophosphates yields C_{5n} isoprenoid pyrophosphates. Based on the number of isoprene units the compounds are classified into monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), triterpenes (C_{30}) and so on. [110–112] These naturally occurring terpenes follow thereby the native isoprene rule. This rule defines terpenoids to be alkene chains containing a characteristic number of linear, head-to-tail condensed, C5 isoprene units. Methyl-migration, rearrangements and proton and hydride transfers affect the diversity of terpenes, e.g., in the cyclization and rearrangement of farnesyl pyrophosphate (FPP) to aristolochene.[113–115] During cyclization of FPP to aristolochene two rings, three chiral centers, and two double bonds are formed regio- and stereospecifically within the active site of aristolochene synthases from *Penicillium roqueforti* and *Aspergillus terreus*.[115] The involved carbon atoms undergo changes in hybridization, configuration or bonding.[115] In particular, the C10-C11 double bond displaces the pyrophosphate leaving group in an S_N2-like reaction to provide a germacryl cation, followed by a proton loss from

C12 to generate germacrene A (Figure 4).[115] The intermediate was postulated to undergo protonation of the C6-C7 double bond and further cyclization to form the bicyclic eudesmane cation. Successive 1,2-hydride shift and methyl migration, followed by loss of hydrogen at C8, leads to the formation of aristolochene.[115]
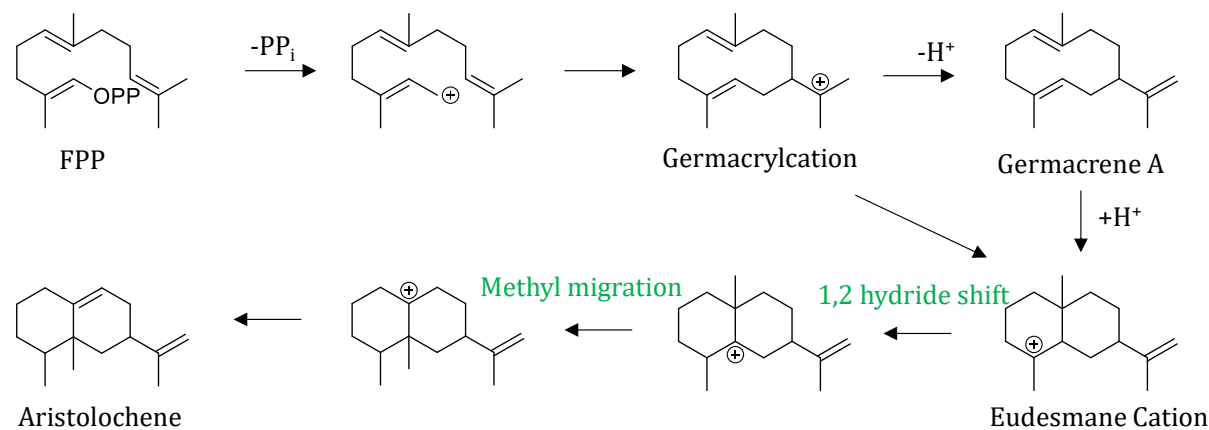
Figure 4. Mechanisms for the cyclization and rearrangement of FPP to aristolochene catalyzed by aristolochene synthases proceeding through germacryl and eudesmane cations.[115]

A   Type I TS
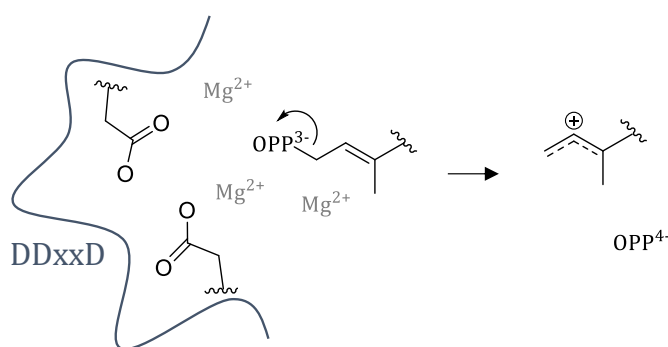
B   Type II TS
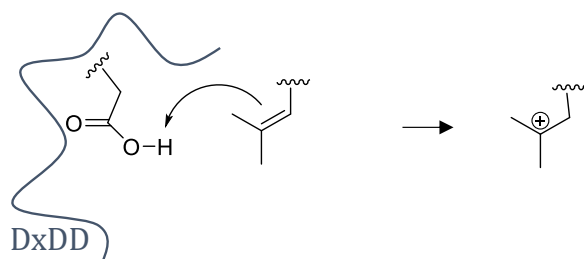
Figure 5. Formation of the first carbocation by (A) heterolytic cleavage mediated by magnesium ions as observed in type I terpene synthases. (B) Electrophilic activation through Brønsted catalysis define type II terpene synthases.

As represented in Figure 4 significant contribution in terpene versatility is observed by cyclization of linear hydrocarbon molecules mediated mostly by terpene synthases (TS) via two pathways that differ in the mechanism of activation of the substrate (Figure 4).[111] In type I TS the active site motif DDxxD is conserved and the reaction is initiated by the formation of an allylic cation via heterolytic cleavage of the hydro-carbon pyrophosphate bond (Figure 5).[114] In type II TS, the reaction is initiated via electrophilic activation at the prenyl end of the substrate by an aspartate from the conserved DxDD motif (Figure 5).[111]

Commonly terpene synthases bind their substrates in relevant conformations, facilitate carbocation generation, shield carbocations from premature quenching and selectively terminate the reaction. Selective stabilization of transitions states is obtained by electrostatic effects for the promotion of quantum mechanical tunneling to bind substrate conformations that resemble key transition state structures.[116,117] For example carbocation–π interactions with aromatic amino acid residues are important to stabilize positively charged species in oxidosqualene cyclases during the formation of lanosterol. These interactions were also shown to substantially lower the reaction energy barrier in mutational studies of tryptophan and phenylalanine residues in the active site of aristolochene synthetase.[118,119] In the reaction from (oxido-) squalene to lanosterol, lupeol and hopene/hopenol, hyperconjugation in triterpene cyclases guides the course of the reaction via multiple carbocations and bridged cyclopropane/carbonium ions within a polycyclization cascade.[120] Additionally, the same starting materials can be converted to different products by precisely controlling reactive carbocations through steric and electrostatic interactions in different monoterpene synthases.[121-123] In these reactions key intermediates determine the final structure of the product.[124-126] Carbocation formation and transition state stabilization by enzymes in terpene biosynthesis are fundamental catalytic principles that contribute to the diversity of terpenes and are important for diverse applications.

1.4.1. Plasticity and evolvability of terpene synthases

The diversity of terpenes is increased during the cyclization reaction in terpene biosynthesis through several new stereocenters formed by variation of substrate pre-organization. In solution, product selectivity, e.g. stereoselectivity and regioselectivity, of terpene cyclizations are hard to control due to premature chemical quenching of the cationic intermediates. In enzymes, initiation of the first carbocation, its propagation, and its quenching can be controlled and spatially separated within the active site.[127-129] A major factor for the diversity is the great plasticity and promiscuity of many TSs for different substrates and reaction trajectories thereby reflecting the significant evolutionary potential.

In the past decades, this plasticity was studied in type I and II TSs and revealed their great potential as biocatalysts for terpene biosynthesis. High promiscuity towards terpene substrates with different conformation, chain length and synthetic modifications were observed.[101,121,130,131] A germacene synthase for example was shown to cyclize FPP,

geranyl pyrophosphate and neryl pyrophosphate.[101] Further, two fungal sesquiterpene synthases were shown to convert FPP geometric isomers via opposite enantiomers of a cyclic carbocation intermediate that were rearranged to different sesquiterpene products.[101] In a δ-cadinene synthase product formation could be shifted towards the alcohol product instead of the bicyclic sesquiterpene δ-cadinene through single amino acid substitutions that enabled coordinated series of loop movements. This allowed bulk water to access to the final carbocation in the active site.[132,133] Similar product interchange from kolavenyl pyrophosphate and ent-copal-8-ol pyrophosphate was observed for class II diterpene synthases.[133] Based on an enzyme design strategy that considered residues in and around the active site, Greenhagen *et al.* demonstrated that the abietadiene synthase could be converted into a pimaradiene synthase and vice versa.[134] Further, β-bisabolene synthase, *E*-β-farnesene synthase, sibirene synthase, longifolene synthase and α-longipinene synthase were designed by recombination from plasticity residues and saturation mutagenesis of a γ-humulene synthase.[135]

For production of valuable products like taxol and artemisinin type I TS were used, frequently.[136,137] In a chemoenzymatic approach valuable germacrene analogues, possessing antimicrobial activity were prepared with a type I germacrene A and germacrene D synthase and modified FPP analogues.[138] These two enzymes demonstrate plasticity by conversion of a variety of modified FPPs to germacrene A and D analogues often with synthetically acceptable conversions. However, type I TSs require activation of substrates with pyrophosphate and are suitable for *in vivo* applications, and less practical for broader and preparative *in vitro* applications. With type II TS efficient production of (+)-ambrein, the most abundant component of the odor ambergris, was demonstrated.[109]

The described plasticity and promiscuity of TSs is essential for the diversity of terpenes and presents a great platform for practical applications. TSs were used and engineered to allow conversion of versatile substrates and formation of multiple products. They produce terpenes under mild conditions in water with highly frequent chemo-, regio- and stereoselectivity. Therefore, they present great synthetic potential and were studied in more detail e.g., in the case of SHCs.

1.5. Squalene hopene cyclases

One of the most prominent class II TS is the bacterial SHC from *Alicyclobacillus acidocaldarius* (*Aac*SHC), which belongs to the SHC superfamily with 325 assigned

members.[139] The physiological function of SHCs is the conversion of squalene into hopene. Hopene plays a crucial role in membrane fluidity by condensing lipid membranes and subsequently improves the structural order at higher temperatures (Figure 6). A side reactivity towards hopanol formation instead of hopene, is observed at hydration of the hopenyl cation instead of deprotonation (ratio 1:5, hopanol:hopene) (Figure 6).[140]
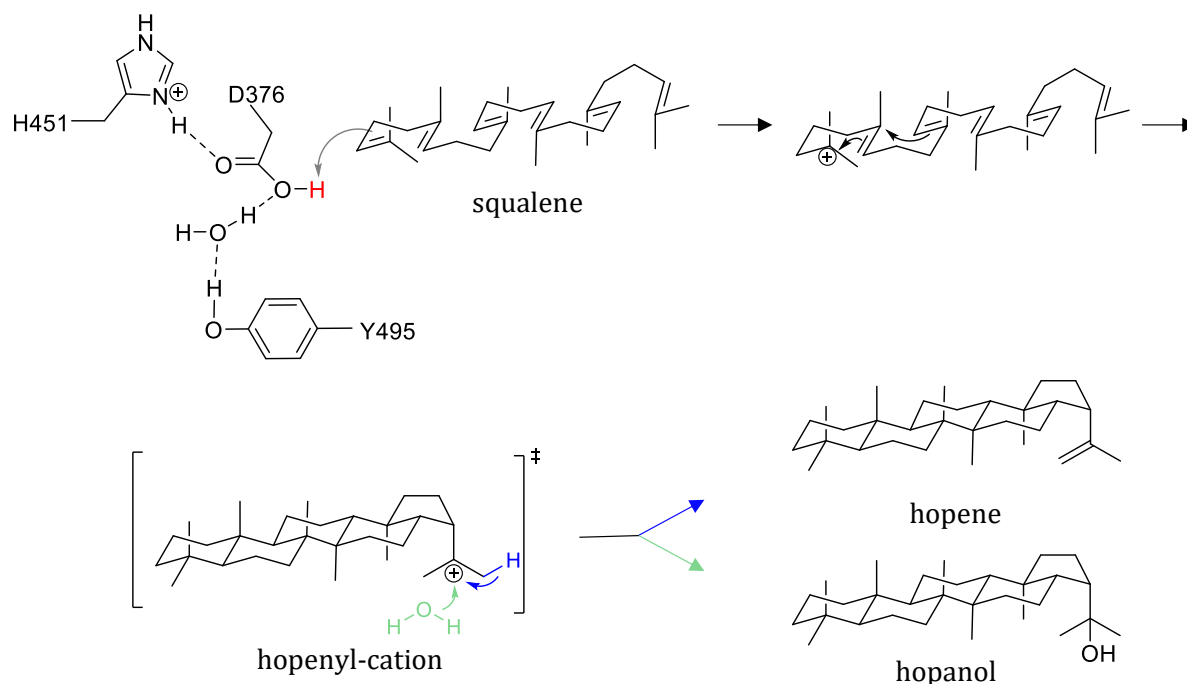
Figure 6. Polycyclization of squalene catalyzed by SHC. Active site residues involved in the initial protonation by SHC are depicted. Rearrangement of the intermediates in the polycyclization cascade yield the hopenyl-cation, which is directly deprotonated or hydrated to yield hopene or hopanol, respectively.[141]

1.5.1. Structure and function of SHC

The first solved crystal structure of *Aac*SHC (PDB: 1SQC) shows the two characteristic βγ-domains with (α/α)$_6$ barrel fold.[85] *Aac*SHC is a homodimer and a monotopic membrane protein with a membrane-associated α-helix within the γ-domain (Figure 7).[85,142] Next to the membrane-associated α-helix a dynamic, hydrophobic channel connects bulk solvent and active site for entrance and release of reaction compounds.[85,143] The active site is located in between the β- and γ-domain and is dominated by aromatic and hydrophobic amino acids (Figure 7).[85,142,144] Catalysis is initiated by the enzyme chaperoning the substrate in an all-chair conformation to establish pre-organization via release of water molecules from the active site through specific channels.[145,146] Protonation proceeds at the terminal double bond by the presumably *anti*-orientated proton of the catalytic acid D376. A hydrogen bond network with Y495 and H451

contributes to the correct orientation of the *anti*-oriented proton to increase the acidity in comparison to *syn*-oriented protons (see section 1.2.2.).[140,142,147] The residues involved in reaction termination are, e.g., E93 and Q262.[141] Notably, a 200 kJ/mol energy release upon product formation by the enzyme was observed.[143] During the energy release, stabilization of the enzyme's secondary structure is assured by highly conserved QW-motifs.[143,148]

The hydrophobic and to a large extent π-electron rich residues in the active site are mainly phenylalanines, tyrosines and tryptophans (Figure 7). These residues are believed to shield reactive intermediates from water molecules and stabilize carbocation intermediates via the confined active site of *Aac*SHC. This supports catalysis via cation-π-interactions and favors SHC product selectivity towards pentacyclization over mono- and tricyclic ring closures.[141] For example, F365 and F605 were identified to be crucial for the stabilization of the bicyclic carbocation and the hopanyl carbocation at C8, C17 and C22 of squalene.[141] Other hydrophobic residues such as L607 were shown to be involved in substrate binding.[141] Monocyclic products were observed by mutation of residues D377, W389, L607 and W489. The residues I261, W420and L607 are involved in folding and pre-organization of the substrate to control the unique stereoselectivity of *Aac*SHC.[141]

A

B
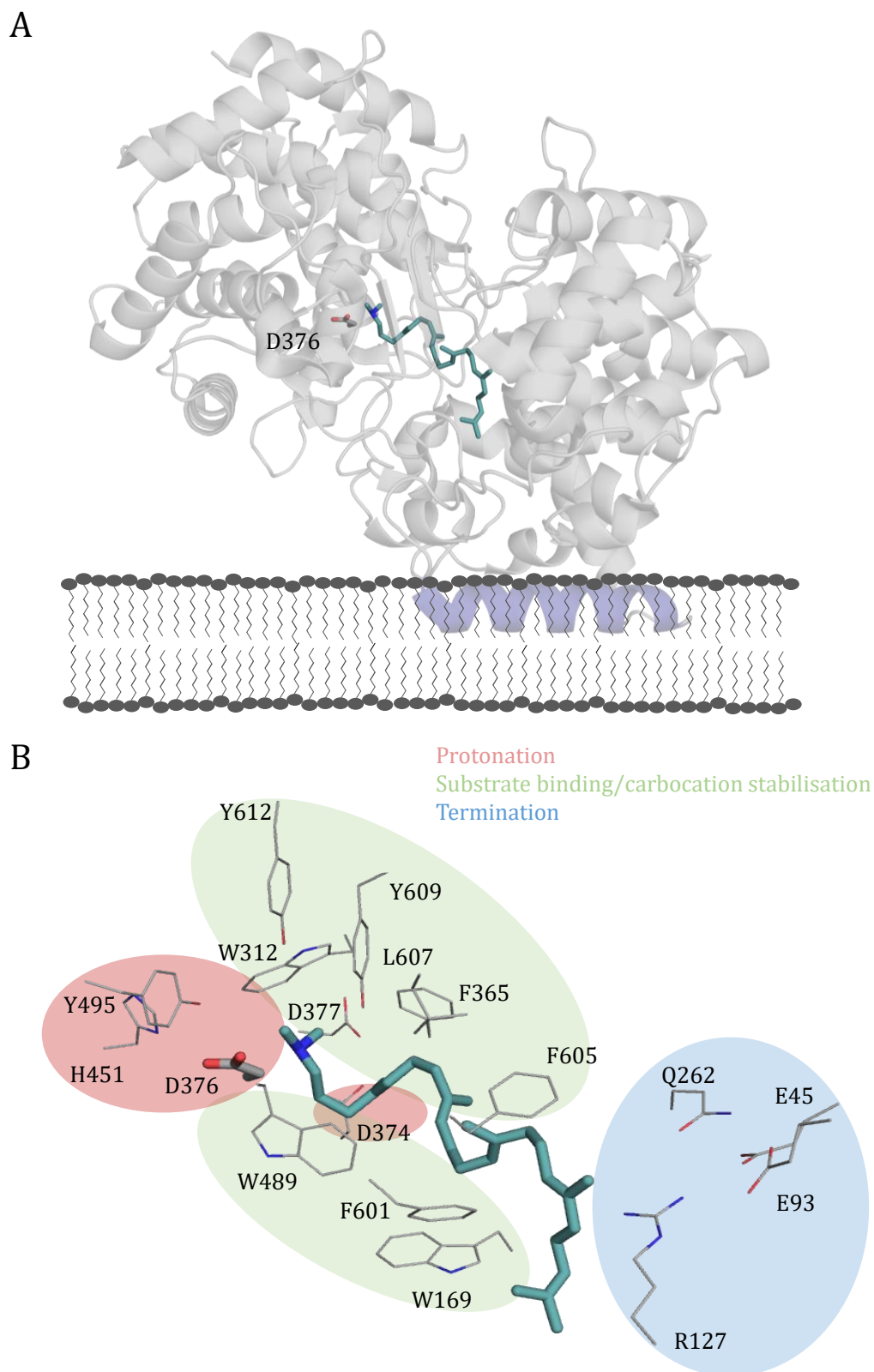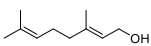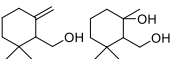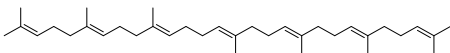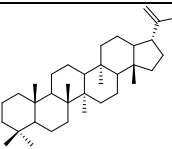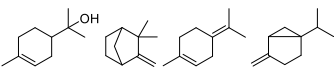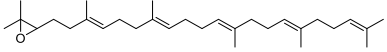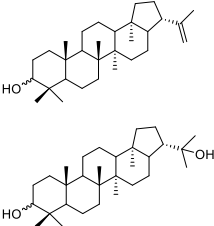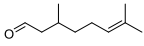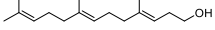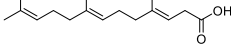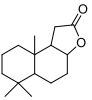
Figure 7. *Aac*SHC crystal structure and important catalytic residues. (A) *Aac*SHC cocrystallized with 2-azasqualene (teal) (PDB: 1UMP) and the membrane associated α-helix (purple) in the γ-domain and the catalytic acid D376 as part of the β-domain. (B) Active site close up with residues involved in protonation (red), substrate binding/carbocation stabilization (green) and termination (blue).[141,149]
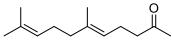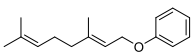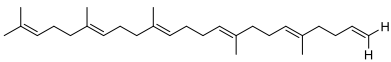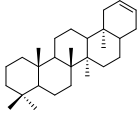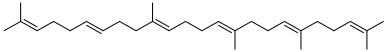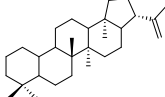
1.5.2. Substrate scope of squalene hopene cyclases

SHCs were intensively studied and demonstrated great plasticity towards various substrates (Table 1).[131,150,151] Several SHC variants were shown to catalyze the conversion of substrates of various sizes and conformations by protonation of a terminal isoprene carbon double bond.[131,141,152] Hammer *et al.* exploited the potential of SHCs as Brønsted acid catalysts and engineered *Aac*SHC to show catalytic activity for the monocyclization of small C_{10} compounds like geraniol and epoxy-geraniol.[82] Overall polyprenoids from C_{10} up to C_{35} can be successfully converted through polycyclization reactions by SHCs (Table 1).[131,153,154] In addition, acidic isomerization of β-pinene with *Aac*SHC was observed. This reaction is catalyzed via several cationic intermediates leading to a non-selective product profile of monocyclic and bicyclic terpenes, including valuable terpenes for the fragrance and perfume industry (Table 1).[155] Direction and control of cations during the reaction progress for selective product formation is of immense interest to efficiently produce value added monoterpenes. By screening SHC mutant libraries for new reactivities and increase in conversion, structural hot spots for functional plasticity were identified and tuned the product selectivity towards α-pinene formation up to 90%.[156] Modifications of the terminal isoprene unit for reaction initiation (initiation unit) were tolerated to a limited extent. Besides more reactive epoxy groups and carbonyl groups in e.g., epoxysqualene and citronellal, hardly any modifications were tolerated (Table 1).[154,157,158] In a highly enantioselective (ee >99%) Prins reaction of (*R*)- and (*S*)-citronellal was converted by *Zmo*SHC1 from *Zymomonas mobilis* to interesting precursors like isopulegol.[82,158,159] Substrate tolerance of SHCs was observed for compounds with variation of functional groups at the final cyclization unit including alcohol, carboxylic acid and ketone motifs to yield cyclic ethers, enol ethers and lactones (Table 1).[9,158–161] For example, the heterocyclic flavor compound ambrox was produced from homofarnesol and three distinct mutations of *Aac*SHC produced ambrox with up to 10-times higher conversion rate.[109] Additionally, homofarnesoic acid could be cyclized to the tricyclic lactone sclareolide and farnesol could be cyclized and linked with a second farnesol via nucleophilic attack to yield an unnatural farnesyl ether.[161] Substrates with carbonyl groups e.g., GA were also cyclized with moderate conversions by *Aac*SHC and the anchoring of the substrate by hydrogen bonding with the active site and the carbonyl group enabled the direction of the cationic cascade to form monocyclic γ-dihydroionones by semi-rational engineering.[128] Moreover, *Aac*SHC catalyzed the intramolecular

alkylation of polyprenyl phenyl ethers by Friedel-Crafts alkylation (Table 1).[162] Typically, the native cyclization via formation of tertiary carbocations is observed. Tolerated modifications at the final cyclization unit are substrates with functional groups such as hydroxy- and carbonyl groups or truncated substrates missing two geminal methyl groups (Table 1).[163]

Table 1. Terpenes with modified isoprene pattern tested in previous studies with SHCs. Tolerated isoprene pattern is defined as cyclization following the reported cyclization cascade of squalene. Protonation is initiated at the first isoprene unit from left to right by SHC.

Isoprene pattern tolerated		
Substrate	**Main Product**	**Enzyme (Conversion/yield)[ref]**
		*Aac*SHC (0.4%/-)[82]
		*Aac*SHC (-/10%)[131]
		*Aac*SHC (28.6%/-)[156]
		*Aac*SHC, *Apa*SHC[a], *Mca*SHC[b] (-/32%)[164,165]
		*Zmo*SHC1[d] (16.4%/-)[154]
		*Aac*SHC (3%/-)[166]
		*Zmo*SHC1[d], *Aac*SHC (23%/-), (2.3%/-)[160]

Isoprene pattern tolerated		
Substrate	**Main Product**	**Enzyme (Conversion/yield)[ref]**
		*Aac*SHC, ZmoSHC1 (23%/-), (23%/-)[128,160]
		*Zmo*SHC1 [d], *Aac*SHC (-/5.4%)[162]
		*Aac*SHC (-/7.7%)[157]
		*Aac*SHC (-/35%)[157]

[a] Squalene hopene cyclase from *Acetobacter pasteurianum* (*Apa*SHC)
[b] Squalene hopene cyclase from *Methylococcus cupsulutus* (*Mca*SHC)
[d] Squalene hopene cyclase from *Zymomonas mobilis* (*Zmo*SHC1)
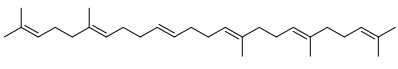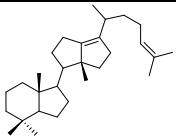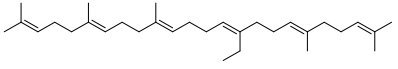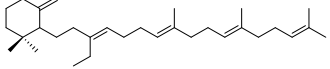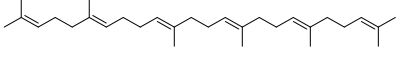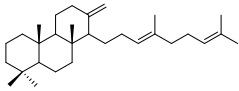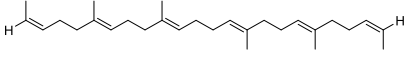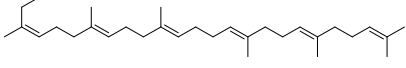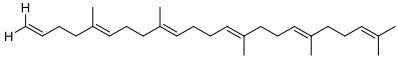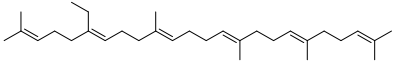
Despite the great promiscuity described above, SHCs do have limitations in terms of substrate scope. Known SHC homologs convert substrates predominantly following the usual and native isoprene rule (see section 1.4.), with variations mostly observed in form of chain length, epoxide groups at the initiation unit or nucleophilic groups at the final cyclization unit. In *Aac*SHC confinement and development of specific binding modes results in restriction to accepted substrates with the common isoprene pattern, since changes, e.g., of the methylation pattern at C10, result in abnormal cyclization (Table 2).[167,168] Furthermore modifications at the two geminal methyl groups at the initiation motif or an ethyl group at C6 resulted in no conversion at all (Table 2).

Hence, the two geminal methyl groups at the initiation unit of squalene are critical to the formation of the A-ring and are required to initiate the cyclization reactions into pentacyclic triterpenes. This further indicates that the formation of a tertiary carbocation is mandatory for the conversion of squalene by SHC.

Table 2. Terpenes with modified isoprene pattern tested in previous studies with SHCs. Abnormal cyclization is considered as cyclization deviating from the reported polycyclization cascade of squalene. Protonation by SHC is initiated at the first isoprene unit from left to right.

Isoprene pattern resulting in abnormal cyclization/not tolerated		
Substrate	**Main Product**	**Enzyme[ref] (Conversion/yield)**
		*Aac*SHC (-/77%)[168]
		*Aac*SHC (-/12%)[167]
		*Aac*SHC (-/7%)[163]
	n.d.	-[169]
	n.d.	-[170]
	n.d.	-[157]
	n.d.	-[167]

Overall, the presented studies demonstrate the great promiscuity of SHC and the applicability for enzyme engineering campaigns to exploit the diverse substrate and reaction scope of SHCs.

1.6. Lycopene cyclases

Besides SHCs several other enzymes are known to catalyze terpene cyclization reactions. Of particular interest are enzymes, which control the cyclization cascade and allow the controlled termination of the cyclization at formation of monocyclic compounds such as carotenoids. Carotenoids are essential building blocks involved in photosynthesis and regulatory networks, widely used in medicine, food, cosmetics, and other fields because of their antioxidant properties.[171,172] The key enzymes for carotenoid synthesis are

lycopene cyclases.[173] Lycopene cyclases are found mainly in bacteria, fungi and plants, but are distributed across all kingdoms of life.[173] Thoroughly studied lycopene cyclases are bacterial CrtY, cyanobacterial CrtL and LCY found in plants. β- and ε-lycopene cyclases (LCY-B and LCY-E) in eukaryotic algae and higher plants are thought to have evolved from CrtL.[174] In the following, focus will be on LCYs found in plants because of their impact in carotenoid biosynthesis.[174,175]

The physiological function of LCYs is the biosynthesis of carotenoids, carotenoid precursors and derivatives thereof (e.g., abscisic acids, provitamin A carotenoids, β-cryptoxanthin, lutein and fucoxanthin) which have biological activities.[172] In particular, LCY cyclizes lycopene in the final step of carotenoid biosynthesis via one of two pathways. In these, a β-ring is formed by LCY-B or an ε-ring formation is catalyzed by LCY-E (Figure 8).[173,176] LCY-E catalyzes the formation of one ring resulting in monocyclic δ-carotene from lycopene whereas LCY-B catalyzes the cyclization of lycopene from both sides to form the bicyclic β-carotene (Figure 8).[177] When combined, the LCY-B and LCY-E convert lycopene to α-carotene which possesses a β-ring at one end and an ε-ring at the other as visible in lutein (Figure 8). The presence of a ring at one end of the molecule prevents or inhibits ring formation by the LCY-E cyclase at the other end.[177] Besides lycopene and the mentioned monocyclic derivatives thereof other accepted substrates for LCYs are neurosporene and β-zeacarotene.[178,179]

LCYs in plants are associated to membranes of plastids and initiate the reaction presumably by electrophilic protons.[180,181] Subsequent ring closure proceeds via a positively charged transition state and a carbonium intermediate or in a concerted reaction.[176,182,183] Since most LCYs are flavoproteins they generally possess an NAD(P)$^+$/FAD-binding motif. It is reported that reduced FAD needs to be non-covalently bound to the enzyme and NADPH functions as the FAD reductant.[176] In the reaction progress, hydrogen from the cofactor is not transferred to lycopene.[176] Instead the role of FAD is proposed to stabilize the transition state carrying a (partial) positive charge via a charge transfer interaction.[176,184] The exact role of the cofactors remains to be determined.

Crystal structures or high similarity templates for homology modelling of LCYs are not available and only a few studies on the protein structure of LCYs are reported. In a β-LCY from *Capsicum annuum* (*Can*LCY-B) several highly conserved residues were determined,

and amino acid substitution to alanine demonstrated decreased relative activities compared to the WT with following residues: D127, D259, E128, E332 and H360.[184] Residues E295 and E196 were assumed to be involved in protic activation.[184]

Currently, LCYs are mainly considered in strain engineering approaches, because of their instability and low yield expression.[173,185] However, LCYs possess tremendous potential for biocatalysis with an expanded substrate scope and applications in medicine, food and the fragrance industry.
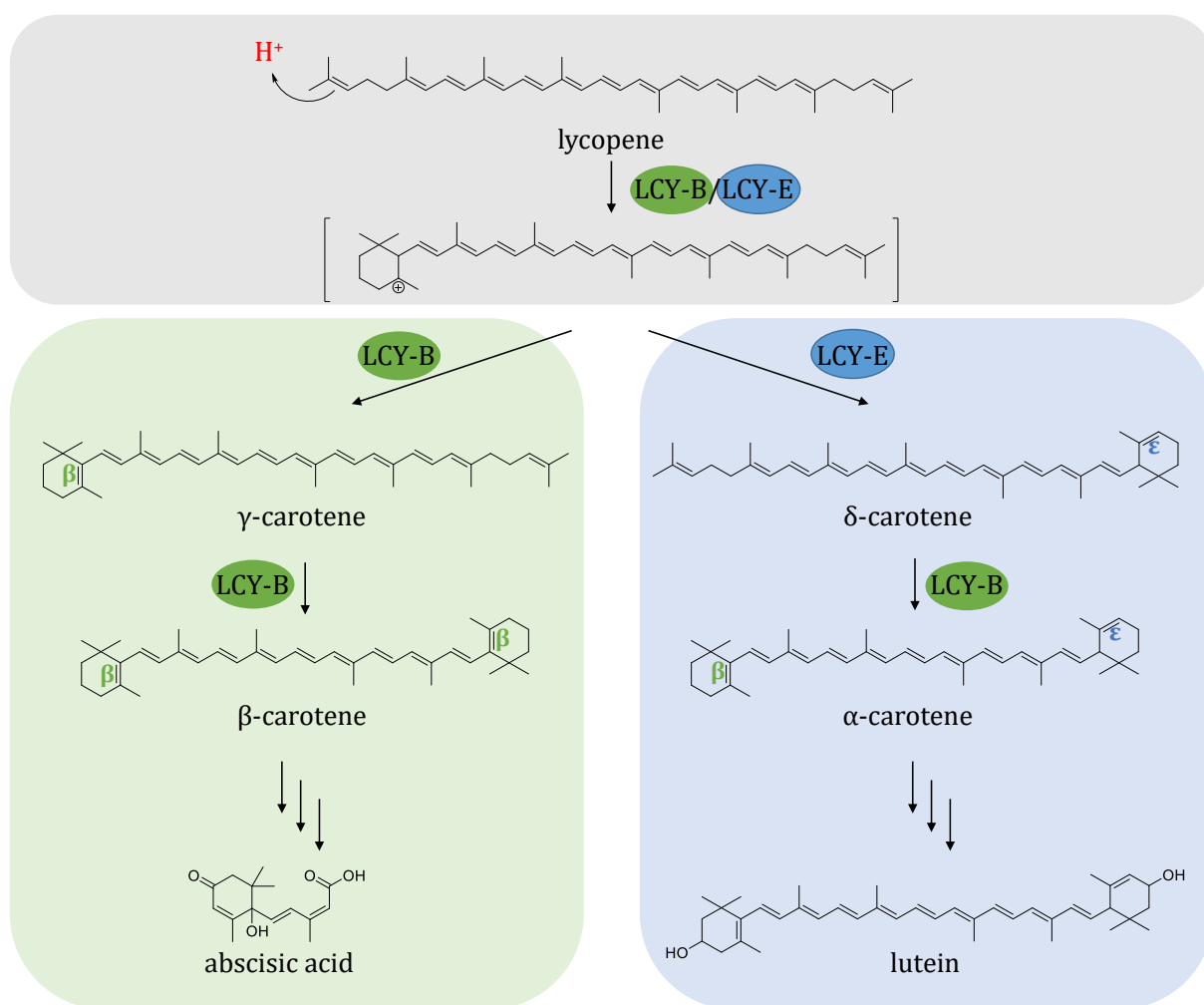
Figure 8. Lycopene cyclization by LCY-B (green) forming a β-ring and LCY-E forming a ε-ring (blue) via a tertiary carbocation intermediate. The LCY products β-carotene and α-carotene are further converted in the biosynthesis pathway to abscisic acid and lutein, respectively.[173]

2. Aim of this work

The aim of this thesis is the development of strategies to obtain pre-organization of alternative, non-native substrates with cyclases. These could form complex terpene structures, interesting products and building blocks, such as irones and terpinen-4-ol, which are in great demand in the fragrance industry (Figure 9). The generation of complex terpenes by reprogramming cyclases could demonstrate an attractive 1-step reaction under mild conditions while reducing byproduct formation. Strategies to create alternative confinement and reprogramming of the active site of cyclases are employed to increase activity and selectivity for substrates with limited pre-organization in the active site (Figure 9). This restraint is due to the evolutionary constraint to efficiently convert the native substrate and discriminate against other molecules to avoid side reactions. For the alternative confinement approach substrates with modified isoprene backbones and substrates forming multiple reactive carbocation intermediates are selected. To develop the most promising strategy for the engineering of confined biocatalysts with the modified substrates, two important principles are focused on: proximity and substrates shape complementarity. The aim is to first establish the proximity of the substrate to the catalytic center. By identifying the most promising engineering approach for proximity, a good initial activity should be achieved, which in this work is defined as a conversion of more than 1%. Rational engineering and semi-rational engineering approaches focusing on e.g., active site binding, loops or tunnel are auspicious with regard to results of previous studies. After establishment of proximity, substrate shape complementarity is employed for two substrates, in order to facilitate improved substrate pre-organization. The generated variants are kinetically characterized and investigated with regard to selectivity. All substrates that demonstrate starting activity with the identified variants are applied for preparative scale product formation. Irreconcilable limitations in this approach in terms of substrates that are unapplicable for generation of starting activity with SHC are addressed by evaluating alternative catalysts. In addition, generation of alternative confinement was tested with monoterpene substrates. Monoterpenes can form multiple transition states and intermediates with diverse reaction trajectories. This results in minor selectivity. Therefore, a different strategy was envisaged, based on cation-π interactions of the carbocation intermediates and aromatic residues to direct the reaction course. In a semi-rational engineering approach, aromatic residues and adjacent

residues are iteratively saturated, and variants with improved conversion or selectivity are characterized.

Overall, the aim of this work is the investigation of the applicability of enzyme engineering strategies to generate alternative confinement. This is to enable the productive pre-organization of substrates and form potentially valuable building blocks or products for industrial applications (Figure 9). In addition, the comprehensive structural analysis of the generated variants together with the characterization provide insights on the enzymes structure-function and expand our knowledge about the mechanisms influencing enzyme activity and selectivity.
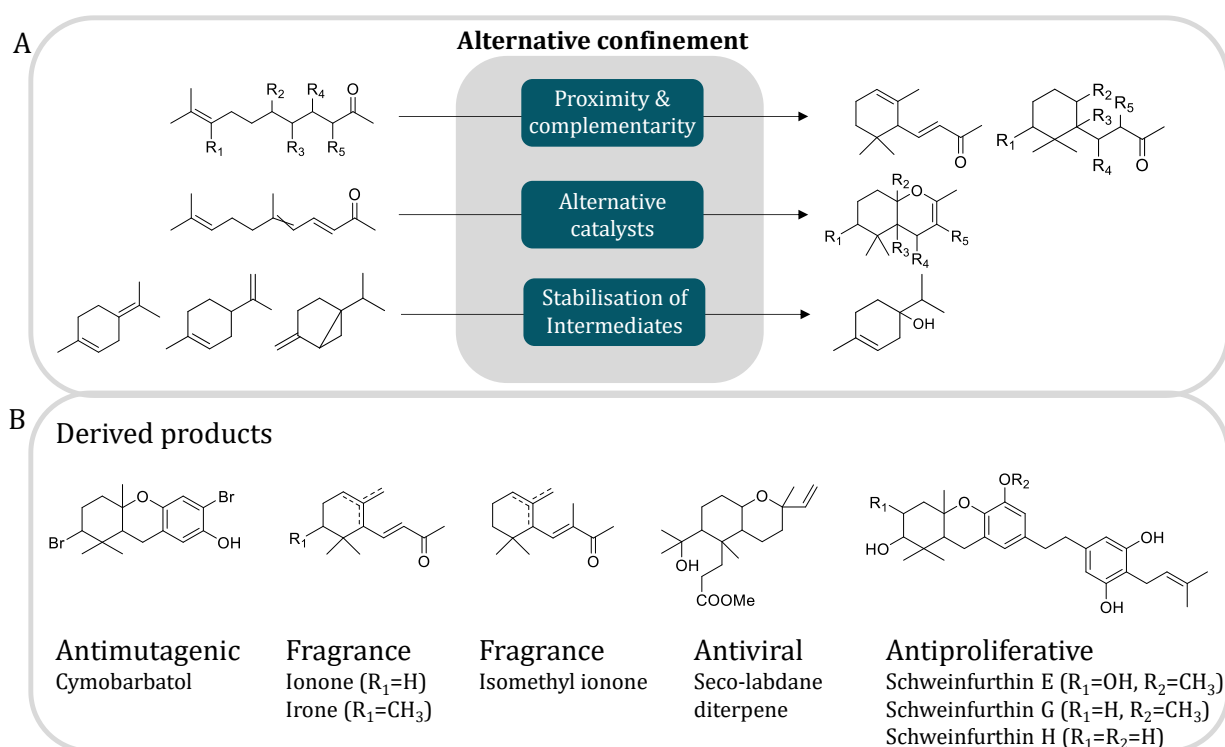
Figure 9. Strategy to create alternative confinement with cyclases for the generation of interesting building blocks. (A) Cyclization of substrates with limited ability to adopt productive pre-organization in the active site was generated via alternative confinement. (B) Industrially relevant compounds derived from possible products generated within this work.

3. Results

The results section of this thesis is divided into three parts each focusing on different approaches for achieving alternative confinement in cyclases. The first part covers the generation of proximity by variants that convert non-native, modified substrates, which are challenging to pre-organize in the active site. Improvement of conversion for two GA analogs was achieved by using a semi-rational enzyme engineering approach with *Aac*SHC, which modified the active site and created substrate shape complementarity to obtain specific secondary interactions tailored for the modified substrate. The successful approach generated variants that enabled the production of novel building blocks on a preparative scale. In the second part, for a GA analogue with conjugated double bonds different cyclases were investigated and starting activity was determined with LCYs. An alignment based semi-rational engineering approach to increase the product formation with LCYs was successfully applied to generate alternative confinement. Lastly, the third part focuses on isomerization reactions with substrates that form various intermediates along the reaction trajectory in the active site of SHC. The potential for achieving alternative confinement through secondary interactions to stabilize transition states or intermediates and increase reactivity and selectivity was investigated. The generated variants in the different approaches and their effects were investigated by computational studies on e.g., tunnel modifications, distances to the catalytic residue and probable secondary interactions in the active site. An overview of the tested substrates in this work is given in Table S 1.

3.1. Cyclization of substrates with modified isoprene pattern

The scope of the studied substrates that deviate from the isoprene pattern is shown in Figure 10. Substrates **1**, **8** and **10** showed good activity in previous studies with SHCs and are interesting in the context of modified cyclic building blocks.[128,153,160] They have different main chain lengths ranging from C11 for GA analogs to C12 for farnesol and C15 for farnesyl acetone analogs. In particular, the main chain length of the carbon skeleton was halved for most of the tested substrates compared to the native substrate squalene with a carbon chain length of 24. The GA analogs have either additional methyl groups at C3 or C9, no methyl group at C6, a hydroxy group at C9 or variation in the position of the double bonds (Figure 10). The farnesyl acetone analog **9** and the farnesol analog **11** both show variation in the position of the double bonds. In addition, **11** has a methyl group at C10.

Geranyl acetone and analogs
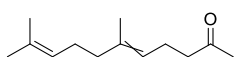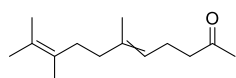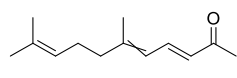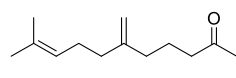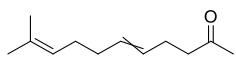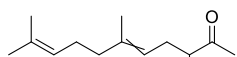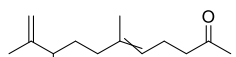
Farnesyl acetone and analog   **Farnesol and analog**
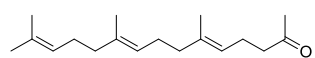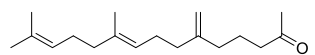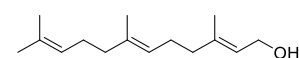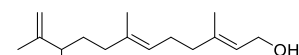
Figure 10. Terpenes with modified isoprene pattern used in this work deviated from GA **1**, farnesyl acetone **8** and farnesol **10**.

Substrate analogs **2** and **11** present a particular challenge because the active site of SHCs is more compact in the early cyclization steps and less compact in the later cyclization steps, which was indicated by studies with ethyl group substitutions at C6.[170] Bulky substituents at the isoprene moiety for initial protonation result in no cyclization, but are

tolerated in the later cyclization steps.[167,186] The conjugated γ,δ-double bond of **3** varies compared to the common isoprene pattern and results in steric interactions with active site residues. This limits the conversion with SHCs and a comprehensive screening panel of SHCs in a previous study failed in converting **3**.[187] For **4**, **5**, and **9**, variations in the cyclization mechanism due to structural challenges were expected; in this regard, for **5** in particular, no tertiary carbocation can be formed at C6, and pre-organization is expected to enable the concerted reaction.

3.1.1. SHC rational engineering and WT screening with compounds 2-6

Multiple rational engineering approaches to modify the activity of SHCs in directed evolution studies were successfully applied in the past.[82,109,128,155,187] Initially, compounds **2-6** were tested with an in-house SHC WT library containing SHCs from different organisms. Knowledge gathered from successful rational engineering studies was applied to either constrain the flexibility of the substrate by anchoring in the active site or to modify the dynamic structures of the enzyme.[128,188,189] Compounds **7**, **9** and **11** were not tested for rational engineering because this work focused on analogs of **1**, initially. SHC variant *Aac*SHC G600R was previously described to enable the binding of smaller substrates in the active site.[128] Moreover, positions Y420 and L607 in *Aac*SHC were described to influence the activity and the selectivity of the reaction. In addition, either Y420 or G600T were postulated to anchor the carbonyl group of **1** within the active site of *Aac*SHC to establish productive pre-organization of the substrate for efficient catalysis.[128] These positions were targeted in a rational engineering approach. The dynamic enzyme structures such as loops are promising engineering targets but pose a major challenge because their structures are often not conserved, and their conformation is difficult to predict due to their high flexibility. Nevertheless, previous studies demonstrated significant impact of inserts and deletions in loop structures on the activity and selectivity of enzymes with terpenes.[188,190] Therefore, residues in loops with different flexibilities were identified and selected for engineering.

The SHC WT in-house library was tested in whole cell biotransformations with compounds **2** and **3** and no product formation was detected (Table S 2). Presumably, the SHC active sites are failing to mimic the pre-organization of the native substrate squalene for the tested substrates. Therefore, conversion with variant *Aac*SHC G600R, the anchoring approach and rational engineering of flexible loops were investigated. The *Aac*SHC was selected because it showed great evolvability in previous directed evolution

studies and the enzyme's crystal structure was available (PDB: 2SQC).[85,159] The Hauer group (University Stuttgart) previously observed approximately 25% conversion with *Aac*SHC WT and 98% conversion with *Aac*SHC G600R of **1** to the bicyclic product**.**[128] Under optimized conditions in this work, *Aac*SHC WT and variant G600R showed conversions of 76% and 97% with **1**, respectively (Figure 11). Minor traces of product were observed in biotransformations with compounds **2**-**5**, whereas both *AacSHC* WT and variant G600R showed conversions of **6** as high as 97% and 98%, respectively (Figure 11). With **7**, 7% conversion was achieved with *Aac*SHC WT, but no conversion or product formation with G600R could be determined (Figure 11).
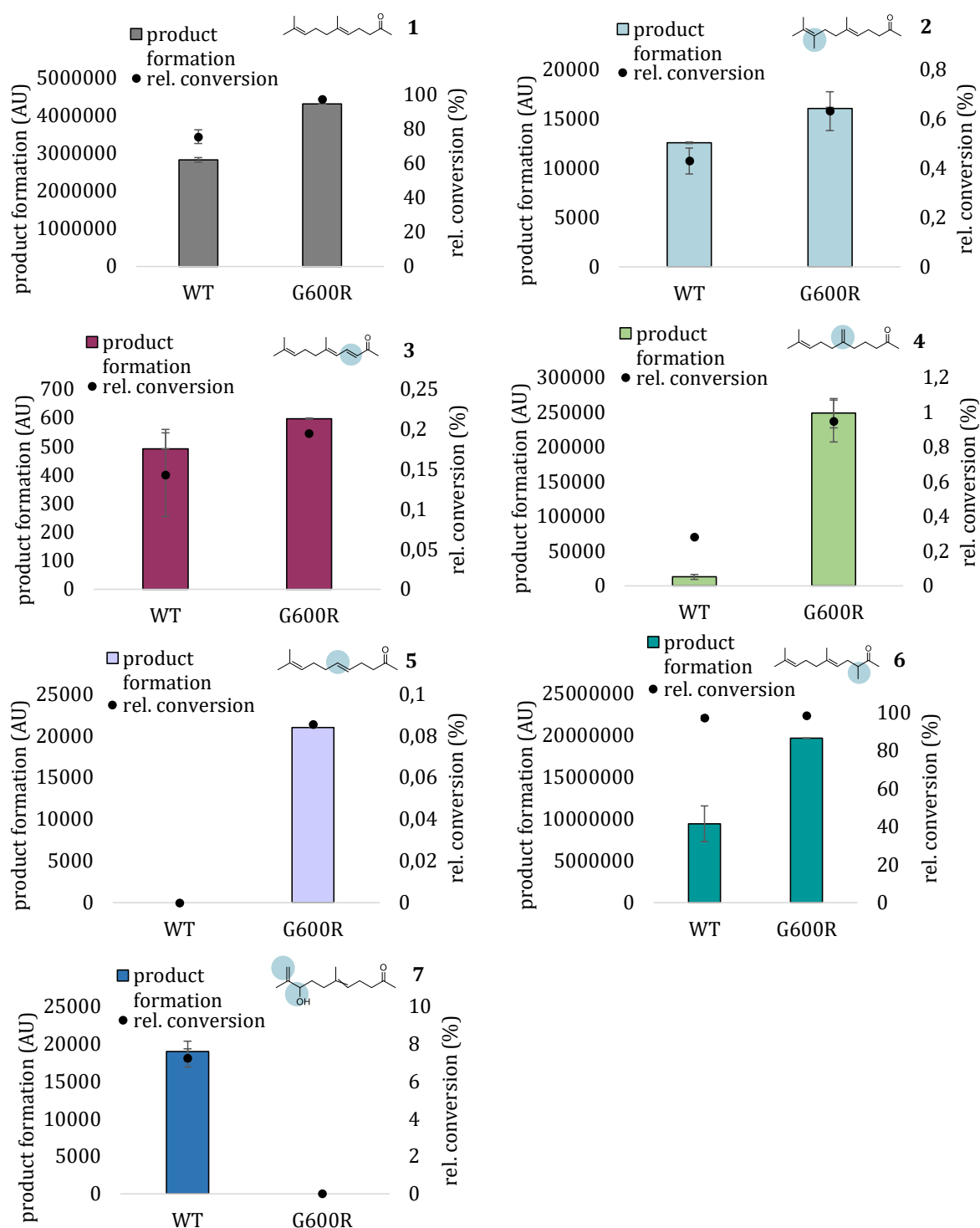
Figure 11. Product formation and rel. conversion of *Aac*SHC WT and variant G600R with **1-7** in duplicates. Error bars represent standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H2O, 1 mM substrate, 20 h, 30 °C, 800 rounds per minute (rpm).

Based on docking studies and a previous publication, positions W169 and S168 in addition to Y420, G600 and L607 were selected for the anchoring approach to allow the hydrogen bonding with the carbonyl group of **1-6** (Figure 12).[128] The product formation

in biotransformations of **2** and **3** with variant Y420F/G600T and Y420F/S168Y/W169G were analyzed. Mutation W169G was suggested to allow interaction of S168Y with the carbonyl group of the substrate and thereby position the substrate close to D376. However, no product formation was observed with any of the tested variants (data not shown).

The rational approach to modify loops by insertion and deletion was based on four spatially separated active site loops analyzed by b-factor analysis. They varied in flexibility and were promising targets to generate starting activity with **2**-**5** (Figure 12), but no product formation was detected with the initially tested compounds **2** and **3** (data not shown). Therefore, the loop variants were not further investigated and not tested with other substrates. Subsequently, modifying the flexibility of the active site loops posed no alternative active site confinement for the pre-organization of the tested substrates.
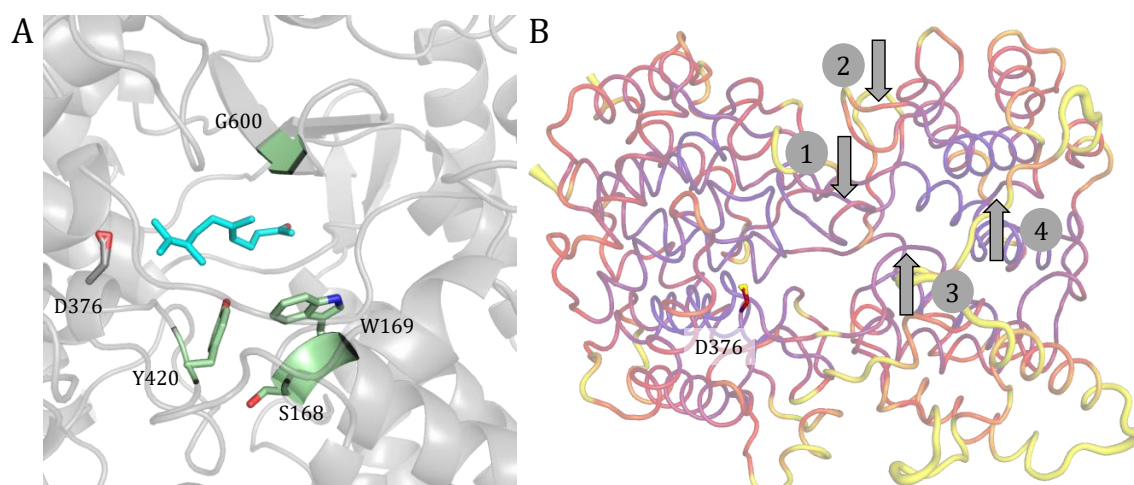
Figure 12. Targets for rational engineering to increase conversion of substrates **2** and **3**. (A) Amino acid residues (green) with positive effects on conversion of **1** observed by Schneider *et al.*[128] (B) B-factors of the residue backbones from 1-20 indicated by the color-coded cartoon structure of *Aac*SHC, from low (purple) to high (yellow), with active site loops altered by insertions and deletions (indicated by arrows numbered from 1-4).

In conclusion, the rational approaches tested in this work with compounds **2**-**6** did not result in product formation. Therefore, a semi-rational approach was planned in the next step. Due to the high conversion of **6** with *Aac*SHC WT the enzyme was chosen as the starting template. To generate the initial activity with **2**-**5**, a semi-rational engineering approach was used based on residues that were thought to have effects on the dynamic tunnel structures. Saturation mutagenesis of these residues could allow pre-organization

of the substrates via various interactions, e.g., steric interactions, π-π interactions, or hydrogen bonding.

3.1.2. Semi-rational engineering resulted in a tunnel modifying variant

A semi-rational engineering approach with *Aac*SHC was employed in the following, using rationally selected positions lining the substrate or product tunnel in the structure simulated by molecular dynamics (MD) modeling. These residues were exchanged to alter substrate preorganization, which has been shown to be relevant for efficient conversion of squalene in WT and is impaired for modified substrates.[146,191] Saturation mutagenesis was performed based on an established protocol for screening SHC variants for the conversion of **1** in 96 deep well plates (DWPs) developed by Andreas Schneider (University Stuttgart).[128] Out of the tested variants, *Aac*SHC G600M (VD1) was the only variant showing conversion of **2** and **3** (Table 3). Despite the different steric requirements of compounds **1**, **4**, **5**, and **9**, biotransformations with VD1 resulted in starting activity or further increase in turnover, with **4** showing the highest increase in turnover of 41-fold compared to the WT enzyme (Table 3). For compound **6,** VD1 did not lead to an increase in product formation (Table 3) and for compound **7** and **11** no product formation with VD1 was measured. Nevertheless, 7% conversion of **7** with *Aac*SHC WT was detected and further screening of SHC variants from the in-house library for activity with **7** identified one variant with 51% conversion (Figure S 1). This variant was the *Zmo*SHC1 Q221S from *Zymomonas mobilis,* which was shown to catalyze the Prins reaction with citronellal in a previous study.[192]

These findings indicate that distinct bulkiness near C1, such as the methyl group in **6**, promotes the correct orientation of substrate in the active site and can lead to almost full conversion. This is probably due to fixation of the substrate in a more productive conformation. With substrates **2**-**5** and **9** the G600M amino acid substitution was necessary to obtain starting activity.

Table 3. Selected terpenes investigated in this work with deviations from the isoprene unit marked in blue circles. The relative conversions with the WT and VD1 variants and the fold increase with variant VD1 compared to the WT in duplicates ± standard deviation are shown. Products are predicted based on the structure of products formed in the conversion of **1**, **8** and **10** in previous studies.[128,153,160]

Substrate		Products		Rel. conversion (%)[a]		fold increase[a]
				WT	**VD1**	
1	(structure)	**1a**	(structure)	75.7 ± 4.0	99.6 ± 0.6	1.3 ± 0.1
2	(structure)	**2a**	(structure)	0.4 ± 0.1	1.9 ± 0.1	4.5 ± 0.7
3	(structure)	**3a**	(structure)	0.14 ± 0.05	0.4 ± 0.1	3.1 ± 0.4
4	(structure)	**4a**	(structure)	0.3 ± 0.1	11.6 ± 0.4	41.3 ± 0.5
5	(structure)	**5a**	(structure)	n.d.	3.7 ± 0.1	
6	(structure)	**6a**	(structure)	97.1 ± 1.2	99.9 ± 0.1	
7	(structure)	**7a**	(structure)	7.2 ± 0.5	-	-
9	(structure)	**9a**	(structure)	11.1 ± 0.4	39.2 ± 6.1	3.5 ± 0.7
11	(structure)	**11a**	(structure)	-	-	-

[a] Reaction conditions: *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H2O, 1 mM substrate, 20 h, 30 °C, 800 rpm.

