ChemBioChem

Combining Chemistry and Biology

Chemistry Europe

European Chemical Societies Publishing

Accepted Article

Title: Generation of cannabigerolic acid derivatives and their precursors using the promiscuity of the aromatic prenyltransferase NphB

Authors: Saskia Spitzer, Jasmin Wloka, Joerg Pietruszka, and Oliver Kayser

To be cited as: *ChemBioChem* **2023**, e202300441

Link to VoR: https://doi.org/10.1002/cbic.202300441

WILEY VCH

10.1002/cbic.202300441

WILEY-VCH

RESEARCH ARTICLE

Generation of Cannabigerolic Acid Derivatives and their Precursors using the Promiscuity of the Aromatic Prenyltransferase NphB

Saskia Spitzer[a], Jasmin Wloka[b], Jörg Pietruszka[b] and Oliver Kayser*[a]

[a] S. Spitzer, Prof. Dr. Dr. h.c. O. Kayser
Technical Biochemistry Laboratory
Technical University Dortmund
44227 Dortmund
E-Mail: oliver.kayser@tu-dortmund.de

[b] J. Wloka, Prof. Dr. J. Pietruszka
Institute of Bioorganic Chemistry
Heinrich-Heine University Düsseldorf
52426 Jülich
E-Mail: j.pietruszka@fz-juelich.de

Supporting information for this article is given via a link at the end of the document.

Abstract: NphB is an aromatic prenyltransferase with high promiscuity for phenolics including flavonoids, isoflavonoids, and plant polyketides. It has been demonstrated that cannabigerolic acid is successfully formed by the catalytic reaction by NphB using geranyl diphosphate and olivetolic acid as substrates. In this study, the substrate specificity of NphB was further determined using olivetolic acid derivatives as potential substrates for the formation of new synthetic cannabinoids. The derivatives differ in the hydrocarbon chain attached to the C6 of the core structure. We performed i*n silico* experiments, including docking of olivetolic acid derivatives, to identify differences in their binding modes. Substrate acceptance were predicted. Based on these results, a library of olivetolic acid derivatives was constructed and synthesized using different organic synthetic routes. Conversion was monitored using *in vitro* assay with purified NphB versions. For the substrates leading to a high conversion olivetolic acid-C8, olivetolic acid-C2 and 2-benzyl-4,6-dihydroxybenzoic acid the products were further elucidated and identified as cannbigerolic acid derivatives. Therefore, these substrates show potential to be adapted in cannabinoid biosynthesis.

