Computer-assisted optimization of cannabinoid biosynthesis catalysts and evaluation of their potential to produce unnatural cannabinoids

Zur Erlangung des akademischen Grades eines

Dr. rer. nat.

von der Fakultät Bio- und Chemieingenieurwesen
der Technischen Universität Dortmund
genehmigte Dissertation

vorgelegt von

M. Sc. Saskia Spitzer

aus

Düren

Tag der mündlichen Prüfung: 09.10.2025

1. Gutachter/-in: Prof. Dr. Oliver Kayser
2. Gutachter/-in: Prof. Dr. Uwe Bornscheuer

Dortmund 2025

Achknowledgements

First, I would like to thank Prof. Dr. Oliver Kayser for giving me the opportunity to work on such a fascinating topic and for guiding me through problems with his valuable advice.

I would also like to thank Jasmin Wloka and Prof. Dr. Pietruszka for collaborating on a significant part of my thesis, providing interesting compounds, and advancing my work, which ultimately resulted in a valuable publication.

I enjoyed my time at the Biozentrum, and it was a pleasure sharing my work experience with everyone, learning many things while keeping up a great atmosphere. I want to thank all the technical and scientific staff who assisted me during my thesis, providing a necessary helping hand when needed. Thanks to Kristine, Jörg, Bettina, Julian, Chantale, Katharina, Don, Sascha, Marco, and Dr. Quentmeier. Thanks to all the PhD students of the Biocentrum for all the special moments and scientific exchanges I will never forget. Special thanks to the TB PhD Students Julia, Leonie, Thanet, Fabian, Olli, Gia-Nam, Selin, Christina, and Sophie.

Additionally, I would like to extend a big thank you to all my supervised students, including Michelle Donner, Sarah Hallek, Lars Reckert, Mara Loddenkemper, and David Kemple. Your work helped me a lot during my thesis and provided interesting insides and results in some of the topics of my PhD.

Last but not least, I would like to thank all of my friends and family who helped me throughout this thesis. I know it was a lot of ups and downs, but you always stood by my side and helped me with everything.

Table of Content

Abstract

Cannabinoids are a compound class that originally arose as secondary metabolites from the plant *Cannabis sativa*. Molecules such as (−)-*trans*-Δ9-tetrahydrocannabinol, cannabidiol, and cannabichromene are highly demanded on the market due to their diverse medical applications. Given the extraction process from plant, the isolation of rare cannabinoids will not be sufficient to meet global demand. Therefore, an alternative production method is required. Additionally, the interest in non-natural cannabinoids, which act on cannabis receptors, continues to rise. Ensuring the sustainable and high-titer production of these novel cannabinoid structures from the initial stages is crucial. A recent approach that has emerged is the heterologous production of cannabinoids in *Saccharomyces cerevisiae*. In addition to the naturally produced products from cannabinoid biosynthesis, the proteins can be used for the formation of non-natural cannabinoids through the exchange of the precursor molecules.

A bottleneck of the production of cannabinoids is the reaction from hexanoyl-CoA towards olivetolic acid (OA). This reaction results in the formation of olivetol, a byproduct that leads to the wastage of precursors including malonyl-CoA, acetyl-CoA, and glucose. Computer-based methods were applied to identify olivetolic acid cyclase (*Cs*OAC) variants, with the result that these variants enhanced the production of OA while decreasing the production of olivetol in *S. cerevisiae*.

Due to the exchange of hexanoic acid, fed through the media, with other fatty acids and the use of already integrated enzymes in an *S. cerevisiae* strain, it is possible to easily form novel OA analogs with only minor changes in the workflow. The following step involves the prenylation of OA using NphB. It has been demonstrated that the high promiscuity of NphB enables the formation of cannabigerol acid derivatives when OA derivatives are applied as substrates. Docking results helped predict structures that are converted to CBGA derivatives.

Zusammenfassung

Cannabinoide sind eine Molekülklasse, die ursprünglich als Sekundärmetaboliten aus der Pflanze *Cannabis sativa* entstanden sind. Moleküle wie (-)*trans*-Δ9-Tetrahydrocannabinol, Cannabidiol und Cannabichromen sind aufgrund ihrer vielfältigen medizinischen Anwendungen auf dem Markt sehr gefragt. Angesichts des Extraktionsverfahrens aus der Pflanze wird die Isolierung seltener Cannabinoide nicht ausreichen, um die weltweite Nachfrage zu decken. Daher ist eine alternative Produktionsmethode erforderlich. Darüber hinaus nimmt das Interesse an nicht natürlichen Cannabinoiden, die auf Cannabisrezeptoren wirken, weiter zu. Die Sicherstellung einer nachhaltigen und hochtitrigen Produktion dieser neuartigen Cannabinoidstrukturen von Anfang an ist von entscheidender Bedeutung. Ein neuer Ansatz ist die heterologe Produktion von Cannabinoiden in *Saccharomyces cerevisiae.* Zusätzlich zu den natürlich hergestellten Produkten der Cannabinoid-Biosynthese können die Proteine durch den Austausch der Präkursormoleküle zur Bildung nicht natürlicher Cannabinoide verwendet werden.

Ein Flaschenhals bei der Herstellung von Cannabinoiden ist die Reaktion von Hexanoyl-CoA zu Olivetolsäure (OA). Diese Reaktion führt zur Bildung von Olivetol, einem Nebenprodukt, das zur Verschwendung von Vorläufern wie Malonyl-CoA, Acetyl-CoA und Glukose führt. Es wurden computergestützte Methoden zur Identifizierung von Olivetolsäure-Cyclase (*Cs*OAC) Varianten angewandt, mit dem Ergebnis, dass diese Varianten die Produktion von OA verstärken, während die Produktion von Olivetol in *S. cerevisiae* verringert wird.

Durch den Austausch von Hexansäure, die über das Medium zugeführt wird, mit anderen Fettsäuren und die Verwendung bereits integrierter Enzyme in einem *S. cerevisiae*-Stamm ist es möglich, mit nur geringfügigen Änderungen des Arbeitsablaufs neue OA-Analoga zu bilden. Der folgende Schritt umfasst die Prenylierung von OA mit NphB. Es wurde gezeigt, dass die hohe Promiskuität von NphB die Bildung von Cannabigerolsäurederivaten ermöglicht, wenn OA-Derivate als Substrate eingesetzt werden. Docking-Ergebnisse halfen bei der Vorhersage von Strukturen, die in CBGA-Derivate umgewandelt werden.

Chapter 1 Introduction

1.1 Phytocannabinoids

Cannabis sativa has been cultivated and utilized for over 12,000 years for the production of hemp or its positive and psychoactive properties (*1*). However, it was not until the 20th century that the first cannabinoid, cannabidiol (CBD-C5), was identified and isolated (*2*). Only a year after the structure elucidation of CBD-C5, the molecule with psychoactivity was identified as (−)-*trans*-Δ9-tetrahydrocannabinol (THC-C5) (*3*). The phytocannabinoid class was introduced based on the structural similarities of these and other secondary metabolites isolated from the plant. These structures are composed of a resorcinyl core equipped with a *para*-positioned alkyl, aralkyl, or isoprenyl moiety. The structure is highly dependent on the polyketide starter molecule. Alkyl chains are constituted of an odd number of carbon atoms (*4*) (Figure 1). To this date, more than 113 distinct phytocannabinoids have been isolated from the plant *C. sativa* (*5*). The phytocannabinoids derived from the majority of the other cannabinoid structures extracted from *C. sativa* are cannabigerolic acid (CBGA-C5), tetrahydrocannabinolic acid (THCA-C5), cannabidiolic acid (CBDA-C5) and cannabichromenic acid (CBCA-C5). From these structures, the four major cannabinoids, namely THC-C5, cannabichromene (CBC-C5), CBD-C5 and cannabigerol (CBG-C5), are derived through a process of heat-induced decarboxylation (*6, 7*). Phytocannabinoids present in low concentrations in the plant are designated as minor or rare cannabinoids (*8*). Examples of minor cannabinoids with a different ring structure are cannabinol (CBN-C5) (*9, 10*) or cannabicyclol (CBL-C5) (*11*). Furthermore, the alkyl moiety may vary, with the pentyl alkyl chain potentially being replaced with a methyl or a propyl chain, resulting in compounds such as cannabigerovarin (CBG-C3), tetrahydrocannabivaranic acid (THCA-C3) or cannabidiresorcin (CBD-C1) (*12*).

In addition to the identification of phytocannabinoids in *C. sativa*, these compounds have also been found in other organisms. Examples of known organisms that produce phytocannabinoids include the plant *Helichrysum umbraculigerum* from South Africa, the liverwort *Radula marginata* from New Zealand, and the fungi *Cylindrocarpon olidum* (*13–15*). Next to the same cannabinoids that are also produced in *C. sativa*, cannabinoids from *H. umbraculigerum* or *R. marginata* also include arakyl phytocannabinoids. Arakylcannabinoids exhibit a cyclic structure, which can be a bibenzyl or a stilbenyl group attached to the resorcinyl core instead of an alkyl group. For *H. umbraculigerum*, isolated and identified examples include cannabibenphenylgerolic acid (CBGA-BB), and for *R. marginata* are tetrahydrocannabibiphenylic acid (THCA-BB) or cannabistyrengerolic acid (CBGA-ST).

Additionally, abnormal arakylcannabinoids, were identified, wherein the isoprenyl moiety is attached at a different position within the resorcinyl core (*13, 14, 4, 16, 17*).

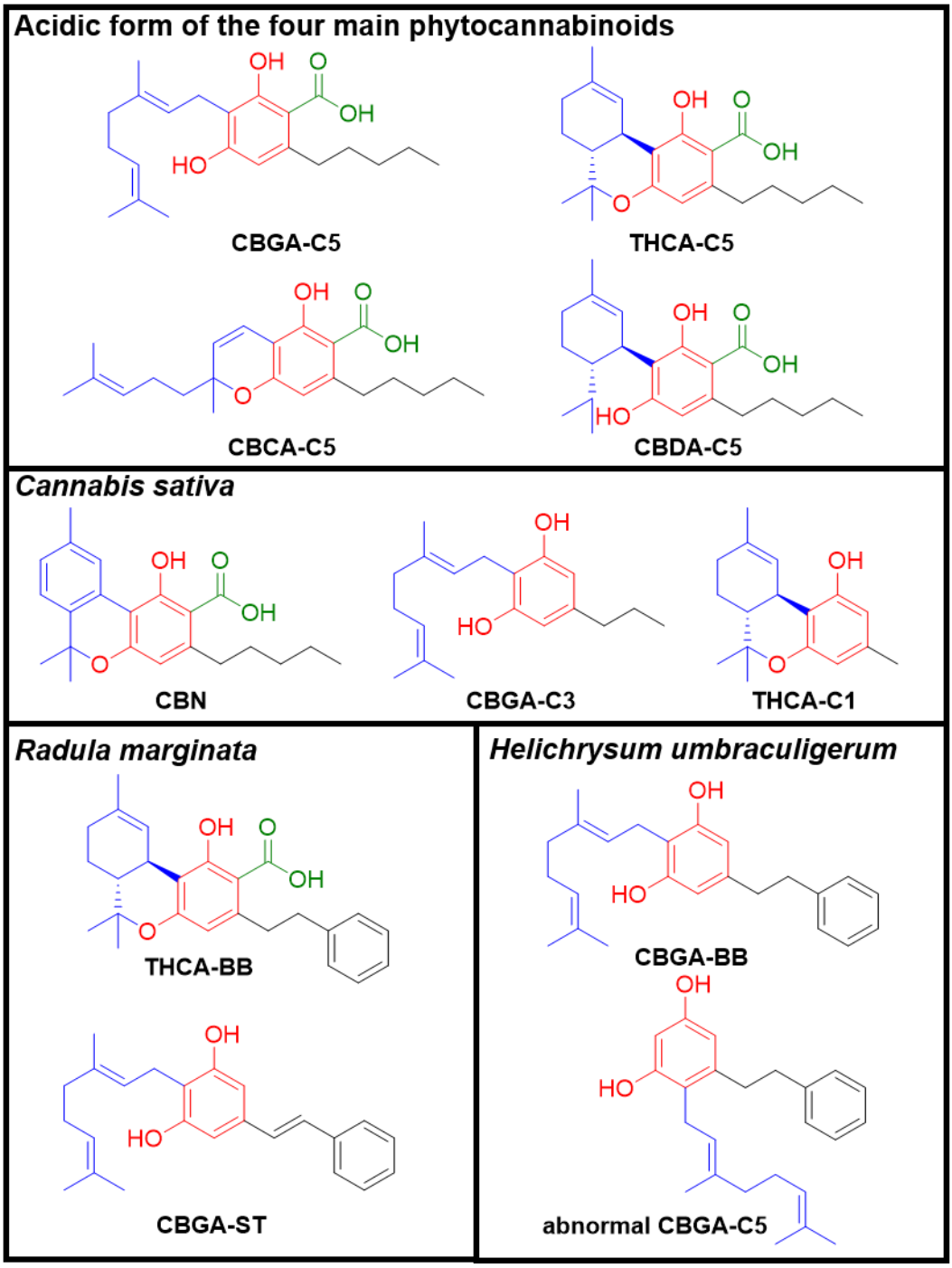


Figure 1: Phytocannabinoid selection including origin organisms. The structure derived from the isoprenoid precursor geranyl diphosphate are marked in blue, the resorcinyl core structure derived from the polyketide starter molecule is marked in red, the altering side chain is marked in black and the carboxylic acid group is marked in green.

1.2 Cannabinoid Biosynthesis

The most well-known and utilized source of cannabinoids remains the isolation of the plant *C. sativa*. The biosynthesis of cannabinoids occurs in the glandular trichrome of the female plants (*18*). For biosynthesis, enzymatic reactions are mainly involved, except for the last step, in which a chemical reaction occurs (Figure 2). The initial formation of the first phytocannabinoid structure is comprised of two precursor molecules, olivetolic acid (OA) and geranyl diphosphate (GPP). These two structures are generated through two different biosynthetic pathways. One pathway starts with the molecule hexanoic acid derived from the fatty acid biosynthesis pathway in *C. sativa* (*19*). The acyl-activating enzyme 1 (*Cs*AAE1) catalyzes the following reaction from hexanoic acid towards hexanoyl-CoA, consuming one acetyl-CoA (*20*). For the biosynthesis of OA, the cooperation of the two proteins, olivetol synthase (*Cs*OLS) and olivetolic acid cyclase (*Cs*OAC), is necessary. The enzyme *Cs*OLS catalyzes the successive condensation of hexanoyl-CoA with three molecules of malonyl-CoA, yielding 3,5,7-trioxododecaneoyl-CoA (*21*). This structure undergoes intramolecular aldol condensation to OA with the assistance of *Cs*OAC (*22*). In *C. sativa* the supply of GPP is ensured by the plastidial 2-methyl-d-erythritol-4-phosphate (MEP) pathway or the mevalonic acid (MVA) pathway (*23*).

The prenylation of OA with GPP is catalyzed by a geranylpyrophosphate:olivetolate geranyltransferase (GOT) to CBGA-C5 (*24*). Two GOTs have been identified in *C. sativa*, which are able to catalyze the reaction in *C. sativa* with *Cannabis* GOT1 (CsPT1) and *Cannabis* GOT4 (*Cs*PT4) (*25*).

The final enzymatic step of cannabinoid biosynthesis is the oxidative cyclization of CBGA-C5 towards the tricyclic structure of THCA-C5 or bicyclic structures like CBDA-C5 or CBCA-C5 (*26*). Three enzymes were identified in *C. sativa*, all of which are capable of catalyzing the final enzymatic reaction. However, the enzymes result in the formation of different products through their different regioselectivity (*18*). The first identified protein is the tetrahydrocannabinolic acid synthase (*Cs*THCAS), which mainly produces THCA-C5 (*27*). Subsequently, the cannabidiolic acid synthase (*Cs*CBDAS) and the cannabichromenic acid synthase (*Cs*CBCAS) were identified, which mainly produce CBDA-C5 or CBCA-C5 respectively (*28, 29*). Due to the thermally unstable nature of cannabinoid acids, the decarboxylation reaction towards the neutral cannabinoids occurs in the plant due to heat or light exposure (*7*).

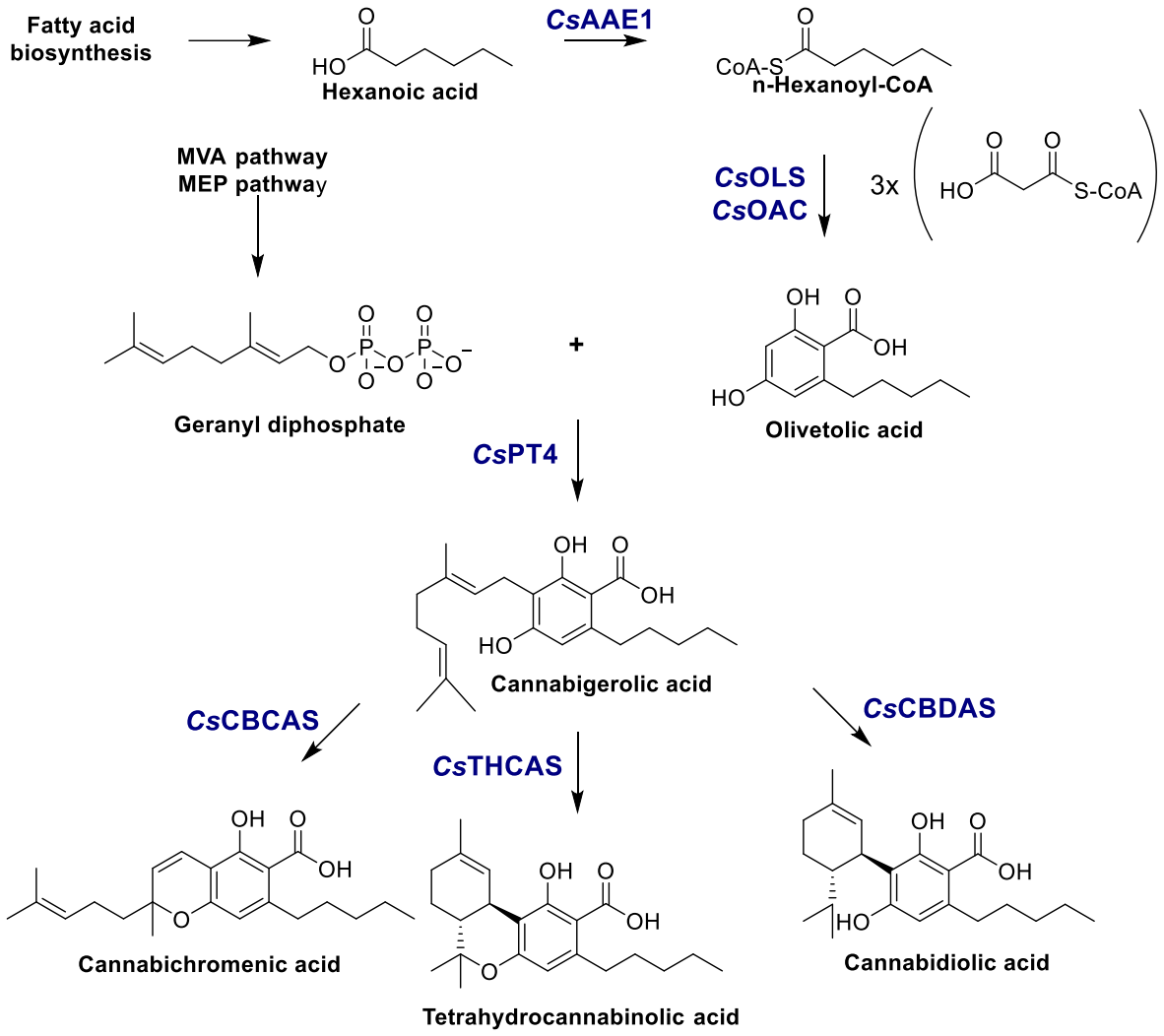


Figure 2: Cannabinoid biosynthesis in *C. sativa*. Hexanoic acid is provided through fatty acid biosynthesis. An enzymatic reaction catalyzed by Acyl activating enzyme 1 (*Cs*AAE1) n-Hexanoyl-CoA is formed. Following a multi-step reaction consuming three malonyl-CoA and using Olivetol synthase (*Cs*OLS) and olivetolic acid cyclase (*Cs*OAC), the polyketide olivetolic acid is generated. For the next catalytic step, the reaction partner of olivetolic acid is geranyl diphosphate, derived from the mevalonate (MVA) and methylerythritol phosphate (MEP) pathways. The prenylation reaction towards cannabigerolic acid (CBGA-C5) is catalyzed by geranylpyrophosphate:olivetolate geranyltransferase 4 (*Cs*PT4). CBGA-C5 is converted by the three enzymes cannabichromenic acid synthase (*Cs*CBCAS), tetrahydocannabinolic acid synthase (*Cs*THCAS) or cannabidiolic acid synthase (*Cs*CBDAS) to yield, respectively, cannabichromenic acid, tetrahydrocannabinolic acid, and cannabidiolic acid.

In addition to the decarboxylation reaction, a small percentage of other chemical reactions may occur, forming minor cannabinoids. For instance, the formation of CBL involves CBC-C5 undergoing an intramolecular stereoselective [2+2] cycloaddition, which leads to an additional C–O bond formation during light exposure (*4*). Furthermore, THC-C5 and THCA-C5 can undergo further degradation under the condition of air oxidation, resulting in aromatization at the level of the menthyl moiety, leading to the formation of CBN-C5 or CBNA-C5, respectively (*30, 10*).

1.2.1 Olivetol Synthase *Cs*OLS

In the biosynthesis of plant secondary metabolites, the class of type III polyketide synthases (PKSs) plays an important role (*31*). A type III PKS has also been identified in cannabinoid biosynthesis with the *Cs*OLS (*21*). Type III PKSs catalyze the condensation of a variable number of molecules onto a starter unit through iterative decarboxylative Claisen condensation reactions (*32*). Interestingly, the *Cs*OLS itself is not able to catalyze the cyclization towards OA without the presence of the *Cs*OAC (*22*).

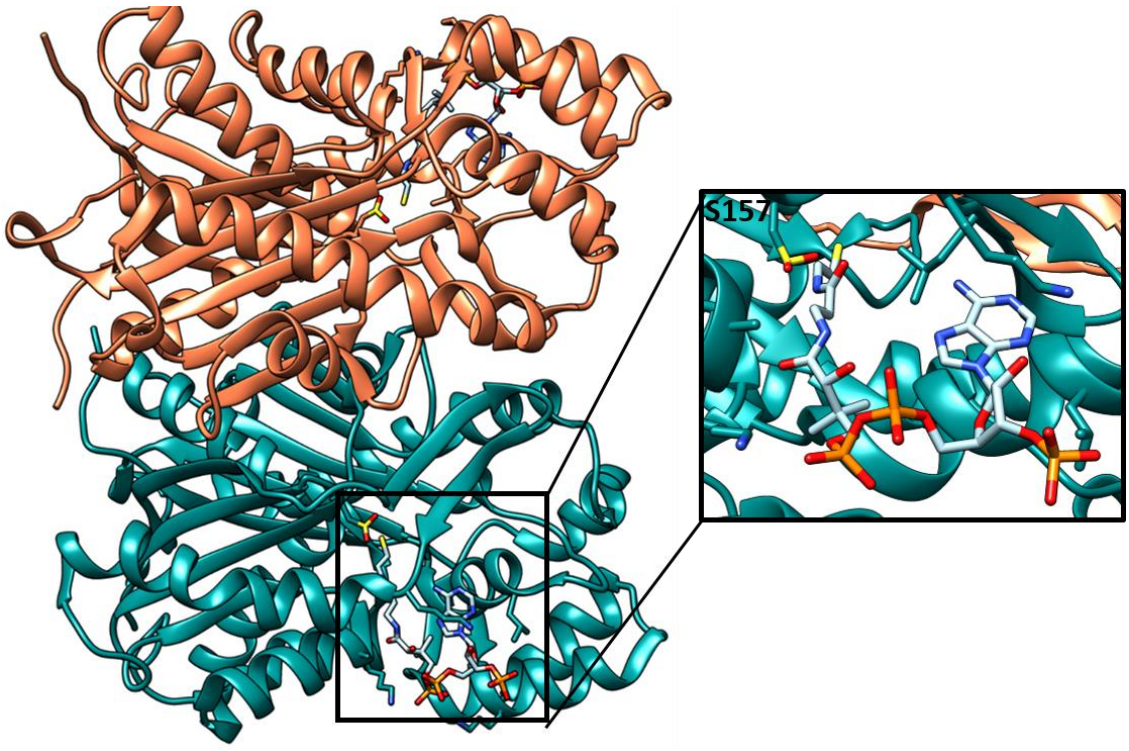


Figure 3: 3D protein structure of *Cs*OLS (PDB-ID: 6GW3) (*33*). Each unit of the protein is colored separately in orange or turquoise. The binding site of CoA (light blue) is depicted highlighting the interaction with S157.

In 2019, the X-ray structure of the *Cs*OLS was first solved. The protein is a homodimer containing two asymmetric units. Hexanoyl-CoA is loaded into the *Cs*OLS and stepwise elongated with three malonyl-CoA units (Figure 3) (*33*). After the elongation with two entities of malonyl-CoA in the first stage, another molecule is added in the second step. This results in forming two intermediate structures, a linear triketide and a linear tetraketide. Next to the release of the linear tetraketide for further reaction towards OA, a hydrolysis reaction can occur on both intermediates. After the first condensation step, the reaction towards pentyl diacetic acid lactone (PDAL) and after the second condensation step, a reaction towards hexanoyltriacetic acid lactone (HTAL) is monitored *in vitro* (*21*). In *C. sativa*, the production of these derailment by-products is not observed (*21*). Since the cyclization step occurs after the

*Cs*OLS reaction, the linear tetraketide has to be released as a product of the *Cs*OLS. In the absence of *Cs*OAC, the linear tetraketide undergoes a non-enzymatic C2→C7 decarboxylative aldol condensation reaction, resulting in olivetol (OL) (Figure 4). This leads to the assumption that the cleavage of the tetraketide intermediate from the CoA moiety must occur first since this cyclization reaction is impossible with the group still attached to the molecule (*33*). The catalytic triad of the *Cs*OLS is composed of Cys157-His297-Asn330 located in the bi-loped cavity responsible for initiation and elongation. The cavity is composed of the starter molecule-binding site, formed by one of the lobes, and a second binding site for the growing polyketide chain. The hexanoyl-CoA's sulfur moiety is close to the Cys157 for reaction, a typical structure element in plant type III PKS (*34*).

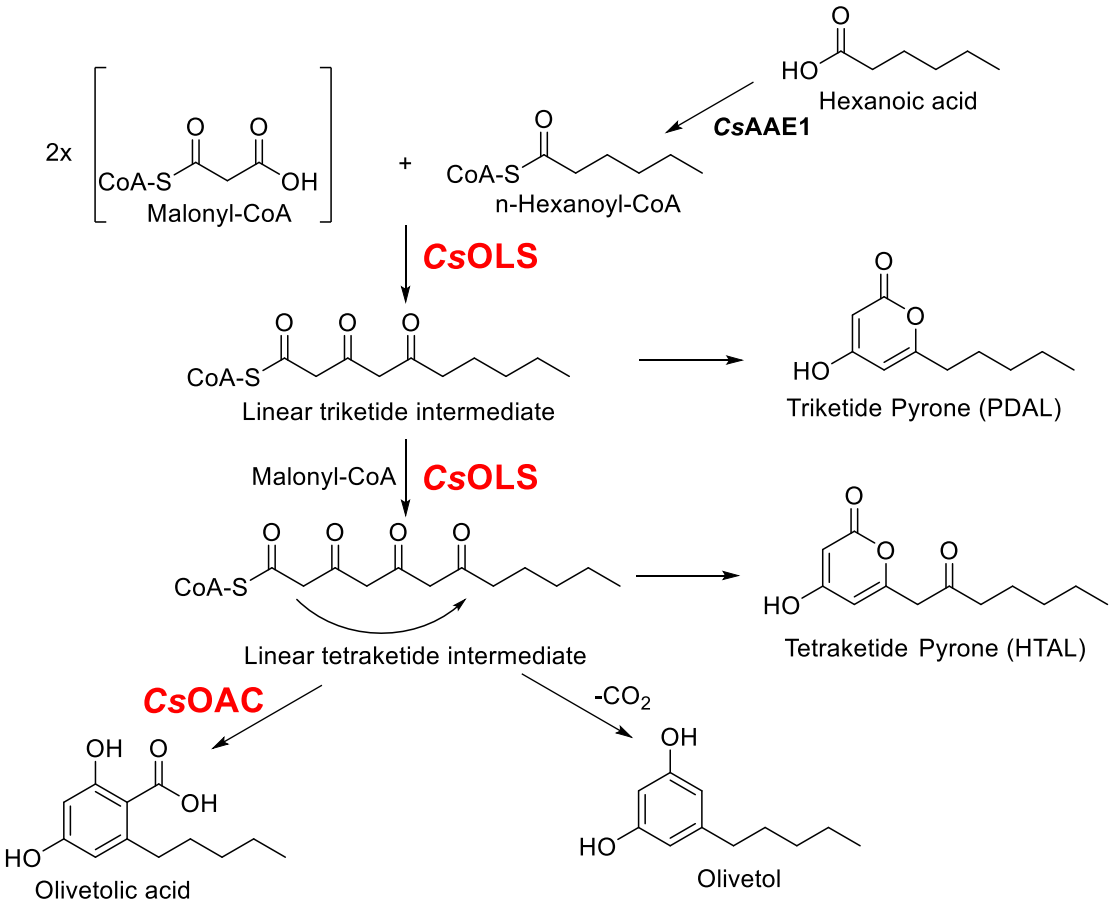


Figure 4: Catalytic reaction of *Cs*OLS and *Cs*OAC including all possible by-reactions. For the formation of the substrate of *Cs*OLS hexanoyl-CoA, the enzyme *Cs*AAE1 needs to catalyze the reaction of hexanoic acid to hexanoyl-CoA. After successive loading of the malonyl-CoA onto the starter molecule hexanoyl-CoA, the first intermediate, a linear triketide, is formed. With the loading of the third malonyl-CoA unit, a linear tetraketide intermediate is formed. The two intermediates can undergo lactonization forming each a derailment by-product with PDAL and HTAL. When *Cs*OAC is present, the linear tetraketide is cyclized, forming olivetolic acid. Next to this reaction, the spontaneous non-enzymatic reaction to Olivetol can occur.

In this enzyme family of type III polyketide synthases, two classes emerged based on the type of cyclization reaction: the Chalcone synthase-(CHS) type or the Stilbene synthase-(STS) type (*31*). The CHS-type catalyzes a C6→C1 intramolecular Claisen cyclization. Therefore, thioester cleavage occurs after cyclization (*35*). In STS-type, the C1 thioester linkage happens before the cyclization step. Therefore, the cyclization occurs in a C2→C7 aldol condensation. An Aldol switch was identified for the thioesterase functionality of the PKS (*36*). *Cs*OLS shows high similarity to both proteins. For CHS the sequence similarity is 66 % and for STS it is 60 % (*33*). As above mentioned the formation of olivetol, leads to a C2→C7 aldol condensation, a STS-like mechanism was expected. Surprisingly the aldol switch was not identified in *Cs*OLS and also the hydrogen bond network of *Cs*OLS resembles more CHS-like proteins. Since *Cs*OLS is not fitting into the known criteria, the role of the adol switch and hydrogen bond network is not as conservative as expected (*36, 33*).

1.2.2 Olivetolic acid cyclase *Cs*OAC

In comparison with other plants, the PKS from *C. sativa* is not capable of producing the needed precursor OA for cannabinoid biosynthesis. Instead, a second protein involved in the formation was identified, the *Cs*OAC (*22*). Despite their close linkage, *Cs*OAC and *Cs*OLS do not have any physical interaction with each other. *Cs*OAC is a small protein with a molecular weight of 12 kDa, comprising two asymmetric units that assemble into a dimeric protein (*37*). The protein's fold, comprising a four-stranded antiparallel β-sheet and two or three α helices per monomeric unit, places it within the family of DABB proteins (Figure 5). Plant DABB proteins are typically induced by stress, but *Cs*OAC is the sole plant DABB protein with a biochemical function (*38*). The binding pocket contains a hydrophobic tunnel known as the pentyl-binding pocket. This tunnel allows the pentyl moiety of the linear tetraketide to bind and interact with the hydrophobic amino acids. In addition to this pocket, another distinctive feature of the protein is the arrangement of the amino acid side chains of His5/Tyr27/Tyr72/His78/Asp96. The crystal structure with OA as a ligand reveals hydrogen bonds between the carboxylic acid group from OA and the side chains of His5 and Tyr72, which enables the hydrogen bond network of the aforementioned amino acid arrangement (Figure 5). This arrangement is of significant importance not only for substrate binding but also for the stability of this protein (*37*).

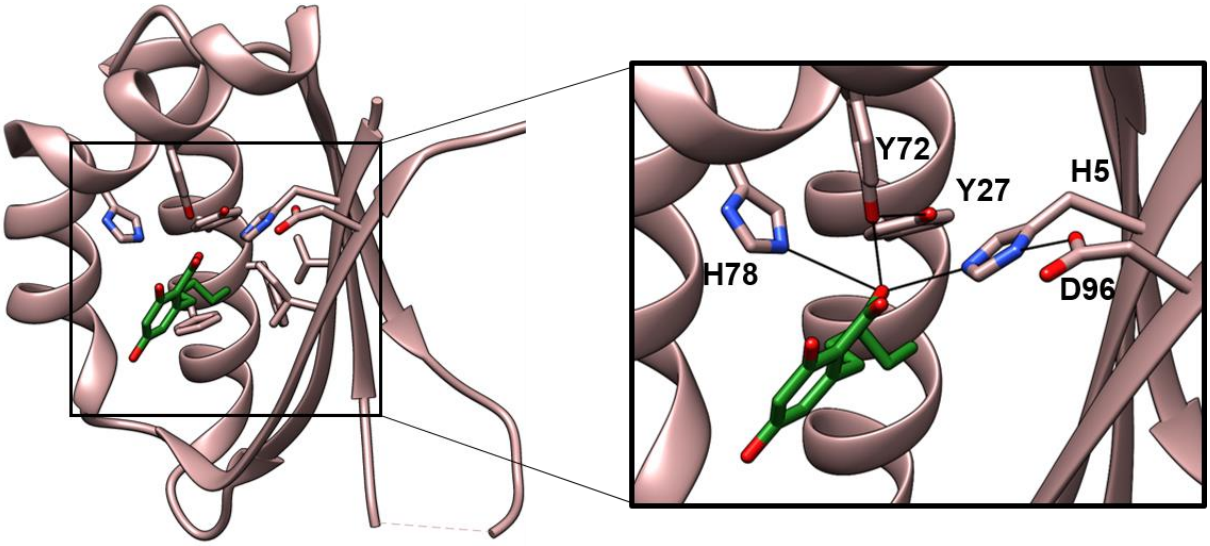


Figure 5: 3D protein structure of *Cs*OAC (PDB-ID:5B09) including the product of the enzymatic reaction OA, colored in green. A close- up of the binding site highlights the hydrogen bond network between the amino acids of the protein and the ligand (*37*).

The X-ray structure enabled the identification of the catalytic mechanism and the involved amino acids (Figure 6). In the first step, after loading the pentyl tetra-β-ketide, Tyr72 activates His78 through a proton abstraction, resulting in a nucleophilic attack of the His78 at the C2 of the linear tetraketide. The intermediate is present as an enolate. The intermediate then proceeds to a nucleophilic attack on C7 through keto/enol tautomerization. To complete the C2-C7 aldol cyclization, the C7 abstracts a proton from His78. Since no catalytic domain for rearomatization and CoA cleavage was identified in the protein, one possibility is that the enzyme releases the cyclized product. The two reactions occur in a non-enzymatic and spontaneous reaction (*37*). As proposed by other sources, it is possible that the CoA cleavage occurred before the enzymatic reaction of the *Cs*OAC and, thus, is not a component of this reaction (*33*).

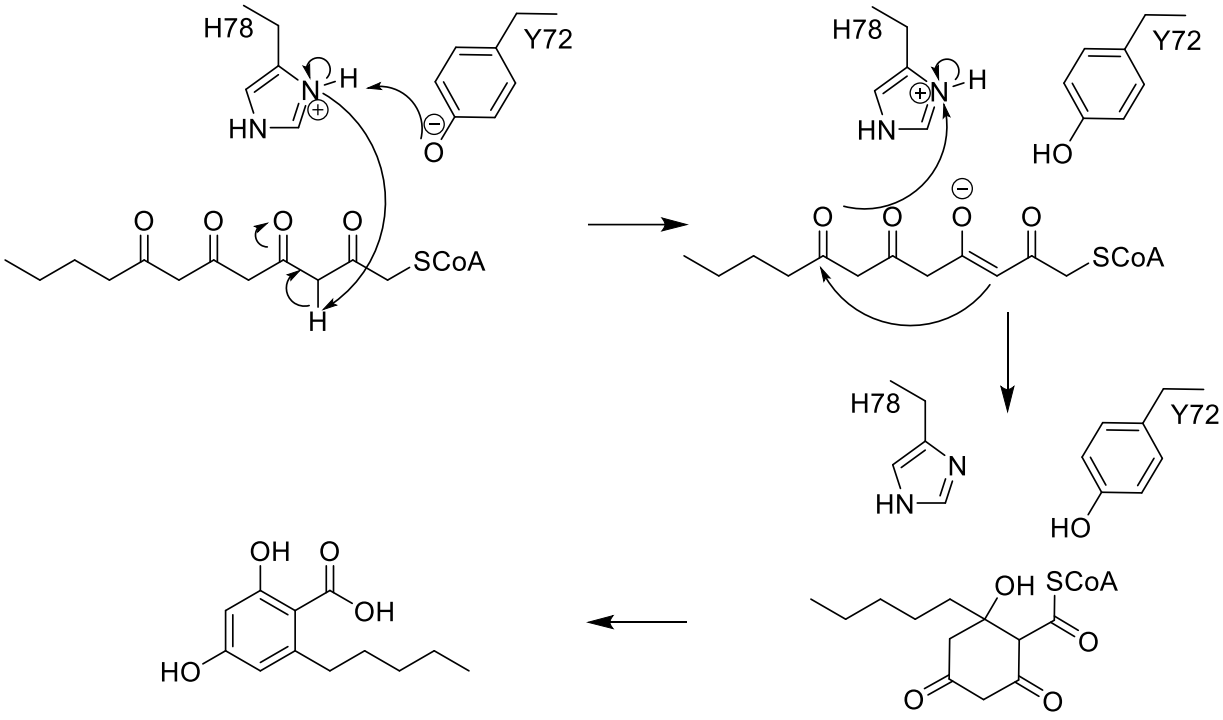


Figure 6: Catalytic mechanism of *Cs*OAC for the formation of olivetolic acid (*37*). First, H78 is activated due to a proton abstraction from Y72 following a nucleophilic attack at the C2 of the substrate. Through keto/enol tautomerization, a nucleophilic attack on C7 leads to a cyclization. In the last step, a proton abstraction from H78 results in rearomatization.

1.2.3 *Cannabis* geranylpyrophosphate:olivetolate geranyltransferase *Cs*PT

Over 25 years ago, the first GOT activity was detected using young *C. sativa* leaves (*24*). Ten years later, the sequence of *Cs*PT1 was published following the latest identification of *Cs*PT4 (*39, 25*). The two proteins share only a 62 % homology with each other, and both proteins are expressed in the *C. sativa* trichomes with a close proximity in gene loci. Consequently, the two proteins may be considered functionally redundant in the plant (*5*). While *Cs*PT4 has demonstrated activity in other organisms, such as *Saccharomyces cerevisiae* or *Nicotiana benthamiana*, *Cs*PT1 has not been observed to catalyze the reaction (*25*). While *Cs*PT1 has a high promiscuity regarding the aromatic substrate, *Cs*PT4 is highly specific for OA. This is also reflected in the kinetic parameters, with *Cs*PT4 exhibiting a Michaelis-Menten constant (K_M) of 6.72 μM, while *Cs*PT1 has a K_M of 60 mM (*40*).

*Cs*PT4 is a transmembrane protein with an N-terminal plastid-targeting sequence catalyzing the C–C Friedel–Craft alkylation of OA, primarily on the C3 position (*41, 42*). The protein belongs to the UbiA superfamily of prenyltransferases (*43*). Family members are predicted to have two characteristics: their three-dimensional structure is predicted to consist of eight to nine transmembrane helices and they have two domains with conserved aspartate-rich motifs, NDxxDxxxD and NQxxDxxxD for the first one and DXXXD for the second one (*44, 45*). The two aspartate-rich motifs coordinate the divalent metal ion, essential for the catalytic reaction. The negatively charged diphosphate interacts with the metal ion, ensuring the substrate can orientate correctly (*44, 46*). Prediction tools were used to assume that *Cs*PT4 consists of eight transmembrane helices (*25, 47*). *Cs*PT4 requires Mg^{2+} for the reaction, and the conserved aspartate-rich motifs were successfully identified with a consensus sequence of NQxxDxxxD and DXXXD. *Cs*PT4 has been shown to accept CBGA-C5 as a substrate, as well as other CBGA derivatives that differ in their alkyl side chain. The absence of available X-ray structures has hindered efforts to elucidate critical aspects such as substrate selectivity (*41*).

1.2.4 Cannabinoid Oxidocyclase *Cs*THCAS, *Cs*CBDAS, *Cs*CBCAS

CBGA-C5 can be accepted from three proteins as a substrate consisting of *Cs*THCAS, *Cs*CBDAS, and *Cs*CBCAS. All three enzymes are similar size, with an approximate molecular weight of 62 kDa. They also exhibit a high degree of similarity in their amino acid sequences, with 92% identity between *Cs*THCAS and *Cs*CBDAS and 84% identity between *Cs*THCAS and *Cs*CBCAS (*48*). These enzymes belong to the family of berbine-bridge like oxidoreductases (*49, 50*). All three proteins require a cofactor: flavin adenine dinucleotide (FAD) (*49, 51*). The X-ray structure of *Cs*THCAS enabled the determination of the FAD and CBGA-C5 binding pocket, and 3D structure elucidation was possible. The FAD binding pocket separates two protein domains. Domain I comprises eight α-helices, eight β-strands, and the covalently bound FAD. Domain II is composed of six α-helices, which are surrounded by eight antiparallel β-strands (*52*). The FAD is covalently bound to the enzyme via a cysteine's sulfur atom and a histidine's nitrogen (Figure 7) (*52, 49*).

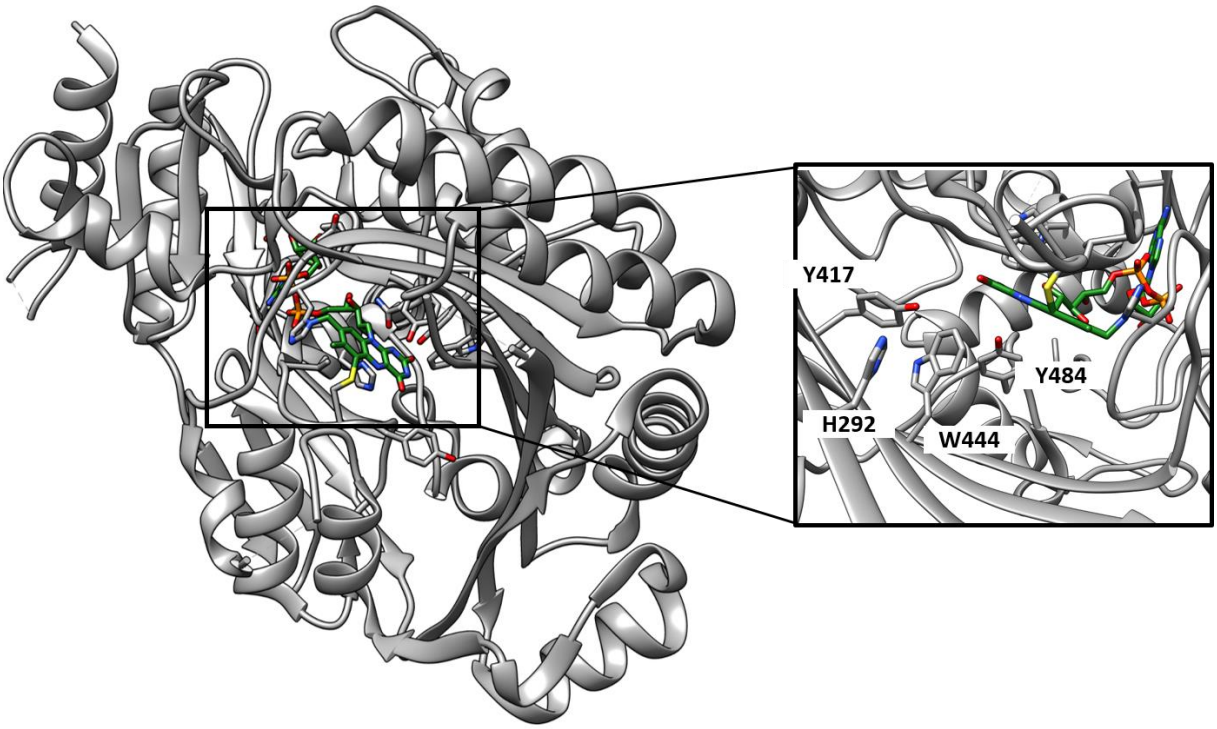


Figure 7: 3D protein structure of *Cs*THCAS (PDB-ID:3VTE) colored in grey, including the covalently bound cofactor FAD colored in green. The interaction between this cofactor and the catalytic amino acid is highlighted in the close-up of the binding site.

The carboxylic acid group of CBGA-C5 must be present for the catalytic reaction to proceed, as CBG-C5 is not converted from any of the three enzymes. The reaction mechanism was proposed using the 3D protein structure, the structural relationship of *Cs*THCAS with related enzymes, and the mutant functional analyses (*52*). In the first step, CBGA-C5 undergoes deprotonation, initiated by the phenolate anion of Y484, followed by the oxidation of the substrate through a hydride transfer from C3 to FAD. This loss results in an intramolecular cyclization within the substrate, resulting in THCA-C5 (Figure 8) (*53, 50, 52*). The key amino acids were successfully identified for the two other proteins, including tyrosine for the start of the catalytic reaction and cysteine and histidine for binding of the FAD (*53, 54*).

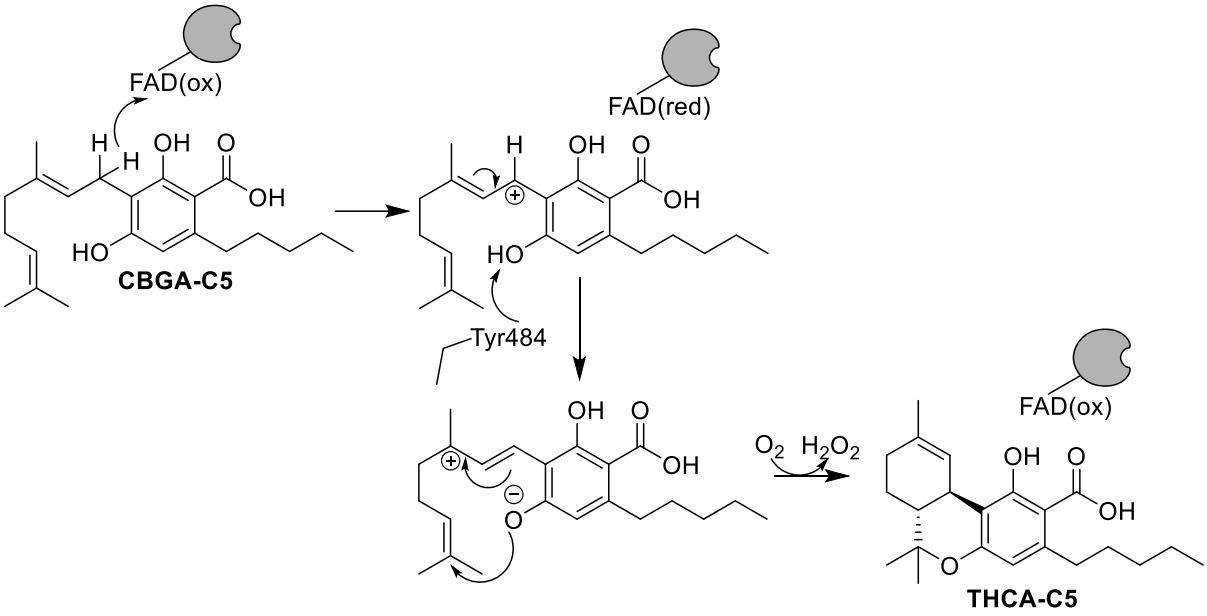


Figure 8: Catalytic mechanism of the THCAS for the reaction of CBGA-C5 towards THCA-C5 including the transition states.

The catalytic reaction of the *Cs*THCAS does not only result in the formation of THCA-C5 but also the formation of CBCA-C5 and CBDA-C5. The product ratio of these three product is strongly depended of the pH value. Increasing the pH values to pH 7 results in an increase of CBCA-C5 production. (*55, 53*)

1.3 Synthetic/unnatural cannabinoids

Cannabinoids produced by plants are natural cannabinoids. Synthetic or unnatural cannabinoids in contrast are specially designed molecules that target the cannabinoid receptors in the human body and can have a higher affinity than natural cannabinoids. This class includes not only artificially created compounds but also modified phytocannabinoids produced through organic chemistry or heterologous production in organisms like *S. cerevisiae*. Their diversity and efficacy offer great potential for developing new therapeutic applications and medications.

1.3.1 Endocannabinoid system

The initial hypothesis regarding the biological effects of cannabinoids was based on their lipophilic structure and the potential for disruption of cell membranes (*56*). Subsequently, the psychoactive effect of cannabinoids was elucidated through their interaction with the cannabinoid receptors (CB) in humans, which were discovered in the 1990s. These receptors comprise two subtypes: CB1, which is predominantly located in the central nervous system (CNS) (*57, 58*), and CB2, which is predominantly located in the peripheral nervous system

(PNS) (*59, 60, 58*). The two receptors belong to the class of G protein-coupled receptor (GPCR) and have a 44 % amino acid similarity with each other (*61–63*). To date, the three-dimensional structure of both receptors has been elucidated, allowing for a more detailed understanding of their mode of action (*64, 65*). This receptor, along with endogenous cannabinoids and the enzymes responsible for their degradation, plays a crucial role in the endocannabinoid system (*66*). In addition to phytocannabinoids, other chemical structures were identified as interacting with these receptors. This expanded the term cannabinoids to include secondary metabolites derived from the plant and endocannabinoids, which naturally interact with the endocannabinoid system (*67*). The majority of these structures are derivatives of amides, esters, and even ethers of long-chain polyunsaturated fatty acids (*68*). The first identified and most extensively biologically analyzed derivatives are those derived from arachidonic acid, anandamide (*69*), and 2-arachidonoyl glycerol (2-AG) (Figure 9) (*70*).

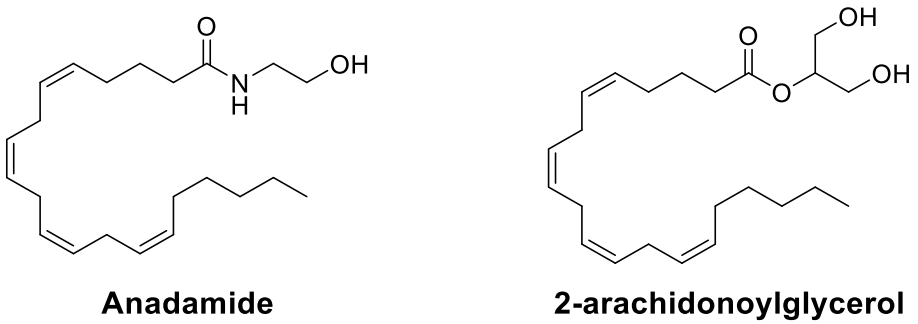


Figure 9: Structure of the first discovered endocannabinoids.

Cannabinoids have a high potential for their application as a drug for several different diseases (*71*). They have shown analgesic and antiemetic properties, causing them to be applied in the palliative care of cancer or Human immunodeficiency virus (HIV) patients for the reduction of nausea and vomiting (*72*). Cannabinoids are also therapeutically administered for neurological disorders such as multiple sclerosis or spinal cord injuries (*73*). For many applications, cannabinoids are still in the clinical phase. The potential for treating different ailments such as inflammatory diseases, cancer, and microbial infections was identified due to the involvement of cannabinoids in the immunity and regulation of apoptotic and angiogenetic signaling pathways (*72*). In the future, applications against the SARS-Covid virus are possible since the inhibitory effects of cannabinoids against the protease were identified (*74, 75*) as also effects of the viral cell entry due to the affinity of cannabinoids for the spike protein (*76*).

1.3.2 Synthetic cannabinoids

With the identification of the cannabinoid receptors and the knowledge about their structural feature, the way has been paved for the exploration of a novel class of synthetic cannabinoids, which are specially designed to interact with the two receptors (*77*). Therefore, the affinity to the CB receptors can be higher than that of THC (*78*). In 2022, up to 237 synthetic cannabinoids were reported. Even though this class of cannabinoids is more diverse compared to phytocannabinoids, structures are constructed of four building blocks. This comprises a core structure to which a tail and a linker within a linked group branch off (*79*). Examples of synthetic cannabinoids are CP47497-C8 (*80*) , representing the group of nonclassical structures, JWH-018 (*81*) representing the group of aminoalkylindoles, AM-4030 (*82*) representing a hybrid structure (Figure 10) (*78*).

Synthetic cannabinoids can also be derivatives of phytocannabinoids. Like the aforementioned structures, they can be formed using organic chemistry (*83, 84*). The first known structure is HU-210, which is a derivative of THC-C5 (*85*). With organic synthesis, a large number of different cannabinoids, consisting of C0, C1, and C3, were synthesized (*84*). Due to heterologous cannabinoid production, another possibility for the formation of novel cannabinoids emerges, where the proteins from cannabinoid biosynthesis are used, but derivatives of the natural structures are applied as substrates (*25*). As one interesting target point for variation, the *n*-pentyl chain was identified as even phytocannabinoids with alteration on this position, such as CBGA-C3, are already known (*41*).

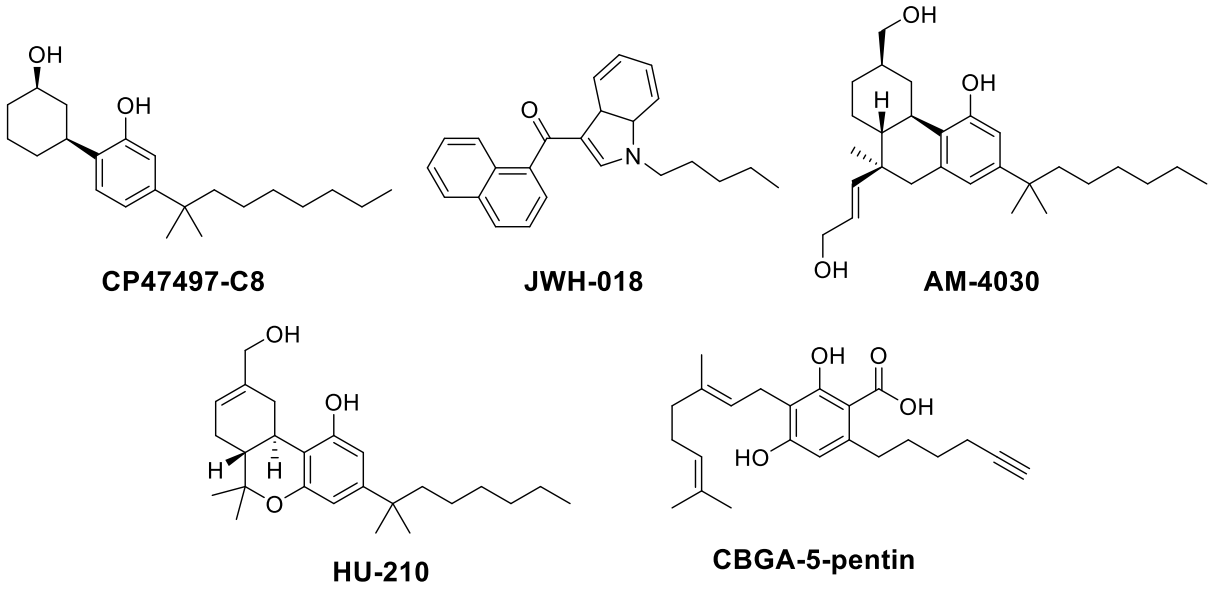


Figure 10: Structures of examples for synthetic cannabinoids.

1.4 Heterologous cannabinoid production

Extracting cannabinoids from the plant *C. sativa* is fraught with challenges due to the regulatory and legislative framework established as a result of the classification of THC-C5 as an illicit drug. The plant's metabolites exhibit pronounced dependence on the geographical and climatic environment (*86*). The chemical synthesis of cannabinoids remains a complex process with many reaction steps, resulting in low yields at a high cost (*87, 84*).

Biotechnological production represents a potential alternative. One advantage of this approach is the production of cannabinoids that are only produced in low amounts in the plant, such as CBCA-C5 (*88, 55*). In the preceding year, a diverse range of hosts was examined, identifying several challenges. The first host to undergo a complete synthesis of cannabinoids, including THCA-C5, THCA-C3, and CBDA-C5, was *Saccharomyces cerevisiae* with a THCA-C5 production of 8 mg L^{-1} (*25*). This organism benefits from the natural presence of the mevalonate pathway, which provides GPP for cannabinoid biosynthesis (*89*). Moreover, the engineering of the production of other precursors, such as the acetyl-CoA pool, can be readily accomplished in *S. cerevisiae*. Additionally, *S. cerevisiae* offers the advantage of posttranslational modification and plant protein expression and is a well-established host organism, which has resulted in the development of numerous genetic tools, including strains, vectors, and promoters, suitable for cannabinoid biosynthesis (*87*). To this date, a CBGA-C5 production of 510 mg L-1 has been achieved through metabolic engineering of *S. cerevisiae*, which only requires glucose and hexanoate feeding (*90*). Using the biosynthesis in *S. cerevisiae*, producing novel cannabinoid structures was also possible while using another precursor, such as hexanoic acid. Using pentanoic acid, the chain was shortened towards C4, using heptanoic acid, the chain was elongated towards C6, and using 6-heptynoic acid, a new feature with the triple bond was added, resulting in 5-pentin (*25*).

Another yeast that has shown promising results is *Komagataella phaffii*, in which the protein *Cs*THCAS for the final reaction from CBGA-C5 to THCA-C5 has been expressed. The use of a crude cell extract (CCE) enabled the identification of THCA-C5 as product, with the addition of CBGA as substrate (*91*).

The polyketide starter OA was also synthesized in the yeast *Yarrowia lipolytica* from glucose with a titer of 9.18 mg L^{-1}, following the engineering of the organism (*92*). In addition to fungal models, the implementation of cannabinoid biosynthesis was also explored in the most commonly utilized model plant, *Nicotiana benthamiana* (*93*). In this instance, the transformation

of infiltrated hexanoic acid towards OA and the transformation of administered OA towards CBGA-C5 were accomplished using the plant as host. However, a transformation towards THCA-C5 has not been achieved to this date. However, this outcome may be attributed to *N. benthamiana's* insufficient production of OA and CBGA-C5. Additionally, it has been observed that the plant's glycosyltransferases accept both molecules as substrates, producing glycosylated OA and CBGA-C5. The two molecules are removed from the cannabinoid biosynthesis pathway (*94*) leading to a low substrate concentration for the following reaction. Additionally, *Escherichia coli* was employed as a biosynthetic host for the production of the starter polyketide OA. This was achieved using hexanoate and a single carbon source, glycerol, with a resulting titer of 80 mg L^{-1} (*95*). However, further steps in cannabinoid biosynthesis were not implementable in *E. coli*. Based on these findings, it can be posited that *S. cerevisiae* is the most promising host organism for cannabinoid synthesis (*96*).

1.4.1 The aromatic prenyltransferase NphB

The natural prenyltransferase involved in cannabinoid biosynthesis is a transmembrane protein. It presents a significant challenge in transforming it into a heterologous host organism, such as *S. cerevisiae* or *E. coli* (*55*). The utilization of a soluble aromatic prenyltransferase would be a preferable approach. Two prenylating enzymes that have been previously employed in cannabinoid biosynthesis include the enzymes AtaPT, isolated from *Aspergillus terreus (97, 98),* and NphB isolated from *Streptomyces* sp. strain CL190 (*99, 100*).

NphB is a promiscuous aromatic prenyltransferase first identified in 2005 and is part of the ABBA family of aromatic prenyltransferases (*101, 102*). The protein's fold motif is an α/β barrel, also referred to as a PT-barrel, comprising ten antiparallel β-strands encased in ten α-helices. The fold forms a spacious, solvent-accessible binding pocket, wherein both the prenyl acceptor and donor can bind (Figure 11) (*101*). *In vitro* assays demonstrated Mg^{2+}-dependency despite the absence of a trinuclear magnesium cluster, a common feature among Mg^{2+}- dependent proteins (*103*).