To understand how the single mutation G600M leads to an improved catalyst for the tested substrates **2**-**5** and **9**, further studies on enzyme dynamics comparing WT and VD1 in terms of flexibility and bottlenecks in tunnels were performed.

Several studies on dynamic structural elements show that tunnels are effective targets to engineer catalysts for valued reactions. Changing flexibility and removing bottlenecks in particular have been shown to be beneficial.[193,194] Additionally, a relationship between enzyme activity and the flexibility/rigidity of residues in the active site has been reported.[189,194]

Flexibility of the WT and variant VD1 without substrate was evaluated by the MD simulated structures and computed b-factors of the backbone for each amino acid. There was no clear difference in flexibility around the active site, including position 600, and apart from the differences between the surface-exposed residues, similar b-factors were observed in the WT and VD1 variants (Figure 13).

Variant	Averaged b-factor
WT	20.13 ± 14.86
VD1	19.46 ± 13.12
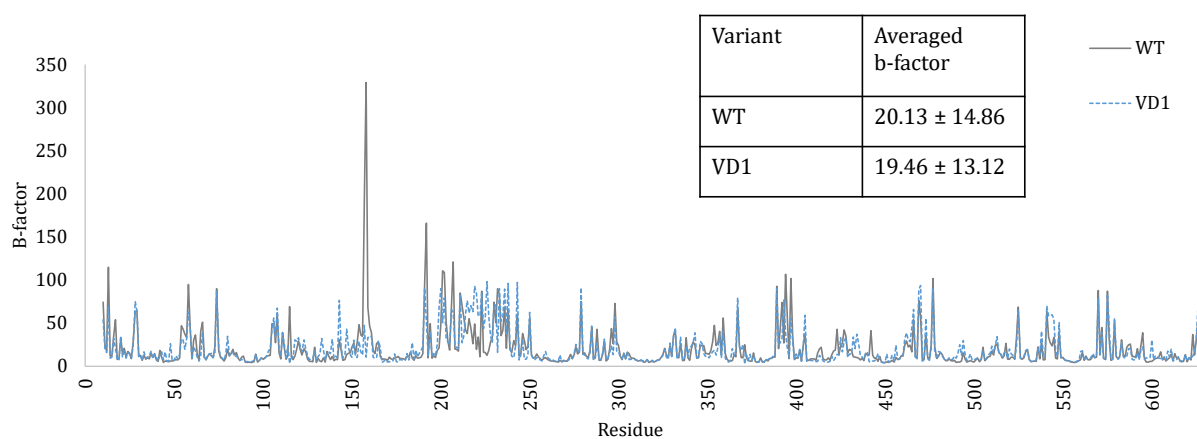
Figure 13. Backbone b-factors for each amino acid of *Aac*SHC WT (gray) and VD1 (dotted blue). B-factors were calculated as the average of the individual MD simulations without ligand present.

Next, tunnel profiles of the simulated tunnels in WT and VD1 were investigated for structural differences and bottlenecks. The starting point of the simulation with CAVER 3.2 software was D376 and the parameters were selected in order to assure the presence of the postulated substrate entrance tunnel (T1).[85,195] The simulation indicated the presence of two additional tunnels in the WT enzyme, named T2 and T3 (Figure 14). Tunnel T2 corresponds to one of three putative water tunnels previously reported.[145] Although a narrower radius of 0.7 to 1 Å is commonly used to simulate water tunnels, tunnel T2 was calculated with the selected probe radius of 1.4 Å.[145,196] The tunnel T3 is partially overlapping with a cavity located adjacent to D376, which is separated from the surrounding solution by a salt bridge and supposedly provides water for the reprotonation of D376 after turnover.[142] In the present work, the focus was laid on tunnel

T1. Changes in enzyme structure were observed when simulating tunnel structures with VD1, as the only tunnel simulated with the same parameters as for WT was T1. (Figure 14). Comparison of the T1 tunnel profiles for WT and VD1 showed minor differences in the radii from 15 Å tunnel length to the enzyme surface (substrate entry/exit). In the range of 5.5-15 Å tunnel length around position 600 large differences in the tunnel structure of T1 were observed. This was due to the reduction of the radius from 3.2 Å to 2.3 Å resulting in a pronounced bottleneck for VD1 compared to WT (Figure 14). Supposedly, this contributes to the productive positioning by reducing the degrees of freedom for substrate positioning. From this structural analysis of the active site tunnel, evidence for improved positioning of substrates **2-5** and **9** by adaptation of the active site tunnel via a bottleneck for the smaller substrates compared to the native substrate squalene with the variant VD1 was provided.
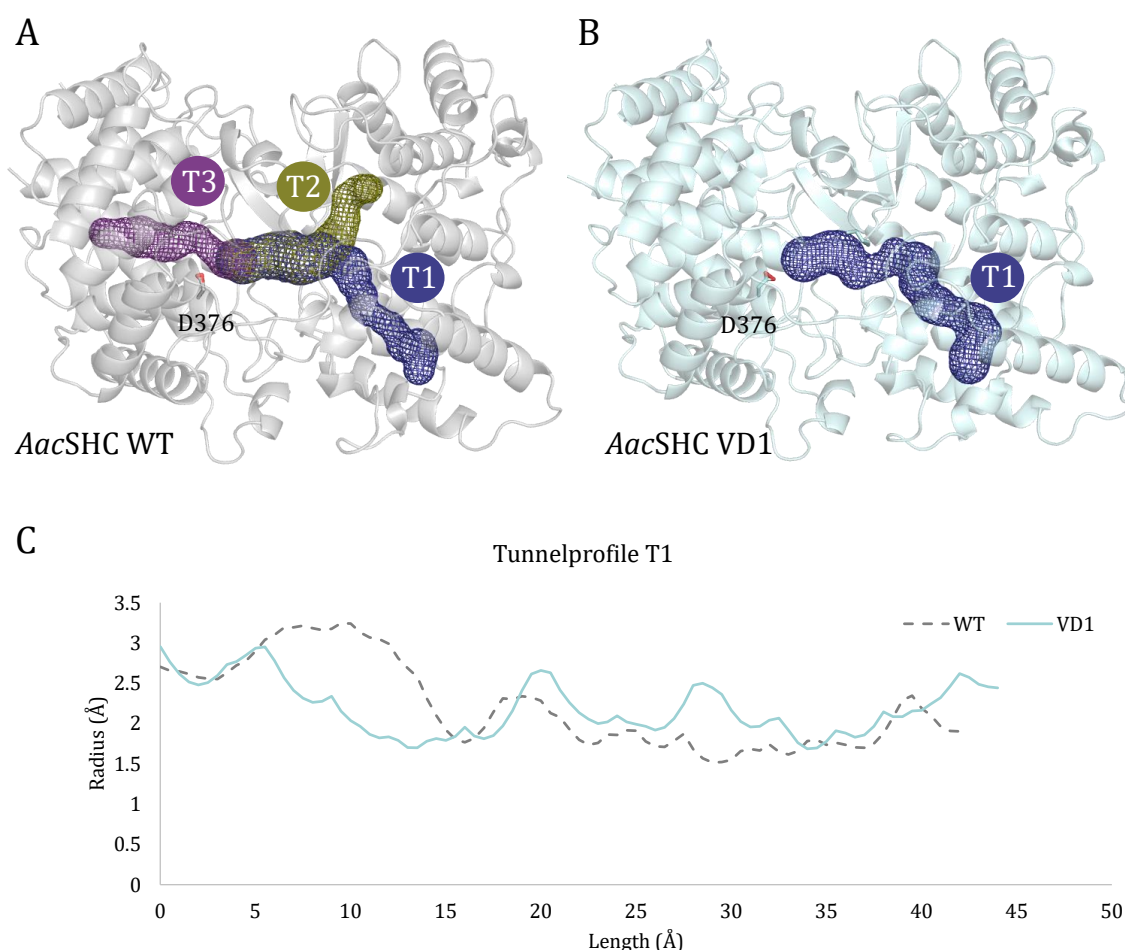
Figure 14. Tunnel structure of the *Aac*SHC WT and variant VD1. The postulated substrate tunnel T1 (blue), water tunnel T2 (olive), and tunnel T3 (purple) are shown for the (A) WT and partially for the (B) variant VD1. (C) Tunnel profiles for the WT (gray) and VD1

(light blue) are shown against the length of tunnel T1 starting from D376. Tunnel calculations were performed using CAVER 3.2 software.[195]

The promiscuity of variant VD1 toward the tested substrates demonstrates the ability of single point mutations in redesigning enzyme tunnels and active sites. This work provides evidence that the unbranched, flexible methionine at position 600 is mandatory for the starting activity with **2**-**5** and **9** by decreasing tunnel radius and that other residues such as valine or isoleucine did not increase product formation.

3.1.3. Structural proximity analysis with molecular docking

Since protonation of the substrate is the rate-limiting step in the SHC cyclization reaction, a decrease in the distance between the substrate and the side chain of D376 could lead to proximity and the observed activity of VD1 with the novel substrates.[145] Docking studies were performed with WT and VD1 for the substrates **1**-**7** and **9,** whose conversions were altered upon biotransformation with the VD1 variant compared to WT. To evaluate the hypothesis of the generated proximity between substrate and D376 the distance was calculated. Therefore, for the most appropriate docking state with proper orientation and favorable binding energy, distances of the double bond of the initiation unit to the side chain of D376 (reaction initiation distance) were determined. The reaction initiation distance, however, was the same for WT and VD1, and no tendency for altered binding parameters with the new substrates was observed (Table S 3). Only when the distance for all 25 simulated docking states was averaged, clear tendencies towards smaller reaction initiation distances in VD1 compared to the WT could be measured for the substrates (Figure 15). Boxplots of the measured distances of the 25 calculated docking states show the distribution of distances around the mean (Figure 15). The distances determined for compounds **2**, **4**, **5**, and **9**, for which the largest effects on conversion and product formation were observed in the biotransformations, show a condensed distribution of distances (Figure 15). On the contrary, examining the distance distribution of **3** and **6**, which revealed minor differences in activity between the WT and the VD1 variant, the opposite effect was observed. Herein, a rather narrow distribution of conformations was observed for the WT, whereas the VD1 variant exhibited a broader range of variation for the distances of the docking states (Figure 15). Notably, with compound **7** no activity was detected with VD1, but the distance was decreased, and the distance distribution was condensed. Overall significant differences in distance distribution between the WT and VD1 were determined for the substrates **2**-**7** (Figure 15, Table S 4).
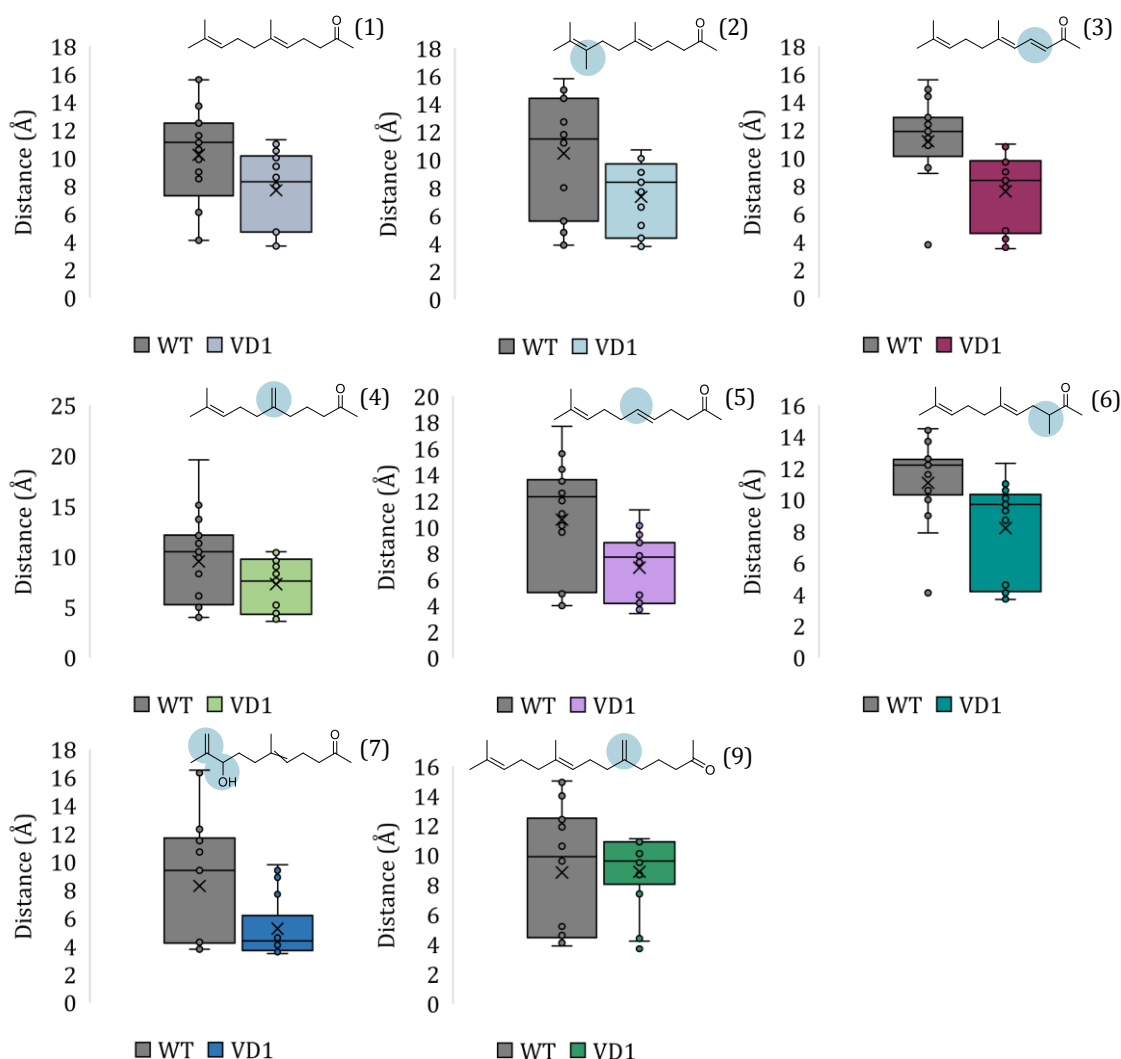
Figure 15. Boxplots of calculated distances from the double bond of the initiation unit of docked substrates **1**-**7** and **9** to D376 of WT and VD1, respectively. For each substrate, 25 energetically favored docking states were evaluated.

Furthermore, clustering the docking states in frames of 2 Å according to their distance from D376 (distance conformation clusters) revealed 2-3 different distance conformation clusters for the WT, while variant VD1 shows only 1-2 clusters for substrates **2**, **4**, **5**, and **9** (Figure 16). This indicates an improved positioning of the initiation unit for catalysis with the catalytic acid D376 and a more productive conformation of the substrates **2**, **4**, **5**, and **9**.

From the reported results, the achieved proximity of the tested substrates to D376 is evident and indicates an improved positioning of the substrates for the protonation reaction in the active site.
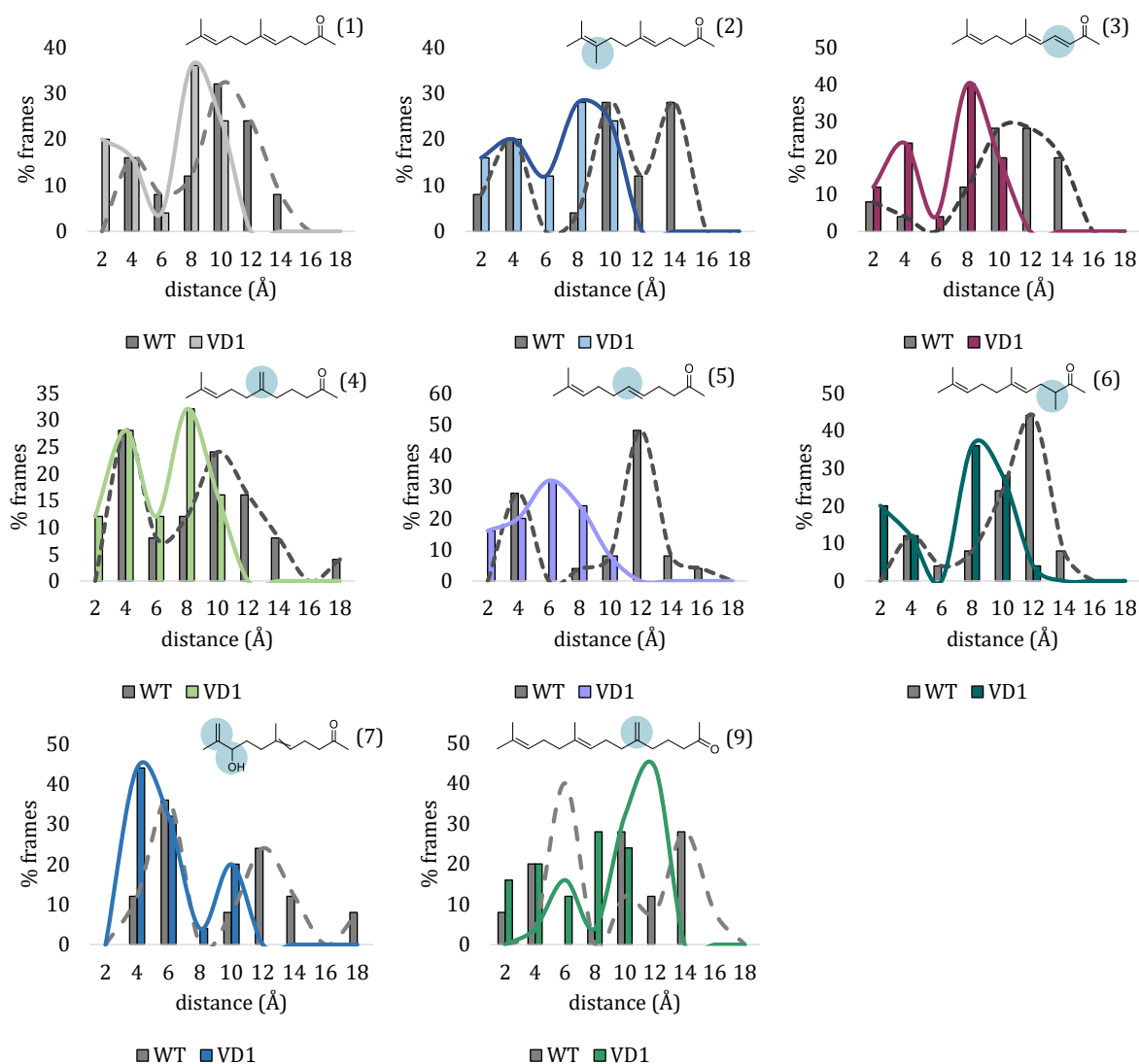
Figure 16. Distribution of distances in 2 Å frames for the distances from the double bond of the initiation unit of docked substrates **2**-**7** and **9** to D376 of WT and variant VD1. For each substrate, 25 energetically favored docking states were evaluated.

3.1.4. Evolvability of *Aac*SHC for the conversion of 2

To investigate the potential for improvement and demonstrate the evolvability of the VD1 variant, iterative saturation mutagenesis was performed in order to improve conversion of **2**. For this, a screening method with VD1 was established and subsequently reaction conditions were optimized. For the semi-rational approach with *Aac*SHC, positions were selected rationally, based on the vicinity to G600M or possible secondary interactions of the active site with the sterically challenging substrates.

<u>3.1.4.1 Screening and optimization of the conversion of **2** by VD1</u>

The establishment of a reliable medium throughput screening method for saturation mutagenesis was aimed with variant VD1 followed by testing for the conversion of **2**.

The procedure was based on a screening method for SHC variants established with **1** and analysis by gas chromatography/mass spectrometry (GC/MS, see section 3.1.1.).[128] The protein expression and biotransformation in *E. coli* whole cells in 96 DWPs was tested and the extracted reactions in a biphasic system were measured directly from the sealed plate. Similar expression of *Aac*SHC in randomly selected wells of the plate was evidenced by SDS-PAGE (Sodium dodecyl sulfate - polyacrylamide gel electrophoresis) and the coefficient of variation (C_V) was determined for the product formation with **2** over 96 wells. (Figure S 2) The measured C_V through the whole plate was 20%, which is at the upper limit for the valid outcome in screening assays (Figure 17). The C_V indicates the variation throughout the plate and implies that approximately 50% increase in product formation is necessary to determine an improved variant reliably. With this screening method positive hits with at least doubled product formation compared to the reaction of VD1 with **2** were determined.
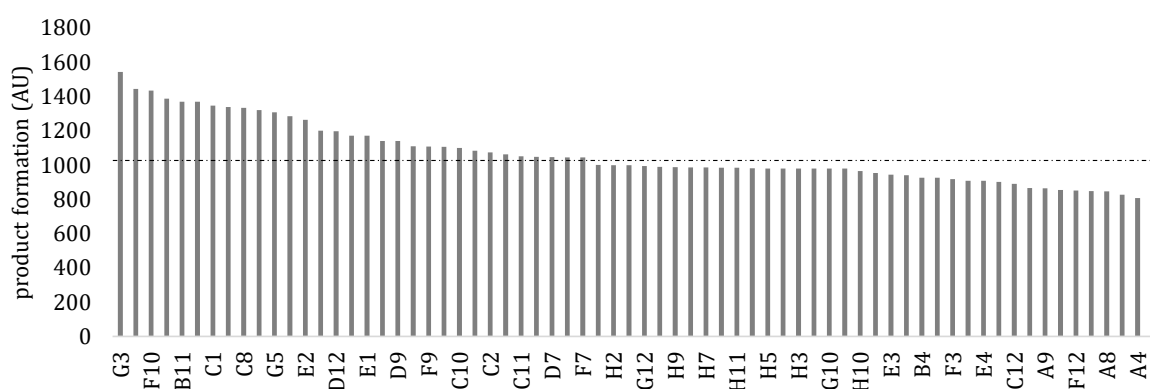
Figure 17. Product formations by variant VD1 with **2** in the medium throughput screening in 96 DWPs. A C_V of 20% was calculated for 96 samples measured directly from a 96 DWP. The dashed line indicates the average product formation. Reaction conditions: *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 0.4 ml 100 mM citric acid, 0.1% SDS pH 6, 1 mM substrate, 20 h, 30 °C, 800 rpm.

Optimization of the conversion was achieved by adjusting reaction conditions for pH, buffer and addition of ions. The enzyme was sensitive for pH values above 6 and showed the best conversion at pH 6 in a citric acid buffer containing 0.1% SDS (Figure 18). Using *dd*H_2O without addition of detergents resulted in the highest conversion (Figure 18). To ensure the comparability between the different conditions, pH was adjusted in the tested samples to pH 6 by 10 mM NaOH. In addition, different ions were selected exemplarily from the Hofmeister series and showed a positive effect on conversion.[197,198] These ions

are lyotropic agents influencing the hydrogen bonding network between water molecules. The *Aac*SHC variants performed best with **2** in 100 mM $Mg(NO_3)_2$ at pH 6 (Figure 18). Overall, 6-fold increase in conversion could be obtained by optimization of the reaction conditions in biotransformations with the best results observed for whole cells in 100 mM $Mg(NO_3)_2$ at pH 6 (Figure 18). Therefore, the reaction composition and selection of ions had a clear impact on conversion of **2** with *Aac*SHC variants and was maintained in the following biotransformations with **2**.
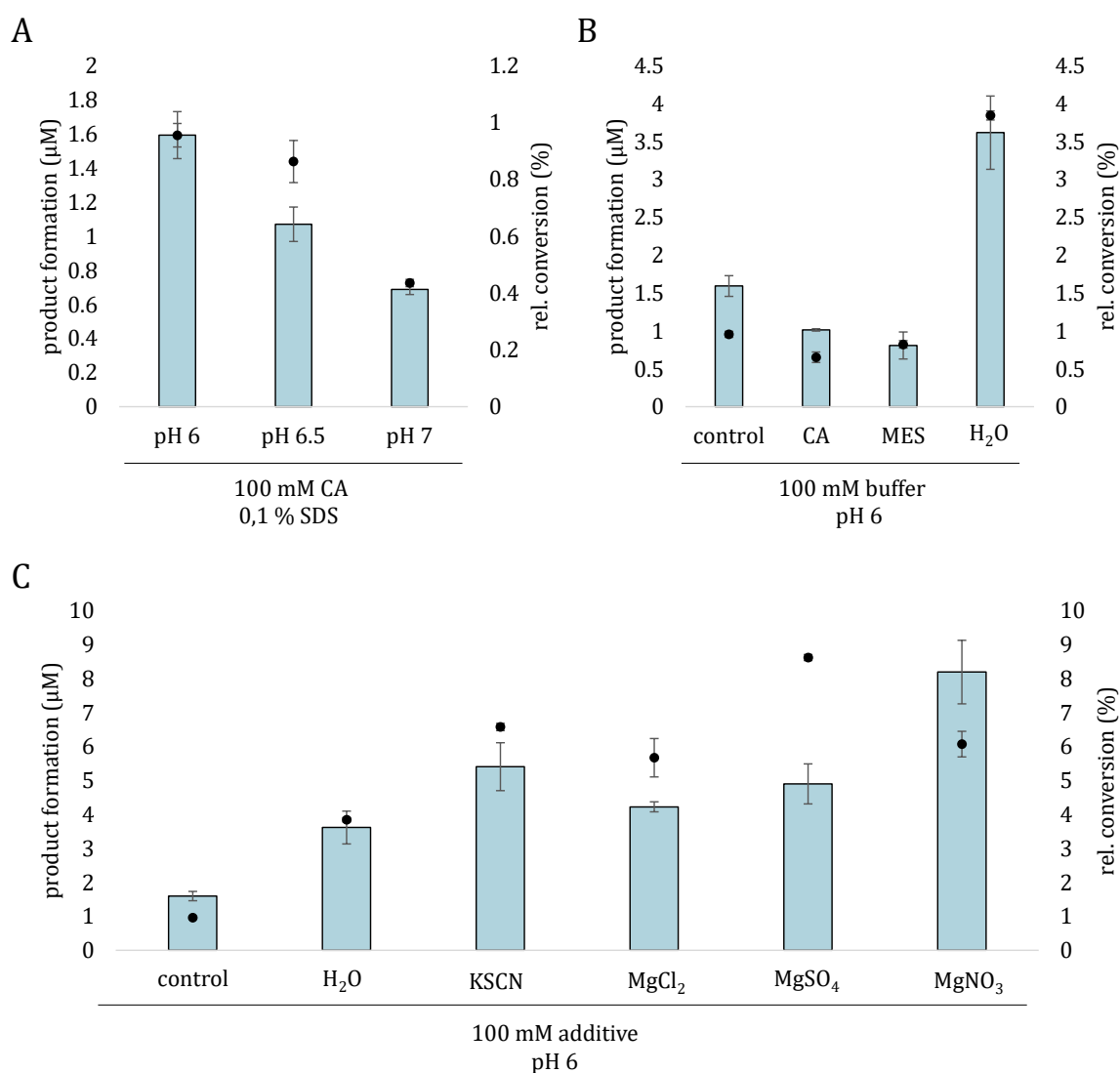
Figure 18. Product formation with **2** at change of (A) pH values in citric acid (CA), (B) buffer composition, and (C) ionic additives. Ionic additives were selected from the Hofmeister series.[197,198] Error bars represent standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in the respective solution, 1 mM substrate, 20 h, 30 °C, 800 rpm. The control reaction contained 100 mM citric acid and 0.1% SDS.

3.1.4.2. Semi-rational engineering of VD1 for the conversion of **2**

VD1 showed the potential to generate proximity of the tested substrates to D376. Refinement of VD1's active site to complement for **2** to increase conversion was therefore the pursued goal. To compensate for possible changes in stability and structural organization caused by mutation G600M, residues at a distance of 5 Å from position 600 in VD1 were initially selected as targets for saturation mutagenesis. Then the focus was on residues around the methyl group at C9 to optimize the positioning of the initiation unit near D376. The residues were located in the first and second shell of the active site within 5 Å distance to the methyl group at C9 of **2**. Variants with mutated residues around G600M within the first and second shell of the active site were screened, the best variants VD2-VD5 were isolated and WT, VD1 and VD5 were purified (Figure S 3). Successive addition of the first-shell mutations R488P and F605C by iterative saturation mutagenesis resulted in a 96-fold increase in TTN for variant *Aac*SHC G600M/R488P/F605C (VD3) compared to the WT (Figure 19). Subsequent beneficial variants Y495D and S445V with mutations around the methyl group at C9 had the greatest effect on TTN, and their addition to VD3 yielded the best variant *Aac*SHC G600M/R488P/F605C/Y495D/S445D (VD5) with 1174-fold increase in TTN compared to WT (Figure 19).

In addition, docking was performed and improved binding (decreased binding energy) of 6.6 kcal mol^{-1} for VD5 for **2** was observed compared to a binding energy of 6.9 kcal mol^{-1} for the WT. For variants VD1, VD3 and *Aac*SHC G600M/R488P/F605C/Y495D (VD4) higher binding energies between 7.1 and 7.4 kcal mol^{-1} were obtained. The best binding energy was computed for *Aac*SHC G600M/R488P (VD2) with 6.4 kcal mol^{-1}. This indicates improved binding for VD2 and VD5 and stabilization of the enzyme-substrate complex compared to the WT. However, the observed differences in binding energy are rather small and can not explain the improved catalysis.

More insights on the impact of the mutations on the confined active site for the conversion of **2** were obtained by examining the distance distribution from the double bond of the initiation unit of **2** to D376. Mutations around G600M resulted in the engineered variants VD2 and VD3 with gradually condensed distance distribution, indicating improved and confined substrate positioning. However, the improved positioning was accompanied by a greater average distance, resulting in a lower proximity to D376 (Figure 19). Thereafter, the addition of mutations around C9 of **2** greatly increased the TTN. This was achieved presumably by first generating altered pre-organization via increasing the variability of

the distances in VD4 and different positioning of the initiation unit. Second, the distance distribution was narrowed in the final variant VD5 to favor more productive substrate pre-organization (Figure 19). A closer investigation on the VD4 variant was conducted in order to determine its potential to provide an alternative catalytic acid with the Y495D amino acid substitution. In a previous study D376A showed residual activity, which was attributed to adjacent D374 and D377.[156] The distances of these residues and Y495D to the terminal double bond of the initiation unit were measured and an average distance of 12.7 Å for the best docking pose was observed for Y495D. For D374 and D377 the distances were 4.4 Å and 8.6 Å, respectively. In comparison to D376 with an average distance of 8.5 Å, Y495D and D377 are more distant for protonation of **2**. Possible protonation by D374 was not further investigated.

Between the WT and variants VD1-VD5 significant differences in TTN and distance distribution were determined for **2** (Figure 19, Table S 5). It is noticeable that variants VD4 and VD5 have similar conversions but very different TTNs. Considering that the TTN is normalized to the enzyme concentration, this indicates the superior efficiency and improved kinetics of VD5 for the reaction with **2** compared to VD4.

Concluding, the semi-rational engineering of VD1 for conversion of **2** generated variant VD5 with 1174-fold increase in TTN and altered distance distribution and positioning of **2** in the active site in 4 rounds of iterative saturation.
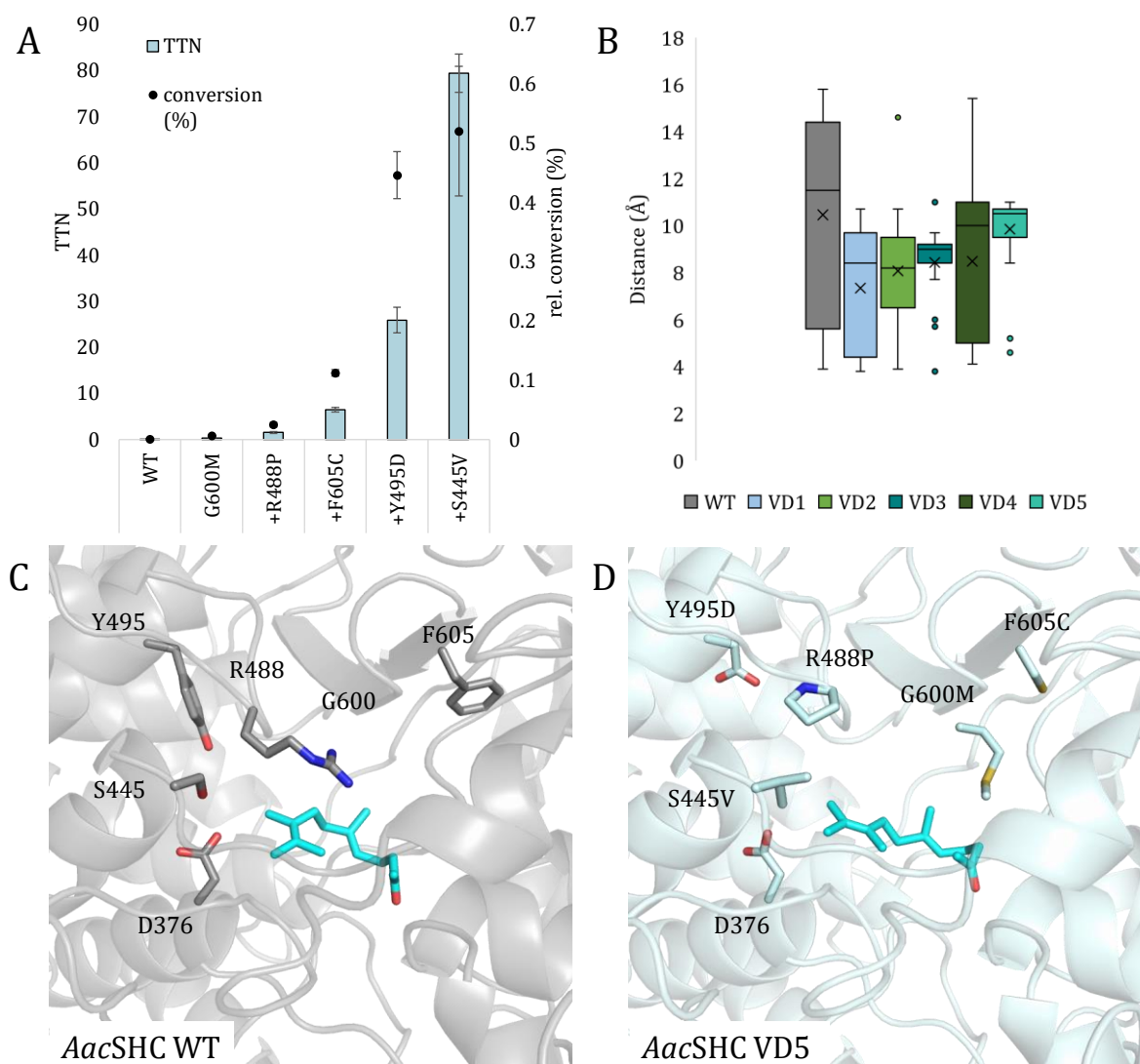
Figure 19. Alternative confinement of *Aac*SHC variants to generate complementarity with **2** by iterative saturation mutagenesis. (A) Improvement of *Aac*SHC variants for conversion of **2**, shown as total turnover number (TTN) and conversion of triplicates with added mutations in each round of saturation mutagenesis, shown under each bar. Error bars represent standard deviation of triplicates. Reaction conditions: 0.5 ml purified enzyme in ion exchange chromatography (IEX) elution buffer (25 mM citrate, 200 mM NaCl, pH 6), 1 mM **2**, 20 h, 30 °C, 800 rpm. (B) Boxplot of distances from the double bond of the initiation unit of 25 docking states from **2** to D376 for each of the generated variants. Energetically preferred docking states of **2** (cyan) in WT (C) and VD5 (D) are shown. D376 and mutant residues are labeled and shown as sticks.

3.1.4.3. Characterization and structural analysis of selected variants

To further characterize the WT, VD1 and VD5 and elucidate substrate affinity and catalytic efficiency, kinetic parameters with **2** were determined. Comparison of turnover number k_{cat} shows little variation between the WT and the variants. The catalytic efficiency k_{cat}/K_m however, is increased 61.4-fold for VD1 and up to 111-fold for variant VD5 compared to WT (Table 4). Here the WT k_{cat}/K_m is reported as apparent k_{cat}/K_m because the calculated

K_m is above the expected solubility limit of **2** (Table 4, Figure S 4). The determined enzyme kinetics provide evidence for improved steady state kinetics, in particular for k_{cat}/K_m.

Table 4. Kinetic Parameters of *Aac*SHC WT and variants VD1 and VD5 for **2**.[a]

Variant	k_{cat} (min^{-1})[b]	K_m (mM)[b]	k_{cat}/K_m (min^{-1} mM^{-1})
WT	0.058	n.d.	0.000081[c]
VD1	0.041	8.2	0.005
VD5	0.060	6.8	0.009

[a] Conditions: 2 mg/ml *Aac*SHC, 0.2 mM to 8 mM **2**, 5% Dimethyl sulfoxide (DMSO), 100 mM MgNO$_3$, pH 6.
[b] Derived from Michaelis–Menten equation.
[c] Determined from the linear slope in the Michaelis-Menten fit.

Furthermore, the influence on conversion of each individual amino acid of VD5 and their combination to double mutants was examined. For the activity, mutation G600M was mandatory and was the only variant showing good starting activity. Minor product formation was observed with F605C or WT (Figure 20). For the variants with amino acid substitutions R488P, Y495D and S445V, no initial activity could be observed. Combination of two amino acids confirmed the significance of G600M on the activity with **2** (Figure 20). When G600M was combined with one of the other individual mutations R488P, F605C, Y495D or S445V, higher product formation and conversion compared to the single mutants and double mutants without G600 were observed. Combination of the remaining amino acids resulted in traces of product or no conversion at all. Moreover, VD2 showed highest product formation and conversion. These results indicate cooperative effects of the mutations of VD5 and highlight the importance of the tunnel modifying variant G600M on the observed increase in conversion (Figure 20).
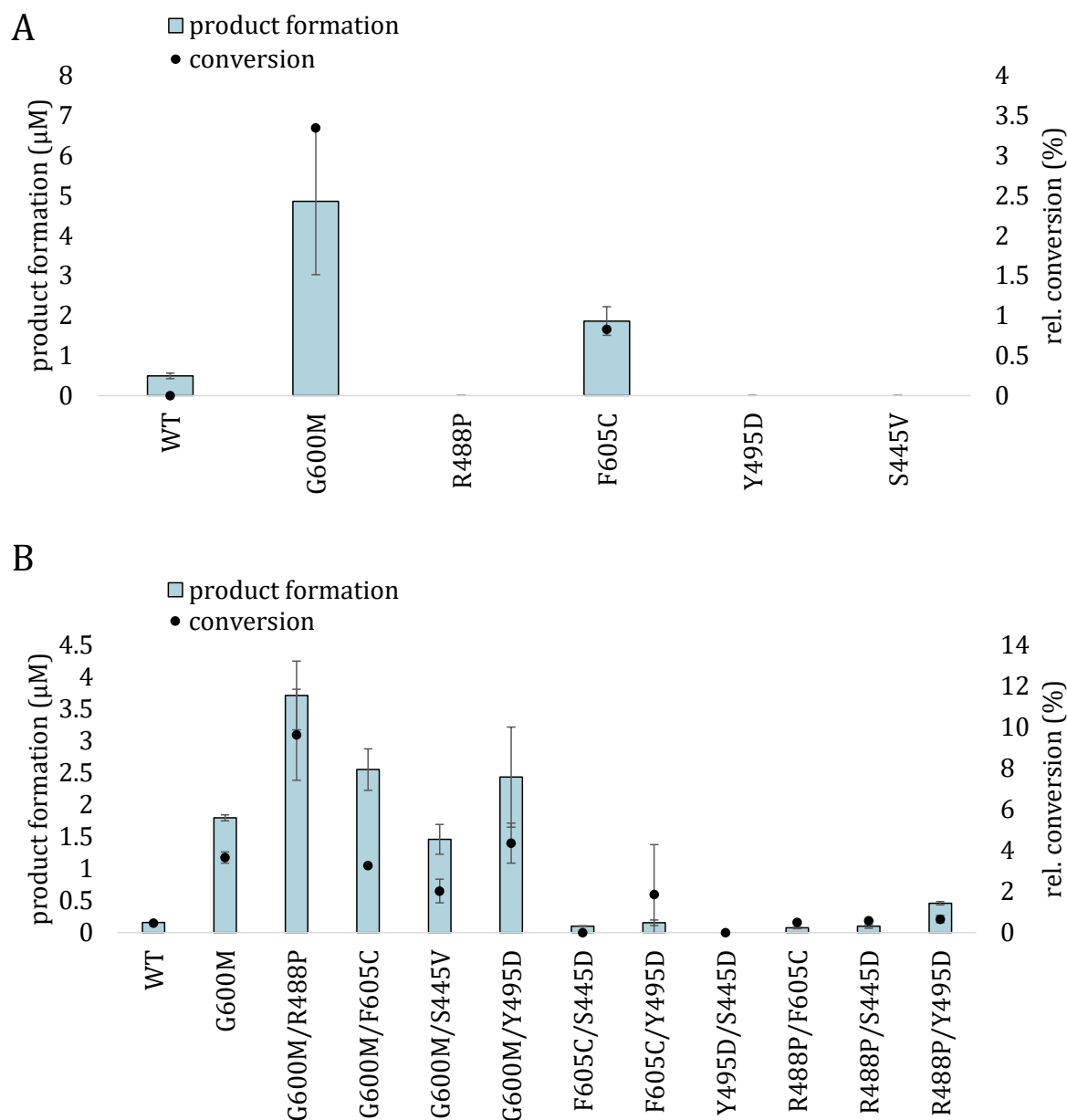
Figure 20. Product formation and rel. conversion of WT and (A) variants with individual amino acid substitutions of mutations in VD5 and (B) combination of individual amino acid substitutions. Error bars represent standard deviation of duplicates. Reaction conditions: 0.2 mg_{CWW} *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml 100 mM Mg(NO$_3$)$_2$, 1 mM substrate, 20 h, 30 °C, 800 rpm.

3.1.4.4. Selectivity of the engineered *Aac*SHC variants

To analyze the selectivity of the engineered *Aac*SHC variants for one stereoisomer of **2** and to analyze product selectivity the *E*-isomer and the *Z*-isomer of **2** were tested separately in biotransformations and products formed by VD1-VD5 were analyzed on a chiral column. For determination of stereoselectivity, *E*- and *Z*-isomers of **2** were separated with a silver nitrate column, characterized by nuclear magnetic resonance (NMR) spectroscopy and tested in biotransformations. Selectivity for the *E*-isomer but not the *Z*-isomer was

observed for all generated variants VD1-VD5 (Figure 21). For the WT no product formation was detected due to the lower sensitivity of this measurement method. In the biotransformation with **2**, two by-products were detected in addition to the main product **2a** (Figure 21). Byproduct formation was 28% with VD1. With VD5, product selectivity for **2a** was increased up to 92%, indicating tighter control of reaction progress and decreased by-product formation in VD5 compared to WT and VD1 (Figure 21). Due to the low conversion, the observed by-product could not be characterized and was not investigated further.
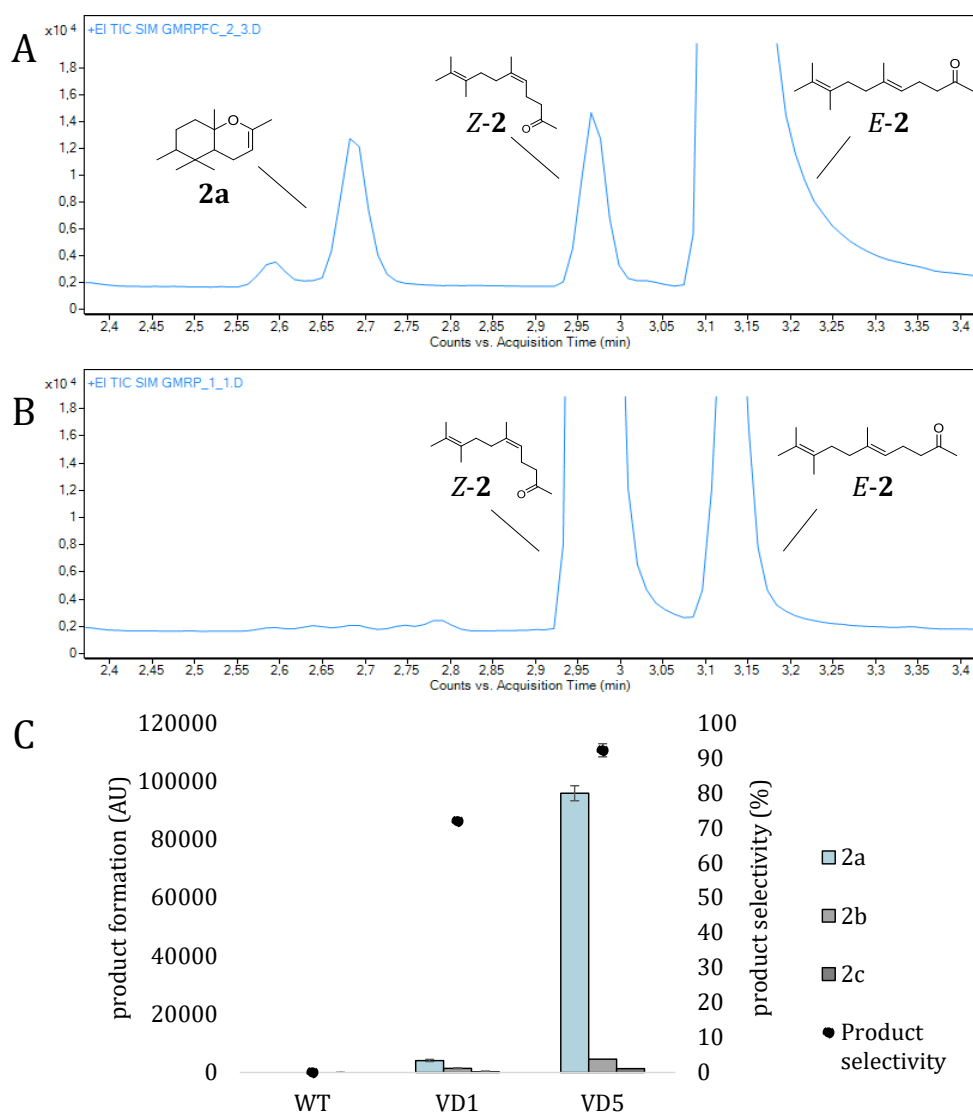
Figure 21. Selectivity of WT and VD1-VD5 with **2** measured by chiral column chromatography. The substrate selectivity is shown as an example in a chromatogram of a VD3 biotransformation with (A) *E*-**2** and (B) *Z*-**2**. (C) Product formation and product selectivity for WT, VD1 and VD5. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H2O, 1 mM substrate, 20 h, 30 °C, 800 rpm.

Overall, variant VD5 possesses stereoselectivity and is specialized for the conversion of the *E*-isomer of **2** and the formation of compound **2a**.

3.1.5. Rational design for directing the cationic cascade with 2

The cationic cascade, in which reactions often allow for the formation of several carbon–carbon bonds, rings, and stereocenters in a single transformation, was described for cyclization reactions with SHCs and **1**.[199] The cationic cascade is initiated by protonation of the substrate by SHC, which forms the first carbocation. Rearrangements within the compound continue via a monocyclic intermediate towards the bicyclic product. The goal was to direct the cationic cascade with **2** to a monocyclic product instead of the bicyclic product. This could extend the structural variability of the generated building blocks with altered isoprene patterns. A GC/MS method to identify the product with the mass fragments consistent with a simulated fragmentation pattern generated with the CFM-ID software was established.[200] The mutagenesis approach to achieve formation of monocyclic products was based on the anchoring of the substrate by hydrogen bonding described in section 3.1.1. in combination with the amino acid substitutions of VD5. It aimed to weaken the hydrogen bonds between Y420 and the carbonyl group of **2**. In addition, hydrogen bonds should be established between S168Y, G600T, or Y609 and the carbonyl group of **2** to achieve a different pre-organization of the substrate and reshape the active site to favor monocyclization. The following mutations and corresponding double and triple mutations were introduced in the *Aac*SHC variant VD5: Y420F; S168Y, W169G, G600T, L607A, F605 and saturation mutagenesis at position G600 was performed.

Initially, variants VD1- VD5 generated by iterative saturation mutagenesis for enhanced conversion of **2** to **2a** were examined for the formation of monocyclic products. Surprisingly, all variants except the WT and VD1 showed formation of monocyclic product in traces, with the variant VD5 showing highest formation of monocyclic product and **2a** (Figure 22). These results for VD5 are consistent with the determined product formation and selectivity with **2**. In the rational approach for positioning **2** in the active site to favor monocyclization via hydrogen bond interactions, single point mutants Y420F, S168Y and W169G were tested with **2**. Y420F was described in a previous study and presumably disturbs the interactions that favor bicyclization.[199] Mutation S168Y was selected based on the docking with **2**. This residue was assumed to form a hydrogen bond with the carbonyl group of **2** to promote monocyclization via different pre-organization of the

substrate. Simultaneously, the addition of W169G could generate void space to position the carbonyl group closer to S168Y (Figure 22). The biotransformations with the single mutants Y420F, S168Y and W169G, the corresponding double mutants and the triple mutant showed no conversion of **2**. Another approach for obtaining the monocyclization of **2** was pursued by adding the mutations Y420F (VD5.3), S168Y (VD5.4) and W169G (VD5.5) to the VD5 variant. Additionally, mutation G600T was tested with variant VD5.2. It was hypothesized that this would maintain shape complementarity to the modified initiation unit and provide a different pre-organization of the final cyclization unit. The variants VD5.2-VD5.5 showed decreased product formation of monocyclic product and **2a** compared to VD5 (Figure 22). In VD5.6-VD5.10 combinations of G600T, L607A and WT amino acid F605 instead of F605C were tested. However, the conversion of **2** towards monocyclic product was not increased. When mutation G600T was present, no formation of monocyclic product was measured, and neither was formation detectable when combined with other mutations (Figure 22).

Moreover, amino acid substitutions R488P, Y495D and S455V which are present in the VD5 variant and are potentially facilitating the substrate shape complementarity with the initiation unit were tested in the triple mutant VD6. The amino acid substitutions Y420F, S168Y, and W169G were added to the triple variant. However, no formation of monocyclic products could be measured with the generated variants (Figure S 5).

In the last attempt to increase the formation of monocyclic product, saturation mutagenesis at position G600 was performed. The product formation for monocyclic product for two variants was slightly higher than for the variant VD5 (Figure S 5). However, the formation of monocyclic product was below 1% for all tested variants and the product selectivity could not be shifted to monocyclic product.
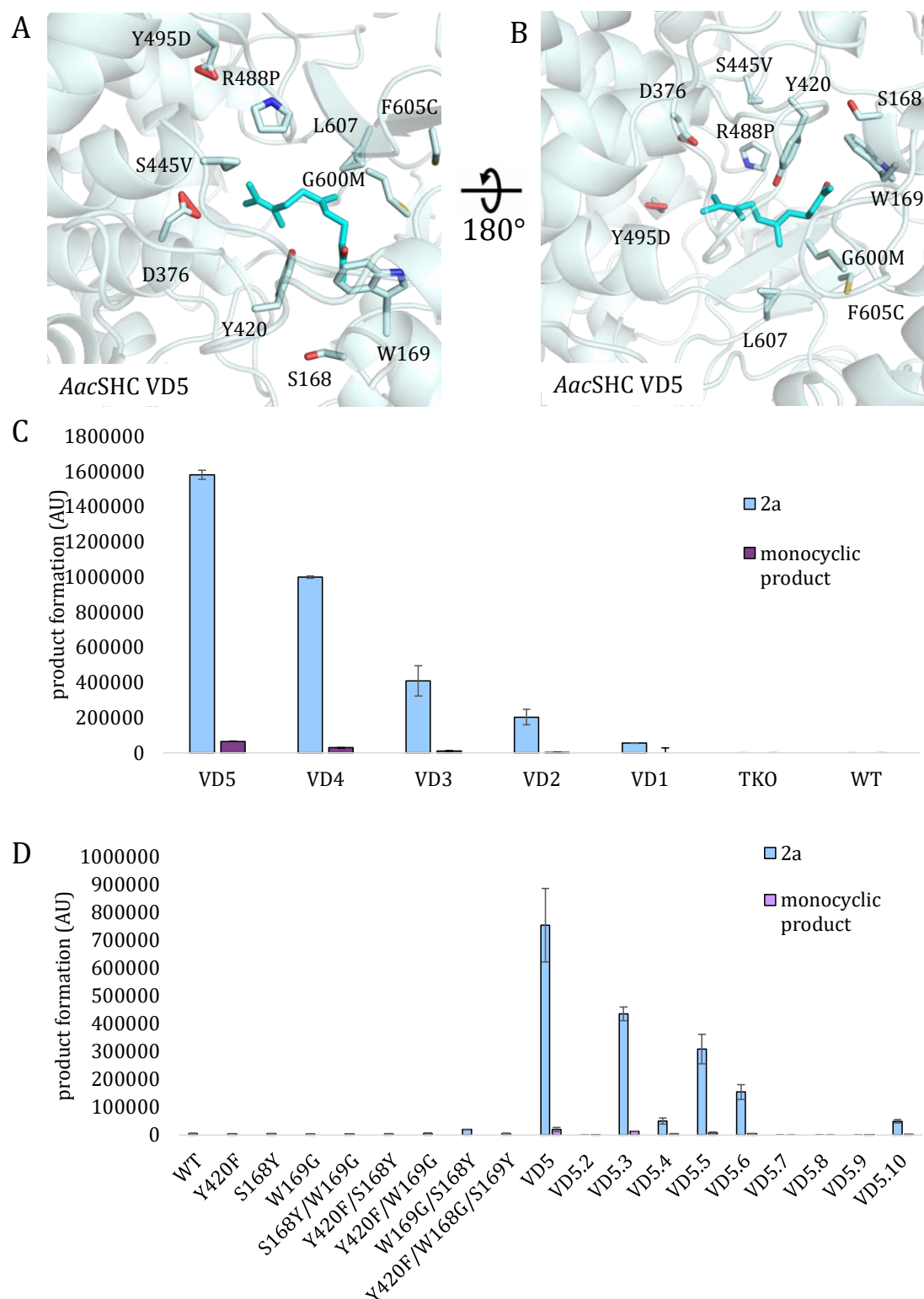
Figure 22. Rational engineering of *Aac*SHC variants for monocyclization of **2** and product formation for variants VD1-VD5. In (A) and (B) active site residues targeted for mutagenesis are labeled and shown as sticks. Formation of **2a** and monocyclic products in biotransformations with (C) VD1-VD5 and (D) *Aac*SHC single, double and triple mutants and variants VD5.2-VD5.10. Error bars represent standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml of *dd*H$_2$O, 1 mM substrate, 20 h, 30 °C, 800 rpm.

3.1.6. Structure based semi-rational engineering with *Aac*SHC for conversion of 3

In addition to **2**, increasing the conversion of **3** with semi-rational engineering of variant VD1 was aimed. Compound **3** has a more rigid structure than the common isoprene backbone due to the conjugated γ,δ-double bond. This structural deviation from the common isoprene pattern was especially challenging for conversion by SHC as shown in this work by only 0.5% conversion with the best variant identified so far. By focusing on residues around 5 Å from the conjugated double bond in the docking with VD1 saturation mutagenesis was performed. After four rounds of iterative saturation mutagenesis, three variants PI3 (F605L/G600M/Y609F), PI4 (F605L/G600M/Y609F/F605L) and PI5 (F605L/G600M/Y609F/F605L/L607F) with stepwise increase in product formation and conversion were isolated (Figure 23). Variant PI3 was identified for the first time by Andreas Schneider (University Stuttgart). The best variant identified in this work was PI5 with 9-fold increased conversion compared to the WT (Figure 23). To assess the contributions of every single mutation, the single mutants of the residues that were substituted in PI5 were tested in biotransformations with **3**. It was observed that variant F365L had the highest impact on the improved catalysis with **3** and little increase in conversion was observed for the other amino acids, individually (Figure 23). Thus, the increase in conversion of **3** with PI5 was a result of cooperative effects of the amino acids.