Introduction

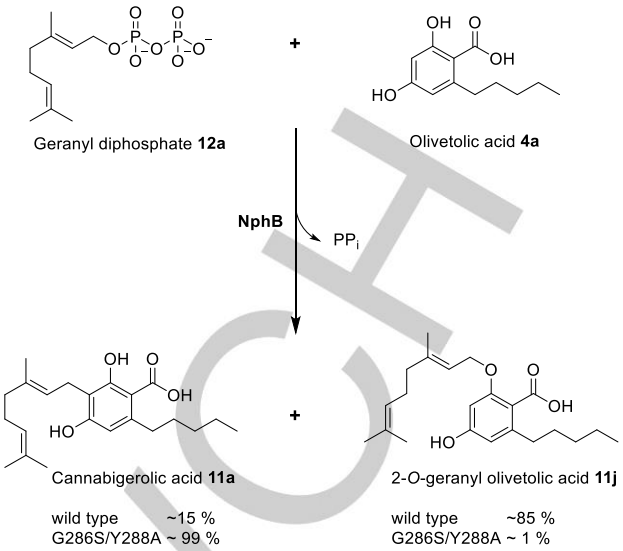
When discovered, cannabinoids were classified as a specific group of secondary metabolites isolated from the plant *Cannabis sativa* L. They consist of an alkylresorcinol and a monoterpene moiety.[1] Today, this particular class of cannabinoids is referred to as phytocannabinoids. The term cannabinoid is more broadly defined to include natural compounds acting on the endocannabinoid receptors CB1 and CB2 with biochemical responses similar to phytocannabinoids.[2] As a result of this discovery, the molecular space of synthetic cannabinoids has been explored in terms of receptor interactions.[3]
Phytocannabinoids are used in various fields and have shown promising results as therapeutic agents for diseases such as epilepsy, neurodegenerative diseases, schizophrenia or multiple sclerosis caused by affective disorders of the central nervous system.[4] The synthesis of phytocannabinoid-like analogs or unnatural synthetic cannabinoids could improve the medical response in human applications. New, unexplored possibilities for therapeutic use could be discovered. [5,6]
Cannabigerolic acid (CBGA-C5) is the committed precursor in the natural biosynthesis of phytocannabinoids. This molecule is converted to Δ^9-tetrahydrocannabinolic acid (THCA-C5) and cannabichromenic acid (CBCA-C5) by enzymes such as tetrahydrocannabinolic acid synthase (THCAS) and cannabichromenic acid synthase (CBCAS). CBGA-C5, therefore, plays a central role in the biosynthesis of cannabinoids, and its formation is crucial.[7–9] CBGA-C5 is formed by a Friedel-Crafts alkylation of olivetolic acid and geranyl diphosphate (GPP). The reaction is catalyzed by the enzyme geranyl diphosphate-olivetolic acid geranyltransferase CsPT4.[10] Transferring the biosynthetic pathway to another organism, such as *Saccharomyces cerevisiae* or *Escherichia coli*, requires biotechnological adaptation. CsPT4 is a plant integral membrane protein; substituting a soluble aromatic prenyltransferase is preferable. Supplying a different substrate than olivetolic acid with a structure similar to the natural one should not dramatically affect cannabinoid biosynthesis. Regarding substrate promiscuity, feeding analogues may open the door to mutasynthesis and new strategies for obtaining a new pool of synthetic cannabinoids.[5,10]

NphB, a protein isolated from *Streptomyces sp.* strain CL190, belongs to the enzyme class of aromatic prenyltransferases and the ABBA superfamily, including enzymes like CloQ or NovQ.[11]This family is characterized by a PT-barrel fold, composed of 10 antiparallel β-strands, which forms a β-barrel surrounded by alternating α-helices, resulting in an αββα pattern. This specific folding creates a spacious binding site, resulting in a relaxed prenyl-acceptor substrate specificity with substrates such as flavonoids, isoflavonoids, plant polyketides, and other naphthols.[12,13] The natural prenyl donor substrate is GPP, but the donor substrate specificity depends on the acceptor substrate.[14] The enzyme catalyzes a C-prenylation at the ortho- or para-position of a hydroxy group, but O-prenylation has also been observed on single hydroxy groups of flavonoids.[15] As mentioned, the reaction mechanism follows a Friedel-Crafts alkylation according to a S_N1-like dissociative mechanism. Within this reaction mechanism the first step is the formation of a carbocation with cleavage of the C-O bond from GPP. The

carbocation is stabilized within the binding pocket by the aromatic substrate. After the prenylation step, the product is rearomatized via a water-mediated proton transfer pathway.[9] The aromatic substrate has fluctuations in its active site orientations resulting in different prenylated products due to the structure of the binding pocket. Three other products, including one major and two minor products, were found for the substance 1,6-dihydroxynaphthalene. There is still an emerging interest in the structures of the isoprenoids. Thus, the relaxed substrate specificity of NphB may be advantageous for producing more isoprenoid structures depending on the substrate.[16] In previous experiments, olivetolic acid was identified as an aromatic acceptor substrate for NphB. Since NphB, like CsPT4, can catalyze the reaction to CBGA-C5, we extrapolated this fact to cannabinoid production. With about ~ 85 % the identified main product of this reaction is 2-*O*-geranyl-olivetolic acid (2-*O*-GOA). NphB carries out O- and C-geranylation with olivetolic acid (Scheme 1).[8,17] Regarding CBGA production a low k_{cat} = 0.0021 ± 0.00008 min^{-1} was observed. Still the conversion of olivetolic acid towards CBGA-C5 implies a high potential for NphB to be used in the biosynthesis of cannabinoids in a new host system, as has already been demonstrated in other studies using *Komagataella phaffi* or *E. coli*.[8,18,19] To exploit the full potential of the enzyme for CBGA-C5 production, protein engineering approaches were performed to shift the product selectivity towards CBGA-C5 and to enhance the catalytic activity. Rational design was performed using various *in silico* tools. It was possible to identify highly selective NphB variants that almost exclusively form CBGA-C5. The Q295F variant identified in our laboratory is one example.[20] A highly product-selective and active NphB variant consisting of two amino acid exchanges with G286S and Y288A was found, with a 1,000-fold higher k_{cat} = 1.58 ± 0.05, min^{-1} and almost exclusive CBGA-C5 formation.[19] The substitution of the indicated two amino acids has changed the space in the binding pocket, which affects the aromatic substrate acceptance or the resulting binding mode.