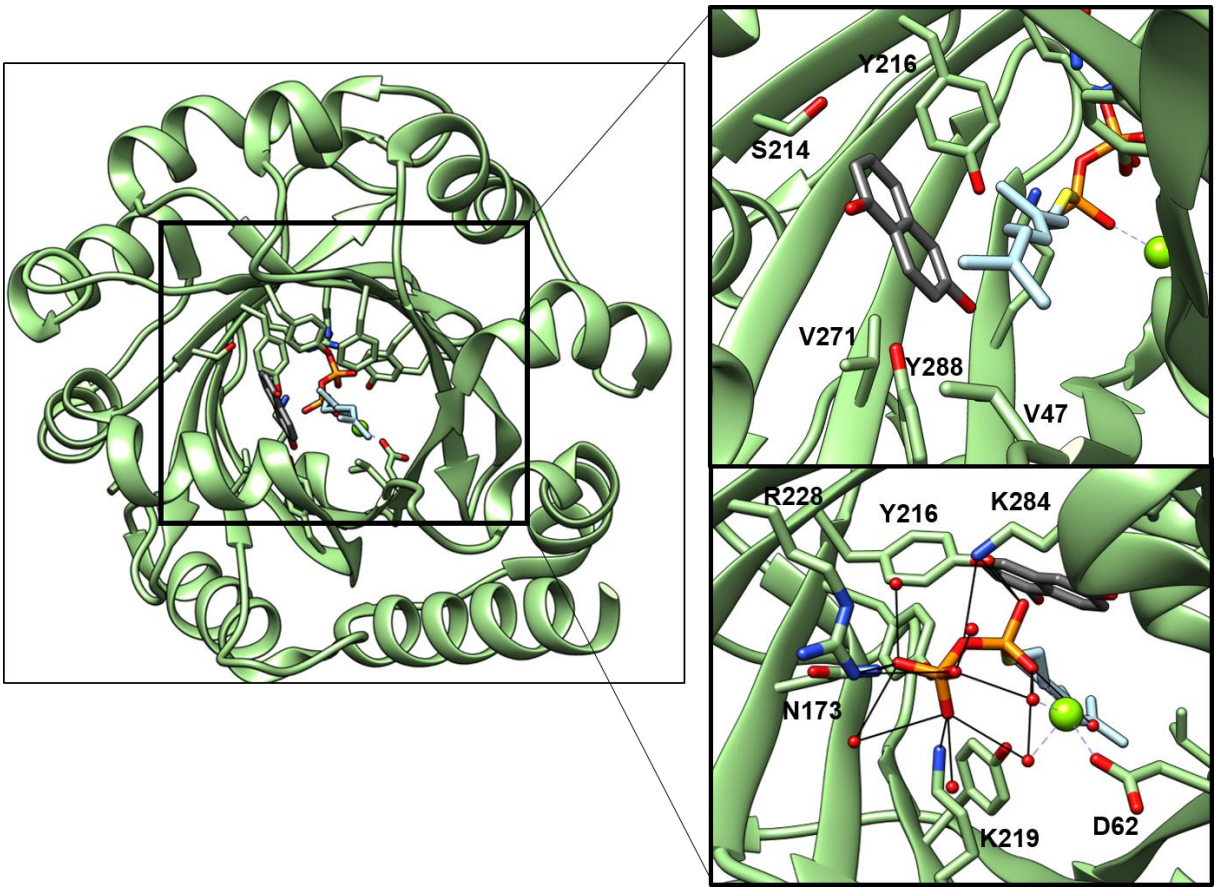


Figure 11: 3D protein structure of NphB, including the two substrates GSPP colored in blue and 1,6-DHN colored in grey (PDB-ID:1zb6). In a close-up of the interaction of the two substrates with the amino acid side chains of NphB, the Mg^{2+} ion and water molecules are presented. Hydrogen bonds are deposited as black lines.

In its original organism, NphB is involved in the biosynthesis of naphterine, where the enzyme catalyzes the transfer of a 10-carbon geranyl group to the aromatic substrate 1,3,6,8-tetrahydroxynaphthalene at the C4 position (*104*). A diverse range of aromatic substrates were identified, including naphthols, flavonoids, sulfonamides, and plant polyketides such as OL or OA (*99, 105*). The primary donor molecule is GPP; thus, other structures, including farnesyl diphosphate, dimethylallyl diphosphate, and 1-methylcyclohex-1-ene diphosphate, were also identified. A correlation was discovered between the acceptor and donor molecules. A change in the acceptor molecule results in a shift in the spectrum of accepted donors. Using 1,6-dihydroxynaphthalene (1,6-DHN) as an aromatic acceptor has led to the identification of longer-chain diphosphates as potential donor structures. Using sulfabenzamide as an aromatic acceptor indicates that shorter-chain diphosphates represent the primary second substrate for NphB (*105*).

It has been observed that NphB C prenylation occurs in ortho- or para-position to a hydroxyl group, as well as O-prenylation at hydroxyl groups (*101, 99*) Recently, an N-alkylation has also

been identified (*105*). The promiscuity of the prenylation positions results in regioselective product formation for some aromatic acceptors. For example, three distinct products were identified for 1,6-DHN (*101, 106*). With regard to OA, NphB is able to catalyze the same reaction towards CBGA-C5 using the prenyl moiety from GPP for an attachment on the C3 position, similar to *Cs*PT4. But using the wild type, the main product of the reaction is not CBGA-C5 but 2-*O*-geranyl-olivetolic acid (2-*O*-GOA), in which the prenylation occurred on the O2 oxygen (Figure 12). The ratio between the two products is 85:15 for 2-*O*-GOA. Prior to the successful application of NphB in cannabinoid biosynthesis, further engineering toward the product regioselectivity is necessary (*100*).

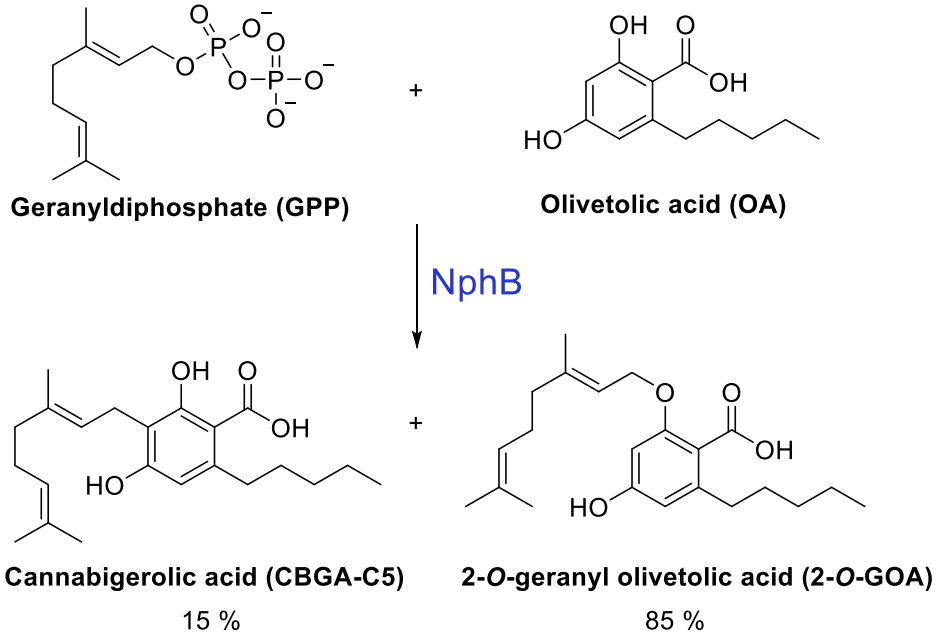


Figure 12: Schematic reaction equation of the NphB catalyzed reaction of GPP and OA towards CBGA-C5 and 2-*O*-GOA including the product ratio (*100*).

The reaction mechanism follows an S_N1 dissociative mechanism similar to the Friedel-Crafts Alkylation (*107*). In this reaction type, the C1-O1 bond is cleaved first, resulting in the formation of a carbocation and the separation of the pyrophosphate. Interactions with Mg^{2+} can facilitate the stabilization of the pyrophosphate. A transitional state is generated following a nucleophilic attack from the geranyl cation on the aromatic acceptor molecule. The transfer of a proton to the pyrophosphate results in the neutralization of the product (Figure 13) (*106*).

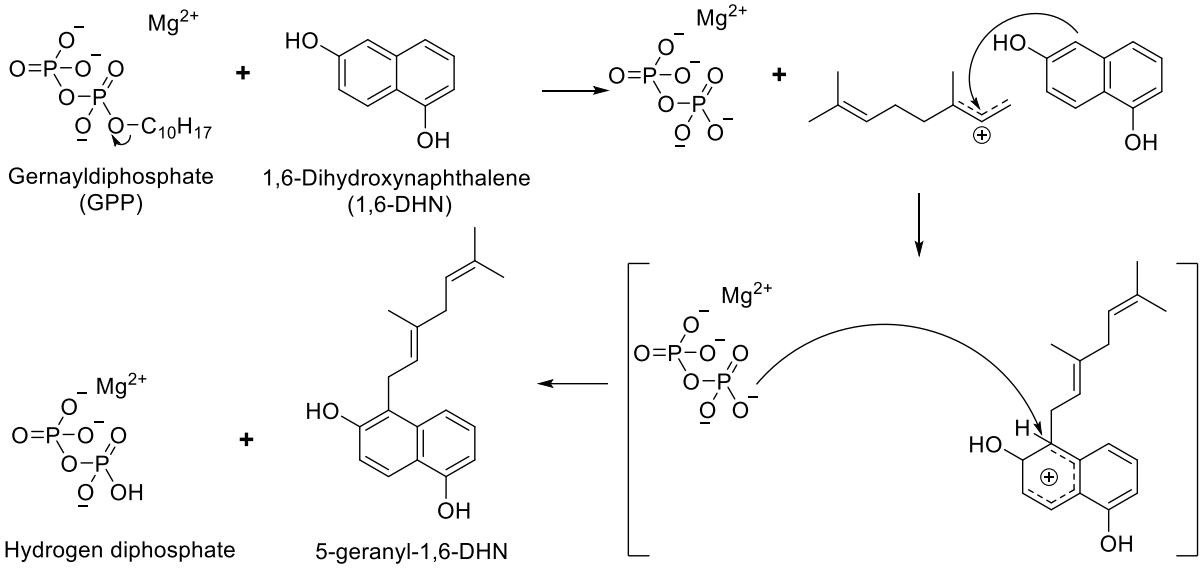


Figure 13: Reaction mechanism of NphB using the example of the reaction with GPP and 1,6-DHN (*106*).

The negative charge of the pyrophosphate is stabilized within the binding pocket through a complex network of interactions. These include the formation of hydrogen bonds with the positively charged amino acid side chains Lys119, Asn173, Arg228, and Lys284, as well as the involvement of a positively charged Mg^{2+} ion and a hydrogen bridge with Tyr216. Subsequently, during the catalytic reaction, the carbon cation is stabilized through cation-π interactions, forming a π-chamber with Tyr 121, Tyr216, and the aromatic acceptor. To date, two X-ray structures have been determined, each containing a different aromatic acceptor substrate: 1,6-DHN and flaviolin. Notably, the two substrates occupy disparate positions within the binding pocket. The high product regioselectivity of 1,6-DHN is illustrated by its high B factor (65.68 Å), which indicates that no single, clearly defined conformation can be identified. This evidence supports the hypothesis that the orientation of the substrate at the beginning of the catalytic process influences the product selectivity of NphB (*106*). The stabilization of the aromatic substrates was monitored through interactions with Met162, Phe213, Ser214, Tyr288, and Gln295, as well as the interaction with the prenyl donor (*101, 106, 108*). The availability of a X-ray structure with the prenyl donor only allows the conclusion that the prenyl donor binds first (*101*).

1.4.2 Protein engineering in cannabinoid biosynthesis

To this date, many examples of successful protein engineering applications to proteins for cannabinoid biosynthesis are known. The goal is to enhance activity for implementing the biosynthesis in another organism for proteins such as *Cs*OAC or NphB or for producing novel cannabinoids, which are analogs of products from biosynthesis.

1.4.2.1 *Cs*OAC and *Cs*OLS engineering

The first variants of the *Cs*OAC were not created to enhance the OA production but to identify the catalytic amino acid in a time before the 3D structure of the protein was solved. Nevertheless during this process, the variant Y27F was identified, which showed *in vitro* 162 % activity compared to the wild type (*37*). To this date, this variant is often used for cannabinoid biosynthesis in other organisms.

Even though it was proven that using *Cs*OLS and *Cs*OAC, both enzymes are able to accept other carboxylic acids as substrates, a limitation towards higher than 6'-hexynyl moiety and bulkier acids than hexanoic acid was visible (*25*). Dual engineering of *Cs*OLS and *Cs*OAC was performed to accept longer chain or bulkier fatty acids. The key amino acids for the pentyl binding side of *Cs*OLS is L190, limiting the size of the accepted substrate. For *Cs*OAC, the amino acids F24 and V59 were identified as size-defining factors for the pentyl binding pocket. With an exchange of amino acid L190 of *Cs*OLS towards glycine and an exchange of amino acid F24 of *Cs*OAC towards isoleucine or leucine, the acceptance of dodecanoic acid towards an OA analog containing a C11 chain instead of a C5 chain was possible (*109*).

1.4.3 NphB engineering

Before the protein NphB can be successfully applied in cannabinoid biosynthesis, two bottlenecks of using this protein need to be solved. The first one is the product specificity of the protein, in which CBGA-C5 is not the main product of the catalytic reaction and the overall low kinetic values (*100*). 2018, the first results were achieved using docking studies to identify promising positions. Since this date, no 3D structure containing OA as a substrate could be elucidated. The effect of chosen amino acid exchanges was tested *in vitro* regarding the CBGA-C5 and 2-*O*-GOA production. While the changes on positions 126, 161, 162, 175, or 213 led to no effect or enhanced 2-*O*-GOA, promising results were achieved with changes on position 295. While an exchange to leucine only led to an enhancement *in vitro* and not *in vivo* using *S. cerevisiae*, the exchange Q295F led to a 20 times higher CBGA-C5 production in

S. cerevisiae (*110, 111*). At the same time, 2-*O*-GOA decreased, which resulted in a shift in the ratio of 20:1 towards CBGA-C5. In 2019, Valliere et al. engineered NphB to achieve a better outcome in their cell-free cannabinoid production. The redesigning dockings were performed as a starting point, and the software Rosetta was used to predict mutations that would improve stabilization of the binding mode of OA, leading to CBGA-C5 formation. After *in vitro* screening of different NphB variants, the two variants, G286S/Y288A and A232S/Y288V, improved both activity and selectivity. CBGA-C5 is produced almost exclusively and the k_{cat} is 1000-fold higher regarding CBGA-C5 production than using the wild type (Wt) (*112*). Parallel to this, Qian et al. also identified the potential for exchanging G286 with another amino acid due to their docking results. The variant G286S led to an improvement in activity and selectivity. Nevertheless, the results were not as advantageous as the ones exceeded by Valliere et al. (*113, 112*). The last known attempt in protein engineering of NphB was performed 2022 from Lim et al.. In this study the variants V49W/Y288P and V49W/Y288P/Q295F were identified using docking and MD-simulations. The variants V49W/Y288P and V49W/Y288P/Q295F were identified using docking and MD simulations. The two variants produced CBGA-C5 exclusively *in vitro* and approximately 13-fold more than the wt, and they also produced more CBGA-C5 than the previously mentioned variants (*114*).

1.4.4 *Cs*THCAS engineering

Since the 3D structure of *Cs*THCAS is solved, protein engineering, including a rational design approach, is easier than for the three other cyclooxidoreductases. For *Cs*THCAS, the correct folding of the protein structure is challenging but crucial. One element involved are the glycosylation sites, consisting of seven amino acids overall (*53*). When these glycosylation were removed, the overall activity was 1.2-fold higher (*115*). The lack of glycosylation sites was analyzed in a mutational study. While some led to a decrease in THCA production, the best combination of N89Q and N499Q led to a 2x-fold increased activity (*53*). For further stabilization of the BBE domain of the protein and the correct positioning of the catalytic amino acid Y484, a further disulfide bridge was introduced with the mutations H494C and R532C. This additional disulfide bridge led to a 1.7-fold increase in activity and a shift in the temperature optimum from 57 °C to 52 °C (*53*).

1.4.5 *Cs*CBDAS engineering

Simultaneously with the engineering of the *Cs*THCAS, protein engineering was performed on the *Cs*CBDAS. The effect of the removal from the glycosylation sites, which had a beneficial effect on the effectivity of the *Cs*THCAS, is not similar for the *Cs*CBDAS. Here, the activity decreased with an exchange of N498Q. To transform the product specificity of the enzyme, amino acids from the binding side were exchanged to mimic the binding site of CsTHCAS. Two amino acid exchanges positively affected the protein activity with S116A and A414V, resulting in a 2.8-fold increase and a 3.3-fold increase in catalytic activity, respectively (*25, 53*).

1.4.6 *Cs*CBCAS engineering

The *Cs*CBCAS has, overall, from all three oxidoreductases from cannabinoids, the significantly lowest catalytic capacity, which results in an unfavorable low conversion of CBGA-C5 at high substrate concentrations. For further application in heterologous cannabinoid production, an increase in turnover number is needed (*116, 54*). Using a homology model and an alpha fold structure, a rational design was performed using the catalytic center next to computational predictions using SSN consensus analysis. This study led to the discovery of the variant C224W, which produces 22-fold more CBCA-C5 than the wild type (*54*).

1.5 Scope of thesis

To date, the production of all cannabinoids, ranging from novel structures that are promising for future medicinal applications to common cannabinoids like CBD-C5, is trending in various application areas. Even with the first success in transforming cannabinoid biosynthesis in a microbial host with the most potential in *S. cerevisiae*, challenges remain (*25, 55*). The overall titer of the produced cannabinoids must be increased to make the heterologous cannabinoid synthesis economic and sustainable. Several bottlenecks are identified. A problem arises with proteins that lack efficiency or form byproducts, which unnecessarily waste the fed carbon sources and reduce product concentration. The potential for novel cannabinoid structures in therapeutic applications is evident. However, it needs to be ensured that these molecules can be produced sustainably without high-cost synthesis. A solution for this can also be heterologous cannabinoid production, where the exchange of the starter molecule allows for the easy production of novel cannabinoid structures with the potential for sustainable production. Nevertheless, a bottleneck appears here with the proteins, for which the novel substrates are not naturally constructed. Here, protein engineering can also be applied to discover the substrate specificity of the proteins and the potential of producing novel cannabinoids while using cannabinoid biosynthesis (*41*).

In the first part of the thesis, the potential in the protein *Cs*OAC regarding OA production is explored further, with the primary goal of decreasing OL production. With the reduction in OL production, resources for cannabinoid production can be saved, increasing the sustainability of cannabinoid biosynthesis in *S. cerevisiae*. The effect of fusing the two proteins is further examined to decrease the physical distance between the two proteins. With computer-based methods, protein variants of CsOAC are identified, which can increase the protein's activity. This effect also increases the OA concentration while the OL concentration decreases. For the later application as protein in the cannabinoid biosynthesis in S. cerevisiae, all variants are screened in vivo regarding their OA and OL production.

Using a microbial host can benefit the formation of novel cannabinoids in a high titer. To produce the common cannabinoids, such as THCA-C5, hexanoic acid is added to the cultivation media. This thesis analyses the change of the fed carboxylic acid and its potential formed products. The substrate acceptance of all three *Cs*AAE1, *Cs*OLS, and *Cs*OAC is analyzed with a focus on forming novel OA analog structures. With a previously developed

OA-producing *S. cerevisiae* strain, screenings are performed to identify the potential for exploring novel cannabinoid structures.

The last part of this thesis analyzes the conversion of OA analogs using NphB. A library of potential OA analogs is identified with docking experiments, which are likely to be converted towards CBGA analogs using NphB. After organic synthesis of the most promising and diverse library of potential substrates, the molecules are tested *in vitro* using two NphB variants. The structure of the highest-produced CBGA analogs is determined in detail, and the possible formation of THCA or CBCA analogs is analyzed using an in vitro reaction containing *Cs*THCAS.

Chapter 2 Results and Discussion

Parts of chapter 2.1 were published in

Spitzer, S.; Aras, M.; Kayser, O. Improving CsOAC Activity in Saccharomyces cerevisiae for Directed Production of Olivetolic Acid through Rational Design. *Chembiochem : a European journal of chemical biology* **2024**, e202400651.

Parts of chapter 2.3 were published in

Spitzer, S.; Wloka, J.; Pietruszka, J.; Kayser, O. Generation of Cannabigerolic Acid Derivatives and Their Precursors by Using the Promiscuity of the Aromatic Prenyltransferase NphB. *Chembiochem : a European journal of chemical biology* **2023,** *24* (22), e202300441.

This chapter is divided into three sub-chapter according to the scope of this thesis

2.1 Increasing olivetolic acid production through protein engineering

2.2 Screening of different carboxylic acids regarding olivetolic acid derivatives production

2.3 Generation of cannabigerolic acid derivatives using NphB

2.1 Increasing olivetolic acid production through protein engineering

For the production of cannabinoids, OA is a natural precursor, serving as the polyketide backbone for all major cannabinoids. When the cannabinoid biosynthesis is transferred to a microbiological host, such as *S. cerevisiae,* the presence of high amounts of OL as a product is observed in addition to OA. OL is synthesized through a chemical reaction of the linear tetraketide, which is released as the catalytic product from *Cs*OLS and is not able to be converted further from the downstream proteins of the cannabinoid synthesis. This chemical reaction is avoided when the linear tetraketide binds as a substrate of *Cs*OAC and is subsequently converted to OA. Hence, a bottleneck of cannabinoid biosynthesis occurs when the linear tetraketide has a high residence time in the cytosol. The formation of OL results in the wastage of significant precursors, including acetyl-CoA, malonyl-CoA, hexanoyl-CoA, and glucose, which are subsequently unavailable for further cannabinoid biosynthesis. Moreover, OL accumulation has the potential to impact the *S. cerevisiae* cell itself, thereby reducing the achievable cannabinoid titer. The elevated residence time may be attributable to inadequate *Cs*OAC concentrations, resulting from insufficient protein expression or deficient protein activity. To overcome this problem, the copy-number of *Cs*OAC was increased (*55*). The increase in the number of proteins led enhancement in the OA to OL ratio, although residual OL was still detected.

2.1.1 Linking *Cs*OLS and *Cs*OAC to increase OA concentration

One potential strategy for reducing the chemical reaction of the linear tetraketide towards OL is to create a fusion protein by connecting the two proteins, *Cs*OAC and *Cs*OLS. The distance between the two proteins is diminished, with one unit of each protein situated adjacently, facilitating substrate transfer from one protein to the other. Consequently, the time the linear tetraketide is present within the cytosol can be reduced (*117*). To establish a connection between two protein domains, amino acid sequences that differ in sequence and length are inserted between the two protein sequences (*118, 119*). In the studies conducted by Luo et al. the two proteins are linked using AATSGSTGSTGSTGSGRSTGSTGSTGSGRSH as the sequence, with the C-termini of *Cs*OLS and the N-termini of *Cs*OAC connected (*120*). The aforementioned fusion protein is then employed for cannabinoid production in *S. cerevisiae* (*25*). In contrast to the aforementioned findings, Ma et al. reported that the fusion of the two proteins in the same direction did not result in any improvement in OA production in

Y. lipolytica (*92*). Accordingly, the present study was conducted with the objective of identifying the optimal connection and sequence length of the linker. Two distinct linker sequences were selected for evaluation. For a flexible linker, a protein sequence consisting of glycine and serine GGGGS was employed. The inclusion of glycine, the most flexible amino acid, results in enhanced flexibility of the linker, which facilitates the movement of the two proteins (*118, 121*). The amino acid sequence EAAAK was selected as a rigid linker for testing purposes. This structure resembles an α-helix due to the salt bridge between the amino acids, which results in a distinct distance and a rigid element between the two proteins. The use of a rigid linker serves to minimize the potential for alteration in the interaction between the proteins (*118, 122, 123*).

In the initial phase of the procedure, the arrangement of the two proteins was conducted following the protocol established by Luo et al. whereby the C-terminus of *Cs*OLS was linked to the N-terminus of *Cs*OAC (*25*). The efficacy of the GGGGS and EAAAK linker sequences was evaluated through the testing of various repetition values, including n = 1, n = 3, and n = 5. To facilitate the screening of the various fusion proteins, the *S. cerevisiae* strain ySP01 was developed. In this strain, the gene for *Cs*AAE1 expression was integrated into the *YPRCτ3* locus of the strain *S. cerevisiae* Δ*pep4* Δ*gal1* Δ*gal80*. This was done to achieve sufficient transformation of HA to hexanoyl-CoA and ensure the supply of the substrate for *Cs*OLS. The strain ySP01 was transformed with each of the plasmids pDio-OLS-$(GGGGS)_n$-OAC and pDio-OLS-$(EAAAK)_n$-OAC. The results were then compared to the results obtained using plasmid pDio-pCCW12-OLS-pTDH3-OAC, in which the two protein sequences have their own promoter and terminator (Figure 14).

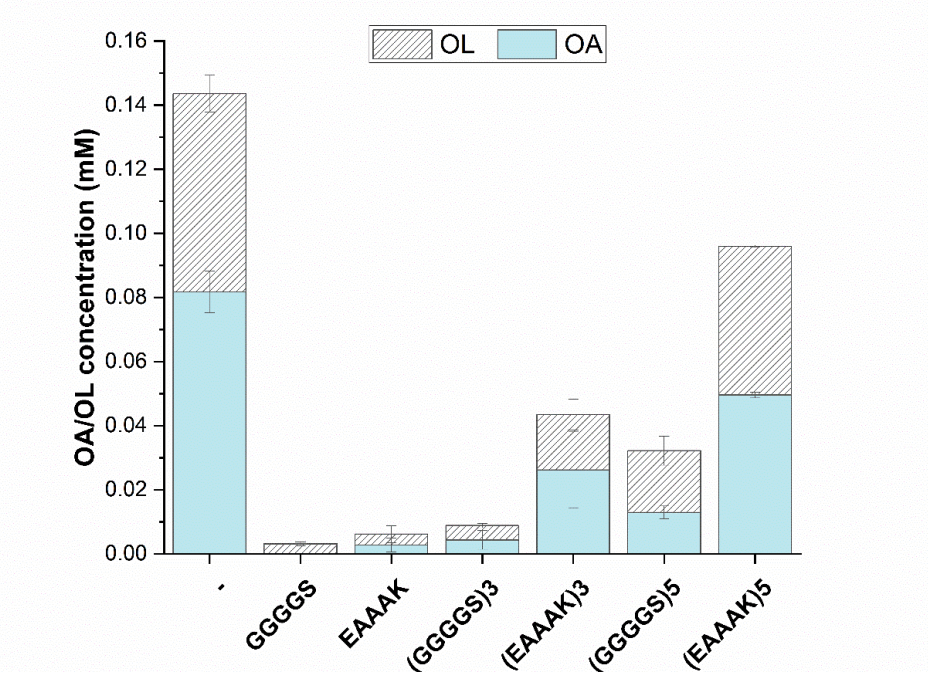


Figure 14: Screening of different linker variations and lengths between *Cs*OLS and *Cs*OAC. Strain ySP01 was transformed with the plasmids pDio-OLS-X-OAC. For X, the different linker sequences were inserted. After 48 h cultivation in YPD + 0.5 mM hexanoic acid, the OA and OL concentrations were determined using HPLC-UV detection at 215 nm. Cultivations were performed in triplicates. For comparison, the plasmid pDio-pCCW12-OLS-pTDH3-OAC were used, where both proteins have their own promotor and terminator.

The first step was to examine the OA concentration as a result of the linker screening. A review of the results revealed two distinct trends. The initial observation was that using the rigid linker sequence EAAAK consistently resulted in a higher product concentration than the flexible linker sequence GGGGS when an equivalent number of repetitions of the sequence were used. Repeating the linker sequence three times, using the fusion protein *Cs*OLS-$(GGGGS)_3$-*Cs*OAC, yielded an OA concentration of 0.004 ± 0.003 mM. In contrast, using the fusion protein *Cs*OLS-$(EAAAK)_3$-*Cs*OAC exhibited an OA concentration of 0.026 ± 0.012 mM, representing a 6.5-fold increase in OA production compared to the other linker. Integrating five repetitions of the linker sequences between the two proteins resulted in a 3.5-fold increase in OA concentration when the fusion protein *Cs*OLS-$(EAAAK)_5$-*Cs*OAC was used in the experiments.

Secondly, it was observed that the shorter the linker sequence, the greater the decrease in product concentration. The linker sequence EAAAK was identified as the more suitable option, demonstrating a notable increase in product concentration with an increase in the number of repetitions when compared with each other. Specifically, a single repetition yielded an OA concentration of 0.0027 ± 0.0022 mM, while five repetitions resulted in a concentration

of 0.0495 ± 0.0009 mM, representing an 18.3-fold increase in concentration. In conclusion, the fusion protein *Cs*OLS-(EAAAK)$_5$-*Cs*OAC yielded the highest OA concentration of all tested linker. However, this value remains significantly lower compared to the OA production, where both proteins were expressed independently, with a concentration of 0.0818 ± 0.0065 mM. Furthermore, the ratio between OA and OL remains unaffected, indicating that introducing a fusion protein does not enhance cannabinoid production in *S. cerevisiae*. Given the substantial influence of linker size and sequence on OA and OL concentration, future studies could investigate a broader range of variations, particularly in the context of extended and more rigid linkers.

The second step involved examining the appropriate arrangement of the proteins, specifically whether the linker should be C-terminal to *Cs*OLS and N-terminal to *Cs*OAC, or *vice versa*, i.e., whether the linker should be C-terminal to *Cs*OAC and N-terminal to *Cs*OLS. Both combinations were tested using the GGGGS and EAAAK linkers, with five repetitions each, as it was hypothesized that a longer linker and fewer interactions between the components would benefit OA production. The ySP01 strain was transformed with each of the plasmids pDio-OAC-(GGGGS)$_5$-OLS, pDio-OAC-(EAAAK)$_5$-OLS, pDio-OLS-(GGGGS)$_5$-OAC and pDio-OLS-(EAAAK)$_5$-OAC (Figure 15).

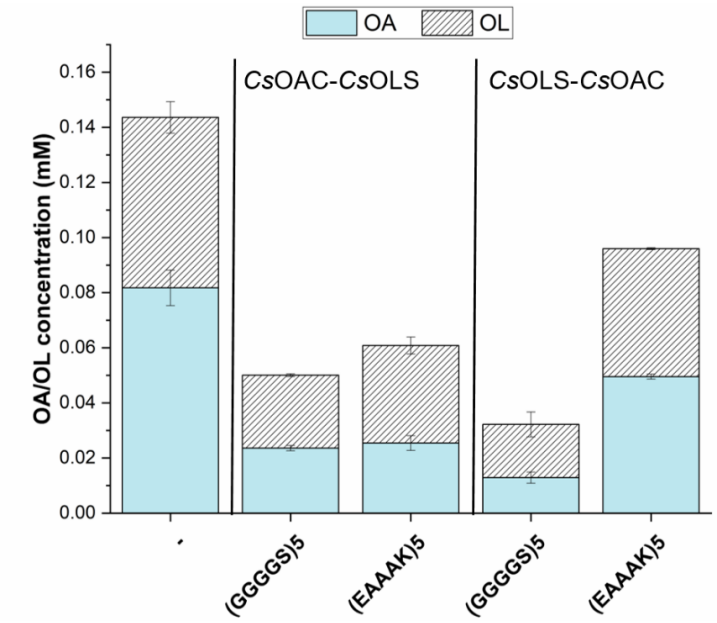


Figure 15: Screening of the linker (EAAAK)$_5$ and (GGGGS)$_5$ between *Cs*OLS and *Cs*OAC with different arrangements of the two proteins. Strain ySP01 was transformed with the plasmid pDio-OLS-X-OAC. For X, the different linker sequences were inserted. After 48 h cultivation in YPD + 0.5 mM hexanoic acid, the OA and OL concentrations were determined using HPLC-UV detection at 215 nm. Cultivations were performed in triplicates. For comparison, the plasmid pDio-pCCW12-OLS-pTDH3-OAC were used, where both proteins have their own promotor and terminator.

In the case of the fusion protein *Cs*OAC-*Cs*OLS, the difference between the two linker types was not as pronounced as in the other cases, with an OA production of approximately 0.02 mM for both linker variants. Consequently, the fusion protein that demonstrated the most optimal performance was *Cs*OLS-(EAAAK)$_5$-*Cs*OAC. The arrangement of the fusion protein, with the connection of the C-termini of *Cs*OLS and the N-termini of *Cs*OAC and a rigid, longer linker sequence, is the most promising one and should be further investigated in future studies.

2.1.2 Binding mode analysis of *Cs*OAC to identify interesting positions for rational design

Enhancing the enzyme activity of *Cs*OAC has been demonstrated to reduce the possibility for the linear tetraketide to react towards OL. A strategy that has been shown to achieve this objective is to target residues in direct contact with the substrate or in proximity to the active site cavity. This approach has been shown to modify the substrate-binding affinity (*124, 125*).

In order to identify potential amino acids for modification, it is first necessary to solve the binding mode of the linear tetraketide intermediate with *Cs*OAC. Binding mode identification was performed using docking experiments, which was an obligatory stage in the procedure given that the available crystal structures encompassed the product OA, but did not include the unstable substrate (*37, 109*). The exact point at which the CoA moiety is cleaved has yet to be determined. Regarding *Cs*OAC, the catalytic domain responsible for the thiolesterase reaction has remained unidentified. Consequently, the cleavage of CoA might occur in two distinct temporal frames: prior to or after the *Cs*OAC catalytic reaction (*33, 37*). In order to facilitate the calculation and results of the docking process, the intermediate was used in the absence of the CoA moiety. A carboxylic acid was designated as the cleavage product (*126*). The protein's overall conformation remains unaltered when OA is bound within the binding pocket; however, the positioning of the amino acid side chains within the binding pocket does change. Accordingly, the holoprotein structure (PDB ID: 5B09) of *Cs*OAC was selected for usage in the docking experiments (*37*).

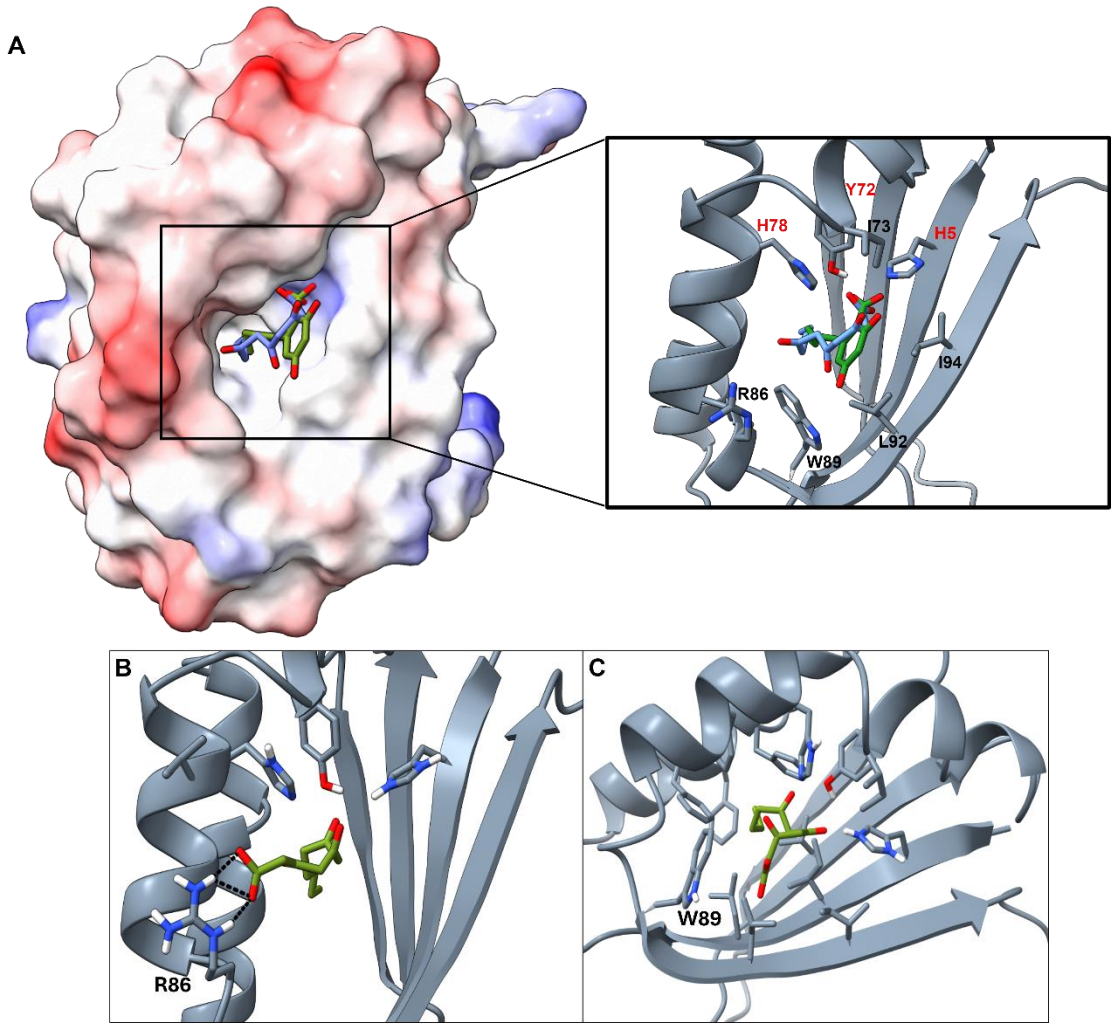


Figure 16: "(A) Depiction of the protein *Cs*OAC, including the X-ray structure 5B09 (grey) with OA bound (green) and the proposed binding mode (blue) from the docking experiments. The catalytic amino acids are highlighted in red, and the hotspot amino acids are highlighted in black. (B) Highest ranked pose from the docking with CsOAC and the linear tetraketide highlighting interactions, which hinder the orientation towards the catalytic amino acids. Including the interaction between R86A (PDB-ID: 5B09) and the carboxylic acid group of the substrate and (C) the interaction between W89A (PDB-ID: 5B09) and the carboxylic acid group of the substrate, if the arginine on position 86 is exchanged to alanine (*37*)." (*127*)

In order to elucidate the required reaction mechanism, it is necessary to ensure that the carboxylic acid group is oriented towards Y72 and H78, as these amino acids are known to initiate the catalytic reaction. However, the initial protein-ligand docking results did not yield the aforementioned position due to a hydrogen bond interaction between the ligand and the side chains of R86 (Figure 16B). To generate a plausible binding mode, alanine was first introduced at position 86. This, in turn, gave rise to an alternative unfavorable binding mode characterized by the interaction of the carboxylic acid group with W89. A plausible binding mode was discovered when both amino acid side chains were exchanged with alanine (Figure 16A). "The interaction between the substrate and amino acids R86 and W89 may hinder

the start of the catalytic reaction and lower enzyme activity. These interactions could also occur *in vitro*, causing issues with substrate positioning in the correct binding mode. However, the examined interactions with amino acids R86 and W89 may also be related to the docking settings. The binding pocket definition during the docking process could be too large, or the cut-off of the CoA moiety results in possible interactions that are not typically feasible due to later cleavage of the CoA (*128*). Nevertheless, these two positions are the first interesting positions for the rational design" (*127*). In the proposed binding mode, the pentyl chain adopts an orientation analogous to that of OA in the X-ray structure, while the carboxylic acid moiety exhibits an interaction and orientation directed towards the catalytic amino acids Y72 and H78.

The newly proposed binding mode has led to the identification of additional hotspot amino acids. These amino acids possess the potential to enhance protein activity through interactions with the substrate or transition state. Additionally, these amino acids may also influence product release. Residues G82, L92, and I94 were identified as promising candidates due to their close proximity to the substrate and the absence of prior examination. (Figure 16A).

"For the second approach, we chose to use an enzyme-engineering-focused web service with the objective of predicting target sites of interest. The HotSpot Wizard 3.1 server was identified as the optimal choice, as it is capable of automating the prediction of functional hot spots (*129*). As an input file, we used the same PDB file (PDB-ID:5B09) as for the docking. The server automates the prediction process by calculating 'mutability' scores for each position from a multiple sequence alignment (MSA) and combining the gathered information to determine the probability of finding 'functional hot spots'. Eight functional hotspots were calculated for *Cs*OAC (Figure S3). After visualizing and evaluating these positions, it was determined that the two amino acids at positions 9 and 23 are located in the pentyl-binding pocket. Therefore, they were excluded from further analysis. The before-mentioned amino acids at positions 82, 86, 89, and 92 were also identified as mutational hotspots. The amino acid at position 94 was classified as having moderate mutability. In addition to these amino acids, the amino acid on position 73 was additionally classified as a mutational hotspot. This amino acid is adjacent to the catalytic ones, which could lead to problems when the amino acid is exchanged but could also have beneficial effects on the protein activity. Therefore, the position is included in our analysis" (*127*).

2.1.3 Generation and evaluation of *S. cerevisiae* strain for *Cs*OAC variant screening

In order to test different protein variants, it is crucial to establish a suitable system to address the specific issues at hand. In pursuing the objective of enhancing OA concentration in *S. cerevisiae*, an *in vivo* screening system was selected as the validation method for the protein variants. A primary benefit of directly testing variants in yeast is that it allows for the avoidance of variants that appear to increase OA production *in vitro* but yield different results when implemented in yeast due to insufficient expression and a lower enzyme concentration can be avoided (*130*). However, it should be noted that the exclusive use of *in vivo* assays precludes the possibility of distinguishing between the effects of each variant on enzyme activity and concentration. Nevertheless, this does not conflict with the findings of our studies (*54*). Regardless of the underlying reason, the highest OA titer and lowest OL titer are desired for implementing the biosynthesis of cannabinoids into *S. cerevisiae*.

A rapid and reliable screening system is crucial for the preliminary characterization of protein variants. To perform a preliminary screening, it was decided that the protein variants should be integrated into a plasmid, with the best-performing variant subsequently selected for genomic integration. The intermediate should be produced at the highest possible titer to ensure an adequate supply of *Cs*OAC substrate and address its instability. *Cs*OLS was integrated multiple times into *S. cerevisiae*, resulting in the creation of strain ySP02 (Table 9). An integration vector construct from the EasyCloneMulti vector set was employed, targeting Ty4 LTR sites within the yeast genome (*131*). Following a screening process, the clone with the highest OL titer was selected for further experimentation. The subsequent step involved the integration of the enzyme *Cs*AAE1 (Table S1) at the YPRCτ3 locus. This was done to achieve sufficient transformation of HA to Hexanoyl-CoA and increase the titer of the linear tetraketide. This integration resulted in the creation of strain ySP03, which exhibited an increase in OL. The enzyme *Cs*AAE1 was identified as a key enzyme in the natural biosynthesis of cannabinoids in *C. sativa* and was found to have the highest conversion rate compared to other enzymes (*55, 20*).

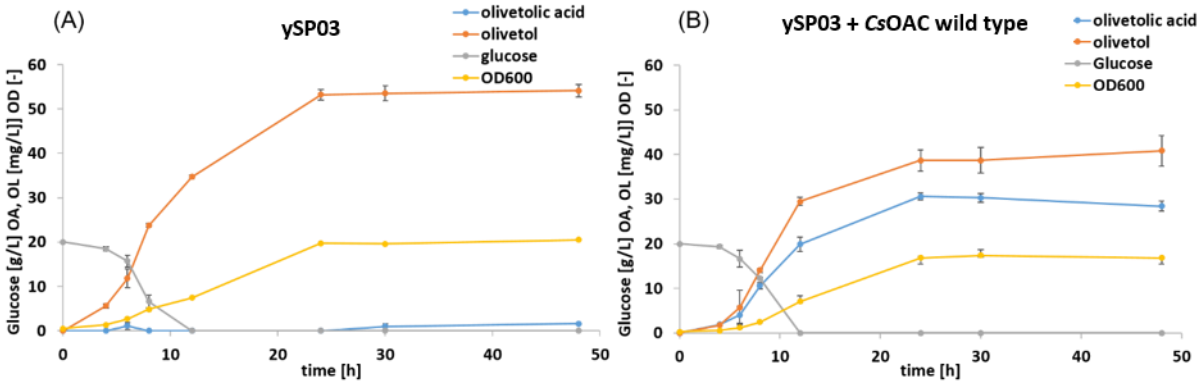


Figure 17: "Olivetol (OL) and Olivetolic acid (OA) production over 48 h for the engineered *Saccharomyces cerevisiae* screening strain ySP03 monitoring of the OL, OA and glucose production and the OD_{600}. (A) Testing strain ySP03. (B) Testing strain ySP03 after transformation with the plasmid pDio-OAC containing the wild type" (*127*).

To further validate strain ySP03, it was cultivated for 48 h in YPD with the addition of 0.5 mM hexanoic acid. In the absence of *Cs*OAC, a titer of 54 mg/L OL can be achieved within a 48-hour period (Figure 17). This titer allows us to conclude that a sufficient production of the linear tetraketide has been achieved. This result is consistent with other published studies where the highest possible production of OL was achieved using the same integration sites (*55*). For the plasmid-based integration, the vector pDio-OAC was used (Figure S1), where the *oac* gene is flanked by the pGAL1 promoter and the CYC1 terminator. Subsequent to the transformation of the plasmid containing the native *Cs*OAC into ySP03, the production of OL and OA was monitored for 48 h. The production of both molecules was observed to occur from the onset of the cultivation. After 4 h, OL production exhibited a more significant increase than OA production. Glucose is completely consumed after 12 hours in both experiments. Upon entering the stationary phase, the production of both molecules also stagnates. During the final hours, a slight decline in OA is observed, which can be attributed to a gradual cleavage of the carboxylic group, leading to a modest increase in OL concentration over time. Overall, a sufficient supply of the unstable intermediate is ensured, and the substrate for OA as a limiting factor in the experimental setup can be excluded (Figure 17).

2.1.4 Alanine Scan

"Though not directly involved in the catalytic mechanism, after the current state of the mechanism, it is first essential to ensure that mutating the chosen six positions does not lead to a complete loss of activity. Therefore, an alanine scan was performed. In these, the amino acids on positions 73, 82, 86, 89, 92, and 94 were exchanged towards alanine (Figure 18).

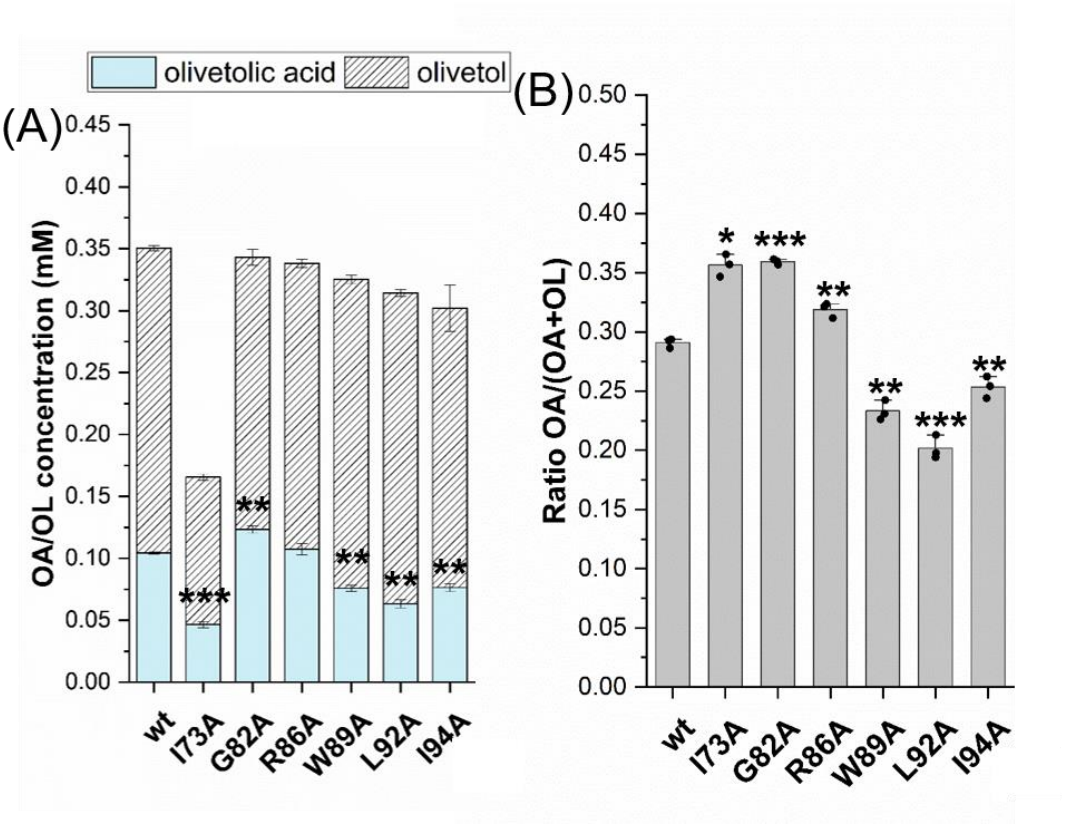


Figure 18: Alanine Screening of *Cs*OAC variants on positions 73, 82, 86, 89, 92, and 94. Variants were transformed into the strain ySP02 with the plasmid pDio-OAC carrying the mutated oac genes each. After 48 h cultivation in YPD + 0.5 mM hexanoic acid, the OA and OL concentration were determined using HPLC-UV detection at 215 nm. For comparison, the production of the wild type was analyzed and the cultivation was performed in triplicates. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***).Detailed statistical reports are shown in the Supplements (Table S9 & S10).

To conduct the analysis, the *S. cerevisiae* strain ySP03 was used for *Cs*OLS and *Cs*AAE1 expression, and the plasmid pDio-OAC, which included the sequence for *Cs*OAC expression, was utilized. After transforming the cells with the corresponding plasmids, the resulting cells were incubated for 48 h at 30 °C and 200 rpm. Following extraction, the concentration of OA and OL in the cell suspension was determined using HPLC-UV (Figure 18A)" (*127*). The production of OA by the *Cs*OAC wild type was set to 100%. A decrease in OA production of up to approximately 50% was observed for the variants I73A, W89A, L92A, and I94A. However, as this decrease in activity does not indicate a complete loss of activity, the amino acids in these positions can be exchanged without disrupting the catalytic reaction. The search for beneficial exchanges in these positions can be continued. In contrast, variants R86A and G82A exhibited an increase in OA production of 13% and 30%, respectively. A comprehensive investigation of all six positions is warranted.

2.1.5 Single exchange screening of *Cs*OAC variants

2.1.5.1 *In silico* determination of first round of variants to test

The search for beneficial amino acid exchanges was approached through two distinct methodologies. The frequency of occurrence for all six positions 73, 82, 86, 89, 92, and 94 was analyzed using HotSpot Wizard (*129*). Based on these results, one amino acid exchange for each position was selected for screening *in vivo*. (Table 1).

Table 1: Single amino acid exchanges of the protein *Cs*OAC tested regarding their OA and OL production in *S. cerevisiae* and their origin.

Origin	Variant
Wild type reference	Wild type
Alanine Scan	I73A, R86A, W89A, L92A, I94A
Alanine Scan /Hotspot Wizard	G82A
HotSpot Wizard	I73V, I73M, I73F, R86L, W89L, L92V, I94V
Docking	I73L
	G82E, G82H, G82I, G82Y
	R86F, R86H, R86S
	W89H, W89N, W89T
	L92F, L92R, L92W
	I94C, I94L, I94Q
***In vivo* assay + Docking**	G82S, G82V
	L92H, L92I, L92Q, L92Y
	I94D, I94S

For position 82, the exchange towards alanine was calculated. This exchange had previously been included in the screening due to an alanine scan performed beforehand; therefore, no amino acid was tested from the HotSpot Wizard experiments.

The second approach in the rational design involved protein-ligand docking, which was performed with all possible single exchanges for each of the six positions to all 19 other natural amino acids of *Cs*OAC and the linear tetraketide using Gold and the ASP scoring function. Initially, the data set was reduced by excluding all *Cs*OAC variants for which the correct binding mode was not calculated within all created poses. Following this, the highest docking score of

a plausible binding mode was identified for each *Cs*OAC variant. A heat map was created to obtain an overview of the changes within the scoring values (Figure 19).

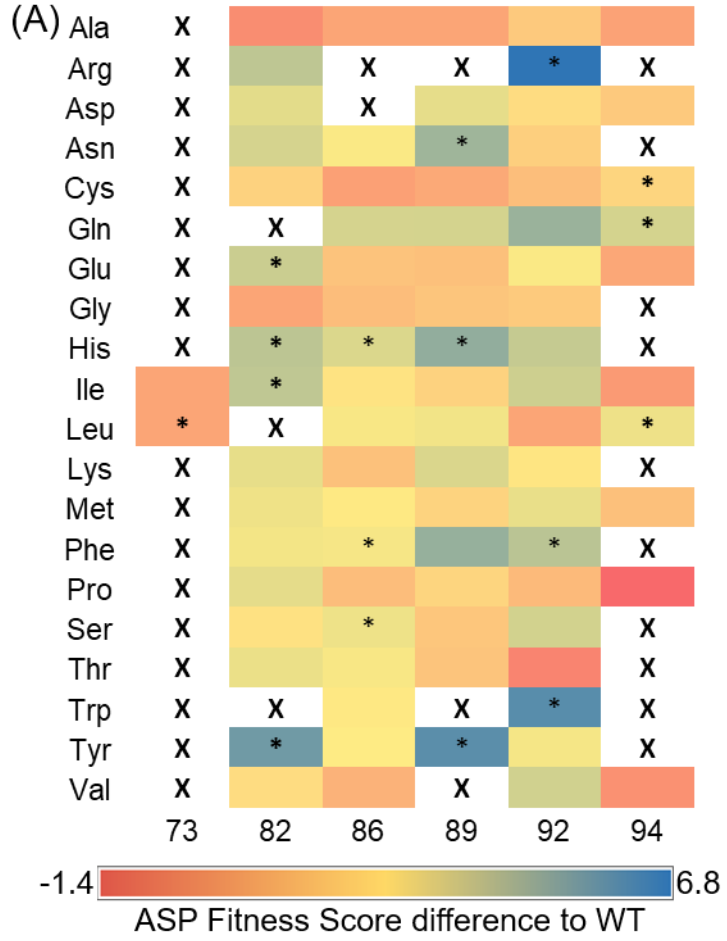


Figure 19: "Heatmap including the comparison of the docking results with *Cs*OAC (PDB-ID:5B09) and 3,5,7 trioxododecanoyl using the docking software Gold and ASP Fitness Score.(*132, 133*) For the position 73, 82, 86, 89, 92 and 94 the amino acid were exchanged to all 19 natural amino acids and the scores were compared according to the fitness score of the wild type. The X marks incorrect binding modes and the * mark the variants with the best docking scores." (*127*).

For positions 82, 86, 89, and 92, most of the variants led to a binding mode comparable to the wild type for the linear tetraketide. Within these variants, the three best-scored variants with the highest shift were chosen for further *in vivo* analysis. "For proteins with an exchange on position 73, most dockings resulted in a pose where the carboxylic acid group is not correctly orientated due to interactions with the newly introduced side chain (Figure 20). The amino acid 73 is located next to the catalytic amino acid, which explains the easy disruption of the carboxylic group's orientation. Therefore, the only amino acid with a plausible binding mode was tested, and two more suggestions from HotSpot Wizard were added for *in vivo* screening.

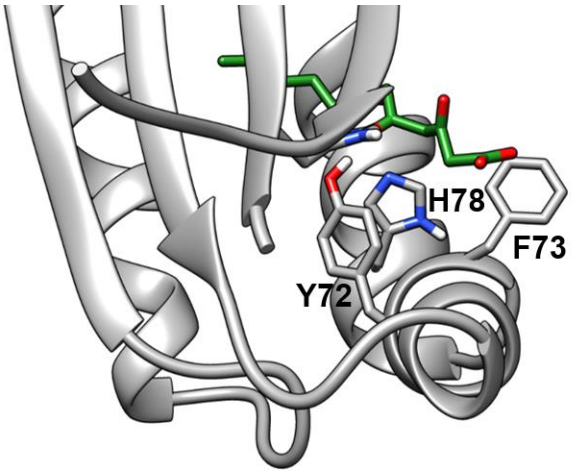


Figure 20: "Highest ranked pose from the docking with *Cs*OAC I73F and the linear tetraketide highlighting interactions, which hinder the orientation towards the catalytic amino acids." (*127*).

The substitutions on position 94 also does not result in the desired binding modes, particularly for amino acids with a high number of atoms or a hydrophilic side chain. This could be explained by the shortened distance to the substrate and potential interfering interactions." (*127*). Nevertheless, the three *Cs*OAC variants leading to the highest scored docking results were also tested for position 94 (Table 1).

2.1.5.2 Plasmid-based screening of first round of variants

"Following the *in silico* analysis, the amount of OA and OL produced from *Cs*OAC variants including the amino acid exchanges was analyzed. Experiments were conducted in a method analogous to that employed in the alanine scan. Two factors were examined: the concentration of the products and the OA content of the overall product formation (Figure 21). The objective is to identify *Cs*OAC variants that result in a significant enhancement or a not significant alteration in OA production in comparison to the *Cs*OAC wild type. It is anticipated that these amino acid exchanges will not have a negative effect on the protein; however, they may prove beneficial if combined with other amino acid exchanges at a later stage in the testing of multiple amino acid exchanges. If a variant has a significant decrease in OA production, it will not be included into further analysis. Also, a highly significant improvement on the ratio of the formed products, without decreasing the OA concentration significant are from interest.

Some variants have a negative impact on OL production, as well as on OA production, resulting in a significant increase in the ratio between the two products. Normally, the two proteins do not interact with each other. Nevertheless, an *in vivo* testing system can be highly sensitive. A change in protein expression, for example, can lead to misfolded protein formation and, consequently, disruption of the cytosol environment. Protein production can be

downregulated to match intracellular demands. Additionally, cell growth can be significantly affected, which also impacts the overall product content (*134*)." (*127*).

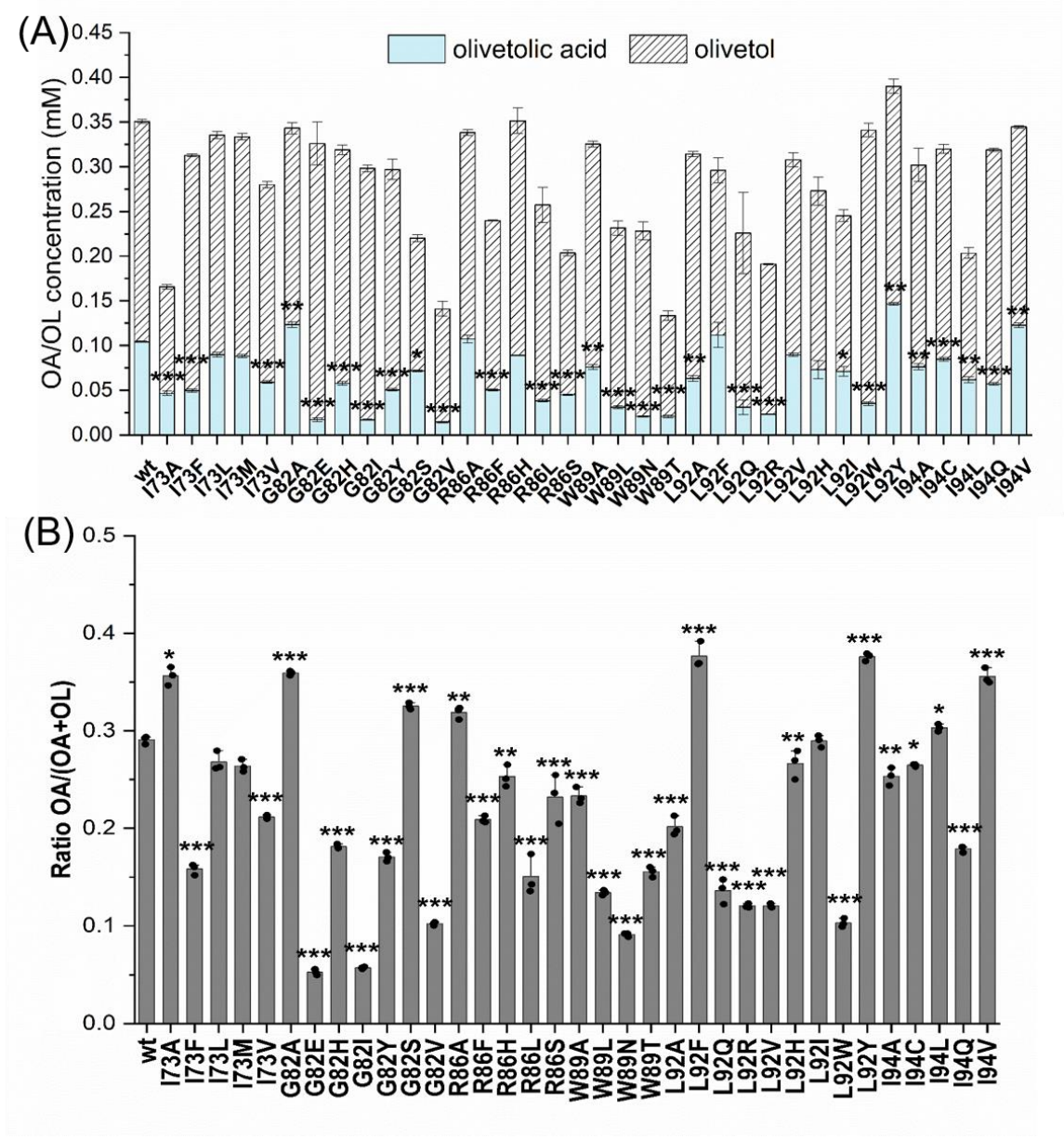


Figure 21: Olivetol (OL) and olivetolic acid (OA) production of all *Cs*OAC variants including a single amino acid exchange. For screening the different *Cs*OAC variants were cloned to pDio-OAC and *S. cerevisiae* strain ySP03 was transformed with the plasmid for each variant. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the Supplements (Table S11 & S12) (*127*).

"Exchanges at positions 73 and 89 did not positively affect OA production (Figure 21). Therefore, further investigations on these positions were postponed, and the focus of this research shifted to the other four positions. Linked to the docking results, exchanges on

position 73 often led to a reaction hindering binding mode. The proximity to the catalytic center and, therefore, the influence on the catalytic process could be too great, limiting changes in the protein environment. Exchanges on position 89, mainly to shorter amino acids such as alanine, result in an increase in distance towards the substrate to approximately 6 Å, which is too far for a considerable interaction (*135*)." (*127*). Although the docking process resulted in an unfavorable binding mode for position 89, this position is nevertheless critical for shaping the binding pocket. A decrease in size at this position leads to a too-wide pocket. However, a number of beneficial exchanges concerning OA production have been identified for all other positions, including 82, 86, 92, and 94. (Figure 22).

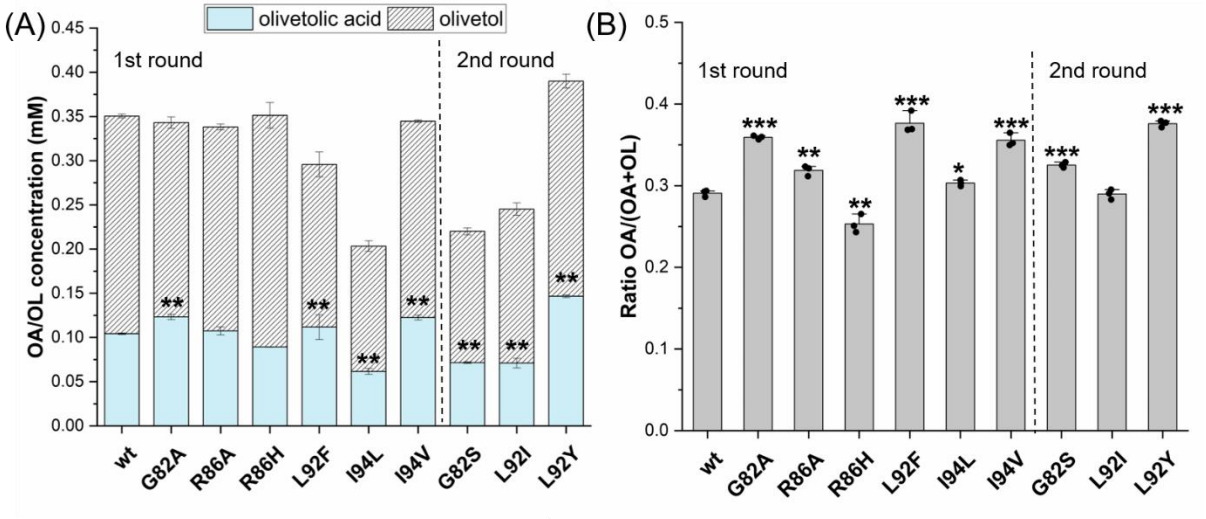


Figure 22: Screening of *Cs*OAC variants with one amino acid exchange. Variants were transformed into the strain ySP03 with the plasmid pDio-OAC carrying the mutated oac genes each. After 48 h cultivation in YPD + 0.5 mM hexanoic acid, the OA and OL concentration were determined using HPLC-UV detection at 215 nm. For comparison, the production of the wild type was analyzed and the cultivation was performed in triplicates. (A) OA and OL concentration of the wild type *Cs*OAC and the best single exchange variants G82A, R86A, R86H, L92F, I94L, and I94V and tested after the analysis of the first results variants G82S, L92I, and L92Y. (B) Ratio between the OA and the OL concentration for the before mentioned variants. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the Supplements (Table S9 & S10) (*127*).