The structural changes of this variant were investigated based on docking studies. It was shown that bulkier residues with no polar functional groups close to the carbonyl group of the docked **3** were present (Figure 23). A notable exception was the less bulky F365L amino acid substitution in the region around the double bond in PI5. This indicates a better positioning for improved catalysis with **3** by interactions with bulky residues at and around G600M and less bulky residue F365L.

A

B

C

*Aac*SHC WT

*Aac*SHC PI5
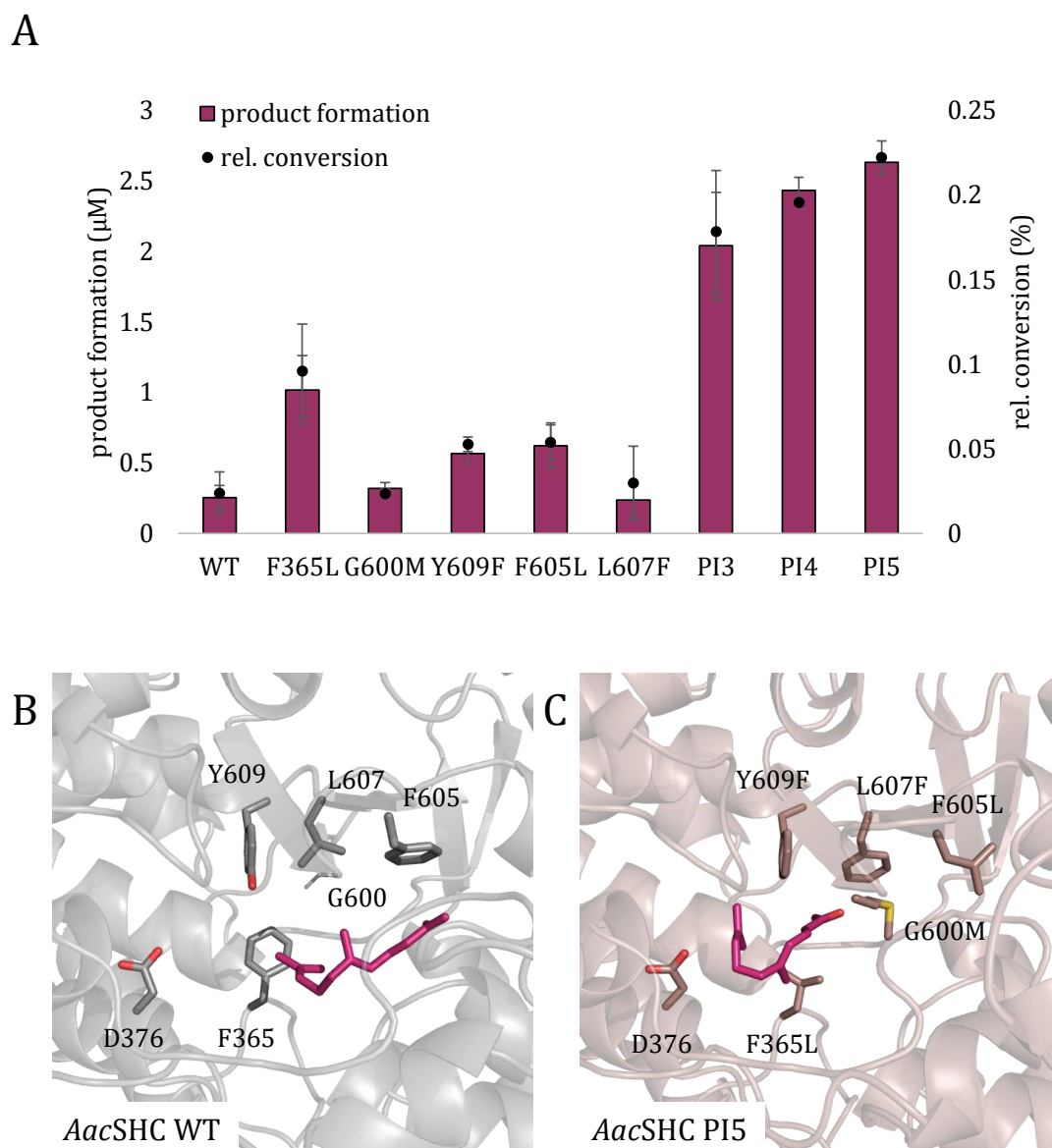
Figure 23. Saturation mutagenesis in variant PI3 for the rel. conversion of **3**. (A) Product formation and conversion with *Aac*SHC WT, single mutants derived from PI5 and PI3-PI5. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H2O, 1 mM **3**, 20 h, 50 °C, 800 rpm. Docking of **3** into the active site of (B) *Aac*SHC WT and (C) PI5. D376 and residues targeted for saturation mutagenesis are labeled and shown as sticks.

3.1.7. Identification of biotransformation products and generation of interesting building blocks in a preparative approach

To elucidate the structure of the products formed and to produce preparative amounts of the building blocks, biotransformations were performed on a preparative scale. VD5 was used for compound **2** and the *Zmo*SHC Q221S variant for **7**. For compounds **3-6** the variants VD1-VD5 were examined for highest conversion and for **9** reaction conditions were adjusted.

Biotransformations with VD1-VD5 for conversion of **3-6**, respectively, were tested for promiscuous cyclization of the variants and increased conversion in *dd*H2O at the optimal reaction conditions determined for compound **2**. Conversions for the tested variants VD2-VD5 with substrates **3-6** were similar to the WT (Figure 24). For **4** and **5**, VD1 was clearly the best variant for conversion and for **6**, the conversion was similar for all variants, hence VD1 was selected as catalyst for the preparative approach with the substrates **4-6** (Figure 24). In conclusion, the VD2-VD5 variants are tailored for the conversion of **2** and are not generally suitable for the conversion of **3-5**.

Notably, for **3** conversions were below 1% and not suitable for preparative scale (Figure 24). Nevertheless, product **3a** formed during the biotransformation of **3** with *Aac*SHC variants could be identified as α-Ionone by comparison with the GC chromatogram of the commercially available product standard. This reaction indicated the ability of *Aac*SHC to catalyze the monocyclization in traces**.**
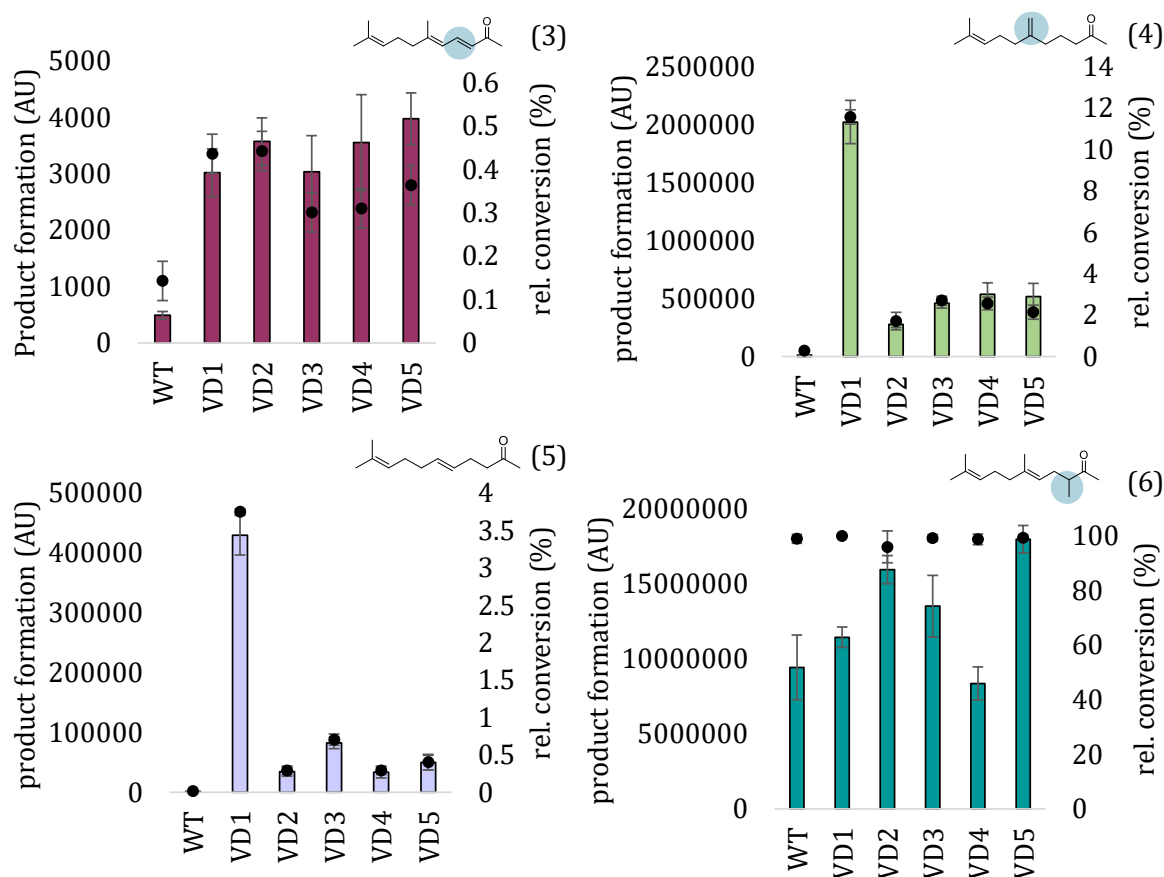
Figure 24. Product formation and rel. conversion of VD1-VD5 with **3-6**, respectively. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H2O, 1 mM substrate, 20 h, 30 °C, 800 rpm.

Due to larger structural differences between **2-6** and **9** the optimal reaction temperature was investigated with *Aac*SHC WT and VD1 in biotransformations of **9**. At 50 °C, product formation and conversion increased for both the WT and VD1 to almost full conversion (Figure 25). Therefore, the preparative biotransformation of **9** was performed with VD1 at 50 °C and otherwise same conditions as for compounds **3-6**.
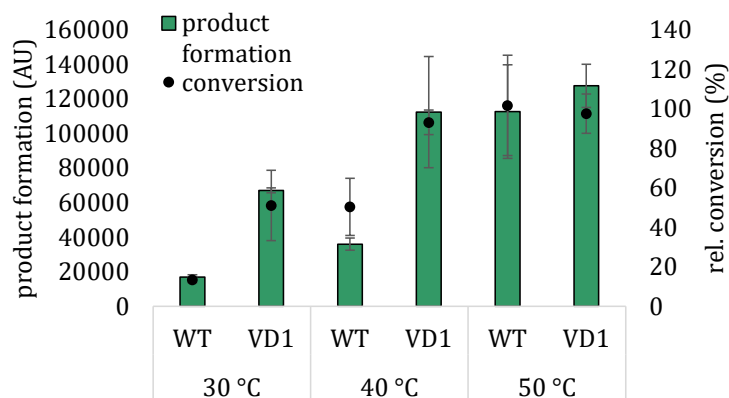
Figure 25. Product formation and rel. conversion of WT and VD1 with **9** at 30 °C, 40 °C and 50°C. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed *Aac*SHC variants dissolved in 1 ml *dd*H$_2$O, 1 mM **9**, 20 h, 800 rpm.

For the upscaling of the reaction products **2a**-**7a** and **9a** using VD1, VD5 and *Zmo*SHC Q221S a protocol with (2-hydroxypropyl)-β-cyclodextrin (CD) as an additive was used and the previously used reaction conditions were applied.[128] CD is known to form micelles, which encapsulate hydrophobic molecules and increased product formation in a previous study.[128] The reaction progress was followed by GC/MS analysis to detect stagnation of product formation and initiate extraction. After successful purification, products at mg scale in yields of 2.7% (2.7 mg **4a**) to 98% (98 mg **6a**) were obtained from 100 mg of substrate for the reaction products, respectively (Figure 26). Except of compound **5a** and **7a** all the other products could be identified and were mainly bicyclic compounds, showing that *Aac*SHC protonation leads to monocyclization and formation of a second carbocation in a concerted reaction to give the bicyclic product after nucleophilic attack on the carbonyl group, as observed previously (Figure 26).[199] In addition, the relative configuration was determined for compound **6a** and either (*R*,*R*) or (*S*,*S*) configuration was observed (Figure 26).
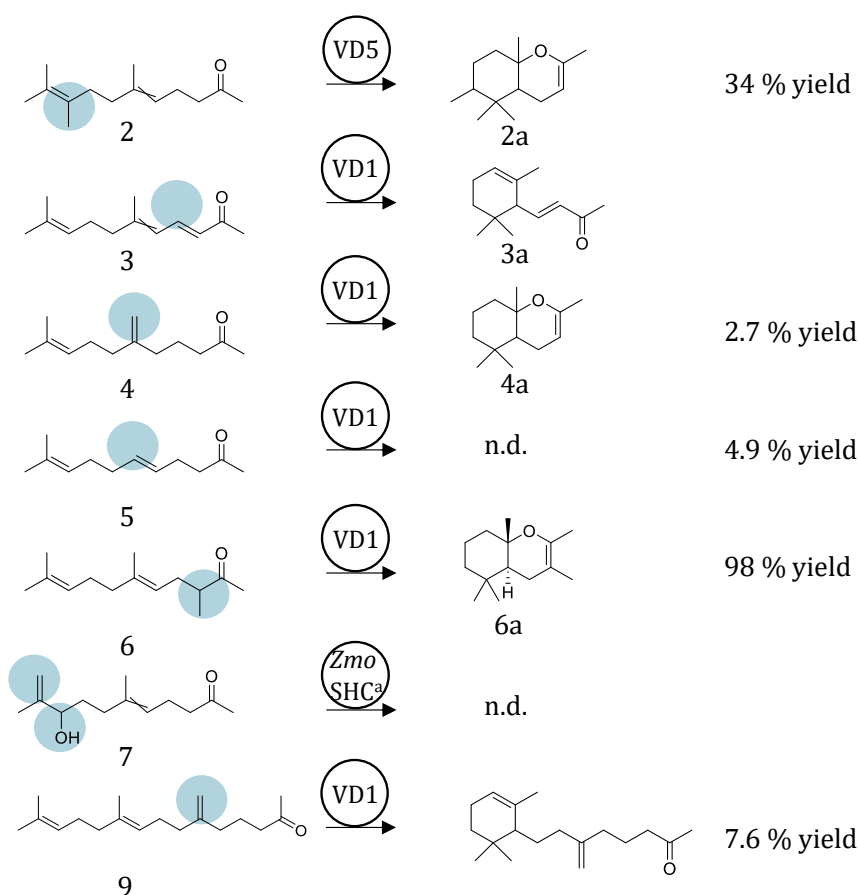
Figure 26. Product structure determination and yield of the preparative scale biotransformations for compounds **2-7** and **9**. Products **3a** and **4a** were identified through comparison to a commercial standard. For compounds **3**, **5** and **7**, the structure of the formed product in preparative scale and the yield of **5a** and **7a** could not be determined due to low conversion or impurities that could not be removed. Relative configuration of **6a** is shown in (*S*,*S*)-configuration. Reaction conditions: 10 g_{CWW} *E. coli* BL21 (DE3) and 1 g of **2** dissolved in 1 l of 100 mM $Mg(NO_3)_2$ or 2 g_{CDW} *E. coli* BL21 (DE3) cells and 100 mg **3-6** dissolved in 100 ml *dd*$H_2$0, respectively, supplemented with 0.4% SDS and equimolar amounts of CD, pH 6, were incubated at 30 °C or 50 °C, 300 rpm.

3.2. Alternative catalysts for 3

The conversion of **3** by variant VD1 and directed evolution of SHC variants appeared as a challenging task and therefore alternative catalysts were explored. To that end, different enzymes with cyclase function were sought and cyclases were identified in the BRENDA enzyme database throughout different enzyme classes and different organisms. From all the different cyclases 57 representatives were selected and their relationship was investigated by phylogenetic analysis.

The obtained phylogeny demonstrates that most of the cyclases are probably not related (Figure 27). However, some enzyme clusters could be discovered. The SHC cluster contained representatives of the group of oxidosqualene cyclases with the ability of

polycyclization and included amyrin synthase, cycloarthenol synthase and lanosterol synthase (Figure 27).[186,201,202] Representatives of the type I terpene cyclases which show primarily activity for mono- and sesquiterpenes were clustered separately (Figure 27).[203] In addition, a cluster of LCYs was found, which catalyze the formation of carotene from lycopene (Figure 27).[204] Different structurally diverse LCYs are known and the two lycopene β-cyclases from *Capsicum annuum* (*Can*LCY-B) and *Arabidopsis thaliana* (*Ath*LCY-B) forming γ-carotene with a β-Ionone end group and one lycopene ε-cyclase from *Arabidopsis thaliana* (*Ath*LCY-E) producing the δ-carotene with α-Ionone end group were selected.[177,184] By their monocyclization reaction in addition to independence from activated substrates and activity for terpenes these LCYs are promising engineering targets for conversion of **3** and were tested for activity.
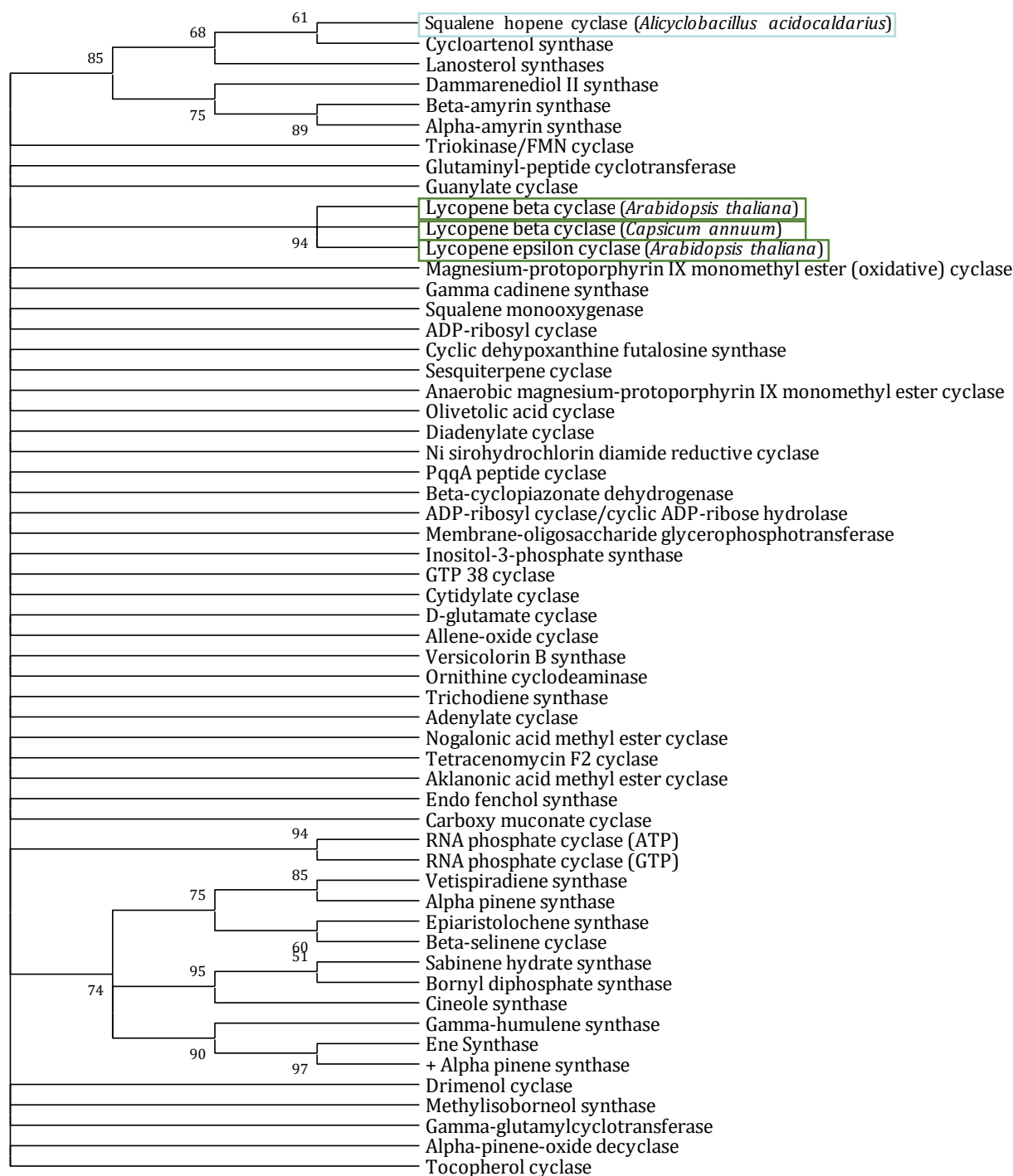
Figure 27. Maximum-likelihood tree of cyclases with *Aac*SHC (blue) and LCYs (green) highlighted built by MegaX.[205] The bootstrap values are given at each branch.

3.2.1. Substrate scope of LCY variants

Little is known on the substrate scope of LCYs. Therefore, the activity of the three LCYs for conversion of **3** was tested first. Second, activity for terpenes with different chain length (C10 – C20) and functional groups e.g., ketones, aldehydes and alcohols or non-functionalized substrates were determined. Reactions of LCYs in whole cell

biotransformations with aldehydes were validated with purified enzyme, because in previous studies the reduction of aldehyde substrates by *E. coli* BL21 (DE3) enzymes was observed.[206,207]
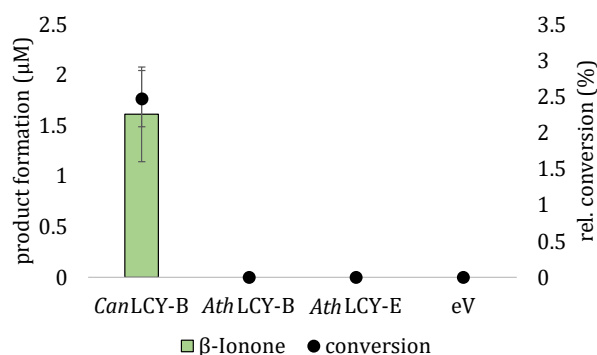
Figure 28. Product formation and rel. conversion of *Can*LCY-B, *Ath*LCY-B and *Ath*LCY-E with **3**. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed LCY homologs dissolved in 1 ml TRIS-Maleate buffer (pH 6.8) with 0.5 mM NADPH, 0.01mM FAD, 1 mM **3**, 20 h, 30 °C, 800 rpm.

The expression and first characterization of the three different LCYs *Can*LCY-B, *Ath*LCY-B and *Ath*LCY-E in whole cells was confirmed by Daniela Ramesohl (University Stuttgart). *Ath*LCY-B and *Ath*LCY-E did not show product formation with **3** as a substrate. Only *Can*LCY-B showed 2.5% conversion to β-Ionone (Figure 28).

Further investigation on the substrate scope of the LCYs resulted in no activity for ketones **1**, **2** and **8** (Table 5). Of two tested aldehydes only compound **12** was converted in whole cells by *Can*LCY-B or *Ath*LCY-E (Table 5). In addition, product formation was detected with *Can*LCY-B in whole cells and the C10 alcohol **14** (Table 5). Notably, this substrate has no conjugated double bonds, was the smallest of the tested alcohols and was poorly converted with *Aac*SHC WT.[82] No product formation was detected with non-functionalized substrates **16** and **17** (Table 5). In biotransformations with purified enzymes the conversion of **3**, **12** and **13** was confirmed, but no conversion of **14** was detected (Table 5, Figure S 6). The formed products were identified by standards and were the monocyclic β-Ionone in the conversion of **3** and the reduced alcohol products geraniol and citronellol for **12** and **13**, respectively. The latter products were unexpected because LCYs were not reported to perform reductions.

In conclusion, **3** was cyclized and **12** and **13** were reduced by *Can*LCY-B. For non-functionalized substrates and alcohols no product formation was measured.

Table 5. Selected terpenes investigated for the conversion with *Can*LCY-B and *Ath*LCY-E and purified *Can*LCY-B. Conversion is indicated by a check mark. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed LCY WT homologs dissolved in 1 ml TRIS-Maleate buffer (pH 6.8) with 0.5 mM NADPH, 0.01mM FAD, 1 mM **3**, 20 h, 30 °C, 800 rpm.

Compound	Structure	Conversion whole cells	Conversion purified *Can*LCY-B
1		✘	-
2		✘	-
3		✓	✓
8		✘	-
12		✓ [a]	✓ [a]
13		✘	✓
14		✓	✘
10		✘	-
15		✘	-
16		✘	-
17		✘	-

[a] converted with *Can*LCY-B and *Ath*LCY-E

3.2.2. Optimization of reaction conditions for the conversion of 3 by LCY-B

To optimize the *Can*LCY-B reaction conditions for the conversion of **3**, different buffers, pHs, detergents and substrate concentrations were tested.

Comparison of product formation in PIPES buffer, *dd*H$_2$O or TRIS (Tris-(hydroxymethyl)-aminomethan)-Maleate buffer, showed highest product formation in TRIS-Maleate buffer (Figure 29). The enzyme was sensitive to the detergents Tween and SDS because addition to the TRIS-maleate buffer decreased the conversion by more than 10-fold, and no increase in product formation was observed when CD (see

3.1.7.) was added (Figure 29). At pH values of 5, 5.5 and 6, similar conversions and product formations were observed (Figure 29). With TRIS-Maleate buffer at pH 6.8 both conversion and product formation halved to 0.7%. At different substrate concentrations and substrate loading the highest conversion was observed at 1 mM substrate concentration, but highest product formation of 14 µM was measured at 10 mM substrate loading (Figure 29).
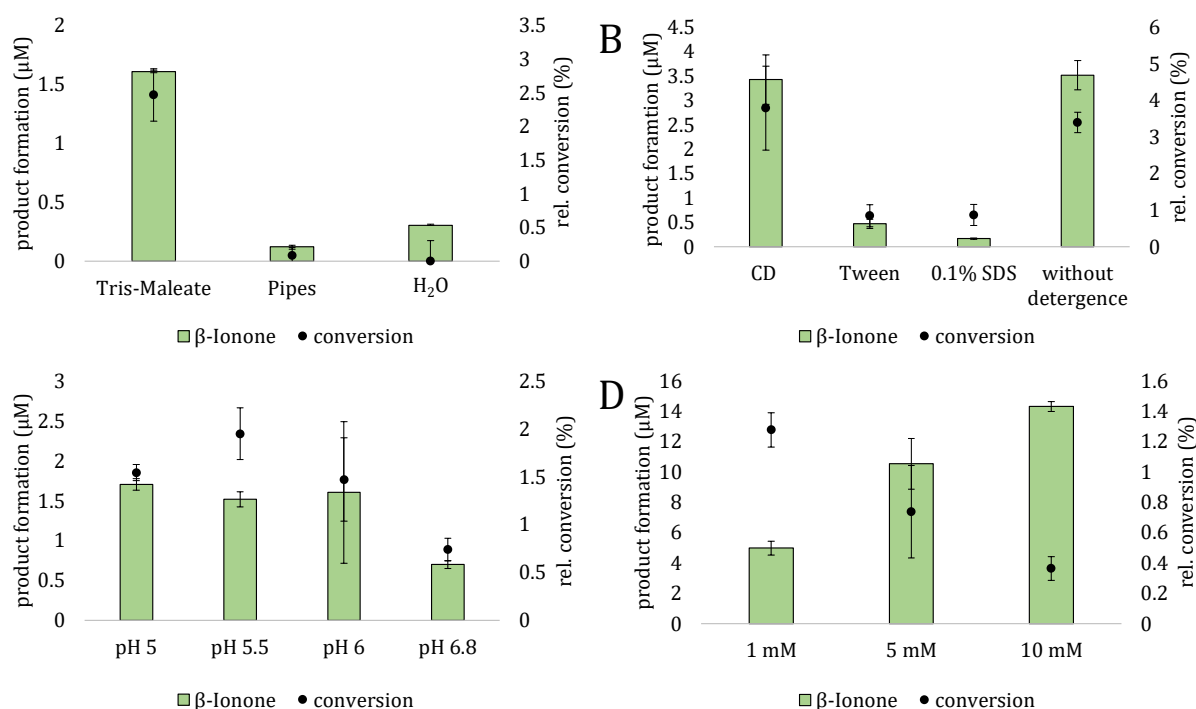
Figure 29. Product formation and rel. conversion of *Can*LCY-B with **3** with different (A) buffers, (B) detergents and addition of CD, (C) pH values and (D) substrate concentration. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Can*LCY-B dissolved in 1 ml buffer, 0.5 mM NADPH, 0.01 mM FAD, 1 mM **3**, unless otherwise stated, 20 h, 30 °C, 800 rpm.

The LCYs were shown to be cofactor dependent in the reaction with the native substrate lycopene and possess NAD(P)/FAD-binding motifs.[173] By application of optimal cofactors screening effort and costs could be reduced. Therefore, investigation of the dependence of the reaction with **3** on cofactor was conducted and higher conversion was observed in reactions with NADPH instead of NADH. No difference with or without FAD in the reaction was determined (Figure 30).

Overall, the best conversion of **3** was observed in TRIS-Maleate buffer at pH 5.5, with 1 mM substrate and addition of NADPH. No influence on product formation was observed with addition of detergents or in the reactions with or without FAD.
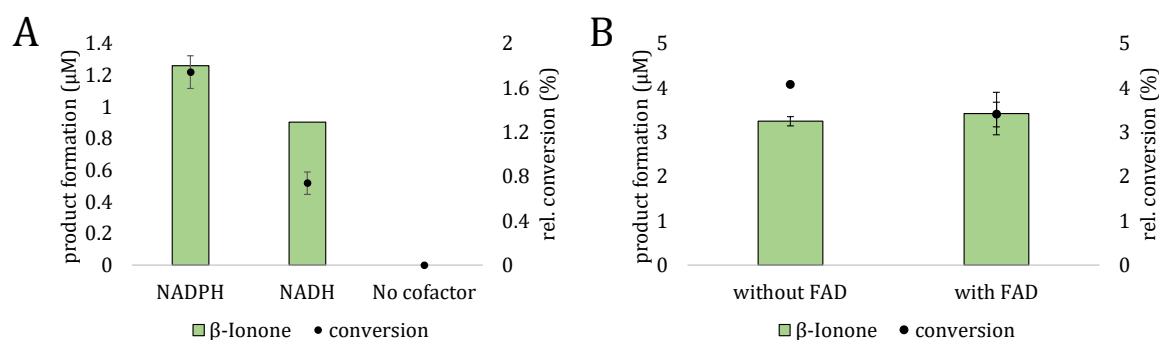
Figure 30. Product formation and rel. conversion of *Can*LCY-B, with **3** and addition of cofactors (A) NADPH and NADH and (B) FAD. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Can*LCY-B dissolved in 1 ml TRIS-Maleate buffer, 1 mM **3**, 0.5 mM NADPH, 0.01 mM FAD, 20 h, 30 °C, 800 rpm.

3.2.3. Identification of key positions for conversion of 3 and 12 by semi-rational engineering of LCY-B and LCY-E

Alternative confinement was used in the next semi-rational engineering approach to increase the conversion of **3** or to change the selectivity for Ionone isomers. With a MSA, the percentual identity between the three tested variants was determined (Figure S 7). This was 78.77% and 36.38% for *Ath*LCY-B and *Ath*LCY-E compared to *Can*LCY-B, respectively. By substrate docking, non-conserved residues between LCY-B and LCY-E, which interact with **3** in the active site (around 5 Å) were identified. Subsequently, *Can*LCY-B was rationally engineered by exchanging the identified amino acid positions to alanine and/or to the amino acids present in *Ath*LCY-E to perform an alanine scan and to identify potential key positions for activity. Amino acids reported in previous studies to be involved in the initiation of the reaction with lycopene were not modified.[184]

For all tested *Can*LCY-B alanine variants and variants with the *Ath*LCY-E amino acid substitutions no activity was measured (Figure S 8). This indicates that the selected positions have a clear influence on the activity and are key positions. Therefore, saturation mutagenesis of the key positions was performed and variant 335 B1 (V335L) with 3.8% higher conversion to β-Ionone compared to the WT was identified (Figure 31). No starting activity for α-Ionone could be detected with the LCY-B variants (Figure 31).

To investigate if one of the positions identified could lead to activity of *Ath*LCY-E with **3** and to change the selectivity for β-Ionone and α-Ionone formation the corresponding amino acid positions were tested by saturation mutagenesis with *Ath*LCY-E. The tested variants showed only traces of β-Ionone (<0.1 µM) and more than 100-fold less β-Ionone

formation compared to *Can*LCY-B (data not shown). For α-Ionone, product formation increased 4.5-fold with the best performing variant being *Ath*LCY-E 359X G3 (S359F) compared to *Ath*LCY-E (Figure 31). This position corresponds to C334 in *Can*LCY-B.

Figure 31. Key positions for saturation mutagenesis of *Can*LCY-B, *Ath*LCY-E to improve conversion of **3**. (A) MSA of *Can*LCY-B, *Ath*LCY-B and *Ath*LCY-E with non-conserved positions highlighted. (B) Product formation and rel. conversion of **3** by (B) *Can*LCY-B and (C) *Ath*LCY-E variants selected by saturation mutagenesis of non-conserved positions with **3**. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed LCY dissolved in 1 ml TRIS-Maleate buffer pH 5.5 with 0.5 mM NADPH, 1 mM **3**, 20 h, 30 °C, 800 rpm.

The selected positions were identified as key positions for the conversion of **3** and the observed amino acid substitutions for *Can*LCY-B and *Ath*LCY-E were bulkier in the variants *Can*LCY-B V335L and *Ath*LCY-E S359 compared to the respective WTs. This indicates that more bulky amino acids in the active site of LCYs are relevant for the conversion of **3**. To elucidate structural changes in more detail the structure of the LCY WTs and variants were modelled and **3** was docked into the active sites. In the LCY structures, amino acid modifications for the improved variants *Can*LCY-B V335L and *Ath*LCY-E S359 occurred in a solvent-exposed β-sheet (Figure 32). The structural changes

by *Ath*LCY-E S359F and *Can*LCY-B V335L are hypothesized to induce improved pre-organization of **3** in the active site, with the methyl group at C6 oriented towards the double bond of the initiation unit (Figure 32). In WT LCYs, this methyl group was oriented in the opposite direction (Figure 32).

This work demonstrated the enhanced formation of ionone isomers by generating an alternative confinement for the pre-folding of substrate **3**. As a result, β-ionone formation was increased for the *Can*LCY-B V335L variant compared to the WT, and the *Ath*LCY-E S359F variant showed moderate α-ionone product formation.
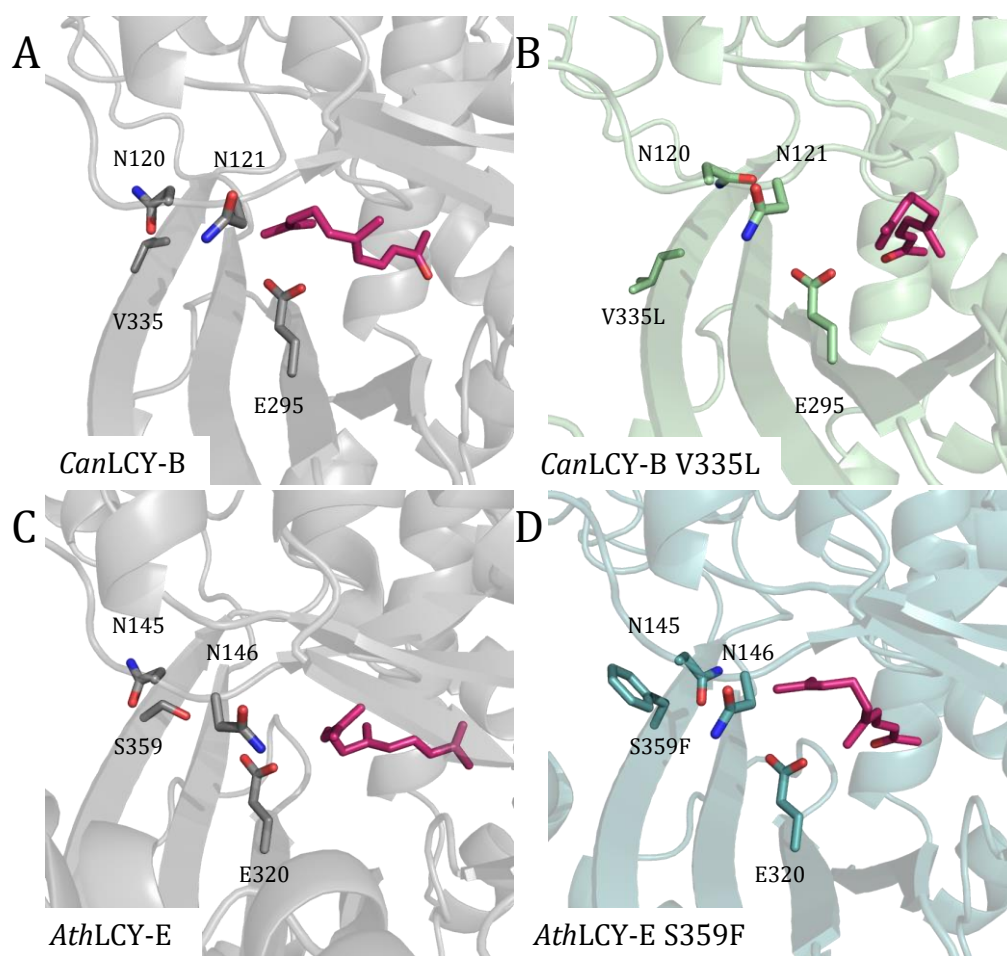
Figure 32. Docking of **3** into the active site of (A) *Can*LCY-B, (B) *Can*LCY-B V335L, (C) *Ath*LCY-E and (D) *Ath*LCY-E S359F. E295 or E320 and key positions mutated during saturation mutagenesis are labeled and shown as sticks.

In addition to **3** reaction conditions for **12** in whole cells were optimized and the key positions identified in this work were tested. Initially activity was observed for *Can*LCY-B and *Ath*LCY-E with product formations of 5.2 µM and 38.9 µM (Figure 33). At increase of the pH from 5 to 6.8, an increase in product formation for both variants with more than 40 µM product at pH 6.8 was observed (Figure 33). No pH above 6.8 was tested because

in previous experiments with **3** activities decreased at pH 6.8. With *Can*LCY-B similar activities were observed with NADH or NADPH and with *Ath*LCY-E, NADPH in the reaction resulted in an increase in product formation of 33% (Figure 33).

With the optimized conditions at hand the key positions identified with **3** were tested with **12** and *Can*LCY-B. It was shown that all the variants increased product formation and the best variant M366S increased product formation compared to the WTs by 2-fold in whole cells compared to *Ath*LCY-E and 17-fold compared to *Can*LCY-B (Figure 33).
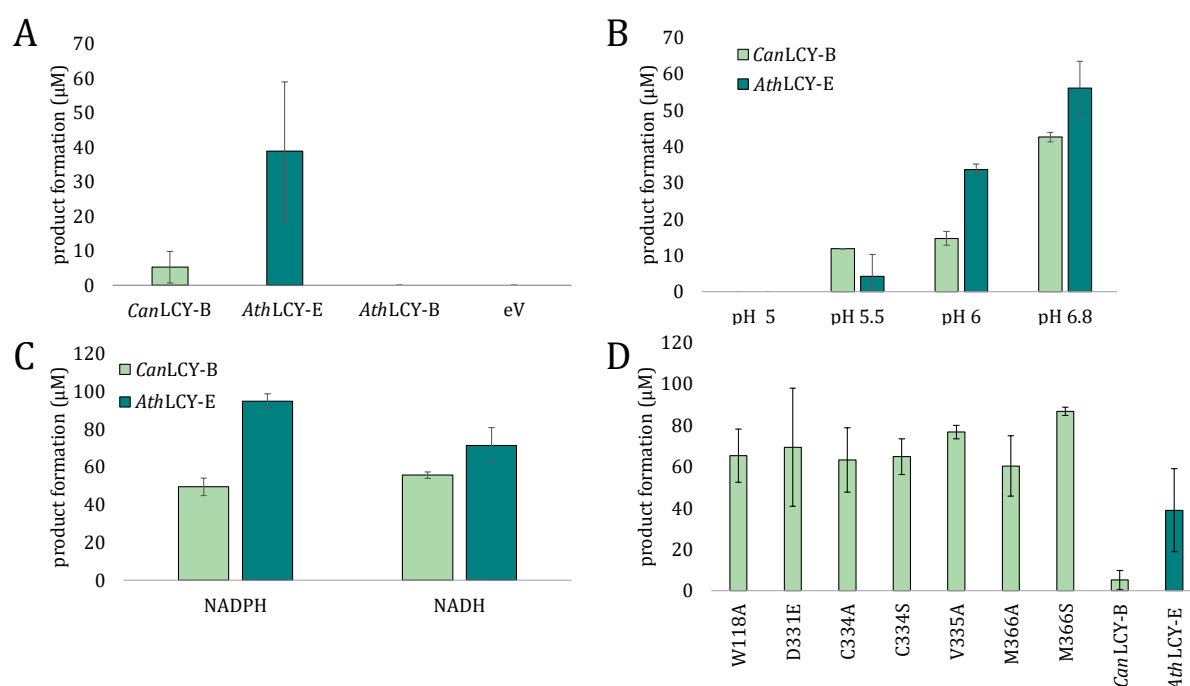
Figure 33. Determination of starting activity with **12** of (A) *Can*LCY-B (mint), *Ath*LCY-E (teal) and *Ath*LCY-B (grey). Product formation and rel. conversion with **12** (B) at change of pH, (C) with NADPH or NADH as cofactor and (D) with amino acid substitutions at key positions for conversion of **3**. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed LCY dissolved in 1 ml TRIS-Maleate buffer, 0.5 mM NADPH of NADH, 1 mM **12**, 20 h, 30 °C, 800 rpm.

The key amino acid positions that influence the activity of LCYs were identified by substrate docking and rational engineering. Saturation mutagenesis was performed on these positions, resulting in the identification of *Can*LCY-B V335L with higher conversion to β-Ionone compared to the WT. Testing of corresponding amino acid positions in *Ath*LCY-E showed increased α-Ionone formation, with the best performing variant being *Ath*LCY-E S359F. The enhanced formation of ionone isomers was attributed to alternative confinement for the pre-folding of substrate **3** induced by amino acid modifications. The

optimized reaction conditions were also tested for **12**, and the key positions identified with **3** were shown to increase product formation, with the best variant M366S resulting in a 2-fold increase in whole cells compared to *Ath*LCY-E WT and 17-fold compared to *Can*LCY-B.

3.3. Conversion of substrates with limited ability for pre-organization

Substrates with limited ability for pre-organization are substrates which are structurally divergent in comparison to the native substrate. An example are very small substrates forming multiple possible intermediates when the native substrate is rather large, or cyclized substrates when the native reaction is the formation of cyclic compounds starting from linear substrates. In this case this limited ability of pre-organization results not only in poor conversions but in a less converged reaction progress and a mixture of products. Monoterpenes are substrates with limited pre-organization when tested with SHCs, because they are often cyclic compounds, and their chain length (C10) is small compared to the native linear substrate squalene (C30). Previous studies with pinenes and *Aac*SHC demonstrated the general feasibility of the reaction but with small activities and selectivity compared to larger terpenes following the isoprene rule.[156,160]

To obtain good activities and selectivity with these substrates the potential of generating alternative confinement for monoterpenes by secondary interactions with the active site of SHCs was tested. In addition, the aim was to stabilize specific transitions states starting from different substrates to enrich the formation of one specific product. Thereby highly confined active sites are generated, which are able to force the reaction progress and rearrangement reactions towards one product.

3.3.1. Detection of starting activity for hydrated compounds

To achieve alternative confinement with *Aac*SHC, commercially available monoterpenes with and without ring stain were selected in collaboration with Julian Ludwig (University Stuttgart). The formation of value-added hydrated compounds was focused because their selective formation is scarce, but in demand.[208]

In total ten different monoterpenes were tested to determine conversion and selectivity (Figure 34). Of the tested monocyclic compounds product formation and conversion in traces was observed only for **18**. With the bicyclic compound **22** the highest conversion of 48% was observed (Figure 34). The selectivity for the main product **20** was 39% and 38% for the hydrated product terpinen-4-ol (Figure 34). The product β-terpineol could

only be measured in traces. For bicyclic **23** and **24** the conversion was 1.9% and 28.6% (Figure 34). The two main products were either the isomerization product **25** and **19** or **25** and **23**, respectively. The hydrated products α-terpineol, borneol and terpinen-4-ol were detected with highest measured α-terpineol selectivity of 9% using **23** as substrate (Figure 34). When using **18** in biotransformations conversion was clearly decreased in comparison to the other substrates with ring strain. The main product could not be characterized or identified. However, terpinen-4-ol could be detected in traces.

Because **22** showed good starting activity with *Aac*SHC and the main product terpinen-4-ol is of high demand for industrial applications it was selected for further enzyme engineering. In addition to the bicyclic substrate **22** the monocyclic substrate **18** with little starting activity was chosen for further engineering and for testing of the potential and limitations of alternative active site confinement design to control carbocation rearrangements for formation of terpinen-4-ol. Substrate **19** showing no activity was tested for evolvability, too, because it is a cheap starting material and conversion to value added products has very promising applications.

	α-terpinolene	limonene	γ-terpinene	α-terpinene	sabinene
Substrate					
	18	**19**	**20**	**21**	**22**
Product distribution		n.d.	n.d.	n.d.	
Conversion	0.05 %[a]	-	-	-	48 %
Main product	unknown	-	-	-	γ-terpinene
Main hydrated product (sel.):	Terpinen-4-ol (21 %)	-	-	-	Terpinen-4-ol (38 %)

	α-pinene	β-pinene	β-camphene	3-carene	4-carene
Substrate					
	23	**24**	**25**	**26**	**27**
Product distribution			n.d.	n.d.	n.d.
Conversion	1,9%	28,6%	-	-	-
Main product	β-camphene	α-pinene	-	-	-
Main hydrated product (sel.):	α-terpineol (9%)	α-terpineol (4%)	-	-	-

Legend:
- terpinen-4-ol
- α-pinene
- β-camphene
- β-pinene
- α-terpinene
- limonene
- γ-terpinene
- β-terpineol (cis-)*
- β-terpineol (trans-)*
- α-terpineol
- n.d.
- n.d.

Figure 34. Tested monoterpenes for product distribution and conversion with *Aac*SHC. The main product and selectivity for the hydrated products are indicated. Reaction

conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC dissolved in 1 ml *dd*H$_2$O, 2 mM substrate, 40 h, 30 °C, 800 rpm.

3.3.2. Optimization of reaction conditions for 18 and 19

Biotranstormations with monoterpenes were performed analogous to the previous ones using substrates with a modified isoprene pattern. However, due to high evaporation of monoterpenes and the partially low conversions, the reaction conditions were optimized to obtain a reliable screening assay via GC/MS. The reaction conditions for the conversion of **18** were optimized with the help of Sandra Schmierer (Hohenheim Agricultural School). Optimization of reaction conditions for conversion of **19** was performed under similar reaction conditions but using variant W312A due to the lack of activity by the WT enzyme with this compound. Previously the Hauer group showed that variant W312A converts **19** to **20** and therefore reaction conditions could be optimized for increased formation of **20**. In the first experiments different buffer conditions were tested with **18** and subsequently the reaction conditions of the biotransformation were adapted for **18** and **19**.

For the improved production of terpinen-4-ol TRIS, MES, citric acid buffer and *dd*H$_2$O were tested with **18** and WT *Aac*SHC (Figure 35). The negative control was performed with variant TKO (triple knock-out) having the catalytic acid D376 and the two adjacent aspartates D374 and D377 exchanged to alanines. No product formation was measured with WT in TRIS buffer and with citric acid buffer terpinen-4-ol concentration was 3.2 µM (Figure 35). With MES and *dd*H$_2$O terpinen-4-ol formation was increased to 4.6 and 4.8 µM, respectively (Figure 35). To determine the pH dependence on the terpinen-4-ol formation the reaction was tested at pH 6, 6.5 and 7. At pH 6.5 and 7 the product formation increased for the WT compared to the TKO variant. At pH 6 product formation was decreased by 75% (Figure 35). Next, lyotropic agents influencing the hydrogen bonding network between water molecules from the Hofmeister series were added to the reaction solution to a concentration of 100 mM. Potassium thiocyanate (KSCN) in the reaction showed highest product formation in comparison to the other tested agents (Figure 35).[197,198] However, with *dd*H$_2$O the highest product formation of 4.5 µM terpinen-4-ol was measured (Figure 35). Reactions in *dd*H$_2$O at pH 7 adjusted with HCl and NaOH resulted in the highest terpinen-4-ol formation and was used as reaction solution.

The following optimization of the biotransformation conditions aimed at further increasing product formation by adapting substrate concentration, reaction temperature and incubation time. Substrate concentrations of 3 mM, 5 mM and 7 mM were tested in a

biotransformation and compared to the TKO variant. With 7 mM and 5 mM substrate an increase in non-catalyzed reaction products, present in the substrate stock, was observed for TKO and the terpinen-4-ol formation by WT was more pronounced with 3 mM substrate (Figure 35). By increasing the temperature from 30 °C to 40 °C up to 50 °C the formation of product decreased (Figure 35). In the case of incubation time, higher product amounts are measured after 40 h in comparison to 20 h reaction time (Figure 35).

In reactions with **19** and the W312A variant or the negative control similar tendencies were observed. High concentrations of the products in non-catalyzed reactions with TKO with higher substrate concentrations, low product formation at 50 °C and higher product amounts after 40 h in comparison to 20 h reaction time were measured (Figure S 9).

Overall, by using H_2O at pH 7 higher product formation was obtained in comparison to standard conditions with citric acid at pH 6. Additionally, biotransformation for 40 h at 30 °C with 3 mM substrate leads to better distinction between enzyme catalyzed and non-catalyzed reactions with the TKO variant. The following biotransformations were performed with *dd*H_2O pH 7 for 40 h at 30 °C with 2 mM **18** or 3 mM **19**.
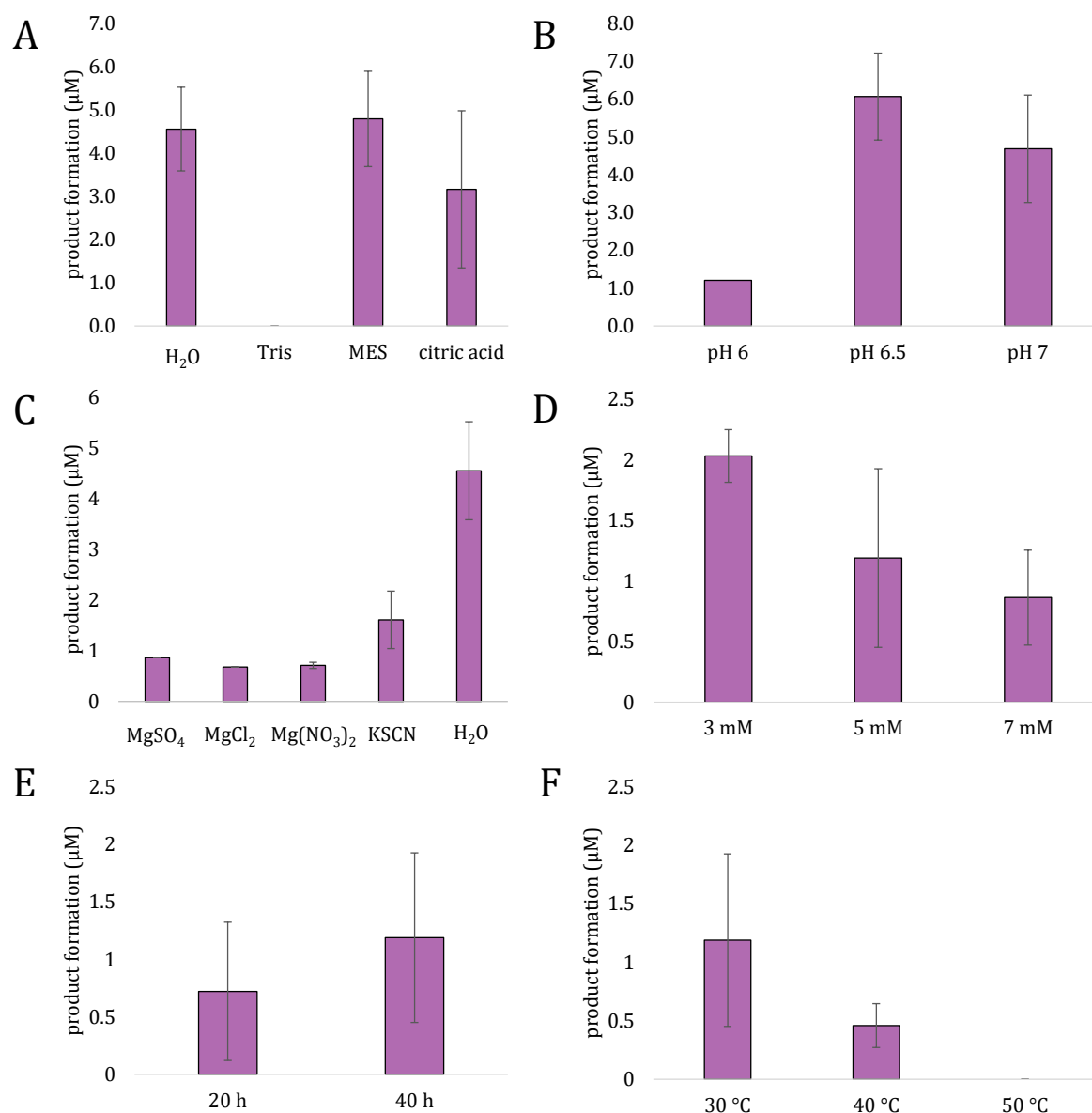
Figure 35. Product formation and rel. conversion of *Aac*SHC with **18** at change of (A) buffer, (B) ionic additives (C) pH values, (D) substrate concentration, (E) biotransformation time and (F) temperature. Error bars represent the standard deviation of duplicates. Ionic additives were selected from the Hofmeister series.[197,198] Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC dissolved in 1 ml buffer or *dd*H$_2$O and 3-5 mM substrate.

With the optimized reaction conditions at hand to allow efficient screening of *Aac*SHC variants during iterative saturation mutagenesis, monoterpene **18** was screened in 96 DWPs, but the product formation was below the detection limit and no product formation was measured. The screening method was adjusted by using 24 DWPs to increase the expression volume from 1 to 4 ml and obtain more cells. Moreover, high volatility of the monoterpenes was observed in catalyzed and non-catalyzed reactions when injected directly from plate (Scheme 1).[209] Applying different sealing materials (Polycarbonate or

Silicone) did not reduce evaporation. Therefore, the biotransformation was extracted in inlets and transferred to airtight GC vials (Scheme 1). By this method a C_v of 23.5 could be established. The C_v indicates the variation throughout the plate and implies that approximately 50% increase in product formation is necessary to determine an improved variant, reliably.
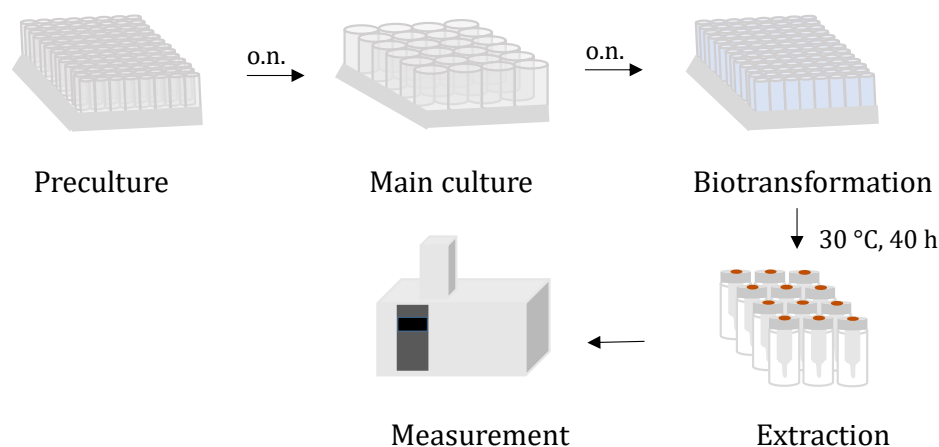
Scheme 1. Adjusted screening procedure for monoterpenes.