However, a clear understanding of mutagenesis is required to control the promiscuity of the enzyme and exploit it for the biosynthesis of semi-natural cannabinoids.

In this study, the substrate specificity of NphB is further evaluated concerning the conversion of various olivetolic acid derivatives. Novel substrates were synthesized by modifying the pentyl chain with different hydrocarbon moieties. *In silico* experiments were performed to generate a diversified substrate library. This substrate library was evaluated with *in vitro* assays regarding their conversion with NphB towards CBGA-derivatives.



Scheme 1. Enzymatic reaction of NphB using geranyl diphosphate and olivetolic acid as substrates.

Results and Discussion

Docking experiments were used to determine the most promising olivetolic acid derivatives in terms of their acceptance as a substrate and C-prenylation on the C3 position. These derivatives were synthesized and analyzed for their conversion using NphB wild type and the variant G286S/Y288A The enzymatic reaction products obtained were resolved by HR-MS. The prenylation pattern of the most promising candidates for implementation in the biosynthesis of cannabinoids has then been elucidated by NMR spectroscopy. Both analytical techniques have enabled the explanation and prediction of the influence of specific amino acid side chains in the designed NphB enzymes.

Docking studies to predict NphB substrate acceptance

Our proposal is that the regioselectivity of NphB depends on the binding mode of the aromatic substrate.[9] Recently, two X-ray structures of NphB were published showing both substrates, the prenyl donor and the acceptor. Thus, the catalytic and substrate-binding amino acids have been successfully identified.[21] The aromatic substrates are stabilized mainly by Met162, Phe213, Ser214, Tyr288, and Gln295 interactions. However, the prenyl donor is also responsible for further stabilization of the aromatic substrate The carbocation stabilization is ensured by forming a π-chamber, composed of the phenolic side chains of the amino acid Tyr121, Tyr216, and the aromatic substrate. The prenylated position is in close proximity to this carbocation and the π-chamber. Therefore the orientation of the aromatic substrate in the binding pocket shapes the prenylation pattern.[9]

With olivetolic acid as substrate, the orientation of the resorcylate core towards the catalytic center is required for the prenylation at C3 position leading to the formation of CBGA-C5. The novel olivetolic acid derivatives differ in the side chain but not in the core structure. Therefore, it is interesting to analyze the variation of the binding modes of the potential substrates. In order to identify a set of promising substrates that are likely to be converted by NphB and cover a large area of chemical space, selected substrates

(Table S1) were analyzed in terms of their proposed binding mode in docking experiments.

Because of their different catalytic activity and prenylation pattern for olivetolic acid, we used the wild type NphB (PDB-ID: 1zdw) and the G286S/Y288A variant. The docking experiments have been performed with Gold (Table S1, S2).[22] To establish a threshold to reduce the number of olivetolic acid analogs to be synthesized and tested *in vitro* for conversion, the docking of the highly characterized substrate olivetolic acid was performed and analyzed first.

The NphB wild type primarily catalyzes the formation of 2-*O*-GOA, and the variant G286S/Y288A predominantly forms CBGA-C5 using olivetolic acid **4a** and GPP as substrates. The resorcylate core is shifted towards the π-chamber by exchanging tyrosine at position 288 to a shorter amino acid with alanine and introducing a polar amino acid side chain with serine at position 286. Two different hydrogen bond networks stabilize **4a** when comparing the two proteins. For the NphB G286S/Y288A variant, a construct consisting of two hydrogen bonds between Ser214 and OH4 and between Ser286 and the carboxyl group was identified. For NphB wild type, this network consists of Ser214 and the carboxyl group, and Tyr288 and OH4 (Figure 1, see **4a** in Table S2). Compared to the wild type, the newly formed hydrogen bonds using the variant promote a 180° rotation of the resorcylate core within the binding pocket. A shift in the proximity between substrate and reaction partner GPP explains the different prenylation pattern of the variant compared to the wild type. Shorter distances were observed between C3 of olivetolic acid and C1 of GPP.