"The substitutions of alanine and histidine for position 86 yielded outcomes comparable to those of the wild type, with no statistically significant change in concentration. For the exchange to alanine, a significant increase in OA amount in the product composition was also observed, with a concentration of 0.319 ± 0.001 mM." (*127*). As anticipated with the protein-ligand docking, the incorporation of a smaller nonpolar amino acid into the system appeared to improve the substrate's binding, thereby ensuring its correct positioning. Given

that the results obtained for OA concentration did not yield a substantial increase, further analysis of Position 86 was not pursued through a second round of individual exchanges.

The *Cs*OAC variant G82A significantly increased OA production, from 0.104 ± 0.001 mM to 0.123 ± 0.003 mM. Furthermore, the ratio OA/(OA+OL) exhibited a significant shift from 0.291 ± 0.004 to 0.359 ± 0.002. The introduction of a side chain to the protein for additional interaction appears to be advantageous, although extensive alterations to the protein, such as the replacement of glycine with a larger amino acid as tyrosine or an acidic or basic amino acid as glutamic acid, resulted in a reduction in OA production.

The two *Cs*OAC variants, L92R and L92W, have been shown to result in a significant decrease in OA production in *S. cerevisiae*. However, a notable observation is that the alteration of position 92, while maintaining a hydrophobic amino acid with and introducing an aromatic group with phenylalanine, has yielded promising outcomes, leading to an increase in OA production to 0.112 ± 0.014 mM and a concomitant significant shift in the product composition towards OA.

"The variant I94L negatively impacts both the OA and OL production, while the ratio significantly changes towards OA. Using the *Cs*OAC I94V, the OA production is increased towards 0.123 ± 0.002 mM and a 1.2-fold increase, also reflecting in a significant shift of the product ratio towards OA." (*127*).

Based on these results, the decision was made to test more amino acid exchanges on position 82, 92 & and 94 (Figure 22). For each position the results of the *in vivo* screening were compared with the *in silico* results. For position 82 further short amino acids were tested including valine and serine, for position 92 other high scoring amino acids were tested, but also tyrosine due to its high similarity in shape with phenylalanine. For position 94 aspartic acid was chosen to test, based on the docking results and serine due to a similar shape in structure with leucine.

"In this new set of variants, the two variants G82S and L92I were identified, which shifted the production ratio of the two products towards OA. However, the overall production of OA and OL was significantly lower. Among all tested variants, *Cs*OAC L92Y produced the highest amount of OA with 0.147 ± 0.006 mM, a 1.4-fold increase of OA production. Together with this increase, the product composition also significantly changed towards OA compared to the product composition of the *Cs*OAC wild type." (*127*).

The selection was made, to continue the research with nine variants—G82A, G82S, R86A, R86H, L92F, L92I, L92Y, I94L, and I94V—wherein amino acids were exchanged at four distinct positions. Among these variants, G82A, L92F, L92Y, and I94V exhibited a statistically significant increase in OA production.

2.1.6 Multiple exchange screening of *Cs*OAC variants

"Following the identification of the first beneficial *Cs*OAC protein variants, the amino acid exchanges were combined to ascertain whether the increase in OA production was additive. A protein-ligand docking was performed for all possible combinations of the nine most promising mutations of the *Cs*OAC, comprising two, three, or four exchanges combined. First, variants leading only to an implausible pose were excluded from further analysis. For each number of exchanges, the variants were then sorted according to their docking score (Table S13 - S15)." (*127*).

A total of ten variants were selected for each number of exchanges based on their scoring values, following the summarization of the docking results. In addition to the *in silico* results, the initial *in vivo* single exchange screening results were also considered when selecting variants for further study. All combinations of the best-performing exchanges (G82A, R86A, L92F, L92Y, and I94V) that had not been high-ranked in the docking experiments were added to the list of multiple exchange variants to screen *in vivo*. Consequently, 43 variants were selected for *in vivo* screening in *S. cerevisiae* to analyze the effect of OA and OL production. These included 16 variants with two amino acid exchanges, 16 variants with three amino acid exchanges, and 12 variants with four amino acid exchanges (Table 2).

Table 2: Multiple amino acid exchanges of the protein *Cs*OAC tested regarding their OA and OL production in *S. cerevisiae* and their origin (*127*).

Origin	Variant
Docking	G82A/R86H, G82A/L92F, G82S/R86A, G82S/R86H, R86A/L92F, R86A/L92Y, R86H/L92F, R86H/L92I, R86H/L92Y, R86H/I94V
	G82A/R86H/L92F, G82A/R86H/L92Y, G82S/R86A/L92F, G82S/R86A/L92Y, G82S/R86H/L92F, G82S/R86H/L92I, G82S/R86H/L92Y, G82S/R86H/I94L, G82S/R86H/I94V, G82S/L92F/I94V
	G82A/R86H/L92Y/I94V, G82S/R86A/L92F/I94L, G82S/R86A/L92F/I94V, G82S/R86A/L92Y/I94L, G82S/R86A/L92Y/I94V, G82S/R86H/L92F/I94L, G82S/R86H/L92F/I94V, G82S/R86H/L92I/I94L, G82S/R86H/L92I/I94V, G82S/R86H/L92Y/I94L
***In vivo* assay**	G82A/R86A, G82A/L92Y, G82A/I94V, R86A/I94V, L92F/I94V, L92Y/I94V
	G82A/R86A/L92F, G82A/R86A/L92Y, G82A/L92F/I94V, G82A/L92Y/I94V, R86A/L92Y/I94V
	G82A/R86A/L92F/I94V, G82A/R86A/L92Y/I94V

"The *in vivo* screening was similar performed compared to screenings with single exchanges. All variants were evaluated with regard to their impact on OA and OL production and the ratio between the two products (Figure 23-26). Among the variants including two amino acid exchanges, six exhibited noteworthy outcomes (Figure 23)." (*127*).

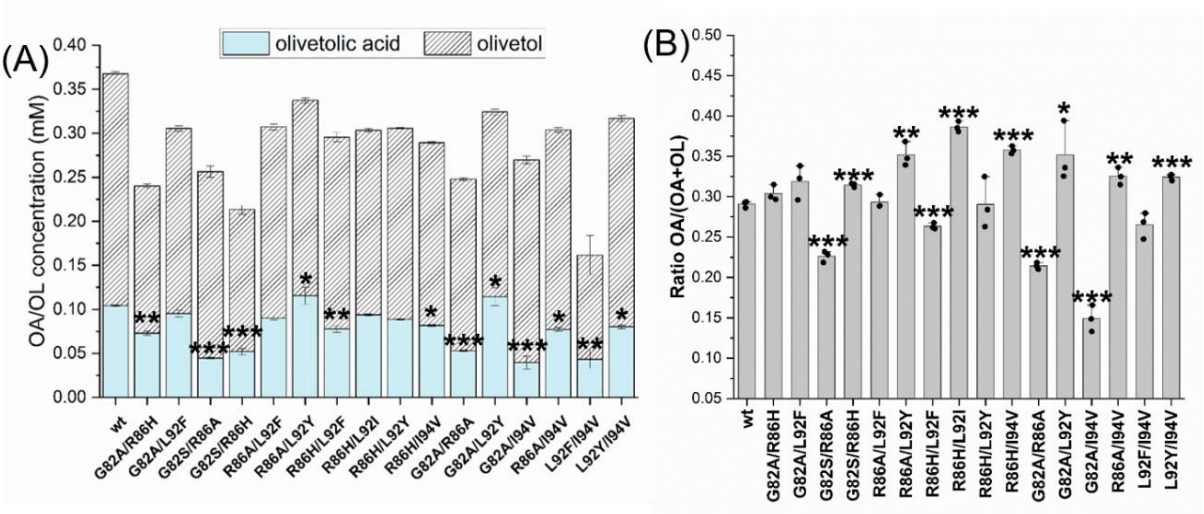


Figure 23: Olivetol (OL) and olivetolic acid (OA) production of all *Cs*OAC variants including two amino acid exchanges. For screening the different *Cs*OAC variants were cloned to pDio-OAC and *S. cerevisiae* strain ySP03 was transformed with the plasmid for each variant. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the Supplements (Table S18 & S19) (*127*).

Out of the six variants, the four *Cs*OAC variants G82A/R86H, G82A/L92F, R86A/L92F, and R86H/L92Y demonstrated comparable levels of OA and OL production compared to the *Cs*OAC wt. The variants G82A/L92Y and R86A/L92Y led to an increase in OA production, including an OA concentration of 0.115 ± 0.010 mM, which is a significant increase compared to the OA concentration yielded by the *Cs*OAC wt. The results showed a significant decrease in OL concentration upon implementation of *Cs*OAC G82A/L92Y and R86A/L92Y, leading to a substantial shift in the ratio toward OA. Intriguingly, the variants that demonstrated the greatest efficacy all incorporated the most effective single exchange L92Y and exhibited a positive effect from this mutation.

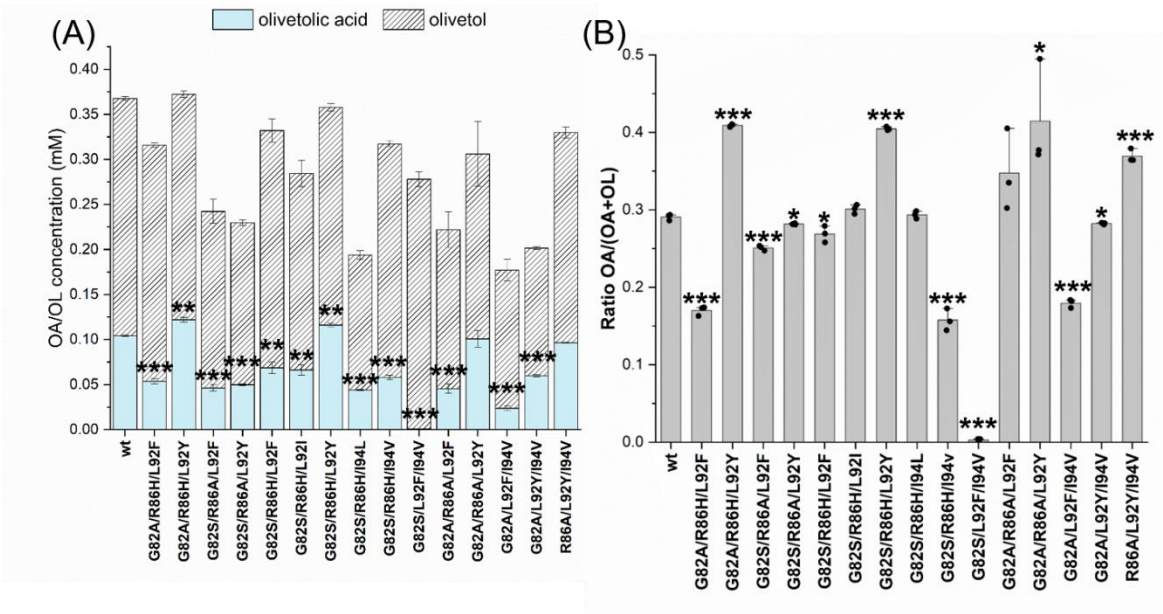


Figure 24: Olivetol (OL) and olivetolic acid (OA) production of all *Cs*OAC variants including three amino acid exchanges. For screening the different CsOAC variants were cloned to pDio-OAC and *S. cerevisiae* strain ySP03 was transformed with the plasmid for each variant. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the Supplements (Table S18 & S19) (*127*).

"Three variants, each comprising three amino acid exchanges, were identified, and a significant increase in OA production was observed. These included G82A/R86A/L92Y, G82A/R86H/L92Y and G82S/R86H/L92Y. The highest increase in OA production was observed using *Cs*OAC G82A/R86H/L92Y with a concentration of 0.122 ± 0.003 mM. Also, the variant R86A/L92Y/I94V is promising of further investigations due to the significant increase of OA proportion in the product composition." (*127*). Similar to the screening with

two amino acid exchanges, all variants include the mutation L92Y, highlighting again the importance of this amino acid exchange.

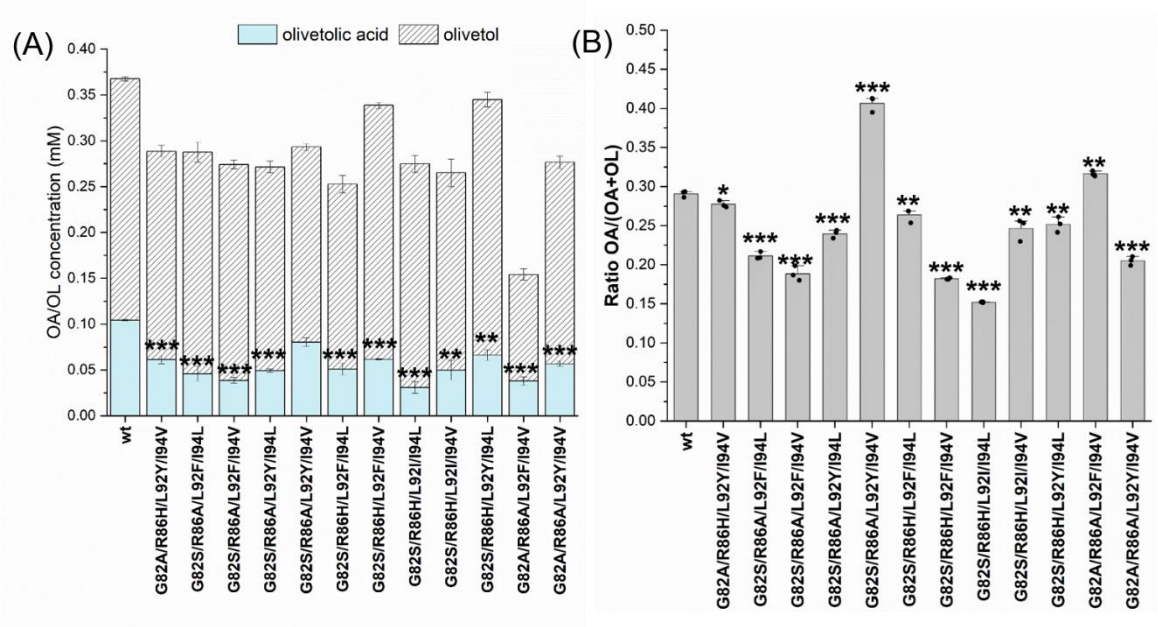


Figure 25: Olivetol (OL) and olivetolic acid (OA) production of all *Cs*OAC variants including four amino acid exchanges. For screening the different *Cs*OAC variants were cloned to pDio-OAC and *S. cerevisiae* strain ySP03 was transformed with the plasmid for each variant. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. For statistical analysis One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the Supplements (Table S20 & S21) (*127*).

"When four amino acids of CsOAC were exchanged, only adverse effects on the OA production were identified. The influence on protein expression or within the binding pocket appeared to be too demanding, resulting in a lower OA production and shift in the product composition towards OL." (*127*).

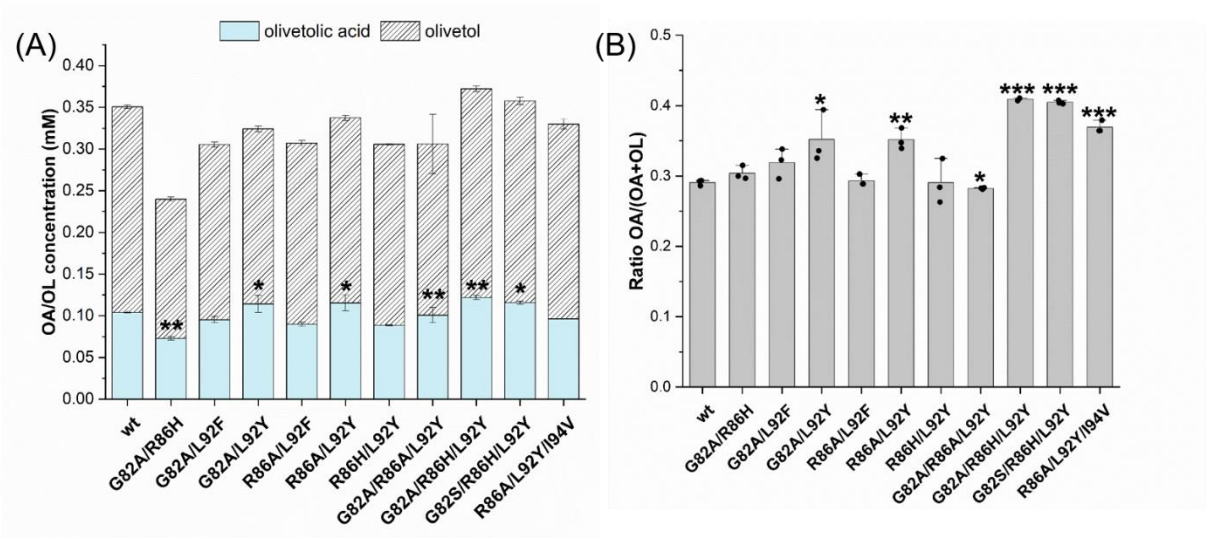


Figure 26: Screening of *Cs*OAC variants with more than one amino acid exchange. The strain ySP03 was transformed with each plasmid pDio-OAC carrying the mutated oac genes. After 48 h cultivation in YPD + 0.5 mM hexanoic acid, the OA and OL concentrations were determined using HPLC-UV detection at 215 nm. For comparison, the production of the wild type was analyzed and cultivation was performed in triplicates. (A) Production of OA and OL using the best producing variants G82A/R86H, G82A/L92F, G82A/L92Y, R86A/L92F, R86A/L92Y, R86H/L92Y, G82A/R86A/L92Y, G82A/R86H/L92Y, G82S/R86H/L92Y and R86A/L92Y/I94V. (B) Ratio between the OA and the OL concentration for the before mentioned variants. One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the supplementary material (Table S16 & S17) (*127*).

When all variants that demonstrated comparable or elevated OA production or an enhanced ratio toward OA involving multiple amino acid exchanges were evaluated, it was observed that certain exchanges were present in various variants. Specifically, R86A and R86H were present in four out of ten variants, G82A was present in five out of ten variants, and L92Y was present in seven out of ten variants. In accordance with the single exchange screening, *Cs*OAC L92Y is identified as the most promising amino acid exchange for OA and cannabinoid production in *S. cerevisiae*.

2.1.7 Genomic integration

"All *Cs*OAC variants that demonstrated comparable or improved performance relative to the wild type in the initial screening were genetically integrated with one copy of *Cs*OLS in *S. cerevisiae* to assess the impact of this integration on OA and OL titers. For this screening, first one copy of *Cs*AAE1 was integrated into *S. cerevisiae* at the YPRCτ3 locus (strain ySP01). In the last step, one copy of *Cs*OLS and one copy of *Cs*OAC were integrated for each variant into *S. cerevisiae* at the trp1 locus, resulting in ySP04. This test further characterized the best variants with respect to a similar copy number with *Cs*OLS, a comparable number of genes to

be expressed for each variant, and a limited substrate concentration. With the decrease of substrate concentration, enzyme kinetics are shifted. We assume that protein expression is less determining than with an overflow of substrate. The results are more reliable, with a genomic integration, than using a plasmid-based approach due to variation in number of integrated plasmids. With the incorporation of the aforementioned results, 19 variants were tested (Figure 27)." (*127*).

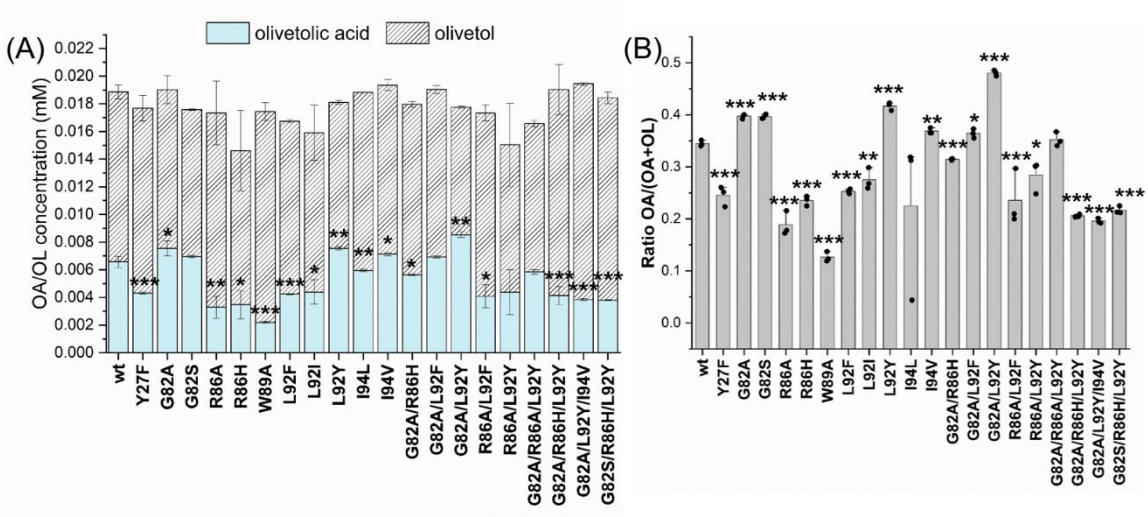


Figure 27: Olivetol (OL) and olivetolic acid (OA) production of all *Cs*OAC variants that were integrated in *S. cerevisiae*. The resulting strain ySP04 was then incubated for 48 h by 37 °C and 200 rpm and the OA and OL production was measured by HPLC-UV. (A) Concentration of the produced OL and OA after 48 h. (B) Ratio between OA and OL. One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of *Cs*OAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the supplementary material (Table S26 & S27) (*127*).

In the first step of the analysis, the known variant *Cs*OAC Y27F was examined to determine its relative production of OA and OL compared to the *Cs*OAC wt. It should be noted that, until this point, only the results of *in vitro* assays have been published. In these assays, the variant demonstrated a relative activity of approximately 162 %, while the wild type exhibited 100 % relative activity (*37*). Consequently, it is of significant interest to ascertain whether the variant exhibits a comparable effect *in S. cerevisiae*. "For the wild type, an OA concentration of 0.0066 ± 0.0004 mM was observed, while for *Cs*OAC Y27F, a concentration of 0.0043 ± 0.0001 mM was measured, resulting in a relative activity of 65% compared to the wild type. Given that the purified enzyme exhibited higher activity, the observed decrease in yeast must be attributed to a lower expression level, correlating to a decreased enzyme concentration. A structural explanation could be the loss of a hydrogen bond in the complex protein system through which a lower protein concentration could be explained. This result underscores the

importance of the *in vivo* screening system for the variants if our primary goal is the transformation of the cannabinoid biosynthesis in *S. cerevisiae*." (*127*). Consequently, the combination of the variants examined in this thesis with Y27F is not a subject of significant interest and has not been pursued.

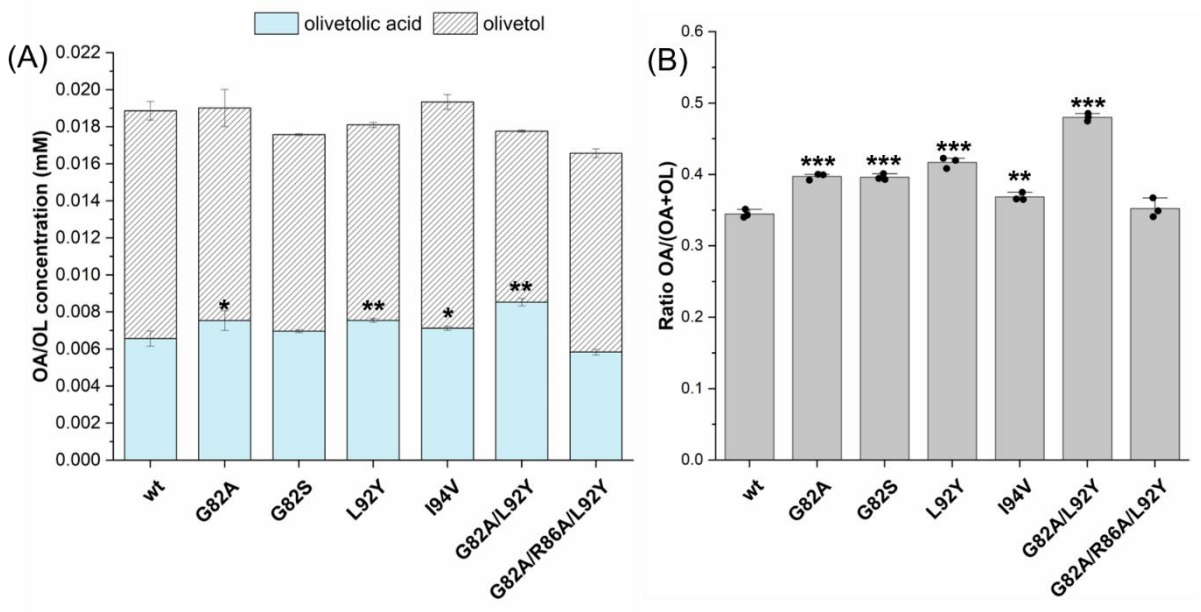


Figure 28: Olivetol (OL) and olivetolic acid (OA) production of the best performing *Cs*OAC variants that were integrated in *S. cerevisiae*. The resulting strain ySP04 was then incubated for 48 h by 37 °C and 200 rpm and the OA and OL production was measured by HPLC-UV. Cultivation was performed in triplicates. (A) Production of OA and OL using the best producing variants G82A, G82S, L92Y, G82A/L92Y and G82A/R86A/L92Y. (B) Ratio between the OA and the OL concentration for the above-mentioned variants. One-way ANOVA tests including post hoc analyses (Bonferroni) were performed. All mean values were compared to the values of CsOAC wt. Asterisks denote statistically significant difference between two mean values: $p \leq 0.05$ (*), $p \leq 0.01$ (**), $p \leq 0.001$ (***). Detailed statistical reports are shown in the supplementary material (Table S24 & S25). (*127*).

"Upon monitoring the values for the single exchanges, the most notable changes were observed for the variants G82A, L92Y and I94V (Figure 6). For the three variants, the OA production after 48 h is significantly increased. Resulting in an OA concentrate increase to 0.0071 ± 0.0001 mM for I94V, 0.0074 ± 0.0001 mM for G82A and 0.0076 ± 0.0001 mM for L92Y, what is an increase of 30 % in relative activity. Additionally, the OL production decreased, indicating the utilization of the linear tetraketide derived from *Cs*OAC. This effect is also visible in product composition of OA and OL. Here, the ratio significantly increased for the three beforementioned variants and for G82S. The most noteworthy change was shown for variant L92Y where the composition significantly increased from 0.345 ± 0.006 for the wild type to 0.417 ± 0.008. In the case of the multiple exchange, the variant G82A/L92Y is the only variant, that is promising for further usage in the cannabinoid biosynthesis. The variant was

found to result in a significant OA concentration increase towards 0.0085 ± 0.0002 mM, representing a 1.7-fold increase overall in OA production. Additionally, the ratio of OA/OL exhibited a shift towards OA, with a value of 0.480 ± 0.006." (*127*)

2.2 Screening of different carboxylic acids regarding olivetolic acid derivatives production

In natural cannabinoid biosynthesis, hexanoic acid is one of the starter molecules. The molecule provides the structural elements for the *n*-pentyl chain of all cannabinoids. In addition to this alkyl chain, natural cannabinoid structures containing a *n*-propyl or a methyl chain are characterized (*12, 4*). It has been demonstrated that the variation of the acid is possible by implementing of the available natural building blocks, thereby facilitating the variation of the respective cannabinoids. In cannabinoid biosynthesis, using *S. cerevisiae* as a host organism facilitates the external feeding approach, wherein hexanoic acid is introduced by adding into the cultivation media. Alternatively, the acid can be produced by the *S. cerevisiae* cell itself (Figure 29). In *S. cerevisiae*, the synthesis of fatty acids is catalyzed by a type I fatty acid synthase (FAS), which is comprised of two subunits, each encoded by a distinct gene, known as *FAS1* and *FAS2 (136, 137)*. By using the mutation *FAS2* G1250S, medium-chain fatty acids are produced (*138, 139*). The synthesis of novel cannabinoid structures can be achieved through the addition of different acids to the media (*41*). With this approach, OA, CBGA, and THCA derivatives were formed containing an elongated C_6, a shortened C_4, a branched alkyl, a 5-pentene, or a 6-heptyn chain (*25*).

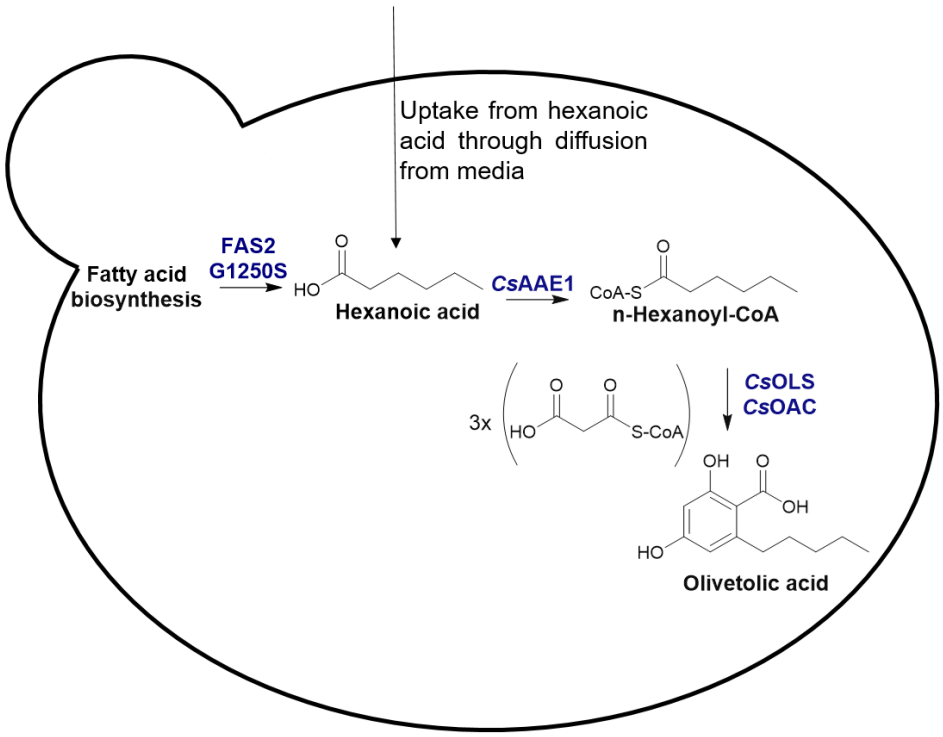


Figure 29: OA biosynthesis in *S. cerevisiae*, highlighting the two different routes for hexanoic acid supply and the proteins used in the *S. cerevisiae* strain yAP01.

The effect of adding different carboxylic acids to the cultivation media was screened using the *S. cerevisiae* strain yAP01 (Table 9). This strain was kindly provided by Alex Prima. In this strain *Cs*AAE1 is integrated into the genome with one copy Additionally, multiple copies of *Cs*OLS and *Cs*OAC Y27F were integrated into the *S. cerevisiae* strain, using an integration vector construct from the EasyCloneMulti vector set, targeting Ty4 LTR sites within the yeast genome (*131*). The integration of the gene *FAS2* G1250S enabled the cell to synthesize hexanoic acid itself. For a first evaluation of how strongly the OA production can be influenced through additional feeding of hexanoic acid, the strain yAP01 was incubated for 48 h at 30 °C with and without adding 0.5 mM hexanoic acid into the media. To determine the concentration of OA and OL, yeast cells with an OD_{600} = 125 were harvested and extracted. The samples were analyzed using HPLC-UV (Table 3).

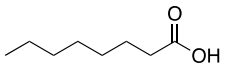
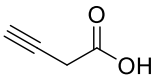
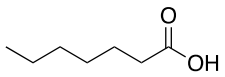
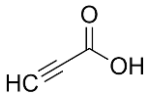
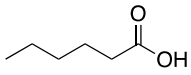
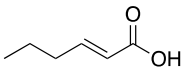
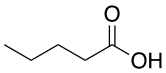
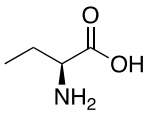
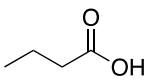
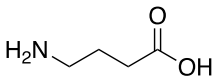
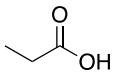
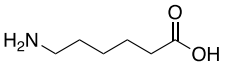
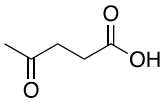
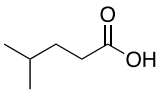
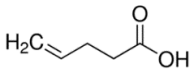
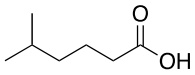
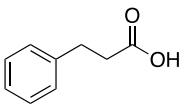
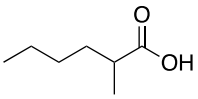
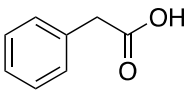
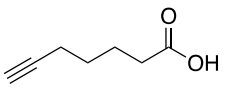
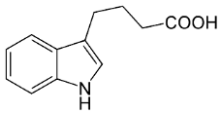
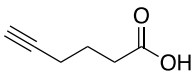
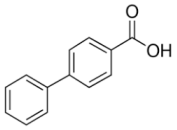
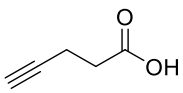
Table 3: Intercellular concentration of OA and OL after cultivation of strain yAP01 for 48 h with the feeding of 0.5 mM hexanoic acid and without.

	Hexanoic acid feeding		**Without hexanoic acid feeding**	
	OA	**OL**	**OA**	**OL**
Concentration [mM]	0.068 ± 0.005	0.152 ± 0.016	0.026 ± 0.002	0.035 ± 0.001

The *FAS2* mutation G1250S and the integration of the proteins from *C. sativa* (*Cs*AAE1, *Cs*OLS & *Cs*OAC Y27F) enable the strain to produce OA and OL without the additional substitution of hexanoic acid through the media, with a concentration of 0.061 ± 0.003 mM for both products. In the experiments involving the addition of hexanoic acid to the media, an overall product concentration of 0.22 ± 0.02 mM was measured. This value indicates a 3.6-fold increase compared to the experiments conducted without the supplementation of hexanoic acid. These results demonstrate the capacity of the strain yAP01 to produce OA and OL from glucose in the medium. In parallel, it was demonstrated that the titer continued to rise in response to the feeding of hexanoic acid. Therefore, the proteins' capacity remains to convert the fed acid. This means that the strain yAP01 can be used for preliminary screening of other fatty acids besides hexanoic acid regarding their conversion to OA derivatives.

To screen alternative carboxylic acid and identify OA or OL derivatives formed with the *S. cerevisiae* strain yAP01, 0.5 mM of each acid was added into the cultivation media. After 48 h of cultivation, the cells were harvested, and extraction was performed. The prepared samples were analyzed in HPLC-UV and HPLC-MS to identify the OA and OL derivatives. A library comprising a variety of carboxylic acids was compiled for this screening. The library contains 25 structures that have the potential to be converted in the first step of cannabinoid biosynthesis due to their carboxylic acid. In addition to this group, the molecules are all constructed with different structural elements. Initially, there are fatty acids that are analogous to hexanoic acid, yet they possess an elongated or reduced alkyl chain. Furthermore, molecules containing a branched alkyl group or featuring double or triple bonds are also evaluated. The library was expanded to encompass molecules incorporating an amino or hydroxy group into the alkyl chain. Finally, the library comprises structures with a carboxylic group and an aromatic structure element. For all molecules, the samples were analyzed for new peaks in the UV spectra and the correlating masses for the novel structures (Table 4, Figure S4 –S9).

Table 4: List of screened fatty acid and the identified masses.

Name	Structure	*m/z* [M+H]$^+$	Name	Structure	*m/z* [M+H]$^+$
Octanoic acid		253.14 209.15	3-Butynoic acid		-
Heptanoic acid		239.13 195.14	2-Propiolic acid		-
Hexanoic acid		225.25 181.11	Trans-2-hexanoic acid		223.09 179.10
Pentanoic acid		-	α-amino-butyric acid		-
Butanoic acid		-	γ-amino-butyric acid		-
Propanoic acid		-	ε-amino-caproic acid		-
Ethanoic acid		-	4-oxo-pentanoic acid		-
4-methyl-pentanoic acid		225.10 181.12	4-Pentenoic acid		-
5-methyl-hexanoic acid		-	Phenyl-propanoic acid		-
2-methyl-hexanoic acid		-	Phenyl-acetic acid		-
6-Heptynoic acid		235.09 191.10	Indole-3-butyric acid		-
5-Hexynoic acid		-	Biphenyl-4-carboxylic acid		-
4-Pentynoic acid		-			

The longer chained acids C_7 and C_8 are converted from all three enzymes, resulting in the identification of the expected masses for the respective OA and the OL derivatives. Similar results were exhibited for 4-methyl-pentanoic acid, showing that another C6 chain is accepted as a substrate for all three enzymes. For the two other non-linear C7 alkyl chain acids masses of the products were not identified, indicating that only minor alterations in structure can disrupt the formation of OA and OL derivatives. The addition of a double bond into the C_6 chain with Tran-2-hexanoic acid also led to the formation of novel OA and OL derivatives, as evidenced by the identification of the correlating masses. Similar to the results from Luo et al. (*25*), the conversion of 6-heptynoic acid to the corresponding OA and OL derivatives was also observed. The shorting of the alkin chain did not result in the identification of OA and OL derivatives. In line with these findings, the shortening of the chain did not result in any observable product, and no new masses were detected. This outcome is incongruent with the results of other studies (*25*). One potential explanation for this observation is the capacity of the strain to synthesize hexanoic acid endogenously, a process that is in direct competition with the substrate provided. In instances where a substrate exhibits a low binding affinity for one of the enzymes involved in OA synthesis, the resulting product concentration may be below the threshold capable of detection. Incorporating a novel structural element comprising a hydroxy or amino group or a phenolic group did not result in identifying any OA or OL derivatives.

A majority of the structures from the carboxylic acid library were not converted. With the applied screening system, the reason for this can be diverse. A notable challenge emerges in hexanoic acid production from *S. cerevisiae* itself, as it consistently competes with all fed acids for the enzymes. In this strain, *Cs*AAE1 was used as the enzyme for the activation of the acid. This enzyme is best suitable for the activation of hexanoic acid and medium-chained fatty acids (*55*). Most of the converted acids also consisted of a C6 body, a structural feature consistent with the recognized substrates of *Cs*AAE1. In the absence of activation, the following reactions with *Cs*OLS and *Cs*OAC remain inaccessible. In the other case, the acids are activated by *Cs*AAE1; however, they are not accepted as substrates by *Cs*OLS. Consequently, the reaction toward OL and OA cannot be catalyzed. It has been demonstrated that structures analogous to hexanoyl-CoA are predominantly accepted as substrates. This observation is consistent with the notion that hexanoyl-CoA serves as the natural substrate of *Cs*OLS. The identification of CBGVA and other cannabinoid structures with a shorter chain than C5 in the plant suggests that the proteins accept shorter-chain fatty acids as a substrate. This observation indicates that these molecules do not undergo the reaction with *Cs*AAE1. For the other molecules, it is

impossible to clearly distinguish whether *Cs*AAE1 did not accept the molecule as a substrate or *Cs*OLS due to the analytical method.

In summary, the strain, yAP01, converted two carboxylic acids that were also converted into OA and OL analogs in other studies (*25*) were converted with our strain, yAP01. Furthermore, three structures with octanoic acid, 4-methylpentanoic acid, and trans-2-hexanoic acid were first converted into their respective OL and OA derivatives. To form cannabinoids with these molecules, it is important to ensure the following reaction toward CBGA-C5. NphB is known as promiscuous regarding its acceptance of the aromatic acceptor, and it is also shown that with protein engineering, product formation can strongly be influenced. Therefore, the substrate acceptance of NphB needs to be investigated in a further step (*112*).

2.3 Generation of cannabigerolic acid derivatives using NphB

In cannabinoid biosynthesis, the pentyl chain is a common feature of the structures, and it is introduced with hexanoic acid. As previously described in chapter 2.2 of this thesis, if the cannabinoid biosynthesis is transferred to *S. cerevisiae* with the feeding of other acids, the formation of analogs containing alteration of the *n*-pently chain is possible. Studies by Luo et al. demonstrated that THCA derivatives are formed through feeding acids and utilizing all proteins derived from *C. sativa*. However, in the context of cannabinoid biosynthesis within *S. cerevisiae*, the prenyltransferase is frequently substituted with alternative proteins, such as NphB. This substitution introduces novel opportunities for synthesizing cannabinoid derivatives, capitalizing on the high promiscuity of the aromatic substrate. With the screening of different olivetolic acid analogs, the synthesis of CBGA analogs is possible, enabling its downstream reaction to their respective THCA or CBCA analogs (Figure 30).

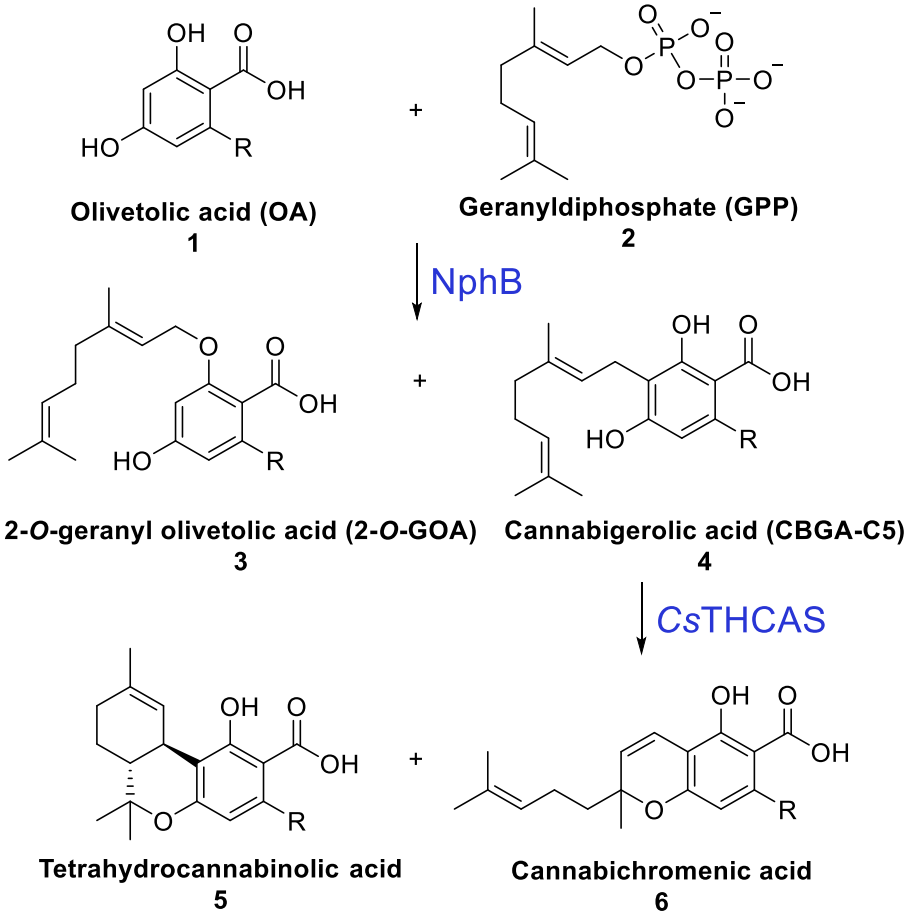


Figure 30: Schematic figure of the variation of OA to gain new products using NphB and the downstream reaction with *Cs*THCAS.

2.3.1 Binding mode identification of OA with NphB

1,6-DHN is a substrate of NphB, for which more than one product has been identified. MD-Simulations have led to the proposal that the substrate position within the binding pocket influences the regioselectivity of the protein. The proposition that this can be extended to OA as an aromatic substrate is put forth (*140*). The availability of three-dimensional protein structures, which have both substrates bound, has enabled the identification and characterization of the catalytic and substrate-binding amino acids (*101*).

Met162, Phe213, Ser214, Tyr288, and Gln295 are the amino acids responsible for stabilizing the aromatic substrate. In addition to the amino acids of NphB, the two substrates also interact with each other; thereby, the prenyl donor provides further stabilization of the aromatic substrate within the binding pocket. The subsequent carbocation is stabilized by a π-chamber comprising the aromatic substrate and the phenolic amino acid side chains of Tyr121 and Tyr216. It is imperative for the prenylation reaction to take place that the aromatic substrate's reactive site be close to the carbocation and, consequently, the π-chamber. In conclusion, the prenylation pattern is significantly influenced by the orientation of the aromatic substrate within the NphB binding pocket (*106*).

In the context of OA as an aromatic substrate and its prenylation pattern, the orientation and position of the resorcylate core in relation to the catalytic center are highly important. The position of the carbon atom at position 3 (C3), which is targeted for the formation of CBGA-C5, and the position oxygen at position 2 (O2), which is targeted for the formation of 2-*O*-GOA, are of particular significance. As no 3D structure is available containing NphB and OA as a substrate, first, the identification of the binding modes of the resorcylate core from OA is necessary. Since NphB wt and NphB G286S/Y288A have different main products, it was expected to yield different binding modes for each protein variant. 2-*O*-GOA is the main product of the reaction using the wild type, while the variant G286S/Y288A almost exclusively produces CBGA-C5. With the information regarding the different binding modes, the evaluation of the OA analogs regarding the reaction towards CBGA analogs or 2-*O*-GOA analogs can be possible. To this objective, protein-ligand docking was conducted using the NphB wild type (PDB-ID: 1zdw) (*101*) and the NphB variant G286S/Y288A (*112*) with the GOLD software (*132*). The advantage of the protein structure with the PDB-ID 1zdw was that GPP was already solved as a substrate of NphB, which could be used for calculations that also included the interaction between OA and GPP.

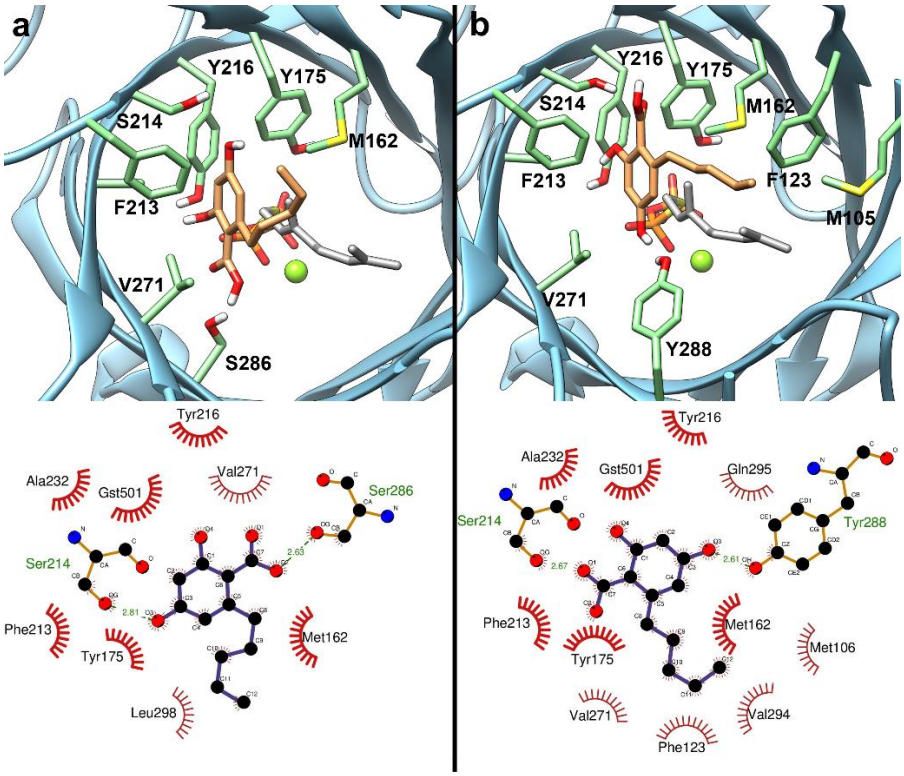


Figure 31: 3D depiction created using UCSF chimera of the docked pose of olivetolic acid **1a** (orange color) inside the binding pocket of NphB G286S/Y288A (a) and NphB wild type (b) with highlighting all interacting amino acid side chains (green color) and 2D depiction created using LigPlot of the interaction profile of **1a** with the amino acids of NphB G286S/Y288A (a) and NphB wild type (b) (*141, 142, 108*).

A comparison of the two binding mode predictions of the docking process reveals a shift of the resorclyate core towards the π-chamber and the carbocation using the variant G286S/Y288A. This is made possible by the two amino acid substitutions. The substitution of glycine with serine at position 286 introduces a polar amino acid, thereby enabling the formation of a distinct hydrogen bond network compared to the one observed between OA and the wild type. In the wild type, two hydrogen bonds were identified between Ser214 and the carboxyl group, and Tyr288 and OH4. In contrast, two hydrogen bonds were identified in the variant between Ser214 and OH4 and between Ser286 and the carboxyl group (Figure 31). This novel hydrogen bond network promotes a 180 ° rotation of the resorcylate core within the binding pocket. The distinct binding mode and the shift in the orientation of OA towards GPP provide insight into the variation in the prenylation pattern observed in the variant compared to the wild type. A reduction in the distance between the C3 of OA and the C1 of GPP results in an increased likelihood of prenylation at this position. Replacing of the bulky tyrosine side chain at position 288 with alanine facilitates not only the formation of a new hydrogen bond network but also the rotation of the resorcylate core. The additional space thus created allows the n-pentyl group to orient towards Met162 and Phe213, facilitating favorable

hydrophobic interactions. The study by Lim et al. supports these findings, demonstrating that the hydrogen bond between OH4 and Tyr288 directly affects the unfavorable orientation of the resorcylate core concerning CBGA-C5 production (*114*). In a manner analogous to 1,6-DHN, two distinct binding modes for OA were identified, representing prenylation at a different position and applicable to other OA derivatives.

2.3.2 *In silico* screening of olivetolic acid derivatives regarding their acceptance as a substrate of NphB

The following step is to design a library of OA derivatives comprising the most promising entries for conversion toward a CBGA derivative. These entries should be diverse and lead towards the formation of novel cannabinoid structures. Accordingly, a library of OA derivatives that can be synthesized using organic chemistry is docked with NphB (Table 4 & 6). Since the structure and space of the binding pocket between the variant G286S/Y288A and wt differ, it was chosen to perform and analyze the docking with both. As previously analyzed, the binding mode of OA was utilized to identify the key points of the binding mode for the OA derivatives. It was observed that the resorcylate core, including the prenylation position C3, is in close proximity to the C1 of the GSPP. At the same time, the side chain is located in the entrance area of the binding pocket. A comparative analysis of X-ray structures, molecular dynamics (MD) simulations, and docking methodologies reveals that the separation between the two reaction positions ranges from approximately 3.0 to 4.5 Å. This is due to the fact that the prenyl acceptor molecule is moving toward the carbocation after the reaction, resulting in a relatively broad distance range (*106, 114, 101*). For all potential substrates of NphB, 50 possible binding poses and their distance from the C3 and C1 of the GSPP were monitored. The resulting data will be divided into two sections: alkyl OA analogs and aromatic OA analogs.

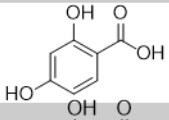
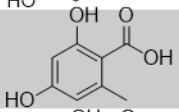
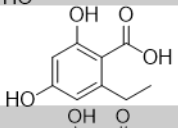
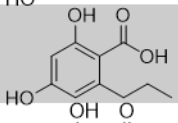
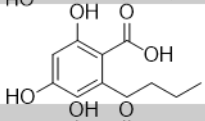
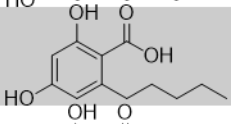
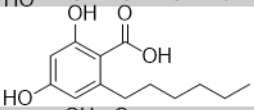
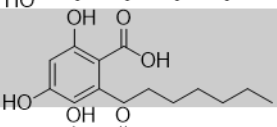
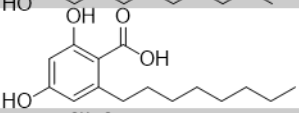
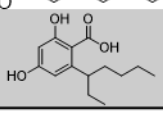
2.3.2.1 *In silico* screening of olivetolic acid derivatives consisting of an alkyl chain

The binding pocket of NphB, which is approximately 500 Å^3 in size, has been observed to accommodate an extended alkyl chain or even a branched alkyl chain. In a study conducted by Lee et al., it was demonstrated that NphB accepts olivetolic acid derivatives with a *n*-propyl, *n*-heptyl, *n*-nonyl, *n*-undecyl ,and *n*-tridecyl chain as a substrate, forming their respective CBGA analog. Additional conversion rates or product ratio details remain unknown (*143*). To date, studies have been conducted with shorter chain derivatives, including orsellinic acid and divarinic acid. The reactions towards CBGA-C1 and CBGA-C3, respectively, are catalyzed using NphB wt. In the case of orsellinic acid, several products, besides CBGA-C1, have been

identified. The use of the NphB variant G286S resulted in a shift in product regioselectivity towards CBGA-C1 (*113*).

The docking was performed with olivetolic acid derivatives consisting of no alkyl chain and a range from a methyl group up to a nonyl group. In addition, an olivetolic acid analog containing a branched alkyl group **1e** was also added for analysis (Table 5).

Table 5: List of all molecules consisting of an alkyl chain that were used in the docking experiments, including the distribution of docking poses where the distance between the C3 core and the C1 GSPP ≤ 4.5 Å . Overall 50 poses were generated during the docking experiment and the percentage of the poses, that fit inro the distance range are shown (*108*).

IUPAC-Name	Structure	Poses where distance between C3 core and C1 GSPP ≤ 4.5 Å	
		NphB G286S/Y288A	NphB wt
2,4-dihydroxybenzoic acid **1j**		100 %	98 %
2,4-dihydroxy-6-methylbenzoic acid **1b**		100 %	100 %
2-ethyl-4,6-dihydroxybenzoic acid **1c**		100 %	100 %
2,4-dihydroxy-6-propylbenzoic acid **1k**		100 %	94 %
2-butyl-4,6-dihydroxybenzoic acid **1l**		100 %	86 %
2,4-dihydroxy-6-pentylbenzoic acid **1a**		100 %	64 %
2-hexyl-4,6-dihydroxybenzoic acid **1m**		100 %	24 %
2-heptyl-4,6-dihydroxybenzoic acid **1n**		100 %	22 %
2,4-dihydroxy-6-octylbenzoic acid **1d**		100 %	30 %
2-(heptan-3-yl)-4,6-dihydroxybenzoic acid **1e**		90 %	100 %

For the NphB variant G286S/Y288A, all OA analogs with an alkyl side chain were correctly oriented in the binding pocket towards the catalytic amino acid with a smaller distance than 4.5 Å between the two possible reaction sites. In contrast, for the wild type, an elongation of

the *n*-pentyl chain led in over 70 % of all created binding poses to a resulting pose, with the C3 not orientated towards the C1 of the GSPP. This resulted in higher distances than 4.5 Å between the two reactive sites. In summary, the analysis indicates that all alkyl chains, irrespective of their length, exhibited the capacity to undergo conversion towards their respective CBGA analogs when engaged with NphB G286S/Y288A. Given the significant variations in the results observed between the two variants, it is of considerable interest to investigate the reduction of the alkyl chains, including orsellinic acid **1b** and an ethyl chain **1c**, as well as the extension of an *n*-octyl **1d** and a branched alkyl chain **1e**.

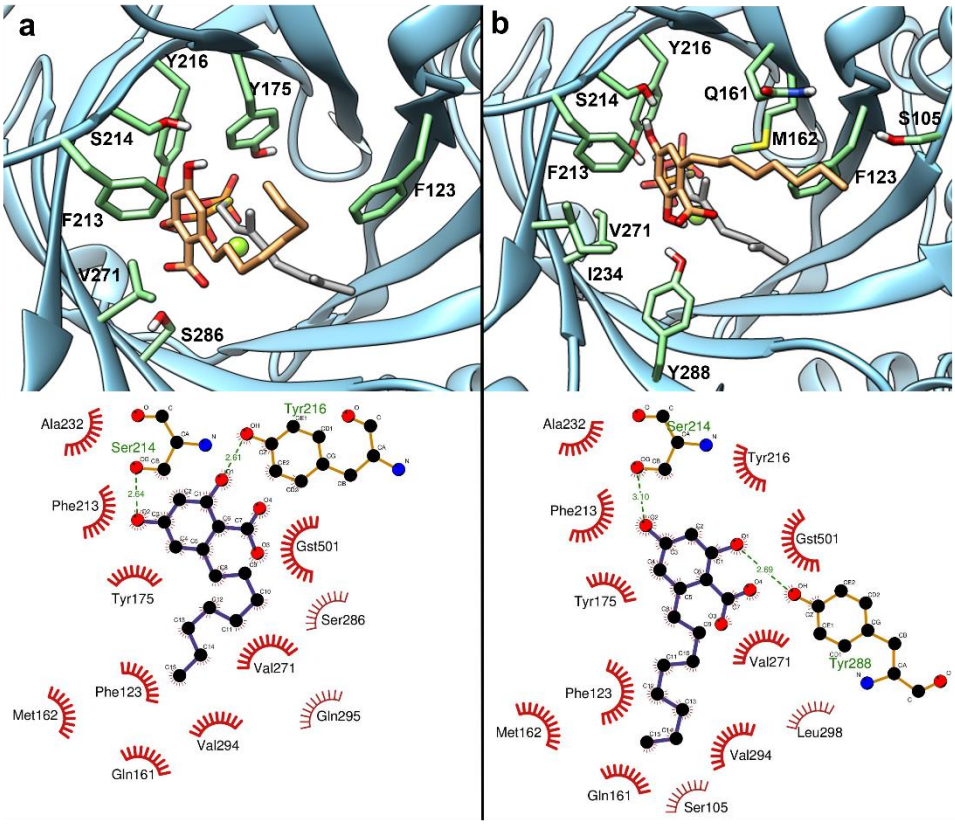


Figure 32: 3D depiction created using UCSF chimera and 2D decpiction created using LigPlot of the docked pose of **1d** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green) (*142, 141, 108*).

The binding mode of both molecules **1d** (Figure 32) and **1e** (Figure S10) indicates that an extension fits into the spacious binding pocket, where the chain is pointing towards the binding pocket entrance. For the variant G286S/Y288A, the calculated binding modes with a focus on the resorcylate core of both substrates were similar to those of OA with the variant.

However, a difference in the ranking of the docking was observed. Although **1d** is among the highest-ranked molecules, **1e** is ranked below olivetolic acid despite the potential for a greater number of atoms to interact within the binding pocket (see **1d**, **1e** Table 6). The *n*-octyl side

exhibits a higher degree of freedom, resulting in a greater number of possible orientations for this side chain. Non-binding interactions with other amino acids can be developed, and distances to amino acids can be reduced.

The resulting binding modes are comparable with each other when **1d** and **1e** are docked to the NphB wt, yet they are not comparable to olivetolic acid **1a**. The formation of hydrogen bonds was observed between Ser214 and O4, as well as between Tyr288 and the carboxyl group for **1d** and **1e**. The hydrogen bond network exhibited by these two molecules differs from that previously observed for **1a**, resulting in a distinct orientation of the benzene-1,3-diol group. This altered orientation creates a significant distance between the amino acids of the π-chamber and the reactive site, which may hinder the formation of the π-chamber and consequently impede the stabilization of the carbocation. Such impediments could potentially affect the enzymatic activity. This binding mode is also different from the one calculated for the variant G286S/Y288A. A key structure element for this change is the exchange of the tyrosine on position 288. This side chain is bulky, and the exchange towards alanine creates more space in the binding pocket, allowing a different orientation of the alkyl chain of the OA derivatives. This altered alkyl chain configuration subsequently influences the resorcylate core's disposition. Given the constrained spatial requirements of NphB wt within the binding pocket, the elongated and spacious alkyl chains favor a distinct positioning of the resorcylate core, which might potentially compromise the catalytic performance. The elongation of the alkyl chain to an octyl chain is a promising approach in terms of conversion efficiency and C3 prenylation using the variant G286S/Y288A, so a conversion equivalent to **1a** can be expected. A branched chain still leads to the desired binding mode in the docking, but in terms of ranking, the substrate conversion is expected to be lower.

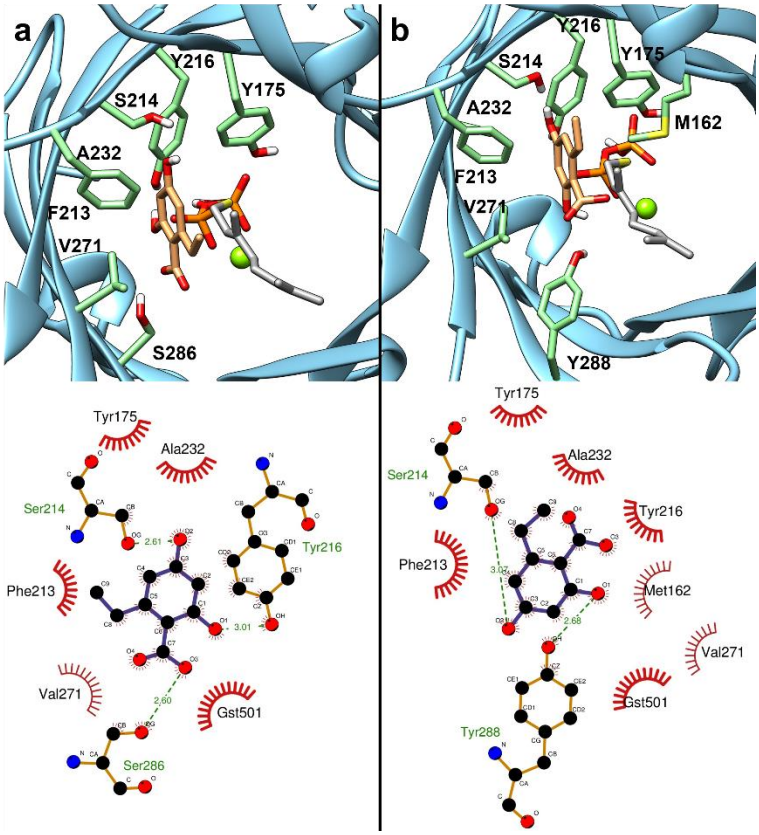


Figure 33: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1c** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

The effects of shortening the alkyl chain have been previously examined; however, to expand the knowledge about NphB and orsellinic acid, it was decided to test the exchange of the *n*-pentyl group towards methyl **1b**, as well as the exchange towards ethyl **1c**, since cannabinoids with this group are not known to be synthesized before. (Figure 33, S11).