3.3.3. Engineering of *Aac*SHC towards terpinen-4-ol formation

The aim was to engineer *Aac*SHC towards terpinen-4-ol formation and to improve selectivity and efficiency starting from three different substrates by controlling the direction of carbocation rearrangement.

With the screening method at hand, the available in-house library was screened with **18**, **19** and **22** to detect a variant with increased formation or selectivity for terpinen-4-ol. With **18,** variant Y420W/G600L (VT1) increased product formation by 4-fold and increased selectivity by 13% compared to the WT (Figure 36). Starting from **22** variant T599M (VS1) showed the highest terpinen-4-ol formation of 184 μM with a minor decrease in selectivity compared to the WT (Figure 36). For **19** no activity could be detected.

To further improve the activity, amino acid residues were mutated that interact with the substrate and have either hydrophobic and π-π interaction properties or are bulky amino acids that can block the substrate in a specific orientation in the active site. With **18**, two rounds of saturation mutagenesis were performed and resulted in variants *Aac*SHC V448Q/Y420W/G600L (VT2) and *Aac*SHC Q366V/V448Q/Y420W/G600L (VT3). The highest terpinen-4-ol formation of 2.9 μM was measured with VT2 (Figure 36). The

highest selectivity of 64% was detected with variant VT3 with a 3-fold increase compared to the WT (Figure 36). The TTN was determined to be 3.4-fold higher for variant VT3 compared to the WT enzyme (Figure 36).

With **19** no variant from the in-house library showed activity. A combinatorial approach was used at positions I261, G600 and L607 and aimed at steering the substrate inside the active site via the formation of a carbocation cage. Within the carbocation cage active site residues stabilize the reactive carbocation intermediates formed in the reaction trajectory and can lead to reaction acceleration and selectivity. Docking studies with MD simulated WT were performed to get insights on the substrate binding. In addition, substrate engineering was tested by ether linkage of limonene to hexane. The hexane-tail was assumed to anchor the substrate in the active site and to facilitate its pre-organization similar to linear substrates (Figure S 10). Nevertheless, no activity was measured with the described approaches.

For **22**, one round of saturation mutagenesis with residues adjacent to T599M resulted in identification of variant T599M/G600T (VS2). This variant showed a selectivity of 48% for terpinen-4-ol and a TTN of 18, thereby 11% increased selectivity, 1.84-fold increased product formation and doubled TTN compared to WT (Figure 36).

To investigate the effects of the identified mutations on the structure and pre-organization of the substrates, substrates **18** and **22** were docked into the active sites of the WT, VS2 and VT3.
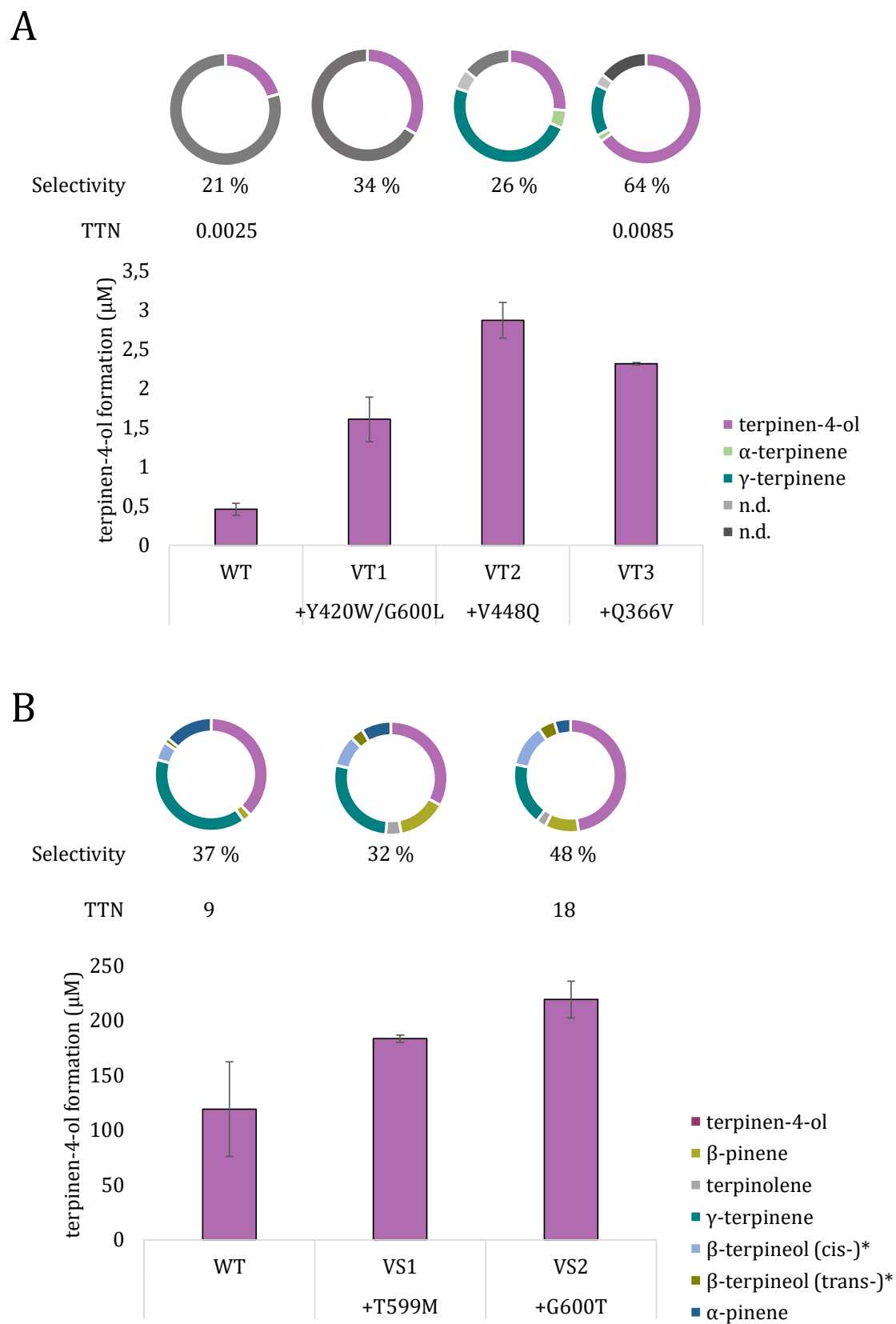
Figure 36. Selectivity, TTN and terpinen-4-ol formation of *Aac*SHC variants generated during saturation mutagenesis for (A) **18** and (B) **22**. Error bars represent the standard deviation of duplicates. Reaction conditions: 1 mg/ml *Aac*SHC variant dissolved in 1 ml *dd*H$_2$O, 1 mM substrate, 40 h, 30 °C, 800 rpm.

3.3.4. Structural analysis of VT3 and VS2

The structural analysis of the generated variants provides explanations for the improved terpinen-4-ol formation. The modelled structures docked with the substrates **18** and **22** are shown in Figure 37.

For monoterpene substrates three residues were identified that could contribute significantly to the stabilization of carbocations by π-π interactions: W312, W489 and F365 (Figure 37). For the variants VT3 and VS2, saturation mutagenesis of these residues did not result in an improved selectivity or product formation. Improved selectivity and product formation was attributed mainly to residues around 3 Å distance to W312, W489 and F365 which were Q366C/V448Q/Y420W/G600L for VT3 and T599M/G600T for VS2 (Figure 37). In VT3 and VS2 the rotation of the side chain of F365 was observed, but not in the WT, possibly indicating an impact of the position of the F365 side chain on the terpinen-4-ol formation with substrates **18** and **22** (Figure 37). For both variants position G600 is modified to bulkier residues leucine and methionine in VT3 and VS2, respectively. It is assumed that G600 is not only involved in positioning of W489 for π-stacking but also responsible for steric effects with the substrates, which can contribute to altered binding of the smaller substrates in the active site. Especially when in combination with Y420W in VT3 interactions with **18** are likely to stall the substrate in a conformation favored for catalysis. In the variant VT2 mutation V448Q was added. This position is adjacent to W489 and to D376 and resulted in an increase in product formation but not in selectivity (Figure 37). The position Q366V added in VT3 adjacent to Y420W could compensate for possible changes in stability and structural organization (Figure 37).

Structural investigation of VS2 indicates a rotation of W489 by mutations T599M and G600T. The docking with VS2 shows that **22** is organized in the active site in the middle of W312, W489 and F365 (Figure 37), whereas in the WT **22** is positioned outwards with larger distance between **22** and W312. Thus, stabilization of carbocation intermediates of **22** could be obtained with higher probability.

In conclusion, structural characteristics in variants VT3 and VS2 were determined such as the presence of bulky residues at position 600 that can influence the arrangement of other aromatic residues in a way to promote the formation of terpinen-4-ol.
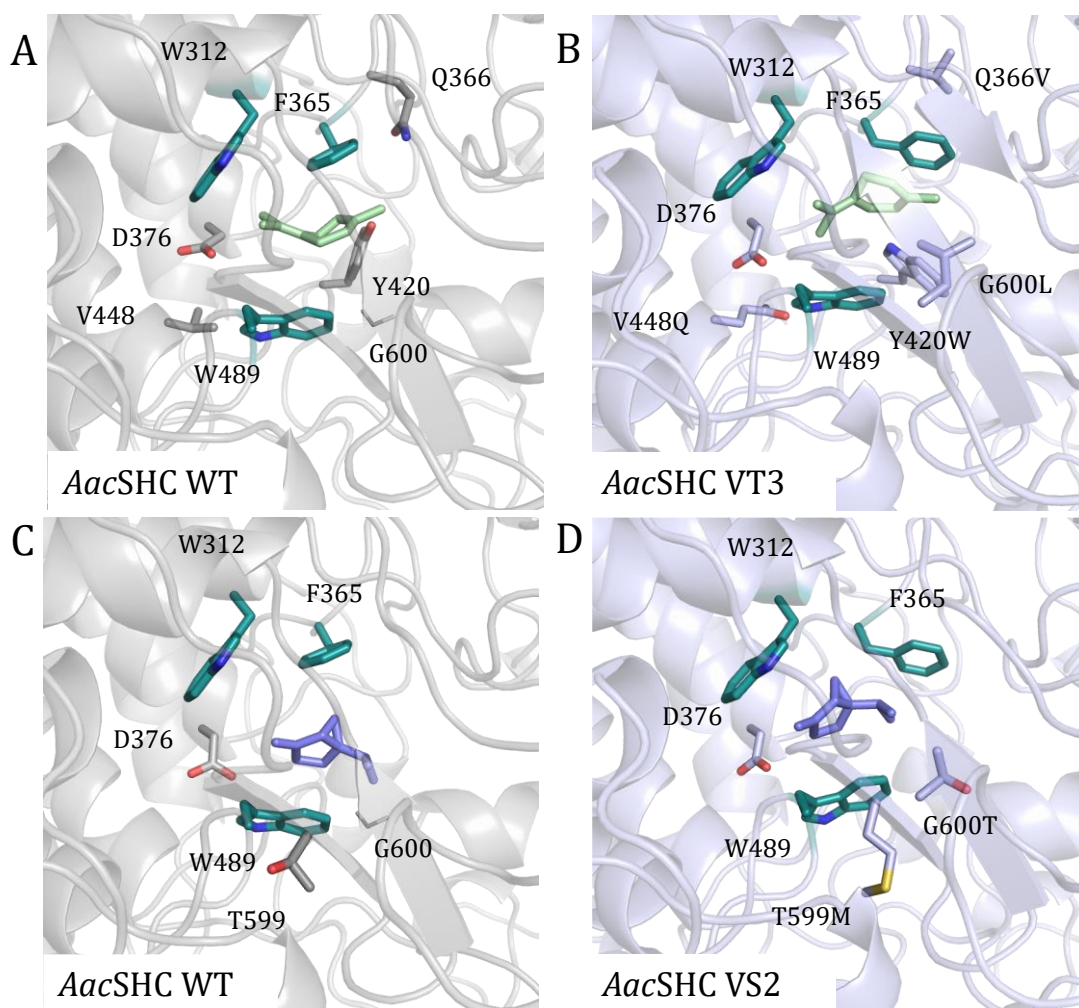
Figure 37. Structure of *Aac*SHC WT and saturation mutagenesis variants VT2 and VT3 with **18** or **22**. Docking of **18** into the active site of (A) *Aac*SHC WT and (B) VT3 and docking of **22** into the active site of (C) *Aac*SHC WT and (D) VS2. D376, residues with the ability to interact via π-π interactions (teal) and key positions mutated during saturation mutagenesis are labeled and shown as sticks.

3.3.5. Acyclic monoterpenes and other substrates converted by the confined active site
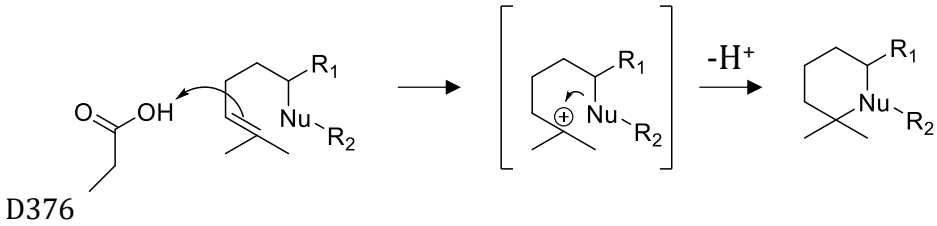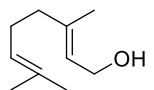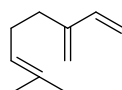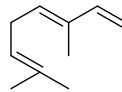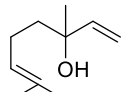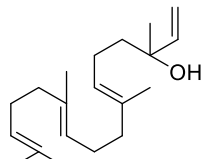
In addition to cyclic monocyclic terpenes, the conversion of acyclic monoterpenes with modified isoprene pattern with *Aac*SHC was investigated. In contrast to cyclic monoterpenes, product selectivity is easier to achieve with acyclic monoterpenes because the active site is configured for conversion of linear molecules. However, the activity of SHCs for conversion of acyclic monoterpenes with C10 chain length was minor or not present. In a previous study **14** showed 1% conversion by *Aac*SHC G600F.[147] The aim was to screen variants for activity with acyclic monoterpenes and monoterpenes with a conjugated π-system. Because of their chain length (C10) acyclic monoterpenes have a high degree of flexibility for positioning in the active site, however the conjugated

π-system results in a more rigid conformation compared to the common isoprene pattern. To meet the requirements of acyclic monoterpenes with conjugated π-systems, the formation of π-π interactions in an alternatively confined active site might be advantageous and would stabilize the compounds for efficient catalysis.

For compounds **28** and **29** with conjugated π-system, conversions of 2.35% and <0.01% were detected with *Aac*SHC WT, respectively (Table 6). The variant G600I/Y420F (VM1) was identified by screening the SHC in-house library and showed the highest conversion of 6.06% with **28** (Table 6). Compound **30** was previously reported to be converted with *Aac*SHC WT and the G600F variant.[147] These results could be confirmed in this work. In addition, a conversion of 0.67% of compound **31** was achieved with variant VM1 in contrast to the WT where no conversion was observed (Table 6).

In summary, variant VM1 was identified as a catalyst with activity in the conversion of **28** and **31**, as an interesting variant for acyclic monoterpenes. Minor activity was observed with **29**. To investigate the origin of the observed activity structural characteristics of variant VM1 were investigated.

Table 6. Overview of acyclic substrates with modified isoprene pattern studied for SHC catalyzed cyclization. Conversion is shown as average ± deviation of measurements in duplicates.

	Substrate	WT rel. conversion (%)	Variant rel. conversion (%)
14		0.40 ± 0.03[a]	68 ± 3.1 (G600F)[a]
28		2.35 ± 0.01	6.06 ± 1.07 (Y420F/G600I)
17		n.d.	n.d.
29		<0.01	<0.01 (Y420F/G600I)
30		0.06 ± 0.01	0.09 ± 0.01 (G600F)
31		n.d.	0.67 ± 0.20 (Y420F/G600I)

[a] results from Hammer *et al.* 2016.[146]
[b] reaction conditions: 0.2 mg$_{CWW}$ *E. coli* BL21 (DE3) whole cells with expressed *Aac*SHC dissolved in 1 ml *dd*H$_2$O and 2 mM substrate, 40 h, 30°C, 800 rpm.
n.d. not determined.

The mutations found in variant VM1, namely Y420F and G600I, have a dual effect on the positioning of the substrate. On one side the bulky side chain of G600I forces the substrate in a position close to the side chain of Y420F (Figure 38). Based on docking studies the conjugated π-system of the substrate might interact with the side chain of Y420F which stabilizes it in a more productive conformation for monocyclization through π-π stacking (Figure 38). A distance between two centroids of 4 Å is possible for π-π interactions.[210] In the depicted docking of VM1 and **28** a distance of 3.9 Å was measured between the conjugated double bond of **28** and Y420F (Figure 38). Thus, the feasibility to form π-π interactions was demonstrated for VM1. With the WT the distance of 6 Å between the conjugated double bond of **28** and Y420 does not support the formation of π-π interactions (Figure 38). Hence, the WT is not providing stabilization of the conjugated π-system and the substrate is therefore less stabilized in the active site. The energetically favored docking state for WT and VM1 depicts **28** with the initiation unit oriented opposite or towards D376, respectively (Figure 38). The distance of the double bond of the initiation unit of **28** to D376 in VM1 is 4 Å and 7.5 Å in the WT (Figure 38). This indicates improved pre-organization of the substrate in the active site of VM1 compared to the WT.
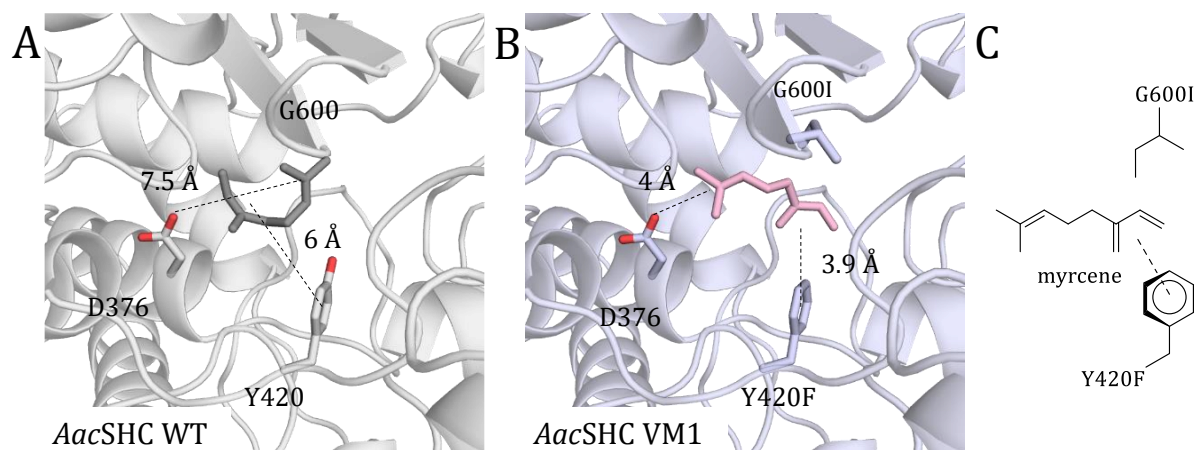
Figure 38. Docking of **28** into the active site of (A) *Aac*SHC WT, (B) VM1 and (C) schematic representation of possible π-π interactions of **28** with Y420F. D376 and key positions identified during screening of *Aac*SHC variants in the in-house library are labeled and shown as sticks.

The engineering resulted in variants with different active site shapes for different functions. The reaction for monocyclic monoterpenes with limited ability to pre-organize in the active site can be directed towards desired products by variants, which presumably stabilize reactive carbocations along the reaction trajectory. For acyclic monoterpene **28** steric interactions and π-π interactions with the variant's active site are hypothesized to

stabilize the substrate's conjugated double bond and result in pre-organization. Overall, alternative confinement was established with acyclic and monocyclic monoterpenes and resulted in starting activity for C10 compounds.

4. Discussion

Confinement contributes to enhanced catalyst performance and improves activity and selectivity. Despite the advantages of confined active sites, they are sensitive towards structural modifications of substrates and pre-organization of the substrate is limited to distinct structural motifs. Subsequently, substrates with hampered or no pre-organization in the active site and subsequent low turnover or selectivity remain challenging to convert efficiently.[3] In Brønsted acid catalysis, site-selective formation of carbocations by protonation of polyprenoids and control of carbocation intermediates is difficult to achieve due to low acidity of general Brønsted acid catalysts, competition between terminal and internal isoprenyl groups and rearrangement reactions.[211] To develop chiral Brønsted acid catalysts and convert unreactive carbohydrate molecules pre-organization is key.[211,212] This pre-organization can vary between confined active sites of Brønsted acid catalysts for transition states resulting in different cyclization products.[213] Within this work, three different approaches to achieve alternative confinement in cyclases SHC and LCY were established, which strongly depend on tightly confined active sites and overcome the described limitations. Productive pre-organization is provided and modified substrates as well as substrates, whose reaction course is difficult to direct, are converted, thereby expanding the substrate versatility and product range. First, starting activities towards challenging substrates with poorly tolerated structural modifications of the isoprene backbone were obtained by tunnel modification and generation of substrate shape complementarity even with initial non-optimized conditions. Second, LCYs were engineered based on residues identified by MSA for the formation of different ionone isomers. Third, SHCs acting as protonases provided access to selective acidic isomerization of small monoterpenes in water by enzyme engineering with focus on amino acids surrounding aromatic residues.

SHC WT homologs and variants were tested towards starting activity in biotransformations with several substrates. SHCs homologs and variants were shown to convert several substrates and despite the observation of promiscuous activity with SHCs, it was shown that SHCs can control the product outcome selectively.[9,159,160] Compared to active sites with more directional interactions, such as hydrogen bonds and salt bridges, the hydrophobic active sites present in SHCs have a degree of conformational variability.[214] Therefore substrates can access reactive and unreactive binding modes within the given structure of the confined active site. Additionally, hydrophobicity is

suggested to contribute to structural flexibility, which allows to find conformations to bind non-natural substrates.[215,216] Finally, hydrophobicity promotes stability towards active site mutations.[217] Nonpolar amino acids can be exchanged by amino acids with similar properties, which disturbs only weak non-bonded interactions. Hence, the SHC's hydrophobic active site was described as a malleable, conformationally flexible and evolvable reaction vessel, equipped with a strong Brønsted acid.[147] Analogous to SHCs, LCYs contain hydrophobic active sites to accommodate the hydrophobic native substrate lycopene.[173] However, for LCYs published data on promiscuity and substrate scope are scarce.

Exploiting the evolvability of SHCs and LCYs to achieve alternative confinement in cyclases increased substrate scope and product formation while maintaining good selectivity. This demonstrates the potential of cyclases to generate interesting products and building blocks for applications in various industries.

4.1. Challenges of substrates with modified isoprene pattern

Modifications of the substrate length from C10 to C35 were tolerated and resulted in the expected cyclization pattern based on cyclization of squalene with SHC. In addition, substrates with epoxide or carbonyl groups at the initiation unit and modifications of the nucleophile at the final cyclization unit (e.g. hydroxyl groups) were converted with moderate to high activities.[155,157,218] Despite this observed promiscuity of SHC, several substrates demonstrated to be particularly challenging. For example, substrates with modification of the backbone, which results in deviation from the common isoprene pattern, were not converted when one or two of the geminal methyl groups at initiation unit were missing.[163] Moreover, the exchange of methyl groups with sterically demanding ethyl groups at C1 or C6 resulted in no detectable conversion.[150,167,170] Abnormal cyclization is observed when the methyl group at C10 is removed, shifted to C11 or when it is replaced by an ethyl group.[163,168] Hence, the active site of cyclases is more compact in early cyclization steps and less compact in later cyclization steps, thus bulkier ethyl groups are rather tolerated in the latter cyclization steps. In this work, analogs of **1**, **8** and **10** were shown to possess similarly challenging backbone modifications at C3 and C6. These 12-18-carbon analogs are smaller than squalene, which has 24 carbons. Especially substrate analogs **2-5** and **11** were poorly or not converted. Therefore, steric modifications at C3 and C6 of the terpene backbone were confirmed to be challenging for

SHC. These observations are true merely for modifications of the hydrocarbon skeleton, because a hydroxy group in **7** instead of a methyl group in **2** at C3 resulted in good starting activity. For substrates with modifications at C9 or C10 a good initial activity was more easily achieved. Based on the docking results the productive pre-organization of the modified substrates **2-5** and **11** was not supported by the active site of *Aac*SHC and the catalysis is impaired due to unfavorable interactions and deteriorated access to the catalytic center with the modified or more rigid substrates. Similar impaired pre-organization in the active site of SHC was previously reported to result in no activity.[128,187]

To obtain starting activity with the challenging substrates **2-5** and **11** alternative active site confinement in *Aac*SHC was created. Beneficial mutations that stabilize reactive binding of hydrophobic substrates in a hydrophobic active site are hard to predict and reactive binding is very sensitive as large changes in activity were observed by minor modifications of the substrates.[82] For example, protonation and substrate selectivity of *Aac*SHC for epoxygeraniol depends on slightly different orientations of the reactive chair-like conformations in the active site.[82]

This work tested different approaches to identify beneficial mutations to generate starting activity. Next to rational engineering based on previous studies, investigating anchoring through hydrogen bonding and loop engineering, a semi rational approach based on tunnel simulations to select positions for subsequent saturation mutagenesis was used. Thereby, the high conversion of 98% from **1** with variant *Aac*SHC G600R was validated, and high conversion with **6** was observed. However, no starting activity for **2-5** and **11** could be detected. With **2**, variant VD1 harboring the G600M mutation was identified, which modified the active site tunnel to generate good starting activity with **2**, **4** and **5**. The ability of single point mutations to redesign the active site of enzymes for new functions is broadly accepted.[219,220] By this variant's promiscuity for the tested substrates the residues along dynamic tunnel structures were identified as key positions to convert modified substrates. The unbranched, flexible methionine is mandatory for the starting activity because all other residues including hydrophobic residues such as valine or isoleucine did not increase product formation. Methionine, among other hydrophobic aliphatic amino acids, is frequently found in proteins and maintains the stability of hydrophobic protein cores.[220,221]

In conclusion *Aac*SHC WT was tolerating modifications in **6** and **9** at C9 and C10 and demonstrated moderate to good activities. Poor or no conversion was observed with *Aac*SHC WT for conversion of **2-5** and **11** due to the compactness of the active site in early cyclization steps and the less compact organization of the active site for cyclization steps later in the cascade. Consequently, not only bulky groups at C3 and C6 resulted in no or minor conversion, but every tested modification of the isoprene backbone around C6 resulted in limited conversion with the *Aac*SHC WT. Hence, the active site around C3 and C6 is not only compact but presumably substrate complementary to the isoprene pattern.

Engineering of tunnel led to the identification of a G600M variant that successfully overcame this limitation with **2**. Additionally, good starting activity with G600M for conversion of **4** and **5** with modification of double bonds was observed. This demonstrated general evolvability towards challenging substrates.

4.1.1. Implications of tunnel modification in variant VD1

Very often, protein engineering addresses the enlargement of the active site to fit bulkier substrates.[222,223] Hence, the changes needed can frequently be estimated by comparing the substrate and the shape of the active site.[24] Small substrates were investigated in this work, which tend to bind in an unreactive manner into the large (native) active site. For small substrates it was postulated that a large unoccupied space is present in the active site of SHCs and results in potentially unproductive binding modes.[82] Usually, rational engineering approaches or saturation mutagenesis is applied to identify positions in the active site or enzyme tunnels, which contribute to better catalyst performance. Growing possibilities of bioinformatics have led to comprehensive computer-based approaches becoming increasingly important in identifying and characterizing positions that impact the catalyst and structural dynamics.[87,143,224] The investigation of mutations on the catalyst-substrate interaction and the influence on dynamic structures is mandatory for data driven enzyme engineering campaigns and prediction of enzyme functions.[188,189,194,225] A comprehensive investigation on the effects of the G600M mutation in variant VD1, which was recognized by tunnel simulation and saturation mutagenesis was conducted within this work. Tunnel simulations were previously applied to modify tunnels for anchoring of substrates e.g. via interaction between a carboxylic acids and arginine in the substrate entrance tunnel.[189] In addition, removal of tunnel bottlenecks was exploited with several enzymes e.g. monooxygenases, hydrolases, halogenases, etc., to improve conversion of more bulky substrates compared to the native

one.[189,194,226] In this work, the variant with improved conversion of substrates with modified isoprene backbone exhibited a bottleneck identified by tunnel simulation. To our knowledge, this is the first time that bottleneck creation in cyclases is attributed to improved conversion and catalysis.

The creation of a bottleneck in variant VD1 resulted in the generation of proximity from the terminal double bond of the substrate's initiation unit to the catalytic D376, thus better positioning of all tested substrates except **9** for catalysis. Condensed distance distribution for substrates **2**, **4**, **5**, **7** and **9** indicates more constrained substrate pre-organization. For **3** and **6**, G600M in VD1 occupies more space in the active site and is suspected to negatively affect catalysis by less productive positioning of the more rigid, conjugated double bond of **3** or the additional methyl group at C9 in **6**. As a result, a broader distance distribution was observed, and the substrates were less compactly pre-organized. Nevertheless, substrates **3** and **6** are positioned in closer proximity to D376. Given the similar conversions of WT and VD1 with **3** and **6**, it is implausible that the positioning was improved in the VD1 variant for these compounds.

Hammer *et al.* observed a correlation between the distances of 2.9 Å or less from the terminal isoprene double bonds of compounds to the catalytic center and the catalytic activity for unnatural substrates.[146] The distances measured for the docked substrates in this work are all above 3.4 Å. This ambiguity may indicate an exception of the reported correlation due to the generated reduction of the tunnel radii or the fact that the distances were measured with substrates docked to the enzyme structure obtained from the MD simulation. In order to obtain adjusted distances MD simulations including substrate and the enzyme would be necessary.

Nevertheless, the decreased distance and condensed conformations of the initiation unit indicate better positioning for modified substrates **2**, **4** and **5**, which were more affected than the other substrates in the biotransformations with VD1.

4.1.2. Substrate complementarity by saturation mutagenesis with 2

Variants with improved conversion of compound **2** obtained during saturation mutagenesis with VD1 demonstrated evolvability of VD1 and generated alternative confinement for conversion of **2**. Initially, residues around G600M were mutated by saturation mutagenesis to compensate for possible changes in stability and structural organization caused by G600M. The added mutations R488P, F605C, Y495D and S455V

increased the TTN stepwise. Addition of mutations Y495D and S455V around C9 of **2** demonstrated the highest impact resulting in overall 1174-fold increased TTN compared to the WT. The amino acids located around C9 of **2** were selected to generate substrate shape complementarity to fit the methyl group. The results demonstrate a high impact of engineered shape complementarity on the catalyst performance regarding TTN and conversion. Specific interactions, e.g., changes in electrostatic effects were not further investigated, but might contribute to the observed improvement of the catalyst.[219] Previous studies reported the increase in efficiency by 5- to 10-fold of a Diels-Alder reaction by an antibody catalyst as consequence of shape complementarity by introducing site-directed mutations that improved packing interactions.[24] Additionally, shape complementarity was successfully used to afford high rate accelerations in artificial enzyme design.[23] The presented multifold increased TTN highlights the efficiency of the shape complementary approach in combination with the tunnel modification to engineer SHCs and increase their catalytic efficiency.

The origin of the improved catalytic performance of VD1 and VD5 was traced back to altered catalytic efficiency k_{cat}/K_m rather than k_{cat}. Steady-state kinetics can be described by the Briggs-Haldane approximation with K_m as $(k_2 + k_{-1})/k_1$ derived from the Michaelis-Menten equation. Due to only small deviations observed in the k_{cat} and the increased k_{cat}/K_m for the variants an effect on the k_2 rate constant, which describes the ES complex dissociating to give free enzyme and product, is not evident in the variants. Therefore, the engineered variants may influence the dissociation of the ES complex to enzyme and substrate or vice versa, thus decrease k_{-1} or increase k_1, respectively. Thereby the substrate affinity and binding would be improved. Further evidence for improved binding in the active site was observed in the docking analysis. The importance of the G600M mutation for activity is highlighted herein, resulting in a 61-fold increase in k_{cat}/K_m. Its significance for the catalyst's performance was evident in the biotransformation with individual single and double variants of the VD5 mutations G600M, R488P, F605C, Y495D and S445V. Here, G600M or combinations with G600M in double variants showed the highest conversion. Minuscule or no product formation was detectable with the other single mutations.

Substrate conversion was also improved by optimizing reaction conditions. A 6-fold increase in conversion was obtained by optimization of reaction conditions in whole cell biotransformations with VD3. Initially, the conversion could be increased by using *dd*H_2O

instead of the tested buffers. This was unexpected, since the performance of enzymes depends on maintaining a constant pH within the range tolerated by the enzyme. Based on previous enzyme studies on pH dependence and the published pH optimum of AacSHC at pH 6.0, the catalytic performance of AacSHC does not improve in ddH$_2$O, presumably.[199,227] More likely, the cellular structure and the membrane were disrupted and therefore the catalyst is more exposed and the diffusion of substrate/product to the enzyme is facilitated. Further, addition of different ions selected exemplarily from the Hofmeister series showed a positive effect on conversion.[197,198] The AacSHC variants performed best with **2** in 100 mM Mg(NO$_3$)$_2$ at pH 6. These ions, among others of the lyotropic series, are postulated to increase the solubility of hydrophobic molecules in water by reducing the viscosity of water and decreasing the order of water.[228] This may lead to better substrate accessibility in solution.

For variants VD4 and VD5 amino acid substitution Y495D is close to C9 of docked **2** and therefore increased the variability of distances to D376 and varied positioning of the initiation unit. Particularly, mutation Y495D, is assumed to generate void space and thus allows improved positioning of the sterically demanding methyl group of **2** at C3. Subsequently, the complementarity between substrate and the active site of VD4 was improved, but more possible conformations in the distance analysis were observed. In the final variant VD5 the distance distribution was condensed again to anchor the initiation unit. This presumably improves pre-organization of the substrate. In addition, VD4 and VD5 showed similar conversions, but very different TTNs. The TTN depends on enzyme concentration and the enzyme concentration of VD5 was lower compared to VD4 in this reaction. This indicates the higher efficiency of VD5 for the reaction with **2** compared to VD4 and supports the hypothesis of a more productive pre-organization of **2** in VD5. Therefore, generation of shape complementarity with VD5 in order to evolve for improved catalyst performance is demonstrated. Furthermore, Y495 is presumed to coordinate a water molecule in the active site to establish hydrogen bonding to D376. Together with H451, Y495 is involved in in the positioning of the D376 in anti-orientation, which was proposed to increase the acidity and contribute significantly to the reactivity.[82,85] However, mutation Y495D did not decrease activity as expected, but resulted in the improved variants VD4 and VD5. Based on previous findings in our group, the aspartate side chain of mutation Y495D in VD4 and VD5 could act as proton shuttle or catalytic acid. It was previously observed that AacSHC activity is completely abolished

only when all three aspartates within the prevalent DxDD motif are replaced by alanines.[156] Residual activity was observed with β-pinene when the D376A variant was used, suggesting that aspartates other than D376 (D374 and D377) can act as catalytic acid.[156] To investigate the possibility of mutation Y495D to act as catalytic acid, the distance of this residue to the terminal double bond of the initiation unit was determined and a distance of 9.5 Å was measured. For D374 and D377 the distances were 4.4 Å and 8.6 Å, respectively. Therefore, no definite conclusion can be drawn as to whether Y495D serves as an alternative residue for Brønsted acid catalysis. QM/MM modelling and structural studies could address the impact of this mutation on e.g., hydrogen bonding in future engineering campaigns to elucidate the origin of the increased activity.

Despite the reported achievements by engineering shape complementarity for the conversion of **2** to bicyclic product, limitations of this approach were observed for the conversion of **3** to monocyclic product and the monocyclization of **2**. The obtained engineered *Aac*SHC variants PI3-PI5 showed only minor improvements in conversion, thus the approach to generate shape complementarity is not suited to engineer *Aac*SHC for conversion of **3**. The rigidity of the conjugated double bonds could impede the productive organization and hinder the evolvability towards conversion of **3** with *Aac*SHC. In a study by Eichenberger *et al.* a similar conclusion was drawn after a comprehensive SHC screening panel with several homologs failed to convert **3**.[187] In the context of monocyclization of **2** the formation of monocyclic dihydroionones was described starting from (*E/Z*)-**1**. In these studies, the SHC variants could differentiate between the *E*- and *Z*-isomer. The *E*-isomer was converted to a bicyclic ether (98%) and the *Z*-isomer yielded the bicyclic (60%) and the monocyclic product (10%).[187] Schneider *et al.* showed that variants, which anchor the substrate by hydrogen bonding increased the formation of monocyclic product.[128] In the present work, no conversion of the *Z*-isomer of **2** could be detected and the generated variants to facilitate hydrogen bonding with the substrate diminished bicyclic product formation, while no or minor amounts of monocyclic product was detected. Based on this high selectivity for the *E*-isomer of **2**, observed with the G600M variant, pre-organization of the *Z*-**2** isomer was not supported. Therefore, the engineering of activity towards monocyclization was not feasible with G600M present and additionally no good starting activity with rationally engineered variants could be obtained.

For monocyclization with rigid compound **3** and engineering of hydrogen bonding to convert the *Z*-Isomer of **2**, generation of shape complementarity was not effective to improve conversion. However, the generation of shape complementarity with *E*-**2** showed a high increase in TTN of 1174-fold and demonstrated great potential for evolving efficient catalysts with alternative confinement in *dd*H_2O. The generated variants show the highest activities when residues around the C3 methyl group of 2 were substituted to generate shape complementarity with high selectivity for *E*-**2**. Addition of salts of the Hofmeister series improved catalysis presumably by increasing the solubility of the hydrophobic substrates. Taken together, the catalysts demonstrate not only the great increase in TTN, but also a highly selective active site for the *E*-isomer of **2** to yield the bicyclic product.

4.1.3. Mechanistic consequences for the characterized products

The formation of building blocks from the tested modified substrates demonstrates the potential for applicability of *Aac*SHC variants in preparative scale synthesis. The formed products in mg scale were mainly bicyclic compounds, indicating that the substrate is bound in an all *pre*-chair state required for the continuous cascade to bicyclic product. This all *pre*-chair conformation is observed already in the dockings of the tested substrates **2**-**7**,**9** and **11** in the active sites of the *Aac*SHC variants. Protonation by SHC would lead to an asynchronous, concerted reaction to monocyclization and formation of a second carbocation to yield the bicyclic product after nucleophilic attack of the carbonyl group, as previously described for **1** (Figure 5).[146,153] For **4** and **5** variation of the cyclization mechanism due to structural challenges was anticipated. Surprisingly **4a** was observed as product of **4**. This product is the same as reported for **1**. Due to the structural variation in the double bonds of **4** compared to **1**, the common cyclization progress is not possible. Thus, the cyclization mechanism or the carbocation formation for this compound vary and the involvement of a hydrate shift was assumed. However, the exact reaction mechanism needs further investigation. Product **5a** could not be detected with the employed methods after extraction, because the activity of VD1 was not sufficient to form enough product for structure characterization. The pre-organization is expected to be mandatory to allow the concerted reaction with variant VD1 because no intermediary tertiary carbocation can be formed at C6 of **5**.

The relative stereochemistry was determined for **6a** to be either (*R*,*R*) or (*S*,*S*) configuration. The cyclic product of **1** and stereochemical information for cyclization by

*Aac*SHC is known from previous studies.[187] In addition, it was observed for SHCs that the absolute configuration of the newly generated stereocenter resulting from the first cyclization step is expected to be defined by the configuration of the double bond of the substrate. Hence, asymmetric induction is determined solely by the geometry of the double bond.[187] Based on the confined active site which forces a similar and constricted pre-organization for substrates **2-7**, **9** and **11** as observed in docking studies and presumably binds only the *E*-isomer in productive conformation, the absolute *S,S*-configuration observed in **1** will likely apply for **6a**. Similar effects are expected for **4a**, due to the confined active site and the observed pre-organization of the substrate. The stereochemistry of further products could not be determined and no comment on the stereochemistry can be made.

Overall, the pre-organization of the substrate in the active site is key for the conversion using cyclases. Establishing proximity and fine tuning of secondary interactions with the active site are mandatory for expansion of the substrate scope and enable the product formation with modified substrates in preparative scale. The hydrophobic active sites of the tested cyclases are tolerant for modifications, because of weak interactions with the substrate (e.g. van der Waals or electrostatic interactions) creating a good platform for engineering. Key residues to fix the substrates in a productive conformation probably promote the stereoselective formation of either (*E,E*) or (*Z,Z*) bicyclic product.

4.2. Lycopene cyclase

The challenge to convert compound **3** efficiently was represented by previous studies where no activity towards **3** could be determined with an expanded SHC library and by this work's limited success in increasing activity by saturation mutagenesis with *Aac*SHC.[160,187] It was argued that the double bond and putative steric effects of the conjugated γ,δ-double bond impose constraints on the substrate backbone conformation, which would not result in productive pre-organization.[160,187] Further enzyme engineering strategies to improve the *Aac*SHC conversion, e.g., combinatorial approaches, random mutagenesis etc., are promising strategies to improve the conversion of **3**. Yet, these strategies were not tested, due to lacking starting activity. Instead, alternative catalysts, of which native substrates are linear terpenes with conjugated double bonds, such as lycopene, were targeted for generation of starting activity with **3**. The enzymes capable of cyclization of lycopene to carotenes are LCYs and are, based on phylogenetic analysis, not

related to SHCs. This was expected due to the origin of the enzymes in prokaryotes and eukaryotes, the different roles in either hopanoid or carotenoid biosynthesis, the different tetra- and triterpene substrates and the different function of SHC and LCY to alter the structural order of phospholipid membranes or as catalysts for accessory pigments, respectively.[172,173,181,199] Moreover, the thermo- and acidophile *Aac*SHC represents an isoprene-based cyclization chemistry that appeared early on Earth.[155] LCYs occur in plants, which evolved in a later period on earth.

Of the three LCYs tested only *Can*LCY-B was able to convert **3** with good starting activity. The formed β-Ionone was the expected product formed by LCY-B, because of the formation of the corresponding β-ring with the native substrate lycopene.[173] The other tested LCYs, *Ath*LCY-B and *Ath*LCY-E, did not show any starting activity with **3**. The latter two enzymes are only distantly related to *Can*LCY-B as was previously shown in a neighbor–joining phylogenetic tree of LCYs.[204] LCY-E and LCY-B were clustered in two different groups and the phylogenetic relation between *Ath*LCY-B and *Can*LCY-B was distant within the LCY-B group.[204] Therefore, the starting activity observed with *Can*LCY-B but not with *Ath*LCY-B and *Ath*LCY-E could be explained by differences during natural evolution. Additionally, the sequence analysis provided identities of 78.77% and 36.38% for *Ath*LCY-B and *Ath*LCY-E to *Can*LCY-B and indicates strong differences between the amino acid sequences. Structural data on the LCY cyclases was not available and based on the sequence-based difference and the modelled structure of the enzymes it was suggested that the pre-organization of **3** differs between the tested LCYs.

4.2.1. Substrate scope and reaction conditions of tested LCYs

Investigation of the substrate scope revealed that next to **3**, **12** and **13** were converted by *Can*LCY-B. Additionally, *Ath*LCY-E converted **12**, but with lower conversion. No product formation could be determined with the other substrates or *Ath*LCY-B. Conjugated double bonds are present in **3** and **12**, but not in **13** and carbonyl groups were present in all the converted compounds. Alcohols **10**, **14** and **15** and unfunctionalized terpenes **16** and **17** with conjugated double bonds were not converted. It was reported that the lycopene cyclases from cyanobacteria, CrtL, are dependent on desaturation of the C7-C8 or C7'-C8' carbon bond for cyclization.[229] However, fungal bifunctional lycopene cyclases are reported, which convert lycopene and phytoene that has mostly saturated bonds and no 7-8 double bond as in lycopene.[230] Nevertheless, due to the reported evolutionary origin of LCYs from CrtL similar restrictions in LCYs were expected.[174] This work shows that in

the case of the respective substrates with no product formation no 7-8 double bond is present. When a double bond was present in **3, 12** and **13** cyclization could be observed. Notably, even compound **13** with a non-conjugated carbonyl double bond present at C7, was converted. Thus, the conjugated double bond in the substrate is assumed to be facultative but not mandatory. To gain more insights on the significance of the 7-8 carbon double bond more substrates could be tested with similar backbone and different chain lengths. Nevertheless, the observed dependence on the 7-8 carbon-carbon bond in CrtL was validated for LCYs and bifunctional activity for mostly unsaturated substrates in addition to lycopene as reported for fungal lycopene cyclases, was refuted for the tested LCYs.

In contrast to **3** the product of the reaction with *Can*LCY-B and **12** and **13** were not cyclic compounds, but reduced products. How the size of the substrates and the aldehyde group instead of the keto group might be involved in this different reaction mechanism has to be investigated further. To our knowledge this is the first-time reduction of aldehydes in the reaction with purified LCY-B was observed.

Biotransformations with whole cells and purified enzyme gave different results for the conversion of **13** and **14**. For **13** and **14** in whole cell biotransformations, reactions with enzymes from the *E. coli* BL21 (DE3) strain were assumed to be involved in the conversion of the substrates and not LCYs. The aldehyde reduction of **13** is expected to occur due to the *E. coli* BL21 (DE3) enzymes with higher catalytic efficiency or affinity for the substrate compared to the tested LCYs. Subsequently, in *E. coli* BL21 (DE3) whole cells no LCY activity could be measured. However, with purified LCY activity could be observed. Several reduction reactions of aldehydes in *E. coli* BL21 (DE3) whole cell approaches were observed previously, and strains were engineered to reduce aldehyde reduction.[206,207]

The reaction optimization to obtain higher β-Ionone concentration and good conversion of **3** in whole cells was partially consistent with references on the conversion of lycopene by purified LCY-B, e.g., the buffer system.[184] However, differences were documented in pH optima, addition of detergence and cofactors. Various studies suggest optimal pH of 7 or 7.6 for LCY-B and a pH of 6.8 for *Can*LCY-B with lycopene.[164] The biotransformation with **3** demonstrated pH tolerance at pH 5 to 6 with highest conversion observed at pH 5.5. Detergent and FAD did not improve conversion or product formation contrary to previous reports.[176] Hence, the whole cell system and/or the smaller size of **3** in comparison to

lycopene is assumed to result in altered reaction conditions for optimal conversion of **3**. Even though it was postulated that FAD is required for the reaction our results suggest the opposite and can be supported by results with a *Erwinia uredovora* lycopene cyclase (CrtY).[231] With this enzyme FAD addition without NADH or NADPH did not result in higher product formation, but NADH or NADPH in the reaction resulted in 10-fold increased activity.[231] A similar trend is observed in the present work, and no activity without addition of NADPH or NADH could be reported.

4.2.2. Saturation mutagenesis with LCYs

The positions selected for saturation mutagenesis were identified by MSA and docked ligand **3** in the LCY active site (around 5 Å). The proposed catalytic acid is E295, because mutations of this residue to E295A, E295K and E295R completely abolished activity.[184] The identified variants with increased product formation for β-Ionone and α-Ionone were *Can*LCY-B V335L and *Ath*LCY-E S359F, respectively and are too distant to directly interact with E295. *Ath*LCY-E S359F corresponds to position C334 in *Can*LCY-B. Hence, the location of the residues is within the same helix in the enzymes. The phenylalanine in comparison to S359 in *Ath*LCY-E WT has a longer carbon chain and could make the active site more compact and therefore better to accommodate the substrate. Possibly, the corresponding cysteine in *Can*LCY-B WT occupies more space in the active site, too and promotes the conversion of **3**. Similar tendencies towards higher conversion with the larger amino acid leucine instead of valine at position 335 in *Can*LCY-B were observed. For both variants *Can*LCY-B V335L and *Ath*LCY-E S359F the modification of the potential amino acid network with N120 and N121 (numbering of positions according to *Can*LCY) based on the simulation and the docking of **3** was considered. These modifications are capable of improving pre-organization of the substrate for catalysis. To investigate possible interactions of N120 and N121 and V335L or S359F for *Can*LCY-B and *Ath*LCY-E with the catalytic acid or the substrate, a crystal structure would be necessary.

In previous studies, a multiple sequence alignment could be used to identify amino acid substitutions with the potential to modify the function of LCYs. Alignment with a bi-ε-cyclase from *Lactuca sativa* var. romaine and the *Ath*LCY-E was used to identify a key position to transform *Ath*LCY-E into a bi-ε-cyclase. The key position for this molecular switch was L448H.[232] This demonstrates the general potential of this method for modifications of LCYs. However, for this work residue L448 was not considered for saturation mutagenesis, because of a 14 Å distance to **3** that was observed in the docking

studies. Minor involvement of this residue in pre-organization of **3** and no effect on catalysis was assumed. However, the multiple sequence alignment demonstrates great potential for future engineering campaigns.

More insights on the activity of LCYs and interactions in the active site with **3** could be obtained. The substrate scope of LCY-B and LCY-E was determined and the evolvability with **3** based on MSA was demonstrated by formation of the products β- and α-Ionone by the LCYs, respectively. This indicates the potential of LCY to complement for the limitations observed by SHCs. Investigation of structural data of the evolved variants revealed analogous positions in both enzymes to be hot spots. Nevertheless, to gain a comprehensive understanding of mechanistic details in LCYs more detailed information on the mechanism and the structure is required.

4.3. Alternative active site confinement for substrates with inherently limited pre-organization

Substrates with inherently limited pre-organization in the active site of SHC showed increased selectivity towards terpinen-4-ol starting from **18** and **22** via alternative confinement. The tested monoterpenes are inherently limited to pre-organize in the active site due to their lacking linearity, their small size, and their ability to adapt several orientations within the large SHC active site. The focus with these substrates was on stabilization of the reactive carbocationic intermediates to direct the product outcome and reduce byproduct formation. Especially reaction trajectories of the carbocation rearrangements with monoterpenes are difficult to control due to multiple possible rearrangement reactions of the highly reactive carbocation intermediates.[113–115] The challenge of selectively converting monoterpenes is immense. In chemical synthesis, usually chiral Brønsted acids in water-free organic solvents at low temperature ($-78\,°C$ to $0\,°C$) are used, because stereoselective Brønsted acid catalysis with monoterpenes in water is extremely challenging owing to protonation by water and subsequent nonselective activation of the substrate.[233] In this work enzymes were applied to enable the selective Brønsted acid catalysis reaction under mild conditions in water.

Physiologically relevant enzymes for conversion of monoterpenes in terpene biosynthesis are terpene synthases and the stabilization of the carbocation intermediates in confined biocatalysts are fundamental for the diversity of terpenes.[234,235] In nature, monoterpenes are mostly generated from linear substrates via diversification through cyclization,

rearrangements, isomerizations and modification by water or pyrophosphate attack or downstream enzymes.[127,236–238] It was suggested that most terpene cyclases are generally promiscuous but have evolved to direct the reaction to specific products via dynamic effects.[239] The hydrophobic active site serves as a template to chaperone the conformations of the substrates and can coordinate water molecules as part of the template. In a bornyl pyrophosphate synthase electrostatic interaction with the pyrophosphate moiety direct the reaction trajectories towards selective product formation.[239] In fact, Weitman *et al.* determined that the lifetime of the bornyl cation could be increased by approximately a factor of 4 in complex with the enzyme in a simulation comparing the enzymatic environment with the reaction in the gas phase.[239] Reactions with monoterpenes catalyzed by SHC can be carried out without pyrophosphate substrates making electrostatic interactions negligible and control of selectivity relies mainly in the stabilizing effects of cation-π interactions with the cationic reaction intermediates. The stabilization of carbocation intermediates via cation-π interactions was applied to design confined heterogenous catalysts by Zhang *et al.* to produce a self-assembled resorcinarene capsule with Brønsted acid functionalities.[19] They demonstrated acid-based cyclization of linear monoterpenes within the confined catalyst and the control of the products by preventing side reactions with undesired nucleophiles.[19,240,241] Confinement allowed steering of the intermediate carbocations and direction of the reaction for selective product formation with up to 39% selectivity depending on the double-bond geometry and/or leaving group.[241] In the present work increased selectivity by SHC variants was observed without the need of different leaving groups. Selectivity in monoterpene isomerization reactions is challenging as shown for pinene isomerization reactions in bulk. Non-aqueous conditions favor ring expansion to bornyl and fenchyl cations, whereas aqueous solvents favor ring opening to the *p*-menthenyl cation.[242,243] With SHC reactions towards different products could be performed under aqueous conditions and inherent promiscuous activity. Conversion of pinenes with *Aac*SHC to different products in previous studies represents a promising starting point to investigate the conversion of other monoterpenes to different products.[155,156]

The initial screening of monoterpenes to identify relevant products resulted in the selection of four potential monoterpene substrates out of ten tested. For the monocyclic monoterpenes, no or very little conversion was observed, probably due to the lack of

bicyclic ring stain, which increases the reactivity of the bicyclic monoterpenes. An exocyclic double bond might be more promising in bicyclic monoterpenes for good initial conversion as observed for **22** and **24**, but not for **25**. For the latter the two adjacent methyl groups next to the initiation motif might block the accessibility to the catalytic center in *Aac*SHC. The initially formed products were mainly deprotonated products and only a few industrially relevant, value-added hydrated compounds could be detected. This indicates that *Aac*SHC excludes water from the active site before reaction initiation to shield the carbocations from nucleophilic attack by water, even if the pre-folding of the substrate cannot be achieved, as already observed for squalene.[146]

Substrates **18** and **19** showed very low to no conversion and were tested under different reaction conditions. High volatility poses a challenge with monoterpenes as substrates. The reaction setup itself already minimized the evaporation of the compounds by performing the reaction and the analysis of the reactions in screw capped vials. Moreover, optimization of the reaction conditions revealed a similar tendency as previous experiments with substrates exhibiting modified isoprene patterns. Deviations from the results with the modified isoprene pattern were observed for the tested monoterpenes in the form of lower product formation at higher substrate concentrations and increased product formation at higher pH. Lower product formation at higher substrate concentrations is expected due to the described cytotoxicity of terpenes with terpinolene rated among the most toxic terpenes and limonene toxicity in *E. coli* BL21 (DE3) observed as consequence of the spontaneously formed common oxidation product limonene hydroperoxide.[244,245] At higher pH the protonation state of the residues differs, thereby modifying the interactions in the active site. Moreover, at higher pH, water activation was demonstrated in the terpene cyclase mechanism of a hedycaryol synthase and was shown to be fundamental for the reaction with nerolidol.[246] Thus, the change in pH and the described optimized conditions could contribute to higher concentration of the hydrated compound terpinen-4-ol for the conversion of **18** and **22** with *Aac*SHC.

4.3.1. Assessment of *Aac*SHC enzyme engineering for selective terpinen-4-ol formation

Natural and synthetic approaches to selectively generate terpinen-4-ol demonstrate challenges for efficient and selective product formation. To date terpinene-4-ol is generated by isolation from tea tree oil or synthetic routes.[247] For extraction from tea trees, breeding programs aimed at increasing the terpinene-4-ol concentration in tea

leaves for its extraction. In leaves terpinene-4-ol is not generated as immediate product of monoterpene synthases but is derived from non-enzymatic rearrangement of sabinene hydrate.[248] In organic synthesis, several methods were developed and patented for industrial application, but present potential for improvement.[249,250] For example high cost for energy and substrate arise in the E2 elimination of 1,4-cineole for terpinene-4-ol synthesis.[249] Some synthesis approaches focused on the photooxidation of terpinolene giving rise to a mixture of hydroperoxides, which could be reduced.[251] More recent studies investigated the epoxidation of terpinolene with peroxides and subsequent formation of terpinene-4-ol.[250,252–254] However, epoxidation of terpinolene suffers from modest overall yields, byproduct formation and multiple chemical steps. To efficiently produce terpinen-4-ol this work addressed the limitations of the current processes by engineering *Aac*SHC. Aromatic residues were shown to be involved in carbocation stabilization and important to place steric constraints on the carbocation intermediate in a linalool synthase and a cineole synthase.[126] The observed cation-π interactions by aromatic residues guide the cation towards a single route down the cyclization cascade and thereby prevent branching. This shows that this approach has the potential to generate selectivity.[126] Therefore, aromatic residues and the neighboring residues in *Aac*SHCs involved in substrate binding were modified in a semi-rational approach.