Both orientations of the resorcylate core reflect the different prenylation position. The loss of a bulky side chain of Tyr288 may be another explanation for the change in orientation. The additional space inside the binding pocket strengthens the hydrophobic interaction between the pentyl chain and Met162 or Phe213. These energetically favorable interactions further favor the rotation of the resorcylate core by the G286S/Y288A variant. This is also supported by a comparison with Lim et al, who concluded that a hydrogen bond between OH4 and Tyr288 leads to an unfavorable orientation for CBGA-C5 production.[23]

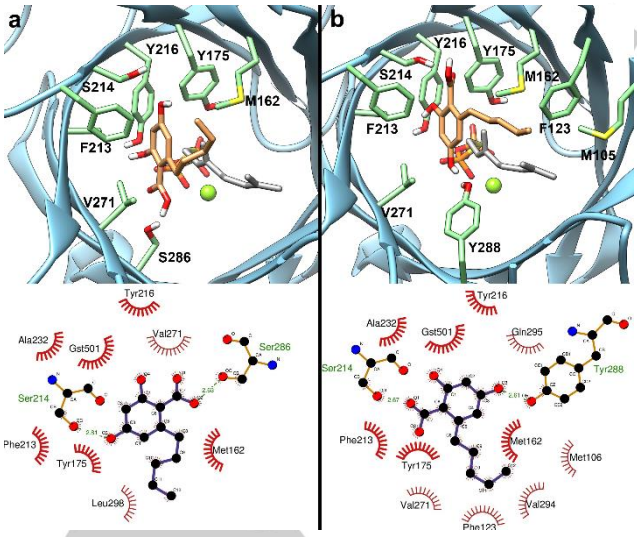


Figure 1. 3D depiction created using UCSF chimera and 2D deception created using LigPlot of the docked pose of **4a** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acids (green).[24,25]

Once the binding mode of olivetolic acid has been analyzed, the potential olivetolic acid derivatives for analysis were docked. The objective is to create a library of molecules that are accepted as substrates of NphB, resulting in the formation of novel products. The most important feature of the proposed binding mode is that the resorcylate core and therefore the prenylation position C3 is located close to C1 of the GSPP and the side chain is oriented towards the entrance of the binding pocket. In X-ray structures, dockings, and MD simulations, a distance of 3 to 4.5 Å between the two atoms was identified.[9,13,23] To evaluate whether a substrate is likely to be converted by NphB, the distance between the C3 of the resorcylate core and C1 of GSPP was calculated for each pose. In total 50 poses were created for each molecule and these poses were analyzed regarding this distance. A distribution of whether the distance of these atoms within the calculated poses matches the distance range is shown in table S1.