As observed for **1b** and **1c**, the shortening of the alkyl chain results in diminished hydrophobic interactions that stabilize the molecule inside the protein binding pocket, consequently leading to the lowest ranking in both docking analyses. Notably, the pose calculated for the variant is analogous to the OA binding mode, suggesting that CBGA derivatives may emerge as a product of the enzymatic reactions. The absence of additional stabilization and the presence of higher fluctuations in the binding mode suggests that the conversion rate may be lower compared to **1a**.

Table 6: Results from the docking experiments using different OA-like substrates, including an alkane moiety and NphB G286S/Y288A and wild type highlighting all amino acids interacting with the substrates. The molecules are sorted after the ranking of the docking results. The interactions were calculated using LigPlot (*108*).

NphB G286S/Y288A			NphB wild type		
Compounds	**Hydrogen bonds**	**Non-bonded contacts**	**Compounds**	**Hydrogen bonds**	**Non-bonded contacts**
1d (*n*-octyl)	S214(Ser:OG-Subs:O2), Y216(Ser:OH-Subs:O1)	GPP, F123, Q161, M162, Y175, F213, S214, Y216, A232, V271, S286, V294, Q295	**1d** (*n*-octyl)	S214(Ser:OG-Subs:O2), Y288(Tyr:OH-Subs:O1)	GPP, S105, F123, Q161, M162, Y175, F213, S214, Y216, A232, V271, Y288, V294, L298
1a (*n*-pentyl)	S214(Subs:O2-Ser:OG), S286(Ser:OG-Subs:O4)	GPP, M162, Y175, F213, S214, Y216, A232, V271, S286, L298	**1e** (hept-3-yl)	S214(Ser:OG-Subs:O2), Y288(Tyr:OH-Subs:O4)	GPP, F123, M162, Y175, F213, S214, Y216, A232, I234, V271, Y288, V294, Q295, L298
1e (hept-3-yl)	S214(Ser:OG-Subs:O2)	GPP, F123, M162, Y175, F213, S214, Y216, A232, I234, V271, S286, Q295, L298	**1a** (*n*-pentyl)	S214(Ser:OG-Subs:O1), Y288(Tyr:OH-Subs:O3)	GPP, M106, F123, M162, Y175, F213, S214, Y216, A232, V271, Y288, V294, Q295
1c (ethyl)	S214(Ser:OG-Subs:O2), Y216(Subs:O1-Tyr:OH), S286(Ser:OG-Subs:O4)	GPP, Y175, F213, S214, Y216, A232, V271, S286	**1c** (ethyl)	S214(Ser:OG-Subs:O2), Y288(Subs:O1-Tyr:OH)	GPP, M162, Y175, F213, S214, Y216, A232, V271, Y288
1b (methyl)	S214(Subs:O2-Ser:OG), Y216(Tyr:OH-Subs:O1), S286(Ser:OG-Subs:O4)	GPP, M162, Y175, S214, Y216, A232, V271, S286	**1b** (methyl)	S214(Ser:OG-Subs:O2), Y288(Tyr:OH-Subs:O3), Y288(Subs:O1-Tyr:OH)	GPP, M162, Y175, F213, Y216, A232, V271, Y288

Overall, it was decided to synthesize four different olivetolic acid derivatives containing an alternative alkyl chain, which were later tested *in vitro* using the NphB wt and the variant NphB G286S/Y288A (Figure 34).

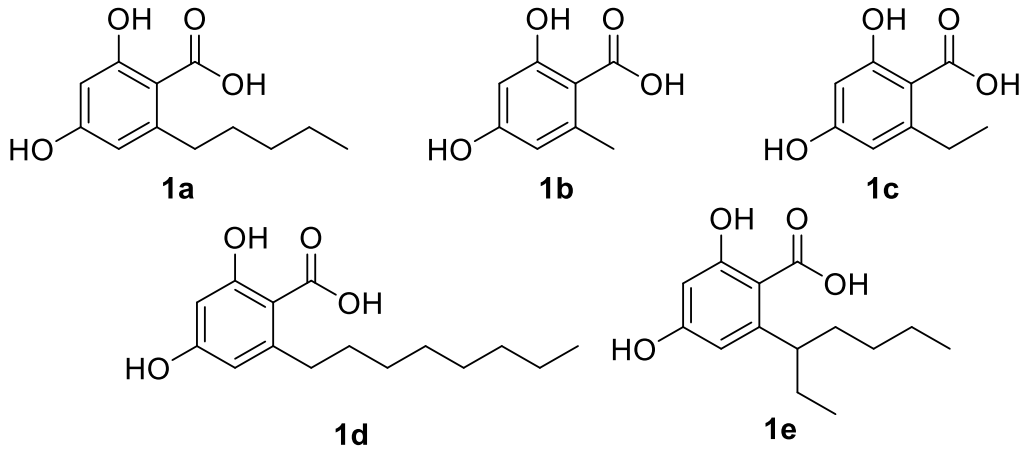
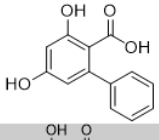
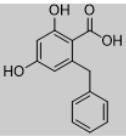
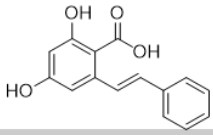
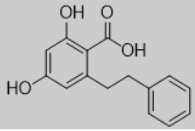
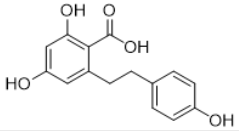
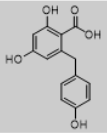
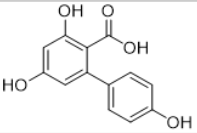
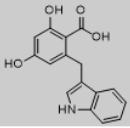
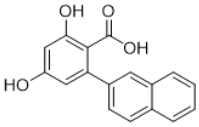
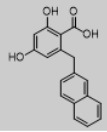
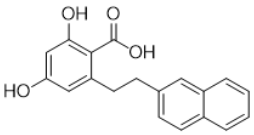


Figure 34: List of all olivetolic acid derivatives containing an alkyl moiety that were synthesized and validated in this study regarding their interaction with NphB (*108*).

2.3.2.2 *In silico* screening of olivetolic acid derivatives consisting of an alkyl aromatic side chains

In other organisms, such as *H. umbraculigerum*, aralkylic cannabinoids have been identified as secondary metabolites, resulting in a possible transferation of these biosynthesis in combination with the cannabinoid biosynthesis in *S. cerevisiae* (144). The prenylation reaction for arakylic CBGA of the respective OA analogs in *S. cerevisiae* could be introduced with the use of NphB or new arakylic cannabinoid structures could be formed. A list of OA analogs containing an aromatic side chain was generated and docked to identify the most promising ones in regard to the conversion to CBGA analogs (Table 7).

Table 7: List of all molecules consisting of an alkyl chain used in the docking experiments, including the distribution of docking poses where the distance between the C3 core and the C1 GSPP ≤ 4.5 Å Overall 50 poses were generated during the docking experiment and the percentage of the poses, that fit inro the distance range are shown (*108*).

IUPAC-Name	Structure	Poses where distance between C3 core and C1 GSPP ≤ 4.5 Å	
		NphB G286S/Y288A	NphB wt
3,5-dihydroxy-[1,1'-biphenyl]-2-carboxylic acid **1f**		74 %	6 %
2-benzyl-4,6-dihydroxybenzoic acid **1i**		100 %	0 %
(E)-2,4-dihydroxy-6-styrylbenzoic acid **1g**		38 %	0 %
2,4-dihydroxy-6-phenethylbenzoic acid **1h**		56 %	0 %
2,4-dihydroxy-6-(4-hydroxy-phenethyl)benzoic acid **1o**		30 %	0 %
2,4-dihydroxy-6-(4-hydroxybenzyl)benzoic acid **1p**		86 %	16 %
3,4',5-trihydroxy-[1,1'-biphenyl]-2-carboxylic acid **1q**		42 %	0 %
2-((1H-indol-3-yl)methyl)-4,6-dihydroxybenzoic acid **1r**		42 %	0 %
2,4-dihydroxy-6-(naphthalen-2-yl)benzoic acid **1s**		0 %	0 %
2,4-dihydroxy-6-(naphthalen-2-ylmethyl)benzoic acid **1t**		32 %	0 %
2,4-dihydroxy-6-(2-(naphthalen-2-yl)ethyl)benzoic acid **1u**		24 %	0 %

The binding mode of all aralkylic OA analogs examined in this study resulted in a distance of more than 4.5 Å between the C3 of the resorcylate core and the C1 of the GSPP using the wt. This finding indicates that the conversion towards CBGA derivatives is not promising when NphB wt is used an enzyme. Therefore, the binding mode of all molecules using NphB wt will not be discussed in detail. However, the variant G286S/Y288A demonstrated different outcomes, depending on the group attached to the resorcylate core. The best results were achieved using a benzyl **1i** or 4-ethyl phenyl group **1p**, where over 85 % of the created poses were within the distance criteria. The exchange towards a phenyl group also led to a primarily favorable orientation of the resorcylate core, with 74% of the poses falling within the distance criteria. Given the fact that the two groups without the hydroxyl group led to an overall better result, it was decided to add **1f** and **1i** into our substrate library. A naphthyl group in different attachments to the resorcylate core (**1s**, **1t** & **1u**) mostly led to an unfavorable binding mode and was excluded for further analysis. The size and inflexibility of this group could explain these results.

The catalytic reaction of NphB using **1g** and **1h** as substrate could result in CBGA-ST and CBGA-BB formation. Therefore, testing of these structures was included in this study. In the following, the highest-ranked pose of the aralkylic molecules **1f—1h** included in the substrate library is analyzed in detail regarding the interaction with NphB wt and G286S/Y288A.

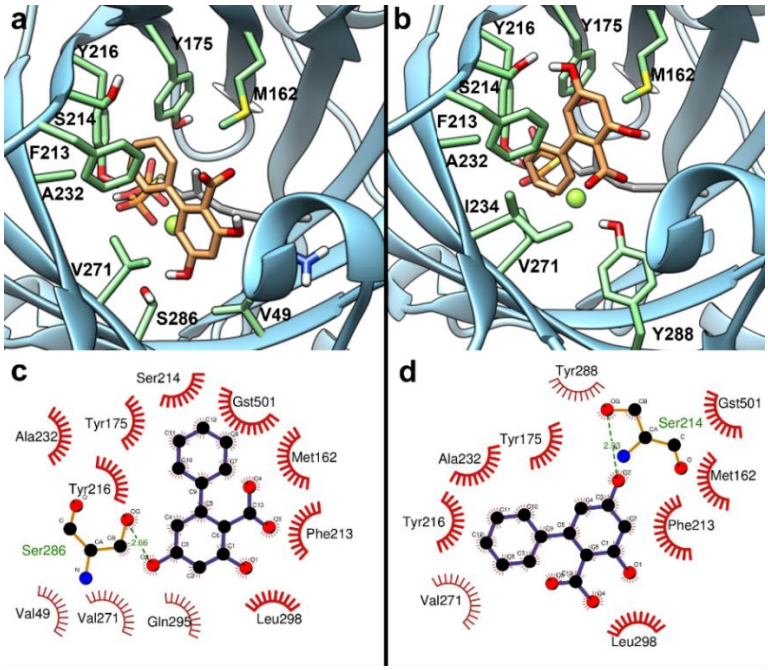


Figure 35: 3D depiction created using UCSF chimera, and 2D depiction created using LigPlot of the docked pose of **1f** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

The highest-ranked poses of **1f**, **1g** and **1h** are similar to each other using both proteins. They highly differ from the previously described position of the resorcylate core and **1a** (Figure 35, Figure S12 & S13). To provide a comprehensive illustration of the proposed binding modes for all three molecules, the binding mode of **1f** will be discussed in detail. In the highest-ranked pose, the newly introduced phenyl group is oriented toward the π-chamber rather than the resorcylate core. The potential prenylation positions are situated at a considerable distance from the carbocation that is formed at a subsequent stage. It is notable that hydrogen bonding can be observed even when the core structure is placed at a distance from the active site of the binding pocket. The resorcylate core, in turn, is capable of forming hydrogen bonds with the same amino acids, such as Ser286, through a change in the amino acid side chain conformation. The introduction of a second aromatic group results in the addition of a structure element that resembles the resorcylate to the molecule. The molecule exhibits limited flexibility due to the shorter linkage between the two aromatic features in comparison to the n-pentyl chain. When the phenylic side chain adopts a position analogous to that of the resorcylate core of the OA binding mode, the side chain finds itself in a more hydrophobic environment, while the resorcylate core maintains its stability through a hydrogen bond network. This finding further constrains the conversion of **1f**, **1g**, and **1h** to the corresponding CBGA derivatives. However, a binding mode analogous to **1a** was also calculated, but with a lower scoring value, indicating that substrate acceptance is still a possibility.

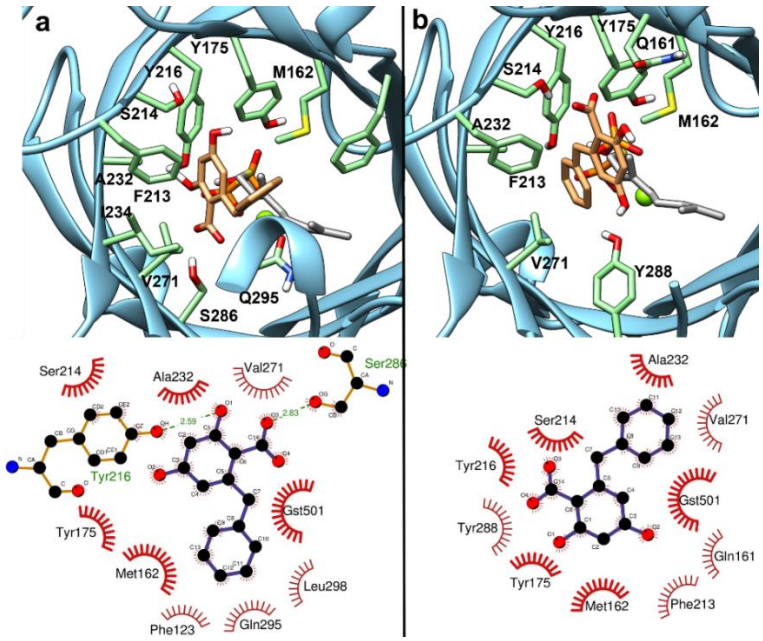


Figure 36: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1i** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green) (*142, 141, 108*).

"A variation of the proposed binding mode was found using a benzyl group in **1i** and the variant G286S/Y288A (Figure 36). The highest ranked pose (Table 8) resembles the olivetolic acid binding mode, but not the other derivatives containing a phenylic moiety. This discrepancy can be explained by the bond between the core structure and the aromatic group. Additional flexibility and rotational freedom are provided by the benzyl group. As a consequence, an orientation of this molecule inside the binding pocket is possible in which the resorcylate core is stabilized in the vicinity of the carbocation, while the phenyl group is in a similar position with respect to the alkyl side chains. This result cannot be transferred to the wild type because the bulky side chain of Tyr288 blocks the above orientation. For this compound, a better conversion compared to **1e** - **1g** with the variant is anticipated." (*108*)

Table 8: Results from the docking experiments using different OA-like substrates including an alkane moiety and NphB G286S/Y288A and wild type highlighting all amino acids, which interact with the substrates. The molecules are sorted after the ranking of the docking results. The interaction were calculated using LigPlot (*108*).

NphB G286S/Y288A			NphB wild type		
Compounds	**Hydrogen bonds**	**Non-bonded contacts**	**Compounds**	**Hydrogen bonds**	**Non-bonded contacts**
1g (styryl)	S286(Ser:OG-Subs:O2) S214(Ser:OG-Subs:O4), Y216(Ser:OH-Subs:O1)	GPP, V49, F123, Q161, M162, Y175, F213, S214, Y216, A232, V271, S286, V294, L298	**1h** (phenethyl)	S214(Subs:O2-Ser:OG)	GPP, S51, M162, Y175, F213, S214, Y216, V271, K284, Y288, L298
1i (benzyl)	Y216(Subs:O1-Tyr:OH), S286(Ser:OG-Subs:O3)	GPP, F123, M162, Y175, S214, Y216, V271, S286, Q295, L298	**1g** (styryl)	S214(Ser:OG-Subs:O2)	GPP, M162, Y175, F213, S214, Y216, V271, K284, Y288, L298
1h (phenethyl)	Y216(Subs:O1-Tyr:OH), S286(Ser:OG-Subs:O2)	GPP, V49, Q161, M162, Y175, F213, S214, Y216, A232, V271, S286, V294, Q295	**1i** (benzyl)	X	GPP, Q161, M162, Y175, F213, S214, Y216, A232, V271, Y288,
1a (*n*-pentyl)	S214(Subs:O2-Ser:OG), S286(Ser:OG-Subs:O4)	GPP, M162, Y175, F213, S214, Y216, A232, V271, S286, L298	**1f** (phenyl)	S214(Ser:OG-Subs:O2)	GPP, M162, Y175, F213, S214, Y216, A232, V271, Y288, L298
1f (phenyl)	S286(Ser:OG-Subs:O2)	GPP, V49, M162, Y175, F213, S214, Y216, V271, S286, A288, Q295, L298	**1a** (*n*-pentyl)	S214(Ser:OG-Subs:O1), Y288(Tyr:OH-Subs:O3)	GPP, M106, F123, M162, Y175, F213, S214, Y216, A232, V271, Y288, V294, Q295

Overall, it was decided to synthesize four different olivetolic acid derivatives containing a second aromatic feature, which were later tested *in vitro* using the NphB wt and the variant NphB G286S/Y288A (Figure 37).

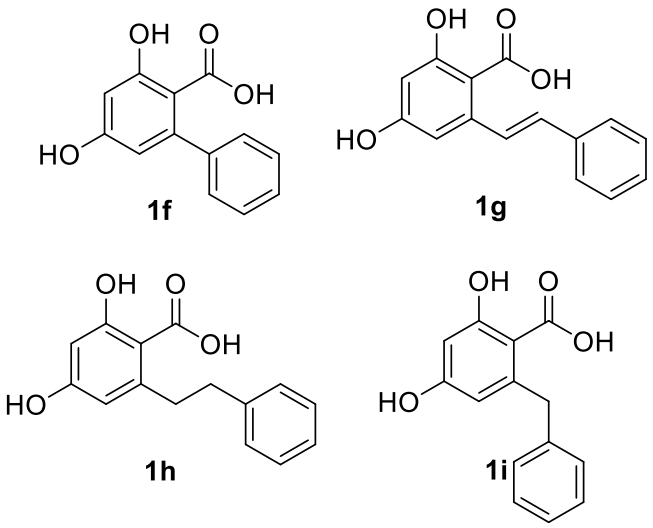


Figure 37: List of all olivetolic acid derivatives containing an alkyl moiety that were synthesized and validated in this study regarding their interaction with NphB (*108*).

2.3.3 *In vitro* screening of olivetolic acid derivatives for the formation of cannabigerolic acid derivatives using NphB

The olivetolic acid derivatives selected by docking were synthesized and characterized and provided by Jasmin Wloka (*108*).

"For a first evaluation of the potential of these structures as alternative substrates, conversion with NphB has to be determined. A high conversion of the substrate is required for implementation into the cannabinoid biosynthesis. The identification of the high-converting substrate is followed by the identification of the prenylation position of these molecules. All synthesized derivatives **1** (Figure 34 & 37) were tested as prenyl acceptors using *in vitro* assays. The reaction mixtures contained 100 µg mL^{-1} purified protein and were started with the addition of substrates (1 mM olivetolic acid derivatives, 2 mM GPP). The assay samples were incubated for 18 h at 1100 rpm. The enzymes were then denatured and precipitated, and the supernatant was analyzed by HPLC-UV. Substrate concentration was calculated using standards and their absorbance at 225 nm to determine the conversion (Figure 38). To ensure the dependency of the substrate conversion on the presence of the active protein, samples were measured without enzyme (data not shown). In this experiment, the total substrate conversion is monitored regardless of the number of products or the prenylation position to get an

overview about the most promising substrates regarding usage in the cannabinoid biosynthesis." (*108*)

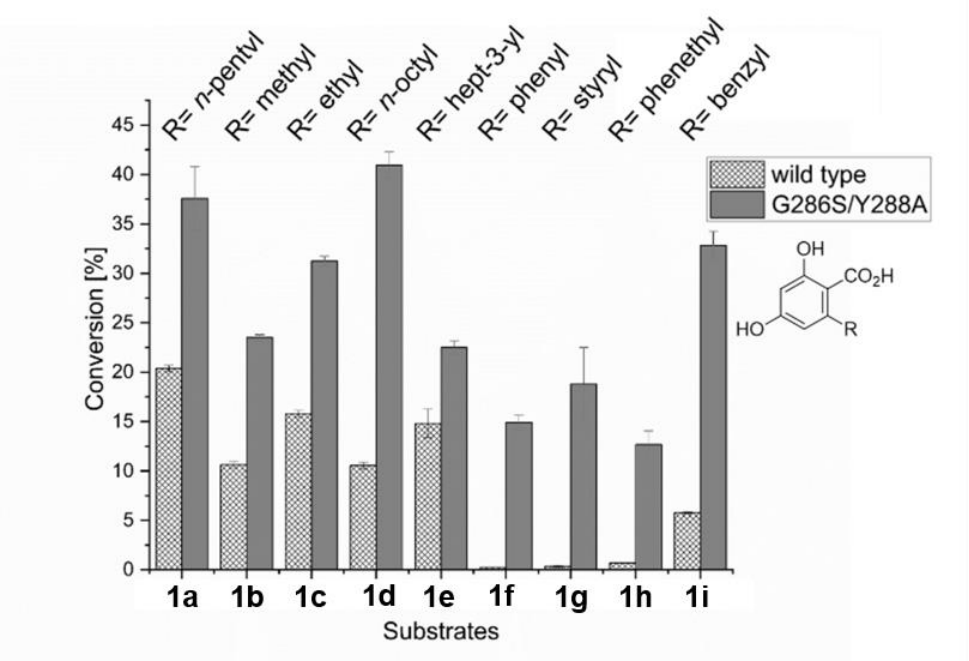


Figure 38: Conversion of all tested substrates **1a**-**1i** using GPP as prenyl donor and NphB wild type (pattern) and NphB G286S/Y288A (grey) for conversion, respectively. All experiments were performed in triplicate (*108*).

"The conversion of OA **1a** has been described previously and was therefore used as a control reaction for the isolated protein and its function (*100*). After 18 h, 40 % of the olivetolic acid was converted by the G286S/Y288A variant and 20 % by the wild type. Identified by HPLC-UV and standard solutions, the major product of the variant was CBGA-C5, and of the wild type 2-*O*-GOA. This result demonstrated that the assay system can be applied to alternative substrates. In the following, the conversion is compared to that of **1a**, marking the 40 % yield obtained with NphB G286S/Y288A as a reference point.

For all substrates **1a** - **1i** product formation was visible in the HPLC chromatogram using the variant G286S/Y288A (Figure S31 – S33). For orsellinic acid **1b** the results were already confirmed before, with the identification of geranylated products (*53, 113, 112*)." (*108*). A comparison of the conversion between the NphB wt and the variant for derivatives containing an alternative alkyl chain (**1a–1e**) reveals a 2-fold increase in conversion using the variant. Given that the variant demonstrates a significantly higher CBGA-C5 production than the wt and the derivatives share a common core structure and analogous building blocks with OA, this outcome aligns with the anticipated results. However, substrate **1d** exhibited an exceptional

outcome, with the variant demonstrating a 4-fold higher conversion than the wt. This observation is further supported by the docking results with NphB wt, which indicate that the elongated alkyl chain adopted a binding mode that positions the reactive site of the aromatic acceptor in a direction that is oriented away from the catalytic amino acid. "The exchange of the tyrosine to an alanine at position 288 generates more space inside the binding pocket, which is beneficial for a more spacious side chain such as the octyl group of **1d**. A trend was observed that shortening the alkyl chain correlated with lower conversion." (*108*). While orsellinic acid **1b** has a conversion of 23.5 ± 0.3 %, **1c** has a conversion of 31.2 ± 0.5 %, and **1d** has a conversion of 41.0 ± 1.3 %. Compared to OA, shortening of the alkyl chain results in lower conversion, while elongation only leads to a comparable conversion. "This result corresponds to the docking experiments: for a shorter alkyl chain, less interactions were identified, and the overall ranking was lower (Table 6), implying a lower binding affinity.

Comparing the substrate acceptance of compounds **1f**, **1g**, and **1h**, almost no conversion was detected with the wild type. These derivatives have the same structural element, the aromatic group attached to C6 of the core structure instead of an alkyl chain. The linkage between the phenyl group and the core structure has no effect on the conversion. This result was expected when combined with the docking results. The proposed binding mode was composed of the resorcylate core not oriented towards the active site (Figure 35)." (*108*). The NphB variant G286S/Y288A demonstrated a conversion rate of ~14% for compounds **1g** – **1h**. In this instance, the impact of the linkage between the phenyl moiety and the core structure is also constrained.

"The only substrate molecule with an aromatic side chain that led to a high conversion is **1i**. A conversion rate of 33% was determined using the variant. In parallel with the *in vitro* results (Figure 38), molecule **1i** had the most promising docking outcome (Figure 36) compared to the other derivatives containing a second phenolic structural element. The hypothesis that structures **1f** - **1h** would not be converted as well as olivetolic acid **1a** confirms the proposed docking studies.

As only a low conversion has been identified for the molecules **1e** - **1h**, the implementation in the cannabinoid biosynthesis by the feeding of these substrates does not seem to be likely at the moment. If there is interest in one of the resulting unnatural cannabinoids, protein engineering of NphB towards a higher conversion of these substrates is required." (*108*)

2.3.4 Structure elucidation of CBGA-C5 derivatives

With the monitoring of the conversion, only hints of the formation of the catalytic products can be concluded. For further identification of the formed products all assay samples were analyzed by High resolution mass spectrometry (HRMS).

"Mass analysis allows confirmation of prenylated products and determination of the number of products formed by catalytic reaction with NphB (Figure S14 – S22). For all substrates, only one mass was identified that correlated with the mass of a prenylated product. In the catalytic reaction, almost exclusively one main product was formed. The retention time and the measured spectra for all geranylated products seemed to be similar for the wild type and the G286S/Y288A variant. This leads to the conclusion that both protein variants exclusively form the same main products, except for olivetolic acid. For derivatives **1a**, **1b**, **1c**, **1d** and **1i**, the reaction products were further analyzed by UPLC-ESI-MS/MS (Figure S15 – S17 & S22). The fragmentation pattern was similar to that of CBGA-C5, suggesting the formation of a C-prenylation. The two patterns distinguish in the formation of fragments smaller then m/z = 219.10. While the fragmentation of 2-*O*-GOA leads to smaller fragments in a high distribution, CBGA-C5 does only have these fragments in traces. This was also observed for the beforementioned substrates (*53, 84*). Prenylation at position C5 can be almost excluded for all products with a side chain comparable to the *n*-pentyl moiety due to the steric hindrance at this position. Prenylation at this position using **1a** was not observed as a product in *in vitro* assays with NphB wild type or G286S/Y288A (*100, 112, 114*). As the space of the group attached to C6 increases, the C5 position becomes even more blocked." (*108*)

"In order to gain more insight into the structure of the newly identified products, a scale-up was carried out with the molecules that led to the highest conversions **1d** and **1i**. These two molecules are most promising for integration into the cannabinoid biosynthesis and a CBGA derivatives of the structures **1d** and **1i** were never synthesized before.

It was decided to perform the scale-up reaction only with the G286S/Y288A variant, since the products for the variant and the wild-type have the same retention time and the fragmentation pattern of the MS-MS measurements exhibited similar results. The structures of these products were further analyzed after isolation by preparative Reversed Phase- (RP-)HPLC. ^{1}H and ^{13}C NMR, COSY, HSQC and HMBC were measured for **1i** (Figure 39, Figure S26 - 30). Due to the HMBC correlation between H8 and C5, the isolated product has to be prenylated at C3 position. For **1i** the conversion towards a CBGA derivative was demonstrated. ^{1}H NMR,

HMBC and HSQC were measured for **1d** (Figure S23 - 25). With this the conformation of a C-prenylation was possible, due to the interaction in the HMBC spectrum between H5 and C8 (Figure S78). This signal is only possible if the prenylation is on C3 position and the H5 is still existent. A prenylation on C5 position can be excluded. (*108*)"

No.	**1**	
	δ_H, M (*J* in Hz)	δ_C
1	-	110.7
2	-	164.5
3	-	111.2
4	-	155.7
5	5.80 s	106.6
6	-	143.7
7	-	172.3
8	4.47 s	39.1
9	-	141.4
10	7.16 m	128.9
11	7.16 m	127.7
12	7.06 m	124.8
13	3.09 d (7.1 Hz)	22.1
14	5.05 t (7.1 Hz)	124.2
15	-	132.1
16	2.00 m	35.1
17	1.97 m	26.5
18	5.17 t (7.1 Hz)	124.3
19	-	130.5
20	1.53 s	25.5
21	1.60 s	17.5
22	1.68 s	16.0

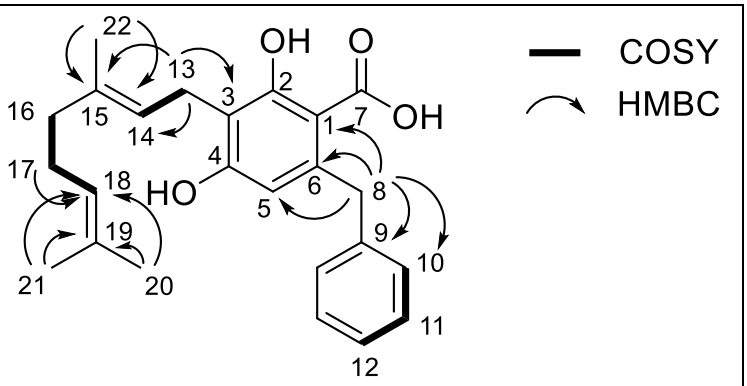


Figure 39: Summary of the NMR data of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid) **4i** in DMSO-d6.

2.3.5 NphB-*Cs*THCAS crude cell extract assays

"For further identification of the prenylation site using **1c**, **1d** and **1i** assays employing the protein *Cs*THCAS were conducted. As described before, the *Cs*THCAS is the final protein in the cannabinoid biosynthesis. It catalyzes the intramolecular cyclization of the monoterpene moiety in CBGA-C5 to THCA-C5. It has already been demonstrated, that this enzymatic reaction also occurs with CBGA derivatives that differ only in their side chain (*112*). With the formation of THCA analogs, we can proof the presence of CBGA analogs as products of the enzymatic reaction of NphB G286S/Y288A. Crude cell extract assay were performed combining cell lysate with overexpressed NphB G286S/Y288A and *Cs*THCAS. As substrates GPP and the most promising analogs **1a**, **1c**, **1d** and **1i** were added to the cell lysate. The assay samples were analyzed through UPLC-ESI-MS/MS after 24 hours of incubation. First the assays were analyzed using **1a** as substrate. With comparison of the fragmentation pattern and retention time of an analytic standard of THCA-C5 and CBCA-C5 the formation of these

products using **1a** and GPP as substrates was confirmed (Figure 40). As described in former reports *Cs*THCAS catalyzes the formation of THCAS-C5 and CBCA-C5 as main products using higher pH values (*91*). Using the substrates **1c**, **1d** and **1i** we also identified two masses for the corresponding THCA and CBCA derivatives, with similar fragmentation patterns to THCA-C5 and CBCA-C5 and masses of the fragments matching different side chains (Figure 40, 41)." (*108*).

Subsequent to the identification of the downstream products, an additional piece of evidence, in addition to the NMR data, for the prenylation on the C3 position for **1c**, **1d**, and **1i** was obtained. The formation of THCA and CBCA derivatives of the substrates using *Cs*THCAS is only possible if the NphB variant G286S/Y288A catalyzes the formation of CBGA derivatives. Consequently, the identified mass and product for each substrate from the *in vitro* assay must be CBGA derivatives.

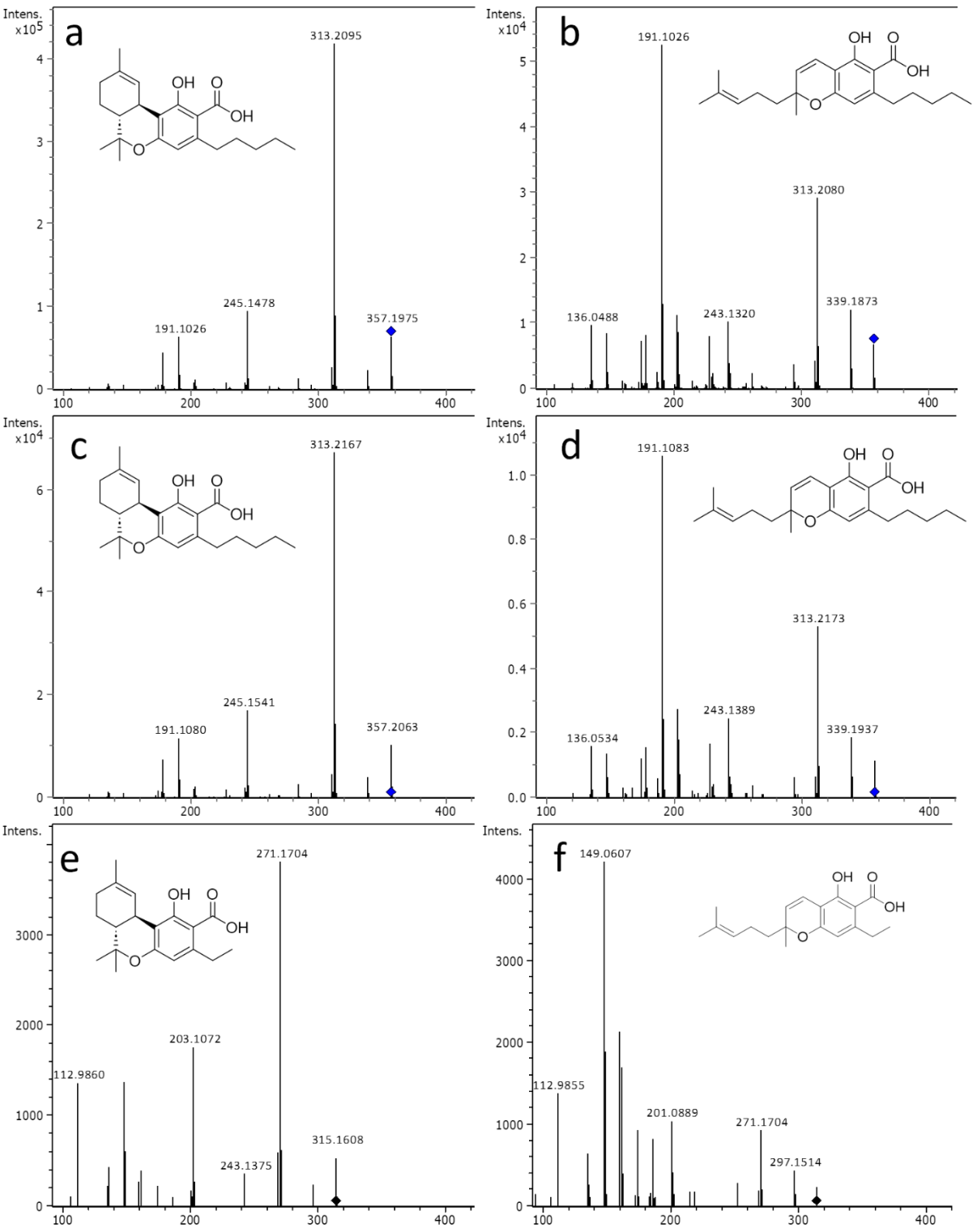


Figure 40: HR-MS-MS spectra of (6aR,10aR)-1-hydroxy-6,6,9-trimethyl-3-pentyl-6a,7,8,10a-tetrahydro-6H-benzo[c]chromene-2-carboxylic acid THCA-C5 **5a** as standard (a) and as assay product (c), of (6aR,10aR)-3-ethyl-1-hydroxy-6,6,9-trimethyl-6a,7,8,10a-tetrahydro-6H-benzo[c]chromene-2-carboxylic acid THCA-C2 **5c** (e), of 5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-7-pentyl-2H-chromene-6-carboxylic acid CBCA-C5 **6a** as standard (b) and as assay product (d), and of 7-ethyl-5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-2H-chromene-6-carboxylic acid CBCA-C2 **6c** (f).

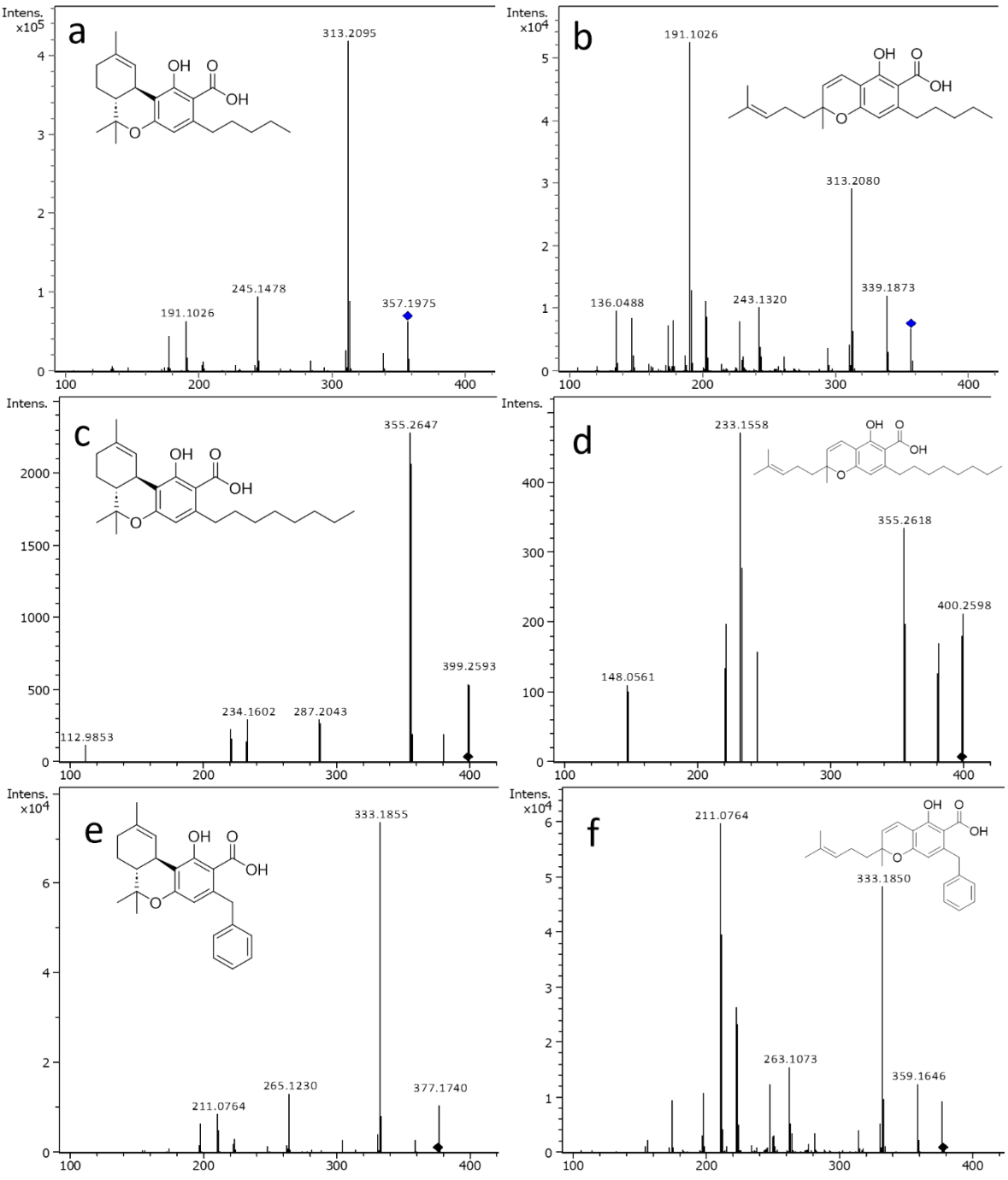


Figure 41: HR-MS-MS spectra of (6aR,10aR)-1-hydroxy-6,6,9-trimethyl-3-pentyl-6a,7,8,10a-tetrahydro-6H-benzo[c]chromene-2-carboxylic acid THCA-C5 **5a** as standard (a), of (6aR,10aR)-1-hydroxy-6,6,9-trimethyl-3-octyl-6a,7,8,10a-tetrahydro-6H-benzo[c]chromene-2-carboxylic acid THCA-C8 **5d** (c), of 5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-7-octyl-2H-chromene-6-carboxylic acid **5i** (e), of 5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-7-pentyl-2H-chromene-6-carboxylic acid CBCA-C5 **6a** as standard (b), of 5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-7-octyl-2H-chromene-6-carboxylic acid CBCA-C8 **6d** (e) and of 7-benzyl-5-hydroxy-2-methyl-2-(4-methylpent-3-en-1-yl)-2H-chromene-6-carboxylic acid **6i** (f).

2.3.6 Establishment of a new system for quantification of CBGA-C5 production

In all experiments, the CBGA-C5 concentration and conversion were measured using HPLC-UV measurements. This method is time-consuming; therefore, an alternative detection method for quantifying CBGA-C5 was conducted and tested. To identify an NphB variant demonstrating high activity, a high-throughput system is required that can analyze NphB assays in parallel without the use of a complex extraction method for the product.

In addition to CBGA-C5, another equimolar formed product of the catalytic reaction of NphB, independent of the substrate used, is a diphosphate group, which is cleaved off from GPP (*101*). A common enzymatic reaction for diphosphate quantification is the quantification over firefly luciferases (*145, 146*). This enzyme catalyzes the adenosine triphosphate (ATP) dependent oxidative decarboxylation of D-luciferin. The product of this reaction is oxyluciferin, and the emitted photon of visible light can be detected and quantified (Figure 42) (*147, 148*).

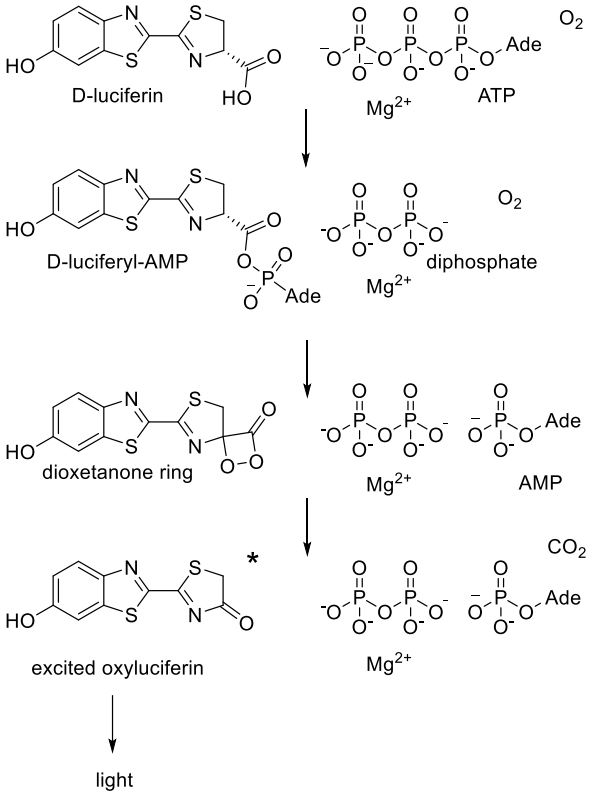


Figure 42: Reaction mechanism of firefly luciferase catalyzed bioluminescence. Adenylation leads to a luciferyl-adenylate intermediate via the release of diphosphate. Nucleophilic attack of molecular oxygen follows, resulting in the formation of an unstable dioxetanone with AMP as the leaving group. Spontaneous breakdown to excited oxyluciferin releases CO2. It decays to the ground state and emits a photon (*149, 150*).

A variety of commercially available kits are available for quantifying diphosphate. These kits are ready-to-use, containing all the necessary materials for converting diphosphate to ATP and the subsequent luciferase assay. In this study, the PPiLight™ inorganic diphosphate assay kit provided by Lonza Group was evaluated to measure diphosphate produced equimolar to CBGA-C5 based on the luminescence generated by luciferase.

2.3.6.1 Compatibility, Adaption, and Optimization of NphB assay substrates

Initially, the previously established conditions for NphB assays were employed in the luminescence measurements to ascertain the applicability of the reaction temperature and shaking frequency to the kit. During this process, the first optimization step arose, namely the different temperature optimums of the two enzymes, NphB and firefly luciferase. Utilizing a temperature of 37 °C, which is normally used in NphB assays due to NphB's temperature optimum, resulted in a decline in signal intensity after 15 minutes, attributable to the aggregation of firefly luciferase above 30 °C (*151*). Conversely, other studies have identified a maximum intensity for firefly luciferase at 25 °C. Therefore, it was determined that this temperature would be employed for the screening method, as the utilized kit has a sensitivity range of 0.02 to 10 µM diphosphate (*152*). In this setting, a stable luminescence signal was monitored.

Next, the optimal substrate concentration for a parallel NphB assay and luciferase assay, without causing interference with each other was determined. It was observed that several substances interfere with the bioluminescence assay due to an inhibitory effect. A structure-activity relationship study discovered that firefly luciferases are inhibited by small, linear, and planar structures (*149, 153*). OA has been found to be consistent with this inhibitor profile. Consequently, all substrates were added to the bioluminescence assay at a concentration similar to that previously used in NphB assays with the addition of a diphosphate standard solution using $Na_2H_2P_2O_7$ (Figure 43).

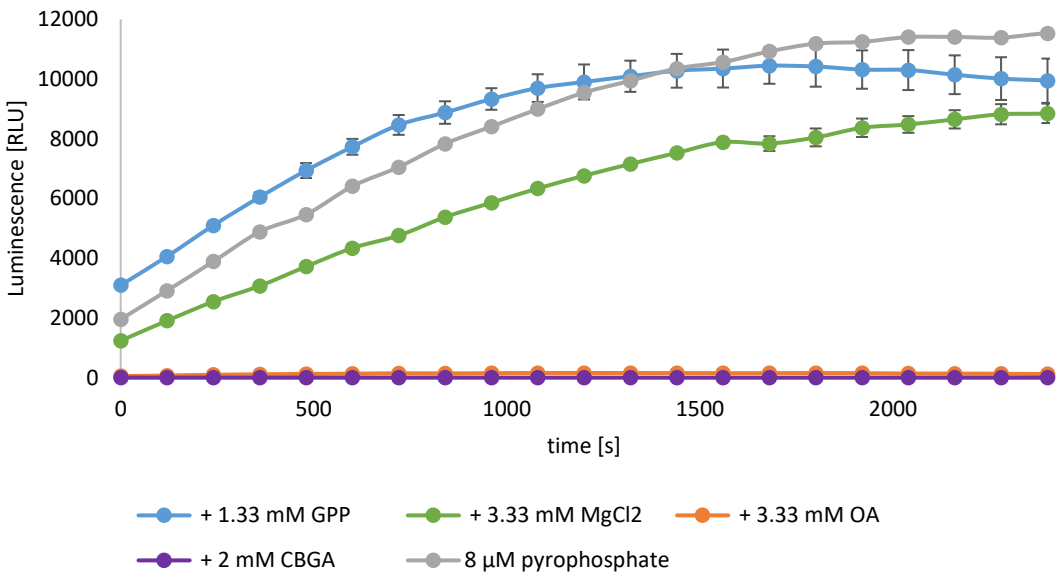


Figure 43: Luminometrically measured diphosphate standard solutions using $Na_2H_2P_2O_7$ (7.2 µM) supplemented with the substrates GPP, $MgCl_2$, olivetolic acid or the product CBGA applying PPiLight™ continuous kinetic protocol. As a reference, a standard with 8 µM diphosphate is shown. Detection at 25 °C and 1000 rpm for 40 min using FLUOstar® Omega Microplate Reader. Each measurements were performed in triplicates.

As anticipated, GPP does not appear to interfere with the luciferase assay; rather, it enhances the signal due to its inherent instability in an aqueous environment, resulting in its hydrolysis into phosphates and subsequent utilization as a phosphate source (*154*). To ensure the assay's reliability, a negative control containing GPP is included in each measurement series, allowing for the consistent comparison of signals.

It has been observed that both OA and CBGA-C5 result in the complete elimination of luminescence. One potential explanation for this phenomenon is the positioning of OA and CBGA-C5 within the active site of the luciferase, which results in a reaction with ATP towards the corresponding adenylate adduct, preventing the conversion of luciferin and the emission of light (*155*). To avoid inhibition, lower OA and CBGA-C5 concentrations were screened (Figure 44).

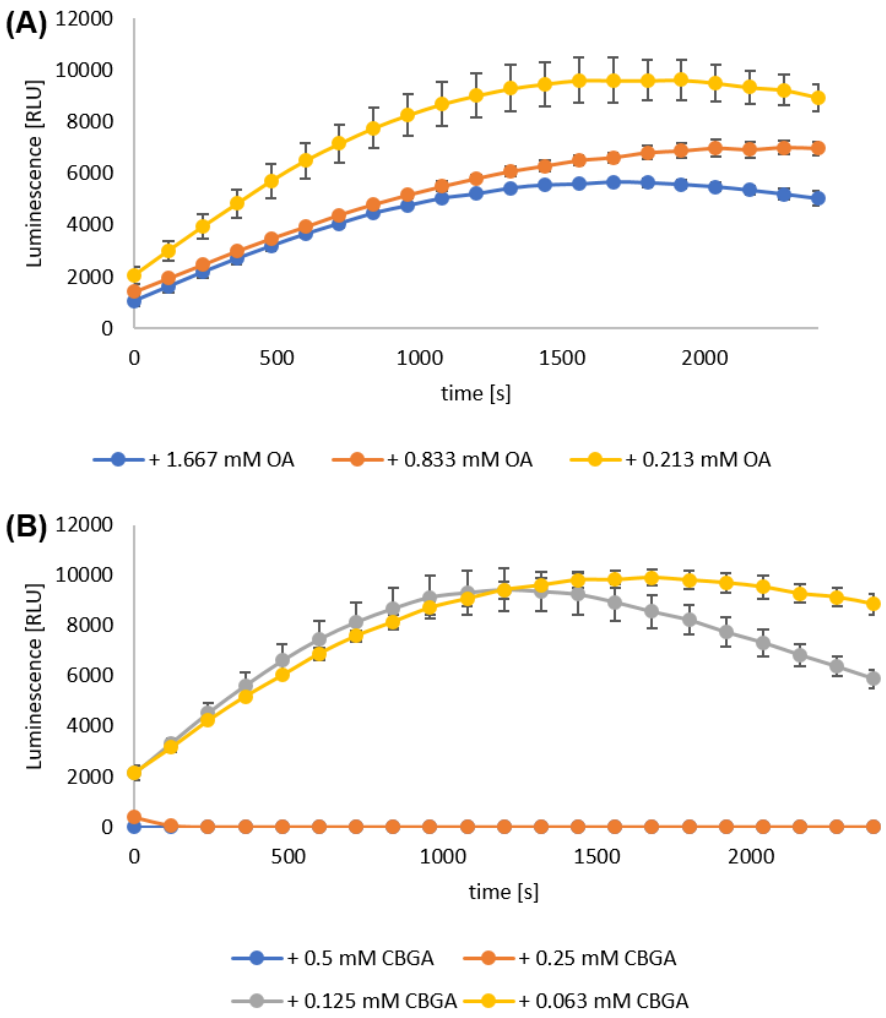


Figure 44: Luminometrically measured diphosphate standard solutions using $Na_2H_2P_2O_7$ (7.2 - 7.6 µM) supplemented with the substrate olivetolic acid (A) and the product CBGA-C5 (B) applying PPiLight™ continuous kinetic protocol. Detection at 25 °C and 1000 rpm for 40 min using FLUOstar® Omega Microplate Reader. Each measurements were performed in triplicates.

For OA, a stable luminescence signal was obtained when the concentration was lowered to 0.2 mM, which is the highest concentration that should be used in NphB assays. Lower concentrations should not be interfere with the luciferase assay kit.

In the case of CBGA-C5, a concentration of 0.125 mM yielded a modest effect on the luminescence signal; however, a more stable signal was observed at 0.063 mM. Given the inherent detection limits of the assay kit, which preclude the detection of 10 µM diphosphate, it can be deduced that CBGA-C5 production levels are likely to be lower. Consequently, a lower product concentration is required to remain the equimolar diphosphate concentration within the detection limit of the assay kit.

2.3.6.2 Product detection of i*n vitro* NphB assay using luminescence detection

Before implementing the luminescence assay in whole cell assays, it is necessary to test the feasibility of running NphB and luciferase assays in parallel in an *in vitro* assay. In order to do so, the buffer in which NphB is stored post-purification must be adjusted to avoid influencing the luciferase reaction. The optimal buffer composition for this purpose was determined to be one containing 50 mM TRIS and 10% glycerol, which exhibited minimal interference with the luciferase reaction while maintaining adequate CBGA-C5 production. In contrast to the buffer utilized in Chapter 2.3.3, NaCl was not added, as NaCl interferes with the luciferase reaction. Although CBGA-C5 production is reduced, it is still within the detection range of the diphosphate kit (Figure 45).

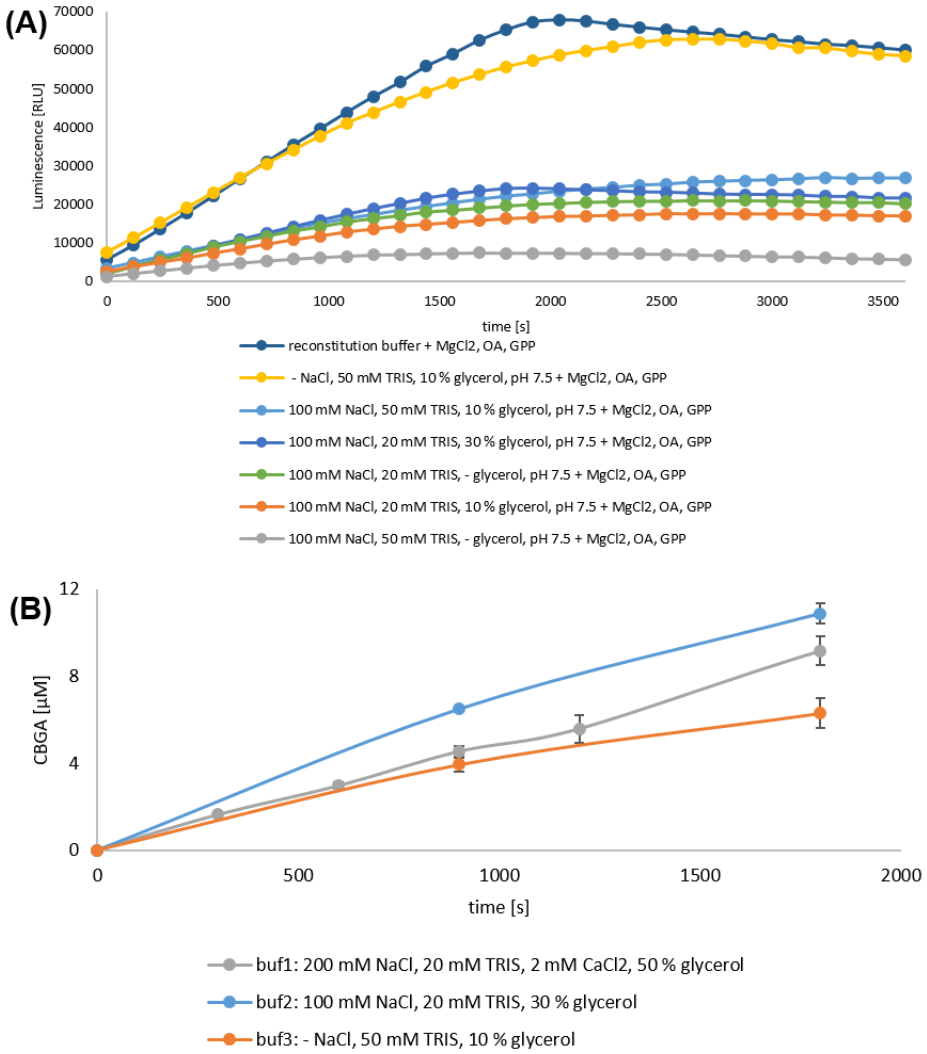


Figure 45: (A) Luminometrically measured buffers free of glycerol or NaCl mixed with reaction components (1 mM GPP, 0.1 mM olivetolic acid, 5 mM $MgCl_2$) applying PPiLight™ continuous kinetic protocol. Detection at 25 °C and 1000 rpm for 1 h using FLUOstar® Omega Microplate Reader. (B) CBGA-C5 production of NphB G286S/Y288A in the course of time within 30 min. Assays were performed using 1 mM GPP, 0.1 mM olivetolic acid, 5 mM $MgCl_2$, 19 ng μL^{-1} protein solved in buffer 1, 2 or 3.

Following the primary elimination of potential errors, the luciferase assay was performed in parallel with the NphB assay using purified enzyme. For the CBGA-C5 production, which is in the detectable sector of the diphosphate quantification kit, a protein concentration of 19 ng μL^{-1} was added to each assay. In the first assay, the detected luminescence of the samples did not differ from the sample without protein but with GPP. Consequently, the natural decay of the substrates still overpowers the amount of diphosphate gained by the assay. A screening was performed using different concentrations of GPP (Figure 46).

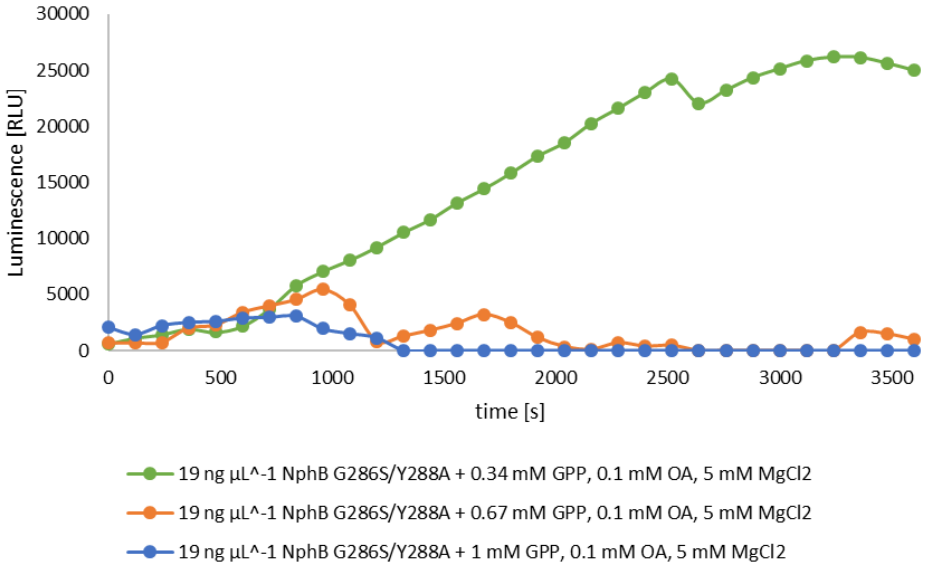


Figure 46: Background subtracted luminescence signal related to diphosphate production in NphB assays using 19 ng μL^{-1} protein in buffer 3 and varying concentrations of GPP, 0.1 mM olivetolic acid, 5 mM $MgCl_2$ within 1 h. Each measurements were performed in triplicates.

A comparison of the background-subtracted luminescence signals reveals that only the assays utilizing 0.34 mM GPP led to a stable signal, exhibiting a linear increase over time. In this assay, the signal is distinctly distinguishable from the signal without protein. The optimal conditions for achieving the best results when running the two assays in parallel are a buffer consisting of 50 mM Tris, 10 % glycerol, a substrate composition of 0.34 mM, 0.1 mM OA, and 5 mM $MgCl_2$.

The next step was to compare different NphB variants to see if they also differ in luciferase assay signals comparable to the CBGA-C5 concentration measured by HPLC-UV. The set of variants to be screened includes the wild type, which produces CBGA-C5 in a low concentration; NphB Q295F, which produces 6.6-fold higher CBGA-C5; and the two best-performing variants from the Crude-Cell-Extract (CCE) assay: NphB G286S/Y288A, which produces 10.6-fold higher CBGA-C5, and A232/Y288V, which produces 10.1-fold higher CBGA-C5 (Figure 47).

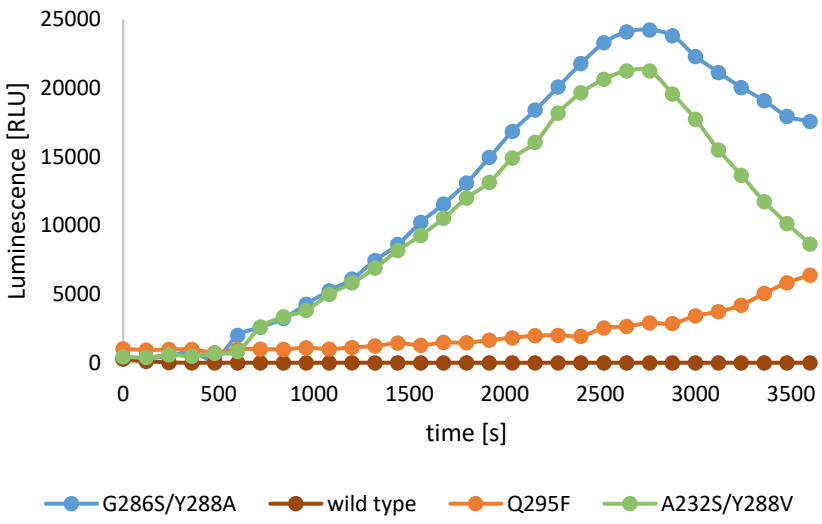


Figure 47: NphB variant screening: background subtracted luminescence signal related to diphosphate production in NphB assays using 20 ng μL^{-1} protein in buffer 3 and 0.34 mM GPP, 0.1 mM olivetolic acid, 5 mM $MgCl_2$ within 1 h. Each measurements were performed in triplicates

The assay revealed a distinct difference between the variants. For the wild type, no signal was detected following the subtraction of background noise. In this instance, the production of CBGA-C5 or 2-*O*-GOA was found to be inadequate, and the cleaved diphosphate could not be measured. Conversely, for the NphB Q295F variant, a signal was detected, suggesting the presence of low product levels. The two variants demonstrating the highest performance level exhibited quantifiable levels of diphosphate, with NphB G286S/Y288A registering the highest values and NphB A232S/Y288V registering the second-highest. The activity of the proteins can be ranked as follows: G286S/Y288A > A232S/Y288V > Q295F > wild type. This ranking is analogous to the *in vitro* results reported by other sources. Consequently, the identification of a high-performing variant is feasible.