Starting from a promiscuous catalyst for acidic isomerization, *Aac*SHC could be engineered with up to four mutations towards more selective formation of terpinen-4-ol from compounds **18** and **22**. Compound **19** is not converted and not a suitable substrate for *Aac*SHC, despite only small structural differences to **18**. Therefore, the different double bond position might be a major bottleneck for conversion with *Aac*SHC. The active site was narrowed by G600L and G600T and resulted in steric constraints. Thereby the observations with monoterpene synthase homologs and an improvement in selectivity based on a narrower and confined active site could be confirmed.[101]

The presence of at least one water molecule in the active site was expected, which is responsible for the nucleophilic attack onto the carbocation to generate terpinen-4-ol. Water molecules in the active site can generate efficient packing densities in combination with the substrate and perform the controlled nucleophilic attack at the terpinen-4-yl carbocation to direct the reaction towards terpinen-4-ol.[94,255,256] However, based on the structural analysis, it was infeasible to identify anchored water molecules that could fulfill this function.[94,256]

The most significant influence on the selectivity and substrate positioning is presumably obtained by the targeted amino acid residues W312, F365 and W489. Rotation of the side chain F365 was shown for both monoterpenes **18** and **22** and might be important for the conversion of small, cyclic substrates. However, other similar structural rearrangements could not be observed for the two monoterpenes. It is likely that a tailored active site is required for each terpene substrate/product combination. Therefore, the generation of alternative confinement and stabilization of intermediates by cation-π interactions is mandatory for the selective conversion of monoterpenes. These interactions in the active site were optimized for terpinen-4-ol formation with a saturation mutagenesis approach focusing on residues around W312, F365 and W489 and resulted in variants with increased TTN and selectivity for the formation of terpinen-4-ol starting from substrates **18** and **19**.

4.4. Enzyme engineering with alternative confinement

Overall, enzyme engineering with alternative confinement demonstrated a promising engineering approach for cyclases and several challenging substrates. With respect to the substrate's properties three different strategies could be developed. For substrates with modified backbone correct pre-organization was difficult to establish. Therefore, an alternative confinement based on proximity via tunnel modification and substrate complementarity could be obtained. For engineering of enzymes with no structural data available and little sequence similarity with known enzymes structure-guided enzyme engineering approaches are not suitable. Therefore, an MSA was used to determine engineering targets and to form α- and β-ionone from **3** with generated LCY variants. For substrates with inherently limited ability to pre-organize an engineering strategy based on the stabilization of the carbocation intermediate through cation-π interactions was selected. It was shown that selectivity and product formation was increased by focusing on residues around aromatic amino acids. Furthermore, variants that formed terpinen-4-ol starting from different substrates were generated.

Engineering of non-native enzyme functions was demonstrated for various enzymes based on promiscuous enzyme activities. Most relevant studies in this context described the lipase-catalyzed kinetic racemate cleavage of amines by selective acetylation, the cytochrome P450-catalyzed carbene and nitrene transfer reactions and the halohydrin dehalogenase-catalyzed epoxide opening reactions with various non-natural

nucleophiles.[257–260] Given a comprehensive understanding of the reaction mechanism and the advances in enzymology, the catalytic functionalities of enzymes can be used to allow various reactions with diverse substrates. This approach is hampered by years of evolution of enzyme activity for specific substrates and reactions. In this work, overcoming this limitation was achieved by creating alternative active site confinement in protonases SHC and LCY. To obtain alternative confinement the following guiding principles were defined: i) Identification of the bottleneck for substrate binding in the active site. In this work these bottlenecks were the limited pre-organization and the limited stabilization of intermediates. ii) Select engineering targets based on structure guided analysis and screen variants for starting activity. iii) Improve the selected traits via saturation mutagenesis with focus on the substrate's modification or the engineering target.

For application of these guiding principles to generate alternative confinement, AacSHC presents a versatile platform, because of the robustness of AacSHC towards active site mutations. Iterative saturation mutagenesis was applied and single-point mutations with focus on residues in the first and second shell of the AacSHC active site were generated.[128,141,147] For all engineering approaches remarkable robustness of the AacSHC activity was observed and deleterious mutations were absent during the iterative saturation mutagenesis. The reason for this is suggested to be the separation between the protonation machinery and the major part of the active site. In that way the amino acids involved in protonation and substrate pre-organization are distinctive structural elements that are independent of each other. The amino acids are mainly part of a rigid, well-ordered α_6-α_6 barrel double fold and the active site is separate and mostly built on loop structures.[85,147] The initial engineering efforts on LCYs demonstrate similarly susceptible enzymes for amino acid substitutions in and around the active site. The plasticity and further investigation could result in a broader substrate scope and versatile applications of this biocatalyst.

The demonstrated approaches to engineer enzymes with focus on alternative confinement were efficiently applied for SHC and LCY. These enzymes exhibit active site plasticity in saturation mutagenesis approaches and tolerate the reshaping of the active site to generate alternative confinement for substrates, which do not pre-organize for efficient catalysis or for stabilization of reaction intermediates. Purposeful enzyme engineering could benefit from the implementation of the reported findings in order to

direct engineering efforts to specific structural elements in the enzyme and efficiently generate variants with the desired traits.

5. Outlook

In this work the ability to generate alternative confinement with SHCs and LCYs was demonstrated. It was shown that SHCs can be engineered to convert linear terpenes with modified isoprene unit and poor pre-organization in the active site and cyclic monoterpenes inherently limited to pre-organize. In addition, the substrates scope of LCYs was investigated and the conversion of **3** to different isomers was obtained by MSA-based engineering.

These results demonstrate the evolvability of SHC for conversion of non-native substrates. The key towards the observed evolvability is the stabilization of substrates in a pre-organized conformation and carbocationic intermediates along the reaction trajectory. Limitations in the conversion of non-native substrates arise through the confined active site, which is evolved and conserved within the SHC superfamily to selectively convert the native substrate by minimizing the byproduct formation. And even though confinement in nature is generally a beneficial trait, it limits the applicability of biocatalysts for new chemistries and substrates. The demonstrated generation of alternative confinement via distinct approaches to improve catalysts for cyclization reaction and acidic isomerization overcomes the limitations of confined active sites and provides comprehensive structural insights to understand the origin of the improvement. This work provides an approach for broader application of these enzymes with non-native substrates and novel reactions. Further approaches to expand the reaction and substrate scope of SHCs and obtain product yield above mg scale could be applied to improve catalysis. Previously, single-point mutations with focus on residues in the first and second shell of the SHC active site were selected as target.[128,141,147] Random or semi-rational engineering approaches could be tested to increase the activity further. In particular, combinatorial approaches could be advantageous for modification of enzyme activities, because active site substitutions have synergistic effects.[261,262] Thereby, residues for tunnel modification and generation of shape complementarity or residues around aromatic amino acids could be addressed in parallel. Besides the structure of the active center, the acid strength could also be varied. The catalytic Brønsted acid (D376 in *Aac*SHC) is activated by a hydrogen bond to a histidine. The introduction of hydrogen bonds of different strengths should affect the acidity of D376 and lead to faster protonation and conversion. Specifically for the tested modified substrates activity could be increased by using only one isomer or generation a tolerance for the different substrate

isomer. Based on the results described above, VD1-5 are solely converting the *E*-isomer of **2** and preparative approaches with this isomer only could increase the product formation. Potentially, inhibitory effects of the *Z*-isomer on the enzyme could result in a slow biotransformation and lower yield. It was shown that SHCs can convert homofarnesol stereoconvergently.[263] Thereby, *E*- and *Z*-isomers can be converted to the same product.[263] However, this behavior was not observed with the tested compounds.

Moreover the up-scaling of reactions in membrane bound or membrane associated enzymes is impaired, because the outer cell membrane acts as a diffusion barrier for hydrophobic substrates and products.[264] Stable and substrate-accessible enzymes were obtained using spheroplasts and SHC activity was increased up to 100-fold in comparison to whole cell biotransformation with squalene, GA or farnesol.[264] In particular for the tested monoterpenes the application of spheroplasts could be promising, because of the reported toxic effects for *E. coli* cells. Moreover, membrane mimicking or enlarging systems could be tested to increase conversion, as the catalytic activity is assumed to be limited to the integration of overexpressed SHC to the membrane. It was shown that induced membrane invagination to a *E. coli* strain increased the squalene production by more than 3-fold.[265]

To complement the presented methods for improving biocatalysis with SHCs to convert terpenes, the expansion of diversity for protonation enzymes is promising. It was experienced that not all substrates and structural modifications are tolerated by SHCs. Thus, no starting activity with **3** could be obtained by the presented approach to generate alternative confinement. However, LCYs show good starting activities towards **3** and further engineering via MSA-analysis gave access to two ionone isomers. These enzymes are expected to act as protonases, similar to SHCs. Further diversification of the protonase superfamily could result in a diverse reaction panel and substrate scope. In this context the substrate molecular docking in the active site could contribute to identifying new potential catalysts.

Besides the application of the catalysts for generation of interesting building blocks the enzymes could be used in natural terpene extracts to enrich the product of interest. The more selective variants for the formation of terpinen-4-ol could be very promising for this application. The generated variants were engineered to generate terpinen-4-ol as the main product starting from two different monoterpene substrates. The variants'

selectivity could be improved, and they could be further engineered to convert other monoterpene substrates. Addition of these variants to essential oils with mixtures of different monoterpenes would then lead to the enrichment of one value added terpene.

Overall, this work contributes to a potentially greener synthesis of organic molecules and understanding of biosynthetic pathways as well as the extension of the reaction portfolio in organic synthesis with regard to chiral Brønsted acid catalysis in water. With deeper insights on the substrate pre-organization and dynamic properties of enzymes, activity hotspots could be more specifically and predictably identified and targeted in the future. This work demonstrates that structure-based engineering for alternative confinement provides higher activities, broader substrate specificities and expands the substrate scope. Exploiting all the past and the future results in the research of structure-activity relationships of catalysts and combining them with emerging digital technologies, such as artificial intelligence, will enable the rational prediction of catalysts on the basis of pattern analysis.

6. Material and Methods

6.1. Material

6.1.1. Chemicals

Reagents were purchased in analytical grade from Sigma-Aldrich (Steinheim, Germany), Carl Roth (Karlsruhe, Germany), VWR (Darmstadt, Germany) and Alfa Aesar (Karlsruhe, Germany), unless otherwise stated. The substrate **2** was obtained from Givaudan (Vernier, Switzerland) and substrates **4-6** were chemically synthesized (see section 6.7). A list of other chemicals used is given in Table 7.

Table 7. Chemicals used with indication of the manufacturer.

Application	Chemical	Manufacturer
Desoxyribonucleic acid (DNA) staining	Midori Green Advance	NIPPON Genetics Europe, Düren, DE
Loading buffer for agarose gel electrophoresis	6x DNA loading buffer	New England BioLabs® GmbH Ipswich, Massachusetts, US
DNA marker for agarose gel electrophoresis	GeneRuler™ 1 kb Plus DNA Ladder	Fermentas, St. Leon-Rot, DE
Protein marker for SDS-PAGE	PageRuler™ Prestained Protein Ladder	Fermentas, St. Leon-Rot, DE
SDS sample buffer (4x)	RunBlue™ SDS Sample Buffer	Expedeon, Heidelberg, ED

6.1.2. Enzymes and molecular biology kits

Commercially available enzymes used in this work are listed in Table 8. The used molecular biology kits Zyppy™ Plasmid Miniprep Kit, DNA Clean and Concentrator™ Kit and the Zymoclean™ Gel DNA Recovery Kit were obtained from Zymo Research (Irvine, USA).

Table 8. Enzymes and kits used with indication of the manufacturer.

Enzyme	Manufacturer
KOD HotStart DNA-Polymerase from *Thermococcus kodakaraensis*	Sigma-Aldrich, Steinheim, DE
*Dpn*I (20 000 U/ml) from *Diplococcus pneumoniae* G41 T5-Exonuclease from Enterobacteriophage T5 *Taq*-DNA-Ligase from *Thermus thermophilus* HB8 Phusion® HF DNA Polymerase	New England BioLabs® GmbH Ipswich, Massachusetts, USA

6.1.3. *Escherichia coli* strains

E. coli strain XL1-Blue was obtained from Stratagene (Agilent, St. Clara, USA) and BL21(DE3) from Novagene Inc. (Madison, Wisconsin, USA) (Table 9). The strains were used for molecular biology methods and biochemical experiments, respectively.

Table 9. *E. coli* strains used with genotype and origin.

Strain	Relevant genotype	Manufacturer
E. coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F *proAB lacIq*Z*ΔM15* Tn*10* (Tetr)]	Stratagene; Agilent, St. Clara, USA
E. coli BL21 (DE3)	*fhu*A2 *[lon] ompT gal (λ* DE3*) [dcm] ΔhsdS* *λ* DE3 = *λ sBamHIo ΔEcoRI-B* *int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Novagen®, Darmstadt, DE

6.1.4. Oligonucleotides and plasmids

For generation of enzyme variants by site-directed mutagenesis and saturation mutagenesis, sequencing and cloning of genes in this work oligonucleotides were designed *in silico* and obtained from Metabion (Planegg, Germany) or Microsynth AG (Balgach, Switzerland) (Table S 6-9). Oligonucleotides were diluted with ddH$_2$O to a final concentration of 0.1 µmol µl^{-1} and stored at -20 °C.

The construction of the pET 22b (+) plasmid encoding *Aac*SHC WT (UniProt: P33247) and variant G600M and G600T was described previously.[192,227] Plasmids with loop insert and deletion SHC variants were generated by Peter Heinemann as described previously.[188] Transcription occurs by T7-RNA polymerase and is induced by 32.4 mM lactose. LCY WT variants (UniProt: Q38933, Q43415, Q38932) were encoded by a pDHE1650 plasmid construct created by Philip Horz (University Stuttgart). Vector constructs used in this work and plasmids, created within this work designed *in silico* with the Snapgene 3.1.4 software are listed in the supplementary information (Table S 10-11).

6.1.5. Media and buffer

Media and buffer for cultivation were adjusted to the required volume with ddH$_2$O and transferred to culture tubes (5 ml) or Erlenmeyer flasks (50 ml, 500 ml) at 20% of the total volume (Table 10). Temperature stable solutions were autoclaved at 120 °C and 100 kPa for 20 min. Temperature sensitive solutions were sterile filtered (0.2 µm diameter). Vector specific antibiotics were added for selection at temperatures below 55 °C. TB media and salts were autoclaved separately and mixed in a 10:1 ratio.

Table 10. Culture media used in this work.

Culture media	Ingredients
Lysogeny broth (LB) medium	10 g/l tryptone, 10 g/l NaCl, 5 g/l yeast extract
Terrific broth defined autoinduction medium (T-DAB) medium[266] T-DAB salts (10 x)	12 g/l tryptone, 24 g/l yeast extract, 2.9 g/l glucose, 11.1 g/l glycerol, 7.6 g/l lactose 23.1 g/l KH_2PO_4, 125.4 g/l K_2HPO_4
Terrific broth (TB) medium TB salts (10 x)	12 g/l tryptone, 24 g/l yeast extract, 4 ml/l glycerol 23.1 g/l KH_2PO_4, 125.4 g/l K_2HPO_4

Buffers and solutions used in this work are listed in Table 11. Stock solutions were aliquoted in 1 ml and stored at -20 °C.

Table 11. Used buffers and stock solutions.

Solution/buffer	Ingredients
Ampicillin (Amp) stock solution	100 mg/ml Ampicillin
Rhamnose stock solution	250 g/l Rhamnose
IPTG stock solution	1 M Isopropyl-β-D-thiogalactopyranosid (IPTG)
TRIS-acetate-EDTA (TAE) buffer (10x)	400 mM TRIS-Acetat, 10 mM EDTA (pH 8)
Electrophorese buffer (10x)	0.25 M TRIS, 1.92 M glycerol, 10% (v/v) Natriumdodecylsulfat (SDS), ad 1000 ml ddH$_2$O
Coomassie staining solution	30% (v/v) Ethanol, 10% (v/v) acetic acid, 0.1% (v/v) Coomassie Brilliant blue, 60% (v/v) ddH$_2$O
Coomassie destaining solution	30% (v/v) Ethanol, 10% (v/v) acetic acid, 60% (v/v) ddH$_2$O
TfBI buffer	30 mM potassium acetate, 100 mM RbCl, 100 mM $CaCl_2 \cdot$ H_2O, 50 mM $MnCl_2 \cdot 4\ H_2O$, 15% (v/v) glycerol (pH 5.8)
TfBII buffer	10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2 \cdot 2\ H_2O$, 15% (v/v) glycerol (pH 6.5)

The buffers used for protein purification and biotransformations are listed in Table 12 and Table 13, respectively.

Table 12. Buffers for protein purification

Buffer	Ingredients
Lysis buffer I	200 mM citrate, 0.1 mM EDTA, pH 6
Solubilization buffer I	100 mM citrate, 1% CHAPS
Ion exchange chromatography (IEX) wash buffer	25 mM citrate, 1% CHAPS, pH 6
IEX elution buffer	25 mM citrate, 200 mM NaCl, 1% CHAPS, pH 6
IEX regeneration buffer	25 mM citrate, 1 M NaCl, 1% CHAPS, pH 6
Lysis buffer II	100 mM TRIS pH 7.4
Binding buffer	50 mM TRIS, 500 mM NaCl, 20 mM Imidazole pH 7.4
Elution buffer	50 mM TRIS, 500 mM NaCl, 500 mM Imidazole pH 7.4
Desalting buffer	50 mM TRIS, 100 mM NaCl pH 7.4

Table 13. Buffers for biotransformation

Buffer	Ingredients
TRIS-Maleate buffer	50 mM Maleic acid, 50 mM TRIS, 1 mM DTT, 1% CHAPS (v/v)
PIPES buffer	50 mM PIPES, 1 mM DTT, 1% CHAPS (v/v)
Cyclodextrin (CD) buffer	0.4% SDS, eqimolar CD (to the substrate), pH 6

6.2. Molecular biological methods

6.2.1. Mutagenesis and plasmid construction

Site-Directed Mutagenesis was performed with the QuikChange® protocol and the employed primers are listed in Table S 6-7. KOD Hot Start DNA polymerase obtained from Merck (Darmstadt, Germany) was used in the reaction and the conditions using the Novagen protocol were adjusted to the size of the target DNA fragment and the annealing temperature of the oligonucleotides shown in Table 14.

For saturation of positions with degenerate codons, the „22c-trick" with overlap extension polymer chain reaction (PCR) was used. Respective oligonucleotides contain 3 forward (fwd) primers including the position of interest and one reverse (rev) primer with complementary overhangs at the 5'-ends (approx. 20 base pairs (bp) long) (Table S 6-7). The non-complementary overhangs at the 3'-end are selected to have a melting temperature (T_m) of approx. 55 °C. The used codons for the forward primer are NDT (N = A/T/C/G, D = no C), VHG (V = no T, H = no G) and TGG. These were mixed in a 12:9:1 ratio to final concentration of 10 µM each. By using this method, the screening effort can be reduced from 96 to 66 variants per position.[267] The used PCR conditions are based on the Novagen protocol with KOD Hot Start DNA polymerase.

During site-directed mutagenesis and saturation mutagenesis template DNA is methylated, in contrast to the amplified DNA, and is digested by treating the PCR products with endonuclease *Dpn*I from NEB GmbH (Frankfurt a. M., Germany) for at least 3 h at 37 °C after PCR.

Table 14. Components and temperature profile for QuikChange® site-directed mutagenesis using the KOD Hot Start DNA polymerase.

PCR assay		PCR program			
	V (µl)		Temperature (°C)	Time (s)	
*dd*H$_2$O	29	Activation	95	120	
Buffer	5	Denaturation	95	30	
Primer fwd. (10 µM)	1	Primer annealing	T_m(Primer) – 10 °C	30	30 cycles
Primer rev. (10 µM)	1	Elongation	70	25 s/kb	
dNTP's (8 mM)	5	Final elongation	70	2x elongation time	
DMSO	2.5	Storage	8		
MgSO$_4$	4.5				
Template DNA (50 ng)	1				
KOD polymerase	1				

For plasmid construction by Gibson assembly, vector and insert DNA were amplified separately with oligonucleotides that generate overhangs and adjusted elongation time (Table S 9). Gibson Assembly Master Mix was prepared according to Table 15 in 15 µl aliquots and stored at -20°C until further use. The purified and concentrated DNA fragments of vector and insert (see 6.1.2.) were added in a 4:1 ratio within 5 µl volume to a Gibson Assembly Master Mix and were

incubated at 50 °C for 60 min. The product obtained was transformed into *E. coli* XL1-Blue for amplification and methylation of the plasmid construct. Colonies grown overnight were used for colony PCR and were processed for sequencing to verify the ligation of the DNA fragments.

Table 15. Composition of Gibson Assembly Master Mix and 5 x ISO reaction buffer.

Gibson Assembly Master Mix		5 x ISO reaction buffer	
5x Iso reaction buffer	100.0 µl	25% (w/v) PEG-8000	1,5 g
T5 Exonuclease (10 U/mL)	0.2 µl	500 mM TRIS-HCl pH 7,5	3000 µl
Phusion polymerase (2U/ml)	6.25 µl	50 mM MgCl$_2$	150 µl
Taq DNA Ligase (40 U/ml)	50.0 µl	50 mM DTT	300 µl
*dd*H2O	218.55 µl	dNTPs (1 mM) each	240 µl
		5mM NAD	300 µl
		*dd*H2O	to 6 ml

6.2.2. Colony PCR

Colony PCR was performed to verify the ligation of DNA products. The Novagen protocol with KOD Hot Start DNA polymerase was prepared as a Master Mix with the oligonucleotides for the amplification of the insert and 1 µl *dd*H$_2$O instead of the template DNA (Table 14). Colonies grown overnight after transformation (see 6.2.6.) of the Gibson assembly product into *E. coli* XL1-Blue were picked from the plate and added separately to the PCR reactions to serve as template DNA. The reaction conditions of the PCR program were adjusted by increasing the activation time to 15 min to lyse the cells and make the template DNA accessible. Analysis of the colony PCR was done by agarose gel electrophoresis.

6.2.3. Agarose gel electrophoresis

Agarose gel electrophoresis confirmed the size of the amplified DNA fragments and was used for analytical separation or preparative separation to purify PCR products. For the preparation, 0.8% agarose was dissolved in TAE buffer by heating. For DNA staining 0.01% Midori Green Advance was added at temperatures below 60 °C. PCR products were mixed with 6 × DNA loading dye (Thermo Scientific, USA) and were loaded onto the gels next to the GeneRuler 1kb Plus DNA ladder to analyze the size of DNA fragments. Samples and ladder were separated electrophoretically at 100 V for 30 min. Laddered fragments were visualized with Blue/Green LED (470 nm) or UV light (366 nm).

6.2.4. Isolation and purification of plasmid DNA and DNA from agarose gels

Plasmid DNA was isolated and purified from 5 ml of overnight grown *E. coli* XL1-Blue in LB medium and 100 µg/ml Amp with the Zyppy™ Plasmid Miniprep Kit following manufacturer's instructions. Plasmids were eluted with 30 µl sterile *dd*H$_2$O and stored until further usage at -20 °C.

DNA from PCR samples was extracted from agarose gels by cutting out the bands with the expected size of the DNA fragment and processed with the Zymoclean Gel DNA Recovery Kit following the manufacturers protocol. For digested PCR products the Zymoclean DNA Clean & Concentrator Kit was used following the manufacturer's instructions. The purified DNA was eluted with 10 µl sterile *dd*H$_2$O.

DNA concentration and purity was determined by a NanoDrop spectrometer (NanoDrop® 2000, Peqlab Biotechnologie GmbH, Erlangen, DE) at a wavelength of 260 nm. *dd*H$_2$O was used as a reference.

6.2.5. Sequencing

DNA sequencing was performed by Eurofins Genomics (Ebersberg, Germany) or Microsynth AG (Balgach, Switzerland). According to the service requirements 15 µl of the purified DNA samples with concentrations of 40-100 ng/µl were send for sequencing with the appropriate oligonucleotides (Table S 9). The obtained sequencing results were analyzed to confirm the mutations before they were transformed to competent *E. coli* BL21(DE3).

6.2.6. Transformation of chemically competent *E. coli*

For the reproduction of the plasmid DNA in chemically competent cells the heat shock method was used. Therefore, 3 µl of purified PCR product was added to 25 µl competent XL1-Blue cells and incubated for 20 min on ice. The heat shock was performed for 60 s at 42 °C. The cells were cooled for approximately 1 min on ice afterwards and 500 µl sterile LB medium was added for recovery at 37°C and 180 rpm for 60 min. For site-directed mutagenesis, the obtained cell suspension was centrifuged at 14000 rpm, dissolved in 150 µl of supernatant, plated and incubated 24 h at 37 °C. Single colonies were used for inoculation of 5 ml LB medium. For saturation mutagenesis cells were directly used to inoculate 5 ml LB medium with 100 µg/ml Amp.

6.3. Microbiological methods

6.3.1. Strain maintenance and cultivation

For strain maintenance glycerol stocks were prepared with 500 µl of 50% glycerol and 500 µl of overnight grown cell culture and stored at -80 °C until further use.

For cultivation of *E. coli* XL1-Blue for molecular biology methods either single colonies grown overnight at 37 °C or glycerol stocks prepared from single colonies and stored at -80 °C were used. The cells were used for inoculation of 5 ml LB-media with the required antibiotic and incubated at 180 rpm and 37 °C overnight. Cultivation of *E. coli* BL21 (DE3) for heterologous expression was conducted on LB agar plates and at least 5 colonies were used for inoculation of 5 ml LB medium with appropriate antibiotic.

6.3.2. Preparation of chemically competent cells

For preparation of chemical competent cells, the Rubidium chloride method was used. Rubidium chloride contributes to the enhanced association of foreign DNA to the cell membrane in the surrounding medium and results in enhanced uptake into the cells by the subsequent heat shock (see 6.2.6.). Therefore, 1 ml *E. coli* culture grown overnight in 5 ml LB medium at 37 °C was used to inoculate 100 ml LB medium. Cells were grown until OD_{600} of 0.5-0.8, cooled on ice for 15 min and kept at 4 °C in the following steps. Centrifugation at 4000 rpm and 4 °C for 15 min, resuspension of the cell pellet with 30 ml volume cold TfBI buffer and incubation for 30-90 min at 4°C followed. Harvesting cells by centrifugation at 4000 rpm and 4 °C for 15 min provided a cell pellet, which was resuspended in 4 ml cold TfBII and incubated on ice for 15 min. Afterwards 50 µl cells were aliquoted in precooled reaction tubes (Eppendorf AG, Hamburg, DE) and snap-frozen in liquid nitrogen. Until further use cells were stored at -80 °C.

6.3.3. Heterologous protein expression

Heterologous expression of SHC variants was conducted with *E. coli* BL21 (DE3) cells in either 50 ml or 500 ml autoinduction medium with the corresponding antibiotic. The medium was inoculated with 1% of freshly grown cells and grown at 37 °C at 180 rpm for 50 ml and 120 rpm for 500 ml cultures. Enzyme expression was induced by 32.4 mM lactose after consumption of the preferred carbon source glucose in the medium. For the expression of LCY variants expression was induced by rhamnose at a final concentration of 250 mg/l. Expression occurred overnight for 18 h. Cells with expressed SHC or LCY

variants were harvested for 20 min at 4000 rpm and used directly for cell lysis and protein purification or biotransformation.

For expression of protein variants in 96 DWPs, 2 mL Riplate®, Carl Roth, Karlsruhe, Germany) the expression protocol was modified as follows. The preculture of 500 µl LB medium with antibiotic was added to each well. 91 freshly transformed colonies were picked and used for inoculation of the medium. The remaining 5 wells were filled with two positive controls and two negative controls, e.g., WT variant and empty vector. One well was not inoculated. The DWPs were sealed with a Breath Easier Sealing Membrane (Sigma-Aldrich, Steinheim, DE) and incubated for 18 h at 37 °C and 300 rpm. For the main culture 990 µl of either T-DAB or TB medium supplemented with antibiotic and rhamnose were added to the wells for SHC or LCY, respectively. To inoculate the media, 10 µl of the preculture were used. The preculture was harvested, and the cell pellets were stored at -80 °C. The main culture was sealed with the Breath Easier Sealing Membrane and incubated for 18 h at 300 rpm and 37 °C or 30 °C for SHC or LCY, respectively. Cells were harvested at 4 °C for 15 min and 4000 rpm. The supernatant was discarded, and the cells used directly for biotransformation.

6.3.4. Cell lysis

Depending on the enzyme and the experiment the cell lysis protocol varied. With SHC variants for IEX chromatography, freshly harvested cells were suspended in lysis buffer I (5 ml per gram of cell pellet) and homogenized using a tissue homogenizer (Radnor, PA, USA). After addition of 10 mg/mL DNAse I (Roche Diagnostics, Switzerland) cells were lysed twice with a high-pressure homogenizer (EmulsiFlex-C5, Avestin, Ottowa, CAN) or with ultrasound (Branson Sonifier 250 with microtip 1/8" diameter, Danbury, Connecticut, USA) under constant cooling in an ice bath for 5 min (duty cycle 30%) with an output of 3-5. The resulting suspension was centrifuged at 40000 x g and 4 °C for 45 min. A sample of the supernatant was taken before discarding the rest. The obtained pellet was washed with wash buffer (3 ml per gram of cell pellet) and centrifuged again. To the obtained pellet containing the membrane bound monotopic enzyme, solubilization buffer I was added and the solution incubated overnight under agitation. The CHAPS (3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate) is a zwitterionic detergent in the solubilization buffer and solubilizes the membrane associated SHC. With the cell solution thermo-lysis at 50 °C for 2 h was performed and precipitated protein was

separated from the supernatant using centrifugation (40000 x g, 1 h, 4 °C). Enzyme preparations of SHC were used for further purification.

For the LCY purification by affinity chromatography, 3 ml per gram of cell pellet lysis puffer II was added after cell harvest. Cells were homogenized using a tissue homogenizer, supplemented with 10 mg/ml DNAse I and lysed twice with a high-pressure homogenizer (EmulsiFlex-C5, Avestin, Ottowa, CAN). The resulting suspension was centrifuged at 50000 x g and 4 °C for 1 h and the obtained supernatant was filtered (0.25 µm) in preparation for purification by affinity chromatography.

6.3.5. Fed-Batch cultivation and lyophilization of cells

For fed-batch cultivation of *E. coli* BL21 (DE3) expressing VD1 to be used in preparative approaches the Labfors 3 bioreactor (Infors AG, Bottmingen, Schweiz) was used. VD1 was produced in a 3 l batch of start salt medium inoculated from 500 ml preculture grown in salt medium pH 6.7 at 30 °C under exponential feeding with 2 growth rates (μ=0.24/0.11) before and after induction, respectively (Table 16). The culture was constantly oxygenated at 6 l/min with NH_4OH as base and expression was induced with 1 mM IPTG at OD=94. After 3 h incubation, the cells were harvested at OD=112 for 1 h at 7500 x g and 4 °C. Cells were frozen and stored at -80 °C overnight. Lyophilization of the cells was

performed in a Christ alpha 2-4 LD plus lyophilizer (Christ, Osterode) for 24 h. Afterwards cells were mortared and stored at -80 °C upon further use.

Table 16. Solutions used for the fed batch process to express VD1.

Solutions	Ingredients
Preculture salt medium*	13.3 g/l KH$_2$PO$_4$, 4.0 g/l (NH$_4$)$_2$HPO$_4$, 1.2 g/l MgSO$_4$·7 H$_2$O, 1.7 g/l citric acid monohydrate, 8.4 mg/l Titriplex III, 2.5 mg/l CoCl$_2$·6 H$_2$O, 1.5 mg/l CuCl$_2$·2 H$_2$O, 3.0 mg/l H$_3$BO$_3$, 15.0 mg/l MnCl$_2$·4 H$_2$O, 2.5 mg/l Na$_2$MoO$_4$·2 H$_2$O, 13.0 mg/l Zn(CH$_3$-COO)$_2$·2 H$_2$O, 100.0 mg/l Fe(III) citrate hydrate, 4.5 mg/l thiamine HCl, 8.0 g/l D-glucose, 0.1 g/l Amp
Start salt medium*	13.3 g/l KH$_2$PO$_4$, 4.0 g/l (NH$_4$)$_2$HPO$_4$, 1.2 g/l MgSO$_4$·7 H$_2$O, 1.7 g/l citric acid monohydrate, 8.4 mg/l Titriplex III, 2.5 mg/l CoCl$_2$·6 H$_2$O, 1.5 mg/l CuCl$_2$·2 H$_2$O, 3.0 mg/l H$_3$BO$_3$, 15.0 mg/l MnCl$_2$·4 H$_2$O, 2.5 mg/l Na$_2$MoO$_4$·2 H$_2$O, 13.0 mg/l Zn(CH$_3$-COO)$_2$·2 H$_2$O, 100.0 mg/l Fe(III) citrate hydrate, 4.5 mg/l thiamine HCl, 0.1 g/l Amp
Feeding solution*	750.0 g/l D-glucose, 20.0 g/l MgSO$_4$·7 H$_2$O, 13.0 mg/l Titriplex III, 4.0 mg/l CoCl$_2$·6 H$_2$O, 2.5 mg/l CuCl$_2$·2 H$_2$O, 5.0 mg/l H$_3$BO$_3$, 23.5 mg/l MnCl$_2$·4 H$_2$O, 4.0 mg/l Na$_2$MoO$_4$·2 H$_2$O, 16.0 mg/l Zn(CH$_3$-COO)2·2 H$_2$O and 40.0 mg/l Fe(III) citrate hydrate, 0.1 g/l Amp

*trace elements autoclaved separately

6.4. Protein biochemical methods

6.4.1. Protein purification

Further impurities could be separated from the SHC variants by ion exchange chromatography with an anion exchange matrix (diethylaminoethyl-Sephacel, Sigma Aldrich, St. Louis, US) of 2 ml column volume (CV). Therefore, the supernatant was diluted 1:5 with *dd*H2O, filtered with 0.45 µm filter and loaded onto the anion exchange matrix, which was equilibrated with IEX wash buffer. The matrix was then washed with 5 CVs IEX wash buffer and the SHC variants were eluted with 7.5 CVs IEX elution buffer. The collected samples during purification and the eluted fractions were analyzed by SDS-PAGE. The column was regenerated with the IEX regeneration buffer. Purified enzymes were used directly in biotransformations.

Affinity chromatography with 2x 2ml HiTrap Chelating columns loaded with Nickel on an ÄKTA pure 25 System (Cytiva, Marlborough, US) was used to purify LCY. The columns were equilibrated with 5 CVs binding buffer at 4 °C before 50 ml of the filtered lysate was loaded onto the column at 5 ml/min. After loading the lysate unspecifically bound proteins were washed out by applying 5 CVs of binding buffer containing 20 mM of

imidazole. The specifically bound His-tagged enzyme was eluted with 200 mM imidazole in the elution buffer. Proteins were desalted using a HiPrep 26/10 with Sephadex G-25 Fine resin column (Cytiva, Marlborough, US) in desalting buffer. Fractions containing the protein were used directly for biotransformation or were concentrated by Vivaspin 20 centrifugal concentrators (Satorius, Göttingen, Germany) with a molecular weight cut-off of 50.000 PES. Buffer was exchanged in the centrifugal concentration to the TRIS maleate buffer, and then the biotransformations were performed.

6.4.2. Determination of protein concentration

Protein concentration was determined in triplicates with a NanoDrop® 2000 spectrometer (Thermo Scientific, USA) with molecular weight (MW) of 71.44 kDa and molar extinction coefficient ε = 185180 for SHC and with MW of 57.43 kDa and ε = 56880 for LCY.

6.4.3. Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein purity was assessed by SDS-PAGE. Samples were diluted 1:10 and 10 μl SDS-sample buffer (RunBlue™ LDS Sample Buffer) was added to 30 μl protein sample and heated at 95 °C for 5 min. 10 μl of the samples were loaded onto a 12% precasted SDS-PAGE SurePAGE gel from Expedeon (Abcam, Berlin, DE) together with 3 μl of the PageRuler™ Prestained Protein Ladder. Separation of the proteins occurred electrophoretically at constant voltage of 80 mV for 10 min following 110 mV for 1.5 h. The gels were stained for 1 h at room temperature with the Coomassie staining solution and destained with the destaining solution afterwards for at least 1 h to visualize the protein bands.

6.5. Biotransformations

6.5.1. Biotransformations in DWPs

After heterologous protein expression and cell harvest, whole cells in DWPs were resuspended with 395 μl ddH$_2$O or TRIS-maleate buffer adjusted to the determined optimal pH for *in vivo* biotransformations with SHCs and LCY, respectively, unless otherwise stated. Cells were transferred to a fresh 96 DWP with 1.2 ml glass inlets and supplemented with 5 μl of a 80 mM or 160 mM substrate solution dissolved in DMSO (1 mM or 2mM final concentration), unless otherwise stated. In reactions with LCY, NAD(P)H and FAD were added in concentrations of 0.5 mM and 0.01 mM (100 mM NAD(P)H and 2.5 mM FAD stock concentration), respectively. Reactions were incubated

in sealed plates for 20 h at 30 °C and 300 rpm with the substrates except for substrate **3** and **9** and SHC. The latter biotransformations were performed at 50 °C. With monoterpenes as substrates biotransformation time was 40 h. Reactions were terminated and reactants extracted by addition of 500 µl cyclohexane/ethylacetate (1:1) and shaking for 40 s. Samples were centrifuged (4000 rpm, 5 min) and analyzed by GC/MS.

6.5.2. Biotransformations in vials

For biotransformations in 9 mm short thread glass vials (Fisher Scientific, Schwerte, DE), cell wet weight (cell$_{WW}$) concentration was adjusted to 0.2 mg/ml with ddH$_2$O or TRIS-maleate buffer adjusted to the optimal pH, for SHCs and LCY, respectively. In the vials 945 µl cell solution of purified enzyme was mixed with 5 µl substrate to a final concentration of 1 mM or 2 mM (190 mM or 380 mM stock solution dissolved in DMSO). In reactions with LCY NAD(P)H and FAD were added to a final concentration of 0.5 mM and 0.01mM (100 mM NAD(P)H and 2.5 mM FAD stock concentration), respectively. Biotransformations with whole cells were performed in duplicates and reactions with purified enzyme were performed in triplicates under conditions described (see 6.5.1). Extraction with 950 µl cyclohexane:ethylacetate (1:1) by 40 s mixing terminated the reaction. The obtained organic phase was analyzed with GC/MS.

6.5.3. Preparative biotransformations

Preparative biotransformations were performed in closed flasks with 2 g freeze dried cells or 10 g fresh cells dissolved in 100 ml ddH$_2$O, 100 ml CD buffer or in 1 l CD buffer with 100 mM Mg(NO$_3$)$_2$ added. Reaction progress was monitored by GC/MS to determine the end of the reaction in order to initiate extraction and purification.

6.5.4. Determination of kinetic parameters

The determination of K$_m$ and k$_{cat}$ was used to characterize the catalytic performance of the tested proteins and was based on the Michaelis-Menten equation. In initial experiments different substrate concentrations and time points were selected to determine the reaction conditions with the critical conversion below 10% and a linear increase in product formation in the steady state. The reactions were performed as described in section 6.5.2. at selected time points of 1 h and 2 h reaction time. Substrate concentrations were selected between 0.1 and 8 mM for **2.** The kinetic parameters were fitted with the GraphPad Prism software to the Michaelis-Menten equation.

6.6. Analytical methods

6.6.1. Gas chromatography

GC/MS measurements were performed on an Agilent GC 7820A coupled to a mass spectrometer MSD 5977B or an Agilent GC 7890A equipped with a mass spectrometer MSD 5975C. The MS detector was operating with 70 eV (EI ionization). The GC 7820A equipped with a PAL autosampler (RSI 120, Agilent, Santa Clara, USA) and a HP-5MS (30 m x 250 µm x 0.25 µm, Agilent, Santa Clara, USA) or a ZB-1HT Inferno™ capillary column (30 m x 250 µm x 0.25 µm, Phenomenex, Aschaffenburg, DE) was used with constant helium pressure of 14.168 ψ or 14.476 ψ and sample injection temperature of 250 °C or 300 °C, respectively. With the Agilent GC 7890A and the autosampler 7693A ALS (Agilent, Santa Clara, USA) a DB-WAX column (30 m x 320 µm x 0.25 µm, Agilent, Santa Clara, USA) was used with sample injection at 280 °C, a split of 5:1 and chromatographic separation of the sample with constant helium gas pressure of 15.754 ψ. The methods developed for the analysis and quantification of the biotransformations with different substrates are listed in Table 17.

For measurements from DWPs, a GC-MS equipped with a PAL-Sampler allowed the sample injection from the organic phase of the sealed plate with 91 variants screened per plate, directly. Only for biotransformations with volatile monoterpenes in DWPs the organic phase was transferred into GC vials with inlets, which were capped to reduce evaporation.

Chiral GC analysis was performed on a Shimadzu GC-2010 system with FID detector equipped with an autosampler (AOC-20i Auto-Injector, Shimadzu, Kyõto, JPN) and CP ChiraSil-Dex CB capillary column (25 m x 250 µm x 0.25 µm, Agilent, Santa Clara, USA). Samples were injected at 250 °C in split mode (15:1) and hydrogen was used as carrier gas with constant linear velocity (33.1 cm/s) (Table 17).

The GC/MS results were analyzed and integrated with the MassHunter Qualitative Analysis Software (Agilent, Santa Clara, USA) or the GC solution software (Shimadzu, Kyõto, JPN). Samples were quantified based on the conversion $(AREA_{product}/(AREA_{substrate}+AREA_{product})*100)$ calculated from the total ion count chromatogram. Products were quantified with a calibration curve from 0.015-1 mM with commercially purchased or synthesized product standards, unless otherwise stated. The total turnover number (TTN) was based on product concentration in mol and enzyme

concentration related to the molecular weight and was calculated from TTN=$(6,02*10^{23}*(1/1000)*(1/1000)*(Conc_{product}))/(Conc_{enzyme}/(74,6*1000*1,66*10^{-27}*1000))$. For screening in DWPs an increase of more than 50% in product formation was the threshold for the selection of hit variants and was based on the determination of the variation coefficient (C_V) determined for VD1 with **2** (see 3.1.4.).

Table 17. GC methods for the substrates and the corresponding products.

| Substrate	Column	Rate (°C/min)	Temperature (°C)	Hold (min)
1	ZB-1HT Inferno™	15	118 128	0.5 5.5
2	HP-5	15	165 185	1 3
2 (chiral)	CP ChiraSil-Dex CB	10	100 190	1 10
3	ZB-1HT Inferno™	30	125 140	0.5 5.6
3	DB-Wax	20	160 175	9
4	ZB-1HT Inferno™	10	40 220	1 1
5	ZB-1HT Inferno™	15	40 200	1 1
6	ZB-1HT Inferno™	15	40 300	1 2
7-10	DB-Wax	70 100 20	75 100 180 210	4.25 0.5 2

6.6.2. Nuclear Magnetic Resonance

The ^{1}H and ^{13}C NMR spectra were measured at 500.15 MHz for ^{1}H- and 125 MHz for ^{13}C using a Bruker Avance 500 spectrometer. The tested substances were dissolved in CDCl$_3$ and sampled at room temperature. Chemical shifts δ refer to tetramethylsilane (TMS) in ppm set to 0. To assign more complex signals 2D NMR methods 1H,1H-COSY (Correlated Spectroscopy), 1H,13C-HSQC (Heteronuclear Single Quantum Coherence), 1H,13C-HMBC (Heteronuclear Multiple Bond Correlation) and 1H,1H-NOESY (Nuclear Overhauser Enhancement Spectroscopy) were used.

6.7. Chemical synthesis and separation of isomers

Compounds **2**-**6** and **9** were synthesized by Heng Li, Lukas Lauterbach, Kizerbo A. Taizoumbe under the supervision of Jeroen S. Dickschat (University Bonn) and were provided for this work.

6.7.1. Synthesis of 7

1   **7$_{I1}$**

7$_{I2}$ **7**
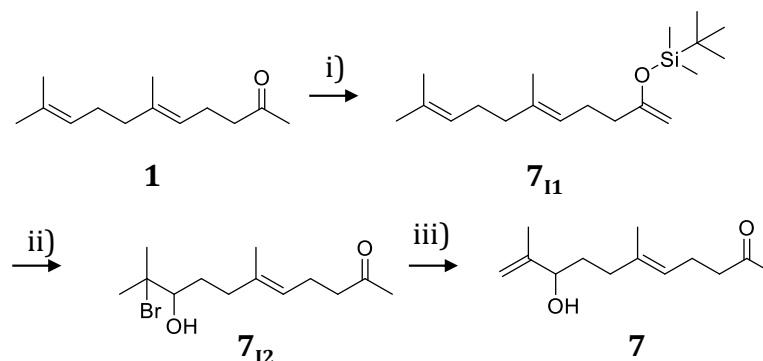
Scheme 2. Preparation of **7** adapted from Hoshino *et al.* (2004).[157] Reagents and conditions: i) TBDMSCl, NEt$_3$, ii) NBS, THF H$_2$O, N$_2$, 0°C and iii) K$_2$CO$_3$, MeOH.

Synthesis of **7** was supervised within the scope of this work and was performed by T. Trinks. Reaction conditions were adapted from Hoshino *et al.* (2004).[157] For the synthesis 4.6 ml (3.4 g, 33 mmol) triethylamine was added oxygen and water free to a 200 ml two-necked flask with 5.6 ml (4.9 g, 25 mmol) **1**. During stirring, 4.9 g (35.50 mmol) *tert*-butyldimethylsilyl chloride (TBDMSCl) was added in the course of 15 min at 30-40 °C. The yellow solution was stirred for an additional 30 min before 35 ml of 0.92 M NaI in dry acetonitrile was added dropwise over the course of 30 min. The solution was stirred for 1 h and warmed to 50 °C. After 4 h incubation under stirring the solution was cooled to room temperature. Next, 25 ml ice-cold hexane and 55 ml ice-cold water were mixed with the solution for extraction with hexane (3x30 ml). The obtained organic phase was dried over sodium sulfate and concentrated *in vacuo*. 5.41 g (17.50 mmol, 70.1% yield) of crude product **7$_{I1}$** was obtained.

In the second step 5.41 g (17.50 mmol) of the crude **7$_{I1}$** was dissolved in 480 ml tetrahydrofuran (THF), stirred under nitrogen atmosphere, and cooled to 0 °C. Water was added until the solution became cloudy. Next, THF was added to the solution until it became clear. During 10 min, 3.38 g (19.00 mmol) N-bromosuccinimide (NBS) was added to the reaction solution and stirred an additional 20 min at 0 °C. Approximately, half of the solvent was evaporated *in vacuo*. The remaining reaction solution was then mixed

with 200 ml water and extracted with hexane (4 x 150 ml). The organic phase was dried over sodium sulfate and the product was concentrated *in vacuo*. 12.12 g of crude **7**I2 was obtained and stored in the fridge for three days under nitrogen atmosphere, until it was stirred with 3.55 g (25.70 mmol) potassium carbonate in 75 ml methanol for 3 h. After extraction with hexane (3 x 100 ml) the solution was filtered and evaporated to remove potassium carbonate and MeOH and obtain **7** (3.55 g, 25.70 mmol).

Yield and spectral data of **7**: Yellow oil, 25.7 mmol, 65.62% yield. **7** ((*E*)-9-hydroxy-6,10-dimethylundeca-5,10-dien-2-one): 1H-NMR (CDCl3, 500 MHz): δ (ppm) 1.15-1.21 (q, 2H), 1.59-1.64 (d, 6H) 2.03-2.06 (t, 2H), 2.13-2.15 (m, 3H), 2.25-2.26 (q, 2H), 2.44-2.51 (m, 2H), 3.30-3.34 (t, 1H), 4.92 (s, 1H), 5.06-5.1 (m, 2H), 5.13-5.16 (t, 1H). Impurities MeOH and THF: δ (ppm): 1.16 (s), 1.6 (m), 3.1 (s). 13C-NMR (CDCl3, 125 MHz): δ (ppm) 18.57 (1C), 20.06 (1C), 22.26 (1C), 23.53 (2C), 30.53 (2C) 33.17 (1C), 44.03 (1C), 123.14 (1C), 124.29 (1C), 127.61 (1C), 136.91 (1C), 208.79 (1C).

6.7.2. Separation of *E*- and *Z*-2

A mixture (1 g) of *E*- and *Z*-isomers of **2** were separated with a silver nitrate column (5.5 g silver nitrate, 50 g mesh silica) protected from direct light with gradually increasing the ratio of polar solvent from 50:1 to 10:1 (petroleum ether/ethyl acetate). Fractions containing the *E*- or the *Z*-isomer were determined by GC/MS measurement and pooled, respectively. After solvent evaporation, the substrates were analyzed by 1H-NMR and tested in a biotransformation.
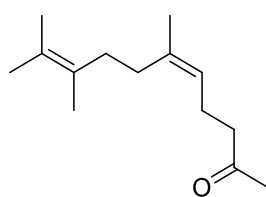

(*Z*)-**2**

Yield and spectral data of (*Z*)-**2**: Colorless oil, 0.911 mmol, 18.99% yield. (*Z*)-**2** ((*Z*)-6,9,10-trimethylundeca-5,9-dien-2-one): 1H-NMR (CDCl3, 500 MHz): δ (ppm) 1.64-1.66 (d, 9H), 1.7 (s, 3H) 2.05-2.07 (s, 4H), 2.13 (s, 3H), 2.24-2.28 (q, 2H), 2.43-2.46 (t, 2H), 5.05-5.07 (t, 1H). 13C-NMR (CDCl3, 125 MHz): δ (ppm) 18.57 (1C), 20.06 (1C), 22.26 (1C), 23.53 (2C), 30.53 (2C) 33.17 (1C), 44.03 (1C), 123.14 (1C), 124.29 (1C), 127.61 (1C), 136.91 (1C), 208.79 (1C).
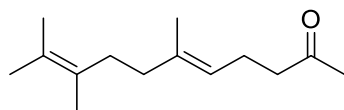
(*E*)-**2**

Yield and spectral data of (*E*)-**2**: Colorless oil, 0.176 mmol, 3.67% yield. (*E*)-**2** ((*E*)-6,9,10-trimethylundeca-5,9-dien-2-one): ^{1}H-NMR (CDCl$_3$, 500 MHz): δ (ppm) 1.63 (s, 12H), 1.96-1.99 (dd, 2H) 2.07-2.10 (dd, 2H), 2.14 (s, 3H), 2.23-2.28 (q, 2H), 2.44-2.47 (t, 2H), 5.06-5.09 (t, 1H). ^{13}C-NMR (CDCl$_3$, 125 MHz): δ (ppm) 16.09 (1C), 20.04 (1C), 20.55 (2C), 29.96 (1C), 33.48 (2C) 38.16 (1C), 43.76 (1C), 122.27 (1C), 124.07 (1C), 127.51 (1C), 136.85 (1C), 208.94 (1C).

6.7.3. Synthesis of the perillyl hexyl ether

The perillyl hexyl ether for substrate engineering of limonene for SHC was synthetized by Williamson ether synthesis with perillyl alcohol and 1-bromohexane adapted from Hammer (2014).[147]
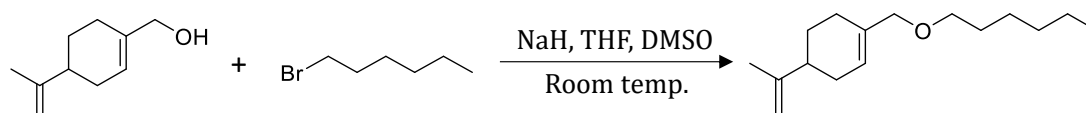
Scheme 3. Synthesis of perillyl hexyl ether adapted from Hammer (2014).[147]

Water and oxygen free to 20 ml anhydrous THF, NaH (1,1 eq., 60% ig, 0,352 g, 8.8 mmol) was added at room temperature (RT). Perillyl alcohol (1 eq., 1.22 g, 8 mmol) was dissolved in 5 ml anhydrous THF and added dropwise in 5 min to the reaction. After 2 h incubation 1-bromohexane (0.8 eq., 1.056 g, 6.8 mM) in 5 ml anhydrous DMSO was added in 5 min to the reaction. The reaction was stirred for 20 h at room temperature until conversion was confirmed by GC/MS. The reaction mixture was then hydrolyzed in ice water and extracted three times with ethyl acetate (50 ml). The combined organic phase was dried with Na$_2$SO$_4$, the solvent was removed on a rotary evaporator and the crude product was purified by SiO$_2$ column chromatography (n-hexane:dichloromethane 2:1). Fractions containing the perillyl butyl ether were determined by GC/MS measurement and pooled. After evaporation of solvents, the substrates were analyzed by ^{1}H-NMR, ^{13}C-NMR and tested in a biotransformation.

Yield and spectral data of perillyl hexyl ether: Colorless oil, 0.47 mmol, 10.23% yield, Rf = 2.37 (n-hexane:dichloromethane 2:1). Perillyl butyl ether (1-((hexyloxy)methyl)-4-(prop-1-en-2-yl)cyclohex-1-ene): ^{1}H-NMR (CDCl$_3$, 500 MHz): δ (ppm) 5.62 (t, 1H), 4.65 (dt,

2H), 3.76 (s, 2H), 3.30 (t, 2H), 2.06 (m, 1H), 1.09 (m, 1H), 1.77 (m, 1H), 1.67 (s, 3H), 1.50 (m, 3H), 1.41 (m, 1H), 1.42 (m, 7H), 0.82 (t, 3H). ^{13}C-NMR (125 MHz, CDCl$_3$) δ (ppm) 13.1 (1C), 19.80 (1C), 21.65 (2C), 24.91 (1C), 25.42 (1C), 28.77 (1C), 29.50 (1C), 30.73 (1C), 40.19 (1C), 69.02 (1C), 74.15 (1C), 107.55 (1C), 122.96 (1C), 134.00 (1C), 148.99(1C).
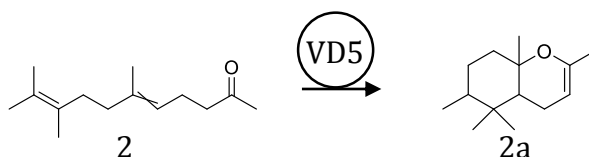
6.8. Enzymatic synthesis

Compound **11** was synthesized by Benjamin Aberle (University Stuttgart) and was provided for this work.[268]

6.8.2. Preparative scale biotransformations

After preparative biotransformations the crude product was first centrifuged to get rid of the cell debris. Second, the aqueous phase containing the product encapsulated by CD was extracted with diethyl ether five times (CD stays in aqueous phase) for **2** and with cyclohexane/ethylacetate (1:1) for **4**-**6**. Next the products were reduced under vacuum, dried over MgSO$_4$, purified over SiO$_2$ column chromatography and characterized by NMR and GC/MS. Due to impurities compound **4a** could not be characterized by NMR and was identified by a commercial product standard via GC/MS.

2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (2a)

Scheme 4. Enzymatic synthesis of **2a** (2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene). Reaction conditions: VD5, 30 °C, 10 mM **2**, 1 l CD buffer with 100 mM Mg(NO$_3$)$_2$.

Yield and spectral data of **2a** (2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene): Yellow oil, 0.163 mmol, 34% yield. 2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (**2a**): ^{1}H-NMR (CDCl$_3$, 500 MHz): δ (ppm) 0.96-1.22 (m, 9H), 1.91-1.93 (m, 6H) 2.72 (s, 5H), 3.38-3.93 (m, 1H), 5.26-5.30 (m, 1H). ^{13}C-NMR (CDCl$_3$, 125 MHz): δ (ppm) 14.65 (1C), 19.00 (1C), 19.28 (1C), 20.52 (1C), 23.74 (1C) 26.63 (1C), 28.77 (1C), 33.64 (1C), 35.36 (1C), 39.29 (1C), 42.21 (1C), 95.28 (1C), 148.02 (1C).

2,3,5,5,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (6a)
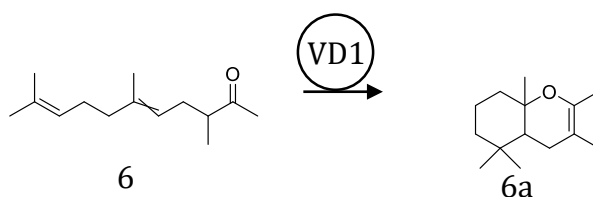
Scheme 5. Enzymatic synthesis of **6a** (2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene). Reaction conditions: VD5, 30 °C, 10 mM **6**, 100 ml CD buffer.