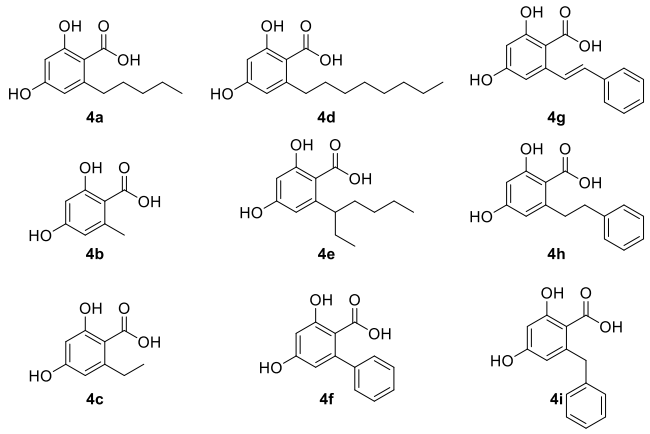
The distributions of poses with matching distances are comparable for all substrates containing alkyl chains of different lengths using the NphB variant G286S/Y288A. In each pose, the distance of the reaction positions is lesser then 4.5 Å. The binding pocket of NphB covers an area of approximately 500 Å^3 for both substrates. Due to its size, an extended or branched alkyl chain attached to the resorcylate core will is likely to fit into the pocket. This has been demonstrated in converting an analog containing a tridecane side chain to CBGA-C13, but the ratio of possible products formed is unknown.[26] Previous studies have reported that olivetolic acid derivatives with shorter alkyl chains, such as the methyl or propyl chain, are converted to CBGA-C1 and CBGA-C3, respectively. Orsellinic acid is converted to various products by the NphB wild type. With the G286S variant, the main product was shifted to CBGA-C1.[18] Regarding these information we decided to include the extension and shortening of the *n*-pentyl chain, resulting in non-natural building blocks such as the ethyl or *n*-octyl chain in **4c** and **4d**. Since introducing a branched side chain still allows a plausible binding mode, we also decided to incorporate **4e** into our library.

Using NphB wild type, all analogs carrying an aromatic side chain mainly led to incorrectly oriented poses of the acceptor inside the active site. Using the NphB variant G286S/Y288A, molecules were excluded from further analysis if at least 50 % of the calculated poses, and thus the most prominent poses have no distance of ≤ 4.5 Å between the C3 from the aromatic acceptor C1 from GSPP (Table S1). Most analogs carrying an aromatic side chain do not meet this criterion. For example each structure containing a naphthyl group led to a misorientation mainly due to its size. We decided to focus on the structures containing a phenyl group rather than a phenolic group since a similar linkage to the core structure led to better results. Three structures **4f**, **4h** and **4i** were selected for further *in silico* and *in vitro* experiments, as they mainly showed the correct orientation. They all consist of a second phenyl group connected to the core by different bonds.

All molecules included are summarized in Scheme 2, including the synthesis yield achieved.

10.1002/cbic.202300441

RESEARCH ARTICLE



Scheme 2. List of chosen compounds **4a-4i** that were synthesized and validated in this study regarding their interaction with NphB.

For all molecules (Scheme 2) the binding mode was further analyzed. For the most interesting substrates regarding differences in the binding mode, the docking result will be discussed in more detail.

The docking results of **4d** (Figure 2) indicate that an extended side chain towards *n*-octyl still fits into the binding pocket. The calculated binding modes of the core structure of substrates **4d** were similar to those of olivetolic acids upon docking into the NphB variant. The hydrogen bond network of the resorcylate core is identical to the ones of **4a**. This network ensures the correct position of the core structure leading to prenylation at the C3 group. For the wild type hydrogen bonds were identified between Ser214 and O4 and between Tyr288 and the carboxyl group. Thus, the position of the resorcylate core is different from the one that was previously reported for **4a** and the wild type. The benzene-1,3-diol group is turned away from the amino acid side chains that are a part of the π chamber. This may cause difficulties in generating the π-chamber and therefore difficulties in stabilization could be caused. This likely hampers the enzymatic activity. One of the reasons for the alternative binding mode is the tyrosine at position 288. To avoid a steric collision between the spacious side chain and the alkyl chain, the resorcylate core is oriented differently. With less space, the longer alkyl chain only allows a different core position of the core.

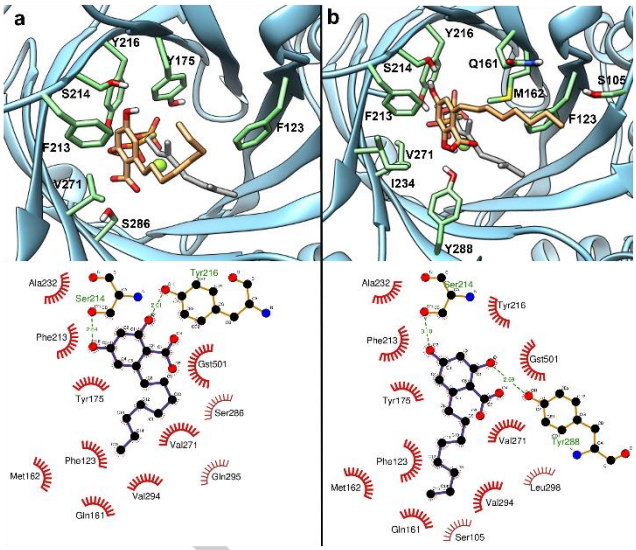


Figure 2. 3D depiction created using UCSF chimera and 2D deciption created using LigPlot of the docked pose of **4d** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green).[24,25]