2.3.6.3 Adaption for Cell Assay for Diphosphate detection

Adapting this assay as a high-throughput system for the utilization of CCE is of high importance, given that the purification of each NphB variant for *in vitro* assays is time-consuming. Consequently, a potential adaptation of CCE assays was investigated. Initially, the signal of a standard was measured in the presence of CCE containing expressed NphB (Figure 48).

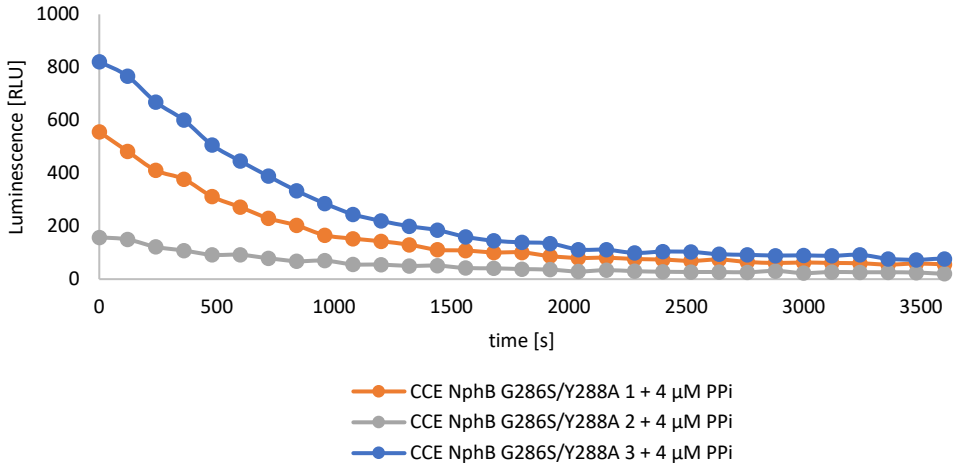


Figure 48: Luminometrically measured CCE supplemented with diphosphate standard solutions using $Na_2H_2P_2O_7$ (4 µM) applying PPiLight™ continuous kinetic protocol. Shown are the triplicates raw data. Detection at 25 °C and 1000 rpm for 1 h using FLUOstar® Omega Microplate Reader.

In this instance, the signal decreases over time in contrast to the other spectra. Firstly, following the separation of the sample from cell debris, the presence of substances such as nucleic acids, additional proteins, and yeast metabolites within the solution persists (*156*). A particular concern is the presence of cellular ATP within the solution, as this molecule is a substrate of the luciferase assay (*157*). Additionally, yeast-expressed enzymes, such as pyrophosphatases, can convert diphosphates prior to quantification (*158*). Other enzymes such as prenyltransferases, could also lead to an increased diphosphate concentration in the assay sample (*159*). The intracellular concentrations of sodium and potassium in *S. cerevisiae* also exert a substantial influence on the luciferase assay. As demonstrated in Chapter 2.3.6.2, the NaCl concentration exerts a strong influence on the stability of the luminescence signal. Therefore, the 30 mM sodium present in the *S. cerevisiae* cell interferes with the diphosphate quantification assay (*160*).

Even though the results were not promising, the parallel execution of CCE NphB assays and diphosphate quantification employing firefly luciferase was conducted (Figure 49)

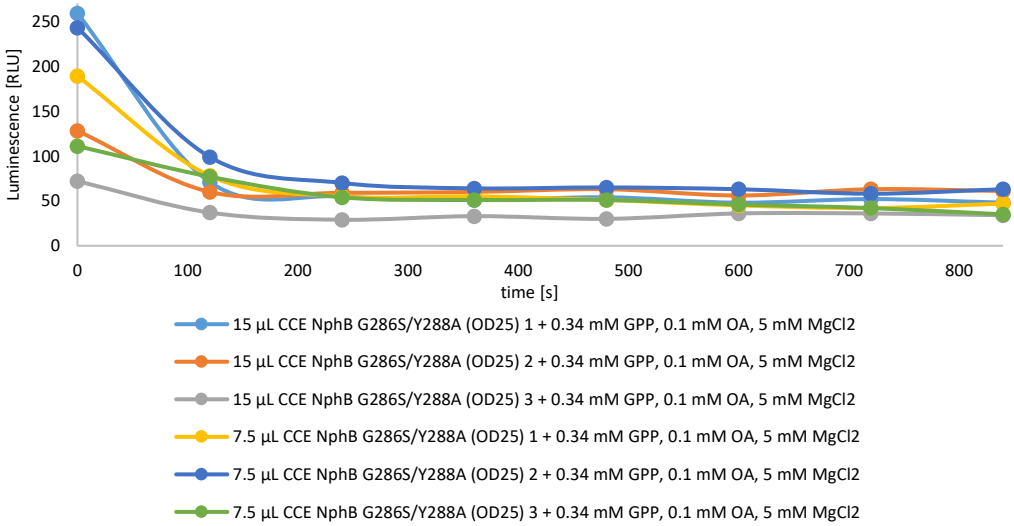


Figure 49: Luminometrically measured CCE assays (0.34 mM GPP, 0.1 mM olivetolic acid, 5 mM $MgCl_2$) using 15 µL or 7.5 µL CCE in buffer 3 applying PPiLight™ continuous kinetic protocol. Shown are the triplicates raw data. Detection at 25 °C and 1000 rpm for 14 min using FLUOstar® Omega Microplate Reader. Each measurements were performed in triplicates.

Here again, the detected luminescence signal was very low, even in the raw data and before any background noise was substituted. Consequently, the quantification of diphosphate or NphB activity remained unmeasurable using the diphosphate quantification with a luciferase. Prior to implementing the assays, a solution for a better preparation protocol of the CCE must be found.

Chapter 3 Conclusion and Outlook

3.1 Increasing olivetolic acid production through protein engineering of *Cs*OAC

In the first step of cannabinoid biosynthesis, olivetolic acid is synthesized from glucose and hexanoic acid precursors, with the two proteins *Cs*OLS and *Cs*OAC. The process entails an initial elongation towards a linear tetraketide through *Cs*OLS, followed by a second step involving the cyclization of this intermediate through *Cs*OAC. Transferring these proteins into *S. cerevisiae*, a chemical reaction from the linear tetraketide towards olivetol is monitored. To prevent this reaction, the goal was to enhance the productivity of the enzyme *Cs*OAC.

The production of OA was increased, and the production of OL was decreased using novel protein variants of *Cs*OAC. Computer-assisted protein engineering was used as a tool to identify potential *Cs*OAC variants. First, the binding pose of the linear tetraketide was identified using protein-ligand docking, followed by the identification of potential positions for mutagenesis. Combining this with a HotSpot wizard analysis led to the identification of six promising positions as HotSpots with I73, G82, R86, W89, L92, and I94 . Using *S. cerevisiae* as a host, an alanine scan was performed, validating all positions for further analysis.

A subsequent examination of these six positions was conducted through *in silico* analysis, with the selection of amino acids for exchange guided by considerations of protein-ligand interaction and homologous proteins. The initial evaluation of the diverse *Cs*OAC variants involved plasmid-based integration into *S. cerevisiae*, yielding a total of 36 *Cs*OAC variants for assessment. To evaluate OA production, a strain was employed that contained a single-copy integration of *Cs*AAE1 and a multi-copy integration of *Cs*OLS, ensuring a high concentration of the *Cs*OAC substrate. Four amino acid exchanges resulted in a significant increase in OA concentrations compared to the native enzyme, nine variants overall that are of interest for subsequent experiments. The variants demonstrating the greatest efficacy were identified as CsOAC G82A, which exhibited a 1.2-fold increase, and L92Y, which exhibited a 1.4-fold increase in OA production.

Following the single screening, the effect of combining the amino acid exchange with a positive effect on OA production was examined. Therefore, all possible combinations, including two, three, and four amino acid exchanges, were analyzed with protein-ligand docking. The results of this analysis, in conjunction with the *in vivo* results previously mentioned, guided the screening of 43 *Cs*OAC variants containing multiple exchanges *in vivo* using *S. cerevisiae* as the

host. The screening process yielded five combinations that demonstrated a substantial enhancement in the ratio between OA and OL. The optimal outcomes were attained with G82A/L92Y and G82A/R86H/L92Y, which resulted in a 1.1-fold increase in OA production.

In the final experiment, the two proteins, *Cs*OAC and *Cs*OLS, were single-copy integrated into *S. cerevisiae* to validate the OA and OL production with a similar number of proteins. Similar to the other screening, a significant increase in OA concentration of 1.2-fold was measured using the variants *Cs*OAC G82A and I94V. The variant that demonstrated the most favorable performance, involving a single amino acid exchange, was identified as *Cs*OAC L92Y, which resulted in a 1.3-fold enhancement in OA production. The variant that exhibited the most optimal performance in this series of experiments was *Cs*OAC, which resulted in a 1.7-fold enhancement in OA production and a shift in the OA amount in the product composition from 0.345 ± 0.006 to 0.480 ± 0.006. The variant fulfilled the effect that OA production in *S. cerevisiae* is improving while OL production is decreased, leading to more efficient utilization of the important precursors such as glucose, acetyl-CoA, malonyl-CoA, or hexanoic acid.

In the future, it would be interesting to integrate the variant into a *S. cerevisiae* strain, where the complete cannabinoid biosynthesis is integrated, to see if the positive effect of a higher OA production also reflects in a higher concentration of the following downstream products, such as CBGA-C5 and THCA-C5. An increase in titer is expected, resulting in a higher yield from the substrates. Since OL concentration was also decreased through the increase in the copy number of *Cs*OAC, this effect needs to be tested in combination with the newly identified variant (*55*). If the two effects are additive, the unwanted production of OL could be further increased, which can positively affect the cannabinoid titer.

Using the chosen setup with direct *in vivo* assay and *S. cerevisiae* as host, it is not possible to distinguish if the positive effect regarding OA production results from an influence of an improved enzyme kinetic or an increase in expression in *S. cerevisiae*. Further experiments on enzyme kinetics using *in vitro* assays could provide interesting aspects (*37*). The expression in *S. cerevisiae* can be further monitored with fusion to a fluorescence tag, which allows monitoring of the expression of the protein during the complete cultivation process. Here, a direct comparison with the wt is also possible.

3.2 Synthesis of non-natural cannabinoids using cannabinoid biosynthesis in *S. cerevisiae*

The interest in novel cannabinoids is still emerging. A notable benefit of the heterologous biosynthesis in *S. cerevisiae* is that the enzymes derived from this biosynthesis can be utilized, while a minimal structural alteration of the substrate results in the formation of different cannabinoid analogs. As demonstrated in the studies conducted by Luo et al., the modification of substrates to be fed from hexanoic acid to another carboxylic acid results in the formation of the analogous structure of all intermediates up to THCA derivatives (*25*). This first experiment evaluates the uptake and conversion of different carboxylic acid structures with a *S. cerevisiae* strain conducted in the chair of technical biochemistry. Only the first enzymes of the biosynthesis including *Cs*AAE1, *Cs*OLS, and *Cs*OAC, are implemented in the strain to focus only on the conversion of these enzymes. Simultaneously, the strain exhibits the capacity to produce hexanoic acid, a consequence of the integration of the *FAS2* G1250S mutation. A broad library of different carboxylic acids, including structure elements of different alkanes, aromatic structure elements, and previously tested structures, is screened.

This study confirmed the conversion of octanoic acid, heptanoic acid, 6-heptynoic acid, trans-2-hexanoic acid, and isocaproic acid towards their respective OA and OL derivatives was confirmed using HPLC-MS. The examination of this strain revealed no occurrence of conversion of structures containing amino groups, halogen moieties, hydroxy groups, or aromatic structural elements.

Several possible limitations could be why the number of converted structures is low. First, structures with a low binding affinity to the proteins compared to hexanoic acid could not be identified since the competition between the two acids is too demanding, and the second product cannot be identified. To avoid this problem, the conversion of all structures without product formation using the strain should be tested again. However, this time, it should be done with a *S. cerevisiae* strain, which cannot produce hexanoic acid.

A second limitation can be the activation of the fatty acid. *Cs*AAE1, the enzyme in this study is known to activate C_6-C_9 fatty acids (*20*). This includes the length of all acids identified for conversion towards OA and OL analogs. For other structures, this enzyme needs to be exchanged. In *C. sativa,* short- and medium-chained fatty acids are also converted towards the respective cannabinoid derivatives. Therefore, the conclusion that the fatty acid activating step

is not catalyzed here is conclusive. In the plant, a second acyl-activating enzyme, converting short- and medium-chained fatty acid structures, is identified with the acyl-activating enzyme 3 (*Cs*AAE3) (*20*). Repeating the screening using *Cs*AAE3 for the not converted fatty acids, especially the short-chained acids, is interesting. Another acyl-activating enzyme that could show potential for the carboxylic acids containing an aromatic structure element is the phenylacetate-CoA ligase (*Pc*PCL) from *Penicillium chrysogenum*. In its host organism, it catalyzes the activation of phenylacetic acid (*161*). Therefore, a high potential for successfully activating aromatic structures seems to be a potential. Since these two enzymes were also already tested in the activation of hexanoic acid in *S. cerevisiae*, activity in this host is already proven and ensured (*55*). Next to these two enzymes, there is a high number of other enzymes that can be tested in a screening. A problem is the difference if the activation step is not catalyzed or if the activated structure is not accepted as a substrate of *Cs*OLS. Therefore, it would be helpful to use only acyl-activating enzymes, which are proven to convert the tested acid.

As described before, longer-chained acids are not converted from *Cs*OAC and *Cs*OLS. However, with the engineering of the binding pocket, dodecanoic acid was successfully converted towards their respective OL and OA derivative (*109*). These variants can also be tested in the conversion of longer-chain fatty acids and bulkier structures. Also, this approach shows the potential of engineering the binding pocket for further conversion. With *in silico* analysis of the binding pocket and screening of different variants, the uptake of a novel set of carboxylic acids could be possible.

In a second step, the downstream steps of cannabinoid biosynthesis were tested for the conversion of novel structures towards cannabinoid derivatives. After the OA synthesis, the formation of CBGA-C5 follows. OA derivatives were tested using the enzymes *Cs*PT4 (*25*) and NphB. For NphB, only structures containing the natural building blocks were tested. In this thesis, the conversion of novel OA analog using *in vitro* NphB assays was analyzed. To identify promising OA derivatives, which are likely converted, a docking study was conducted using two different NphB variants, namely the wt and the best-performing variant G286S/Y288A. The binding mode was analyzed, and only structures with the correct orientation of the resorcylate core were chosen for *in vitro* testing. Interestingly, a side chain variation from C_0 to C_8 did not significantly influence the binding mode with the variant G286S/Y288A. The binding mode of molecules of OA derivatives containing a second aromatic structure element was not that promising regarding orientation, except for the benzyl group. Eight structures were chosen to be synthesized and were then screened with *in vitro*

NphB assays regarding their conversion. Here, the highest conversion of novel OA analogs was achieved using the analog containing ethyl, *n*-octyl, and a benzyl group. The potential CBGA derivatives were further analyzed using HPLC-MS/MS and NMR, confirming the prenylation on the correct position. For another confirmation and the testing of the next cannabinoid biosynthesis step, the three most promising OA derivatives were incubated with NphB and *Cs*THCAS in parallel, resulting in the formation of their respective THCA and CBCA derivatives.

As a next step, the feeding of the most promising OA derivatives to a *S. cerevisiae* strain containing NphB and *Cs*THCAS should be tested to see if an *in vivo* conversion of these structures is also possible, which is of high importance for an implementation of these structures into the cannabinoid biosynthesis in yeast. When the respective fatty acid to the OA derivatives is also converted and identified in the experiments before, synthesis of the novel THCA derivative from the fitting carboxylic acids for these three structures could also be possible. With this, the complex synthesis of the OA derivatives can be avoided, and a cost-efficient and fast way to gain novel cannabinoid structures can be used.

For these derivatives, which are only converted in low percentages from the NphB variant G286S/Y288A, it would be helpful to start another approach of protein engineering to identify a variant in which the binding pocket is adjusted to the OA derivative. To simplify the screening and engineering of the NphB variant, a novel high-throughput system was tested, where the CBGA-C5 concentration should not be measured with HPLC-UV but over a diphosphate quantification using firefly luciferase. After adjusting the screening parameters, it was possible to measure the differences in CBGA-C5 production of the NphB variants *in vitro*. Still, the transmission towards CCE assays is limited due to many factors inside the *S. cerevisiae* cell that interfere with the luciferase assay. Here a different extraction method of the diphosphate is needed.

Chapter 4 Materials and Methods

4.1 Materials

The chemicals used for this work were mainly purchased from Carl Roth GmbH & Co. KG (Karlsruhe, Germany), Sigma-Aldrich® (Darmstadt, Germany), Invitrogen (Karlsruhe, Germany), Merck (Darmstadt, Germany), and VWR (Darmstadt, Germany) if not stated elsewhere.

Restriction enzymes were obtained from New England Biolabs® (Frankfurt am Main, Germany), and the protocols provided were used for application.

DNA isolation and purification were performed using Kits from Macherey-Nagel GmbH & Co. KG (Düren, Germany). For plasmid isolation from *E. coli*, the Nucleospin® Plasmid (No Lid) Kit, for purification of PCR product or of DNA from agarose gel, the NucleoSpin® Gel and PCR Clean-Up Kit and for gDNA isolation from *S. cerevisiae* the NucleoSpin® DNA Yeast Kit was used.

GPP was synthesized in-house after Woodside et al. (*162*).

The olivetolic acid analogs were kindly synthesized and provided by Jasmin Wloka (*108*). Purity and structure were analyzed by ^{1}H-NMR and ^{13}C-NMR.

Diphosphate assays were carried out using the PPiLight™ inorganic pyrophosphate assay kit from Lonza (Basel, Switzerland).

4.2 Methods

4.2.1 Microorganisms

For all experiments, including plasmid amplification, the *E. coli* strain DH5α was used. The cells were cultivated at 37 °C and 200 rpm for 16 -18 h in Luria Bertani (LB) liquid medium (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 10 g L^{-1} NaCl, pH 7, for plates: 20 g L^{-1} agar) with the addition of the antibiotic Ampicillin (100 µg mL^{-1}) if needed for selection.

For all other experiments, the host organism was *S. cerevisiae*. The base strain on whom all other strains used in this thesis are based is *S. cerevisiae* Δpep4 Δgal1 Δgal80. The cells were cultivated at 30 °C and 200 rpm for 48 h using Yeast Peptone Dextrose medium (YPD, 20 g L^{-1} peptone, 10 g L^{-1} yeast extract, 20 g L^{-1} glucose (20 g L^{-1} agar for plates)). For cells containing a plasmid, cultivation conditions were similar. However, the media was exchanged towards synthetic mineral salt medium without leucine or uracil (6.7 g L^{-1} yeast nitrogen base without amino acids,

1.92 g L^{-1} yeast synthetic drop-out medium, 20 g L^{-1} glucose, for plates: 20 g L^{-1} agar) was used. All strains used in this thesis are listed in Table 9.

Table 9: List of all strains used in this study including their relevant genotype and the application. The strains are sorted according to their usage in their certain projects.

Strain	Relevant genotype	Application	Reference
***E.coli* DH5α**	*F−endA1, hsd*R17(*rk*−*mk*+) *supE44, thi-l λ−recAI gyrA96 relA1 φ80⊿lacAm15*	Plasmid amplification	Life Technologies; Darmstadt, Germany
***S. cerevisiae* Δpep4 Δgal1 Δgal80**	*S. cerevisiae* CEN.PK2-1C *⊿pep4 ⊿gal1 ⊿gal80*	Start organism for strain development, NphB expression	Euroscarf
Chapter 2.1			
ySP01	*S. cerevisiae* CEN.PK2-1C *⊿pep4 ⊿gal1 ⊿gal80 yprcτ3::pTEF1-**CsAAE1**-tCYC1*	*Cs*OAC-*Cs*OLS Linker screening	This study
ySP02	*S. cerevisiae ⊿pep4 ⊿gal1 ⊿gal80 ty4::pCCW12-**CsOLS**-tENO2 URA3d*		This study
ySP03	ySP02 *yprcτ3::pTEF1-**CsAAE1**-tCYC1*	*Cs*OAC variant screening	This Study
ySP04	ySP01 *trp1::pCCW12-**CsOLS**-tENO2 pTDH3-**CsOAC**-tTDH1*	*Cs*OAC variant screening	This Study
Chapter 2			
yAP01	*S. cerevisiae* CEN.PK2-1C *⊿pep4 ⊿gal1 ⊿gal4mig1b ty4::**CsOLS/CsOAC-CsOAC** YPRC⊿15::pGal10-**ERG20ww**-tADH1.pGal1-tHMG1-tCYC1 **FAS2-G1250S** URA3::KanMX pCCW12-**CsOLS**-tENO2 pTDH3-**CsOAC Y27F**-tTDH1 yprcτ3::pTEF1-**CsAAE1**-tCYC1*	Screening feeding of acids	A. Prima
Chapter 3			
HAC1	*S. cerevisiae ⊿pep4 ⊿gal1 ⊿gal80 yprcτ3::pHHF2-**ScHAC1s**-tADH1*	Expression of *Cs*THCAS	C. Schmidt

4.2.2 Plasmid construction

In this study, two plasmid backbones were used for all conducted experiments. For plasmid-based protein expression, the pDioynsos (*163*) plasmid was utilized, and for genomic integration and multi-copy integration, plasmid backbones from pCfB2791 (Addgene plasmid #63654 was a kind gift from Irina Borodina and Jerome Maury) was utilized (*131*). All plasmids are listed in Table 10. The construction of the plasmids was performed using the Gibson Assembly method (*164*). For the creation of mutants, site-directed mutagenesis was applied (*165*). The composition of the plasmid backbones as vector maps, the amino acid and nucleotide sequence of *Cs*AAE1, *Cs*OLS, *Cs*OAC, NphB, and *Cs*THCAS, and the applied primers are shown in Figure S1 and S2 and in Table S1 - S8.

4.2.2.1 Gibson assembly

For Gibson assembly, the chosen DNA insert and vector fragments were amplified by polymerase chain reaction (PCR) using Q5® High-Fidelity 2X Master Mix (NEB, Frankfurt am Main, Germany). The primer for insert amplification were constructed, resulting in a 25 bp overhang to the vector. The PCR products were digested using DpnI and then purified using a PCR purification kit. The Gibson assembly was performed using the 2x Gibson Assembly Master Mix (NEB, Germany). The reaction mix was transformed into *E. coli* DH5α using heat shock transformation using the heat shock method from Inoue et al (*166*). After plasmid isolation, correct sequences were verified with sequencing (Microsynth Seqlab GmbH, Göttingen, Germany).

4.2.2.2 Site-directed mutagenesis

To obtain the *Cs*OAC variants, a one-step site-directed mutagenesis protocol was applied for cloning (*165*). According to the given primer design guideline, up to six nucleotides can be exchanged in one PCR. PCR was executed with Q5® High-Fidelity 2X Master Mix (NEB, Frankfurt am Main, Germany) using a special protocol, where 12-18 cycles of elongation at annealing temperature 5 °C lower of non-complementary part of primer and 1-12 cycles of elongation with annealing temperature of 5 °C lower than Tm of the complementary part was performed. PCR products were digested using DpnI at 37 °C overnight. Competent *E. Coli* DH5α was transformed with 2 µL of the PCR reaction mixtures via heat shock. The resulting plasmids were isolated and, for validation of the sequence, sent to routine sequencing.

2.2.2.3 Plasmid transformation

Plasmids were transformed into *S. cerevisiae* using the LiAC/PEG-Method (*167*). For- URA transformations, the cells were subsequently plated on a synthetic mineral salt medium lacking uracil and incubated at 30 °C for 2-3 days. For -Leu transformations, the cells were resuspended in 1 mL YPD medium and incubated for 4 h at 30 °C. Afterward, the cells were plated out on a synthetic mineral salt medium without leucine and incubated at 30 °C for 4-5 days.

A list of all plasmids transformed into *S. cerevisiae* including the strain and the Expression cassette are shown in Table 10.

Table 10: List of all plasmids, that were transformed into *S. cerevisiae* including the transformed strain and the expression cassette.

Plasmid	**strain**	**Expression cassette**	**Reference**
pDionysos	-	-	Stehle(*163*)
Chapter 2.1			
pDio-OAC	-	pGal1-OAC	This study
pDio-OLS	-	pCCW12-OLS	This study
pDio-pCCW12-OLS-pTDH3-OAC	ySP01	pCCW12-OLS pTDH3-OAC	L. Reckert
pDio-OLS-OAC	ySP01	pCCW12-OLS-*Cs*OAC	L. Reckert
pDio-OLS-GGGGS-OAC	ySP01	pCCW12-OLS-GGGGS-OAC	L. Reckert
pDio-OLS-EAAAK-OAC	ySP01	pCCW12-OLS-EAAAK-OAC	L. Reckert
pDio-OLS-$(GGGGS)_3$-OAC	ySP01	pCCW12-OLS$(GGGGS)_3$OAC	L. Reckert
pDio-OLS-$(EAAAK)_3$-OAC	ySP01	pCCW12-OLS$(EAAAK)_3$OAC	L. Reckert
pDio-OLS-$(GGGGS)_5$OAC	ySP01	pCCW12-OLS$(GGGGS)_5$OAC	D. Kempel

pDio-OLS-$(EAAAK)_5$-OAC	ySP01	pCCW12-OLS $(EAAAK)_5$OAC	D. Kempel
pDio-OAC-OLS	ySP01	pCCW12-OAC-OLS	D. Kempel
pDio-OAC-$(GGGGS)_5$-OLS	ySP01	pCCW12-OAC $(GGGGS)_5$OLS	D. Kempel
pDio-OAC-$(EAAAK)_5$-OLS	ySP01	pCCW12-OAC$(EAAAK)_5$OLS	D. Kempel
Chapter 2.3			
pDio-NphB	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB	This study
pDioynsos-NphB-His-Tag	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB His-Tag	This study
pDio-NphB G286S/Y288A	*S. cerevisiae* Δpep4 Δgal1 Δgal80	*S. cerevisiae* Δpep4 Δgal1 Δgal80	This study
pDio-NphB G286S/Y288A-eGFP	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB G286S/Y288A-eGFP	This study
pDio-*Cs*THCAS	*S. cerevisiae* HAC1	pGal1-THCAS	C. Schmidt
pDio-NphB G286S/Y288A-His-Tag	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB G286S/Y288A His-Tag	This study
pDio-NphBQ295F-His-Tag	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB Q295F His-Tag	M. Donner
pDio-NphB A232S/Y288V-His-Tag	*S. cerevisiae* Δpep4 Δgal1 Δgal80	pGal1-NphB A232S/Y288V His-Tag	M. Donner

4.2.3 Genomic integration

For single-copy integration into *S. cerevisiae* strains, modified plasmids from the EasyCloneMulti systems were used. With PCR and Q5® High-Fidelity 2X Master Mix (NEB, Frankfurt am Main, Germany), the integration cassette was amplified, including primers that created 50 bp overhangs homologous to the upstream and downstream regions of the flanking sequence.. From yeast gDNA, isolated with the NucleoSpin® DNA Yeast Kit, 500 - 1000 bp fragments from the upstream and the downstream flanking regions of the integration site were amplified using Q5® High-Fidelity 2X Master Mix. After the clean-up of the PCR fragments using the NucleoSpin® Gel and PCR Clean-Up Kit, concentration was determined using Nanodrop™. *S. cerevisiae* was transformed with 1 µg of equimolar amounts of the integration cassettes as well as the upstream and downstream flanking regions.

For multi-copy integration into the Ty4 locus, the integration cassette and the plasmids were digested with Not1 overnight (*131*). After digestion of the integration cassette, the DNA was purified using NucleoSpin® Gel and PCR Clean-Up Kit. 10 µg of the linearized integration cassette was transformed into yeast using LiAC/PEG method (*167*). For cell selection, the transformation mix was plated on a synthetic mineral salt medium without uracil and incubated at 30 °C for 2-3 days. For a comparison of copy numbers, the production of olivetol was analyzed using HPLC-UV.

All plasmids used for genomic integration are listed in Table 11.

Table 11: List of all plasmids used for genomic integration including the strain to integrated and the integration cassette.

Plasmid	Strain	Integration Casette	Reference
pCFB2791	-	-	AddGene (*131*)
pCFB2791-OLS	*S. cerevisiae* Δpep4 Δgal1 Δgal80	*pCCW12-**CsOLS**-tENO2 URA3d*	This study
pCFB2791-AAE1	*S. cerevisiae* Δpep4 Δgal1 Δgal80; ySP02	*yprcτ3::pTEF1-**CsAAE1**-tCYC1*	C. Schmidt
pCFB2791-OLS/OAC	ySP01	*trp1::pCCW12-**CsOLS**-tENO2 pTDH3-**CsOAC**-tTDH1*	C. Schmidt

4.2.4 *Cs*OAC & *Cs*OLS expression in *S. cerevisiae*

4.2.4.1 Plasmid-based Expression

For the expression of the different *Cs*OAC variants in *S. cerevisiae*, a cultivation approach was employed that involved two precultures and one main culture, using the transformed cells containing the relevant plasmid. For the first preculture, 10 mL synthetic mineral salt medium without leucine was inoculated with colonies from the plate. Following a 24-hour cultivation period at 30 °C and 200 rpm, 25 mL of synthetic mineral salt medium without leucine was inoculated to an optical density (OD_{600}) of 0.1 –0.2 using the first preculture. The second preculture was cultivated overnight at 30 °C and 200 rpm. For the main culture, 30 mL of YPD and 0.5 mM hexanoic acid were inoculated in a flask containing baffles to an OD_{600} of 0.5 using the second preculture. All main cultures were cultivated in triplicate for 48 h at 30 °C and 200 rpm. Afterward, the cell suspension was stored at -20 °C or directly prepared for High-Performance Liquid Chromatography (HPLC) measurements.

4.2.4.2 Expression of genomically integrated genes

For the cultivation of strain ySP04, the main procedure follows the method described in chapter 4.2.4.1. It differs in the media used for the preculture, where the cells were cultivated with YPD as the medium. The volume for the main culture also differs; here, 10 mL of YPD was used as the total volume and inoculated to an OD600 of 0.5.

4.2.5 Feeding of different acids into olivetolic acid producing strain

For testing the influence of the feeding of different acidic structures in the production of secondary metabolites, the strain yAP01 was used. The acids were dissolved in ddH_2O or ethanol in dependency on their solubility (Table 4). The cultivation protocol in Chapter 4.2.4.2 only differs in that, instead of 0.5 mM hexanoic acid, other acids were added to the main culture. Afterward, the cell suspension was stored at -20 °C or directly prepared for HPLC analysis.

4.2.6 *Cs*THCAS & NphB expression in *S. cerevisiae*

For the expression of NphB or *Cs*THCAS in *S. cerevisiae*, the transformed cells containing the correlating plasmid were cultivated using two precultures and one main culture. Within the first preculture, 10 ml of mineral salt medium without leucine was inoculated using multiple colonies from the plate and cultivated for 24 h at 200 rpm and 30 °C. The second preculture was inoculated using the first preculture to an OD600 of 0.1–0.2 and cultivated overnight at

30 °C for 24 h. In the main culture, the YPD medium was inoculated to OD600 = 0.6 using the second preculture and cultivated for 48 h at 150 rpm and 20 °C. Afterward, the cell suspension was stored at -20 °C or directly prepared HPLC measurements or protein purification.

4.2.7 Purification of NphB with Immobilized metal ion affinity chromatography (IMAC)

For the purification of NphB, the yeast cells were harvested by centrifugation (3000 g, 5 min, 4 °C), and the supernatant was discarded. For each gram of the cell dry pellet, 5 mL lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 30 mM Imidazole, pH 7,5) was used to resuspend the cells. To the cells, 2 mg/mL DNAse II was added. The cells were lysed using glass beads (0.75 - 1.00 mm) and vortexing (20 min, 4 °C, 2.000 g) or by French press® at a pressure of 20,000 psi. Glass beads were removed by centrifugation (3000 g, 5 min, 4 °C) following a second centrifugation (15.000 rpm, 30 min, 4 °C) to remove cell debris. The proteins were purified using the ÄktaPure System (Ge Healthcare, Solingen, Germany) and a HisTrap™ HP column or HisTrapTM FF column (Cytiva Europa, Freiburg im Breisgau, Germany). Depending on the amount of protein to purify, a 1 mL or 5 mL column was used. The pre-equilibration of the column was performed using 10 mL or 50 mL of binding buffer (50 mM Tris-HCl, 500 mM NaCl, 30 mM Imidazole, pH 7.5). Afterward, the filtered cell solution was loaded onto the column (1 mL/min or 5 mL/min flow rate). To remove other proteins from the column, a washing step with 10 mL binding buffer was performed afterward. The protein was eluted using a linear gradient of imidazole between 30 mM and 500 mM using the binding buffer and elution buffer (50 mM Tris-HCl, 500 mM NaCl, 500 mM Imidazole, pH 7.5). The protein was collected in 1 mL fractions, and the fractions with the highest UV absorbance were pooled. As the last step, the desalting was performed using a PD10 column (Cytiva Europa, Freiburg im Breisgau, Germany) to replace the buffer with protein storage buffer (50 mM Tris-HCl, 200 mM NaCl, 50% (v/v) glycerol, pH 7.5). Protein concentration was determined using NanoDrop™ One (Thermo Scientific™). The proteins were frozen in liquid nitrogen and then stored at -80 °C until use.

4.2.8 NphB *in vitro* activity assay

For *in vitro* assays, purified protein solved in storage buffer (50 mM Tris-HCl, 200 mM NaCl, 50% (v/v) glycerol, pH 7.5) was used. All assays were performed in triplicates. GPP was dissolved in ddH_2O, and the prenyl acceptor substrates were dissolved in DMSO. The assay mixture with a total of 120 µL contained 5 mM $MgCl_2$, 2 mM GPP, 1 mM prenyl acceptor, and

100 µg purified protein. The reaction mixture was incubated for 18 h at 37 °C and 1100 rpm in the dark. For the termination of the assays, 2.9 volumes of acetonitrile and 0.1 volume of formic acid were added. After incubation for 15 min on ice and centrifugation by 13,000 g and 4 °C for 30 min, the supernatant was stored at -20 °C until usage.

For the scale-up of the reactions, the concentrations of all components of the assay mixture were similar but adjusted to a higher volume. The chosen volume for this reaction was dependent on the amount of purified protein.

4.2.9 Crude cell extract NphB/THCAS

For crude cell extract assays, cultivation leading to the expression of NphB and cultivation leading to the expression of *Cs*THCAS are described in Chapter 4.2.6, both in separate flasks. From both cultivation approaches, each cells with an OD_{600} = 125 were harvested for one assay. Harvested cells were resuspended in 400 µL assay buffer (50 mM Tris-HCl, 100 mM NaCl pH = 7.5). Lysation was performed using glass beads (0.75 – 1.00 mm) and vortexing for 20 min at 4 °C and 2.000 *g*. To remove the cell debris, the cells were centrifuged for 20 min at 4 °C and 8.000 rpm. For the crude cell extract assays, 95 µL of each crude cell extract was combined, and substrates were added with 5 mM MgCl2, 2 mM GPP, and 1 mM of the aromatic acceptor, resulting in a total assay volume of 200 µL for each assay. After incubation for 24 h at 37 °C and 1100 rpm in the dark, the assays were terminated through the addition of 2.9 assay volumes of acetonitrile and 0.1 assay volumes of formic acid. "After incubation on ice for 15 min, the samples were centrifuged (13,000 g, 4 °C, 30 min), and the supernatant was stored by -20 °C until usage. All assays were performed in triplicates" (*108*)

4.2.10 Luminescence measurements

Luminescence measurements were carried out using the manual for continuous kinetics provided by Lonza Group (PPiLight™ inorganic pyrophosphate assay, Lonza Nottingham, Ltd.). For this measurement, the bioluminescence signal is directly proportional to the amount of diphosphate present in the protein assay. The signal increases at a steady and constant rate proportional to the diphosphate concentration over 60 minutes, resulting in a linear signal over time. The detection limit is 10 µM diphosphate in total.

The measurements were performed using the FLUOstar® Omega Microplate Reader (BMG Labtech, Ortenberg, Germany) with the following settings: The emission filter was set to 535 nm with a gain of 3600, top optics were applied, and a normalization time of 0.10 sec was

used. Before each measurement cycle, the plate was shaken in double orbital mode at 1000 rpm for 30 s. Overall, in one hour, 31 cycles with a time of 120 s were performed.

For simultaneously carried out NphB and luciferase assays, the reactions were performed in a 96-well-plate with a reaction volume of 40 µL and 20 µL of the converting and detection reagent (PPiLight™ inorganic pyrophosphate assay, Lonza Nottingham, Ltd.).

4.2.11 Preparative RP-HPLC

The isolation of the NphB reaction products was performed using extraction with acetonitrile. For a reduction of the solvent volume, probes were pooled, and the solvent was minimized under vacuum. Purification of the assay products was performed using preparative HPLC (Shimadzu, Duisburg, Germany), and separation was carried out using Nucleodur EC-C18 (125 × 21 mm, 5 µm; Macherey-Nagel, Düren, Germany). ddH_2O and acetonitrile were applied as mobile phases A and B, respectively. "The following method was used with a flow rate of 6 mL min 1 and an oven temperature of 40 °C: 30% B for 2 min, gradient from 30% to 100% B in 15 min, gradient from 100 % to 30% B in 2 min, 30% B for 6 min. The separated product peaks were detected in the DAD at 225 nm and collected. The solvent was removed, and the products were dried using SpeedVac. The compounds were submitted for ^{1}H- and ^{13}C-NMR measurements." (*108*)

4.2.12 Analytics - HPLC-UV detection and quantification

For sample preparation, a sample was taken from cell suspension, the cultivation medium, or a cell pellet with an OD600 = 125 resuspended in 100 µL ddH_2O and mixed with 2.9 volumes of acetonitrile and 0.1 volumes of formic acid.

Samples were incubated on ice for 15 minutes and then centrifuged at 4°C, 16,100 × *g*, for 30 minutes. The supernatant was then transferred into HPLC vials and, if needed, filtered (0.45 µm, nylon). Prepared assay samples were analyzed by HPLC-DAD (Agilent Infinity II 1260, Waldbronn, Germany), and the separation was achieved using a Poroshell 120 EC-C18 (2.1 × 100 mm, 2.7 µm; Agilent, Waldbronn, Germany) column. For the application, identification, and quantification of molecules, different methods were performed.

For the quantification of OA and OL, the mobile phase consisted of ddH2O + 0.1 % formic acid (A) and Acetonitrile (B). The flow rate was adjusted to 0.6 mL min-1 and the oven temperature to 40 °C. The following solvent composition was used in a method of 11 min: 30 % B for 1 min, gradient from 30 % to 100 % B in 8 min, gradient from 100 % to

30 % in 0.5 min, 30 % C for 1.5 min (70/30). For identifying and quantifying the produced compounds, a standard mixture containing OA (95 %) purchased from Santa Cruz Biotechnology and OL (95 %) purchased from Sigma Aldrich was measured. Analyzation was performed using the acquired absorbance signals at 215 nm.

For the quantification of OA, OL, and CBGA-C5 analogs, the mobile phases A and B were ddH_2O + 0.1% formic acid and acetonitrile, respectively. A flow rate of 0.6 mL min-1 and an oven temperature of 40 °C was applied using the following method and solvent composition for 17 min: 10% B for 3 min, gradient from 10% to 100% B in 8 min, gradient from 100 % to 10 % B in 2.5 min, 10% B for 2.5 min. For the detection of OA, OL, and CBGA-C5 analogs, the acquired absorbance signals at 215 nm, 225, and 235 nm were analyzed.

For the quantification of THCA-C5, CBCA-C5, and their respective analogs, ddH_2O + 0.1 % formic acid (A) and Acetonitrile (B) were applied as mobile phase. The oven temperature was set to 40 °C, and the flow rate was set to 0.6 mL min-1. The following method and solvent composition were used: 30 % B for 1 min, gradient from 30 % to 85.5 % B in 21 min, gradient from 85.5 % to 100 % B in 1.5 min, 100 % B for 1 min, gradient from 100 % to 30 % B in 1 min, 30 % B for 3 min. For THCA-C5 and analogs, the acquired absorbance signal at 225 nm, and for CBCA-C5 and the analogs, the acquired absorbance signal at 255 was analyzed.

4.2.13 Analytics UPLC-MS and MS/MS

For structure elucidation and assay product identification, MS and MS/MS fragmentation measurements were performed. For these measurements, an Agilent 1290 Infinity II UPLC system combined with a Bruker compact MS/qTOF system equipped with an electrospray ionization source (ESI) was used. For the separation of compounds, the same methods and conditions as described in Chapter 4.2.12 were carried out. Samples were measured using positive and negative mode acquisition.

"For the source conditions in negative mode, the following settings were used: end plate offset 500 V, capillary voltage 2200 V, nebulizer pressure 4.0 bar, dry gas flow rate and temperature 12.0 L min−1 and 220°C, respectively. The source conditions in positive mode were as follows: end plate offset 500 V, capillary voltage 3800 V, nebulizer pressure 4.0 bar, dry gas flow rate and temperature 12.0 L min−1 and 220°C respectively. MS data were acquired in a range from 90 to 700 m/z. MS/MS fragmentation was carried out in the multireaction monitoring mode (MRM) using a collision energy of 36 eV and −36 eV in positive and negative mode, respectively. Identification of the compounds was carried out by comparison of the acquired

MS data and MS/MS fragmentation patterns to the corresponding data of reference standards." (*108*)

4.2.14 Analytics - HPLC-RID

For the determination of glucose and ethanol concentration, the samples were measured with an HPLC-RID (Agilent 1260 Infinity HPLC system) using a Metab-AAC column (300 × 7.8 mm, 10 µm, ISERA). Detection was performed using a using a refractive index detector (RID). For the mobile phase ddH_2O with 5 mM sulfuric acid was applied at a flow of 0.5 mL min^{-1}. The column temperature was set to 40 °C.

4.2.15 Analytics - NMR measurement

For the confirmation of the CBGA derivatives, NMR measurements were carried out at ambient temperature using a Bruker AV 600 Avance III HD spectrometer, resulting in 1D (^1H, ^{13}C) and 2D (^1H-^1H COSY, HSQC, and HMBC) NMR spectroscopy data. NMR spectra were recorded with deuterated DMSO (MeOD) as a solvent and internal standard (DMSO: δ_H 2.50 ppm and δ_C 39.51 ppm). The resulting NMR spectra were analyzed using Bruker TopSpin® software.

4.2.16 Molecular Docking

For the preparation of the protein, like the addition of Hydrogen atoms and the energy minimization, UCSF Chimera (*141*) was used. For the creation of protein variants based on existing protein structure, the amino acids in the chosen positions were exchanged towards the amino acids with the lowest energetic conformation of the amino acid.

For the creation of files consisting of the substrates for docking, the program KNIME was used (*168*). There, the 2D structure was implemented, 3D coordinates were created, the energy was minimized, protons were added, and the molecules were saved as mol2-file.

The docking experiments were performed using Gold (*132*). For *Cs*OAC, the binding pocket was defined using the olivetolic acid bound in the PDB structure 5B09 and a 5 Å radius around this molecule. Afterward, the substrate was deleted from the protein structure. For NphB, the binding pocket was defined using the bound aromatic acceptor molecule and a 5 Å radius around this molecule. While the aromatic acceptor was deleted from the protein structure file, the prenyl donor remained in the binding pocket, as potential substrates also interact with this molecule; otherwise, calculations do not include this interaction. As a scoring function,

ASP_score was used. For each ligand, 50 GA runs were performed. The search efficiency for the GA runs was set to 200 %.

References

1. Crocq, M.-A. History of cannabis and the endocannabinoid system. *Dialogues in clinical neuroscience* **2020,** *22* (3), 223–228. DOI: 10.31887/DCNS.2020.22.3/mcrocq.
2. Burstein, S. Cannabidiol (CBD) and its analogs: a review of their effects on inflammation. *Bioorganic & medicinal chemistry* **2015,** *23* (7), 1377–1385. DOI: 10.1016/j.bmc.2015.01.059.
3. Pertwee, R. G. Cannabinoid pharmacology: the first 66 years. *British journal of pharmacology* **2006,** *147 Suppl 1* (Suppl 1), S163-71. DOI: 10.1038/sj.bjp.0706406.
4. Hanuš, L. O.; Meyer, S. M.; Muñoz, E.; Taglialatela-Scafati, O.; Appendino, G. Phytocannabinoids: a unified critical inventory. *Natural product reports* **2016,** *33* (12), 1357–1392. DOI: 10.1039/c6np00074f.
5. Gülck, T.; Møller, B. L. Phytocannabinoids: Origins and Biosynthesis. *Trends in plant science* **2020,** *25* (10), 985–1004. DOI: 10.1016/j.tplants.2020.05.005.
6. Pellati, F.; Borgonetti, V.; Brighenti, V.; Biagi, M.; Benvenuti, S.; Corsi, L. Cannabis sativa L. and Nonpsychoactive Cannabinoids: Their Chemistry and Role against Oxidative Stress, Inflammation, and Cancer. *BioMed research international* **2018,** *2018*, 1691428. DOI: 10.1155/2018/1691428.
7. Wang, M.; Wang, Y.-H.; Avula, B.; Radwan, M. M.; Wanas, A. S.; van Antwerp, J.; Parcher, J. F.; ElSohly, M. A.; Khan, I. A. Decarboxylation Study of Acidic Cannabinoids: A Novel Approach Using Ultra-High-Performance Supercritical Fluid Chromatography/Photodiode Array-Mass Spectrometry. *Cannabis and cannabinoid research* **2016,** *1* (1), 262–271. DOI: 10.1089/can.2016.0020.
8. Walsh, K. B.; McKinney, A. E.; Holmes, A. E. Minor Cannabinoids: Biosynthesis, Molecular Pharmacology and Potential Therapeutic Uses. *Frontiers in pharmacology* **2021,** *12*, 777804. DOI: 10.3389/fphar.2021.777804.
9. Levine, J. Origin of Cannabinol. *J. Am. Chem. Soc.* **1944,** *66* (11), 1868–1870. DOI: 10.1021/ja01239a018.
10. Maioli, C.; Mattoteia, D.; Amin, H. I. M.; Minassi, A.; Caprioglio, D. Cannabinol: History, Syntheses, and Biological Profile of the Greatest "Minor" Cannabinoid. *Plants (Basel, Switzerland)* **2022,** *11* (21). DOI: 10.3390/plants11212896.

11. Crombie, L.; Ponsford, R.; Shani, A.; Yagnitinsky, B.; Mechoulam, R. Hashish components. Photochemical production of cannabicyclol from cannabichromene. *Tetrahedron Letters* **1968,** *9* (55), 5771–5772. DOI: 10.1016/s0040-4039(00)76346-5.

12. Shoyama, Y.; Hirano, H.; Oda, M.; SOMEHARA, T.; NISHIOKA, I. Cannabichromevarin and cannabigerovarin, two new propyl homologues of cannabichromene and cannabigerol. *Chemical & pharmaceutical bulletin* **1975,** *23* (8), 1894–1895. DOI: 10.1248/cpb.23.1894.

13. Bohlmann, F.; Hoffmann, E. Cannabigerol-ähnliche verbindungen aus Helichrysum umbraculigerum. *Phytochemistry* **1979,** *18* (8), 1371–1374. DOI: 10.1016/0031-9422(79)83025-3.

14. Toyota, M.; Shimamura, T.; Ishii, H.; Renner, M.; Braggins, J.; Asakawa, Y. New bibenzyl cannabinoid from the New Zealand liverwort Radula marginata. *Chemical & pharmaceutical bulletin* **2002,** *50* (10), 1390–1392. DOI: 10.1248/cpb.50.1390.

15. Quaghebeur, K.; Coosemans, J.; Toppet, S.; Compernolle, F. Cannabiorci- and 8-chlorocannabiorcichromenic acid as fungal antagonists from Cylindrocarpon olidum. *Phytochemistry* **1994,** *37* (1), 159–161. DOI: 10.1016/0031-9422(94)85016-X.

16. Hussain, T.; Espley, R. V.; Gertsch, J.; Whare, T.; Stehle, F.; Kayser, O. Demystifying the liverwort Radula marginata, a critical review on its taxonomy, genetics, cannabinoid phytochemistry and pharmacology. *Phytochem Rev* **2019,** *18* (3), 953–965. DOI: 10.1007/s11101-019-09638-8.

17. Berman, P.; Haro, L. A. de; Jozwiak, A.; Panda, S.; Pinkas, Z.; Dong, Y.; Cveticanin, J.; Barbole, R.; Livne, R.; Scherf, T.; Shimoni, E.; Levin-Zaidman, S.; Dezorella, N.; Petrovich-Kopitman, E.; Meir, S.; Rogachev, I.; Sonawane, P. D.; Aharoni, A. Parallel evolution of cannabinoid biosynthesis. *Nature plants* **2023,** *9* (5), 817–831. DOI: 10.1038/s41477-023-01402-3.

18. Tahir, M. N.; Shahbazi, F.; Rondeau-Gagné, S.; Trant, J. F. The biosynthesis of the cannabinoids. *Journal of cannabis research* **2021,** *3* (1), 7. DOI: 10.1186/s42238-021-00062-4.

19. Marks, M. D.; Tian, L.; Wenger, J. P.; Omburo, S. N.; Soto-Fuentes, W.; He, J.; Gang, D. R.; Weiblen, G. D.; Dixon, R. A. Identification of candidate genes affecting Delta9-tetrahydrocannabinol biosynthesis in Cannabis sativa. *Journal of experimental botany* **2009,** *60* (13), 3715–3726. DOI: 10.1093/jxb/erp210.

20. Stout, J. M.; Boubakir, Z.; Ambrose, S. J.; Purves, R. W.; Page, J. E. The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in Cannabis sativa trichomes. *The Plant journal : for cell and molecular biology* **2012,** *71* (3), 353–365. DOI: 10.1111/j.1365-313X.2012.04949.x.

21. Taura, F.; Tanaka, S.; Taguchi, C.; Fukamizu, T.; Tanaka, H.; Shoyama, Y.; Morimoto, S. Characterization of olivetol synthase, a polyketide synthase putatively involved in cannabinoid biosynthetic pathway. *FEBS letters* **2009,** *583* (12), 2061–2066. DOI: 10.1016/j.febslet.2009.05.024.

22. Gagne, S. J.; Stout, J. M.; Liu, E.; Boubakir, Z.; Clark, S. M.; Page, J. E. Identification of olivetolic acid cyclase from Cannabis sativa reveals a unique catalytic route to plant polyketides. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (31), 12811–12816. DOI: 10.1073/pnas.1200330109.

23. Mansouri, H.; Salari, F. Influence of mevinolin on chloroplast terpenoids in Cannabis sativa. *Physiology and molecular biology of plants : an international journal of functional plant biology* **2014,** *20* (2), 273–277. DOI: 10.1007/s12298-014-0222-x.

24. Fellermeier, M.; Zenk, M. H. Prenylation of olivetolate by a hemp transferase yields cannabigerolic acid, the precursor of tetrahydrocannabinol. *FEBS letters* **1998,** *427* (2), 283–285. DOI: 10.1016/S0014-5793(98)00450-5.

25. Luo, X.; Reiter, M. A.; d'Espaux, L.; Wong, J.; Denby, C. M.; Lechner, A.; Zhang, Y.; Grzybowski, A. T.; Harth, S.; Lin, W.; Lee, H.; Yu, C.; Shin, J.; Deng, K.; Benites, V. T.; Wang, G.; Baidoo, E. E. K.; Chen, Y.; Dev, I.; Petzold, C. J.; Keasling, J. D. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **2019,** *567* (7746), 123–126. DOI: 10.1038/s41586-019-0978-9.

26. Taura, F.; Tanaya, R.; Sirikantaramas, S. Recent advances in cannabinoid biochemistry and biotechnology. *ScienceAsia* **2019,** *45* (5), 399. DOI: 10.2306/scienceasia1513-1874.2019.45.399.

27. Taura, F.; Morimoto, S.; Shoyama, Y.; Mechoulam, R. First direct evidence for the mechanism of .DELTA.1-tetrahydrocannabinolic acid biosynthesis. *J. Am. Chem. Soc.* **1995,** *117* (38), 9766–9767. DOI: 10.1021/ja00143a024.

28. Taura, F.; Morimoto, S.; Shoyama, Y. Purification and characterization of cannabidiolic-acid synthase from Cannabis sativa L. Biochemical analysis of a novel enzyme that

catalyzes the oxidocyclization of cannabigerolic acid to cannabidiolic acid. *The Journal of biological chemistry* **1996,** *271* (29), 17411–17416. DOI: 10.1074/jbc.271.29.17411.

29. Morimoto, S.; Komatsu, K.; Taura, F.; Shoyama, Y. Purification and characterization of cannabichromenic acid synthase from Cannabis sativa. *Phytochemistry* **1998,** *49* (6), 1525–1529. DOI: 10.1016/s0031-9422(98)00278-7.

30. Garrett, E. R.; Gouyette, A. J.; Roseboom, H. Stability of tetrahydrocannabinols II. *Journal of pharmaceutical sciences* **1978,** *67* (1), 27–32. DOI: 10.1002/jps.2600670108.

31. Flores-Sanchez, I. J.; Verpoorte, R. Plant polyketide synthases: a fascinating group of enzymes. *Plant physiology and biochemistry : PPB* **2009,** *47* (3), 167–174. DOI: 10.1016/j.plaphy.2008.11.005.

32. Katsuyama, Y.; Ohnishi, Y. Type III polyketide synthases in microorganisms. *Methods in enzymology* **2012,** *515*, 359–377. DOI: 10.1016/B978-0-12-394290-6.00017-3.

33. Kearsey, L. J.; Prandi, N.; Karuppiah, V.; Yan, C.; Leys, D.; Toogood, H.; Takano, E.; Scrutton, N. S. Structure of the Cannabis sativa olivetol-producing enzyme reveals cyclization plasticity in type III polyketide synthases. *The FEBS journal* **2020,** *287* (8), 1511–1524. DOI: 10.1111/febs.15089.

34. Weng, J.-K.; Noel, J. P. Structure-function analyses of plant type III polyketide synthases. *Methods in enzymology* **2012,** *515*, 317–335. DOI: 10.1016/B978-0-12-394290-6.00014-8.

35. Jez, J. M.; Bowman, M. E.; Noel, J. P. Structure-guided programming of polyketide chain-length determination in chalcone synthase. *Biochemistry* **2001,** *40* (49), 14829–14838. DOI: 10.1021/bi015621z.

36. Austin, M. B.; Bowman, M. E.; Ferrer, J.-L.; Schröder, J.; Noel, J. P. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chemistry & biology* **2004,** *11* (9), 1179–1194. DOI: 10.1016/j.chembiol.2004.05.024.

37. Yang, X.; Matsui, T.; Kodama, T.; Mori, T.; Zhou, X.; Taura, F.; Noguchi, H.; Abe, I.; Morita, H. Structural basis for olivetolic acid formation by a polyketide cyclase from Cannabis sativa. *The FEBS journal* **2016,** *283* (6), 1088–1106. DOI: 10.1111/febs.13654.

38. Dastmalchi, M. Elusive partners: a review of the auxiliary proteins guiding metabolic flux in flavonoid biosynthesis. *The Plant journal : for cell and molecular biology* **2021,** *108* (2), 314–329. DOI: 10.1111/tpj.15446.

39. Page, J. E.; Boubakir, Z. Aromatic Prenyltransferase from Cannabis. US201013389815 20100804.

40. Sands, L. B.; Haiden, S. R.; Ma, Y.; Berkowitz, G. A. Hormonal control of promoter activities of Cannabis sativa prenyltransferase 1 and 4 and salicylic acid mediated regulation of cannabinoid biosynthesis. *Scientific reports* **2023,** *13* (1), 8620. DOI: 10.1038/s41598-023-35303-4.

41. Lim, K. J. H.; Lim, Y. P.; Hartono, Y. D.; Go, M. K.; Fan, H.; Yew, W. S. Biosynthesis of Nature-Inspired Unnatural Cannabinoids. *Molecules (Basel, Switzerland)* **2021,** *26* (10). DOI: 10.3390/molecules26102914.

42. Schultz, E. E.; Braffman, N. R.; Luescher, M. U.; Hager, H. H.; Balskus, E. P. Biocatalytic Friedel-Crafts Alkylation Using a Promiscuous Biosynthetic Enzyme. *Angewandte Chemie (International ed. in English)* **2019,** *58* (10), 3151–3155. DOI: 10.1002/anie.201814016.

43. Li, W. Bringing Bioactive Compounds into Membranes: The UbiA Superfamily of Intramembrane Aromatic Prenyltransferases. *Trends in biochemical sciences* **2016,** *41* (4), 356–370. DOI: 10.1016/j.tibs.2016.01.007.

44. Huang, H.; Levin, E. J.; Liu, S.; Bai, Y.; Lockless, S. W.; Zhou, M. Structure of a membrane-embedded prenyltransferase homologous to UBIAD1. *PLoS biology* **2014,** *12* (7), e1001911. DOI: 10.1371/journal.pbio.1001911.

45. Sasaki, K.; Mito, K.; Ohara, K.; Yamamoto, H.; Yazaki, K. Cloning and characterization of naringenin 8-prenyltransferase, a flavonoid-specific prenyltransferase of Sophora flavescens. *Plant Physiology* **2008,** *146* (3), 1075–1084. DOI: 10.1104/pp.107.110544.

46. Cheng, W.; Li, W. Structural insights into ubiquinone biosynthesis in membranes. *Science (New York, N.Y.)* **2014,** *343* (6173), 878–881. DOI: 10.1126/science.1246774.

47. Krogh, A.; Larsson, B.; Heijne, G. von; Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology* **2001,** *305* (3), 567–580. DOI: 10.1006/jmbi.2000.4315.

48. van Velzen, R.; Schranz, M. E. Origin and Evolution of the Cannabinoid Oxidocyclase Gene Family. *Genome biology and evolution* **2021,** *13* (8). DOI: 10.1093/gbe/evab130.

49. Sirikantaramas, S.; Morimoto, S.; Shoyama, Y.; Ishikawa, Y.; Wada, Y.; Shoyama, Y.; Taura, F. The gene controlling marijuana psychoactivity: molecular cloning and

heterologous expression of Delta1-tetrahydrocannabinolic acid synthase from Cannabis sativa L. *The Journal of biological chemistry* **2004,** *279* (38), 39767–39774. DOI: 10.1074/jbc.M403693200.

50. Taura, F.; Sirikantaramas, S.; Shoyama, Y.; Shoyama, Y.; Morimoto, S. Phytocannabinoids in Cannabis sativa: recent studies on biosynthetic enzymes. *Chemistry & biodiversity* **2007,** *4* (8), 1649–1663. DOI: 10.1002/cbdv.200790145.

51. Daniel, B.; Wallner, S.; Steiner, B.; Oberdorfer, G.; Kumar, P.; van der Graaff, E.; Roitsch, T.; Sensen, C. W.; Gruber, K.; Macheroux, P. Structure of a Berberine Bridge Enzyme-Like Enzyme with an Active Site Specific to the Plant Family Brassicaceae. *PloS one* **2016,** *11* (6), e0156892. DOI: 10.1371/journal.pone.0156892.

52. Shoyama, Y.; Tamada, T.; Kurihara, K.; Takeuchi, A.; Taura, F.; Arai, S.; Blaber, M.; Shoyama, Y.; Morimoto, S.; Kuroki, R. Structure and function of Δ1-tetrahydrocannabinolic acid (THCA) synthase, the enzyme controlling the psychoactivity of Cannabis sativa. *Journal of molecular biology* **2012,** *423* (1), 96–105. DOI: 10.1016/j.jmb.2012.06.030.

53. Zirpel, B.; Kayser, O.; Stehle, F. Elucidation of structure-function relationship of THCA and CBDA synthase from Cannabis sativaL. *Journal of biotechnology* **2018,** *284*, 17–26. DOI: 10.1016/j.jbiotec.2018.07.031.

54. Thomas, F.; Kayser, O. Improving CBCA synthase activity through rational protein design. *Journal of biotechnology* **2023,** *363*, 40–49. DOI: 10.1016/j.jbiotec.2023.01.004.

55. Schmidt, C.; Aras, M.; Kayser, O. Engineering cannabinoid production in Saccharomyces cerevisiae. *Biotechnology journal* **2024,** *19* (2), e2300507. DOI: 10.1002/biot.202300507.

56. Zou, S.; Kumar, U. Cannabinoid Receptors and the Endocannabinoid System: Signaling and Function in the Central Nervous System. *International journal of molecular sciences* **2018,** *19* (3). DOI: 10.3390/ijms19030833.

57. Matsuda, L. A.; Lolait, S. J.; Brownstein, M. J.; Young, A. C.; Bonner, T. I. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* **1990,** *346* (6284), 561–564. DOI: 10.1038/346561a0.

58. Galiègue, S.; Mary, S.; Marchand, J.; Dussossoy, D.; Carrière, D.; Carayon, P.; Bouaboula, M.; Shire, D.; Le Fur, G.; Casellas, P. Expression of central and peripheral cannabinoid

receptors in human immune tissues and leukocyte subpopulations. *European journal of biochemistry* **1995,** *232* (1), 54–61. DOI: 10.1111/j.1432-1033.1995.tb20780.x.

59. Munro, S.; Thomas, K. L.; Abu-Shaar, M. Molecular characterization of a peripheral receptor for cannabinoids. *Nature* **1993,** *365* (6441), 61–65. DOI: 10.1038/365061a0.

60. Howlett, A. C.; Barth, F.; Bonner, T. I.; Cabral, G.; Casellas, P.; Devane, W. A.; Felder, C. C.; Herkenham, M.; Mackie, K.; Martin, B. R.; Mechoulam, R.; Pertwee, R. G. International Union of Pharmacology. XXVII. Classification of cannabinoid receptors. *Pharmacological reviews* **2002,** *54* (2), 161–202. DOI: 10.1124/pr.54.2.161.

61. Howlett, A. C. Cannabinoid inhibition of adenylate cyclase. Biochemistry of the response in neuroblastoma cell membranes. *Molecular pharmacology* **1985,** *27* (4), 429–436.