Yield and spectral data of **6a**: Yellow oil, 0.48 mmol, 98% yield. 2,3,5,5,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (**6a**): ^{1}H-NMR (CDCl$_3$, 500 MHz): δ (ppm) 0.82 (s, 4H), 0.92 (s, 4H), 1.12 (s, 3H), 1.25-1.29 (m, 2H), 1.40-1.42 (m, 2H), 1.48-1.52 (m, 2H), 1.56 (s, 2H), 1.60 (s, 3H), 1.69 (s, 3H). ^{13}C-NMR (CDCl$_3$, 125 MHz): δ (ppm) 16.67 (1C), 17.95 (1C), 19.21 (1C), 19.84 (1C), 20.60 (1C) 25.65 (1C), 32.17 (1C), 33.11 (1C), 39.87 (1C), 41.61 (1C), 49.01 (1C), 75.54 (1C), 100.99 (1C), 141.21 (1C).

6.9. Computational studies

6.9.1. *In silico* mutagenesis, ligand and receptor preparation

Models of *Aac*SHC variants were constructed *in silico* by SWISS-MODEL19 or RoseTTAFold.[269,270] The target sequences with the corresponding amino acid substitution in FASTA format were used in the template search to build the model on the crystal structure of the WT *Aac*SHC obtained from the RCSB Protein Data Bank (PDB entry 2SQC). For LCY the model structure with the highest identity (≥ 60%) was used as template. The protonation state of the obtained variant structures was computed by H++ 3.0 web tool at pH 6, 0.01 molar salinity and with an internal dielectric constant of 10 and external of 80 when models were constructed with SWISS-MODEL19.[269] Subsequent energy minimization of all obtained variant structures and tested substrates with YASARA (version 19.12.14.W.64) gave the primed models.[271]

6.9.2. MD refinement and simulation

The obtained enzyme models were subjected to MD refinement for 500 ps. The resulting lowest energy snapshots were used for MD simulation in the AMBER14 force field over a period of 25 ns using the default settings with a density of 0.997 g/ml, 0.9% NaCl and pH 7.4 at 298 K in a water filled simulation cell. A total of 250 snapshots were generated. Calculation of root-mean-square deviation (RMSD) values and b-factors by YASARA was based on AMBER14.[271]

6.9.3. Molecular docking in the active site

The substrates were docked into the crystal structure of *Aac*SHC WT (PDB entry 2SQC) and the simulated variants by YASARA docking tools (version 19.12.14.W.64) based on AutoDock VINA algorithms.[271] The simulation cell was defined by a box of 20x20x20 Å covering the active site of the *Aac*SHC variants. The resulting binding energies of 25 runs were clustered. Best rated ligand structures based on the binding energies with the isoprene unit oriented towards D376 were selected for visualization.

6.9.4. Calculation of molecular tunnel

CAVER 3.0 plugin for PyMOL was used to identify protein tunnels in the *Aac*SHC WT or variant's average structure of all snapshots calculated by YASARA.[195,271] Tunnels were calculated using the default setting and a shell radius of 3 Å and shell depth of 4 Å. D376 was set as the starting point for the tunnel calculation.

6.9.5. Multiple sequence alignment

Homologous sequences of *Aac*SHC and *Can*LCY-B, *Ath*LCY-B and *Ath*LCY-E were searched using the Protein Blast sequence similarity search with Point Accepted Mutation Matrix (PAM 30) in the target database of UniprotKB reference genomes and Swiss-Prot. The database was screened for 250 sequences and the E-value to limit the hits to the most significant ones was set to 10.

6.9.6. Phylogenetic tree construction

Sequences of different cyclase representatives from different enzyme classes and families were retrieved from BRENDA and combined with the sequences of interest for this work. In the MEGA11 software ClustalW sequence alignment was performed.[272] Phylogenetic analysis was performed by selection of default settings and the Maximum Likelihood method. The bootstrap consensus tree represents the evolutionary history of the analyzed taxa. Original trees for the heuristic search were automatically generated by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated with a Jones-Taylor-Thornton (JTT) model and the topology with the superior log likelihood value. The analysis included 57 amino acid sequences.

References

1.	Schell, K. *et al.* Alternative Active Site Confinement in Squalene–Hopene Cyclase Enforces Substrate Preorganization for Cyclization. *ACS Catal* **13**, (2023).

2.	Rothenberg, G. *Catalysis: Concepts and Green Applications. Catalysis: Concepts and Green Applications* (2008).

3.	Mitschke, B., Turberg, M. & List, B. Confinement as a Unifying Element in Selective Catalysis. *Chem* **6**, (2020).

4.	William P. Jencks. Catalysis in Chemistry and Enzymology. *Courier Corporation, Science* (1987).

5.	Thomas, J. M. & Williams, R. J. P. Catalysis: Principles, progress, prospects. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* vol. 363 (2005).

6.	Knowles, R. R. & Jacobsen, E. N. Attractive noncovalent interactions in asymmetric catalysis: Links between enzymes and small molecule catalysts. *Proc Natl Acad Sci U S A* **107**, (2010).

7.	Gaeta, C. *et al.* The Hexameric Resorcinarene Capsule at Work: Supramolecular Catalysis in Confined Spaces. *Chemistry - A European Journal* **25**, (2019).

8.	Kaphan, D. M., Toste, F. D., Bergman, R. G. & Raymond, K. N. Enabling New Modes of Reactivity via Constrictive Binding in a Supramolecular-Assembly-Catalyzed Aza-Prins Cyclization. *J. Am. Chem. Soc* **137**, (2015).

9.	Bastian, S. A., Hammer, S. C., Kreß, N., Nestl, B. M. & Hauer, B. Selectivity in the Cyclization of Citronellal Introduced by Squalene Hopene Cyclase Variants. *ChemCatChem* **9**, (2017).

10.	Hargittai, I. The 2021 chemistry Nobel laureates and asymmetric organocatalysis. *Struct Chem* **33**, (2022).

11.	Lelais, G. & MacMillan, D. W. C. Modern strategies in organic catalysis: The advent and development of iminium activation. *Aldrichimica Acta* vol. 39 (2006).

12.	Mukherjee, S., Yang, J. W., Hoffmann, S. & List, B. Asymmetric enamine catalysis. *Chemical Reviews* vol. 107 (2007).

13.	Simón, L. & Goodman, J. M. Enzyme catalysis by hydrogen bonds: The balance between transition State binding and substrate binding in oxyanion holes. *Journal of Organic Chemistry* **75**, (2010).

14.	Ballester, P. Effects of Nanoconfinement on Catalysis. Edited by Rinaldo Poli. *Angewandte Chemie International Edition* **56**, (2017).

15.	Zou, Y., Li, C., Brunzelle, J. S. & Nair, S. K. Molecular basis for substrate selectivity and specificity by an LPS biosynthetic enzyme. *Biochemistry* **46**, (2007).

16. Weisz, P. B. Molecular shape selective catalysis. *Pure and Applied Chemistry* **52**, (1980).

17. Li, Q., Levi, S. M., Wagen, C. C., Wendlandt, A. E. & Jacobsen, E. N. Site-selective, stereocontrolled glycosylation of minimally protected sugars. *Nature* **608**, (2022).

18. Takezawa, H., Kanda, T., Nanjo, H. & Fujita, M. Site-Selective Functionalization of Linear Diterpenoids through U-Shaped Folding in a Confined Artificial Cavity. *J. Am. Chem. Soc* **141**, (2019).

19. Zhang, Q., Rinkel, J., Goldfuss, B., Dickschat, J. S. & Tiefenbacher, K. Sesquiterpene cyclizations catalysed inside the resorcinarene capsule and application in the short synthesis of isolongifolene and isolongifolenone. *Nat Catal* **1**, (2018).

20. Properzi, R. *et al.* Catalytic enantiocontrol over a non-classical carbocation. *Nat Chem* **12**, (2020).

21. Gaeta, C. *et al.* Supramolecular Catalysis with Self-Assembled Capsules and Cages: What Happens in Confined Spaces. *ChemCatChem* **13**, (2021).

22. Kobayashi, T., Hoppmann, C., Yang, B. & Wang, L. Using Protein-Confined Proximity to Determine Chemical Reactivity. *J Am Chem Soc* **138**, (2016).

23. Blomberg, R. *et al.* Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **503**, (2013).

24. Xu, J. *et al.* Evolution of shape complementarity and catalytic efficiency from a primordial antibody template. *Science* **286**, (1999).

25. Mouarrawis, V., Plessius, R., van der Vlugt, J. I. & Reek, J. N. H. Confinement effects in catalysis using well-defined materials and cages. *Frontiers in Chemistry* vol. 6 (2018).

26. Fiedler, D. *et al.* Supramolecular Catalysis of a Unimolecular Transformation: Aza-Cope Rearrangement within a Self-Assembled Host. *Angew. Chem. Int. Ed* **116**, (2004).

27. Hastings, C. J., Pluth, M. D., Bergman, R. G. & Raymond, K. N. Enzymelike Catalysis of the Nazarov Cyclization by Supramolecular Encapsulation. **132**, (2010).

28. Steitz, T. A., Harrison, R., Weber, I. T. & Leahy, M. Ligand-Induced Conformational Changes in Proteins. in *Ciba Foundation Symposium 93 - Mobility and Function in Proteins and Nucleic Acids* (2008).

29. Pompliano, D. L., Peyman, A. & Knowles, J. R. Stabilization of a Reaction Intermediate as a Catalytic Device: Definition of the Functional Role of the Flexible Loop in Triosephosphate Isomerase. *Biochemistry* **29**, (1990).

30. Pluth, M. D., Bergman, R. G. & Raymond, K. H. Acid catalysis in basic solution: A supramolecular host promotes orthoformate hydrolysis. *Science* **316**, (2007).

31. Koshland, D. E. The Key–Lock Theory and the Induced Fit Theory. *Angewandte Chemie International Edition in English* vol. 33 (1995).

32. Weikl, T. R. & Paul, F. Conformational selection in protein binding and function. *Protein Science* vol. 23 (2014).

33. Wendt, K. U., Feil, C., Lenhart, a, Poralla, K. & Schulz, G. E. Crystallization and preliminary X-ray crystallographic analysis of squalene-hopene cyclase from *Alicyclobacillus acidocaldarius*. *Protein Sci* **6**, (1997).

34. Dusselier, M., Van Wouwe, P., Dewaele, A., Jacobs, P. A. & Sels, B. F. Shape-selective zeolite catalysis for bioplastics production. *Science* **349**, (2015).

35. De Clercq, R. *et al.* Confinement Effects in Lewis Acid-Catalyzed Sugar Conversion: Steering Toward Functional Polyester Building Blocks. *ACS Catal* **5**, (2015).

36. Tsuji, N. *et al.* Activation of olefins via asymmetric Brønsted acid catalysis. *Science* **359**, (2018).

37. Monnereau, L., Sémeril, D. & Matt, D. Synthesis of calixarene-based bis(iminophosphoranes) and their use in Suzuki-Miyaura cross-coupling. *European J Org Chem* (2012).

38. Bhatia, S. *Zeolite Catalysis: Principles and Applications. Zeolite Catalysis: Principles and Applications* (2020).

39. Venuto, P. B. & Landis, P. S. Organic reactions catalyzed by crystalline aluminosilicates. III. Condensation reactions of carbonyl compounds. *J Catal* **6**, (1966).

40. Corma, A., Nemeth, L. T., Renz, M. & Valencia, S. Sn-zeolite beta as a heterogeneous chemoselective catalyst for Baeyer-Villiger oxidations. *Nature* **412**, (2001).

41. Metrano, A. J. & Miller, S. J. Peptide-Based Catalysts Reach the Outer Sphere through Remote Desymmetrization and Atroposelectivity. *Acc Chem Res* **52**, (2019).

42. Breslow, R. & Campbell, P. Selective Aromatic Substitution within a Cyclodextrin Mixed Complex. *Journal of the American Chemical Society* vol. 91 (1969).

43. Breslow, R. & Overman, L. E. An "Artificial Enzyme" Combining a Metal Catalytic Group and a Hydrophobic Binding Cavity. *Journal of the American Chemical Society* vol. 92 (1970).

44. Caulder, D. L., Powers, R. E., Parac, T. N. & Raymond, K. N. The self-assembly of a predesigned tetrahedral M4L4 supramolecular cluster. *Angewandte Chemie - International Edition* **37**, (1998).

45. Hart-Cooper, W. M., Clary, K. N., Toste, F. D., Bergman, R. G. & Raymond, K. N. Selective monoterpene-like cyclization reactions achieved by water exclusion from reactive intermediates in a supramolecular catalyst. *J Am Chem Soc* **134**, (2012).

46. Zhang, P., Tsuji, N., Ouyang, J. & List, B. Strong and Confined Acids Catalyze Asymmetric Intramolecular Hydroarylations of Unactivated Olefins with Indoles. *J Am Chem Soc* **143**, (2021).

47. Tsui, G. C., Liu, L. & List, B. The organocatalytic asymmetric prins cyclization. *Angewandte Chemie - International Edition* **54**, (2015).

48. Taylor, M. S. Catalysis based on reversible covalent interactions of organoboron compounds. *Acc Chem Res* **48**, (2015).

49. Peng, B. *et al.* A Powerful Chiral Super Brønsted C-H Acid for Asymmetric Catalysis. *J Am Chem Soc* **144**, (2022).

50. Merad, J., Lalli, C., Bernadat, G., Maury, J. & Masson, G. Enantioselective Brønsted Acid Catalysis as a Tool for the Synthesis of Natural Products and Pharmaceuticals. *Chemistry - A European Journal* **24**, (2018).

51. Akiyama, T. Stronger Bronsted Acids. *Chem Rev* **107**, (2007).

52. Uraguchi, D. & Terada, M. Chiral Brønsted Acid-Catalyzed Direct Mannich Reactions via Electrophilic Activation. *J Am Chem Soc* **126**, (2004).

53. Akiyama, T., Itoh, J., Yokota, K. & Fuchibe, K. Enantioselective Mannich-type reaction catalyzed by a chiral Brønsted acid. *Angewandte Chemie - International Edition* **43**, (2004).

54. Čorić, I. & List, B. Asymmetric spiroacetalization catalysed by confined Brønsted acids. *Nature* **483**, (2012).

55. Akiyama, T. & Mori, K. Stronger Brønsted Acids: Recent Progress. *Chemical Reviews* vol. 115 (2015).

56. Akiyama, T., Itoh, J. & Fuchibe, K. Recent progress in chiral Brønsted acid catalysis. *Advanced Synthesis and Catalysis* vol. 348 (2006).

57. Bordwell, F. G. *Equilibrium Acidities in Dimethyl Sulfoxide Solution. Acc. Chem. Res* vol. 21 (1988).

58. Kampen, Daniela; Reisinger, Corinna M.; List, B. Chiral Brønsted Acids for Asymmetric Organocatalysis. in *Topics in Current Chemistry* (2009).

59. Doyle, A. G. & Jacobsen, E. N. Small-molecule H-bond donors in asymmetric catalysis. *Chemical Reviews* vol. 107 (2007).

60. Jakab, G., Tancon, C., Zhang, Z., Lippert, K. M. & Schreiner, P. R. (Thio)urea organocatalyst equilibrium acidities in DMSO. *Org Lett* **14**, (2012).

61. Sigman, M. S. & Jacobsen, E. N. Schiff base catalysts for the asymmetric strecker reaction identified and optimized from parallel synthetic libraries. *J Am Chem Soc* **120**, (1998).

62. Christ, P. *et al.* PKa values of chiral Brønsted acid catalysts: Phosphoric acids/amides, sulfonyl/sulfuryl imides, and perfluorinated TADDOLs (TEFDDOLs). *Chemistry - A European Journal* **17**, (2011).

63. Malerich, J. P., Hagihara, K. & Rawal, V. H. Chiral Squaramide Derivatives are Excellent Hydrogen Bond Donor Catalysts. *J Am Chem Soc* **130**, (2008).

64.　Hoffmann, S. *et al.* A Powerful Brønsted Acid Catalyst for the Organocatalytic Asymmetric Transfer Hydrogenation of Imines. *Angewandte Chemie - International Edition* **44**, (2005).

65.　Guo, Q. S., Du, D. M. & Xu, J. The development of double axially chiral phosphoric acids and their catalytic transfer hydrogenation of quinolines. *Angewandte Chemie - International Edition* **47**, (2008).

66.　Mori, K., Ehara, K., Kurihara, K. & Akiyama, T. Selective activation of enantiotopic C(sp3)-hydrogen by means of chiral phosphoric acid: Asymmetric synthesis of tetrahydroquinoline derivatives. *J Am Chem Soc* **133**, (2011).

67.　Hashimoto, T. & Maruoka, K. Design of Axially Chiral Dicarboxylic Acid for Asymmetric Mannich Reaction of Arylaldehyde N-Boc Imines and Diazo Compounds. *J Am Chem Soc* **129**, (2007).

68.　Zhou, W., Xu, L.-W., Li, L., Yang, L. & Xia, C.-G. Enantioselective Michael-Type Friedel-Crafts Reactions of Indoles to Enones Catalyzed by a Chiral Camphor-Based Brønsted Acid. *Eur. J. Org. Chem. 2006* (2006).

69.　Chen, L. Y., He, H., Chan, W. H. & Lee, A. W. M. Chiral sulfonimide as a brønsted acid organocatalyst for asymmetric Friedel-Crafts alkylation of indoles with imines. *Journal of Organic Chemistry* **76**, (2011).

70.　Rueping, M. *et al.* Chiral Brønsted Acids in the Catalytic Asymmetric Nazarov Cyclization-The First Enantioselective Organocatalytic Electrocyclic Reaction. *Angew. Chem. Int. Ed* **46**, (2007).

71.　Sun, F.-L. *et al.* Enantioselective Synthesis of Fluorene Derivatives by Chiral Phosphoric Acid Catalyzed Tandem Double Friedel-Crafts Reaction. *Chem. Eur. J* **15**, (2009).

72.　Hong Cheon, C. & Yamamoto, H. A Brønsted Acid Catalyst for the Enantioselective Protonation Reaction. *J. Am. Chem. Soc.* **130**, (2008).

73.　Schwengers, S. A. *et al.* Unified Approach to Imidodiphosphate-Type Brønsted Acids with Tunable Confinement and Acidity. *J Am Chem Soc* **143**, (2021).

74.　Zou, Y. *et al.* Enzyme-catalyzed cationic epoxide rearrangements in quinolone alkaloid biosynthesis. *Nat Chem Biol* **13**, (2017).

75.　Pronin, S. V. & Shenvi, R. A. Synthesis of highly strained terpenes by non-stop tail-to-head polycyclization. *Nat Chem* **4**, (2012).

76.　Schowen, K. B., Limbach, H. H., Denisov, G. S. & Schowen, R. L. Hydrogen bonds and proton transfer in general-catalytic transition-state stabilization in enzyme catalysis. *Biochimica et Biophysica Acta - Bioenergetics* vol. 1458 (2000).

77.　Shan, S. O. & Herschlag, D. The change in hydrogen bond strength accompanying charge rearrangement: Implications for enzymatic catalysis. *Proc Natl Acad Sci U S A* **93**, (1996).

78.　Fersht, A. *Structure and mechanism in protein science* vol. 13409 (1999).

79. Bornscheuer, U. T. & Kazlauskas, R. J. *Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations: Second Edition*. *Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations: Second Edition* (2006).

80. Hedstrom, L. Serine protease mechanism and specificity. *Chem Rev* **102**, (2002).

81. Torbeev, V. Y. *et al.* Protein conformational dynamics in the mechanism of HIV-1 protease catalysis. *Proc Natl Acad Sci U S A* **108**, (2011).

82. Hammer, S. C., Marjanovic, A., Dominicus, J. M., Nestl, B. M. & Hauer, B. Squalene hopene cyclases are protonases for stereoselective Brønsted acid catalysis. *Nat Chem Biol* **11**, 1–39 (2014).

83. Gandour, R. D. On the importance of orientation in general base catalysis by carboxylate. *Bioorg Chem* **10**, (1981).

84. Gao, J. & Pavelites, J. J. Aqueous Basicity of the Carboxylate Lone Pairs and the C-0 Barrier in Acetic Acid: A Combined Quantum and Statistical Mechanical Study. *J Am Chem Soc* **114**, (1992).

85. Wendt, K. U., Poralla, K. & Schulz, G. E. Structure and function of a squalene cyclase. *Science* **277**, (1997).

86. Olah, G. A. The General Concept and Structure of Carbocations Based on Differentiation of Trivalent ("Classical") Carbenium Ions from Three-Center Bound Penta- or Tetracoordinated ("Nonclassical") Carbonium Ions. the Role of Carbocations in Electrophilic Reactions. *J Am Chem Soc* **94**, (1972).

87. Sato, H., Mitsuhashi, T., Yamazaki, M., Abe, I. & Uchiyama, M. Computational Studies on Biosynthetic Carbocation Rearrangements Leading to Quiannulatene: Initial Conformation Regulates Biosynthetic Route, Stereochemistry, and Skeleton Type. *Angewandte Chemie - International Edition* **57**, (2018).

88. Naredla, R. R. & Klumpp, D. A. Contemporary carbocation chemistry: Applications in organic synthesis. *Chemical Reviews* vol. 113 (2013).

89. Aue, D. H. Carbocations. *Wiley Interdiscip Rev Comput Mol Sci* **1**, 487–508 (2011).

90. Wendlandt, A. E., Vangal, P. & Jacobsen, E. N. Quaternary stereocentres via an enantioconvergent catalytic SN1 reaction. *Nature* **556**, (2018).

91. Ofial, A. R. & Mayr, H. Reactivities of carbocations and carbanions. in *Macromolecular Symposia* vol. 215 (2004).

92. Boronat, M., Viruela, P. M. & Corma, A. Reaction Intermediates in Acid Catalysis by Zeolites: Prediction of the Relative Tendency to Form Alkoxides or Carbocations as a Function of Hydrocarbon Nature and Active Site Structure. *J Am Chem Soc* **126**, (2004).

93. Arnett, E. M. & Hofelich, T. C. *Stabilities of Carbocations in Solution. 14. An Extended Thermochemical Scale of Carbocation Stabilities in a Common Superacid*. *J. Am. Chem. Soc* vol. 105 (1983).

94. Tantillo, D. J. The carbocation continuum in terpene biosynthesis—where are the secondary cations? *Chem Soc Rev* **39**, (2010).

95. Gutta, P. & Tantillo, D. J. Theoretical studies on farnesyl cation cyclization: Pathways to pentalenene. *J Am Chem Soc* **128**, (2006).

96. Hess, B. A. Concomitant C-ring expansion and D-ring formation in lanosterol biosynthesis from squalene without violation of Markovnikov's rule. *J Am Chem Soc* **124**, (2002).

97. Zhang, X. & Tan, C. H. Stereospecific and stereoconvergent nucleophilic substitution reactions at tertiary carbon centers. *Chem* vol. 7 (2021).

98. Isomura, M., Petrone, D. A. & Carreira, E. M. Coordination-Induced Stereocontrol over Carbocations: Asymmetric Reductive Deoxygenation of Racemic Tertiary Alcohols. *J Am Chem Soc* **141**, (2019).

99. De Rosa, M. *et al.* Carbocation catalysis in confined space: activation of trityl chloride inside the hexameric resorcinarene capsule. *Chem Sci* **13**, (2022).

100. Mischko, W., Hirte, M., Fuchs, M., Mehlmer, N. & Brück, T. B. Identification of sesquiterpene synthases from the Basidiomycota Coniophora puteana for the efficient and highly selective β-copaene and cubebol production in *E. coli. Microb Cell Fact* **17**, (2018).

101. Lopez-Gallego, F., Agger, S. A., Abate-Pella, D., Distefano, M. D. & Schmidt-Dannert, C. Sesquiterpene synthases Cop4 and Cop6 from Coprinus cinereus: Catalytic promiscuity and cyclization of farnesyl pyrophosphate geometric isomers. *ChemBioChem* **11**, (2010).

102. Justicia, J. *et al.* Total synthesis of 3-hydroxydrimanes mediated by titanocene(III) - Evaluation of their antifeedant activity. *European J Org Chem* **2005**, (2005).

103. Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural world. *Nat Chem Biol* **3**, (2007).

104. Hill, R. A. & Connolly, J. D. *Dictionary of Terpenoids.* (1991).

105. Chen, X. *et al.* Terpene synthase genes in eukaryotes beyond plants and fungi: Occurrence in social amoebae. *Proc Natl Acad Sci U S A* **113**, (2016).

106. Crowell, P. L. Prevention and Therapy of Cancer by Dietary Monoterpenes. *J Nutr* **129**, (1999).

107. Kris-Etherton, P. M. *et al.* Bioactive compounds in foods: Their role in the prevention of cardiovascular disease and cancer. *American Journal of Medicine* **113**, (2002).

108. Armaka, M., Papanikolaou, E., Sivropoulou, A. & Arsenakis, M. Antiviral properties of isoborneol, a potent inhibitor of herpes simplex virus type 1. *Antiviral Res* **43**, (1999).

109. Eichhorn, E. *et al.* Biocatalytic Process for (-)-Ambrox Production Using Squalene Hopene Cyclase. *Adv. Synth. Catal.* **360**, (2018).

110. Poulter, C. D. & Rilling, H. C. The Prenyl Transfer Reaction. Enzymatic and Mechanistic Studies of the 1'-4 Coupling Reaction in the Terpene Biosynthetic Pathway. *Acc Chem Res* **11**, (1978).

111. Wendt, K. U. & Schulz, G. E. Isoprenoid biosynthesis: Manifold chemistry catalyzed by similar enzymes. *Structure* **6**, (1998).

112. Oldfield, E. & Lin, F. Y. Terpene biosynthesis: Modularity rules. *Angewandte Chemie - International Edition* **51**, (2012).

113. Dale Poulter, C. Biosynthesis of Non-Head-to-Tail Terpenes. Formation of l'-l and 1'-3 Linkages. *Acc Chem Res* **23**, (1990).

114. Oldfield, E., Lin, F.-Y., Oldfield, E. & Lin, F. Terpen-Biosynthese: Modularitätsregeln Angewandte Aufsätze. **124**, (2012).

115. Allemann, R. K. Chemical wizardry? The generation of diversity in terpenoid biosynthesis. in *Pure and Applied Chemistry* vol. 80 (2008).

116. Hur, S. & Bruice, T. C. The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proc Natl Acad Sci U S A* **100**, (2003).

117. Hay, S. & Scrutton, N. S. Good vibrations in enzyme-catalysed reactions. *Nature Chemistry* vol. 4 (2012).

118. Spencer, T. A. & Ditchfield, R. Organic & Biomolecular Chemistry A simpler method affords evaluation of π stabilization by phenylalanine of several biochemical carbocations. *Org. Biomol. Chem* **18**, (2020).

119. Allemann, R. K., Young, N. J., Ma, S., Truhlar, D. G. & Gao, J. Synthetic Efficiency in Enzyme Mechanisms Involving Carbocations: Aristolochene Synthase. (2007).

120. Matsuda, S. P. T., Wilson, W. K. & Xiong, Q. Mechanistic insights into triterpene synthesis from quantum mechanical calculations. Detection of systematic errors in B3LYP cyclization energies. *Org. Biomol. Chem.* **4**, (2006).

121. Pazouki, L. & Niinemetst, U. Multi-substrate terpene synthases: Their occurrence and physiological significance. *Front Plant Sci* **7**, (2016).

122. Martin, D. M. *et al.* Functional Annotation, Genome Organization and Phylogeny of the Grapevine (*Vitis vinifera*) Terpene Synthase Gene Family Based on Genome Assembly, FLcDNA Cloning, and Enzyme Assays. *BMC Plant Biol* **10**, (2010).

123. Jones, C. G. *et al.* Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (TPS)-a and TPS-b subfamilies, including santalene synthases. *Journal of Biological Chemistry* **286**, (2011).

124. Hayashi, K. ichiro *et al.* Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. *FEBS Lett* **580**, (2006).

125. Karunanithi, P. S. & Zerbe, P. Terpene Synthases as Metabolic Gatekeepers in the Evolution of Plant Terpenoid Chemical Diversity. *Frontiers in Plant Science* vol. 10 (2019).

126. Leferink, N. G. H. *et al.* Molecular Determinants of Carbocation Cyclisation in Bacterial Monoterpene Synthases. *ChemBioChem* **23**, (2022).

127. Tantillo, D. J. Importance of Inherent Substrate Reactivity in Enzyme-Promoted Carbocation Cyclization/Rearrangements. *Angew. Chem. Int. Ed* **56**, (2017).

128. Schneider, A., Jegl, P. & Hauer, B. Stereoselective Directed Cationic Cascades Enabled by Molecular Anchoring in Terpene Cyclases. *Angew. Chem. Int. Ed* **60**, (2021).

129. Rabe, P., Rinkel, J., Klapschinski, T. A., Barra, L. & Dickschat, J. S. A method for investigating the stereochemical course of terpene cyclisations. *Org. Biomol. Chem.* **14**, (2016).

130. Pazouki, L., Memari, H. R., Kännaste, A., Bichele, R. & Niinemets, Ü. Germacrene A synthase in yarrow (Achillea millefolium) is an enzyme with mixed substrate specificity: Gene cloning, functional characterization and expression analysis. *Front Plant Sci* **6**, (2015).

131. Abe, I., Tanaka, H. & Noguchi, H. Enzymatic formation of an unnatural hexacyclic C35 polyprenoid by bacterial squalene cyclase. *J Am Chem Soc* **124**, (2002).

132. Loizzi, M., González, V., Miller, D. J. & Allemann, R. K. Nucleophilic Water Capture or Proton Loss: Single Amino Acid Switch Converts δ-Cadinene Synthase into Germacradien-4-ol Synthase. *ChemBioChem* **19**, (2018).

133. Hansen, N. L., Nissen, J. N. & Hamberger, B. Two residues determine the product profile of the class II diterpene synthases TPS14 and TPS21 of *Tripterygium wilfordii*. *Phytochemistry* **138**, (2017).

134. Greenhagen, B. T., O'Maille, P. E., Noel, J. P. & Chappell, J. Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. *Proceedings of the National Academy of Sciences* **103**, (2006).

135. Yoshikuni, Y., Ferrin, T. E. & Keasling, J. D. Designed divergent evolution of enzyme function. *Nature* **440**, (2006).

136. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, (2006).

137. Ajikumar, P. K. *et al.* Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**, (2010).

138. Cascón, O. *et al.* Chemoenzymatic preparation of germacrene analogues. *Chemical Communications* **48**, (2012).

139. Racolta, S., Juhl, P. B., Sirim, D. & Pleiss, J. The triterpene cyclase protein family: A systematic analysis. *Proteins: Structure, Function and Bioinformatics* **80**, (2012).

140. Sato, T. & Hoshino, T. Functional Analysis of the DXDDTA Motif in Squalene-Hopene Cyclase by Site-directed Mutagenesis Experiments: Initiation Site of the Polycyclization Reaction and Stabilization Site of the Carbocation Intermediate of the Initially Cyclized A-Ring. *Biosci Biotechnol Biochem* **63**, (1999).

141. Hoshino, T. & Sato, T. Squalene-hopene cyclase: catalytic mechanism and substrate recognition. *Chemical Communications* (2002).

142. Wendt, K. U., Lenhart, a & Schulz, G. E. The structure of the membrane protein squalene-hopene cyclase at 2.0 A resolution. *J Mol Biol* **286**, (1999).

143. Schwab, F., Van Gunsteren, W. F. & Zagrovic, B. Computational study of the mechanism and the relative free energies of binding of anticholesteremic inhibitors to squalene-hopene cyclase. *Biochemistry* **47**, (2008).

144. Abe, I., Rohmer, M. & Prestwich, G. D. Enzymatic Cyclization of Squalene and Oxidosqualene to Sterols and Triterpenes. *Chem Rev* **93**, (1993).

145. Syrén, P.-O., Hammer, S. C., Claasen, B. & Hauer, B. Entropy is key to the formation of pentacyclic terpenoids by enzyme-catalyzed polycyclization. *Angewandte Chemie - International Edition* **53**, (2014).

146. Hammer, S. C., Syrén, P.-O. & Hauer, B. Substrate Pre-Folding and Water Molecule Organization Matters for Terpene Cyclase Catalyzed Conversion of Unnatural Substrates. *ChemistrySelect* **1**, (2016).

147. Hammer, S. C. Zur Anwendbarkeit von Squalen-Hopen-Zyklasen als chirale Brønsted-Säuren in der asymmetrischen Katalyse. *Thesis* (2014).

148. Sato, T. & Hoshino, T. Kinetic studies on the function of all the conserved tryptophans involved inside and outside the QW motifs of squalene-hopene cyclase: stabilizing effect of the protein structure against thermal denaturation. *Biosci. Biotechnol. Biochem.* **63**, (1999).

149. Reinert, D. J., Balliano, G. & Schulz, G. E. Conversion of Squalene to the Pentacarbocyclic Hopene. *Chem Biol* **11**, (2004).

150. Tanaka, H., Noguchi, H. & Abe, I. Enzymatic cyclization of 26- and 27-methylidenesqualene to novel unnatural C31 polyprenoids by squalene:hopene cyclase. *Tetrahedron Lett* **45**, (2004).

151. Tanaka, H., Noma, H., Noguchi, H. & Abe, I. Enzymatic formation of pyrrole-containing novel cyclic polyprenoids by bacterial squalene:hopene cyclase. *Tetrahedron Lett* **47**, (2006).

152. Pale-Grosdemange, C., Merkofer, T., Rohmer, M. & Poralla, K. Production of bicyclic and tricyclic triterpenes by mutated squalene- hopene cyclase. *Tetrahedron Lett* **40**, (1999).

153. Hoshino, T., Kumai, Y., Kudo, I., Nakano, S. I. & Ohashi, S. Enzymatic cyclization reactions of geraniol, farnesol and geranylgeraniol, and those of truncated squalene analogs having C20 and C25 by recombinant squalene cyclase. *Org Biomol Chem* **2**, (2004).

154. Seitz, M. Characterization of the substrate specificity of squalene-hopene cyclases. *Thesis* (2012).

155. Syrén, P. O., Henche, S., Eichler, A., Nestl, B. M. & Hauer, B. Squalene-hopene cyclases - evolution, dynamics and catalytic scope. *Curr Opin Struct Biol* **41**, 73–82 (2016).

156. Svenja Diether. Squalene-Hopene Cyclase Catalyzed Isomerization of Monoterpenes. *Thesis* (2020).

157. Hoshino, T., Nakano, S. I., Kondo, T., Sato, T. & Miyoshi, A. Squalene - Hopene cyclase: Final deprotonation reaction, conformational analysis for the cyclization of (3*R*,*S*)-2,3-oxidosqualene and further evidence for the requirement of an isopropylidene moiety both for initiation of the polycyclization cascade and for the formation of the 5-membered E-ring. *Org Biomol Chem* **2**, (2004).

158. Seitz, M. *et al.* Substrate specificity of a novel squalene-hopene cyclase from *Zymomonas mobilis*. *J Mol Catal B Enzym* **84**, (2012).

159. Hammer, S. C., Syrén, P.-O., Seitz, M., Nestl, B. M. & Hauer, B. Squalene hopene cyclases: Highly promiscuous and evolvable catalysts for stereoselective CC and CX bond formation. *Curr Opin Chem Biol* **17**, (2013).

160. Seitz, M. *et al.* Synthesis of Heterocyclic Terpenoids by Promiscuous Squalene – Hopene Cyclases. *ChemBioChem* **14**, (2013).

161. Kühnel, L. C., Nestl, B. M. & Hauer, B. Enzymatic Addition of Alcohols to Terpenes by Squalene Hopene Cyclase Variants. *ChemBioChem* **18**, (2017).

162. Hammer, S. C., Dominicus, J. M., Syrén, P.-O., Nestl, B. M. & Hauer, B. Stereoselective Friedel-Crafts alkylation catalyzed by squalene hopene cyclases. *Tetrahedron* **68**, (2012).

163. Nakano, S. I., Ohashi, S. & Hoshino, T. Squalene-hopene cyclase: Insight into the role of the methyl group on the squalene backbone upon the polycyclization cascade. Enzymatic cyclization products of squalene analogs lacking a 26-methyl group and possessing a methyl group at C(7) or C(11). *Org Biomol Chem* **2**, (2004).

164. Rohmer, M., Bouvier, P. & Ourisson, G. *Non-specific Lanosterol and Hopanoid Biosynthesis by a Cell-Free System from the Bacterium Methylococcus cupsulatus. Eur. J. Biochem* vol. 112 (1980).

165. Rohmer, M., Anding, C. & Ourisson, G. Non-specific Biosynthesis of Hopane Triterpenes by a Cell-Free System from *Acetobacter pasteurianum*. *Eur J Biochem* **112**, (1980).

166. Neumann, S. & Simon, H. Purification, Partial Characterization and Substrate Specificity of a Squalene Cyclase from *Bacillus acidocaldarius*. *Biol Chem Hoppe Seyler* **367**, (1986).

167. Takahashi, K., Sasaki, Y. & Hoshino, T. Squalene-Hopene Cyclase : On the Polycyclization Reactions of Squalene Analogues Bearing Ethyl Groups at Positions C-6, C-10, C-15, and C-19. *European J Org Chem* **2018**, (2018).

168. Hoshino, T. & Ohashi, S. Importance of the Methyl Group at C(10) of Squalene for Hopene Biosynthesis and Novel Carbocyclic Skeletons with 6/5 + 5/5 + (6) Ring System(s). *Org. Lett* **4**, (2002).

169. Hoshino, T. & Kondo, T. The cyclization mechanism of squalene in hopene biosynthesis: the terminal methyl groups are critical to the correct folding of this substrate both for the formation of the five-membered E-ring and for the initiation of the polycyclization reaction. *Chem. Commun* (1999).

170. Hoshino, T., Kaneko, I. & Terasawa, Y. Squalene-hopene cyclase: mechanistic insights into the polycyclization cascades of squalene analogs bearing ethyl and hydroxymethyl groups at the C-2 and C-23 positions. *Chemistry - A European Journal* **24**, (2018).

171. Britton, G. Biosynthesis of carotenoids. in *Carotenoids in Photosynthesis* (1993).

172. Kolašinac, S. M., Stevanović, Z. P. D., Kilibarda, S. N. & Kostić, A. Carotenoids: New applications of "old" pigments. *Phyton* vol. 90 Preprint at https://doi.org/10.32604/phyton.2021.015996 (2021).

173. Zhao, Z., Liu, Z. & Mao, X. Biotechnological advances in lycopene β-Cyclases. *Journal of Agricultural and Food Chemistry* vol. 68 (2020).

174. Sandmann, G. Molecular evolution of carotenoid biosynthesis from bacteria to plants. *Physiologia Plantarum* vol. 116 (2002).

175. Song, W. *et al.* Functional characterization and comparison of lycopene epsilon-cyclase genes in Nicotiana tabacum. *BMC Plant Biol* **22**, (2022).

176. Yu, Q. *et al.* The lycopene cyclase CrtY from Pantoea ananatis (formerly Erwinia uredovora) catalyzes an FADred-dependent non-redox reaction. *Journal of Biological Chemistry* **285**, (2010).

177. Cunningham, F. X. *et al.* Functional Analysis of the B and E Lycopene Cyclase Enzymes of Arabidopsis Reveals a Mechanism for Control of Cyclic Carotenoid Formation. *Plant Cell* **8**, (1996).

178. An, G. H., Cho, M. H. & Johnson, E. A. Monocyclic carotenoid biosynthetic pathway in the yeast *Phaffia rhodozyma* (Xanthophyllomyces dendrorhous). *J Biosci Bioeng* **88**, (1999).

179. Ramaprasad, E. V. V., Sasikala, C. & Ramana, C. V. Neurosporene is the major carotenoid accumulated by *Rhodobacter viridis* JA737. *Biotechnol Lett* **35**, (2013).

180. Moreno, J. C. *et al.* Levels of Lycopene β-Cyclase 1 Modulate Carotenoid Gene Expression and Accumulation in Daucus carota. *PLoS One* **8**, (2013).

181. Britton, G., Liaaen-Jensen, S. & Pfänder, H. *Carotenoids. Vol 3. Biosynthesis and metabolism. Carotenoids* vol. 3 (1998).

182. Armstrong, G. Carotenoid Genetics and Biochemistry. in *Comprehensive Natural Products Chemistry* (1999).

183. Britton, G., Lockley, W. J. S., Patel, N. J. & Goodwin, T. W. The use of deuterium from deuterium oxide as a label in studies of biosynthetic pathways Carotenoid transformations in a Flavobacterium species. *FEBS Lett* **79**, (1977).

184. Mialoundama, A. S. *et al.* Characterization of plant carotenoid cyclases as members of the flavoprotein family functioning with no net redox change. *Plant Physiol* **153**, (2010).

185. Elleuch, F. *et al.* Carotenoids overproduction in Dunaliella sp.: Transcriptional changes and new insights through lycopene cyclase regulation. *Applied Sciences* **9**, (2019).

186. Hoshino, T. β-Amyrin biosynthesis: catalytic mechanism and substrate recognition. *Org. Biomol Chem.* **15**, (2017).

187. Eichenberger, M. *et al.* Asymmetric Cation-Olefin Monocyclization by Engineered Squalene-Hopene Cyclases. *Angew. Chem. Int. Ed* **60**, (2021).

188. Heinemann, P. M., Armbruster, D. & Hauer, B. Active-site loop variations adjust activity and selectivity of the cumene dioxygenase. *Nature Communication* (2019).

189. Rapp, L. R. *et al.* Substrate Anchoring and Flexibility Reduction in CYP153AM.aqLeads to Highly Improved Efficiency toward Octanoic Acid. *ACS Catal* **11**, (2021).

190. Nestl, B. M. & Hauer, B. Engineering of flexible loops in enzymes. *ACS Catalysis* vol. 4 (2014).

191. Kaneko, I., Te Rasawa, U., Utomu Hoshino, T. Squalene-Hopene Cyclase: Mechanistic Insights into the Polycyclization Cascades of Squalene Analogs Bearing Ethyl and Hydroxymethyl Groups at the C-2 and C-23 Positions.

192. Henche, S. Squalen-Hopen Zyklasen vermittelte Friedel-Crafts Alkylierung. *Thesis*(2019).

193. Lu, Z. *et al.* Tunnel engineering to accelerate product release for better biomass-degrading abilities in lignocellulolytic enzymes. *Biotechnol Biofuels* **12**, (2019).

194. Kokkonen, P., Bednar, D., Pinto, G., Prokop, Z. & Damborsky, J. Engineering enzyme access tunnels. *Biotechnology Advances* vol. 37 (2019).

195. Chovancová, Eva; Pavelka, Antonín; Beneš, Petr; Strnad, Ondrej; Brezovský, Jan; Kozlikova, Barbora; Gora, Artur; Šustr, Vilém; Klvana, Martin; Medek, Petr; Biedermannová, Lada; Sochor, Jirí; Damborský, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput Biol* **8**, (2012).

196. Oprea, T. I., Hummer, G. & García, A. E. Identification of a functional water channel in cytochrome P450 enzymes. *Proc Natl Acad Sci U S A* **94**, (1997).

197. Zhang, Y. & Cremer, P. S. Interactions between macromolecules and ions: the Hofmeister series. *Curr Opin Chem Biol* **10**, (2006).

198. Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Archiv für Experimentelle Pathologie und Pharmakologie* **25**, (1888).

199. Siedenburg, G. & Jendrossek, D. Squalene-hopene cyclases. *Appl Environ Microbiol* **77**, (2011).

200. Wang, F. *et al.* CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification. *Anal Chem* **93**, (2021).

201. Kolesnikova, M. D., Xiong, Q., Lodeiro, S., Hua, L. & Matsuda, S. P. T. Lanosterol biosynthesis in plants. *Arch Biochem Biophys* **447**, (2006).

202. Gas-Pascual, E., Berna, A., Bach, T. J. & Schaller, H. Plant oxidosqualene metabolism: Cycloartenol synthase-dependent sterol biosynthesis in nicotiana benthamiana. *PLoS One* **9**, (2014).

203. Degenhardt, J., Köllner, T. G. & Gershenzon, J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* vol. 70 (2009).

204. Koc, I., Filiz, E. & Tombuloglu, H. Comparative analysis of plant lycopene cyclases. *Comput Biol Chem* **58**, (2015).

205. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* **35**, (2018).

206. Kunjapur, A. M., Tarasova, Y. & Prather, K. L. J. Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli. J Am Chem Soc* **136**, (2014).

207. Rodriguez, G. M. & Atsumi, S. Toward aldehyde and alkane production by removing aldehyde reductase activity in *Escherichia coli. Metab Eng* **25**, (2014).

208. Masyita, A. *et al.* Terpenes and terpenoids as main bioactive compounds of essential oils, their roles in human health and potential application as natural food preservatives. *Food Chem X* **13**, (2022).

209. De Alwis, R., Fujita, K., Ashitani, T. & Kuroda, K. Volatile and non-volatile monoterpenes produced by elicitor-stimulated *Cupressus lusitanica* cultured cells. *J Plant Physiol* **166**, (2009).

210. Avasthi, K., Shukla, L., Kant, R. & Ravikumar, K. Folded conformations due to arene interactions in dissymmetric and symmetric butylidene-linker models based on pyrazolo[3,4-d]pyrimidine, purine and 7-deazapurine. *Acta Crystallogr C Struct Chem* **70**, (2014).

211. Merlini, V., Luparia, M., Porta, A., Zanoni, G. & Vidari, G. Biomimetic cyclization of geraniol derivatives, a useful tool in the total synthesis of bioactive monocyclic terpenoids. *Natural Product Communications* vol. 6 (2011).

212. Brunoldi, E., Luparia, M., Porta, A., Zanoni, G. & Vidari, G. Biomimetic Cyclizations of Functionalized Isoprenoid Polyenes: A Cornucopia of Synthetic Opportunities. *Curr Org Chem* **10**, (2006).

213. Raptis, C., Lykakis, I. N., Tsangarakis, C. & Stratakis, M. Acid-Catalyzed Cyclization of Terpenes Under Homogeneous and Heterogeneous Conditions as Probed Through Stereoisotopic Studies: A Concerted Process with Competing Preorganized Chair and Boat Transition States. *Chemistry – A European Journal* **15**, (2009).

214. Jung, S. T., Lauchli, R. & Arnold, F. H. Cytochrome P450: taming a wild type enzyme. *Curr Opin Biotechnol* **22**, (2011).

215. De Luca, V. & Mandrich, L. Enzyme Promiscuous Activity: How to Define it and its Evolutionary Aspects. *Protein Pept Lett* **27**, (2020).

216. Tokuriki, N. & Tawfik, D. S. Protein dynamism and evolvability. *Science* **324**, (2009).

217. Romero, M. L., Garcia Seisdedos, H. & Ibarra-Molero, B. Active site center redesign increases protein stability preserving catalysis in thioredoxin. *Protein Science* **31**, (2022).

218. Henche, S., Nestl, B. M. & Hauer, B. Enzymatic Friedel-Crafts Alkylation Using Squalene-Hopene Cyclases. *ChemCatChem* **13**, (2021).

219. Ishida, T. Effects of point mutation on enzymatic activity: Correlation between protein electronic structure and motion in chorismate mutase reaction. *J Am Chem Soc* **132**, (2010).

220. Toscano, M. D., Woycechowsky, K. J. & Hilvert, D. Minimalist active-site redesign: Teaching old enzymes new tricks. *Angewandte Chemie - International Edition* **46**, (2007).

221. Dill, K. A. Dominant forces in protein folding. *Biochemistry* **29**, 7133–55 (1990).

222. Harms, V. *et al.* Methyl-Shifted Farnesyldiphosphate Derivatives Are Substrates for Sesquiterpene Cyclases. *Org. Lett* **22**, (2020).

223. Ignea, C. *et al.* Synthesis of 11-carbon terpenoids in yeast using protein and metabolic engineering. *Nat Chem Biol* **14**, (2018).

224. Smentek, L. & Hess, B. A. Compelling computational evidence for the concerted cyclization of the ABC rings of hopene from protonated squalene. *J Am Chem Soc* **132**, (2010).

225. Kokkinidis, M., Glykos, N. M. & Fadouloglou, V. E. Protein flexibility and enzymatic catalysis. in *Advances in Protein Chemistry and Structural Biology* vol. 87 (2012).

226. Pavlova, M. *et al.* Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat Chem Biol* **5**, (2009).

227. Ochs, D., Kaletta, C., Entian, K. D., Beck-Sickinger, A. & Poralla, K. Cloning, expression, and sequencing of squalene-hopene cyclase, a key enzyme in triterpenoid metabolism. *J Bacteriol* **174**, (1992).

228. Robert A. Horton, Moran, L. A., Scrimgeour, G., Perry, M. & Rawn, D. 2.4 Unpolare Substanzen sind in Wasser schwer löslich. in *Biochemie. Pearson Deutschland GmbH* vol. 4 (2008).
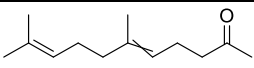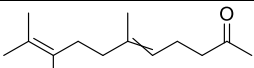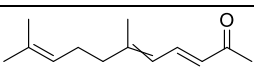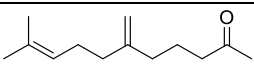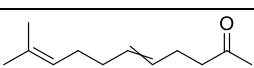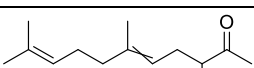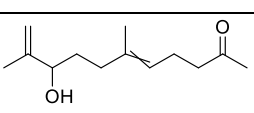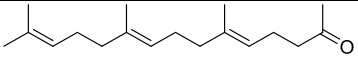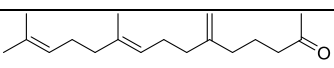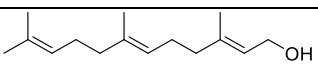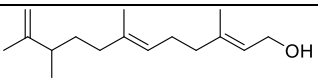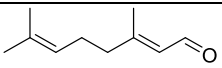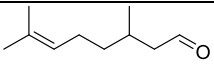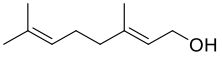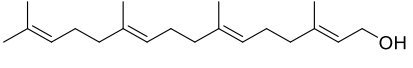
229. Cunningham, F. X., Sun, Z., Chamovitz, D., Hirschberg, J. & Gantt, E. Molecular Structure and Enzymatic Function of Lycopene Cyclase from the Cyanobacterium Synechococcus sp Strain PCC7942. *Plant Cell* **6**, (1994).

230. Velayos, A., Eslava, A. P. & Iturriaga, E. A. A bifunctional enzyme with lycopene cydase and phytoene synthase activities is encoded by the carRP gene of Mucor circinelloides. *Eur J Biochem* **267**, (2000).

231. Schnurr, G., Misawa, N. & Sandmann, G. Expression, purification and properties of lycopene cyclase from Erwinia uredovora. *Biochemical Journal* **315**, (1996).

232. Cunningham, J. & Gantt, E. One ring or two? Determination of ring number in carotenoids by lycopene ε-cyclases. *Proc Natl Acad Sci U S A* **98**, (2001).

233. Rueping, M. & Theissmann, T. Asymmetric Brønsted acid catalysis in aqueous solution. *Chem Sci* **1**, (2010).

234. Christianson, D. W. Roots of biosynthetic diversity. *Science* **316**, (2007).

235. Christianson, D. W. Structural and Chemical Biology of Terpenoid Cyclases. *Chem Rev* **117**, (2017).

236. Lange, B. M. & Srividya, N. Enzymology of monoterpene functionalization in glandular trichomes. *J Exp Bot* **70**, (2019).

237. Parvin, R., Shahrokh, K. O., Mozafar, S., Hassan, E. & Mehrdad, B. Biosynthesis, regulation and properties of plant monoterpenoids. *Journal of Medicinal Plants Research* **8**, (2014).

238. Hernandez-Ortega, A., Vinaixa, M., Zebec, Z., Takano, E. & Scrutton, N. S. A Toolbox for Diverse Oxyfunctionalisation of Monoterpenes. *Sci. Rep.* **8**, (2018).

239. Major, D. T. & Weitman, M. Electrostatically guided dynamics-the root of fidelity in a promiscuous terpene synthase? *J Am Chem Soc* **134**, (2012).

240. Zhang, Q. & Tiefenbacher, K. Terpene cyclization catalysed inside a self-assembled cavity. *Nat Chem* **7**, (2015).

241. Zhang, Q., Catti, L., Pleiss, J. & Tiefenbacher, K. Terpene Cyclizations inside a Supramolecular Catalyst: Leaving-Group-Controlled Product Selectivity and Mechanistic Studies. *J Am Chem Soc* **139**, (2017).

242. Williams, C. M. & Whittaker, D. Rearrangements of pinane derivatives. Part II. Products of acid-catalysed rearrangement of α-pinene and β-pinene in acetic acid. *Journal of the Chemical Society B: Physical Organic* (1971).

243. Williams, C. M. & Whittaker, D. Rearrangements of pinane derivatives. Part I. Products of acid catalysed hydration of α-pinene and β-pinene. *Journal of the Chemical Society B: Physical Organic* (1971).

244. Agus, H. H. Terpene toxicity and oxidative stress. *Toxicology* 33–42 (2021).

245. Chubukov, V. *et al.* Acute Limonene Toxicity in *Escherichia coli* Is Caused by Limonene Hydroperoxide and Alleviated by a Point Mutation in Alkyl Hydroperoxidase AhpC. *Appl Environ Microbiol* **81**, (2015).

246. Baer, P. *et al.* Hedycaryol synthase in complex with nerolidol reveals terpene cyclase mechanism. *ChemBioChem* **15**, 213–216 (2014).

247. Doran, J. C., Baker, G. R., Williams, E. R. & Southwell, I. A. Genetic gains in oil yields after nine years of breeding *Melaleuca alternifolia* (*Myrtaceae*). *Aust J Exp Agric* **46**, (2006).

248. Cornwell, C. P., Leach, D. N. & Wyllie, S. G. The origin of terpinen-4-ol in the steam distillates of melaleuca argentea, *M. dissitiflora* and *M. linariifolia*. *Journal of Essential Oil Research* **11**, (1999).

249. Peter, M. W. D. Preparation of terpinen-4-ol. (1985).

250. Wolf, B., Rack, M., Benson, S., Goetz, R. & Kraus, H. Process for preparing terpinene-4-ol. (2018).

251. Scharf, H. D., Esser, P., Kuhn, W. & Pelzer, R. Method for the photooxidation of terpene olefins. (1995).

252. Leffingwell, J. C. Preparation of terpene alkohol. (1971).

253. Edwards, C. L. & Wilson, S. E. Epoxide isomerization process. (1985).

254. Wolf, B. *et al.* A process for the preparation of terpinolene epoxide. (2016).

255. Jia, M. *et al.* Changing Face: A Key Residue for the Addition of Water by Sclareol Synthase. *ACS Catal* **8**, (2018).

256. Whittington, D. A. *et al.* Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A* **99**, (2002).

257. Brandenberg, O. F., Chen, K. & Arnold, F. H. Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. *J Am Chem Soc* **141**, (2019).

258. McIntosh, J. A. *et al.* Enantioselective intramolecular C-H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. *Angewandte Chemie - International Edition* **52**, (2013).

259. De Jong, R. M. *et al.* Structural basis for the enantioselectivity of an epoxide ring opening reaction catalyzed by halo alcohol dehalogenase HheC. *J Am Chem Soc* **127**, (2005).

260. Balkenhohl, F., Ditrich, K., Hauer, B. & Ladner, W. Optically active amines via lipase-catalyzed methoxyacetylation. *Journal fur Praktische Chemie - Chemiker - Zeitung* **339**, (1997).

261. Voigt, C. A., Kauffman, S. & Wang, Z. G. Rational evolutionary design: the theory of in vitro protein evolution. *Adv Protein Chem* **55**, (2000).

262. Chica, R. A., Doucet, N. & Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: Combining the benefits of directed evolution and rational design. *Curr Opin Biotechnol* **16**, (2005).

263. Schneider, A., Curado, C., Lystbaek, T. B., Osuna, S. & Hauer, B. Tailoring the squalene-hopene cyclase for stereoconvergent and efficient cationic cyclization cascades. *ChemRxiv* (2022).

264. Benítez-Mateos, A. I., Schneider, A., Hegarty, E., Hauer, B. & Paradisi, F. Spheroplasts preparation boosts the catalytic potential of a squalene-hopene cyclase. *Nature Communications 2022 13:1* **13**, (2022).