Besides the variation of the alkyl chain, a complete shift towards a cyclic structure is of great interest for discovering new potential drugs. Different linkers between the resorcylate core and the phenyl group were used to synthesize **4f**, **4g**, **4h** and **4i**. The resulting binding mode changes were analyzed to identify the most promising derivatives.

The highest ranked pose for molecules **4f**, **4g** and **4h** differs from the results obtained for the previously analyzed derivatives **4a** - **4e**, using either the variant or the wild type (Figure S5, S6). As a representative example, the proposed binding mode of **4f** is further discussed (Figure 3). The newly introduced phenyl group faces the π-chamber instead of the resorcylate core in the highest ranked pose. Potential prenylation positions are distant from the later formed carbocation. Nevertheless, the resorcylate core forms hydrogen bonds with the same amino acids, for example Ser286. Hydrogen bonding can be observed even when the core structure is far from the active site. This is achieved by changing the conformation of the amino acid side chain. Compared to an alkyl moiety, the phenylic group is more similar in shape to the resorcylate core. A short linkage between the core and the phenyl moiety restricts flexibility. When the positions of the phenyl group and the resorcylate core are swapped, the phenyl group is in a more hydrophobic environment.

In contrast, the core structure is still able to form hydrogen bonds. Consequently, the kinetic energy of the protein substrate system is not maintained by a binding mode similar to that of olivetolic acid. With this conclusion, the conversion of **4f**, **4g** and **4h** to the corresponding CBGA derivatives is also constrained. However, a binding mode comparable to **4a** was also calculated, but with a lower scoring value (Table S1), suggesting that substrate acceptance is still possible.

10.1002/cbic.202300441

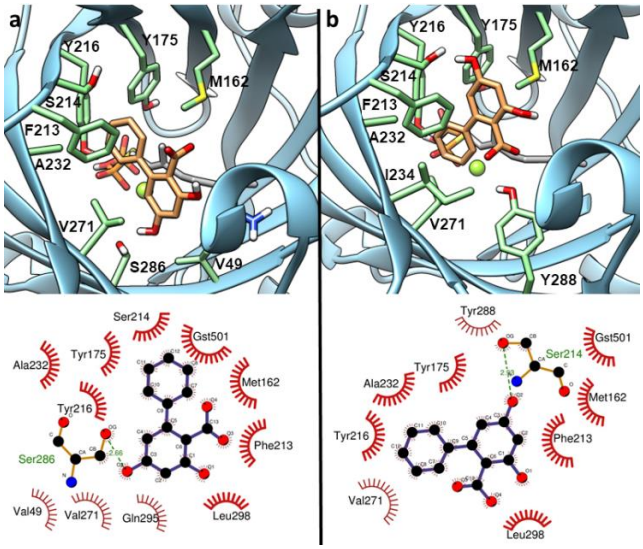


Figure 3 3D depiction created using UCSF chimera and 2D deciption created using LigPlot of the docked pose of **4f** (orange) inside the binding pocket of NphB G286S/Y288A (a) and wild type (b) with highlighted interacting amino acid (green).[24,25]

A variation of the proposed binding mode was found using a benzyl group in **4i** and the variant G286S/Y288A (Figure S7). The highest ranked pose (Table S2) resembles the olivetolic acid binding mode in terms of the hydrogen bond network, but not the other derivatives containing a phenylic moiety. The bond between the core structure and the aromatic group can explain this discrepancy. The benzyl group provides additional flexibility and rotational freedom. 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. This result cannot be transferred to the wild type because the bulky side chain of Tyr288 blocks the above orientation. A better substrate acceptance is anticipated for **4i** compared to **4e** - **4g** with the variant.