62. Jin, W.; Brown, S.; Roche, J. P.; Hsieh, C.; Celver, J. P.; Kovoor, A.; Chavkin, C.; Mackie, K. Distinct domains of the CB1 cannabinoid receptor mediate desensitization and internalization. *The Journal of Neuroscience* **1999,** *19* (10), 3773–3780. DOI: 10.1523/JNEUROSCI.19-10-03773.1999.

63. Console-Bram, L.; Marcu, J.; Abood, M. E. Cannabinoid receptors: nomenclature and pharmacological principles. *Progress in neuro-psychopharmacology & biological psychiatry* **2012,** *38* (1), 4–15. DOI: 10.1016/j.pnpbp.2012.02.009.

64. Hua, T.; Vemuri, K.; Pu, M.; Qu, L.; Han, G. W.; Wu, Y.; Zhao, S.; Shui, W.; Li, S.; Korde, A.; Laprairie, R. B.; Stahl, E. L.; Ho, J.-H.; Zvonok, N.; Zhou, H.; Kufareva, I.; Wu, B.; Zhao, Q.; Hanson, M. A.; Bohn, L. M.; Makriyannis, A.; Stevens, R. C.; Liu, Z.-J. Crystal Structure of the Human Cannabinoid Receptor CB1. *Cell* **2016,** *167* (3), 750-762.e14. DOI: 10.1016/j.cell.2016.10.004.

65. Li, X.; Hua, T.; Vemuri, K.; Ho, J.-H.; Wu, Y.; Wu, L.; Popov, P.; Benchama, O.; Zvonok, N.; Locke, K.; Qu, L.; Han, G. W.; Iyer, M. R.; Cinar, R.; Coffey, N. J.; Wang, J.; Wu, M.; Katritch, V.; Zhao, S.; Kunos, G.; Bohn, L. M.; Makriyannis, A.; Stevens, R. C.; Liu, Z.-J. Crystal Structure of the Human Cannabinoid Receptor CB2. *Cell* **2019,** *176* (3), 459-467.e13. DOI: 10.1016/j.cell.2018.12.011.

66. Mackie, K. Cannabinoid receptors: where they are and what they do. *Journal of neuroendocrinology* **2008,** *20 Suppl 1*, 10–14. DOI: 10.1111/j.1365-2826.2008.01671.x.

67. Bisogno, T. Endogenous cannabinoids: structure and metabolism. *Journal of neuroendocrinology* **2008,** *20 Suppl 1*, 1–9. DOI: 10.1111/j.1365-2826.2008.01676.x.

68. Di Marzo, V.; Bifulco, M.; Petrocellis, L. de. The endocannabinoid system and its therapeutic exploitation. *Nature reviews. Drug discovery* **2004,** *3* (9), 771–784. DOI: 10.1038/nrd1495.

69. Mechoulam, R.; Ben-Shabat, S.; Hanus, L.; Ligumsky, M.; Kaminski, N. E.; Schatz, A. R.; Gopher, A.; Almog, S.; Martin, B. R.; Compton, D. R. Identification of an endogenous 2-monoglyceride, present in canine gut, that binds to cannabinoid receptors. *Biochemical pharmacology* **1995,** *50* (1), 83–90. DOI: 10.1016/0006-2952(95)00109-D.

70. Sugiura, T.; Kondo, S.; Sukagawa, A.; Nakane, S.; Shinoda, A.; Itoh, K.; Yamashita, A.; Waku, K. 2-Arachidonoylglycerol: a possible endogenous cannabinoid receptor ligand in brain. *Biochemical and biophysical research communications* **1995,** *215* (1), 89–97. DOI: 10.1006/bbrc.1995.2437.

71. Schrot, R. J.; Hubbard, J. R. Cannabinoids: Medical implications. *Annals of medicine* **2016,** *48* (3), 128–141. DOI: 10.3109/07853890.2016.1145794.

72. Śledziński, P.; Nowak-Terpiłowska, A.; Zeyland, J. Cannabinoids in Medicine: Cancer, Immunity, and Microbial Diseases. *International journal of molecular sciences* **2020,** *22* (1). DOI: 10.3390/ijms22010263.

73. Di Marzo, V.; Petrocellis, L. de. Plant, synthetic, and endogenous cannabinoids in medicine. *Annual review of medicine* **2006,** *57*, 553–574. DOI: 10.1146/annurev.med.57.011205.135648.

74. Pitakbut, T.; Nguyen, G.-N.; Kayser, O. Activity of THC, CBD, and CBN on Human ACE2 and SARS-CoV1/2 Main Protease to Understand Antiviral Defense Mechanism. *Planta medica* **2022,** *88* (12), 1047–1059. DOI: 10.1055/a-1581-3707.

75. Nguyen, L. C.; Yang, D.; Nicolaescu, V.; Best, T. J.; Gula, H.; Saxena, D.; Gabbard, J. D.; Chen, S.-N.; Ohtsuki, T.; Friesen, J. B.; Drayman, N.; Mohamed, A.; Dann, C.; Silva, D.; Robinson-Mailman, L.; Valdespino, A.; Stock, L.; Suárez, E.; Jones, K. A.; Azizi, S.-A.; Demarco, J. K.; Severson, W. E.; Anderson, C. D.; Millis, J. M.; Dickinson, B. C.; Tay, S.; Oakes, S. A.; Pauli, G. F.; Palmer, K. E.; Meltzer, D. O.; Randall, G.; Rosner, M. R. Cannabidiol inhibits SARS-CoV-2 replication through induction of the host ER stress and innate immune responses. *Science advances* **2022,** *8* (8), eabi6110. DOI: 10.1126/sciadv.abi6110.

76. van Breemen, R. B.; Muchiri, R. N.; Bates, T. A.; Weinstein, J. B.; Leier, H. C.; Farley, S.; Tafesse, F. G. Cannabinoids Block Cellular Entry of SARS-CoV-2 and the Emerging Variants. *Journal of natural products* **2022,** *85* (1), 176–184. DOI: 10.1021/acs.jnatprod.1c00946.

77. Antoniou, T.; Juurlink, D. N. Synthetic cannabinoids. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* **2014,** *186* (3), 210. DOI: 10.1503/cmaj.130510.

78. Roque-Bravo, R.; Silva, R. S.; Malheiro, R. F.; Carmo, H.; Carvalho, F.; da Silva, D. D.; Silva, J. P. Synthetic Cannabinoids: A Pharmacological and Toxicological Overview. *Annual review of pharmacology and toxicology* **2023,** *63*, 187–209. DOI: 10.1146/annurev-pharmtox-031122-113758.

79. Andrews, R.; Jorge, R.; Christie, R.; Gallegos, A. From JWH-018 to OXIZIDS: Structural evolution of synthetic cannabinoids in the European Union from 2008 to present day. *Drug testing and analysis* **2023,** *15* (4), 378–387. DOI: 10.1002/dta.3422.

80. Bileck, A.; Ferk, F.; Al-Serori, H.; Koller, V. J.; Muqaku, B.; Haslberger, A.; Auwärter, V.; Gerner, C.; Knasmüller, S. Impact of a synthetic cannabinoid (CP-47,497-C8) on protein expression in human cells: evidence for induction of inflammation and DNA damage. *Archives of toxicology* **2016,** *90* (6), 1369–1382. DOI: 10.1007/s00204-015-1569-7.

81. Atwood, B. K.; Huffman, J.; Straiker, A.; Mackie, K. JWH018, a common constituent of 'Spice' herbal blends, is a potent and efficacious cannabinoid CB receptor agonist. *British journal of pharmacology* **2010,** *160* (3), 585–593. DOI: 10.1111/j.1476-5381.2009.00582.x.

82. Pertwee, R. G. Pharmacological actions of cannabinoids. *Handbook of experimental pharmacology* **2005** (168), 1–51. DOI: 10.1007/3-540-26573-2_1.

83. Nguyen, G.-N.; Jordan, E. N.; Kayser, O. Protecting-Group-Free Synthesis of Novel Cannabinoid-Like 2,5-Dihydrobenzoxepines. *Synthesis* **2022,** *54* (24), 5540–5550. DOI: 10.1055/s-0042-1751361.

84. Nguyen, G.-N.; Jordan, E. N.; Kayser, O. Synthetic Strategies for Rare Cannabinoids Derived from Cannabis sativa. *Journal of natural products* **2022,** *85* (6), 1555–1568. DOI: 10.1021/acs.jnatprod.2c00155.

85. Mechoulam, R.; Feigenbaum, J. J.; Lander, N.; Segal, M.; Järbe, T. U.; Hiltunen, A. J.; Consroe, P. Enantiomeric cannabinoids: stereospecificity of psychotropic activity. *Experientia* **1988,** *44* (9), 762–764. DOI: 10.1007/BF01959156.

86. Wiles, D.; Shanbhag, B. K.; O'Brien, M.; Doblin, M. S.; Bacic, A.; Beddoe, T. Heterologous production of Cannabis sativa-derived specialised metabolites of medicinal significance - Insights into engineering strategies. *Phytochemistry* **2022,** *203*, 113380. DOI: 10.1016/j.phytochem.2022.113380.

87. Carvalho, Â.; Hansen, E. H.; Kayser, O.; Carlsen, S.; Stehle, F. Designing microorganisms for heterologous biosynthesis of cannabinoids. *FEMS yeast research* **2017,** *17* (4). DOI: 10.1093/femsyr/fox037.

88. Thomas, F.; Schmidt, C.; Kayser, O. Bioengineering studies and pathway modeling of the heterologous biosynthesis of tetrahydrocannabinolic acid in yeast. *Applied microbiology and biotechnology* **2020,** *104* (22), 9551–9563. DOI: 10.1007/s00253-020-10798-3.

89. Liao, P.; Hemmerlin, A.; Bach, T. J.; Chye, M.-L. The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnology advances* **2016,** *34* (5), 697–713. DOI: 10.1016/j.biotechadv.2016.03.005.

90. Zhang, Y.; Guo, J.; Gao, P.; Yan, W.; Shen, J.; Luo, X.; Keasling, J. D. Development of an efficient yeast platform for cannabigerolic acid biosynthesis. *Metabolic engineering* **2023,** *80*, 232–240. DOI: 10.1016/j.ymben.2023.10.004.

91. Zirpel, B.; Degenhardt, F.; Zammarelli, C.; Wibberg, D.; Kalinowski, J.; Stehle, F.; Kayser, O. Optimization of Δ9-tetrahydrocannabinolic acid synthase production in Komagataella phaffii via post-translational bottleneck identification. *Journal of biotechnology* **2018,** *272-273*, 40–47. DOI: 10.1016/j.jbiotec.2018.03.008.

92. Ma, J.; Gu, Y.; Xu, P. Biosynthesis of cannabinoid precursor olivetolic acid in genetically engineered Yarrowia lipolytica. *Communications biology* **2022,** *5* (1), 1239. DOI: 10.1038/s42003-022-04202-1.

93. Reddy, V. A.; Leong, S. H.; Jang, I.-C.; Rajani, S. Metabolic Engineering of Nicotiana benthamiana to Produce Cannabinoid Precursors and Their Analogues. *Metabolites* **2022,** *12* (12). DOI: 10.3390/metabo12121181.

94. Gülck, T.; Booth, J. K.; Carvalho, Â.; Khakimov, B.; Crocoll, C.; Motawia, M. S.; Møller, B. L.; Bohlmann, J.; Gallage, N. J. Synthetic Biology of Cannabinoids and Cannabinoid

Glucosides in Nicotiana benthamiana and Saccharomyces cerevisiae. *Journal of natural products* **2020,** *83* (10), 2877–2893. DOI: 10.1021/acs.jnatprod.0c00241.

95. Tan, Z.; Clomburg, J. M.; Gonzalez, R. Synthetic Pathway for the Production of Olivetolic Acid in Escherichia coli. *ACS synthetic biology* **2018,** *7* (8), 1886–1896. DOI: 10.1021/acssynbio.8b00075.

96. Shujat, S.; Robinson, G. I.; Norouzkhani, F.; Kovalchuk, I. Using Advanced Biotechnological Techniques to Improve Cannabis Cultivars. *Biocatalysis and Agricultural Biotechnology* **2024**, 103250. DOI: 10.1016/j.bcab.2024.103250.

97. Chen, R.; Gao, B.; Liu, X.; Ruan, F.; Zhang, Y.; Lou, J.; Feng, K.; Wunsch, C.; Li, S.-M.; Dai, J.; Sun, F. Molecular insights into the enzyme promiscuity of an aromatic prenyltransferase. *Nature chemical biology* **2017,** *13* (2), 226–234. DOI: 10.1038/nchembio.2263.

98. Kearsey, L. J.; Yan, C.; Prandi, N.; Toogood, H. S.; Takano, E.; Scrutton, N. S. Biosynthesis of cannabigerol and cannabigerolic acid: the gateways to further cannabinoid production. *Synthetic biology (Oxford, England)* **2023,** *8* (1), ysad010. DOI: 10.1093/synbio/ysad010.

99. Kumano, T.; Richard, S. B.; Noel, J. P.; Nishiyama, M.; Kuzuyama, T. Chemoenzymatic syntheses of prenylated aromatic small molecules using Streptomyces prenyltransferases with relaxed substrate specificities. *Bioorganic & medicinal chemistry* **2008,** *16* (17), 8117–8126. DOI: 10.1016/j.bmc.2008.07.052.

100. Zirpel, B.; Degenhardt, F.; Martin, C.; Kayser, O.; Stehle, F. Engineering yeasts as platform organisms for cannabinoid biosynthesis. *Journal of biotechnology* **2017,** *259*, 204–212. DOI: 10.1016/j.jbiotec.2017.07.008.

101. Kuzuyama, T.; Noel, J. P.; Richard, S. B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **2005,** *435* (7044), 983–987. DOI: 10.1038/nature03668.

102. Tello, M.; Kuzuyama, T.; Heide, L.; Noel, J. P.; Richard, S. B. The ABBA family of aromatic prenyltransferases: broadening natural product diversity. *Cellular and molecular life sciences : CMLS* **2008,** *65* (10), 1459–1463. DOI: 10.1007/s00018-008-7579-3.

103. Cui, G.; Li, X.; Merz, K. M. Understanding the substrate selectivity and the product regioselectivity of Orf2-catalyzed aromatic prenylations. *Biochemistry* **2007,** *46* (5), 1303–1311. DOI: 10.1021/bi062076z.

104. Murray, L. A. M.; McKinnie, S. M. K.; Pepper, H. P.; Erni, R.; Miles, Z. D.; Cruickshank, M. C.; López-Pérez, B.; Moore, B. S.; George, J. H. Total Synthesis Establishes the Biosynthetic Pathway to the Naphterpin and Marinone Natural Products. *Angewandte Chemie (International ed. in English)* **2018,** *57* (34), 11009–11014. DOI: 10.1002/anie.201804351.

105. Johnson, B. P.; Scull, E. M.; Dimas, D. A.; Bavineni, T.; Bandari, C.; Batchev, A. L.; Gardner, E. D.; Nimmo, S. L.; Singh, S. Acceptor substrate determines donor specificity of an aromatic prenyltransferase: expanding the biocatalytic potential of NphB. *Applied microbiology and biotechnology* **2020,** *104* (10), 4383–4395. DOI: 10.1007/s00253-020-10529-8.

106. Yang, Y.; Miao, Y.; Wang, B.; Cui, G.; Merz, K. M. Catalytic mechanism of aromatic prenylation by NphB. *Biochemistry* **2012,** *51* (12), 2606–2618. DOI: 10.1021/bi201800m.

107. Organic chemistry. *J. Chem. Soc.* **1877,** *32*, 725. DOI: 10.1039/js8773200725.

108. Spitzer, S.; Wloka, J.; Pietruszka, J.; Kayser, O. Generation of Cannabigerolic Acid Derivatives and Their Precursors by Using the Promiscuity of the Aromatic Prenyltransferase NphB. *Chembiochem : a European journal of chemical biology* **2023,** *24* (22), e202300441. DOI: 10.1002/cbic.202300441.

109. Lee, Y.-E.; Nakashima, Y.; Kodama, T.; Chen, X.; Morita, H. Dual Engineering of Olivetolic Acid Cyclase and Tetraketide Synthase to Generate Longer Alkyl-Chain Olivetolic Acid Analogs. *Organic letters* **2022,** *24* (1), 410–414. DOI: 10.1021/acs.orglett.1c04089.

110. Kayser, O.; Stehle, F. O. BIOTECHNOLOGICAL PRODUCTION OF CANNABINOIDS. US201917257867 20190717.

111. Kayser, O.; Stehle, F. O. Biotechnologische Herstellung von Cannabinoiden. DE201810117233 20180717.

112. Valliere, M. A.; Korman, T. P.; Woodall, N. B.; Khitrov, G. A.; Taylor, R. E.; Baker, D.; Bowie, J. U. A cell-free platform for the prenylation of natural products and

application to cannabinoid production. *Nature communications* **2019,** *10* (1), 565. DOI: 10.1038/s41467-019-08448-y.

113. Qian, S.; Clomburg, J. M.; Gonzalez, R. Engineering Escherichia coli as a platform for the in vivo synthesis of prenylated aromatics. *Biotechnology and bioengineering* **2019,** *116* (5), 1116–1127. DOI: 10.1002/bit.26932.

114. Lim, K. J. H.; Hartono, Y. D.; Xue, B.; Go, M. K.; Fan, H.; Yew, W. S. Structure-Guided Engineering of Prenyltransferase NphB for High-Yield and Regioselective Cannabinoid Production. *ACS Catal.* **2022,** *12* (8), 4628–4639. DOI: 10.1021/acscatal.2c00786.

115. Taura, F.; Dono, E.; Sirikantaramas, S.; Yoshimura, K.; Shoyama, Y.; Morimoto, S. Production of Delta(1)-tetrahydrocannabinolic acid by the biosynthetic enzyme secreted from transgenic Pichia pastoris. *Biochemical and biophysical research communications* **2007,** *361* (3), 675–680. DOI: 10.1016/j.bbrc.2007.07.079.

116. Laverty, K. U.; Stout, J. M.; Sullivan, M. J.; Shah, H.; Gill, N.; Holbrook, L.; Deikus, G.; Sebra, R.; Hughes, T. R.; Page, J. E.; van Bakel, H. A physical and genetic map of Cannabis sativa identifies extensive rearrangements at the THC/CBD acid synthase loci. *Genome research* **2019,** *29* (1), 146–156. DOI: 10.1101/gr.242594.118.

117. Sun, C.; Li, G.; Li, H.; Lyu, Y.; Yu, S.; Zhou, J. Enhancing Flavan-3-ol Biosynthesis in Saccharomyces cerevisiae. *Journal of agricultural and food chemistry* **2021,** *69* (43), 12763–12772. DOI: 10.1021/acs.jafc.1c04489.

118. Chen, X.; Zaro, J. L.; Shen, W.-C. Fusion protein linkers: property, design and functionality. *Advanced drug delivery reviews* **2013,** *65* (10), 1357–1369. DOI: 10.1016/j.addr.2012.09.039.

119. Zhao, H. L.; Xue, C.; Wang, Y.; Li, X. Y.; Xiong, X. H.; Yao, X. Q.; Liu, Z. M. Circumventing the heterogeneity and instability of human serum albumin-interferon-alpha2b fusion protein by altering its orientation. *Journal of biotechnology* **2007,** *131* (3), 245–252. DOI: 10.1016/j.jbiotec.2007.04.016.

120. Keasling, J. D.; D, E. L.; Wong, J.; Luo, X.; Reiter, M.; Denby, C.; Lechner, A. Microorganisms and methothods for producing cannabinoids and cannabinoid derivatives. EP20180728259 20180427.

121. Liu, W.; Zhao, H.; Jia, B.; Xu, L.; Yan, Y. Surface display of active lipase in Saccharomyces cerevisiae using Cwp2 as an anchor protein. *Biotechnology letters* **2010,** *32* (2), 255–260. DOI: 10.1007/s10529-009-0138-7.

122. Aurora, R.; Creamer, T. P.; Srinivasan, R.; Rose, G. D. Local interactions in protein folding: lessons from the alpha-helix. *The Journal of biological chemistry* **1997,** *272* (3), 1413–1416. DOI: 10.1074/jbc.272.3.1413.

123. Arai, R.; Ueda, H.; Kitayama, A.; Kamiya, N.; Nagamune, T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein engineering* **2001,** *14* (8), 529–532. DOI: 10.1093/protein/14.8.529.

124. Chica, R. A.; Doucet, N.; Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Current opinion in biotechnology* **2005,** *16* (4), 378–384. DOI: 10.1016/j.copbio.2005.06.004.

125. Song, Z.; Zhang, Q.; Wu, W.; Pu, Z.; Yu, H. Rational design of enzyme activity and enantioselectivity. *Frontiers in bioengineering and biotechnology* **2023,** *11*, 1129149. DOI: 10.3389/fbioe.2023.1129149.

126. Hunt, M. C.; Alexson, S. E. H. The role Acyl-CoA thioesterases play in mediating intracellular lipid metabolism. *Progress in lipid research* **2002,** *41* (2), 99–130. DOI: 10.1016/s0163-7827(01)00017-0.

127. Spitzer, S.; Aras, M.; Kayser, O. Improving CsOAC Activity in Saccharomyces cerevisiae for Directed Production of Olivetolic Acid through Rational Design. *Chembiochem : a European journal of chemical biology* **2024**, e202400651. DOI: 10.1002/cbic.202400651.

128. Feinstein, W. P.; Brylinski, M. Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. *J. Cheminform.* **2015,** *7*, 18. DOI: 10.1186/s13321-015-0067-5.

129. Sumbalova, L.; Stourac, J.; Martinek, T.; Bednar, D.; Damborsky, J. HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information. *Nucleic acids research* **2018,** *46* (W1), W356-W362. DOI: 10.1093/nar/gky417.

130. Chen, Y.; Nielsen, J. In vitro turnover numbers do not reflect in vivo activities of yeast enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **2021,** *118* (32). DOI: 10.1073/pnas.2108391118.

131. Maury, J.; Germann, S. M.; Baallal Jacobsen, S. A.; Jensen, N. B.; Kildegaard, K. R.; Herrgård, M. J.; Schneider, K.; Koza, A.; Forster, J.; Nielsen, J.; Borodina, I. EasyCloneMulti: A Set of Vectors for Simultaneous and Multiple Genomic Integrations in Saccharomyces cerevisiae. *PloS one* **2016,** *11* (3), e0150394. DOI: 10.1371/journal.pone.0150394.

132. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *Journal of molecular biology* **1997,** *267* (3), 727–748. DOI: 10.1006/jmbi.1996.0897.

133. Li, L.; Guo, D.; Huang, Y.; Liu, S.; Xiao, Y. ASPDock: protein-protein docking algorithm using atomic solvation parameters model. *BMC bioinformatics* **2011,** *12*, 36. DOI: 10.1186/1471-2105-12-36.

134. La Kastberg, L. B.; Ard, R.; Jensen, M. K.; Workman, C. T. Burden Imposed by Heterologous Protein Production in Two Major Industrial Yeast Cell Factories: Identifying Sources and Mitigation Strategies. *Frontiers in fungal biology* **2022,** *3*, 827704. DOI: 10.3389/ffunb.2022.827704.

135. Mohamed, R.; Degac, J.; Helms, V. Composition of Overlapping Protein-Protein and Protein-Ligand Interfaces. *PloS one* **2015,** *10* (10), e0140965. DOI: 10.1371/journal.pone.0140965.

136. Lomakin, I. B.; Xiong, Y.; Steitz, T. A. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* **2007,** *129* (2), 319–332. DOI: 10.1016/j.cell.2007.03.013.

137. Tehlivets, O.; Scheuringer, K.; Kohlwein, S. D. Fatty acid synthesis and elongation in yeast. *Biochimica et biophysica acta* **2007,** *1771* (3), 255–270. DOI: 10.1016/j.bbalip.2006.07.004.

138. Inokoshi, J.; Tomoda, H.; Hashimoto, H.; Watanabe, A.; Takeshima, H.; Omura, S. Cerulenin-resistant mutants of Saccharomyces cerevisiae with an altered fatty acid synthase gene. *Molecular & general genetics : MGG* **1994,** *244* (1), 90–96. DOI: 10.1007/BF00280191.

139. Schäfer, K. J.; Aras, M.; Boles, E.; Kayser, O. Optimizing hexanoic acid biosynthesis in Saccharomyces cerevisiae for the de novo production of olivetolic acid. *Biotechnology for biofuels and bioproducts* **2024,** *17* (1), 141. DOI: 10.1186/s13068-024-02586-2.

140. Chakravorty, D. K.; Merz, K. M. Role of substrate dynamics in protein prenylation reactions. *Accounts of chemical research* **2015,** *48* (2), 439–448. DOI: 10.1021/ar500321u.

141. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004,** *25* (13), 1605–1612. DOI: 10.1002/jcc.20084.

142. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein engineering* **1995,** *8* (2), 127–134. DOI: 10.1093/protein/8.2.127.

143. Lee, Y.-E.; Kodama, T.; Morita, H. Novel insights into the antibacterial activities of cannabinoid biosynthetic intermediate, olivetolic acid, and its alkyl-chain derivatives. *Journal of natural medicines* **2023,** *77* (2), 298–305. DOI: 10.1007/s11418-022-01672-9.

144. Pollastro, F.; Petrocellis, L. de; Schiano-Moriello, A.; Chianese, G.; Heyman, H.; Appendino, G.; Taglialatela-Scafati, O. Amorfrutin-type phytocannabinoids from Helichrysum umbraculigerum. *Fitoterapia* **2017,** *123*, 13–17. DOI: 10.1016/j.fitote.2017.09.010.

145. Leitão, J. M. M.; Da Esteves Silva, J. C. G. Firefly luciferase inhibition. *Journal of photochemistry and photobiology. B, Biology* **2010,** *101* (1), 1–8. DOI: 10.1016/j.jphotobiol.2010.06.015.

146. Karamohamed, S.; Nordström, T.; Nyrén, P. Real-time bioluminometric method for detection of nucleoside diphosphate kinase activity. *BioTechniques* **1999,** *26* (4), 728–734. DOI: 10.2144/99264rr02.

147. Inouye, S. Firefly luciferase: an adenylate-forming enzyme for multicatalytic functions. *Cellular and molecular life sciences : CMLS* **2010,** *67* (3), 387–404. DOI: 10.1007/s00018-009-0170-8.

148. Auld, D. S.; Thorne, N.; Nguyen, D.-T.; Inglese, J. A specific mechanism for nonspecific activation in reporter-gene assays. *ACS chemical biology* **2008,** *3* (8), 463–470. DOI: 10.1021/cb8000793.

149. Auld, D. S.; Southall, N. T.; Jadhav, A.; Johnson, R. L.; Diller, D. J.; Simeonov, A.; Austin, C. P.; Inglese, J. Characterization of chemical libraries for luciferase inhibitory activity. *Journal of medicinal chemistry* **2008,** *51* (8), 2372–2386. DOI: 10.1021/jm701302v.

150. Marques, S. M.; Da Esteves Silva, J. C. G. Firefly bioluminescence: a mechanistic approach of luciferase catalyzed reactions. *IUBMB life* **2009,** *61* (1), 6–17. DOI: 10.1002/iub.134.

151. Noori, A. R.; Hosseinkhani, S.; Ghiasi, P.; Heydari, A.; Akbari, J. Water-miscible ionic liquids as novel effectors for the firefly luciferase reaction. *Engineering in Life Sciences* **2013,** *13* (2), 201–209. DOI: 10.1002/elsc.201100214.

152. Lohrasbi-Nejad, A.; Torkzadeh-Mahani, M.; Hosseinkhani, S. Hydrophobin-1 promotes thermostability of firefly luciferase. *The FEBS journal* **2016,** *283* (13), 2494–2507. DOI: 10.1111/febs.13757.

153. Thorne, N.; Shen, M.; Lea, W. A.; Simeonov, A.; Lovell, S.; Auld, D. S.; Inglese, J. Firefly luciferase in chemical biology: a compendium of inhibitors, mechanistic evaluation of chemotypes, and suggested use as a reporter. *Chemistry & biology* **2012,** *19* (8), 1060–1072. DOI: 10.1016/j.chembiol.2012.07.015.

154. Rudge, E. S.; Chan, A. H. Y.; Leeper, F. J. Prodrugs of pyrophosphates and bisphosphonates: disguising phosphorus oxyanions. *RSC medicinal chemistry* **2022,** *13* (4), 375–391. DOI: 10.1039/d1md00297j.

155. Auld, D. S.; Thorne, N.; Maguire, W. F.; Inglese, J. Mechanism of PTC124 activity in cell-based luciferase assays of nonsense codon suppression. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** *106* (9), 3585–3590. DOI: 10.1073/pnas.0813345106.

156. Otto, A.; Bernhardt, J.; Hecker, M.; Becher, D. Global relative and absolute quantitation in microbial proteomics. *Current opinion in microbiology* **2012,** *15* (3), 364–372. DOI: 10.1016/j.mib.2012.02.005.

157. Crouch, S. P.; Kozlowski, R.; Slater, K. J.; Fletcher, J. The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity. *Journal of immunological methods* **1993,** *160* (1), 81–88. DOI: 10.1016/0022-1759(93)90011-u.

158. Moe, O. A.; Butler, L. G. Yeast Inorganic Pyrophosphatase. *The Journal of biological chemistry* **1972,** *247* (22), 7308–7314. DOI: 10.1016/S0021-9258(19)44630-9.

159. Sato, M.; Fujisaki, S.; Sato, K.; Nishimura, Y.; Nakano, A. Yeast Saccharomyces cerevisiae has two cis-prenyltransferases with different properties and localizations. Implication for their distinct physiological roles in dolichol synthesis. *Genes to cells : devoted to molecular & cellular mechanisms* **2001,** *6* (6), 495–506. DOI: 10.1046/j.1365-2443.2001.00438.x.

160. van Eunen, K.; Bouwman, J.; Daran-Lapujade, P.; Postmus, J.; Canelas, A. B.; Mensonides, F. I. C.; Orij, R.; Tuzun, I.; van den Brink, J.; Smits, G. J.; van Gulik, W. M.; Brul, S.; Heijnen, J. J.; Winde, J. H. de; Mattos, M. J. T. de; Kettner, C.; Nielsen, J.; Westerhoff, H. V.; Bakker, B. M. Measuring enzyme activities under standardized in vivo-like conditions for systems biology. *The FEBS journal* **2010,** *277* (3), 749–760. DOI: 10.1111/j.1742-4658.2009.07524.x.

161. Koetsier, M. J.; Jekel, P. A.; van den Berg, M. A.; Bovenberg, R. A. L.; Janssen, D. B. Characterization of a phenylacetate–CoA ligase from Penicillium chrysogenum. *Biochemical Journal* **2009,** *417* (2), 467–476. DOI: 10.1042/BJ20081257.

162. Huang, Z.; Poulter C. D. Trisammonium geranyl diphosphate. *Org. Synth.* **1988,** *66*, 211. DOI: 10.15227/orgsyn.066.0211.

163. Stehle, F.; Stubbs, M. T.; Strack, D.; Milkowski, C. Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. *The FEBS journal* **2008,** *275* (4), 775–787. DOI: 10.1111/j.1742-4658.2007.06244.x.

164. Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **2009,** *6* (5), 343–345. DOI: 10.1038/nmeth.1318.

165. Liu, H.; Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC biotechnology* **2008,** *8*, 91. DOI: 10.1186/1472-6750-8-91.

166. Inoue, H.; Nojima, H.; Okayama, H. High efficiency transformation of Escherichia coli with plasmids. *Gene* **1990,** *96* (1), 23–28. DOI: 10.1016/0378-1119(90)90336-p.

167. Gietz, R. D.; Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* **2007,** *2* (1), 31–34. DOI: 10.1038/nprot.2007.13.

168. Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Thiel, K.; Wiswedel, B. KNIME - the Konstanz information miner. *SIGKDD Explor. Newsl.* **2009,** *11* (1), 26–31. DOI: 10.1145/1656274.1656280.

Supplements

I. Abbreviations

2-AG	2-arachidonoyl glycerol
Amp	Ampicilin
a.u.	arbitrary unit
C. sativa	*Cannabis sativa*
CB1	Cannabinoid receptor 1
CB2	Cannabinoid receptor 2
CBC-C5	Cannabichromen
CBCA-C5	Cannabichromenic acid
CBD-C1	Cannabidiresorcin
CBD-C5	Cannabidiol
CBDA-C5	Cannabidiolic acid
CBG-C3	Cannabigerovarin
CBG-C5	Cannabigerol
CBGA-BB	Cannabibenphenylgerolic acid
CBGA-C3	Cannabigerovarin
CBGA-C5	Cannabigerolic acid
CBGA-ST	Cannabistyrengerolic acid
CHS	Chalcone synthase
CoA	Coenzyme A
*Cs*AAE1	Acyl activating enzyme 1
*Cs*AAE3	Acyl activating enzyme 3
*Cs*CBCAS	Cannabichromenic acid synthase
*Cs*CBDAS	Cannabindiolic acid synthase

CNS	Central nervous system
*Cs*OAC	Olivetolic acid cyclase
*Cs*OLS	Olivetol synthase
*Cs*PT4	Cannabis sativa prenyltransferase 4
DAD	Diode array detector
DMAPP	Dimethylallyl diphosphate
E. coli	*Escherichia coli*
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FAD	Flavine adenine dinucleotide
FAS	Fatty acid synthase
FPP	Farnesyl diphosphate
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
GPP	Geranyl diphosphate
GSPP	Geranyl thiolodiphosphate
HIV	Human immunodeficiency virus
HPLC	High performance liquid chromatoraphy
HRMS	High resolution mass spectrometry
HTAL	Hexanoyltriacetic acid lacton
IMAC	Immobilized metal ion affinity chromatography
IPP	Isopentenyl diphosphate
LB	Lysogeny broth
MD	Molecular dynamics

MEP	2-methyl-d-erythritol-4-phosphate
MVA	Mevalonic acid
N. benthamiana	*Nicotiana benthamiana*
OA	Olivetolic acid
2-*O*-GOA	2-O-geranyl olivetolic acid
OL	Olivetol
*Pc*PCL	Phenylacetate-CoA ligase
PDAL	Pentyl diacetic acid lactone
PNS	Peripheral nervous system
S. cerevisiae	*Saccharomyces cerevisiae*
STS	Stilbene synthase
THC-C5	(−)-*trans*-Δ9-tetrahydrocannabinol
THCA-BB	Tetrahydrocannabibiphenylic acid
THCA-C3	Tetrahydrocannabivaranic acid
THCA-C5	Tetrahydrocannabinolic acid
Wt	Wild type
YNB	Yeast nitrogen base
YPD	Yeast Extract Pepton Dextrose Medium

II. List of Figures

III. List of Tables

IV. Supplementary materials

i. Molecular biology, Primer, Plasmids, Amino acid sequence

Table S1: List of all primer used for strain generation.

Primer	Sequence (5' → 3')	Application
YPRCtau3 upstream fw	CAGCACCTTGAATTTCATGTTG	Generation of ySP01 & ySP03
YPRCtau3 upstream rv	AACAACCTCTTGCTATCAAACTT	Generation of ySP01 & ySP03
YPRCtau3 downstream fw	CTCTCCAATACAGCGTTACC	Generation of ySP01 & ySP03
YPRCtau3 downstream rv	TCCCCTAGCTGAACAACT	Generation of ySP01 & ySP03
YPRCtau3 AAE1 ins fw	TTGGCAACCCAAGACTCGGCATACCATATTGGTA ACGCTGTATTGGAGAGACTCACTATAGGGCGAAT TGGG	Generation of ySP01 & ySP03
YPRCtau3 AAE1 ins rv	CGAGATATCTGCAATAAAAGCAAAAGTAAGTTTG ATAGCAAGAGGTTGTTCATGGAATGCGTGCGATG AG	Generation of ySP01 & ySP03
pCFB2791 cut OAC out fw	GTGCTTCCGAAGATCGCGTCAGCTGAA	Generation of ySP02
pCFB2791 cut OAC out rv	GCGATCTTCGGAAGCACCCATGAACCAC	Generation of ySP02
Trp uptream fw	GCTGAATTGCCACTGCTATC	Generation of ySP04
Trp upstream rv	TTTGAAAAATTGTGCATTAGCATGAGGTCGCTCA TCGCACGCATTCCATGTTTGTGTGCTTAATCACG TATACTC	Generation of ySP04
Trp downstream fw	GATATCAGATCCACTAGTGGCCTATGCACCCAAT TCGCCCTATAGTGAGTATTTAAGTATTGTTTGTG CACTTG	Generation of ySP04
Trp downstream rv	AACCGCTAAATGTTTTTGTTC	Generation of ySP04

Trp OLS-OAC int fw	GGGCATTGGTGACTATTGAGCACGTGAGTATACG TGATTAAGCACACAAACATGGAATGCGTGCGATG AG	Generation of ySP04
Trp OLS-OAC int rv	TGCTTGCTTTTCAAAAGGCCTGCAGGCAAGTGCA CAAACAATACTTAAATACTCACTATAGGGCGAAT TGGG	Generation of ySP04

Table S2: List of all primer used to conduct the plasmid pDio-OAC.

Primer	**Sequence (5' → 3')**
Vec-OAC-in-pDio fw	CTAGAGGGCCGCATCATGTAATTAG
Vec-OAC-in-pDio rv	GGTTTTTTCTCCTTGACGTTAAAGT
Ins-OAC-in-pDio-fw	ACTTTAACGTCAAGGAGAAAAAACCATGGCCGTTAAGCACTTA ATTG
Ins-OAC-in-pDio-rv	CTAATTACATGATGCGGCCCTCTAGTTACTTTCTAGGGGTGTA ATCGAA

Table S3: Primer used for mutational studies of *Cs*OAC (single exchange variants).

Primer	**Sequence (5' → 3')**
I73A_fw	AGACTACGCAATTCACCCAGCTCACGTTGGTTTCGGT
I73A_rv	GGTGAATTGCGTAGTCTTGAATAGTTTCGACAGATTCGAAGG
I73F_fw	AGACTACTTTATTCACCCAGCTCACGTTGGTTTCGGT
I73F_rv	GGGTGAATAAAGTAGTCTTGAATAGTTTCGACAGATTCGAAGG
I73L_fw	AGACTACTTAATTCACCCAGCTCACGTTGGTTTCGGT
I73L_rv	GGTGAATTAAGTAGTCTTGAATAGTTTCGACAGATTCGAAGG
I73M_fw	AGACTACATGATTCACCCAGCTCACGTTGGTTTCGGT
I73M_rv	GGTGAATCATGTAGTCTTGAATAGTTTCGACAGATTCGAAGG
I73V_fw	AGACTACGTTATTCACCCAGCTCACGTTGGTTTCGGT
I73V_rv	GGTGAATAACGTAGTCTTGAATAGTTTCGACAGATTCGAAGG

G82A_fw	TTTCGCTGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATCTTCG
G82A_rv	TCTGTAGACATCAGCGAAACCAACGTGAGCTGGGTGAATG
G82I_fw	TTTCATTGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATCTTCGAT
G82I_rv	GATCTGTAGACATCAATGAAACCAACGTGAGCTGGGTGAATG
G82Y_fw	GGTTTCTATGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATCT
G82Y_rv	TCTGTAGACATCATAGAAACCAACGTGAGCTGGGTGAATG
G82H_fw	TGGTTTCCATGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATC
G82H_rv	GTAGACATCATGGAAACCAACGTGAGCTGGGTGAATGA
G82E_fw	TGGTTTCGAAGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATC
G82E_rv	TAGACATCTTCGAAACCAACGTGAGCTGGGTGAATGAT
G82S_fw	TGGTTTCTCTGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATC
G82S_rv	TCTGTAGACATTCTCGAAACCAACGTGAGCTGGGTGAATG
R86A_fw	CTACGCTTCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCCCT
R86A_rv	TTTCCCAGAAAGAAGCGTAGACATCACCGAAACCAACGTGAGC
R86F_fw	CTACTTCTCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCC
R86F_rv	TTTCCCAGAAAGAGAAGTAGACATCACCGAAACCAACGTG
R86H_fw	CTACCACTCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCC
R86H_rv	TTTCCCAGAAAGAGTGGTAGACATCACCGAAACCAACGTG
R86L_fw	CTACTTGTCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCC
R86L_rv	TTTCCCAGAAAGACAAGTAG ACATCACCGAAACCAACGTG
R86S_fw	CTACTCTTCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCC
R86S_rv	TTTCCCAGAAAGAAGAGTAGACATCACCGAAACCAACGTG
W89A_fw	ATCTTCGATTACACCCCTAGAAAGTCTTTCGCTGAAAAGTTGTTG
W89A_rv	CAACAACTTTTCAGCGAAAGATCTGTAGACATCACCGAAACCAAC
W89H_fw	TCTTTCCATGAAAAGTTGTTGATCTTCGATTACACCCCTAGAA
W89H_rv	AACAACTTTTCATGGAAAGATCTGTAGACATCACCGAAACC
W89L_fw	TCTTTCTTAGAAAAGTTGTTGATCTTCGATTACACCCCTAGAAAG

W89L_rv	TCAACAACTTTTCTAAGAAAGATCTGTAGACATCACCGAAACC
W89N_fw	TCTTTCAATGAAAAGTTGTTGATCTTCGATTACACCCCTAGAAAG
W89N_rw	CAACAACTTTTCATTGAAAGATCTGTAGACATCACCGAAACC
W89T_fw	TCTTTCACTGAAAAGTTGTTGATCTTCGATTACACCCCTAGAA
W89T_rv	AACAACTTTTCAGTGAAAGATCTGTAGACATCACCGAAACC
L92A_fw	TGGGAAAAGGCTTTGATCTTCGATTACACCCCTAGAAAGTAACTAG
L92A_rv	AGATCAAAGCCTTTTCCCAGAAAGATCTGTAGACATCACCGAAAC
L92F_fw	CTGGGAAAAGCATTTGATCTTCGATTACACCCCTAGAAAGTAAC
L92F_rv	AGATCAAATGCTTTTCCCAGAAAGATCTGTAGACATCACCGAA
L92H_fw	TCTGGGAAAAGCATTTGATCTTCGATTACACCCCTAGAAAG
L92H_rv	TCAAATGCTTTTCCCAGAAAGATCTGTAGACATCACCGAAA
L92Q_fw	GGGAAAAGCAATTGATCTTCGATTACACCCCTAGAAAGTAACTA
L92Q_rv	AAGATCAATTGCTTTTCCCAGAAAGATCTGTAGACATCACCG
L92V_fw	TCTGGGAAAAGGTTTTGATCTTCGATTACACCCCTAGAAAG
L92V_rv	TCAAAACCTTTTCCCAGAAAGATCTGTAGACATCACCGAAA
L92I_fw	CTGGGAAAAGATTTTGATCTTCGATTACACCCCTAGAAAGTAAC
L92I_rv	AGATCAAAATCTTTTCCCAGAAAGATCTGTAGACATCACCGAA
L92Y_fw	CTGGGAAAAGTACTTGATCTTCGATTACACCCCTAGAAAGTAAC
L92Y_rv	AGATCAAGTACTTTTCCCAGAAAGATCTGTAGACATCACCGAA
L92W_fw	CTGGGAAAAGTGGTTGATCTTCGATTACACCCCTAGAAAGTAAC
L92W_rv	AGATCAACCTCTTTTCCCAGAAAGATCTGTAGACATCACCGAA
L92R_fw	CTGGGAAAAGAGATTGATCTTCGATTACACCCCTAGAAAGTAAC
L92R_rv	AGATCAATCTCTTTTCCCAGAAAGATCTGTAGACATCACCGAA
I94A_fw	AAGTTGTTGGCTTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94A_rv	AATCGAAAGCCAACAACTTTTCCCAGAAAGATCTGTAGACATCA
I94S_fw	AAGTTGTTGTCTTTCGATTACACCCCTAGAAAGTAACTAGAGG
I94S_rv	AATCGAATCCCAACAACTTTTCCCAGAAAGATCTGTAGACATCA

I94D_fw	AAGTTGTTGGATTTCGATTACACCCCTAGAAAGTAACTAGAGG
I94D_rv	TAATCGAAATCCAACAACTTTTCCCAGAAAGATCTGTAGACATC
I94C_fw	AAGTTGTTGTGTTTCGATTACACCCCTAGAAAGTAACTAGAGG
I94C_rv	TAATCGAAACACAACAACTTTTCCCAGAAAGATCTGTAGACATC
I94Q_fw	AAGTTGTTGCAATTCGATTACACCCCTAGAAAGTAACTAGAGG
I94Q_rv	TAATCGAATTGCAACAACTTTTCCCAGAAAGATCTGTAGACATC
I94L_fw	AAGTTGTTGCATTTTCGATTACACCCCTAGAAAGTAACTAGAGG
I94L_rv	TAATCGAACAACAACAACTTTTCCCAGAAAGATCTGTAGACATC
I94V_fw	AAGTTGTTGGTTTTCGATTACACCCCTAGAAAGTAACTAGAGG
I94V_rv	TAATCGAAAACCAACAACTTTTCCCAGAAAGATCTGTAGACATC
F27Y_fw	AGACCTATGTTAACTTGGTTAACATTATTCCAGCTATGAAGGACG
F27Y_rv	AACCAAGTTAACATAGGTCTTGAAGAATTCCTCTTTTTGAGCTTCAGTGA

Table S4: Primer used for mutational studies of *Cs*OAC (multiple exchange variants).

Primer	**Sequence (5' → 3')**
R86A in G82A rv	TTTCCCAGAAAGAAGCGTAGACATCAGCGAAACCAACGTGAGC
R86H in G82A rv	TTTCCCAGAAAGAGTGGTAGACATCAGCGAAACCAACGTGAGC
R86H in G82S rv	TTTCCCAGAAAGAGTGGTAGACATCAGAGAAACCAACGTGAGC
G82A in L92F fw	TTTCGCTGATGTCTACAGATTTTCTGGGAAAAGTTTTTGATCTTCG
G82A in L92Y fw	TTTCGCTGATGTCTACAGATTTTCTGGGAAAAGTACTTGATCTTCG
G82A in I94V fw	TTTCGCTGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGGTCTTCG
R86A in L92F fw	CTACGCTTCTTTCTGGGAAAAGTTTTTGATCTTCGATTACACCCCT
R86A in L92Y fw	CTACGCTTCTTTCTGGGAAAAGTACTTGATCTTCGATTACACCCCT
R86A in I94V fw	CTACGCTTCTTTCTGGGAAAAGTTGTTGGTCTTCGATTACACCC
R86H in L92F fw	CTACCACTCTTTCTGGGAAAAGTTTTTGATCTTCGATTACACCCC

R86H in L92I fw	CTACCACTCTTTCTGGGAAAAGATCTTGATCTTCGATTACACCCC
R86H in L92Y fw	CTACCACTCTTTCTGGGAAAAGTACTTGATCTTCGATTACACCCC
R86H in I94V fw	CTACCACTCTTTCTGGGAAAAGTTGTTGGTCTTCGATTACACCCCT
L92F/I94V fw	GGAAAAGTTTTTGCAATTCGATTACACCCCTAGAAAGTAACTAGAG
L92F/I94V rv	CGAAGACCAAAAACTTTTCCCAGAAAGATCTGTAGACATCACCGA
I94L in L92F fw	AAGTTTTTGTTGTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94L in L92I fw	AAGATTTTGTTGTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94L in L92Y fw	AAGTATTTGTTGTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94V in L92F fw	AAGTTTTTGGTTTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94V in L92Y fw	AAGTATTTGGTTTTCGATTACACCCCTAGAAAGTAACTAGAGGG
I94V in L92I fw	AAGATTTTGGTTTTCGATTACACCCCTAGAAAGTAACTAGAGGG
94L in R86H rv	TAATCGAACAACAACAACTTTTCCCAGAAAGAGTGGTAGACATC
G82A in R86A/L92F fw	TGGTTTCGCTGATGTCTACGCTTCTTTCTGGGAAAAGTTTTTGATC
I94V in R86H/L92Y rv	TAATCGAAAACCAAGTACTTTTCCCAGAAAGAGTGGTAGACATC
I94L in R86H/L92I rv	TAATCGAACAACAAAATCTTTTCCCAGAAAGAGTGGTAGACATC
I94L in R86H/L92F rv	TAATCGAACAACAAAAACTTTTCCCAGAAAGAGTGGTAGACATC
I94V in R86A/L92F rv	TAATCGAAAACCAAAAACTTTTCCCAGAAAGAAGCGTAGACATC
I94L in R86H/L92Y rv	TAATCGAACAACAAGTACTTTTCCCAGAAAGAGTGGTAGACATC
I94V in R86H/L92F rv	TAATCGAAAACCAAAAACTTTTCCCAGAAAGAGTGGTAGACATC
I94L in R86A/L92F rv	TAATCGAACAACAAAAACTTTTCCCAGAAAGAAGCGTAGACATC

I94V in R86H/L92I rv	TAATCGAAAACCAAAATCTTTTCCCAGAAAGAGTGGTAGACATC
I94L in G82S/R86H rv	TAATCGAACAACAACAACTTTTCCCAGAAAGAGTGGTAGACATC
I94V in R86A/L92Y rv	TAATCGAAGACCAAATACTTTTCCCAGAAAGAAGCGTAGACATC
G82A in R86A/L92F fw	TGGTTTCGCTGATGTCTACGCTTCTTTCTGGGAAAAGTTTTTGATC
G82A in R86H/L92Y fw	TGGTTTCGCTGATGTCTACCACTCTTTCTGGGAAAAGTACTTGATC
G82A in L92F/I94V fw	TGGTTTCGCGATGTCTACAGTTCTTTCTGGGAAAAGTTTTTGGTT
G82S in R86A/L92F fw	TGGTTTCTCGATGTCTACGCTTCTTTCTGGGAAAAGTTTTTGATC
G82S in L92F/I94V fw	TGGTTTCTCGATGTCTACAGTTCTTTCTGGGAAAAGTTTTTGGTT
G82S in R86H/L92I fw	TGGTTTCTCTGATGTCTACCACTCTTTCTGGGAAAAGATTTTGATC
G82S in R86H/L92Y fw	TGGTTTCTCTGATGTCTACCACTCTTTCTGGGAAAAGTACTTGATC
G82S in R86H/L92F fw	TGGTTTCTCTGATGTCTACCACTCTTTCTGGGAAAAGTTTTTGATC
G82S in R86H/I94V fw	TGGTTTCTCTGATGTCTACAGATCTTTCTGGGAAAAGTTTTTGGTT
G82S in R86A/L92Y/I94L fw	TGGTTTCTCTGATGTCTACGCTTCTTTCTGGGAAAAGTACTTGTTG
G82S in R86A/L92Y/ I94L fw	TGGTTTCTCTGATGTCTACGCTTCTTTCTGGGAAAAGTACTTGGTT
G82A in R86A/L92Y fw	TGGTTTCGCTGATGTCTACGCTTCTTTCTGGGAAAAGTACTTGATC

G82A in R86A/L92F fw	TGGTTTCGCTGATGTCTACGCTTCTTTCTGGGAAAAGTTTTTGATC
I94L in R86A/L92Y rv	TAATCGAACAACAAATACTTTTCCCAGAAAGAAGCGTAGACATC
I94V in G82A/L92Y rv	TAATCGAAGACCAAATACTTTTCCCAGAAAGATCTGTAGACATC
L92Y in G82A/R86A rv	AGATCAAATACTTTTCCCAGAAAGAAGCGTAGACATCAGAGAA
G82A in L92F/I94V fw	TTTCGCZGATGTCTACAGATCTTTCTGGGAAAAGTTTTTGGTCTTCG

Table S5: List of primer used for the construction of *Cs*OAC and *Cs*OLS fusion proteins.

Primer	Sequence (5' → 3')	Application
Ins pTDH3-OAC-tTDH1 fw	TTTTCTTGAATTTGCAGTTTGAAAAATCAGTTCGAGTTTATCATTATCAATACTGCCATT	Generation of pDio OLS-OAC
Ins pTDH3-OAC-tTDH1 rv	TTGGCCGATTCATTAATGCAGGGCCCGTTCAGGGTAATATATTTTAACCGCCGACG	
Vec pDio-OLS fw	GGCCCTGCATTAATGAATCGGCCAA	
Vec pDio-OLS rv	ATTTTTCAAACTGCAAATTCAAGAAAAAGCC	
OLS GGGGS OAC fw	CGGTGGTGGTGGTTCTATGGCCGTTAAGCACTTAATTGT	Generation of pDio OLS-GGGGS-OAC
OLS GGGGS OAC rv	TAGAACCACCACCACCGTACTTGATAGGAACAGACCTAACAA	
OLS EAAAK OAC fw	TATCAAGTACGAAGCTGCTGCTAAAATGGCCGTTAAGCACTTA	Generation of pDio OLS-EAAAK-OAC
OLS EAAAK OAC rv	TAACGGCCATTTTAGCAGCAGCTTCGTACTTGATAGGAACAGACCT	
Ins OLS-(GGGGS)3-OAC fw	TTGGTTTTGGTCCAGGTTTGACTGTCGAAAGAGTTGTTGTTAGGTCTGTT	

Ins OLS-(GGGGS)3-OAC rv	CTCTTTTTGAGCTTCAGTGATTTCATCCTTGAACTTCAAGACAATTAAGTG	Generation of pDio OLS-$(GGGGS)_3$-OAC
Ins OLS-(EAAAK)3-OAC fw	TTTGGTTTTGGTCCAGGTTTGACTGTCGAAAGAGTTGTTGTTAGGTCTG	Generation of pDio OLS-$(EAAAK)_3$-OAC
Ins OLS-(EAAAK)3-OAC rv	CTCTTTTTGAGCTTCAGTGATTTCATCCTTGAACTTCAAGACAATTAAGTG	
Ins OLS-(EAAAK)5-OAC fw	GTTTGACTGTCGAAAGAGTTGTTGTTAGGTCTGTTCCTATCAAGTACGA	Generation of pDio OLS-$(EAAAK)_5$-OAC
Ins OLS-(EAAAK)5-OAC rv	GTGATTTCATCCTTGAACTTCAAGACAATTAAGTGCTTAACGGCCATT	
Ins OLS-(GGGGS)5-OAC fw	GTTTGACTGTCGAAAGAGTTGTTGTTAGGTCTGTTCCTATCAAGTACG	Generation of pDio OLS-$(GGGGS)_5$-OAC
Ins OLS-(GGGGS)5-OAC rv	GTGATTTCATCCTTGAACTTCAAGACAATTAAGTGCTTAACGGCCATA	
Vec OLS-OAC fw	TGAAATCACTGAAGCTCAAAAAGAGGAA	Generation of pDio OLS-Linker-OAC
Vec OLS-OAC rv	ACAGTCAAACCTGGACCAAAACCAA	
Ins OAC fw	ATTCGCTTAAACACTATATCAATAAAAAAAAATGGCCGTTAAGCACTTAATTGTC	Generation of pDio OAC-OLS
Ins OAC rv	GACCTTCAGCTCTCAAGTGGTTCATCTTTCTAGGGGTGTAATCGAAGATCA	
Vec_pDio OLS fw	ATGAACCACTTGAGAGCTGAAGGTC	
Vec_pDio OLS rv	TTTTTTTTATTGATATAGTGTTTAAGCGAATGA	
Ins OAC-(Linker)5-OLS fw	ATCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACCCCTAGAAAGGA	Generation of pDio OAC-(Linker)5)OLS
Ins OAC-(Linker)5-OLS rv	ACCAATAGCCAAAACAGAAGCTGGACCTTCAGCTCTCAAGTGGTT	
Vec pDio OAC-OLS fw	TCCAGCTTCTGTTTTGGCTATTGGT	
Vec pDio OAC-OLS rv	GATCAACAACTTTTCCCAGAAAGATCTGTA	

Table S6: List of primer used for the construction of the plasmid for NphB expression.

	Sequence (5' → 3')
NphB_GFP_out fw	TTTGATTCTTTGGAAGATCACCACCACCATCATCATCATC
NphB_GFP_out rv	ATGATGATGGTGGTGGTGATCTTCCAAAGAATCAAAAGCCTTCA

Table S7: Amino acid sequences of all proteins involved in cannabinoid biosynthesis used in this study.

Name	**Sequence**
***Cs*AAE1**	MGKNYKSLDSVVASDFIALGITSEVAETLHGRLAEIVCNYGAAT PQTWINIANHILSPDLPFSLHQMLFYGCYKDFGPAPPAWIPDPE KVKSTNLGALLEKRGKEFLGVKYKDPISSFSHFQEFSVRNPEVY WRTVLMDEMKISFSKDPECILRRDDINNPGGSEWLPGGYLNSA KNCLNVNSNKKLNDTMIVWRDEGNDDLPLNKLTLDQLRKRV WLVGYALEEMGLEKGCAIAIDMPMHVDAVVIYLAIVLAGYVVV SIADSFSAPEISTRLRLSKAKAIFTQDHIIRGKKRIPLYSRVVEAKS PMAIVIPCSGSNIGAELRDGDISWDYFLERAKEFKNCEFTAREQ PVDAYTNILFSSGTTGEPKAIPWTQATPLKAAADGWSHLDIRKG DVIVWPTNLGWMMGPWLVYASLLNGASIALYNGSPLVSGFAKF VQDAKVTMLGVVPSIVRSWKSTNCVSGYDWSTIRCFSSSGEASN VDEYLWLMGRANYKPVIEMCGGTEIGGAFSAGSFLQAQSLSSFS SQCMGCTLYILDKNGYPMPKNKPGIGELALGPVMFGASKTLLN GNHHDVYFKGMPTLNGEVLRRHGDIFELTSNGYYHAHGRAD DTMNIGGIKISSIEIERVCNEVDDRVFETTAIGVPPLGGGPEQLV IFFVLKDSNDTTIDLNQLRLSFNLGLQKKLNPLFKVTRVVPLSSL PRTATNKIMRRVLRQQFSHFE
***Cs*OLS**	MNHLRAEGPASVLAIGTANPENILLQDEFPDYYFRVTKSEHMT QLKEKFRKICDKSMIRKRNCFLNEEHLKQNPRLVEHEMQTLDA RQDMLVVEVPKLGKDACAKAIKEWGQPKSKITHLIFTSASTTD MPGADYHCAKLLGLSPSVKRVMMYQLGCYGGGTVLRIAKDIAE NNKGARVLAVCCDIMACLFRGPSESDLELLVGQAIFGDGAAAVI VGAEPDESVGERPIFELVSTGQTILPNSEGTIGGHIREAGLIFDL HKDVPMLISNNIEKCLIEAFTPIGISDWNSIFWITHPGGKAILDK VEEKLHLKSDKFVDSRHVLSEHGNMSSSTVLFVMDELRKRSLEE GKSTTGDGFEWGVLFGFGPGLTVERVVVRSVPIKY
***Cs*OAC**	MAVKHLIVLKFKDEITEAQKEEFFKTYVNLVNIIPAMKDVYWG KDVTQKNKEEGYTHIVEVTFESVETIQDYIIHPAHVGFGDVYRS FWEKLLIFDYTPRK

NphB wt	MSEAADVERVYAAMEEAAGLLGVACARDKIYPLLSTFQDTLVE GGSVVVFSMASGRHSTEDFSISVPTSHGDPYATVVEKGLFPATG HPHLPVSMFAIDGEVTGGFKKTYAFFPTDNMPGVAELSAIPSPP AVAENAELFARYGLDKVQMTSMDYKKRQVNLYFSELSAQTLE AESVLALVRELGLHVPNELGLKFCKSFSVYPTLNWETGKIDRLC FAVISNDPTLVPSSDEGDIEKFHNYATKAPYAYVGEKRTLVYGL TLSPKEYYKLGAYYHITDVQRGLLKAFDSLED
NphB G286S/Y288A	MSEAADVERVYAAMEEAAGLLGVACARDKIYPLLSTFQDTLVE GGSVVVFSMASGRHSTEDFSISVPTSHGDPYATVVEKGLFPATG HPHLPVSMFAIDGEVTGGFKKTYAFFPTDNMPGVAELSAIPSPP AVAENAELFARYGLDKVQMTSMDYKKRQVNLYFSELSAQTLE AESVLALVRELGLHVPNELGLKFCKSFSVYPTLNWETGKIDRLC FAVISNDPTLVPSSDEGDIE KFHNYATKAPYAYVGEKRTLVYGLTLSPKEYYKLSAAYHITDV QRGLLKAFDSLED
***Cs*THCAS fused with a N-terminal vacuole tag**	MSSLKALLPLALLLVSANQVAAISLQRPLGLDKDVLLQANPREN FLKCFSKHIPNNVANPKLVYTQHDQLYMSILNSTIQNLRFISDTT PKPLVIVTPSQNSHIQATILCSKKVGLQIRTRSGGHDAEGMSYIS QVPFVVVDLRNMHSIKIDVHSQTAWVEAGATLGEVYYWINEK NENLSFPGGYCPTVGVGGHFSGGGYGALMRNYGLAADNIIDA HLVNVDGKVLDRKSMGEDLFWAIRGGGGENFGIIAAWKIKLV AVPSKSTIFSVKKNMEIHGLVKLFNKWQNIAYKYDKDLVLMTH FITKNITDNHGKNKTTVHGYFSSIFHGGVDSLVDLMNKSFPEL GIKKTDCKEFSWIDTTIFYSGVVNFNTANFKKEILLDRSAGKKT AFSIKLDYVKKPIPETAMVKILEKLYEEDVGAGMYVLYPYGGIM EEISESAIPFPHRAGIMYELWYTASWEKQEDNEKHINWVRSVYN FTTPYVSQNPRLAYLNYRDLDLGKTNCASPNQYTQARIWGEKY FGKNFNRLVKVKTKVDPNNFFCNEQSIPPLPPHHH

Table S8: Nucleotide sequences of all proteins involved in cannabinoid biosynthesis used in this study.