265. Meng, Y. *et al.* Extension of cell membrane boosting squalene production in the engineered *Escherichia coli*. *Biotechnol Bioeng* **117**, (2020).

266. Li, Z., Kessler, W., Van Den Heuvel, J. & Rinas, U. Simple defined autoinduction medium for high-level recombinant protein production using T7-based *Escherichia coli* expression systems. *Appl Microbiol Biotechnol* **91**, (2011).

267. Kille, S. *et al.* Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth Biol* **2**, (2013).

268. Aberle, B. *et al.* Methylation of Unactivated Alkenes with Engineered Methyltransferases To Generate Non-natural Terpenoids. *Angewandte Chemie - International Edition* **62**, (2023).

269. Waterhouse, A. *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, (2018).

270. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, (2021).

271. Land, H. & Humble, M. S. YASARA: A tool to obtain structural guidance in biocatalytic investigations. in *Methods in Molecular Biology* vol. 1685 (2018).

272. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol* **38**, (2021).

273. Schmid, J. Regioselective hydration of terpenoids using cofactor-independent hydratases. *Thesis* (2019).

Supplementary

Table S 1. Numbered substrates with trivial names used in this work.

No.	Compound	Trivial name
1		Geraniol
2		Dihydropseudoirone
3		Pseudoionone
4		-
5		-
6		-
7		-
8		Farnesylacetone
9		-
10		Farnesol
11		-
12		Citral
13		Citronellal
14		Geraniol
15		Geranylgeraniol
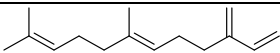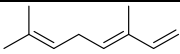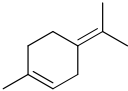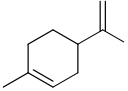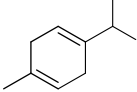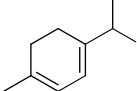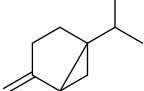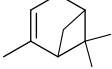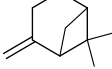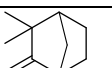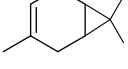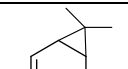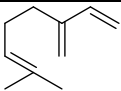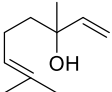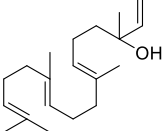
No.	Compound	Trivial name
16		Farnesene
17		Ocimene
18		α-terpinolene
19		Limonene
20		γ-terpinene
21		α-terpinene
22		Sabinene
23		α-pinene
24		β-pinene
25		β-camphene
26		3-carene
27		4-carene
28		Myrcene
29		Alloocimene
30		Linalool
31		Geranyllinalool

Table S 2. List of SHC WT variants from the in-house library, the organism of origin, and the corresponding NCBI identifier.

| Name	Organism	NCBI No.
*Aac*SHC	*Alicyclobacillus acidocaldarius*	WP_014465455.1
*Bam*SHC1	*Burkholderia ambifaria*	WP_011659891.1
*Zmo*SHC1	*Zymomonas mobilis*	P33990.2
*Sfu*SHC	*Syntrophobacter fumaroxidans*	ABK17672.1
*Tel*SHC	*Thermosynechococcus elongatus*	WP_011058142.1
*Rpa*SHC	*Rhodopseudomonas palusTRIS*	WP_011665849.1
*Sco/Sfu*SHC	*Streptomyces coelicolor*	WP_011031166.1
*Bja*SHC	*Bradyrhizobium japonicum*	WP_012041632.1
*Apa*SHC	*Acetobacter pasteurianus*	WP_012812952.1
*Ttu*SHC	*Teredinibacter turnerae*	WP_015819476.1
*Pca*SHC	*Pelobacter carbinolicus*	ABA87615.1
*Bam*SHC2	*Burkholderia ambifaria*	WP_011660979.1
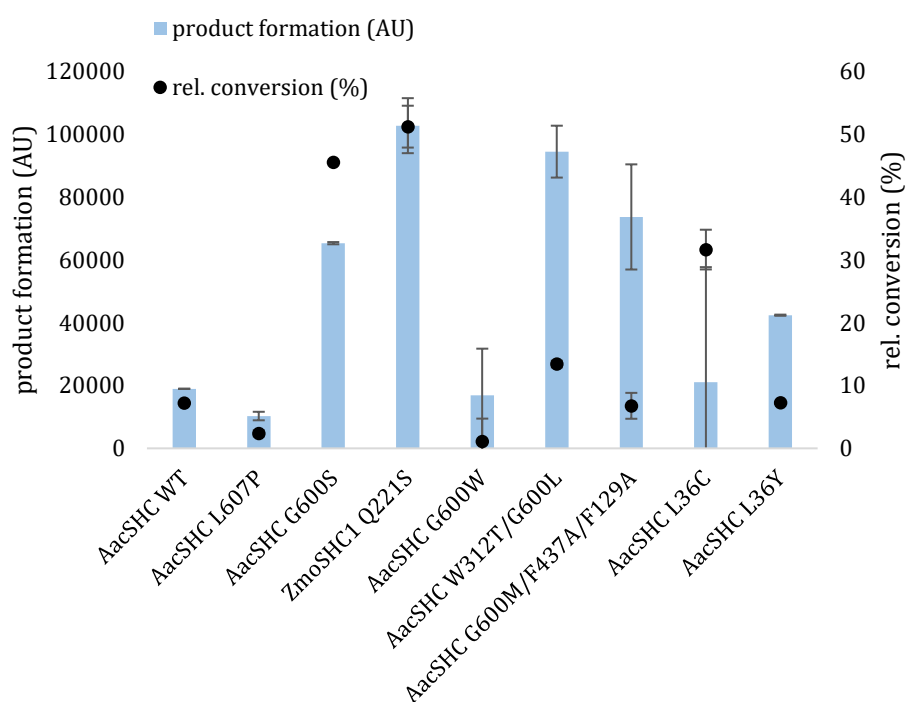
Figure S 1. Product formation and rel. conversion of in-house library hits with **7** in duplicates. Error bars represent standard deviation of replicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed *Aac*SHC and *Zmo*SHC1 variants dissolved in 1 ml *dd*H2O, 2 mM substrate, 20 h, 30 °C, 800 rpm.

Table S 3. Reaction initiation distance, binding energy and dissociation constant for the best docking pose of **1**-**7** and **9** to the active site of WT and VD1. The best docking pose was selected based on the orientation if the substrate to the catalytic center and the binding parameters.

Substrate	Reaction initiation distance (Å)		Binding energy (kcal/mol)		Dissociation constant (mM)	
	WT	VD1	WT	VD1	WT	VD1
1	13.8	9.4	6.82	6.48	9.97	17.91
2	3.9	8.5	7.00	7.41	7.55	3.70
3	14.9	4.8	7.39	7.35	3.86	4.10
4	8.3	10.5	6.71	6.46	12.00	18.33
5	13.7	7.4	6.20	6.55	12.28	15.78
6	12.5	9.7	7.00	6.57	7.38	15.20
7	10.9	4.7	6.93	7.25	8.35	4.84
9	4.1	9.5	7.38	7.40	3.86	3.77

Table S 4. Two sample t-test statistics for the docking distances of **2-7** and **9** in *Aac*SHC WT and VD1 with 95% confidence interval.

Substrate	Average		p-value	t-statistic	d.f.[1]
	WT	VD1			
2	10.460	7.340	0.002	3.243	48
3	11.188	7.608	0.000	4.204	48
4	9.540	7.272	0.027	2.279	48
5	10.552	6.904	0.000	3.812	48
6	11.080	8.212	0.002	3.318	48
7	8.296	5.268	0.003	3.140	48
9	8.844	8.884	0.967	0.041	48

[1] degrees of freedom
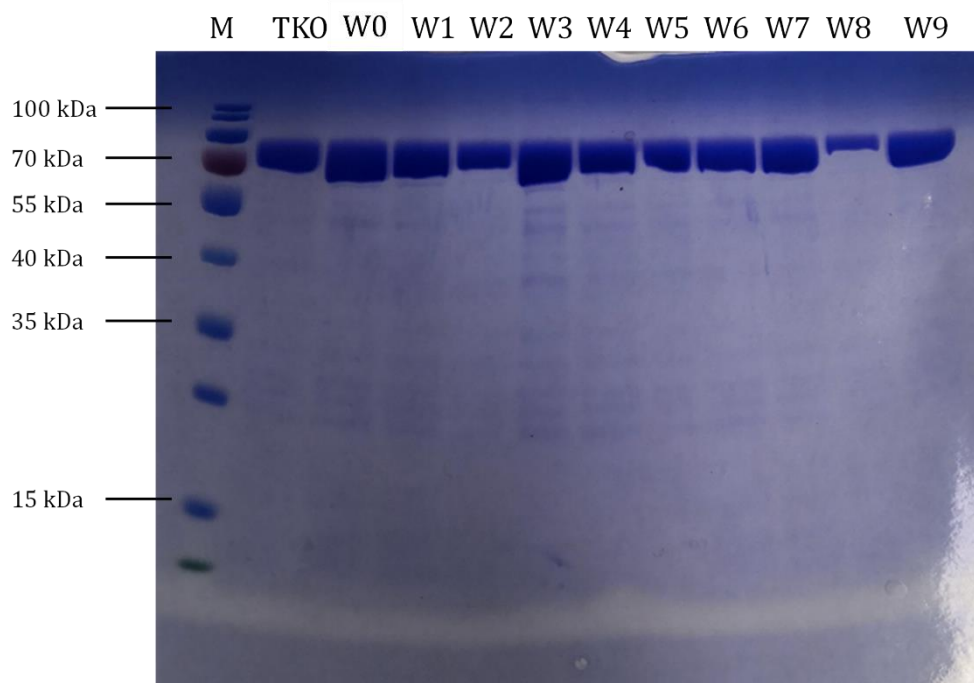
Figure S 2. SDS-PAGE from 10 randomly selected wells of a 96 DWP expressing *Aac*SHC and one sample of expressed TKO. M is the size standard.
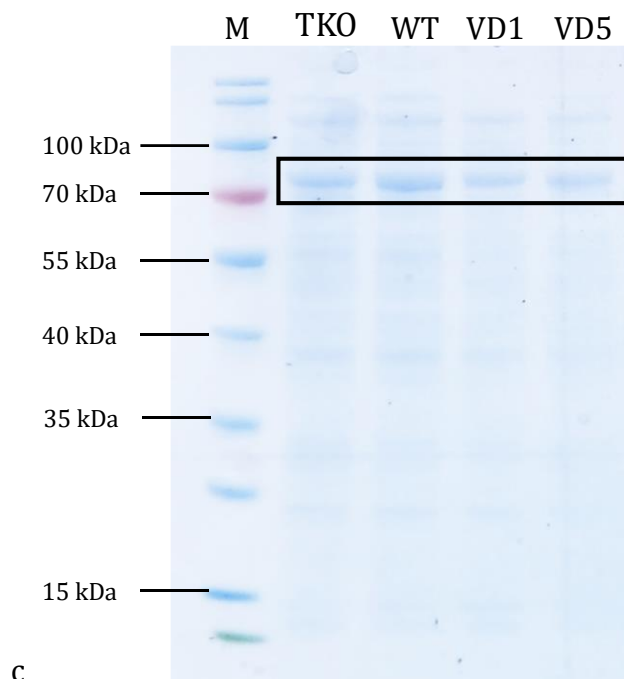
Figure S 3. SDS-PAGE after IEX purification of *Aac*SHC WT, VD1 and VD5. M stands for size standard.

Table S 5. ANOVA results for the measured distances between D376 and the terminal substrate double bond and the TTN for 5 different variants.

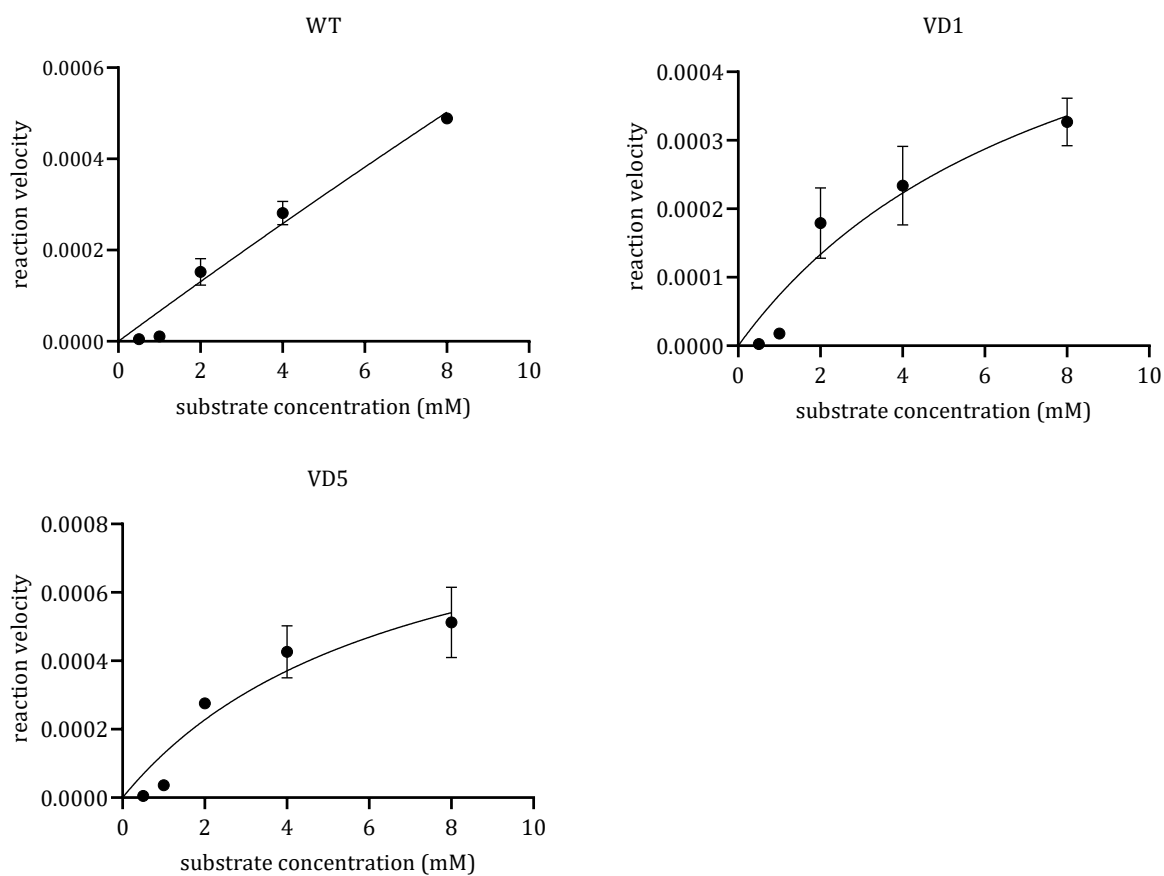| Parameter	p-value	F value	d.f.[1]
Distance	<0.001	4.4	5 (144)
TTN	<0.001	592.3	5 (11)

[1] degrees of freedom

Figure S 4. Michaelis-Menten fit to data from kinetic measurements of *Aac*SHC WT, VD1 and VD5. Reaction conditions: 0.5 ml purified enzyme in IEX elution buffer (25 mM citrate, 200 mM NaCl at pH 6), 0.5-8 mM **2**, 30 °C, 800 rpm.
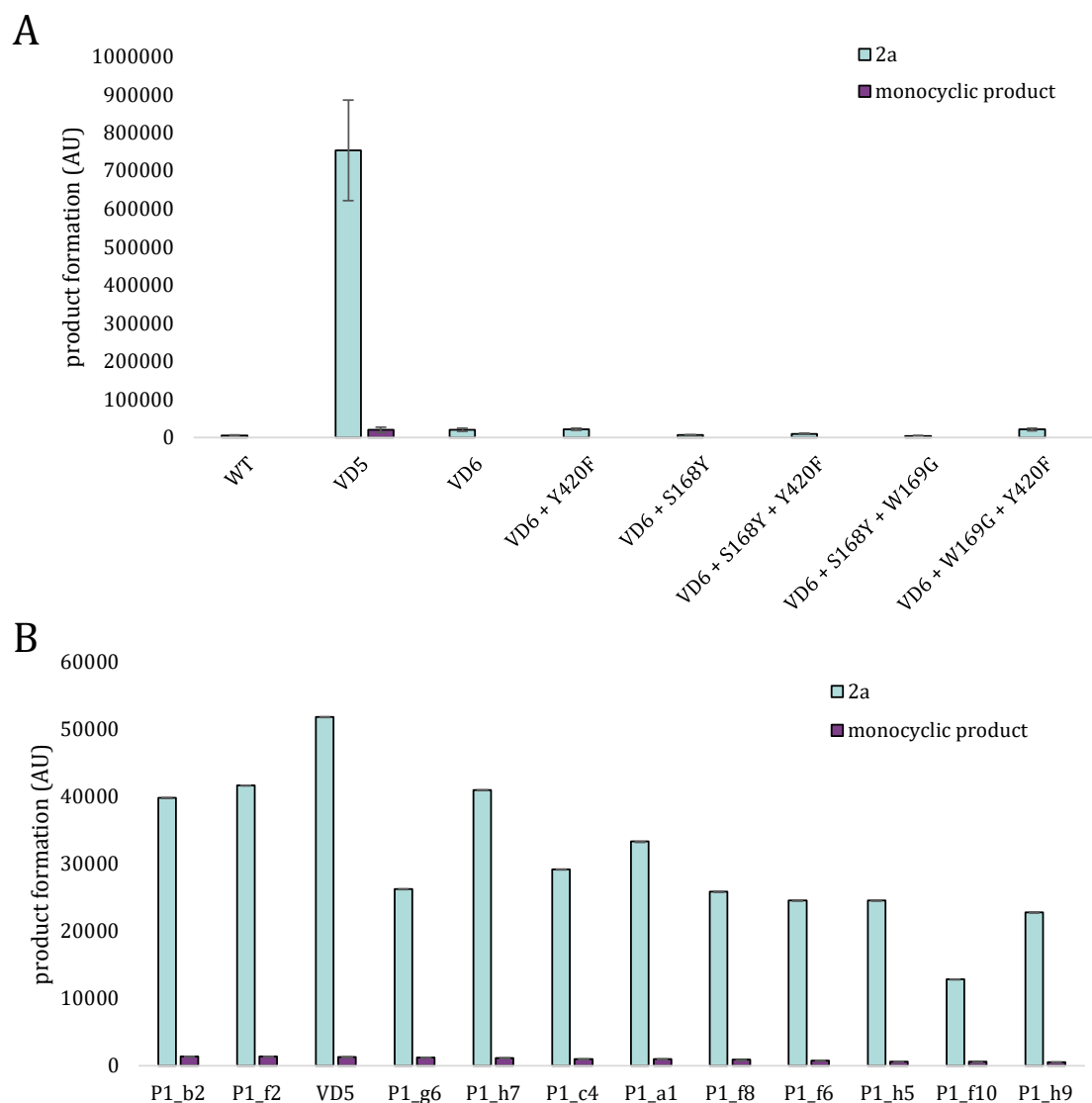
Figure S 5. Rational engineering of *Aac*SHC variants for monocyclization of **2**. Formation of **2a** and monocyclic product in biotransformations with (C) WT, VD5 and VD6 in combination with Y420F, S168Y, and W169G and (D) variants of the G600 saturation mutagenesis shown from highest (left) to lowest (right) formation of monocyclic product. For all other variants no monocyclic product was measured. Error bars represent standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed *Aac*SHC variants dissolved in 1 ml of *dd*H$_2$O, 1 mM substrate, 20 h, 30 °C, 800 rpm.

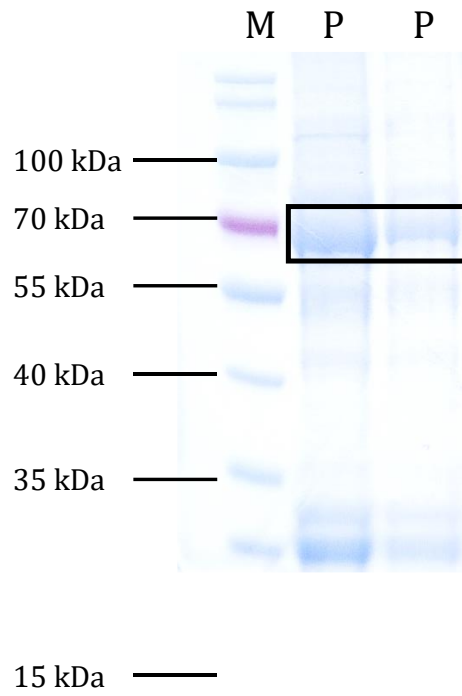**Figure S 6.** SDS-PAGE after purification with affinity chromatography of *Can*LCY-B. M is the size standard.

```
sp|Q38932.2|LycEA/1-524    1 MECV-GARNFAAMAVSTFPSWSC-RRKFPVVKRYSYRNIRFGLCSVRASGGG  50
sp|Q43415.2|LycBC/1-498    1 MDTLLRTPNNLEFLH----GFGVKVSAFSSV-----KSQKFGAKKFCE-G--  40
sp|Q38933.1|LycBA/1-501    1 MDTLLKTPNKLDFFIPQFHGFERLCSNNPYH-----SRVRLGVKKRAI-K--  44

sp|Q38932.2|LycEA/1-524   51 SSGSESCVAVREDFADEEDFVKAGGSEILFV-QMQQNKDMDEQSKLVDKLPP 101
sp|Q43415.2|LycBC/1-498   41 ------------LGSRSVCVKASSSALLELVPETKKENLD------FELPM  73
sp|Q38933.1|LycBA/1-501   45 ------------I---VSSVVSGSAALLDLVPETKKENLD------FELPL  74

sp|Q38932.2|LycEA/1-524  102 IS-IGDGALDLVVIGCGPAGLALAAESAKLGLKVGLIG--PDLPFTNNYGVW 150
sp|Q43415.2|LycBC/1-498   74 YDPSKGVVVDLAVVGGGPAGLAVAQQVSEAGLSVCSIDPNPKLIWPNNYGVW 125
sp|Q38933.1|LycBA/1-501   75 YDTSKSQVVDLAIVGGGPAGLAVAQQVSEAGLSVCSIDPSPKLIWPNNYGVW 126

sp|Q38932.2|LycEA/1-524  151 EDEFNDLGLQKCIEHVWRETIVYLDDDKPITIGRAYGRVSRRLLHEELLRRC 202
sp|Q43415.2|LycBC/1-498  126 VDEFEAMDLLDCLDATWSGATVYIDDNTTKDLNRPYGRVNRKQLKSKMMQKC 177
sp|Q38933.1|LycBA/1-501  127 VDEFEAMDLLDCLDTTWSGAVVYVDEGVKKDLSRPYGRVNRKQLKSKMLQKC 178

sp|Q38932.2|LycEA/1-524  203 VESGVSYLSSKVDSITEASDGLRLVACDDNNVIPCRLATVASGAASGKLLQY 254
sp|Q43415.2|LycBC/1-498  178 ILNGVKFHQAKVIKVIHEESK-SMLICNDGITIQATVVLDATG-FSRSLVQY 227
sp|Q38933.1|LycBA/1-501  179 ITNGVKFHQSKVTNVVHEEAN-STVVCSDGVKIQASVVLDATG-FSRCLVQY 228

sp|Q38932.2|LycEA/1-524  255 EVGGPRVCVQTAYGVEVEVENSPYDPDQMVFMDYRDYTNEK---VRSLEAEY 303
sp|Q43415.2|LycBC/1-498  228 DKP-YNPGYQVAYGILAEVEEHPFDVNKMVFMDWRDSHLKNNVELKERNSRI 278
sp|Q38933.1|LycBA/1-501  229 DKP-YNPGYQVAYGIVAEVDGHPFDVDKMVFMDWRDKHLDSYPELKERNSKI 279

sp|Q38932.2|LycEA/1-524  304 PTFLYAMPMTKSRLFFEETCLASKDVMPFDLLKTKLMLRLDTLGIRILKTYE 355
sp|Q43415.2|LycBC/1-498  279 PTFLYAMPFSSNRIFLEETSLVARPGLGMDDIQERMVARLSHLGIKVKSIEE 330
sp|Q38933.1|LycBA/1-501  280 PTFLYAMPFSSNRIFLEETSLVARPGLRMEDIQERMAARLKHLGINVKRIEE 331

sp|Q38932.2|LycEA/1-524  356 EEWSYIPVGGSLPNTEQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIA 407
sp|Q43415.2|LycBC/1-498  331 DEHCVIPMGGPLPVLPQRVVGIGGTAGMVHPSTGYMVARTLAAAPVVANAII 382
sp|Q38933.1|LycBA/1-501  332 DERCVIPMGGPLPVLPQRVVGIGGTAGMVHPSTGYMVARTLAAAPIVANAIV 383

sp|Q38932.2|LycEA/1-524  408 EILREETT-KQINSNISRQAWDTLWPPERKRQRAFFLFGLALIVQFDTEGIR 458
sp|Q43415.2|LycBC/1-498  383 QYLSSER--SHSGDELSAAVWKDLWPIERRRQREFFCFGMDILLKLDLPATR 432
sp|Q38933.1|LycBA/1-501  384 RYLGSPSSNSLRGDQLSAEVWRDLWPIERRRQREFFCFGMDILLKLDLDATR 435

sp|Q38932.2|LycEA/1-524  459 SFFRTFFRLPKWMWQGFLGSTLTSGDLVLFALYMFVISPNNLRKGLI----- 505
sp|Q43415.2|LycBC/1-498  433 RFFDAFFDLEPRYWHGFLSSRLFLPELIVFGLSLFSHASNTSRLEIMTKGTL 484
sp|Q38933.1|LycBA/1-501  436 RFFDAFFDLQPHYWHGFLSSRLFLPELLVFGLSLFSHASNTSRLEIMTKGTV 487

sp|Q38932.2|LycEA/1-524  506 ------NHLISDPTGATMIKTYLKV                             524
sp|Q43415.2|LycBC/1-498  485 PLVHMINNLLQDKE-----------                             498
sp|Q38933.1|LycBA/1-501  488 PLAKMINNLVQDRD-----------                             501
```

Figure S 7. MSA of *Ath*LCA-E (LycEA), *Can*LCY-B (LycBC) and *Ath*LCY-B (LycBA) by ClustalW.
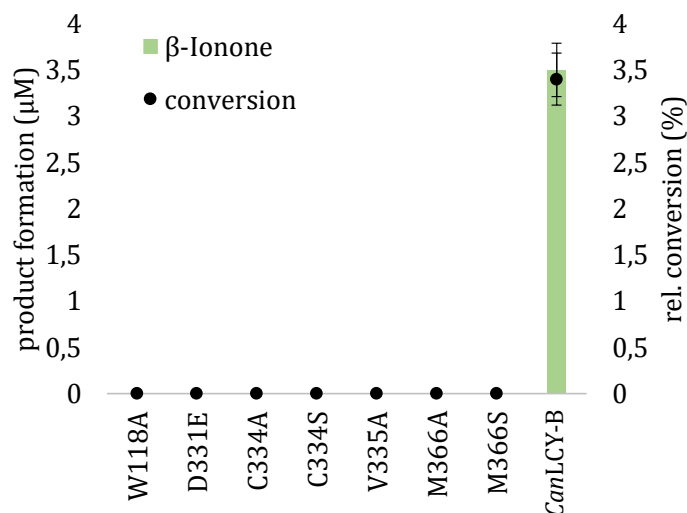
Figure S 8. Product formation and rel. conversion of *Can*LCY-B alanine and *Ath*LCY-E amino acid substitutions variants from **3**. Positions for substitution were selected based on an MSA. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed LCY variants dissolved in 1 ml TRIS-Maleate buffer (pH 6.8), 0.5 mM NADPH, 0.01mM FAD, 1 mM **3**, 20 h, 800 rpm.
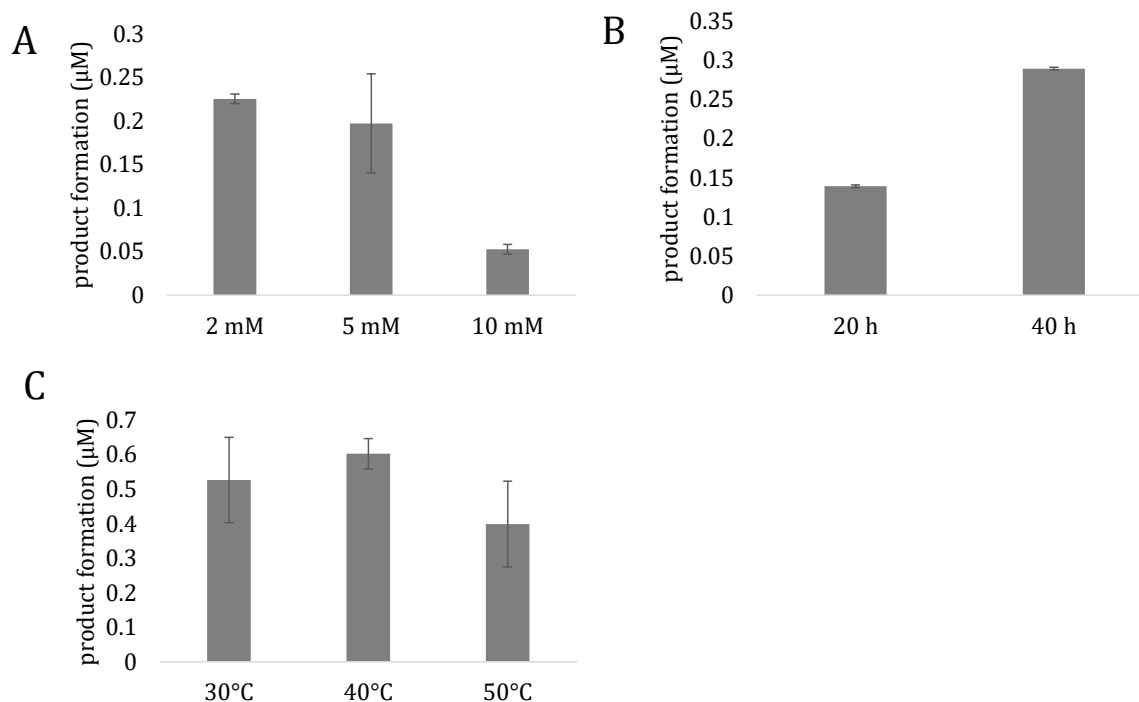

Figure S 9. Product formation and rel. conversion of *Aac*SHC with **19** at change of (A) substrate concentration, (B) biotransformation time and (C) biotransformation temperature. Error bars represent the standard deviation of duplicates. Reaction conditions: 0.2 mg$_{CWW}$ *E. coli* whole cells with expressed *Aac*SHC dissolved in 1 ml buffer or *dd*H$_2$O and 3-5 mM substrate.
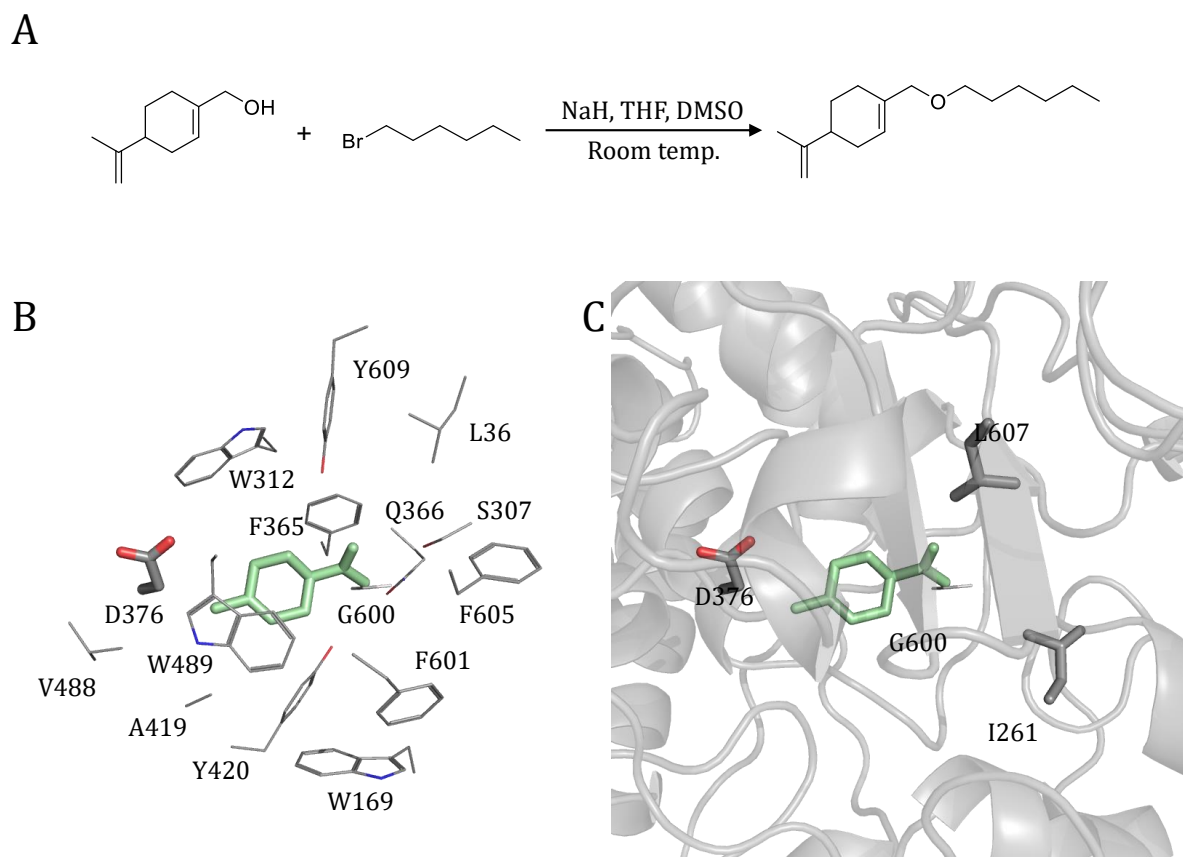
A

B

C

Figure S 10. Approaches to convert **19** by SHCs. In (A) schematic synthesis of a limonene ether for substrate engineering, (B) docking studies with MD simulated *Aac*SHC WT and (C) combinatorial approach at positions I261, G600 and L607 in *Aac*SHC WT.

Table S 6. Oligonucleotides for saturation mutagenesis and site-directed mutagenesis with *Aac*SHC for substrates with modified isoprene pattern.

Variant		Sequence (5'- to 3'-direction)
R488X	fwd 1	CGGATGGCAGCTGGTTTGGCNDTTGGGGCGTGA
	fwd 2	CGGATGGCAGCTGGTTTGGCVHGTGGGGCGTGA
	fwd 3	CGGATGGCAGCTGGTTTGGCTGGTGGGGCGTGA
	rev	CGCGAACAGAAACCGGATGGCAGCTGGTTTGGCCGC
G600M/T599X	fwd 1	ATGAACCGTATTATACCGGCNDTATGTTCCCGGGCGA
	fwd 2	ATGAACCGTATTATACCGGCVHGATGTTCCCGGGCGA
	fwd 3	ATGAACCGTATTATACCGGCTGGATGTTCCCGGGCGA
	rev	GCCGGTATAATACGGTTCATCCCAGCCGCCATCCGGGCGCTGGGTTTC
G600M/Y606X	fwd 1	CCATGTTCCCGGGCGATTTTNDTCTGGGCTATACCATGTATCG
	fwd 2	CCATGTTCCCGGGCGATTTTVHGCTGGGCTATACCATGTATCG
	fwd 3	CCATGTTCCCGGGCGATTTTTGGCTGGGCTATACCATGTATCG
	rev	AAAATCGCCCGGGAACATGGTGCCGGTATAATACGGTTCAT
G600M/F605X	fwd 1	GCACCATGTTCCCGGGCGATNDTTATCTGGGCTATACCATGTATCG
	fwd 2	GCACCATGTTCCCGGGCGATVHGTATCTGGGCTATACCATGTATCG
	fwd 3	GCACCATGTTCCCGGGCGATTGGTATCTGGGCTATACCATGTATCG
	rev	ATCGCCCGGGAACATGGTGCCGGTATAATACGGTTCATCCCAG

Variant		Sequence (5'- to 3'-direction)
G600M/P602X	fwd 1	ATTATACCGGCACCATGTTCNDTGGCGATTTTTATCTGG
	fwd 2	ATTATACCGGCACCATGTTCVHGGGCGATTTTTATCTGG
	fwd 3	ATTATACCGGCACCATGTTCTGGGGCGATTTTTATCTGG
	rev	GAACATGGTGCCGGTATAATACGGTTCATCCCAGC
G600M/F605C/Y609X	fwd 1	CGGGCGATTGTTATCTGGGCNDTACCATGTATCGCCATGTG
	fwd 2	CGGGCGATTGTTATCTGGGCVHGACCATGTATCGCCATGTG
	fwd 3	CGGGCGATTGTTATCTGGGCTGGACCATGTATCGCCATGTG
	rev	GCCCAGATAACAATCGCCCGGGAACATGGTGCCGGTA
G600M/F605C/L607X	fwd 1	TGTTCCCGGGCGATTGTTATNDTGGCTATACCATGTATCG
	fwd 2	TGTTCCCGGGCGATTGTTATVHGGGCTATACCATGTATCG
	fwd 3	TGTTCCCGGGCGATTGTTATTGGGGCTATACCATGTATCG
	rev	ATAACAATCGCCCGGGAACATGGTGCCGGTATAATACGGTTC
R488P/Y495X	fwd 1	CGTGGGGCGTGAACTATCTGNDTGGCACCGGCGCGGTG
	fwd 2	CGTGGGGCGTGAACTATCTGVHGGGCACCGGCGCGGTG
	fwd 3	CGTGGGGCGTGAACTATCTGTGGGGCACCGGCGCGGTG
	rev	CAGATAGTTCACGCCCCACGGGCCAAACCAGCTGC
G600M/F605C/Y612X	fwd 1	GTTATCTGGGCTATACCATGNDTCGCCATGTGTTTCCGAC
	fwd 2	GTTATCTGGGCTATACCATGVHGCGCCATGTGTTTCCGAC
	fwd 3	GTTATCTGGGCTATACCATGTGGCGCCATGTGTTTCCGAC
	rev	CATGGTATAGCCCAGATAACAATCGCCCGGGAACA
G600M/D377X	fwd 1	TGTATTATCCGGATGTGGATNDTACCGCGGTggtG
	fwd 2	TGTATTATCCGGATGTGGATVHGACCGCGGTggtG
	fwd 3	TGTATTATCCGGATGTGGATTGGACCGCGGTggtG
	rev	ATCCACATCCGGATAATACACGTTATCAAACTGAAACGCAAAG
G600M/D374X	fwd 1	TTGATAACGTGTATTATCCGNDTGTGGATGATACCGCGGTg
	fwd 2	TTGATAACGTGTATTATCCGVHGGTGGATGATACCGCGGTg
	fwd 3	TTGATAACGTGTATTATCCGTGGGTGGATGATACCGCGGTg
	rev	CGGATAATACACGTTATCAAACTGAAACGCAAAGC
R488P/W489X	fwd 1	ATGGCAGCTGGTTTGGCCCGNDTGGCGTGAACTATCTGTA
	fwd 2	ATGGCAGCTGGTTTGGCCCGVHGGGCGTGAACTATCTGTA
	fwd 3	ATGGCAGCTGGTTTGGCCCGTGGGGCGTGAACTATCTGTA
	rev	CGGGCCAAACCAGCTGCCATCCGGTTTCTGTTCGCGTTTC
S445X	fwd 1	GAAGTGACCGATCCGCCGNDTGAAGATGTGACCGC
	fwd 2	GAAGTGACCGATCCGCCGVHGGAAGATGTGACCGC
	fwd 3	GAAGTGACCGATCCGCCGTGGGAAGATGTGACCGC
	rev	CGGCGGATCGGTCACTTCGCCAAAATCGCAAAACGGAA
F605C/G600X	fwd 1	AGATAACAATCGCCCGGGAANDTGGTGCCGGTATAATAC
	fwd 2	AGATAACAATCGCCCGGGAAVHGGGTGCCGGTATAATAC
	fwd 3	AGATAACAATCGCCCGGGAATGGGGTGCCGGTATAATAC
	rev	TTCCCGGGCGATTGTTATCTGGGCTATACCATGTATCGC
G600M/Y609F/F605X	fwd 1	GCACCATGTTCCCGGGCGATNDTTATCTGGGCTTTACCATGTATCG
	fwd 2	GCACCATGTTCCCGGGCGATVHGTATCTGGGCTTTACCATGTATCG

L

Variant		Sequence (5'- to 3'-direction)
	fwd 3	GCACCATGTTCCCGGGCGATTGGTATCTGGGCTTTACCATGTATCG
	rev	ATCGCCCGGGAACATGGTGCCGGTATAATACGGTTCATCCCAG
G600M/Y609F/L605X	fwd 1	TGTTCCCGGGCGATTTTTATNDTGGCTTTACCATGTATCG
	fwd 2	TGTTCCCGGGCGATTTTTATVHGGGCTTTACCATGTATCG
	fwd 3	TGTTCCCGGGCGATTTTTATTGGGGCTTTACCATGTATCG
	rev	ATAAAAATCGCCCGGGAACATGGTGCCGGTATAATACGGTTC
R488P	fwd	CTGGTTTGGCCCGTGGGGCGTGAAC
	rev	GTTCACGCCCCACGGGCCAAACCAG
F605C	fwd	ATACCGGCACCGGGTTCCCGGGCGATTTTTATCTGGGCT
	rev	ATCGCCCGGGAACCCGGTGCCGGTATAATACGGTTCATCC
Y495D	fwd	GCTGGGGCGTGAACTATCTGGATGGCACCGGCGCGGTG
	rev	CAGATAGTTCACGCCCCAGCGGCCAAACCAGCTGC
S445V	fwd	CGAAGTGACCGATCCGCCGGTCGAAGATGTGACC
	rev	GCACATGCGCGGTCACATCTTCGACCGGCGGATCG
S168Y	fwd	CATTTATGAATTTGGCTATTGGGCGCGCGCGACCGTG
	rev	CGGTCGCGCGCGCCCAATAGCCAAATTCATAAATG
W169G/S168Y	fwd	CATTTATGAATTTGGCTATGGCGCGCGCGCGACCGTG
	rev	CGGTCGCGCGCGCGCCATAGCCAAATTCATAAATG
G600M/Y609L	fwd	AGATAAAAATCGCCCGGGAACATGGTGCCGGTATAATAC
	rev	TTCCCGGGCGATTTTTATCTGGGCCTGACCATGTATCGC
G600T/F607A	fwd	ACCTTCCCGGGCGATTGTTATGCGGGCTATACCATG
	rev	GATACATGGTATAGCCCGCATAACAATCGCCCGGGAA
G600T/F605C	fwd	CGTATTATACCGGCACCACCTTCCCGGGCGATTGTTATC
	rev	GATAACAATCGCCCGGGAAGGTGGTGCCGGTATAATACG
G600T/F605C/F607A	fwd	ACCTTCCCGGGCGATTTTTATGCGGGCTATACCATG
	rev	GATACATGGTATAGCCCGCATAAAAATCGCCCGGGAA
F365X	fwd 1	GAAACCGGGCGGCTTTGCGNDTCAGTTTGATAACGTGTATTATCCGG
	fwd 2	GAAACCGGGCGGCTTTGCGVHGCAGTTTGATAACGTGTATTATCCGG
	fwd 3	GAAACCGGGCGGCTTTGCGTGGCAGTTTGATAACGTGTATTATCCGG
	rev	CGCAAAGCCGCCCGGTTTCAGGTTCGGGCGTTTCAC
F365L/Q366X	fwd 1	CCGGGCGGCTTTGCGCTGNDTTTTGATAACGTGTATTATCCGGATGTG
	fwd 2	CCGGGCGGCTTTGCGCTGVHGTTTGATAACGTGTATTATCCGGATGTG
	fwd 3	CCGGGCGGCTTTGCGCTGTGGTTTGATAACGTGTATTATCCGGATGTG
	rev	CAGCGCAAAGCCGCCCGGTTTCAGGTTCGGGCGTTTC
G600M/F605L/Y609X	fwd 1	CGGGCGATCTGTATCTGGGCNDTACCATGTATCGCCATGTG
	fwd 2	CGGGCGATCTGTATCTGGGCVHGACCATGTATCGCCATGTG
	fwd 3	CGGGCGATCTGTATCTGGGCTGGACCATGTATCGCCATGTG
	rev	GCCCAGATACAGATCGCCCGGGAACATGGTGCCGGTA
G600M/Y609F/F605X /L607X	fwd 1	TGTTCCCGGGCGATCTGTATNDTGGCTTTACCATGTATCG
	fwd 2	TGTTCCCGGGCGATCTGTATVHGGGCTTTACCATGTATCG
	fwd 3	TGTTCCCGGGCGATCTGTATTGGGGCTTTACCATGTATCG
	rev	ATACAGATCGCCCGGGAACATGGTGCCGGTATAATACGGTTC

Variant		Sequence (5'- to 3'-direction)
G600M/F605L/Y612X	fwd 1	TGTATCTGGGCTTTACCATGNDTCGCCATGTGTTTCCGAC
	fwd 2	TGTATCTGGGCTTTACCATGVHGCGCCATGTGTTTCCGAC
	fwd 3	TGTATCTGGGCTTTACCATGTGGCGCCATGTGTTTCCGAC
	rev	CATGGTAAAGCCCAGATACAGATCGCCCGGGAACA
S307X	fwd 1	GTATCCCACACCGGGCTAATNDTCGCCTGAAACATCC
	fwd 2	GTATCCCACACCGGGCTAATVHGCGCCTGAAACATCC
	fwd 3	GTATCCCACACCGGGCTAATTGGCGCCTGAAACATCC
	rev	ATTAGCCCGGTGTGGGATACCGGCCTgGCGGTGC
G600M/F605L/601X	fwd 1	CGTATTATACCGGCACCATGNDTCCGGGCGATCTGTA
	fwd 2	CGTATTATACCGGCACCATGVHGCCGGGCGATCTGTA
	fwd 3	CGTATTATACCGGCACCATGTGGCCGGGCGATCTGTA
	rev	CATGGTGCCGGTATAATACGGTTCATCCCAGCCGCCA
A306X	fwd 1	ATGGCGGCTGGATGTTTCAGNDTAGCATTAGCCCGGT
	fwd 2	ATGGCGGCTGGATGTTTCAGVHGAGCATTAGCCCGGT
	fwd 3	ATGGCGGCTGGATGTTTCAGTGGAGCATTAGCCCGGT
	rev	CTGAAACATCCAGCCGCCATAATCCAGTTCCACGCCAT
F434X	fwd 1	CGGAATATGGTTCGGCAGATCGCTGGTGTTATCCACATC
	fwd 2	ATCTGCCGAACCATATTCCGNDTTGCGATTTTGGCGAAG
	fwd 3	ATCTGCCGAACCATATTCCGVHGTGCGATTTTGGCGAAG
	rev	ATCTGCCGAACCATATTCCGTGGTGCGATTTTGGCGAAG

Table S 7. Oligonucleotides for saturation mutagenesis and site-directed mutagenesis with *Can*LCY-B and *Ath*LCY-E for conversion of **3**.

Variant		Sequence (5'- to 3'-direction)
M366A	fwd	TGCATCCCTCTACCGGCTACGCGGTGGCACGGACAC
	rev	GTAGCCGGTAGAGGGATGCACCATACCTGCAGTACCG
M366S	fwd	TGCATCCCTCTACCGGCTACAGCGTGGCACGGACAC
	rev	GTAGCCGGTAGAGGGATGCACCATACCTGCAGTACC
F220A	fwd	CTTAGACGCCACGGGAGCGTCTCGTTCATTGGTA
	rev	TACCAATGAACGAGACGCTCCCGTGGCGTCTAAG
W118A	fwd	ATCCGAATCCCAAATTGATTGCGCCAAATAATTACGGTGTA
	rev	AATCAATTTGGGATTCGGATCGATGCTGCATACGCT
W118F	fwd	ATCCGAATCCCAAATTGATTTTCCCAAATAATTACGGTGTA
	rev	AATCAATTTGGGATTCGGATCGATGCTGCATACGCT
C334A	fwd	GCATCGAGGAAGACGAGCACGCGGTCATTCCTATGGGC
	rev	GTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAGGTGTG
C334S	fwd	GCATCGAGGAAGACGAGCACAGCGTCATTCCTATGGGC
	rev	GTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAGGTG
V335A	fwd	TCGAGGAAGACGAGCACTGCGCGATTCCTATGGGCGGGCCGTTA
	rev	GCAGTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAG
V335Y	fwd	TCGAGGAAGACGAGCACTGCTATATTCCTATGGGCGGGCCGTTA

Variant		Sequence (5'- to 3'-direction)
	rev	GCAGTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAGGT
D331E	fwd	AGGTGAAAAGCATCGAGGAAGAAGAGCACTGCGTCA
	rev	TTCCTCGATGCTTTTCACCTTTATACCCAGGTGTGAAAGACG
D331A	fwd	AGGTGAAAAGCATCGAGGAAGCGGAGCACTGCGTCA
	rev	TTCCTCGATGCTTTTCACCTTTATACCCAGGTGTGAAAGACG
M366X	fwd 1	TGCATCCCTCTACCGGCTACNDTGTGGCACGGACAC
	fwd 2	TGCATCCCTCTACCGGCTACVHGGTGGCACGGACAC
	fwd 3	TGCATCCCTCTACCGGCTACTGGGTGGCACGGACAC
	rev	GTAGCCGGTAGAGGGATGCACCATACCTGCAGTACCG
F220X	fwd 1	CTTAGACGCCACGGGANDTTCTCGTTCATTGGTA
	fwd 2	CTTAGACGCCACGGGAVHGTCTCGTTCATTGGTA
	fwd 3	CTTAGACGCCACGGGATGGTCTCGTTCATTGGTA
	rev	TACCAATGAACGAGACGCTCCCGTGGCGTCTAAG
W118X	fwd 1	ATCCGAATCCCAAATTGATTNDTCCAAATAATTACGGTGTA
	fwd 2	ATCCGAATCCCAAATTGATTVHGCCAAATAATTACGGTGTA
	fwd 3	ATCCGAATCCCAAATTGATTTGGCCAAATAATTACGGTGTA
	rev	AATCAATTTGGGATTCGGATCGATGCTGCATACGCT
C334X	fwd 1	GCATCGAGGAAGACGAGCACNDTGTCATTCCTATGGGC
	fwd 2	GCATCGAGGAAGACGAGCACVHGGTCATTCCTATGGGC
	fwd 3	GCATCGAGGAAGACGAGCACTGGGTCATTCCTATGGGC
	rev	GTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAGGTGTG
V335X	fwd 1	TCGAGGAAGACGAGCACTGCNDTATTCCTATGGGCGGGCCGTTA
	fwd 2	TCGAGGAAGACGAGCACTGCVHGATTCCTATGGGCGGGCCGTTA
	fwd 3	TCGAGGAAGACGAGCACTGCTGGATTCCTATGGGCGGGCCGTTA
	rev	GCAGTGCTCGTCTTCCTCGATGCTTTTCACCTTTATACCCAG
D331X	fwd 1	AGGTGAAAAGCATCGAGGAANDTGAGCACTGCGTCA
	fwd 2	AGGTGAAAAGCATCGAGGAAVHGGAGCACTGCGTCA
	fwd 3	AGGTGAAAAGCATCGAGGAATGGGAGCACTGCGTCA
	rev	TTCCTCGATGCTTTTCACCTTTATACCCAGGTGTGAAAGACG
Q85X	fwd 1	TTTTGTTCGTCCAAATGCAANDTAACAAGGACATGGACG
	fwd 2	TTTTGTTCGTCCAAATGCAAVHGAACAAGGACATGGACG
	fwd 3	TTTTGTTCGTCCAAATGCAATGGAACAAGGACATGGACG
	rev	TTGCATTTGGACGAACAAAATCTCGCTCCCGCC
A359X	fwd 1	AAACATATGAAGAGGAATGGNDTTACATCCCAGTAGGTGGA
	fwd 2	AAACATATGAAGAGGAATGGVHGTACATCCCAGTAGGTGGA
	fwd 3	AAACATATGAAGAGGAATGGTGGTACATCCCAGTAGGTGGA
	rev	CCATTCCTCTTCATATGTTTTCAGGATGCGAATACCCAAG
A360X	fwd 1	CATATGAAGAGGAATGGTCANDTATCCCAGTAGGTGGAAG
	fwd 2	CATATGAAGAGGAATGGTCAVHGATCCCAGTAGGTGGAAG
	fwd 3	CATATGAAGAGGAATGGTCATGGATCCCAGTAGGTGGAAG
	rev	TGACCATTCCTCTTCATATGTTTTCAGGATGCGAATACCCAA
A387X	fwd 1	CTGCATCCATGGTCCACCCANDTACCGGGTATAGTG

Variant		Sequence (5'- to 3'-direction)
	fwd 2	CTGCATCCATGGTCCACCCAVHGACCGGGTATAGTG
	fwd 3	CTGCATCCATGGTCCACCCATGGACCGGGTATAGTG
	rev	TGGGTGGACCATGGATGCAGCTGCTCCGAAAGCCAG
A391X	fwd 1	TCCACCCAGCCACCGGGTATNDTGTGGTCCGATCGCTTA
	fwd 2	TCCACCCAGCCACCGGGTATVHGGTGGTCCGATCGCTTA
	fwd 3	TCCACCCAGCCACCGGGTATTGGGTGGTCCGATCGCTTA
	rev	ATACCCGGTGGCTGGGTGGACCATGGATGCAGCTGC

Table S 8. Oligonucleotides for saturation mutagenesis and site-directed mutagenesis with *Aac*SHC for monoterpenes as substrates.

Variant		Sequence (5'- to 3'-direction)
G600L/L607X	fwd 1	TCTTCCCGGGCGATTTTTATNDTGGCTATACCATGTATCG
	fwd 2	TCTTCCCGGGCGATTTTTATVHGGGCTATACCATGTATCG
	fwd 3	TCTTCCCGGGCGATTTTTATTGGGGCTATACCATGTATCG
	rev	ATAAAAATCGCCCGGGAAGAGGGTGCCGGTATAATACG
G600L/Y609X	fwd 1	GGCGATTTTTATCTGGGCNDTACCATGTATCGCCATGTGTTTC
	fwd 2	GGCGATTTTTATCTGGGCVHGACCATGTATCGCCATGTGTTTC
	fwd 3	GGCGATTTTTATCTGGGCTGGACCATGTATCGCCATGTGTTTC
	rev	GCCCAGATAAAAATCGCCCGGGAAGAGGGTGCC
Y420W/A419X	fwd 1	GCAGCAACGGCGGCTGGGGCNDTTGGGATGTGGATAACACC
	fwd 2	GCAGCAACGGCGGCTGGGGCVHGTGGGATGTGGATAACACC
	fwd 3	GCAGCAACGGCGGCTGGGGCTGGTGGGATGTGGATAACACC
	rev	GCCCCAGCCGCCGTTGCTGCTCTGCATGC
Q366X	fwd 1	CCGGGCGGCTTTGCGTTTNDTTTTGATAACGTGTATTATCCGGATGTG
	fwd 2	CCGGGCGGCTTTGCGTTTVHGTTTGATAACGTGTATTATCCGGATGTG
	fwd 3	CCGGGCGGCTTTGCGTTTTGGTTTGATAACGTGTATTATCCGGATGTG
	rev	AAACGCAAAGCCGCCCGGTTTCAGGTTCGGGCGTTTC
V448X	fwd 1	GATCCGCCGAGCGAAGATNDTACCGCGCATGTGCTGG
	fwd 2	GATCCGCCGAGCGAAGATVHGACCGCGCATGTGCTGG
	fwd 3	GATCCGCCGAGCGAAGATTGGACCGCGCATGTGCTGG
	rev	ATCTTCGCTCGGCGGATCGGTCACTTCGCCAAAATCGCAAAAC
Q366V/F365X	fwd 1	GAAACCGGGCGGCTTTGCGNDTGTTTTTGATAACGTGTATTATCCGG
	fwd 2	GAAACCGGGCGGCTTTGCGVHGGTTTTTGATAACGTGTATTATCCGG
	fwd 3	GAAACCGGGCGGCTTTGCGTGGGTTTTTGATAACGTGTATTATCCGG
	rev	CGCAAAGCCGCCCGGTTTCAGGTTCGGGCGTTTCAC
G600L/T599X	fwd 1	ATGAACCGTATTATACCGGCNDTCTCTTCCCGGGCGA
	fwd 2	ATGAACCGTATTATACCGGCVHGCTCTTCCCGGGCGA
	fwd 3	ATGAACCGTATTATACCGGCTGGCTCTTCCCGGGCGA
	rev	GCCGGTATAATACGGTTCATCCCAGCCGCCATCCGGGCGCTGGGTTTC
T599M/F605X	fwd 1	GCATGGGGTTCCCGGGCGATNDTTATCTGGGCTATACCATGTATCG
	fwd 2	GCATGGGGTTCCCGGGCGATVHGTATCTGGGCTATACCATGTATCG

Variant		Sequence (5'- to 3'-direction)
	fwd 3	GCATGGGGGTTCCCGGGCGATTGGTATCTGGGCTATACCATGTATCG
	rev	ATCGCCCGGGAACCCCATGCCGGTATAATACGGTTCATCCCAG
T599M/L607X	fwd 1	GGTTCCCGGGCGATTTTTATNDTGGCTATACCATGTATCGC
	fwd 2	GGTTCCCGGGCGATTTTTATVHGGGCTATACCATGTATCGC
	fwd 3	GGTTCCCGGGCGATTTTTATTGGGGCTATACCATGTATCGC
	rev	ATAAAAATCGCCCGGGAACCCCATGCCGGTATAATACGG
T599M/G600X	fwd 1	AGATAAAAATCGCCCGGGAANDTCATGCCGGTATAATAC
	fwd 2	AGATAAAAATCGCCCGGGAAVHGCATGCCGGTATAATAC
	fwd 3	AGATAAAAATCGCCCGGGAATGGCATGCCGGTATAATAC
	rev	TTCCCGGGCGATTTTTATCTGGGCTATACCATGTATCGC
G600I/T599X	fwd 1	ATGAACCGTATTATACCGGCNDTATTTTCCCGGGCGA
	fwd 2	ATGAACCGTATTATACCGGCVHGATTTTCCCGGGCGA
	fwd 3	ATGAACCGTATTATACCGGCTGGATTTTCCCGGGCGA
	rev	GCCGGTATAATACGGTTCATCCCAGCCGCCATCCGGGCGCTGGGTTTC
G600I/F605X	fwd 1	GCACCATTTTCCCGGGCGATNDTATCTGGGCTATACCATGTATCG
	fwd 2	GCACCATTTTCCCGGGCGATVHGTATCTGGGCTATACCATGTATCG
	fwd 3	GCACCATTTTCCCGGGCGATTGGTATCTGGGCTATACCATGTATCG
	rev	ATCGCCCGGGAAAATGGTGCCGGTATAATACGGTTCATCCCAG
Y420F/A419X	fwd 1	CAGCAACGGCGGCTGGGGCNDTTTTGATGTGGATAACACCAG
	fwd 2	CAGCAACGGCGGCTGGGGCVHGTTTGATGTGGATAACACCAG
	fwd 3	CAGCAACGGCGGCTGGGGCTGGTTTGATGTGGATAACACCAG
	rev	GCCCCAGCCGCCGTTGCTGCTCTGCATGCCCACAATCCA

Table S 9. Oligonucleotides for cloning of *Can*LCY-B in the pDHE1650 Vektor with His-tag and sequencing primer.

Gibson_ *Can*LCY-B	fwd	TGGTGCCGCGCGGCAGCCATATGGACACCCTCTTGC
Gibson_ *Can*LCY-B	rev	CAAAACAGCCGGTACCCTCATTCCTTGTCTTGAAGAAGGTTGTTAATC
Gibson_pDHE-His	fwd	CTTCAAGACAAGGAATGAGGGTACCGGCTGTTTTGG
Gibson_pDHE-His	rev	GTGCGCAAGAGGGTGTCCATATGGCTGCCGCGCGGCA
*Aac*SHC_SeqPR 1	fwd	CGCTGAGCATTGTGATGAGCCGC
*Aac*SHC_SeqPR 2	fwd	GCGTGAACTATCTGTATGGCACCGGCG
*Aac*SHC_SeqPR 3	rev	GTTAGCAGCCGGATCTC
RV 4	fwd	GACGATAGTCATGCCCCGC
pKK223 - 3 -for	rev	GGCGTTTCACTTCTGAGTTC

Table S 10. Vector constructs used in this work.

ITB No.	Vector	Relevant insert	Reference
	pET22b(+)		Vlada Urlacher, University Stuttgart
pITB285	pET22b(+)	*Aac*SHC	Stephan Hammer, University Stuttgart
pITB5137	pDHE1650	*Rg*CrtC[a]	Jens Schmid, University Stuttgart

	pDHE1650	*Can*LCY-B	Philip Horz, University Stuttgart
	pDHE1650	*Ath*LCY-B	Philip Horz, University Stuttgart
	pDHE1650	*Ath*LCY-E	Philip Horz, University Stuttgart
	pDHE1650-His	*Can*LCY-B	This work
	pDHE1650-His	*Ath*LCY-B	This work
	pDHE1650-His	*Ath*LCY-E	This work

[a]Carotenoid-1,2-hydratase from *Rubrivivax gelatinosus*[273]

Table S 11. Plasmids used in this work.

ITB No.	Plasmid	Vector	Insert	Mutations	Origin
pIITB6032	pET-22b::*Aac*SHC G600M R488P	pET-22b(+)	*Aac*SHC	G600M R488P	this work
pIITB6033	pET-22b:: *Aac*SHC G600M F605C	pET-22b(+)	*Aac*SHC	G600M F605C	this work
pIITB6034	pET-22b:: *Aac*SHC G600M R488P F605C	pET-22b(+)	*Aac*SHC	G600M R488P F605C	this work
pIITB6042	pET-22b:: *Aac*SHC G600M R488P F605C Y495D	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D	this work
pIITB6043	pET-22b:: *Aac*SHC G600M R488P F605C Y495D S445V	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V	this work
pIITB6044	pBAD:: *Aac*SHC G600M R488P F605C Y495D S445V	pBAD18	*Aac*SHC	G600M R488P F605C Y495D S445V	this work
pIITB6045	pET-22b: *Aac*SHC G600M R488P F605C Y495D S445V Y420F	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V Y420F	this work
pIITB6046	pET-22b:: *Aac*SHC G600M R488P F605C Y495D S445V I261V	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V I261V	this work
pIITB6047	pET-22b:: *Aac*SHC G600L Y420W	pET-22b(+)	*Aac*SHC	G600L Y420W	this work
pIITB6048	pET-22b:: *Aac*SHC G600T R488P F605C Y495D S445V	pET-22b(+)	*Aac*SHC	G600T R488P F605C Y495D S445V	this work
pIITB6049	pET-22b:: *Aac*SHC Y420F G600T R488P F605C Y495D S445V	pET-22b(+)	*Aac*SHC	Y420F G600T R488P F605C Y495D S445V	this work
pIITB6050	pET-22b:: *Aac*SHC G600L Y420W V448Q	pET-22b(+)	*Aac*SHC	G600L Y420W V448Q	this work

ITB No.	Plasmid	Vector	Insert	Mutations	Origin
pIITB6051	pET-22b:: *Aac*SHC G600L Y420W V448M	pET-22b(+)	*Aac*SHC	G600L Y420W V448M	this work
pIITB6052	pET-22b:: *Aac*SHC L607A G600T R488P Y495D S445V	pET-22b(+)	*Aac*SHC	L607A G600T R488P Y495D S445V	this work
pIITB6053	pET-22b:: *Aac*SHC L607A G600T R488P F605C Y495D S445V	pET-22b(+)	*Aac*SHC	L607A G600T R488P F605C Y495D S445V	this work
pIITB6054	pET-22b:: *Aac*SHC L36A Y420F T599M	pET-22b(+)	*Aac*SHC	L36A Y420F T599M	this work
pIITB6055	pET-22b:: *Aac*SHC G600L Y420W V448Q Q366V	pET-22b(+)	*Aac*SHC	G600L Y420W V448Q Q366V	this work
pIITB6056	pET-22b:: *Aac*SHC T599M G600T	pET-22b(+)	*Aac*SHC	T599M G600T	this work
pIITB6058	pET-22b:: *Aac*SHC G600M R488P F605C Y495D S445V W169G	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V W169G	this work
pIITB6059	pET-22b:: *Aac*SHC G600M Y420F R488P F605C Y495D S445V W169G	pET-22b(+)	*Aac*SHC	G600M Y420F R488P F605C Y495D S445V W169G	this work
pIITB6060	pET-22b:: *Aac*SHC Y420F W169G	pET-22b(+)	*Aac*SHC	Y420F W169G	this work
pIITB6061	pET-22b:: *Aac*SHC G600M R488P F605C Y495D S445V S168Y	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V S168Y	this work
pIITB6062	pET-22b:: *Aac*SHC G600M Y420F R488P F605C Y495D S445V S168Y	pET-22b(+)	*Aac*SHC	G600M Y420F R488P F605C Y495D S445V S168Y	this work
pIITB6063	pET-22b:: *Aac*SHC Y420F S168Y	pET-22b(+)	*Aac*SHC	Y420F S168Y	this work
pIITB6064	pET-22b:: *Aac*SHC Y420F R488P Y495D S445V W169G	pET-22b(+)	*Aac*SHC	Y420F R488P Y495D S445V W169G	this work
pIITB6065	pET-22b:: *Aac*SHC S445V	pET-22b(+)	*Aac*SHC	S445V	this work
pIITB6066	pET-22b:: *Aac*SHC R488P	pET-22b(+)	*Aac*SHC	R488P	this work
pIITB6067	pET-22b:: *Aac*SHC G600M Y420F R488P F605C Y495D S445V W168G S169Y	pET-22b(+)	*Aac*SHC	G600M Y420F R488P F605C Y495D S445V W168G S169Y	this work
pIITB6068	pET-22b:: *Aac*SHC Y420F W168G S169Y	pET-22b(+)	*Aac*SHC	Y420F W168G S169Y	this work

ITB No.	Plasmid	Vector	Insert	Mutations	Origin
pIITB6069	pET-22b:: *Aac*SHC G600M R488P F605C Y495D S445V W169G S168Y	pET-22b(+)	*Aac*SHC	G600M R488P F605C Y495D S445V W169G S168Y	this work
pIITB6070	pET-22b:: *Aac*SHC F605C	pET-22b(+)	*Aac*SHC	F605C	this work
pIITB6071	pET-22b:: *Aac*SHC R488P Y495D S445V	pET-22b(+)	*Aac*SHC	R488P Y495D S445V	this work
pIITB6072	pET-22b:: *Aac*SHC Y420F R488P Y495D S445V	pET-22b(+)	*Aac*SHC	Y420F R488P Y495D S445V	this work
pIITB6073	pET-22b:: *Aac*SHC R488P Y495D S445V S168Y	pET-22b(+)	*Aac*SHC	R488P Y495D S445V S168Y	this work
pIITB6074	pET-22b:: *Aac*SHC Y420F R488P Y495D S445V S168Y	pET-22b(+)	*Aac*SHC	Y420F R488P Y495D S445V S168Y	this work
pIITB6075	pET-22b:: *Aac*SHC Y420F R488P Y495D S445V W168G S168Y	pET-22b(+)	*Aac*SHC	Y420F R488P Y495D S445V W168G S168Y	this work
pIITB6076	pET-22b:: *Aac*SHC S168Y	pET-22b(+)	*Aac*SHC	S168Y	this work
pIITB6077	pET-22b:: *Aac*SHC W169G	pET-22b(+)	*Aac*SHC	W169G	this work
pIITB6078	pET-22b:: *Aac*SHC S168Y W169G	pET-22b(+)	*Aac*SHC	S168Y W169G	this work
pIITB6079	pET-22b:: *Aac*SHC R488P Y495D S445V W169G S168Y	pET-22b(+)	*Aac*SHC	R488P Y495D S445V W169G S168Y	this work
pIITB6080	pET-22b:: *Aac*SHC Y495D	pET-22b(+)	*Aac*SHC	Y495D	this work
pIITB6082	pET-22b:: *Aac*SHC G600M Y609L	pET-22b(+)	*Aac*SHC	G600M Y609L	this work
pIITB6083	pET-22b:: *Aac*SHC G600I Y420F	pET-22b(+)	*Aac*SHC	G600I Y420F	this work
pIITB6084	pET-22b:: *Aac*SHC F365L G600M Y609F F605L	pET-22b(+)	*Aac*SHC	AacSHC F365L G600M Y609F F605L	this work
pIITB6085	pET-22b:: *Aac*SHC F365L G600M Y609F F605R	pET-22b(+)	*Aac*SHC	AacSHC F365L G600M Y609F F605R	this work
pIITB6086	pET-22b:: *Aac*SHC G600M S445V	pET-22b(+)	*Aac*SHC	G600M S445V	this work
pIITB6087	pET-22b:: *Aac*SHC G600M Y495D	pET-22b(+)	*Aac*SHC	G600M Y495D	this work
pIITB6088	pET-22b:: *Aac*SHC S445V F605C	pET-22b(+)	*Aac*SHC	S445V F605C	this work
pIITB6089	pET-22b:: *Aac*SHC Y495D F605C	pET-22b(+)	*Aac*SHC	Y495D F605C	this work
pIITB6090	pET-22b:: *Aac*SHC S445V Y495D	pET-22b(+)	*Aac*SHC	S445V Y495D	this work

ITB No.	Plasmid	Vector	Insert	Mutations	Origin
pIITB6091	pET-22b:: *Aac*SHC R488P F605C	pET-22b(+)	*Aac*SHC	R488P F605C	this work
pIITB6092	pET-22b:: *Aac*SHC R488P S445V	pET-22b(+)	*Aac*SHC	R488P S445V	this work
pIITB6093	pET-22b:: *Aac*SHC R488P Y495D	pET-22b(+)	*Aac*SHC	R488P Y495D	this work
pIITB6094	pET-22b:: *Aac*SHC F365L G600M F605L Y609F L607F	pET-22b(+)	*Aac*SHC	F365L G600M F605L Y609F L607F	Bachelor thesis Daniela Ramelsoh[1]
pIITB6095	pET-22b:: *Aac*SHC F365L G600M Y609F	pET-22b(+)	*Aac*SHC	F365L G600M Y609F	this work
pIITB6096	pET-22b:: *Aac*SHC F365L G600M Y609F Y420F	pET-22b(+)	*Aac*SHC	F365L G600M Y609F Y420F	this work
pIITB6097	pET-22b:: *Aac*SHC G600M Y609F	pET-22b(+)	*Aac*SHC	G600M Y609F	this work
pIITB6098	pDHE::LycE	pDHE1650	*Ath*LCY-E		Philip Horz[1]
pIITB6099	pDHE::LycC	pDHE1650	*Can*LCY-B		Philip Horz[1]
pIITB6100	pDHE::HisLycC	pDHE1650	*Can*LCY-B		this work
pIITB6101	pHDHE1650::LycBC V335L	pDHE1650	*Can*LCY-B	V335L	Bachelor thesis Daniela Ramelsoh[1]
pIITB6102	pHDHE1650::LycBC W118A	pDHE1650	*Can*LCY-B	W118A	Bachelor thesis Daniela Ramelsoh[1]
pIITB6103	pHDHE1650::LycBC W118F	pDHE1650	*Can*LCY-B	W118F	Bachelor thesis Daniela Ramelsoh[1]
pIITB6104	pHDHE1650::LycBC D331A	pDHE1650	*Can*LCY-B	D331A	Bachelor thesis Daniela Ramelsoh[1]
pIITB6105	pHDHE1650::LycBC C334A	pDHE1650	*Can*LCY-B	C334A	Bachelor thesis Daniela Ramelsoh[1]
pIITB6106	pHDHE1650::LycBC C334S	pDHE1650	*Can*LCY-B	C334S	Bachelor thesis Daniela Ramelsoh[1]
pIITB6107	pHDHE1650::LycBC V335A	pDHE1650	*Can*LCY-B	V335A	Bachelor thesis Daniela Ramelsoh[1]
pIITB6108	pHDHE1650::LycBC M366A	pDHE1650	*Can*LCY-B	M366A	Bachelor thesis Daniela Ramelsoh[1]

ITB No.	Plasmid	Vector	Insert	Mutations	Origin
pIITB6109	pHDHE1650::LycBC M366S	pDHE1650	*Can*LCY-B	M366S	Bachelor thesis Daniela Ramelsoh[1]
pIITB6110	pHDHE1650::LycE S359F	pDHE1650	*Ath*LCY-E	S359F	this work

[1] University Stuttgart

NMR spectra of substrates and products
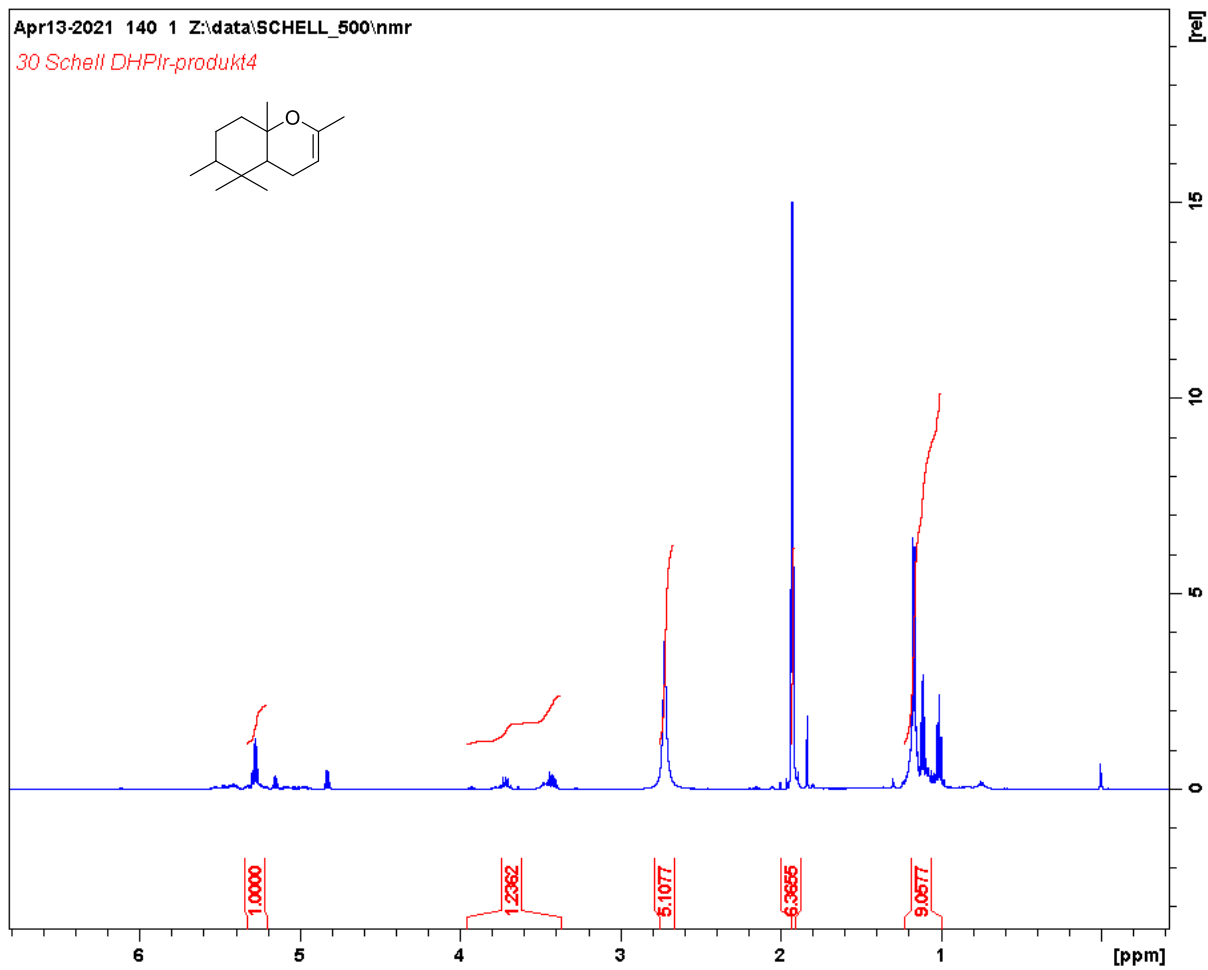
(*E*)-9-hydroxy-6,10-dimethylundeca-5,10-dien-2-one (7)

^{1}H-NMR

¹³C-NMR

30 Trinks Gdiol
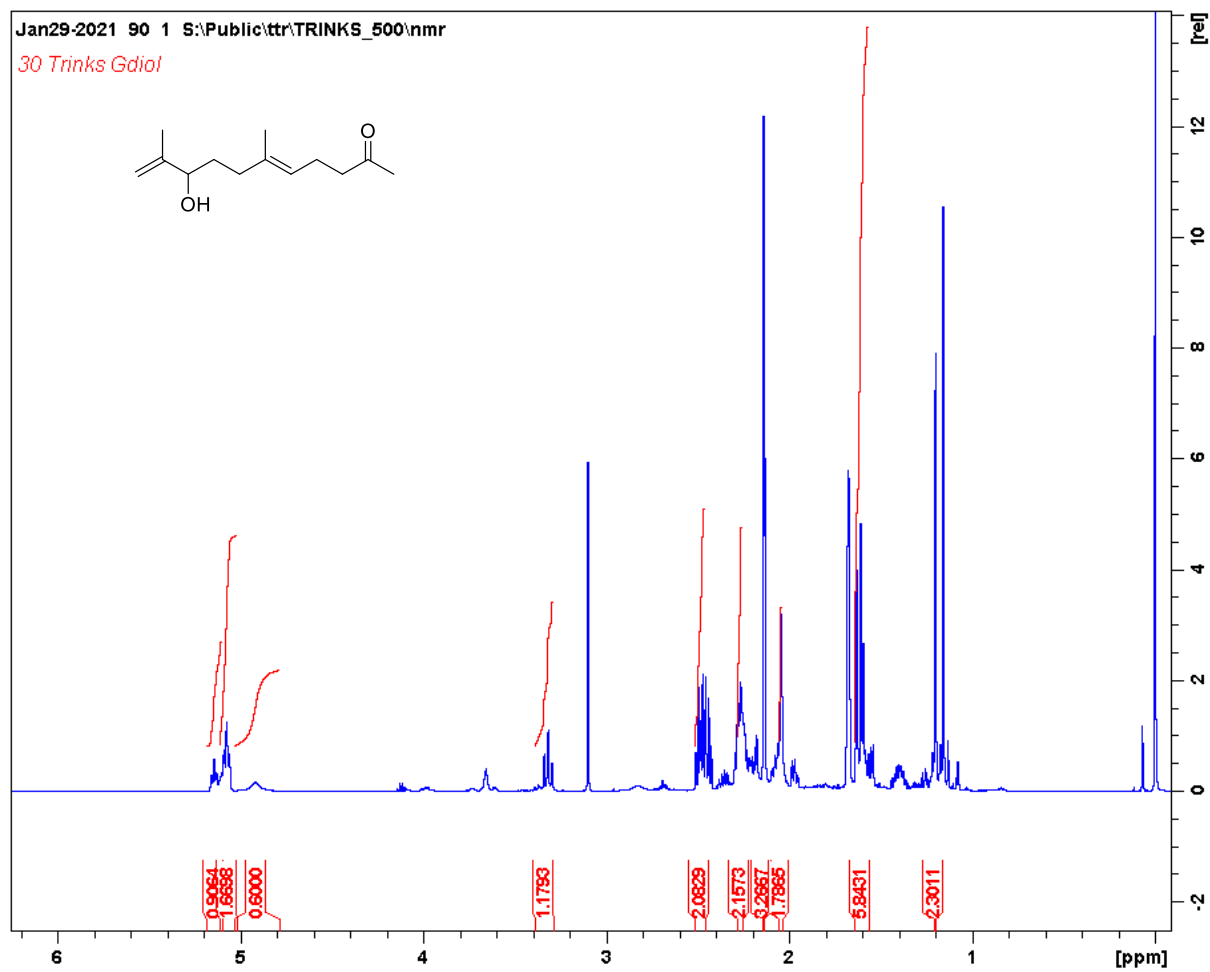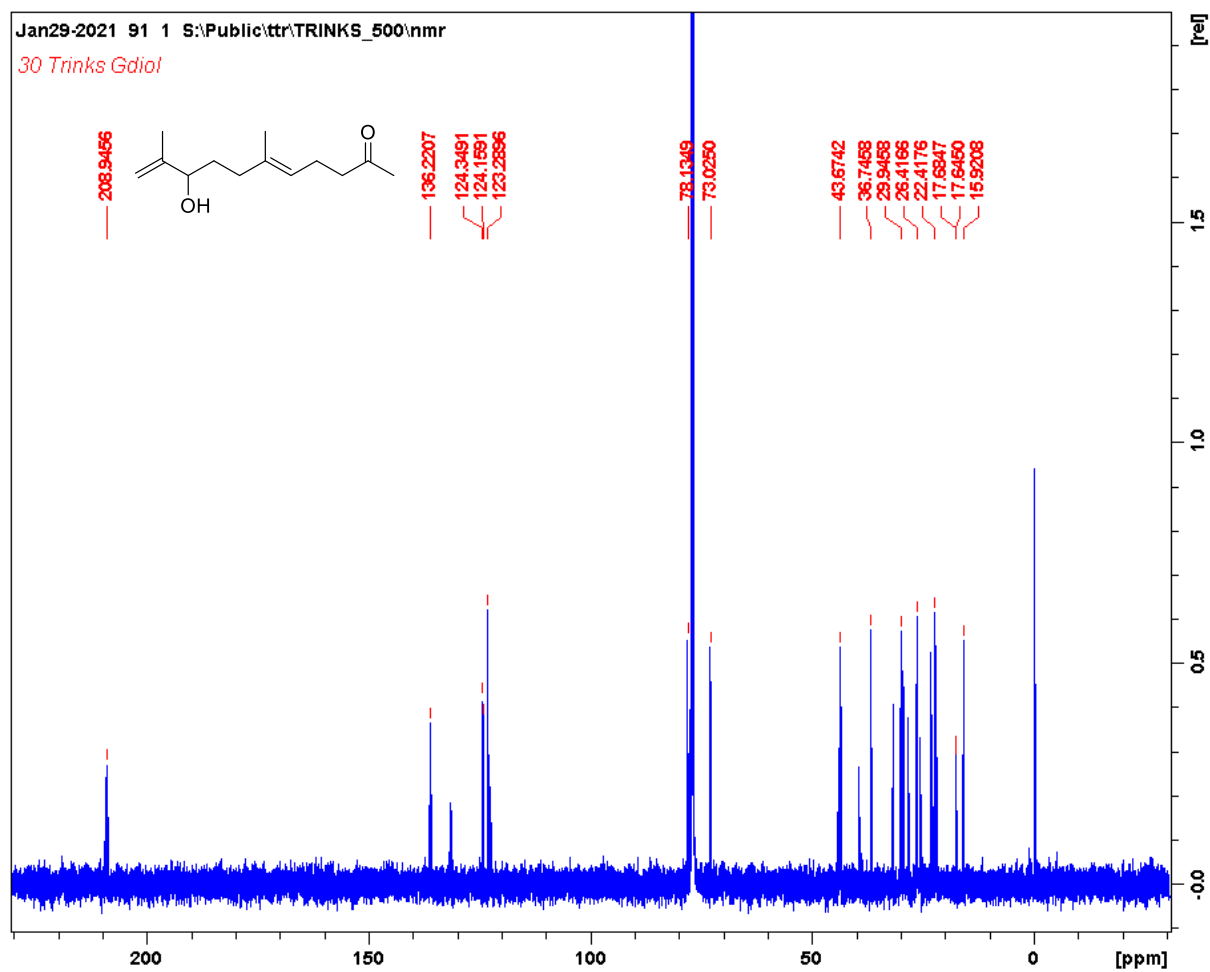
(*Z*)-6,9,10-trimethylundeca-5,9-dien-2-one (*Z*-2)

¹H-NMR
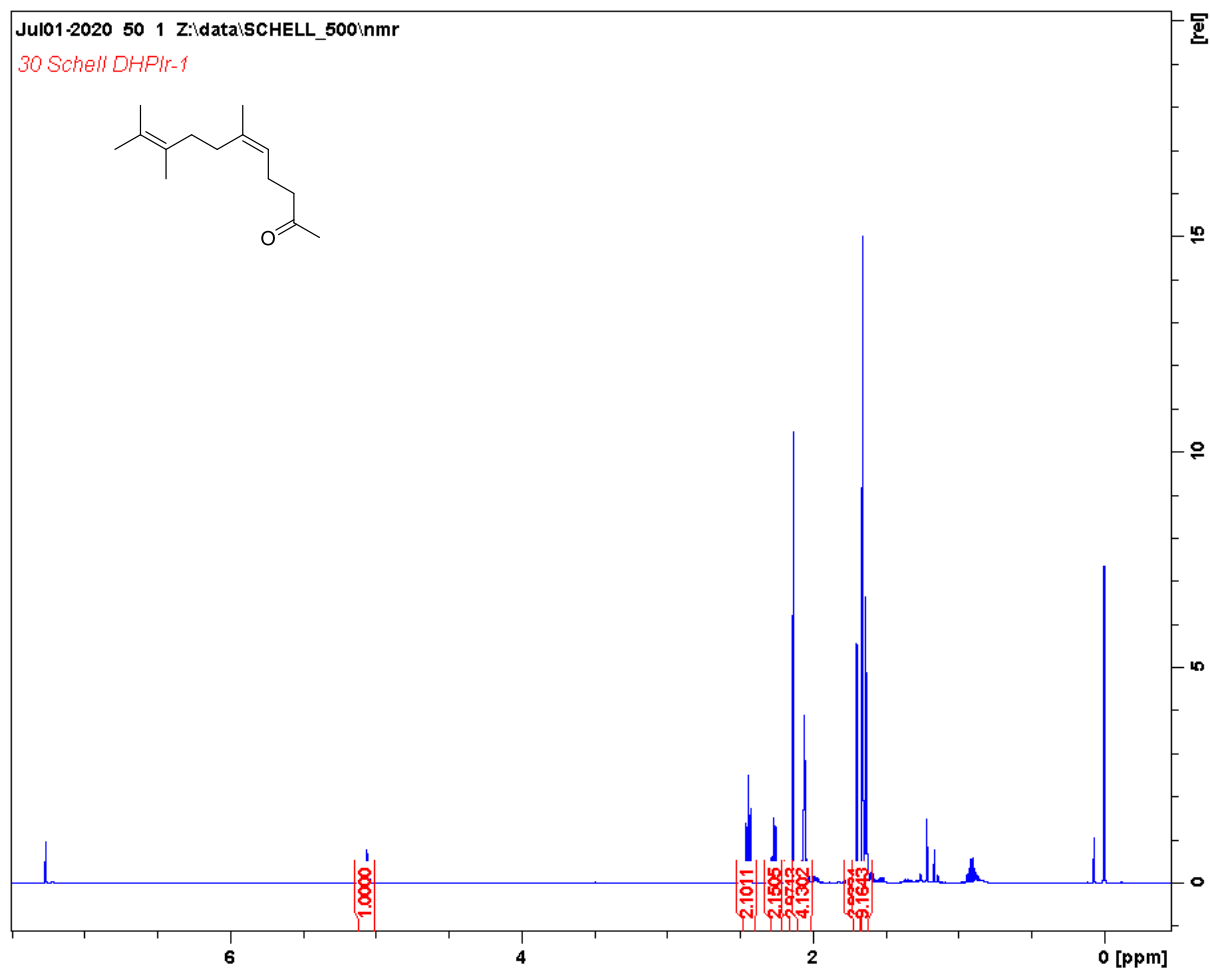
^{13}C-NMR
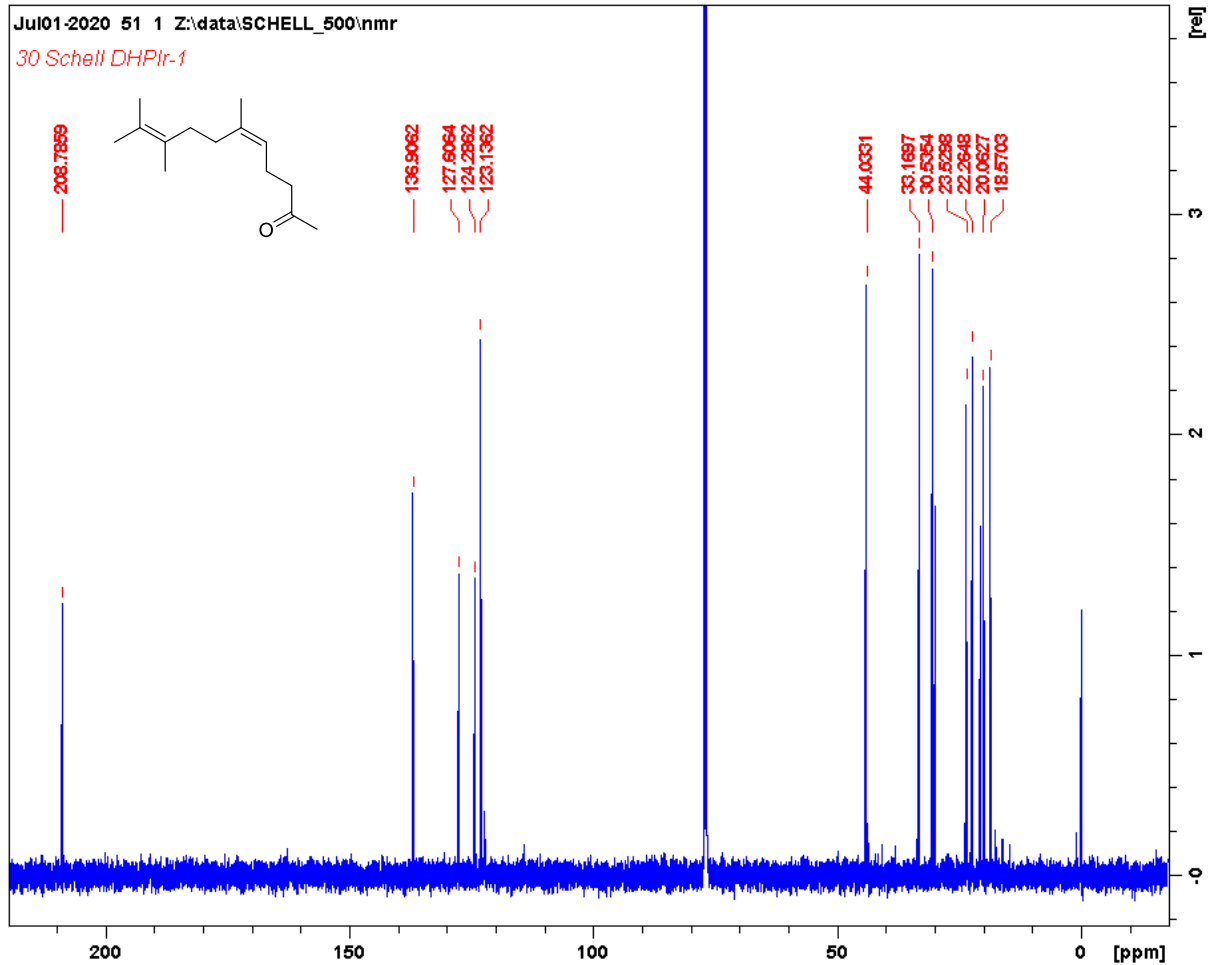
30 Schell DHPIr-1

(*E*)-6,9,10-trimethylundeca-5,9-dien-2-one (*E*-2)

¹H-NMR
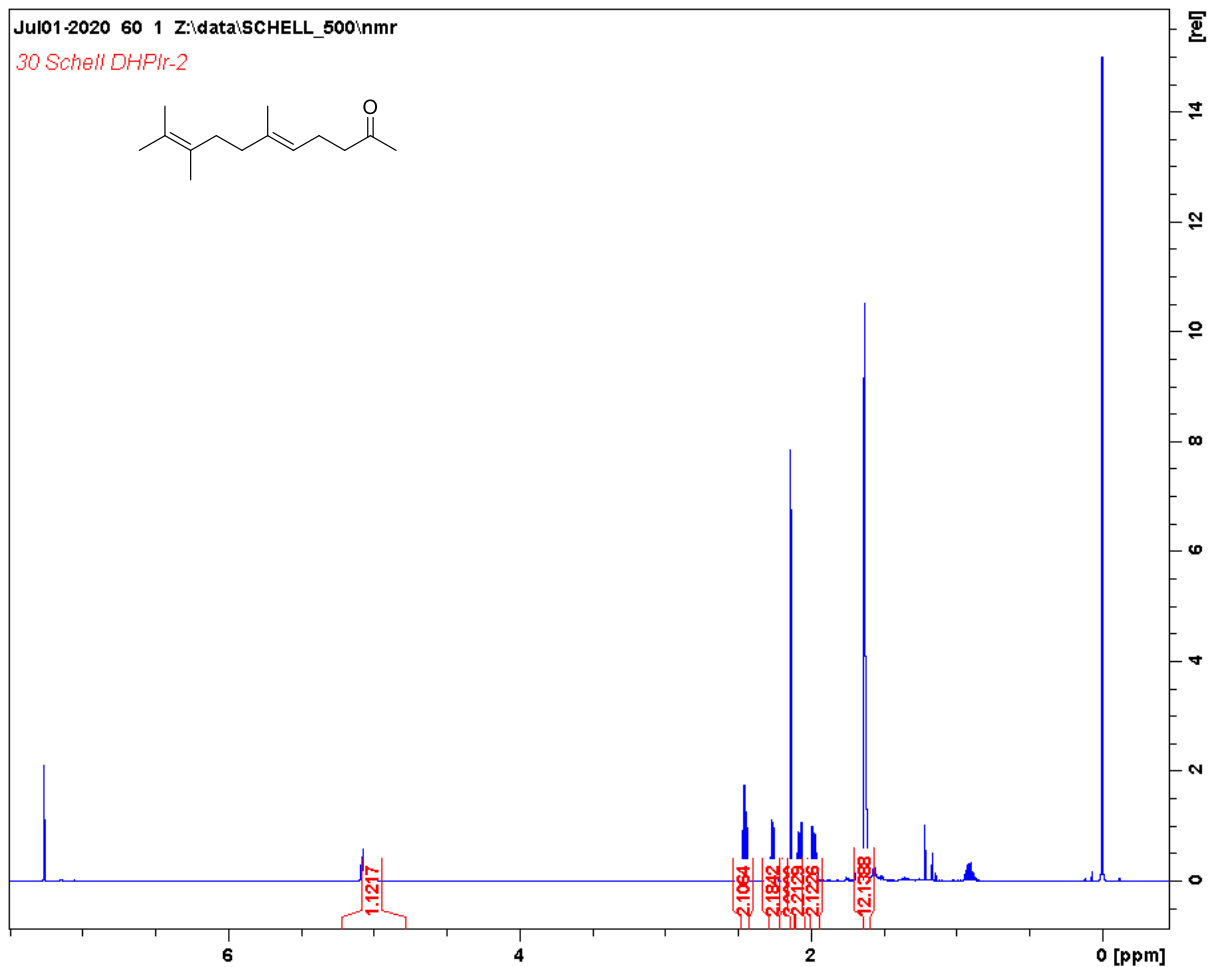
^{13}C-NMR
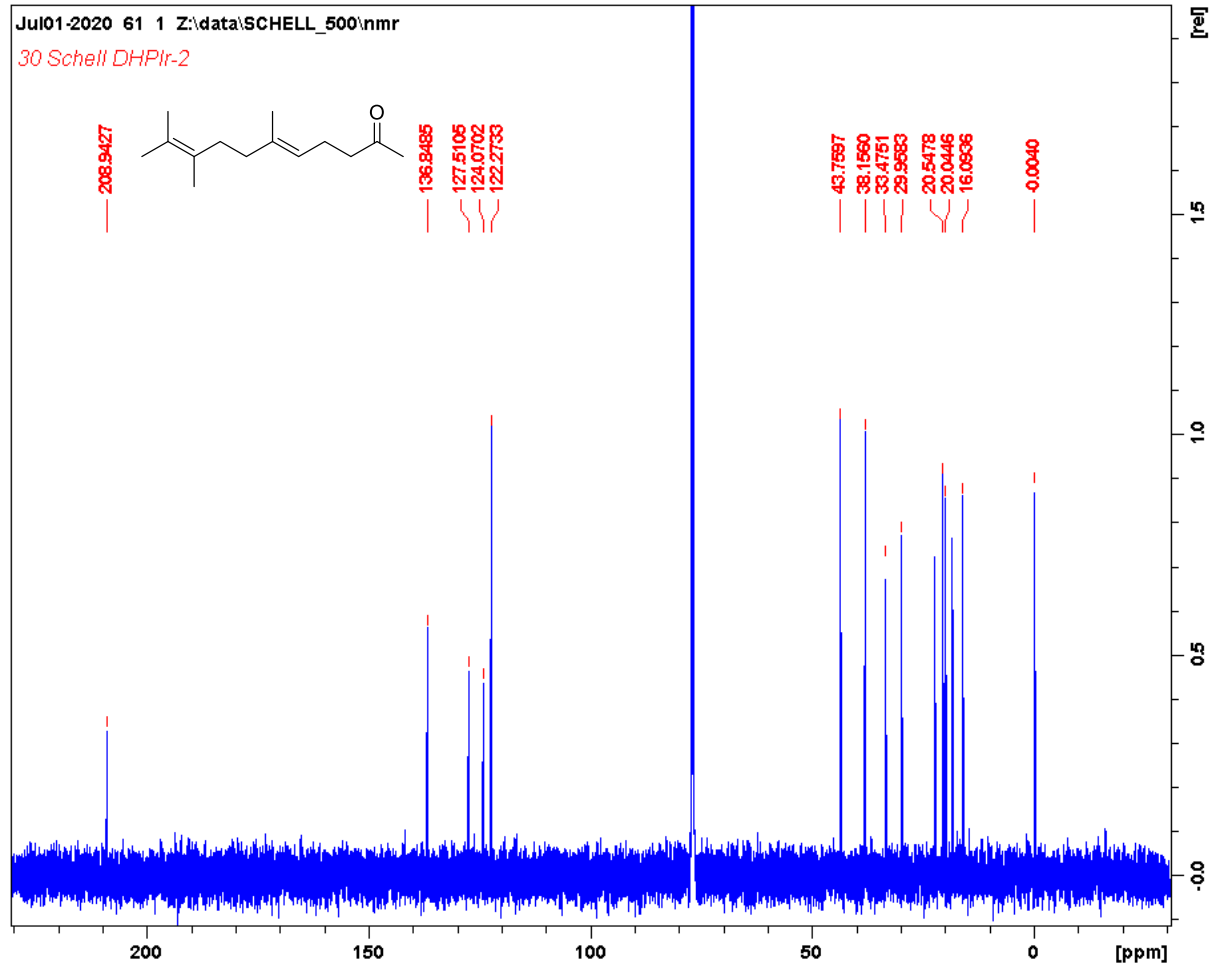
30 Schell DHPIr-2

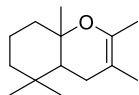## 2,5,5,6,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (2a)

¹H-NMR

^{13}C-NMR

¹H,¹³C-HSQC-NMR

Feb15-2021 24 1 Z:\data\SCHELL_700\nmr

30 Schell DHPr-Produkt-3

^{1}H,^{13}C-HMBC-NMR

Feb15-2021 23 1 Z:\data\SCHELL_700\nmr

30 Schell DHPr-Produkt-3

¹H,¹³C-COSY-NMR

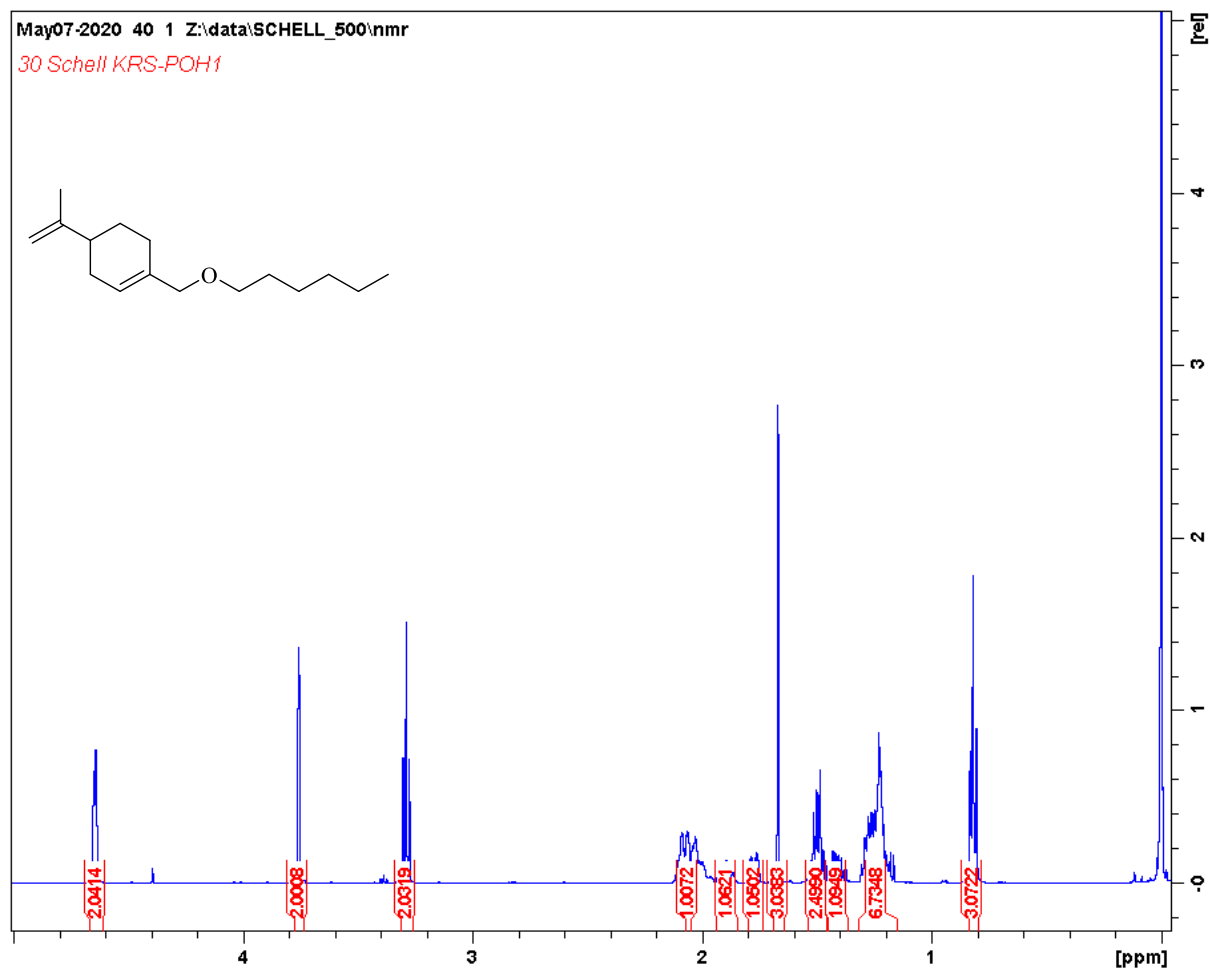Feb15-2021 22 1 Z:\data\SCHELL_700\nmr

30 Schell DHPr-Produkt-3
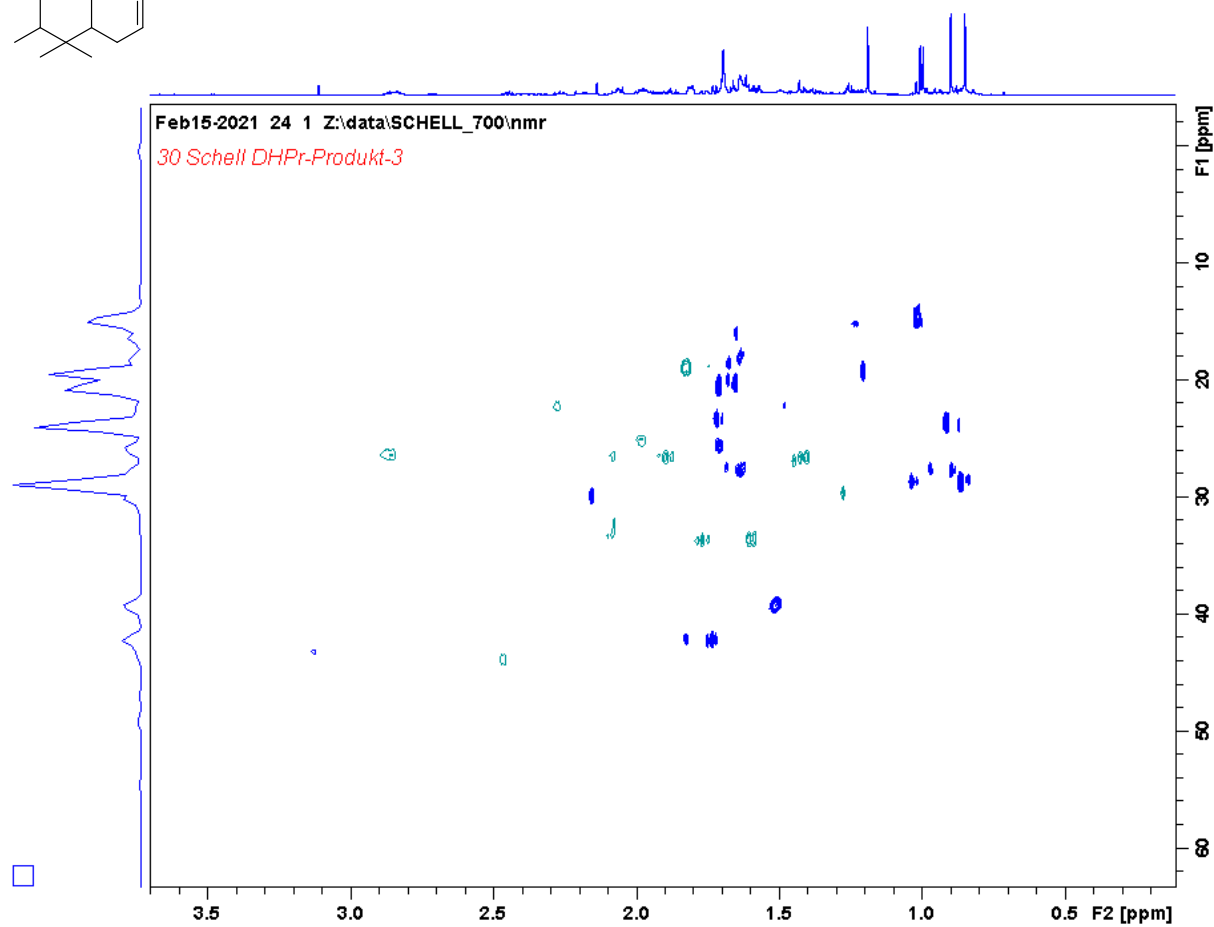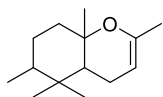
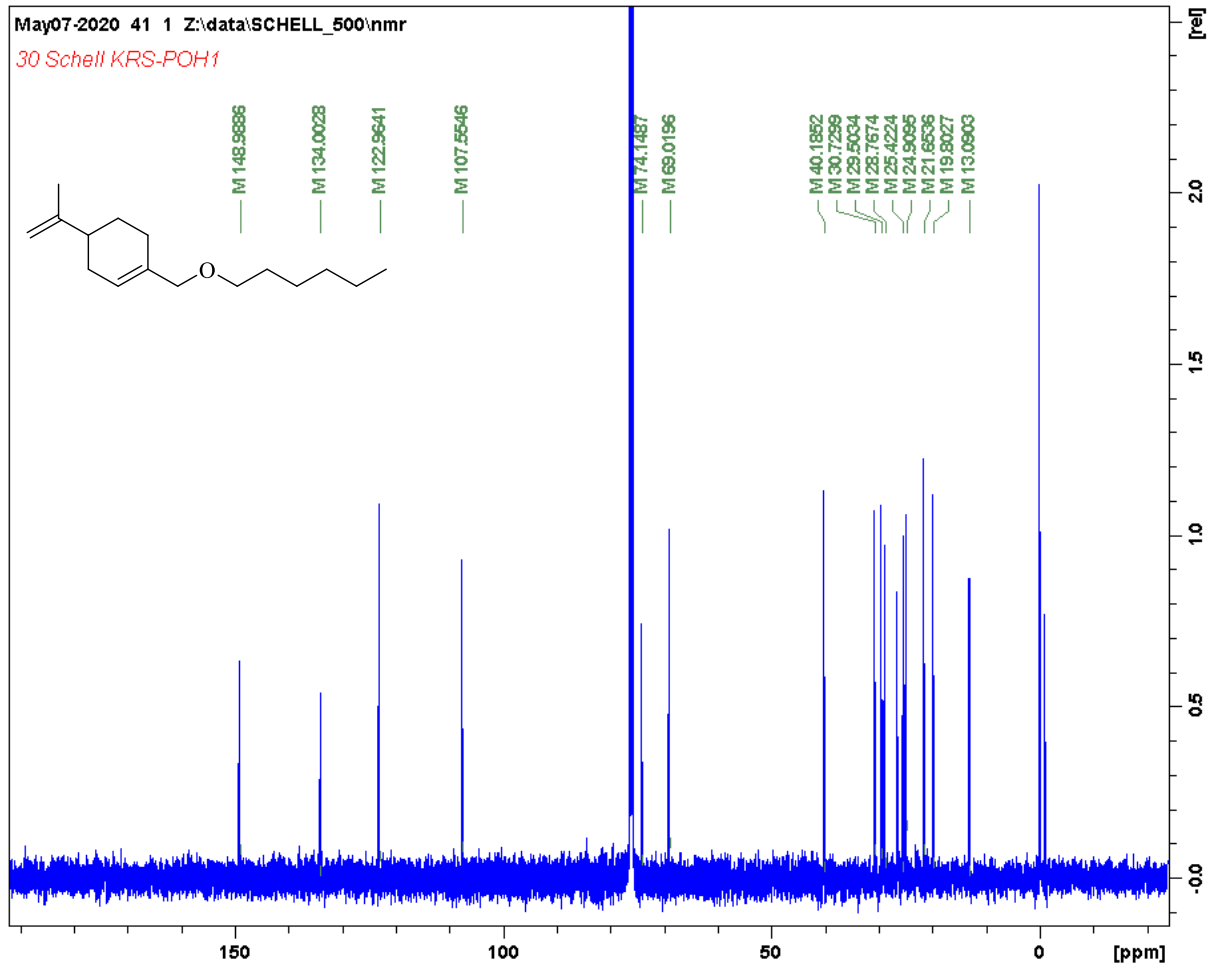## 2,3,5,5,8a-pentamethyl-4a,5,6,7,8,8a-hexahydro-4H-chromene (6a)
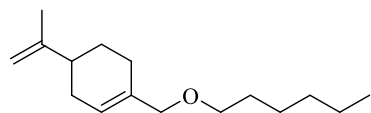
¹H-NMR

^{13}C-NMR

Nov19-2021 51 1 Z:\data\SCHELL_500\nmr

30 Schell C3-Produkt2

141.2098 · 100.9879 · 75.5390 · 49.0141 · 41.6069 · 39.8965 · 33.1134 · 32.1705 · 25.6494 · 20.5903 · 19.8381 · 19.2087 · 17.9482 · 16.6726

¹H,¹³C-HSQC-NMR

¹H,¹³C-HMBC-NMR

1-((hexyloxy)methyl)-4-(prop-1-en-2-yl)cyclohex-1-ene

¹H-NMR

May07-2020 40 1 Z:\data\SCHELL_500\nmr

30 Schell KRS-POH1

^{13}C-NMR

May07-2020 41 1 Z:\data\SCHELL_500\nmr

30 Schell KRS-POH1

M 148.9886
M 134.0028
M 122.9641
M 107.5546
M 74.1487
M 69.0196
M 40.1852
M 30.7299
M 29.5034
M 28.7674
M 25.4224
M 24.9095
M 21.6536
M 19.8027
M 13.0903

¹H,¹³C-COSY-NMR

May07-2020 42 1 Z:\data\SCHELL_500\nmr

30 Schell KRS-POH1

^{1}H,^{13}C-HSQC-NMR

May07-2020 44 1 Z:\data\SCHELL_500\nmr
30 Schell KRS-POH1

^{1}H,^{13}C-HMBC-NMR