Name	Sequence
***Cs*AAE1**	ATGGGCAAGAACTACAAGTCCTTGGATTCTGTTGTTGCCTCTGATTTTA TTGCCTTGGGTATTACTTCTGAAGTTGCCGAAACTTTACATGGTAGATT GGCTGAAATTGTCTGTAATTATGGTGCTGCTACTCCACAAACCTGGATT AACATTGCTAACCATATCTTGTCTCCAGACTTGCCATTTTCCTTGCATC AAATGTTGTTCTACGGTTGCTACAAGGATTTTGGTCCAGCTCCACCAGC TTGGATTCCAGATCCAGAAAAAGTTAAGTCTACCAACTTGGGTGCTTTG TTGGAAAAGAGAGGCAAAGAATTTTTGGGCGTTAAGTACAAGGATCCCA TCTCTTCATTTTCCCACTTCCAAGAATTCTCCGTTAGAAACCCAGAAGT TTATTGGAGAACTGTCTTGATGGACGAGATGAAGATTTCATTCTCTAAG GATCCAGAGTGCATCTTGAGAAGAGATGATATTAACAATCCAGGTGGTT CTGAATGGTTGCCTGGTGGTTATTTGAATTCTGCTAAGAACTGCTTGAA CGTCAACTCCAACAAAAAGTTGAACGACACTATGATCGTTTGGAGAGAT

	GAGGGTAATGATGATTTGCCATTGAACAAGTTGACCTTGGACCAATTGA
	GAAAGAGAGTTTGGTTAGTTGGTTACGCCTTGGAAGAAATGGGTTTAGA
	AAAAGGTTGCGCTATTGCTATCGATATGCCAATGCATGTTGATGCCGTT
	GTTATCTATTTGGCTATAGTTTTGGCTGGTTACGTCGTTGTTTCAATTG
	CCGATTCATTTTCTGCTCCAGAAATCTCTACTAGGTTGAGATTGTCTAA
	GGCTAAGGCTATTTTCACCCAAGATCATATCATCAGAGGTAAGAAGAGA
	ATTCCCTTGTACTCCAGAGTTGTTGAAGCTAAATCTCCAATGGCTATCG
	TTATTCCATGTTCCGGTTCTAATATTGGTGCCGAATTGAGAGATGGTGA
	TATCTCTTGGGATTATTTCTTGGAAAGGGCCAAAGAATTCAAGAACTGT
	GAGTTTACTGCTAGAGAACAACCAGTTGACGCTTACACTAATATCTTGT
	TCTCTTCTGGTACTACCGGTGAACCTAAAGCTATTCCATGGACTCAAGC
	TACTCCATTGAAAGCTGCTGCTGATGGTTGGTCACATTTGGATATTAGA
	AAGGGTGATGTTATCGTCTGGCCAACTAACTTAGGTTGGATGATGGGTC
	CTTGGTTGGTTTATGCTTCTTTGTTGAATGGTGCATCCATTGCCTTGTA
	TAATGGTTCTCCATTGGTTTCTGGTTTCGCCAAGTTTGTTCAAGATGCT
	AAGGTTACTATGTTGGGTGTTGTTCCATCTATCGTTAGATCTTGGAAAT
	CCACCAATTGTGTCTCTGGTTATGATTGGTCTACTATCAGATGCTTCTC
	ATCTTCTGGTGAAGCTTCTAACGTTGATGAATACTTGTGGTTGATGGGT
	AGAGCTAATTACAAGCCAGTTATTGAAATGTGCGGTGGTACTGAAATTG
	GTGGTGCTTTTTCTGCTGGTTCATTCTTGCAAGCTCAATCCTTGTCATC
	CTTTTCTTCTCAATGTATGGGTTGCACCTTGTACATCTTGGATAAGAAT
	GGTTATCCCATGCCAAAAAACAAACCAGGTATTGGTGAATTGGCTTTGG
	GTCCAGTTATGTTTGGTGCTTCTAAAACCTTGTTGAACGGTAACCATCA
	CGATGTTTACTTTAAGGGTATGCCAACTTTGAACGGTGAAGTTTTGAGA
	AGGCACGGTGATATTTTCGAATTGACTTCTAACGGTTACTACCATGCTC
	ACGGTAGAGCTGATGATACAATGAATATCGGTGGTATCAAGATCTCCTC
	CATCGAAATTGAAAGAGTCTGCAACGAAGTTGACGACAGAGTTTTTGAA
	ACTACCGCTATTGGTGTTCCACCATTAGGTGGTGGTCCAGAACAATTGG
	TTATTTTCTTCGTTCTGAAGGACTCCAACGATACCACCATTGATTTGAA
	TCAACTGAGGCTGTCTTTTAACTTGGGCTTGCAAAAGAAGTTGAACCCC
	TTGTTTAAAGTTACCAGAGTCGTTCCATTGTCCTCATTGCCAAGAACTG
	CAACTAACAAGATCATGAGAAGAGTCTTGAGACAACAGTTCTCTCACTT
	CGAATGA
***Cs*OLS**	ATGAACCACTTGAGAGCTGAAGGTCCAGCTTCTGTTTTGGCTATTGGTA
	CTGCTAATCCAGAGAACATCTTGTTGCAAGATGAATTTCCAGACTACTA
	CTTCAGAGTCACTAAGTCTGAACACATGACCCAGTTGAAAGAGAAGTTC
	AGAAAGATTTGCGACAAGTCCATGATCAGAAAGAGAAACTGTTTCTTGA
	ACGAGGAACACTTGAAGCAAAACCCAAGATTGGTTGAACACGAAATGCA
	AACATTGGATGCCAGACAAGATATGTTGGTTGTTGAAGTTCCAAAGTTG
	GGTAAAGATGCTTGTGCTAAGGCTATCAAAGAATGGGGTCAACCTAAAT
	CCAAGATCACCCATTTGATTTTCACCTCTGCTTCTACTACTGATATGCC
	AGGTGCTGATTACCATTGTGCCAAATTATTGGGTTTGTCCCCATCTGTA
	AAAAGGGTCATGATGTATCAATTGGGTTGTTATGGTGGTGGTACTGTTT
	TGAGAATTGCTAAGGATATTGCCGAGAACAACAAAGGTGCTAGAGTTTT
	GGCTGTTTGCTGTGATATTATGGCTTGTTTGTTCAGAGGTCCATCCGAA
	TCTGATTTGGAATTATTGGTTGGTCAAGCCATCTTTGGTGATGGTGCTG
	CTGCTGTTATAGTTGGTGCTGAACCAGATGAATCTGTTGGTGAAAGACC
	AATCTTCGAATTGGTTTCTACTGGTCAAACTATCTTGCCAAACTCCGAA
	GGTACTATTGGTGGTCATATCAGAGAAGCCGGTTTGATTTTTGACTTGC
	ATAAGGATGTTCCCATGCTGATCTCTAACAACATTGAAAAGTGTTTGAT

	CGAGGCTTTCACCCCAATTGGTATTTCTGATTGGAACTCCATTTTCTGG ATTACTCATCCAGGTGGTAAAGCTATCTTGGATAAGGTAGAAGAAAAGC TGCATCTGAAGTCCGATAAGTTCGTTGATTCAAGACACGTCTTGTCTGA ACATGGTAACATGTCATCTTCTACCGTTTTGTTCGTTATGGACGAATTG AGAAAGAGGTCTTTGGAAGAAGGTAAATCTACTACCGGTGATGGTTTTG AATGGGGTGTTTTGTTTGGTTTTGGTCCAGGTTTGACTGTCGAAAGAGT TGTTGTTAGGTCTGTTCCTATCAAGTACTGA
***Cs*OAC**	ATGGCCGTTAAGCACTTAATTGTCTTGAAGTTCAAGGATGAAATCACTG AAGCTCAAAAAGAGGAATTCTTCAAGACCTATGTTAACTTGGTTAACAT TATTCCAGCTATGAAGGACGTCTACTGGGGTAAGGACGTTACCCAAAAA AACAAGGAAGAGGGTTACACTCACATCGTCGAAGTTACCTTCGAATCTG TCGAAACTATTCAAGACTACATCATTCACCCAGCTCACGTTGGTTTCGG TGATGTCTACAGATCTTTCTGGGAAAAGTTGTTGATCTTCGATTACACC CCTAGAAAGTAA
NphB wt	ATGAGCGAAGCGGCGGATGTGGAACGCGTGTATGCGGCGATGGAAGAAG CGGCGGGCCTGCTGGGCGTGGCGTGCGCGCGCGATAAAATTTATCCGCT GCTGAGCACCTTTCAGGATACCCTGGTGGAAGGCGGCAGCGTGGTGGTG TTTAGCATGGCGAGCGGCCGCCATAGCACCGAAGATTTTAGCATTAGCG TGCCGACCAGCCATGGCGATCCGTATGCGACCGTGGTGGAAAAAGGCCT GTTTCCGGCGACCGGCCATCCGCATCTGCCGGTGAGCATGTTTGCGATT GATGGCGAAGTGACCGGCGGCTTTAAAAAAACCTATGCGTTTTTTCCGA CCGATAACATGCCGGGCGTGGCGGAACTGAGCGCGATTCCGAGCCCGCC GGCGGTGGCGGAAAACGCGGAACTGTTTGCGCGCTATGGCCTGGATAAA GTGCAGATGACCAGCATGGATTATAAAAAACGCCAGGTGAACCTGTATT TTAGCGAACTGAGCGCGCAGACCCTGGAAGCGGAAAGCGTGCTGGCGCT GGTGCGCGAACTGGGCCTGCATGTGCCGAACGAACTGGGCCTGAAATTT TGCAAAAGCTTTAGCGTGTATCCGACCCTGAACTGGGAAACCGGCAAAA TTGATCGCCTGTGCTTTGCGGTGATTAGCAACGATCCGACCCTGGTGCC GAGCAGCGATGAAGGCGATATTGAAAAATTTCATAACTATGCGACCAAA GCGCCGTATGCGTATGTGGGCGAAAAACGCACCCTGGTGTATGGCCTGA CCCTGAGCCCGAAAGAATATTATAAACTGGGCGCGTATTATCATATTAC CGATGTGCAGCGCGGCCTGCTGAAAGCGTTTGATAGCCTGGAAGAT
NphB G286S/Y288A	ATGAGCGAAGCGGCGGATGTGGAACGCGTGTATGCGGCGATGGAAGAAG CGGCGGGCCTGCTGGGCGTGGCGTGCGCGCGCGATAAAATTTATCCGCT GCTGAGCACCTTTCAGGATACCCTGGTGGAAGGCGGCAGCGTGGTGGTG TTTAGCATGGCGAGCGGCCGCCATAGCACCGAAGATTTTAGCATTAGCG TGCCGACCAGCCATGGCGATCCGTATGCGACCGTGGTGGAAAAAGGCCT GTTTCCGGCGACCGGCCATCCGCATCTGCCGGTGAGCATGTTTGCGATT GATGGCGAAGTGACCGGCGGCTTTAAAAAAACCTATGCGTTTTTTCCGA CCGATAACATGCCGGGCGTGGCGGAACTGAGCGCGATTCCGAGCCCGCC GGCGGTGGCGGAAAACGCGGAACTGTTTGCGCGCTATGGCCTGGATAAA GTGCAGATGACCAGCATGGATTATAAAAAACGCCAGGTGAACCTGTATT TTAGCGAACTGAGCGCGCAGACCCTGGAAGCGGAAAGCGTGCTGGCGCT GGTGCGCGAACTGGGCCTGCATGTGCCGAACGAACTGGGCCTGAAATTT TGCAAAAGCTTTAGCGTGTATCCGACCCTGAACTGGGAAACCGGCAAAA TTGATCGCCTGTGCTTTGCGGTGATTAGCAACGATCCGACCCTGGTGCC GAGCAGCGATGAAGGCGATATTGAAAAATTTCATAACTATGCGACCAAA GCGCCGTATGCGTATGTGGGCGAAAAACGCACCCTGGTGTATGGCCTGA

	CCCTGAGCCCGAAAGAATATTATAAACTGAGCGCGGCGTATCATATTAC CGATGTGCAGCGCGGCCTGCTGAAAGCGTTTGATAGCCTGGAAGAT
***Cs*THCAS fused with a N-terminal vacuole tag**	**ATGTCCAGCTTGAAAGCATTATTGCCATTGGCCTTGTTGTTGGTCAGCG** **CCAACCAAGTTGCTGCAATCTCATTGCAAAGACCGTTGGGTCTAGATAA** **GGACGTTTTGCTGCAAGCT**AACCCTAGAGAAAACTTTTTGAAATGCTTC TCCAAACACATCCCAAACAACGTTGCCAACCCAAAATTGGTCTATACCC AACACGATCAATTGTATATGTCCATCTTGAACTCCACCATCCAAAACTT GAGATTCATATCCGACACTACCCCTAAACCATTGGTAATAGTCACCCCT TCCCAAAACTCACACATCCAAGCCACCATCTTGTGCTCTAAAAAAGTCG GTTTACAAATCAGAACCAGATCCGGTGGTCATGACGCCGAAGGTATGTC CTATATCTCACAAGTCCCATTCGTTGTCGTCGATTTGAGAAACATGCAC TCCATCAAAATCGATGTTCACTCTCAAACTGCCTGGGTAGAAGCTGGTG CCACATTGGGTGAAGTCTATTATTGGATCAACGAAAAAAATGAAAACTT GTCCTTCCCTGGTGGTTATTGTCCAACTGTCGGTGTTGGTGGTCACTTT TCTGGTGGTGGTTATGGTGCTTTAATGAGAAATTATGGTTTAGCCGCCG ATAACATCATAGACGCCCACTTGGTAAACGTCGACGGTAAAGTCTTGGA TAGAAAATCCATGGGTGAAGATTTGTTCTGGGCCATTAGAGGTGGTGGT GGTGAAAATTTCGGTATTATCGCCGCCTGGAAAATCAAATTGGTTGCCG TCCCTTCCAAATCCACCATCTTTTCCGTAAAAAAAAACATGGAAATCCA CGGTTTAGTCAAATTGTTCAACAAATGGCAAAACATTGCCTATAAGTAT GACAAAGACTTGGTCTTAATGACCCACTTCATCACCAAAAACATCACCG ATAATCACGGTAAAAACAAAACAACCGTTCACGGTTATTTCTCTTCCAT CTTCCACGGTGGTGTCGATTCTTTGGTCGACTTGATGAACAAATCCTTT CCTGAATTGGGTATCAAAAAAACCGATTGTAAGGAATTTTCCTGGATCG ATACTACCATCTTTTATTCCGGTGTCGTCAACTTCAATACTGCCAACTT CAAAAAGGAAATCTTATTGGACAGATCCGCCGGTAAAAAAACTGCCTTC TCCATCAAATTGGATTATGTCAAAAAACCAATCCCTGAAACCGCTATGG TCAAAATCTTGGAAAAATTGTATGAAGAAGATGTCGGTGCCGGTATGTA TGTCTTGTATCCATATGGTGGTATAATGGAAGAAATCTCCGAATCCGCC ATCCCATTCCCTCATAGAGCTGGTATCATGTATGAATTGTGGTATACCG CTTCCTGGGAAAAACAAGAAGATAACGAAAAACACATCAACTGGGTTAG ATCCGTCTATAACTTCACAACCCCATATGTCTCCCAAAATCCTAGATTG GCCTATTTGAACTATAGAGACTTGGACTTGGGTAAAACAAACTGTGCTT CCCCAAACCAATATACTCAAGCCAGAATCTGGGGTGAAAAGTATTTCGG TAAAAACTTCAATAGATTAGTCAAAGTCAAGACTAAAGTCGACCCAAAC AACTTCTTTTGTAACGAACAATCCATCCCTCCATTGCCTCCACATCACC ATTAA

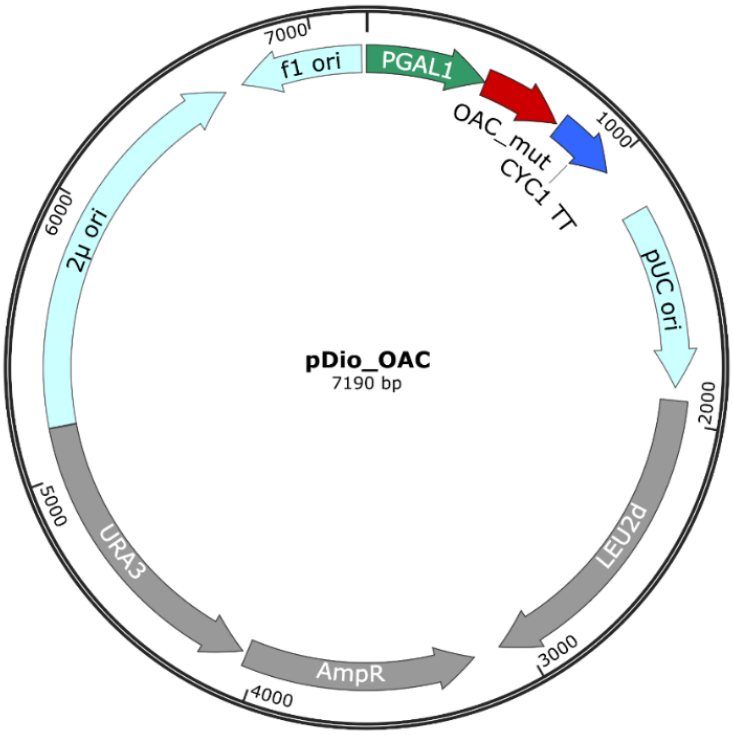


Figure S1: Expression vector pDionysos *Cs*OAC (pDio-OAC). The plasmid uses the pYES2 yeast expression vector (Fisher Scientific GmbH, Schwerte, Germany) as background and is modified by an additional Leu2d gene. The *oac* gene is flanked by the pGAL1 promoter and the CYC1 terminator (*163*).

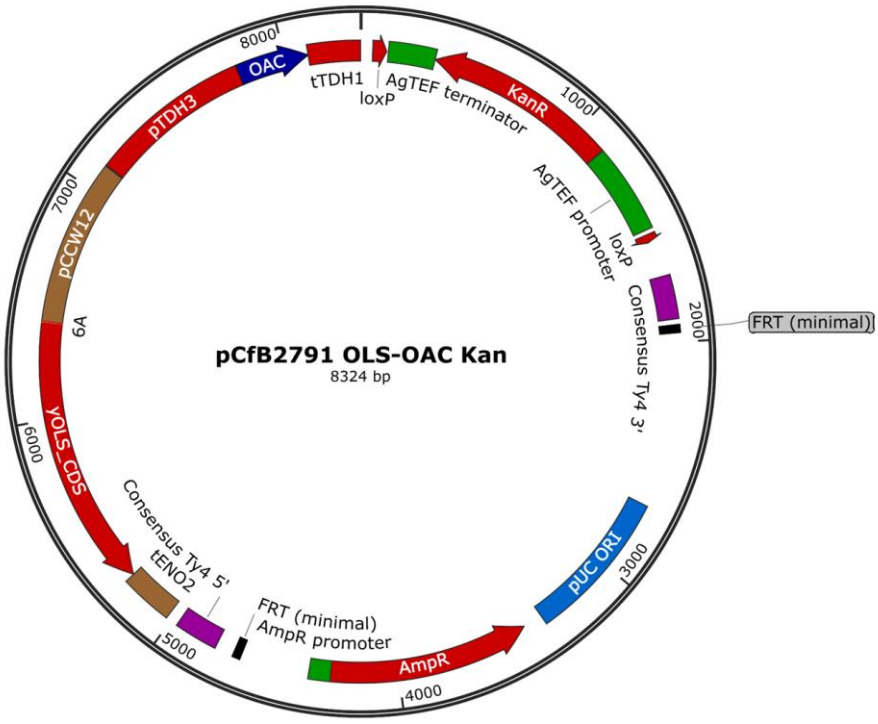


Figure S2: Vector pCfB2791 OLS-OAC KAN for genomic integration (*131*).

ii. Supplements – Results and Discussion

	chain	position	residue	mutable	non-essential	in tunnel	in catalytic pocket	HotSpot
Chain A								
	A	86	Arg	✓	✓	✓	✗	✓
	A	9	Leu	✓	✓	✓	✓	✓
	A	23	Phe	✓	✓	✗	✓	✓
	A	73	Ile	✓	✓	✓	✓	✓
	A	82	Gly	✓	✓	✓	✗	✓
	A	89	Trp	✓	✓	✓	✓	✓
	A	92	Leu	✓	✓	✓	✓	✓
	A	93	Leu	✓	✓	✓	✗	✓
	A	7	Ile	?	✓	✓	✓	?
	A	27	Tyr	?	✓	✓	✓	?
	A	49	Lys	?	✓	✓	✓	?
	A	59	Val	?	✓	✗	✓	?
	A	81	Phe	?	✓	✓	✓	?
	A	94	Ile	?	✓	✓	✓	?

Figure S3: List of all functional HotSpot of the protein *Cs*OAC predicted using HotSpot Wizard (*129*).

Table S9: **Post hoc analysis data for OA concentration in Figure 18A &22A.** The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

***Cs*OAC variant**	**Mean Difference**	**SEM**	**t-value**	**p-value**	**Alpha**	**Sig**
I73A	-0.05204	0.01181	-11.54	3.22E-04	0.05	1
G82A	0.02472	0.00591	5.31	0.00606	0.05	1
R86A	0.00926	0.00301	1.89	0.13137	0.05	0
W89A	-0.02268	0.00546	-5.04	0.00729	0.05	1
L92A	-0.03517	0.00814	-7.43	0.00175	0.05	1
I94A	-0.0222	0.0054	-4.67	0.00951	0.05	1
R86H	-0.00931	0.00279	-2.25	0.08796	0.05	0
L92F	0.01316	0.00563	1.23	0.28721	0.05	0
I94L	-0.03687	0.00852	-7.68	0.00154	0.05	1
I94V	0.02402	0.00575	5.26	0.00627	0.05	1
G82S	-0.02695	0.00303	-6.42	0.00303	0.05	1
L92I	-0.02756	0.00646	-6.38	0.00309	0.05	1
L92Y	0.04812	0.01111	7.80	0.00146	0.05	1

Table S10: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 18B& 22B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

***Cs*OAC variant**	**Mean Difference**	**SEM**	**t-value**	**p-value**	**Alpha**	**Sig**
I73A	0.06541	0.00756	8.65	0.01311	0.05	1
G82A	0.06837	0.01533	64.09	1.37E-05	0.05	1
R86A	0.02789	0.00431	6.47	0.00294	0.05	1
W89A	-0.05771	0.00536	-10.77	0.00208	0.05	1
L92A	-0.08925	0.02015	-14.26	1.41E-04	0.05	1
I94A	-0.03741	0.00875	-6,49	0.00291	0.05	1
R86H	-0.00691	0.00897	-5.46	0.00548	0.05	1
L92F	0.08571	0.0195	10.69	4.33E-04	0.05	1
I94L	0.0123	0.00309	3.89	0.01765	0.05	1
I94V	0.06475	0.01466	12.53	2.33E-04	0.05	1
G82S	0.03447	0.00782	11.40	3.37E-04	0.05	1
L92I	-0.00125	0.00191	-0.29	0.78299	0.05	0
L92Y	0.08511	0.01909	25.97	1.31E-05	0.05	1

Table S11: Post hoc analysis data for OA concentration in Figure 21A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
I73F	-0.13253	0.02969	-33.50	4.73E-06	0.05	1
I73L	-0.0229	0.00585	-3.63	0.02218	0.05	1
I73M	-0.02694	0.00633	-6.22	0.00339	0.05	1
I73V	-0.07932	0.01777	-30.85	6.58E-06	0.05	1
G82E	-0.23838	0.05332	-82.58	1.29E-07	0.05	1
G82H	-0.10953	0.02452	-40.90	2.14E-06	0.05	1
G82I	-0.23381	0.05229	-97.55	6.62E-08	0.05	1
G82Y	-0.12019	0.02692	-33.88	4.53E-06	0.05	1
G82V	-0.18864	0.0422	-74.85	1.91E-07	0.05	1
R86F	-0.08181	0.01835	-26.17	1.27E-05	0.05	1
R86L	-0.14016	0.03179	-11.76	2.99E-04	0.05	1
W89L	-0.15669	0.03506	-56.42	5.91E-07	0.05	1
W89N	-0.19971	0.04467	-77.68	1.65E-07	0.05	1
W89T	-0.1353	0.0303	-35.39	3.80E-06	0.05	1
L92Q	-0.15462	0.03475	-19.87	3.79E-05	0.05	1
L92R	-0.17026	0.03809	-64.92	3.37E-07	0.05	1
L92V	9.02E-05	0.00279	0.01	0.98915	0.05	0
L92H	-0.02441	0.00676	-2.73	0.05225	0.05	0
L92W	-0.18774	0.04201	-53.83	7.13E-07	0.05	1
I94C	-0.02609	0.00593	-10.65	4.40E-04	0.05	1
I94Q	-0.11175	0.02502	-37.19	3.12E-06	0.05	1

Table S12: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 21B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
I73F	-0.04901	0.01113	-11.35	3.44E-04	0.05	1
I73L	-0.00868	0.00282	-1.89	0.13118	0.05	0
I73M	-0.01051	0.00304	-2.43	0.07195	0.05	0
I73V	-0.03932	0.009	-9.22	7.68E-04	0.05	1
G82E	-0.08138	0.0183	-18.47	5.06E-05	0.05	1
G82H	-0.0407	0.00433	-9.39	7.17E-04	0.05	1
G82I	-0.08152	0.01832	-19.65	3.96E-05	0.05	1
G82Y	-0.0479	0.01088	-11.17	3.65E-04	0.05	1
G82V	-0.08412	0.0189	-20.16	3.58E-05	0.05	1
R86F	-0.04844	0.01099	-11.60	3.16E-04	0.05	1
R86L	-0.06006	0.01356	-14.04	1.49E-04	0.05	1
W89L	-0.06746	0.01521	-15.77	9.45E-05	0.05	1
W89N	-0.07774	0.01748	-18.70	4.82E-05	0.05	1
W89T	-0.07774	0.01749	-18.11	5.47E-05	0.05	1
L92Q	-0.06767	0.01547	-9.34	7.33E-04	0.05	1
L92R	-0.07548	0.01698	-18.22	5.34E-05	0.05	1
L92V	-0.00905	0.0028	-2.10	0.10408	0.05	0
L92H	-0.02551	0.0068	-3.08	0.03692	0.05	1
L92W	-0.06335	0.0143	-14.33	1.37E-04	0.05	1
I94C	-0.01379	0.00366	-3.12	0.03532	0.05	1
I94Q	-0.04141	0.00945	-9.73	6.24E-04	0.05	1

Table S13: List of docking results from all *Cs*OAC variants including two exchanges, that are possible based on previous experiments and the linear tetraketide. The variants are sorted based on their calculated ASP fitness Score (*127*).

*Cs*OAC Variant	ASP Fitness Score	*Cs*OAC Variant	ASP Fitness Score
G82S/R86A	35.73	**G82A/L92I**	29.68
G82S/R86H	35.07	**L92F/I94V**	29.67
R86H/L92Y	32.7	**G82A/L92Y**	29.23
R86H/L92I	32.66	**L92I/I94V**	28.74
R86H/L92F	32.19	**G82A/I94V**	28.13
G82A/R86H	32.09	**G82A/I94L**	26.26
R86A/L92Y	32.00	**G82S/I94L**	x
R86H/I94V	31.90	**G82S/I94V**	x
G82A/L92F	31.70	**G82S/L92F**	x
R86A/L92F	31.64	**G82S/L92I**	x
R86H/I94L	30.82	**G82S/L92Y**	x
R86A/L92I	30.63	**L92F/I94L**	x
G82A/R86A	29.77	**L92I/I94L**	x
R86A/I94V	29.75	**R86A/I94L**	x

Table S14: List of docking results from all *Cs*OAC variants including three exchanges, that are possible based on previous experiments and the linear tetraketide. The variants are sorted based on their calculated ASP fitness Score (*127*).

*Cs*OAC Variant	ASP Fitness Score	*Cs*OAC Variant	ASP Fitness Score
G82S/R86H/L92F	36.86	**G82A/R86H/I94V**	31.03
G82S/R86H/L92Y	36.07	**R86A/L92F/I94L**	30.98
G82S/R86H/L92I	35.68	**G82A/R86H/I94L**	30.70
G82S/R86H/I94V	35.53	**R86A/L92F/I94V**	30.35
G82S/R86H/I94L	34.68	**G82A/R86A/L92F**	30.26
G82A/R86H/L92Y	33.52	**G82A/R86A/L92I**	30.24
G82S/R86A/L92F	33.32	**G82S/R86A/I94V**	30.19
G82S/L92F/I94V	33.22	**G82A/L92F/I94L**	30.08
G82A/R86H/L92F	33.08	**R86A/L92Y/I94V**	30.02
G82S/R86A/L92Y	32.90	**G82A/L92Y/I94L**	29.91
G82A/L92F/I94L	32.80	**R86A/L92I/I94V**	29.87
R86H/L92F/I94L	32.58	**G82S/R86A/I94L**	29.82
R86H/L92I/I94V	32.49	**G82A/L92I/I94V**	29.66
R86H/L92F/I94V	32.38	**G82A/L92I/I94L**	29.36
R86H/L92Y/I94L	32.28	**G82A/R86A/I94L**	29.16
R86H/L92Y/I94V	32.04	**G82A/R86A/I94V**	27.55
G82A/R86A/L92Y	31.73	**R86A/L92I/I94L**	27.16
G82S/R86A/L92I	31.67	**G82S/L92I/I94L**	x
G82A/R86H/L92I	31.53	**G82S/L92I/I94V**	x
R86H/L92I/I94L	31.52	**G82S/L92Y/I94L**	x
G82A/L92Y/I94V	31.25	**G82S/L92Y/I94V**	x
G82A/L92F/I94V	31.20	**R86A/L92Y/I94L**	x

Table S15: List of docking results from all *Cs*OAC variants including four exchanges, that are possible based on previous experiments and the linear tetraketide. The variants are sorted based on their calculated ASP fitness Score (*127*).

*Cs*OAC Variant	ASP Fitness Score	*Cs*OAC Variant	ASP Fitness Score
G82S/R86H/L92I/I94V	37.42	**G82S/R86H/L92Y/I94V**	32.93
G82S/R86A/L92F/I94L	36.27	**G82A/R86H/L92F/I94V**	32.36
G82S/R86H/L92F/I94V	36.21	**G82A/R86A/L92Y/I94V**	31.50
G82S/R86H/L92Y/I94L	36.14	**G82A/R86H/L92I/I94V**	31.33
G82S/R86A/L92F/I94V	35.87	**G82A/R86A/L92F/I94L**	31.08
G82S/R86H/L92F/I94L	35.82	**G82A/R86H/L92I/I94L**	30.96
G82S/R86H/L92I/I94L	34.67	**G82A/R86A/L92F/I94V**	30.78
G82S/R86A/L92Y/I94V	34.62	**G82S/R86A/L92I/I94V**	30.78
G82S/R86A/L92Y/I94L	34.20	**G82A/R86A/L92Y/I94L**	30.40
G82A/R86H/L92Y/I94V	34.11	**G82A/R86A/L92I/I94L**	28.99
G82A/R86H/L92Y/I94L	33.55	**G82A/R86A/L92I/I94V**	28.11
G82S/R86A/L92I/I94L	33.03	**G82A/R86H/L92F/I94L**	x

Table S16: Post hoc analysis data for OA concentration in Figure 26A: The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H	-0.02561	0.00606	-5.81	0.00437	0.05	1
G82A/L92F	-0.00319	0.00231	-0.65	0.55242	0.05	0
G82A/L92Y	0.012342	0.00596	3.67	0.02136	0.05	1
R86A/L92F	-0.00842	0.00288	1.73	0.15829	0.05	0
R86A/L92Y	0.01787	0.00478	3.05	0.03819	0.05	1
R86H/L92Y	-0.00981	0.00374	-1.45	0.2215	0.05	0
G82A/R86A/L92Y	-0.04846	0.011	-11.56	3,20E-04	0.05	1
G82A/R86H/L92Y	0.02357	0.00563	5.35	0.00587	0.05	1
G82S/R86H/L92Y	-0.01741	0.00435	-4,01	0.01601	0.05	1
R86A/L92Y/I94V	0.00219	0.00192	0.53	0.62615	0.05	0

Table S17: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 26B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H	0.01306	0.004	2.14	0.09943	0.05	0
G82A/L92F	0.01254	0.00843	2.25	0.08812	0.05	0
G82A/L92Y	0.06103	0.0167	2.83	0.04712	0.05	1
R86A/L92F	0.0025	0.0024	0.48	0.65711	0.05	0
R86A/L92Y	0.06091	0.01419	6.82	0.00242	0.05	1
R86H/L92Y	-2.13E-04	0.00821	-0.01	0.99129	0.05	0
G82A/R86A/L92Y	-0.00846	0.00218	-3.51	0.02462	0.05	1
G82A/R86H/L92Y	0.11842	0.0265	46.17	1.32E-06	0.05	1
G82S/R86H/L92Y	0.11395	0.02551	41.86	1.95E-06	0.05	1
R86A/L92Y/I94V	0.07855	0.01774	14.31	1.39E-04	0.05	1

Table S18: Post hoc analysis data for OA concentration in Figure 23A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H	-0.02561	0.00441	-5.81	0.00437	0.05	1
G82A/L92F	-0.00319	0.00492	-0.65	0.55242	0.05	0
G82S/R86A	-0.05469	0.00424	-12.90	2.08E-04	0.05	1
G82S/R86H	-0.04638	0.00431	-10.76	4.24E-04	0.05	1
R86A/L92F	-0.00842	0.00486	-1.73	0.15829	0.05	0
R86A/L92Y	0.01787	0.00587	3.05	0.03819	0.05	1
R86H/L92F	-0.02072	0.00418	-4.96	0.00772	0.05	1
R86H/L92I	-0.00545	0.00426	-1.28	0.26946	0.05	0
R86H/L92Y	-0.00981	0.00678	-1.45	0.2215	0.05	0
R86H/I94V	-0.01699	0.0042	-4.05	0.01551	0.05	1
G82A/R86A	-0.04542	0.00417	-10.90	4.02E-04	0.05	1
G82A/L92Y	0.012342	0.00596	3.67	0.02136	0.05	1
G82A/I94V	-0.05887	0.00663	-8.87	8.91E-04	0.05	1
R86A/I94V	-0.0214	0.005	-4.28	0.01284	0.05	1
L92F/I94V	-0.05531	0.00825	-6.70	0.00257	0.05	1
L92Y/I94V	-0.01843	0.00438	-4.21	0.01358	0.05	1

Table S19: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 23B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H	0.01306	0.00611	2.14	0.09943	0.05	0
G82A/L92F	0.02815	0.01254	2.25	0.08812	0.05	0
G82S/R86A	-0.06462	0.00463	-13.95	1.53E-04	0.05	1
G82S/R86H	0.0237	0.00272	8.72	9.54E-04	0.05	1
R86A/L92F	0.0025	0.00521	0.48	0.65711	0.05	0
R86A/L92Y	0.06091	0.00893	6,82	0.00242	0.05	1
R86H/L92F	-0.02744	0.00318	-8.64	9.86E-04	0.05	1
R86H/L92I	0.09563	0.00444	21.55	2.74E-05	0.05	1
R86H/L92Y	-2.13E-04	0.01835	-0.01	0.99129	0.05	0
R86H/I94V	0.06696	0.00361	18.56	4.96E-05	0.05	1
G82A/R86A	-0.07661	0.00329	-23.26	2.03E-05	0.05	1
G82A/L92Y	0.06103	0.02153	2.83	0.04712	0.05	1
G82A/I94V	-0.14166	0.00973	-14.55	1.30E-04	0.05	1
R86A/I94V	0.0345	0.00655	5.27	0.00623	0.05	1
L92F/I94V	-0.02558	0.00967	-2.65	0.05718	0.05	0
L92Y/I94V	0.03352	0.00324	10.35	4.92E-04	0.05	1

Table S20: Post hoc analysis data for OA concentration in Figure 24A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H/L92F	-0.04628	0.00475	-9.74	6.22E-04	0.05	1
G82A/R86H/L92Y	0.02357	0.0044	5.35	0.00587	0.05	1
G82S/R86A/L92F	-0.05222	0.00483	-10.81	4.15E-04	0.05	1
G82S/R86A/L92Y	-0.04888	0.00418	-11.69	3.06E-04	0.05	1
G82S/R86H/L92F	-0.03001	0.0061	-4.92	0.00793	0.05	1
G82S/R86H/L92I	-0.03223	0.00579	-5.57	0.0051	0.05	1
G82S/R86H/L92Y	0.01741	0.00434	4.01	0.01601	0.05	1
G82S/R86H/I94L	-0.05463	0.00416	-13.12	1.95E-04	0.05	1
G82S/R86H/I94V	-0.04911	0.00486	-10.11	5.39E-04	0.05	1
G82S/L92F/I94V	-0.09798	0.00415	-23.63	1.90E-05	0.05	1
G82A/R86A/L92F	-0.05073	0.00567	-8.94	8.64E-04	0.05	1
G82A/R86A/L92Y	0.00232	0.00778	0.30	0.78079	0.05	0
G82A/L92F/I94V	-0.07481	0.00449	-16.66	7.60E-05	0.05	1
G82A/L92Y/I94V	-0.04846	0.00419	-11.56	3.20E-04	0.05	1
R86A/L92Y/I94V	-0.00219	0.00415	-0.53	0.62615	0.05	0

Table S21: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 24B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H/L92F	-0.12077	0.00419	-28.81	8.65E-06	0.05	1
G82A/R86H/L92Y	0.11842	0.00256	46.17	1.32E-06	0.05	1
G82S/R86A/L92F	-0.03989	0.00293	-13.62	1.68E-04	0.05	1
G82S/R86A/L92Y	-0.00903	0.00236	-3.83	0.01866	0.05	1
G82S/R86H/L92F	-0.02195	0.00659	-3.33	0.02915	0.05	1
G82S/R86H/L92I	0.01016	0.00411	2.47	0.06905	0.05	0
G82S/R86H/L92Y	0.11395	0.00272	41.86	1.95E-06	0.05	1
G82S/R86H/I94L	0.00281	0.00364	0.77	0.48217	0.05	0
G82S/R86H/I94V	-0.13304	0.00853	-15.60	9.85E-05	0.05	1
G82S/L92F/I94V	-0.28807	0.00269	-107.17	4.55E-08	0.05	1
G82A/R86A/L92F	0.05674	0.03044	1.86	0.13578	0.05	0
G82A/R86A/L92Y	0.12375	0.04031	3.07	0.0373	0.05	1
G82A/L92F/I94V	-0.11122	0.00395	-28.14	9.49E-06	0.05	1
G82A/L92Y/I94V	-0.00846	0.00241	-3.51	0.02462	0.05	1
R86A/L92Y/I94V	0.07855	0.00549	14.31	1.49E-04	0.05	1

Table S22: Post hoc analysis data for OA concentration in Figure 25A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H/L92Y/I94V	-0.03709	0.00418	-8.87	8.94E-04	0.05	1
G82S/R86A/L92F/I94L	-0.05236	0.00418	-12.54	2.33E-04	0.05	1
G82S/R86A/L92F/I94V	-0.05987	0.00456	-13.12	1.95E-04	0.05	1
G82S/R86A/L92Y/I94L	-0.04916	0.00435	-11.29	3.50E-04	0.05	1
G82S/R86A/L92Y/I94V	-0.00393	0.00539	-0.73	0.5058	0.05	0
G82S/R86H/L92F/I94L	-0.0476	0.00448	-10.61	4.46E-04	0.05	1
G82S/R86H/L92F/I94V	-0.03683	0.00416	-8.86	8.97E-04	0.05	1
G82S/R86H/L92I/I94L	-0.06767	0.0042	-16.10	8.71E-05	0.05	1
G82S/R86H/L92I/I94V	-0.04866	0.00608	-8.01	0.00132	0.05	1
G82S/R86H/L92Y/I94L	-0.03236	0.00496	-6.53	0.00284	0.05	1
G82A/R86A/L92F/I94V	-0.0606	0.00436	-13.90	1.56E-04	0.05	1
G82A/R86A/L92Y/I94V	-0.04174	0.00467	-8.95	8.64E-04	0.05	1

Table S23: Post hoc analysis data for Ratio OA/(OA/OL) in Figure 25B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A/R86H/L92Y/I94V	-0.01336	0.00345	-3.88	0.0179	0.05	1
G82S/R86A/L92F/I94L	-0.07938	0.00361	-22.02	2.52E-05	0.05	1
G82S/R86A/L92F/I94V	-0.10223	0.00593	-17.23	6.56E-05	0.05	1
G82S/R86A/L92Y/I94L	-0.05107	0.00378	-13.51	1.74E-04	0.05	1
G82S/R86A/L92Y/I94V	0.11588	0.00625	18.54	4.98E-05	0.05	1
G82S/R86H/L92F/I94L	-0.02715	0.00552	-4.92	0.00795	0.05	1
G82S/R86H/L92F/I94V	-0.10873	0.00241	-45.14	1.44E-06	0.05	1
G82S/R86H/L92I/I94L	-0.13874	0.00234	-59.32	4.83E-07	0.05	1
G82S/R86H/L92I/I94V	-0.04443	0.00861	-5.16	0.0067	0.05	1
G82S/R86H/L92Y/I94L	-0.03909	0.00618	-6.32	0.0032	0.05	1
G82A/R86A/L92F/I94V	0.02562	0.00305	8.39	0.0011	0.05	1
G82A/R86A/L92Y/I94V	-0.08569	0.00405	-21.13	2,96E-05	0.05	1

Table S24: Post hoc analysis data for OA concentration in Figure 28A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A	9.23E-04	2.50E-04	2.94	0.04237	0.05	1
G82S	4.92E-04	1.48E-04	2.21	0.09116	0.05	0
L92Y	0.00108	2.61E-04	4.72	0.00917	0.05	1
I94V	6.53E-04	2.28E-04	2.87	0.04569	0.05	1
G82A/L92Y	0.00206	4.75E-04	7.98	0.00134	0.05	1
G82A/R86A/L92Y	6.33E-04	1.78E-04	2.62	0.05878	0.05	0

Table S25: Post hoc analysis data for Ratio OA/(OA/OL)in Figure 28B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
G82A	0.05239	0.01187	12.34	2.48E-04	0.05	1
G82S	0.05132	0.01163	12.24	2.56E-04	0.05	1
L92Y	0.07209	0.01631	13.10	1.96E-04	0.05	1
I94V	0.0237	0.0057	5.06	0.0072	0.05	1
G82A/L92Y	0.13534	0.03033	29.66	7.69E-06	0.05	1
G82A/R86A/L92Y	0.00756	0.00417	0.89	0.42573	0.05	0

Table S26: Post hoc analysis data for OA concentration in Figure 27A. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
Y27F	-0.00215	2.25E-04	-9.57	6.66E-04	0.05	1
R86A	-0.00317	6.03E-04	-5.26	0.00623	0.05	1
R86H	-0.003	7.47E-04	-4.02	0.01587	0.05	1
W89A	-0.00427	2.21E-04	-19.33	4.22E-05	0.05	1
L92F	-0.00224	2.17E-04	-10.33	4.95E-04	0.05	1
L92I	-0.00208	6.48E-04	-3.21	0.03262	0.05	1
I94L	-5.13E-04	2.21E-04	-2.33	0.08068	0.05	0
G82A/R86H	-8.26E-04	2.19E-04	-3.78	0.01946	0.05	1
G82A/L92F	4.81E-04	2.21E-04	2.17	0.09586	0.05	0
R86A/L92F	-0.00238	6.21E-04	-3.84	0.01847	0.05	1
R86A/L92Y	-0.00209	0.00116	-1.80	0.14688	0.05	0
G82A/R86H/L92Y	-0.00267	2.16E-04	-12.38	2.45E-04	0.05	1
G82A/L92Y/I94V	-0.00269	2.25E-04	-11.95	2.81E-04	0.05	1
G82S/R86H/L92Y	-0.00234	5.78E-04	-11.59	2.21E-04	0.05	1

Table S27 Post hoc analysis data for Ratio OA/(OA/OL) in Figure 27B. The F-test of the one-way ANOVA resulted in a significant difference in the group means at the level of 0.05. Bonferroni post hoc analysis was performed to explore significant difference between means compared to the *Cs*OAC wt (*127*).

*Cs*OAC variant	Mean Difference	SEM	t-value	p-value	Alpha	Sig
Y27F	-0.09979	0.02292	-8.55	0.00103	0.05	1
R86A	-0.15615	0.03547	-11.14	3.69E-04	0.05	1
R86H	-0.10953	0.00645	-16.97	7.07E-05	0.05	1
W89A	-0.21836	0.04891	-33.87	4.53E-06	0.05	1
L92F	-0.09197	0.02065	-21.66	2.69E-05	0.05	1
L92I	-0.06961	0.01653	-5.61	0.00497	0.05	1
I94L	-0.11986	0.04859	-1.32	0.2565	0.05	0
G82A/R86H	-0.03056	0.00701	-8.77	9.31E-04	0.05	1
G82A/L92F	0.01929	0.0051	3.18	0.03347	0.05	1
R86A/L92F	-0.13534	0.03033	-29.66	7.69E-06	0.05	1
R86A/L92Y	-0.10945	0.02814	-3.52	0.02439	0.05	1
G82A/R86H/L92Y	-0.139	0.00376	36.96	3.20E-06	0.05	1
G82A/L92Y/I94V	-0.14918	0.0044	33.94	4.50E-06	0.05	1
G82S/R86H/L92Y	-0.12832	0.00536	23.92	1.81E-05	0.05	1

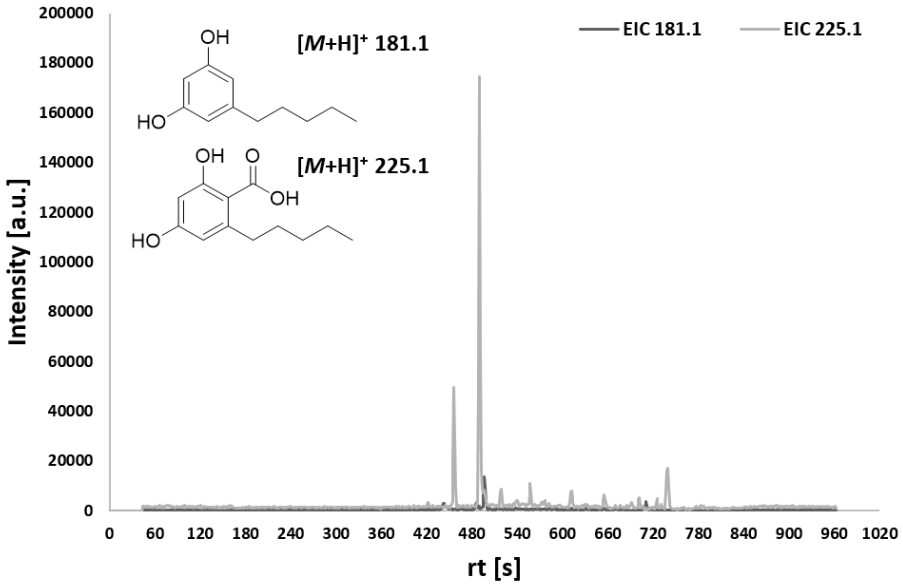


Figure S4: Extracted ion chromatograms for m/z 181.1 and 225.1 of cells from a cultivation using yAP01 and 0.5 mM hexanoic acid.

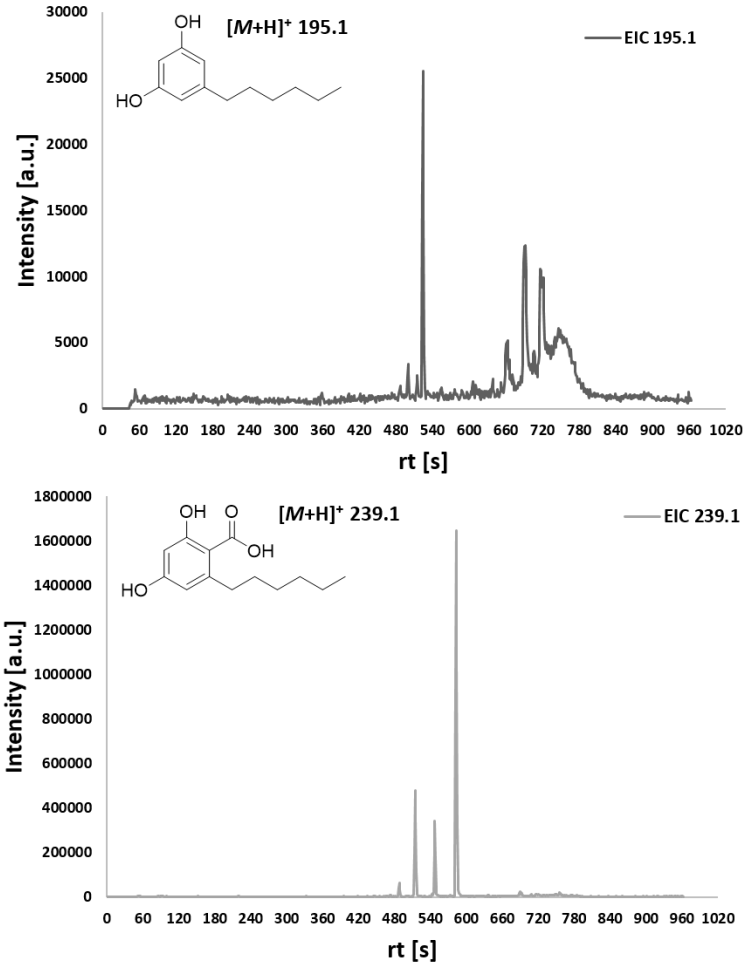


Figure S5: Extracted ion chromatograms for m/z 195.1 and 239.1 of cells from a cultivation using yAP01 and 0.5 mM heptanoic acid.

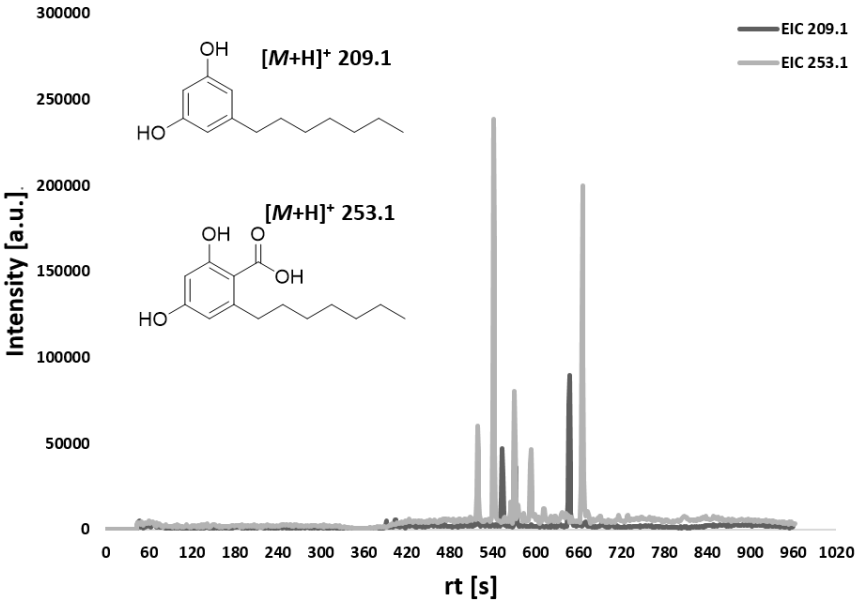


Figure S6: Extracted ion chromatograms for m/z 209.1 and 253.1 of cells from a cultivation using yAP01 and 0.5 mM octanoic acid.

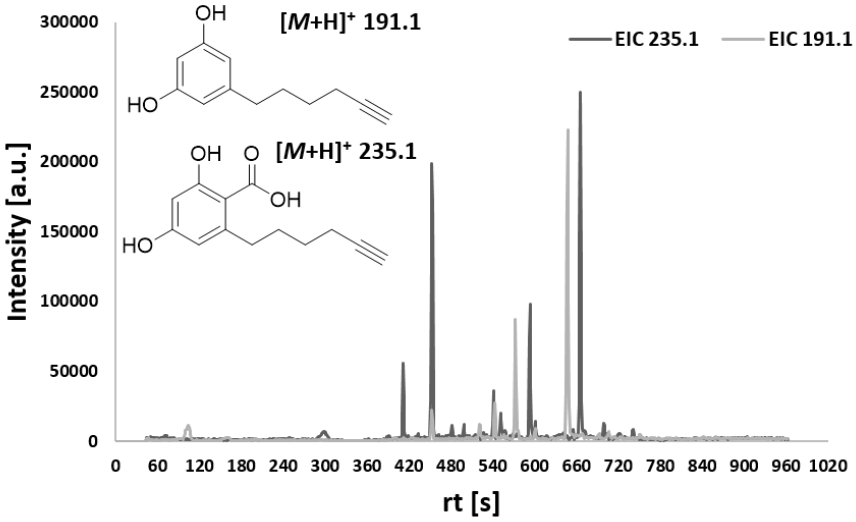


Figure S7: Extracted ion chromatograms for m/z 191.1 and 235.1 of cells from a cultivation using yAP01 and 0.5 mM 6-Heptynoic acid.

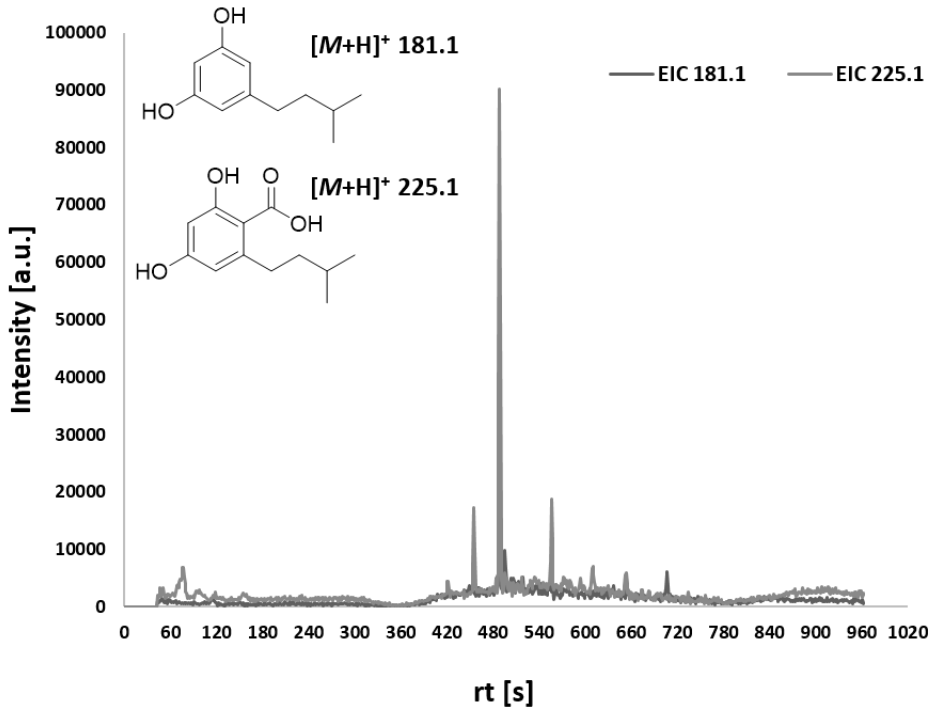


Figure S8: Extracted ion chromatograms for m/z 181.1 and 225.1 of cells from a cultivation using yAP01 and 0.5 mM 4-methyl-pentanoic acid

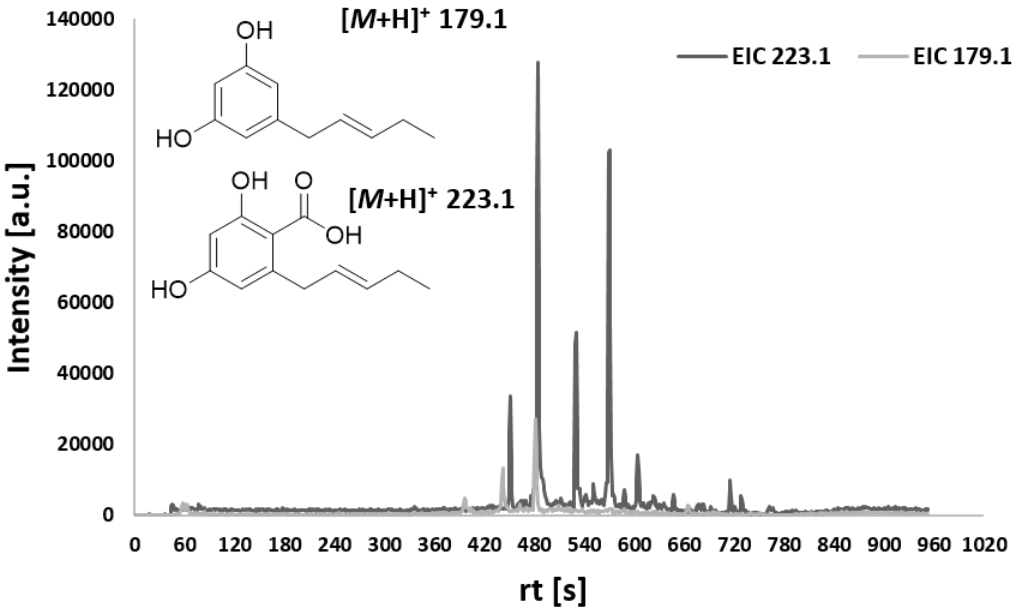


Figure S9: Extracted ion chromatograms for m/z 179.1 and 223.1 of cells from a cultivation using yAP01 and 0.5 mM trans-2-hexanoic acid

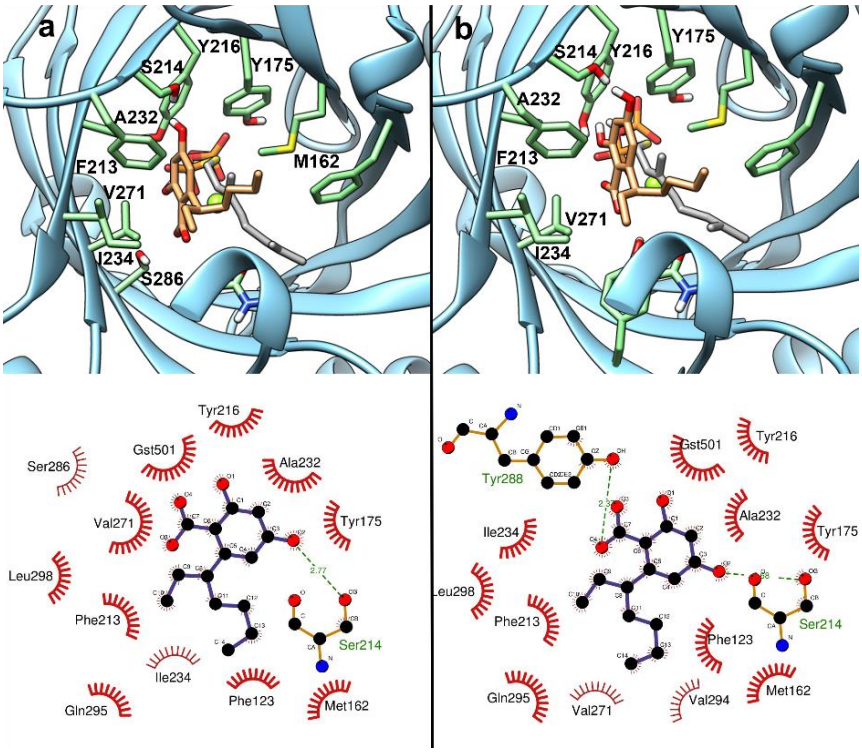


Figure S10: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1e** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

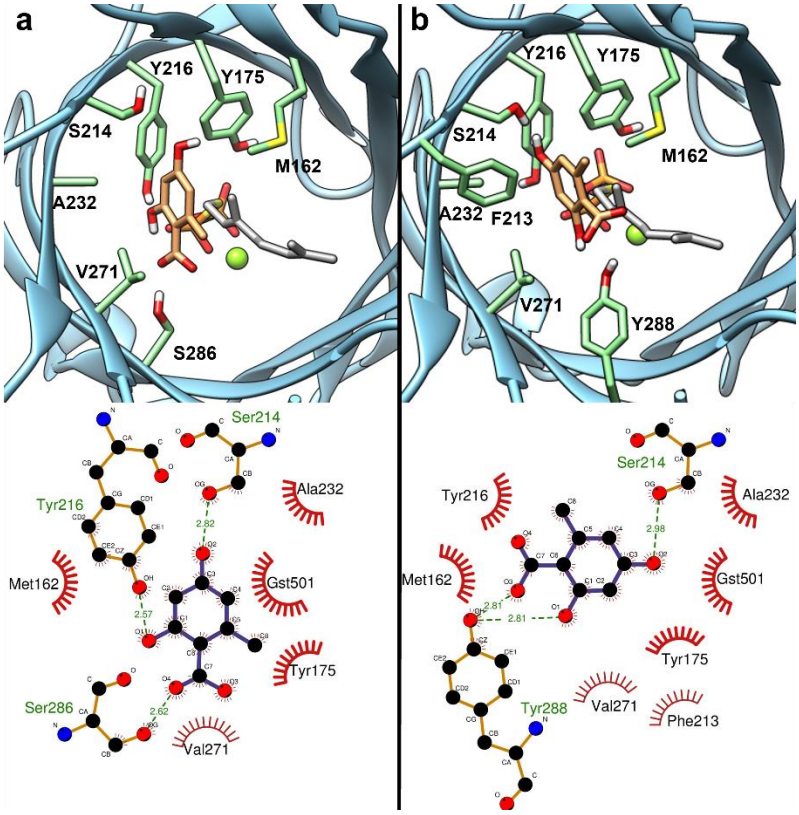


Figure S11: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1b** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

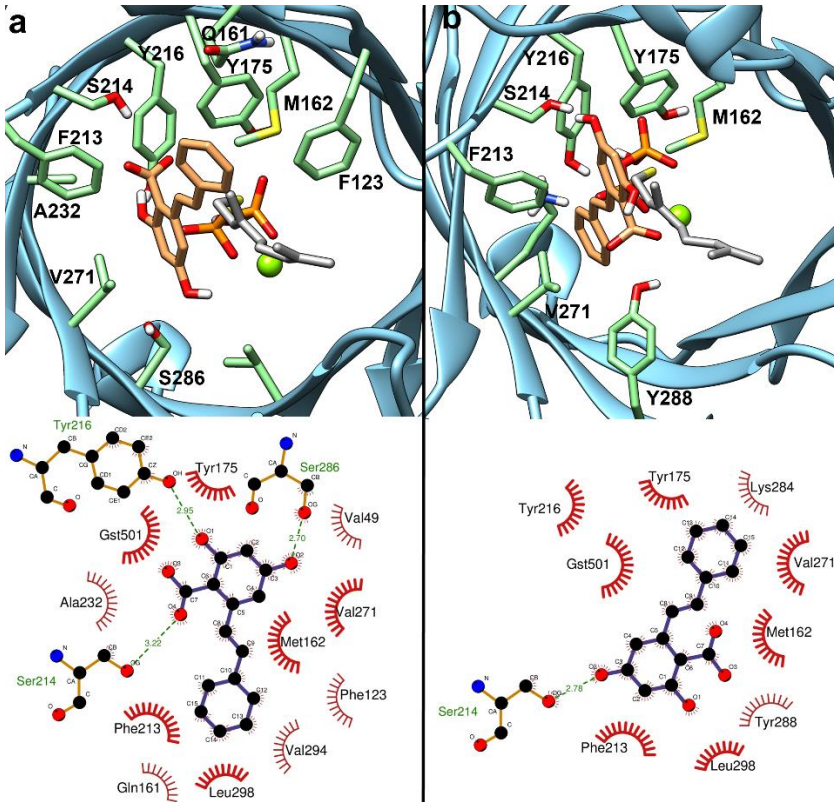


Figure S12: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1g** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

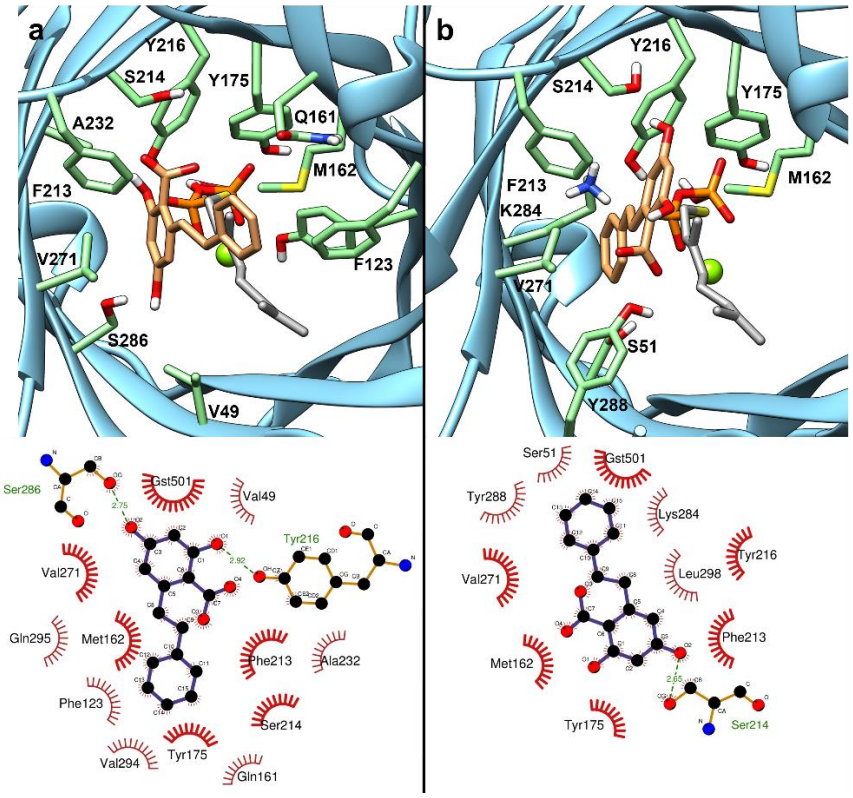


Figure S13: 3D depiction created using UCSF chimera and 2D depiction created using LigPlot of the docked pose of **1h** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green). (*142, 141, 108*)

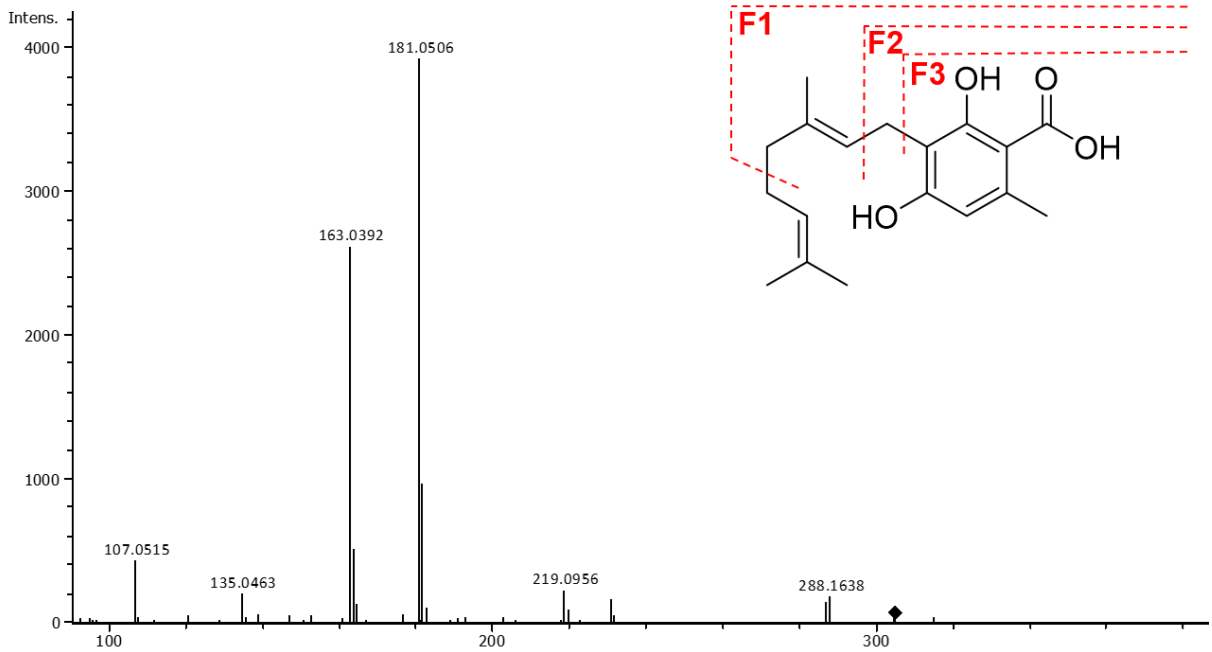


Figure S14: HR-MS-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-methylbenzoic acid **4b**. *m/z* calcd. for $C_{18}H_{24}O_4$+H$^+$: 305.1680 [*M*+H]$^+$; found: 305.1698.

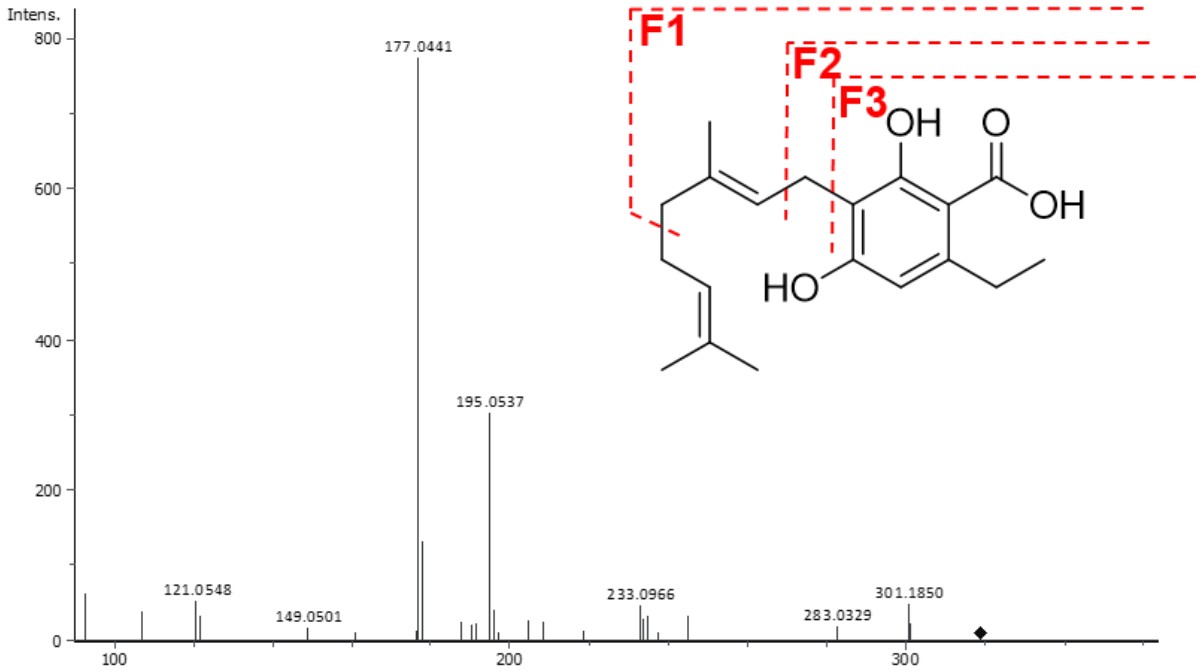


Figure S15: HR-MS-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-6-ethyl-2,4-dihydroxybenzoic acid 4**c**. *m/z* calcd. for $C_{19}H_{26}O_4$-H$^+$: 317.1837 [*M*-H]$^-$; found: 317.1879.

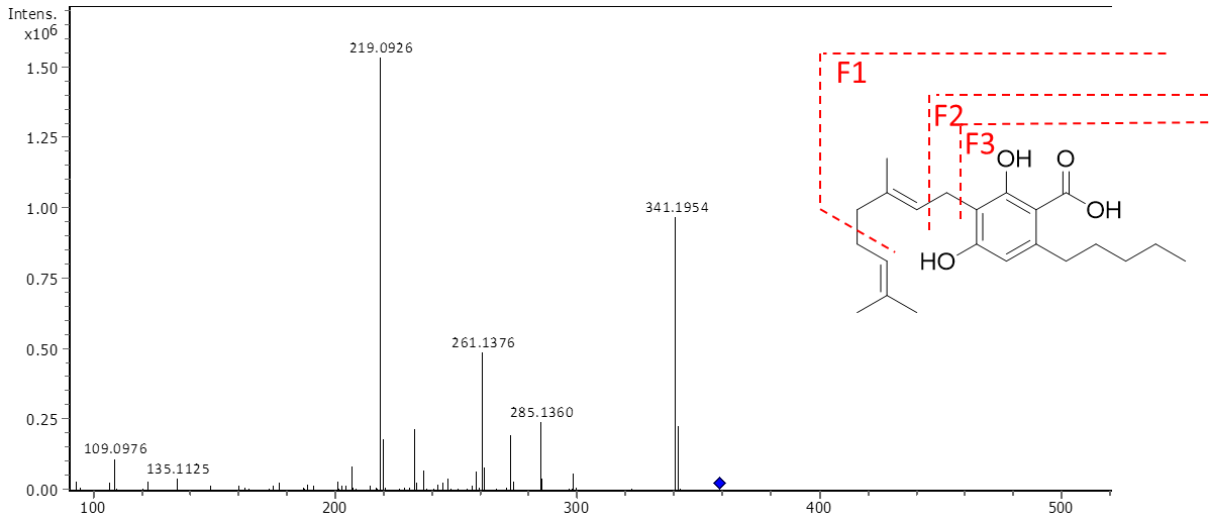


Figure S16: HR-MS-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-pentylbenzoic acid **4a**. *m/z* calcd. for $C_{22}H_{32}O_4$+H^+: 360.2306 [*M*+H]$^+$; found: 317.1879.

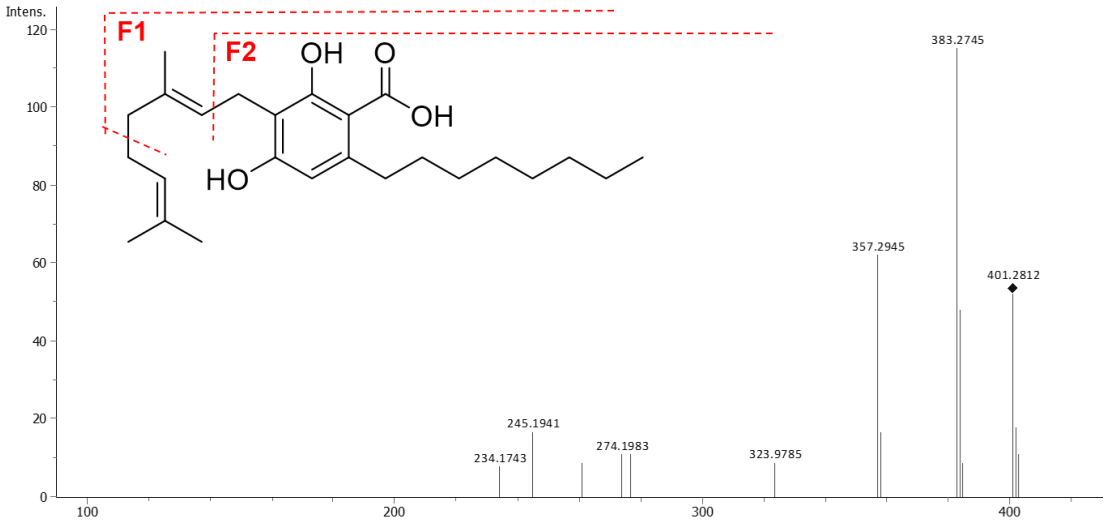


Figure S17: HR-MS-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-octylbenzoic acid **4d**. *m/z* calcd. for $C_{25}H_{38}O_4$-H^+: 401.2776 [*M*-H]$^-$; found: 401.2812.

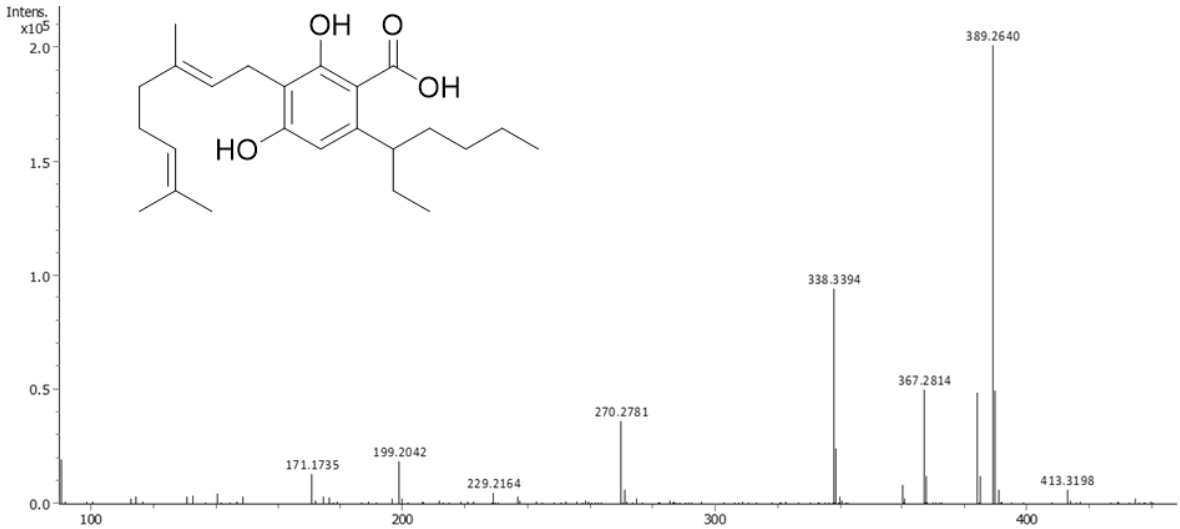


Figure S18: HR-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-6-(heptan-3-yl)-2,4-dihydroxybenzoic acid **4e**. *m/z* calcd. for $C_{24}H_{36}O_4$+H^+: 389.2619 [*M*+H]$^+$; found: 389.2640.

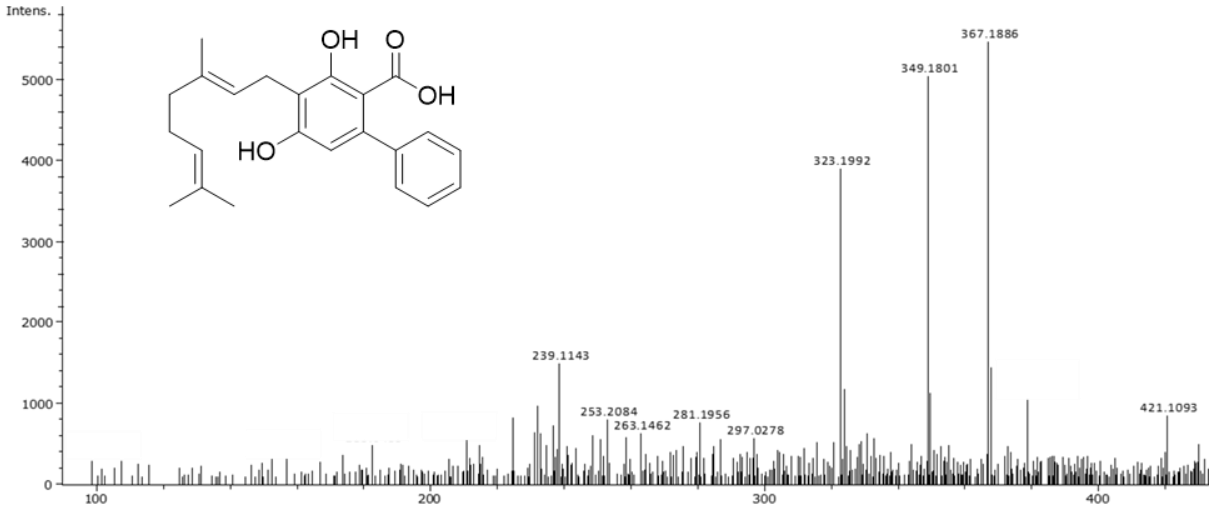


Figure S19: HR-MS spectrum of (E)-4-(3,7-dimethylocta-2,6-dien-1-yl)-3,5-dihydroxy-[1,1'-biphenyl]-2-carboxylic acid **4f**. *m/z* calcd. for $C_{23}H_{26}O_4+H^+$: 367.1837 $[M+H]^+$; found: 367.1886.

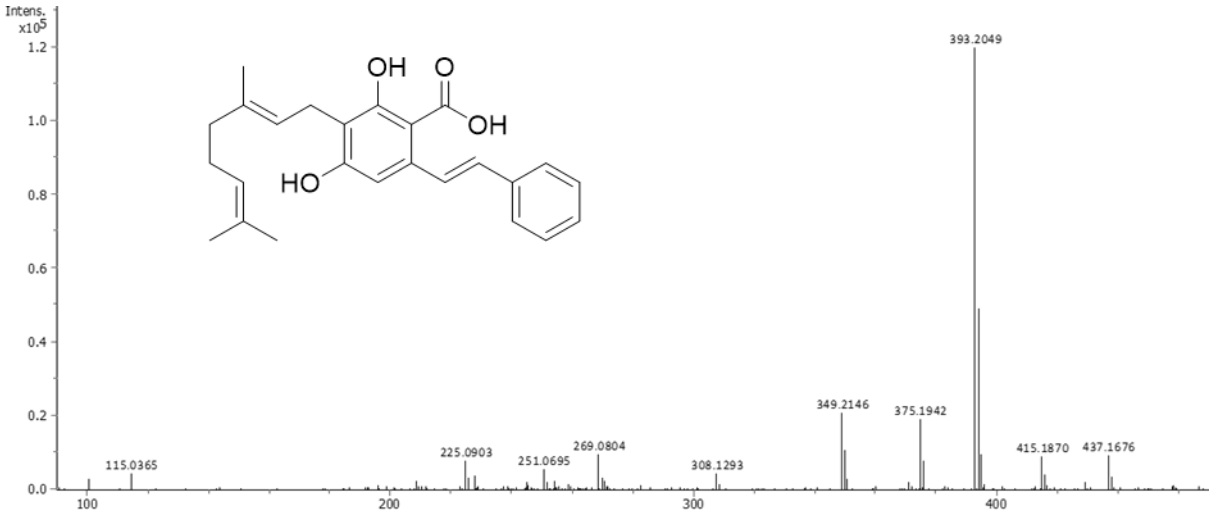


Figure S20: HR-MS spectrum of 3-((E)-3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-((E)-styryl)benzoic acid **4g**. *m/z* calcd for $C_{25}H_{28}O_4+H^+$: 393.2071 $[M+H]^+$; found: 393.2049.

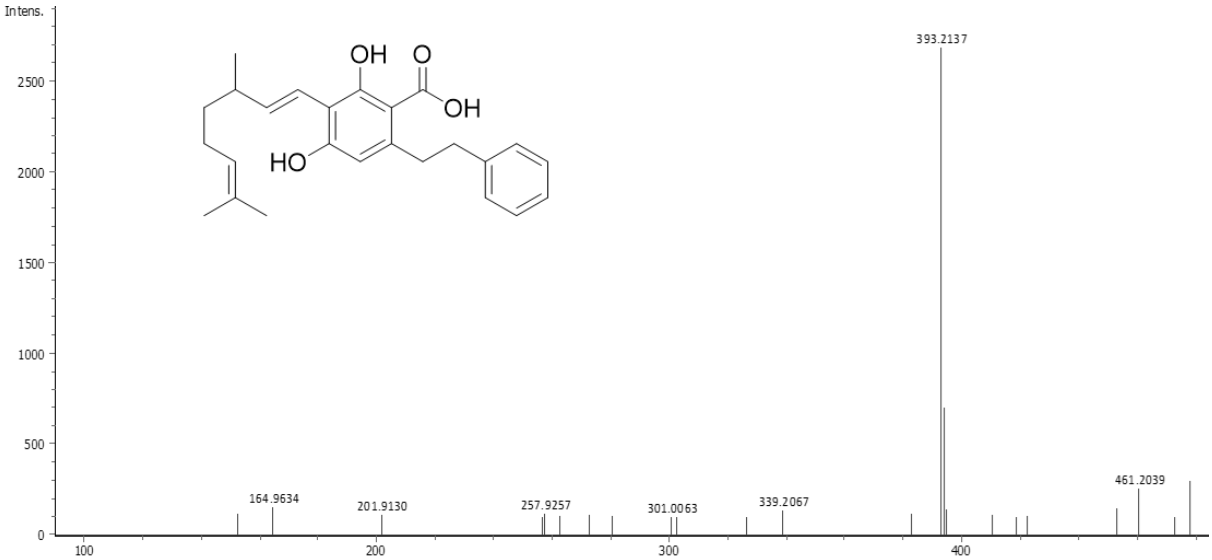


Figure S21: HR-MS spectrum of (E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-phenethylbenzoic acid **4h**. *m/z* calcd for $C_{25}H_{30}O_4+H^+$: 393.2150 $[M-H]^-$; found: 393.2137.

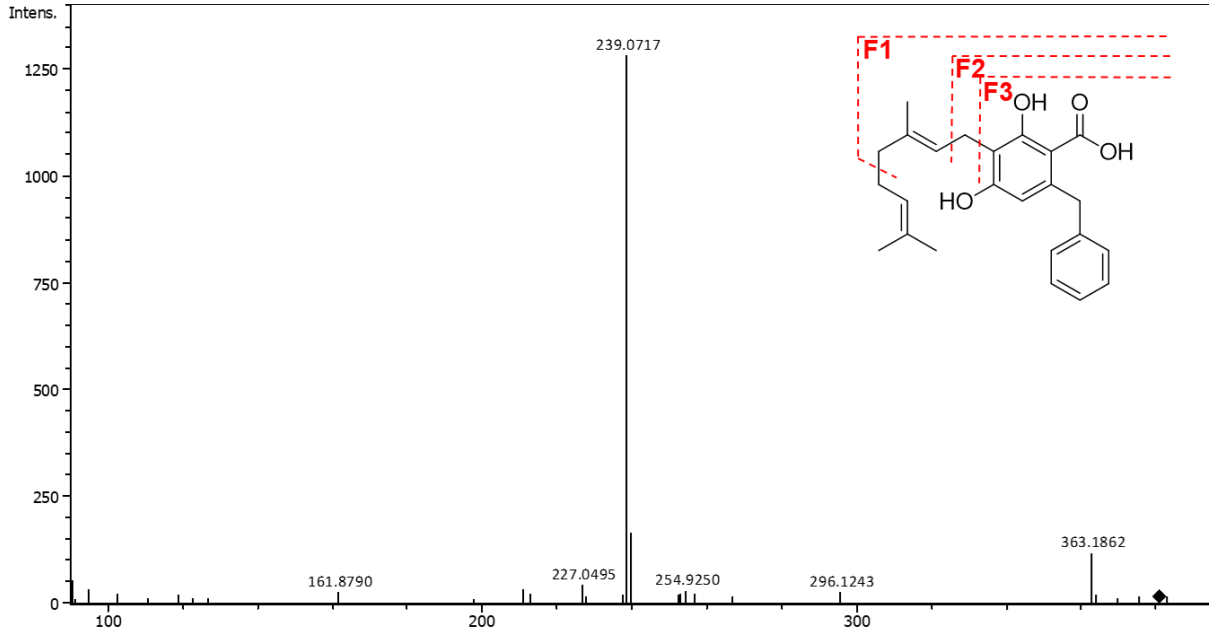


Figure S22: HR-MS spectrum of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i**. *m/z* calcd. for $C_{24}H_{28}O_4$+H$^+$: 380.1993 [*M*+H]$^+$; found: 381.2029.

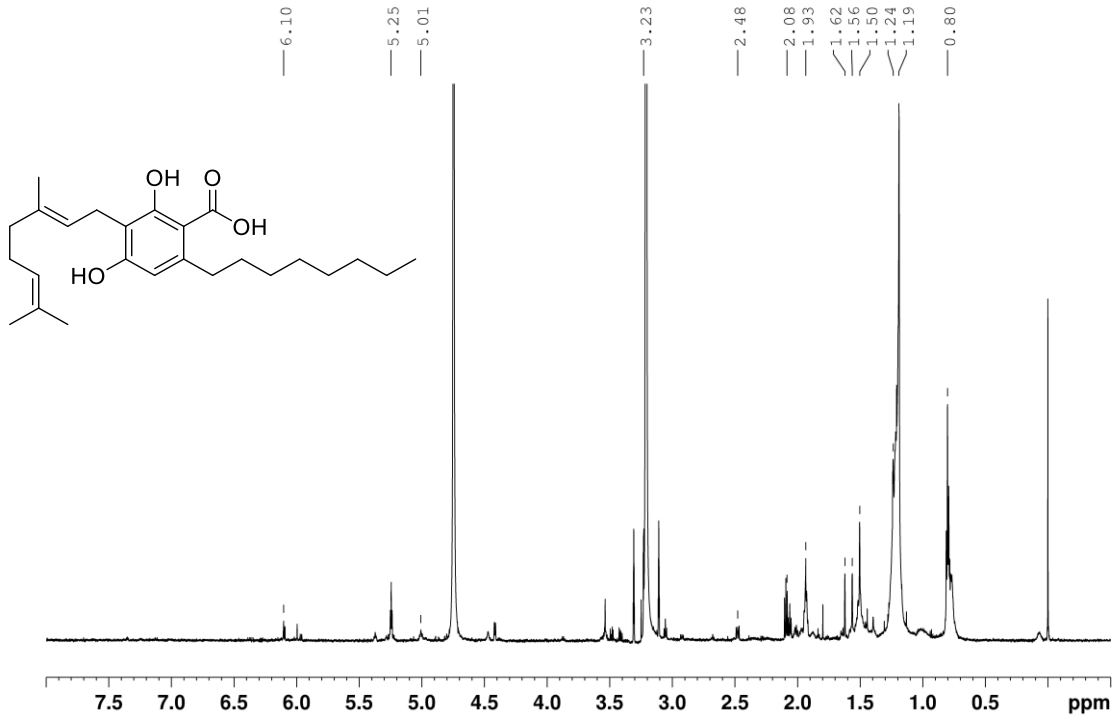


Figure S23: ^{1}H-NMR of **(**E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-octylbenzoic acid **4d** in MeOD-d_4 (400 MHz). **^{1}H NMR** (CD_3OD, 400 MHz) δ 0.80 (3H, t, *J* = 6.9 Hz), 1.19(6H, s), 1.20 – 1.30 (6H, m), 1.50 (2H, q, *J* = 7.2 Hz), 1.56 (3H, s), 1.62 (3H, s), 1.93 (3H, s), 2.01 – 2.11 (4H, m), 2.48 (2H, m, *J* = 7.8 Hz), 3.19 – 3.26 (1H, d, *J* = 7.1 Hz), 5.01 (1H, dt, *J* = 7.3), 5.25 (1H, dt, *J* = 7.1 Hz), 6.10 (1H, s).

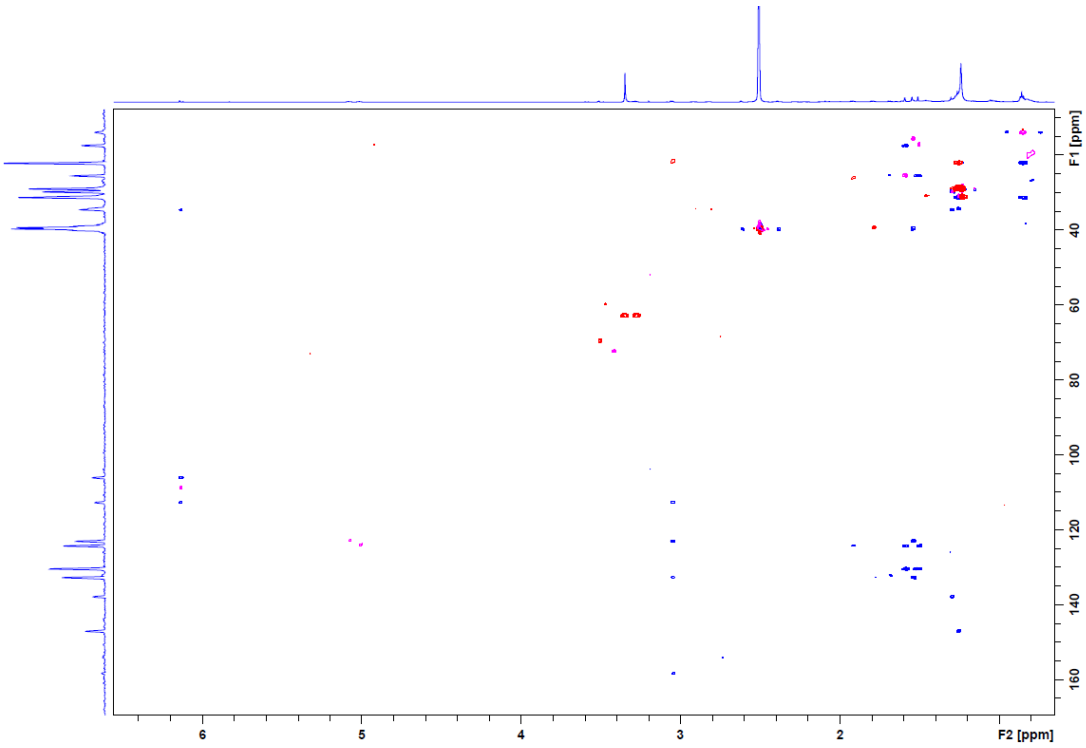


Figure S24: HMBC (blue) and HSQC (red) of **(**E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-octylbenzoic acid **4d** DMSO-d_6 (600 MHz).

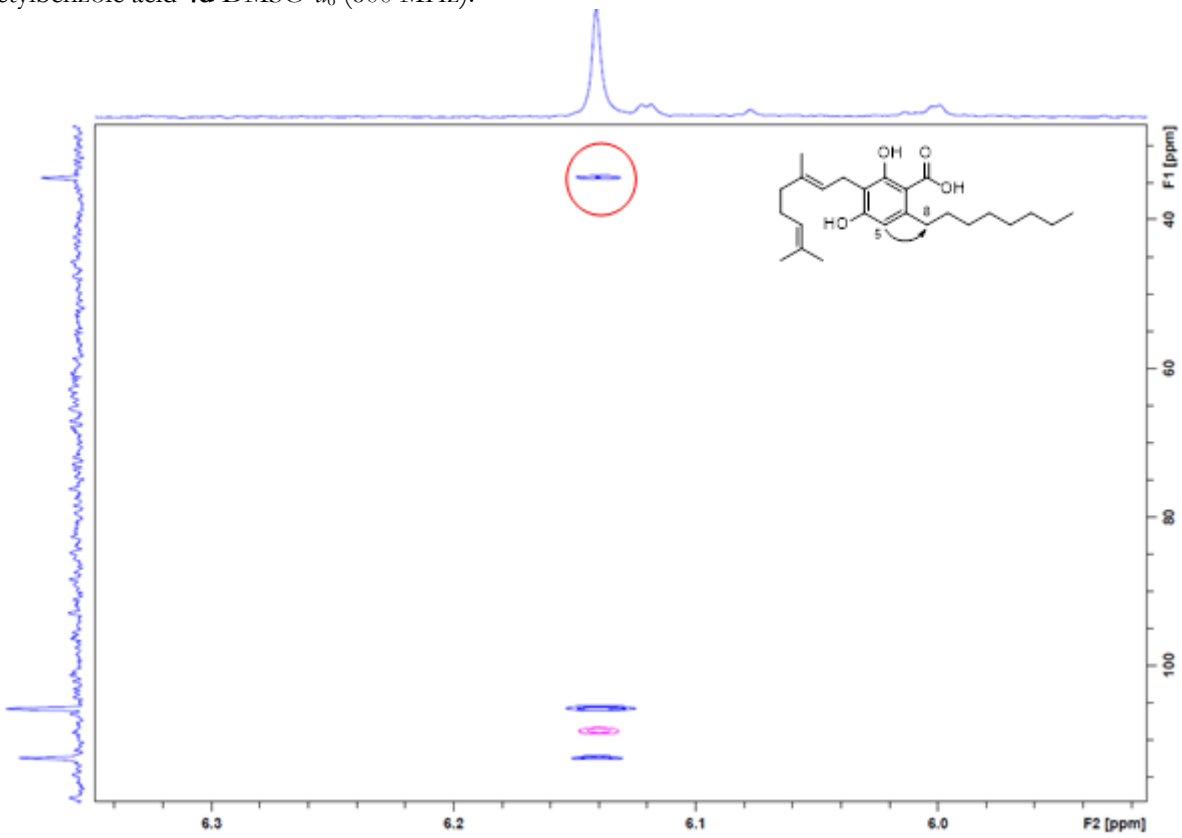


Figure S25: Focus on the signal of the HMBC (blue) and HSQC (red) of **(**E)-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxy-6-octylbenzoic acid **4d** DMSO-d_6 (600 MHz).

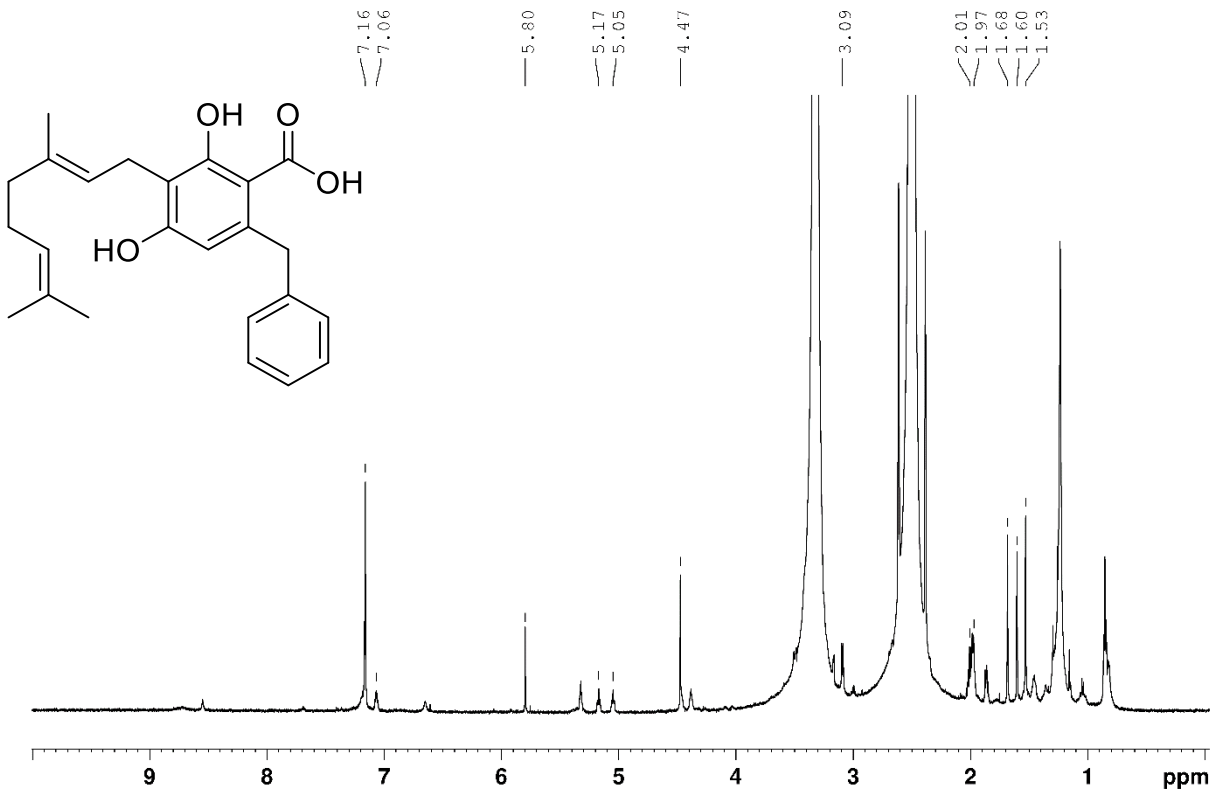


Figure S26. ^{1}H-NMR of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i** (600 MHz, DMSO-d_6, 25 °C).

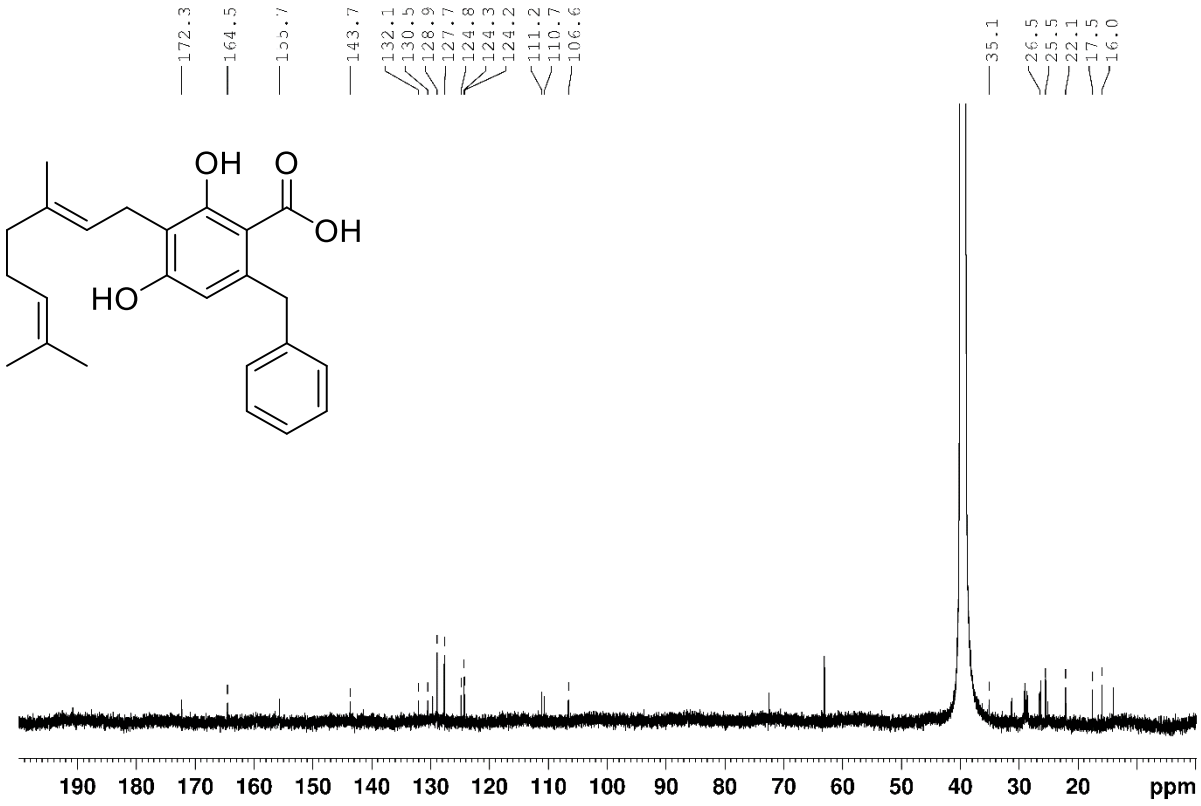


Figure S27. ^{13}C-NMR of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i** (150 MHz, DMSO-d_6, 25 °C).

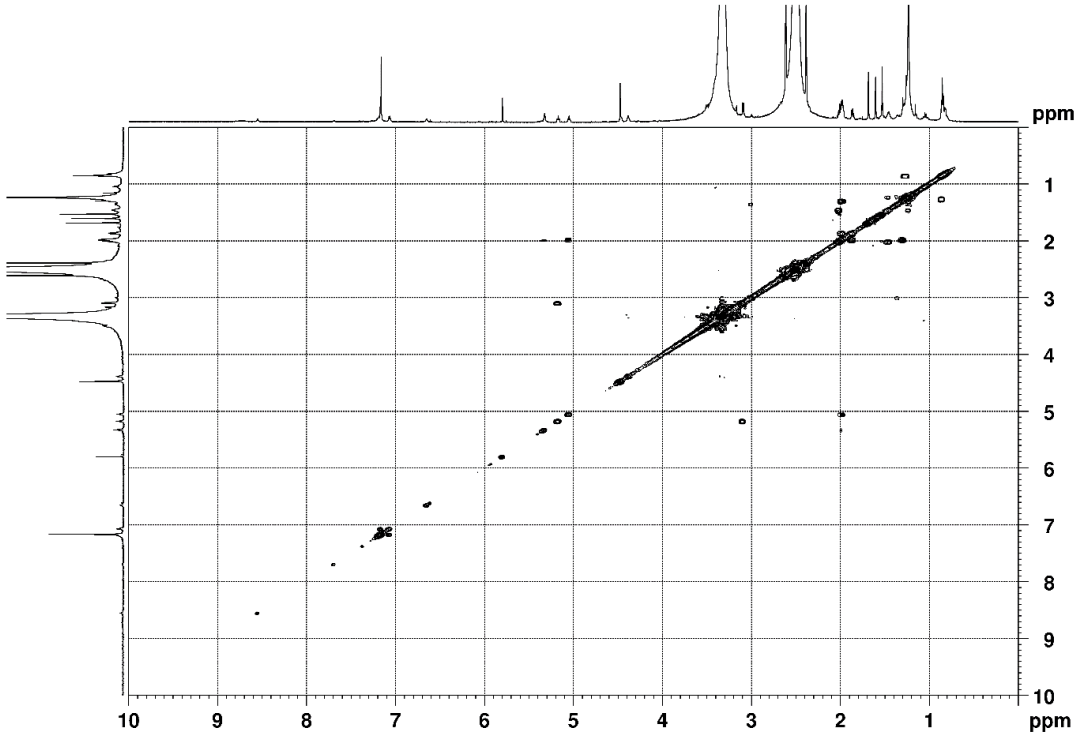


Figure S28: COSY spectrum of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i** (600 MHz, DMSO-d6, 25 °C).

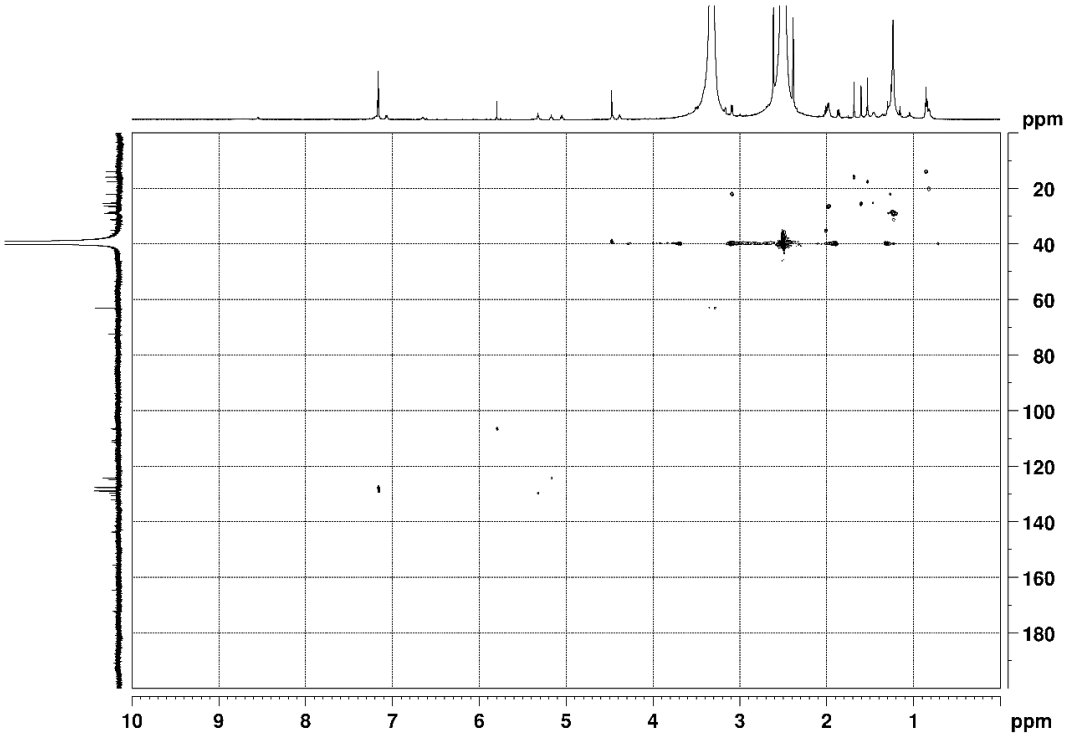


Figure S29: HSQC spectrum of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i** (600 MHz, DMSO-d6, 25 °C).

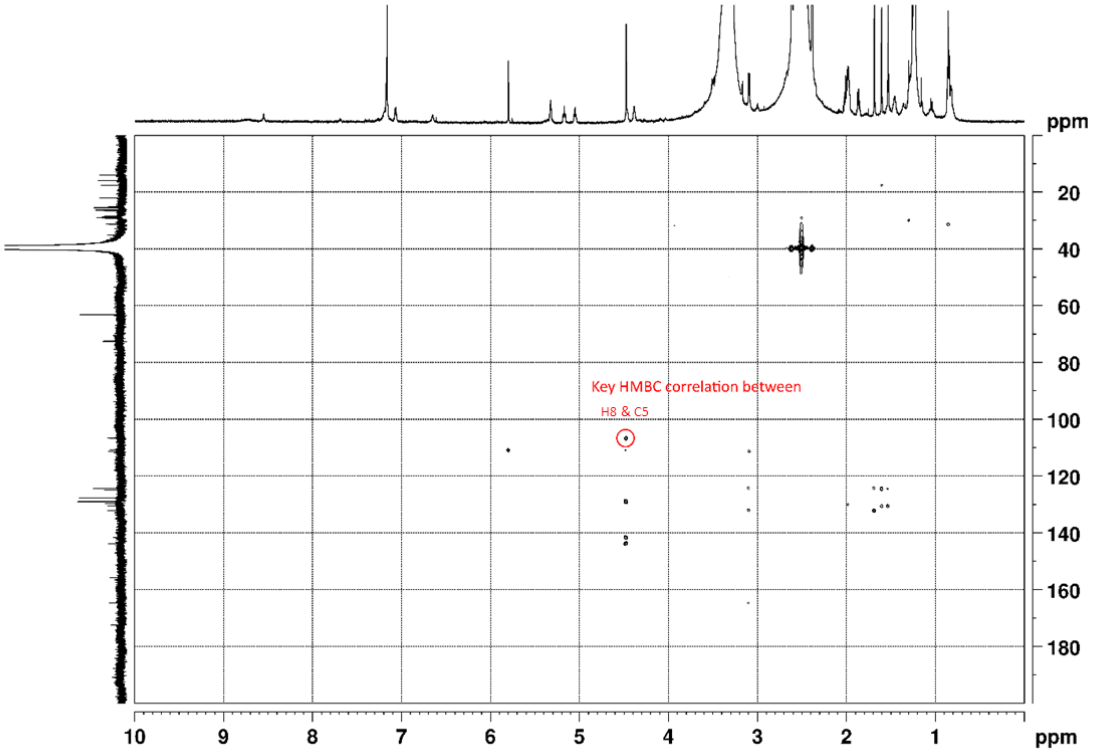


Figure S30: HMBC spectrum of (E)-6-benzyl-3-(3,7-dimethylocta-2,6-dien-1-yl)-2,4-dihydroxybenzoic acid **4i** (600 MHz, DMSO-d6, 25 °C).

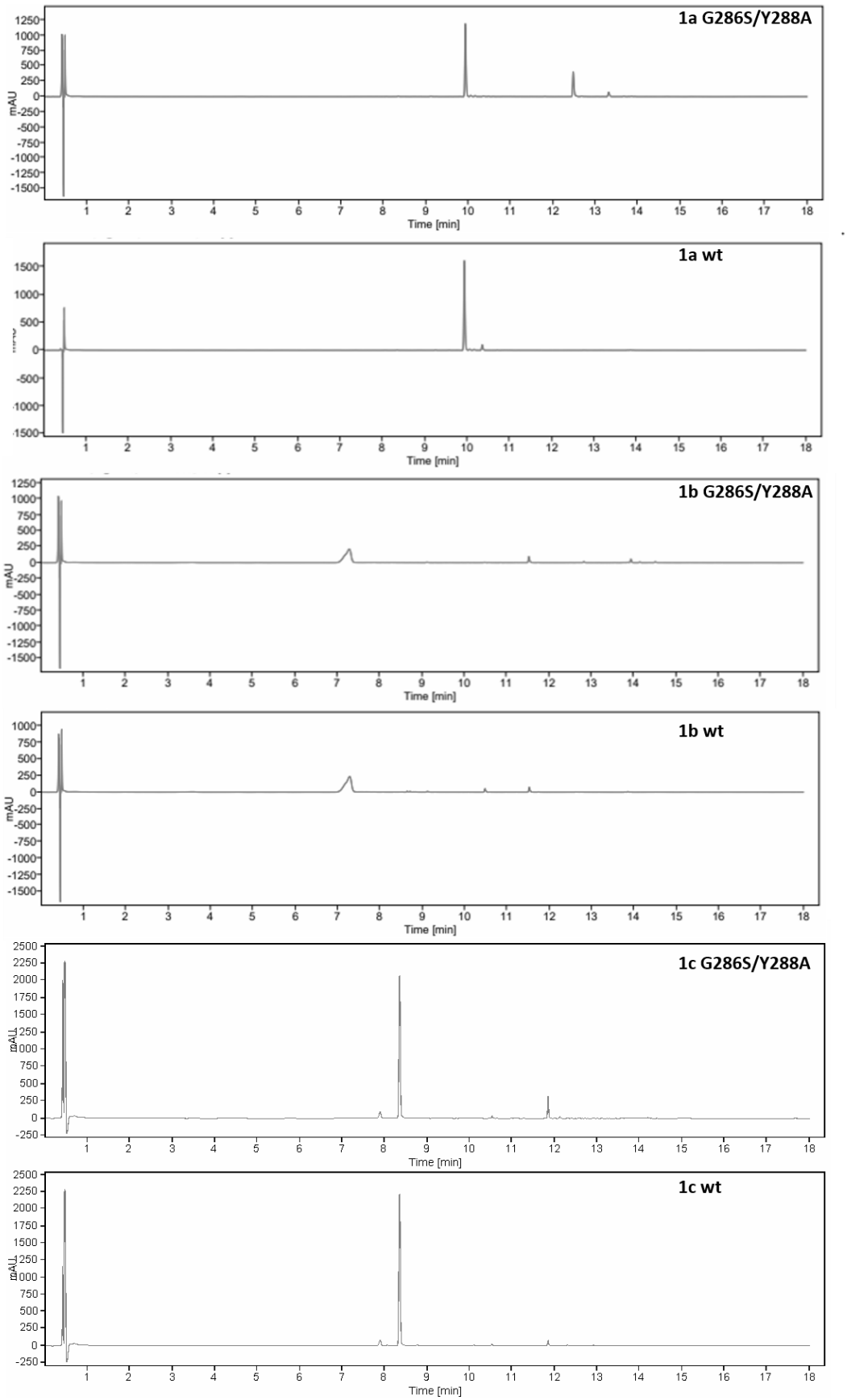


Figure S31: HPLC-UV measurement at 225 nm of the products from the conversion assays using NphB G286S/Y288A and wt and the substrates **1a**-**1c**.

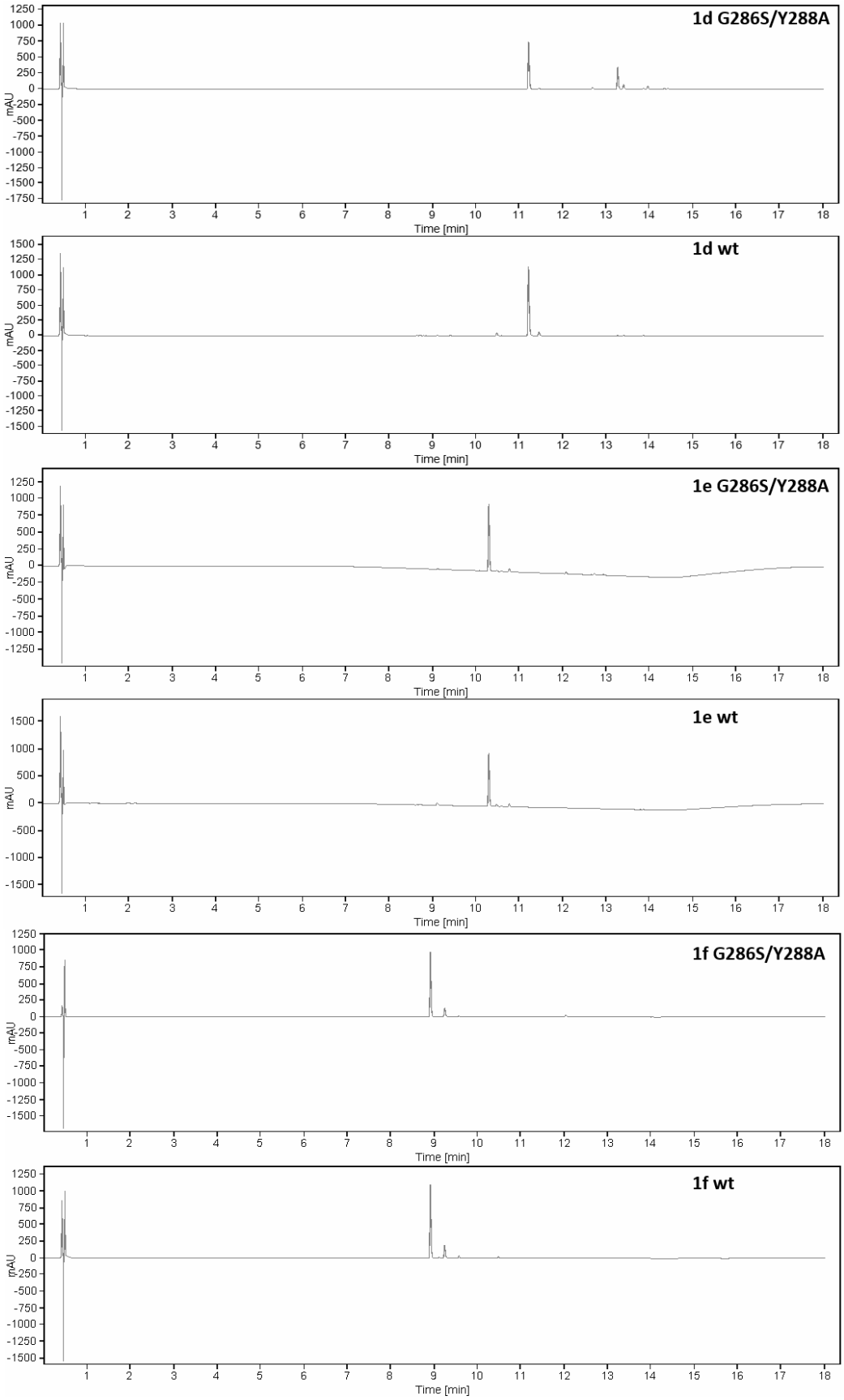


Figure S32: HPLC-UV measurement at 225 nm of the products from the conversion assays using NphB G286S/Y288A and wt and the substrates **1d**-**1f**.

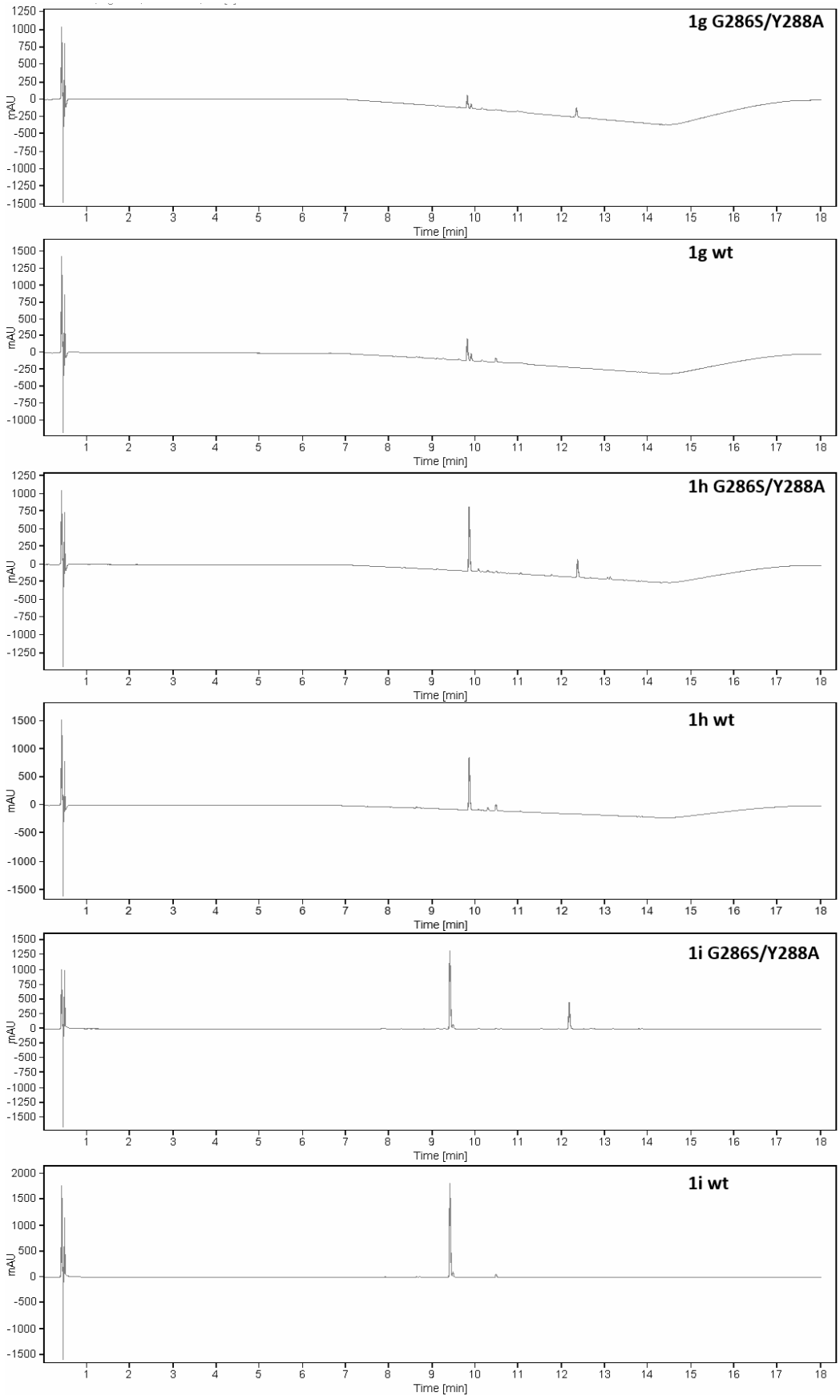


Figure S33: HPLC-UV measurement at 225 nm of the products from the conversion assays using NphB G286S/Y288A and wt and the substrates **1g**-**1i**.

Curriculum Vitae

Personal Information

Name	Saskia Spitzer
Date of birth	21.12.1995
Place of birth	Düren
Nationality	German

Academic education

since 01/2020 PhD student Technical biochemistry, Biochemical and chemical engineering, TU Dortmund, Germany

04/2017 – 08/2019 M.Sc. Chemical Biology, TU Dortmund, Germany

10/2013 – 12/2016 B.Sc. Chemical Biology ,TU Dortmund, Germany

School education

08/2005 – 06/2013 Abitur Franken-Gymnasium Zülpich, Germany

Working experience

since 08/2024 Area Sales Manager Carl Roth, Germany

Publications

Spitzer, S.; Aras, M.; Kayser, O. Improving CsOAC Activity in Saccharomyces cerevisiae for Directed Production of Olivetolic Acid through Rational Design. *Chembiochem : a European journal of chemical biology* **2024**, e202400651.

Spitzer, S.; Wloka, J.; Pietruszka, J.; Kayser, O. Generation of Cannabigerolic Acid Derivatives and Their Precursors by Using the Promiscuity of the Aromatic Prenyltransferase NphB. *Chembiochem : a European journal of chemical biology* **2023,** *24* (22), e202300441.

Humbeck, L.; Pretzel, J.; **Spitzer, S**.; Koch, O. Discovery of an Unexpected Similarity in Ligand Binding between BRD4 and PPARγ. *ACS chemical biology* **2021**, *16* (7), 1255–1265. DOI: 10.1021/acschembio.1c00323

Erklärung zur Reproduktion vorab veröffentlichter Inhalte

Die Inhalte dieser Forschungsarbeit sind in dem Projekt program VIPplus CannaCell (03VP06370) erarbeitet worden.

Teile dieser Arbeit sind bereits vom Autor veröffentlicht und präsentiert worden oder basieren auf Messdaten, die im Rahmen betreuter studentischer Arbeiten am Lehrstuhl für Technische Biochemie (Fakultät Bio- und Chemieingenieurwesen, TU Dortmund) entstanden sind.

Kapitel 2	2.1	in Teilen modifiziert aus	[A]
	2.1.1	in Teilen modifiziert aus	[a], [b]
	2.2	in Teilen modifiziert aus	[c]
	2.3	in Teilen modifiziert aus	[B]
	2.3.6	in Teilen modifiziert aus	[d]

Publikationen

[A] **Spitzer, S**.; Aras, M.; Kayser, O. Improving CsOAC Activity in Saccharomyces cerevisiae for Directed Production of Olivetolic Acid through Rational Design. *Chembiochem : a European journal of chemical biology* **2024**, e202400651.

[B] **Spitzer, S**.; Wloka, J.; Pietruszka, J.; Kayser, O. Generation of Cannabigerolic Acid Derivatives and Their Precursors by Using the Promiscuity of the Aromatic Prenyltransferase NphB. *Chembiochem : a European journal of chemical biology* **2023,** *24* (22), e202300441

Studentische Arbeiten

[a] Reckert, L. (2023) **Konstruktion eines Fusionsproteins zur Optimierung der Olivetolsäureproduktion in *Saccharomyces cerevisiae***

[b] Kempel, D. (2023) **Optimierung der Olivetolsäureproduktion durch Fusionsproteine in *Saccharomyces cerevisiae***

[c] Hallek, S. (2020) **Untersuchung von verschiedenen Carbonsäuren als alternative Substrate der Cannabinoidbiosynthese in *Saccharomyces cerevisiae***

[d] Donner, M. (2023) **Establishment of an alternative quantification system for cannabigerolic acid production**