The docking study describes the differences between prenylation catalytic activity and the different patterns of the two protein versions. In the binding mode calculation, the exchange of two amino acids within the binding pocket can lead to different results. The binding mode of **4a** – **4e** is similar with respect to the hydrogen bonds formed and therefore the position of the resorcylate core. Only the number of unbound contacts varies with the length of the alkyl chain. A substrate acceptance of all molecules (Scheme 2) is likely using NphB G286S/Y288A. With the absence of the hydrogen bond to Y288, the formation of CBGA derivatives with the protein and alkyl chains of different sizes can be expected. By introducing a more similar structural element compared to the resorcylate core in **4f** - **4h**, the substrate can be stabilized in an unfavorable position. For these molecules, the conversion is likely to be lower than when using **4a**. This does not apply to **4i**, which is expected to show high conversion and correct prenylation.
The *in silico* selected substrates (Scheme 2) were studied for conversion by NphB wild type and variant.

Substrate conversion of novel substrates with NphB

The olivetolic acid derivatives selected by docking were synthesized and characterized (Supporting Information). 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 requiered 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 **4** (Scheme 2) 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 5). 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 a overview about the most promising substrates regarding usage in the cannabinoid biosynthesis.

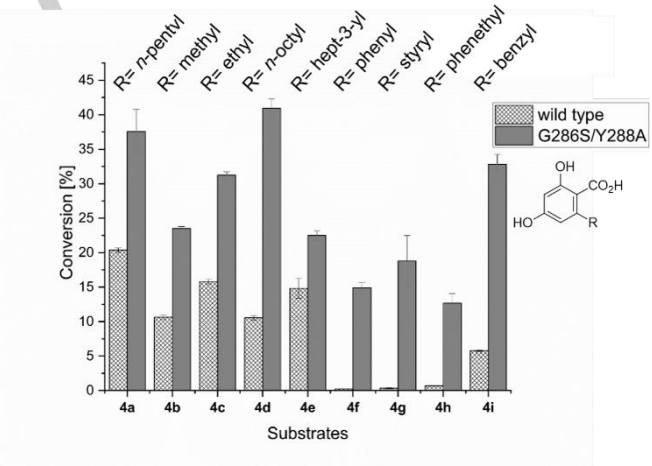


Figure 4. Conversion of all tested substrates **4a**-**4i** using GPP as prenyl donor and NphB wild type (pattern) and NphB G286S/Y288A (grey) for conversion, respectively. All experiments were performed in triplicate.

The conversion of olivetolic acid has been described previously and was therefore used as a control reaction for the isolated protein and its function.[8] 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 **4a**, marking the 40 % yield obtained with NphB G286S/Y288A as a reference point.
For all substrates **4a** - **4i** product formation was visible in the HPLC chromatogram using the variant G286S/Y288A (Figure S90 – S92). Overall, a higher conversion was determined for each substrate using the variant compared to the wild type. For orsellinic acid **4b** the results were already confirmed before,

with the identification of geranylated products.[8,18,19] In general, for molecules with various alkyl chains, regardless of length, the variant resulted in a higher conversion of up to 40 %. A trend was observed that shortening the alkyl chain correlated with lower conversion. This result corresponds to the docking experiments: for a shorter alkyl chain, less interactions were identified, and the overall ranking was lower (Table S2), implying a lower binding affinity.

The highest conversion for wild type was observed with olivetolic acid with up to 20 %. Using the wild type, conversion was also detected for the derivatives with different alkyl chain lengths. A considerable difference between the two variants was observed using **4d**. While a high conversion of 40 % was determined for the G286S/Y288A variant, only a low conversion of 10 % was observed for the wild type. 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 **4d**, like predicted in docking experiments (Figure 2).

Comparing the substrate acceptance of compounds **4e**, **4f**, **4g**, and **4h**, 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 3). The linkage between the core structure and the phenyl moiety influences the conversion for the variant. The only substrate molecule with an aromatic side chain that led to a high conversion is **4i**. A conversion rate of 33% was determined. In parallel with the *in vitro* results (Figure 4), molecule **4i** had the most promising docking outcome (Figure S7) compared to the other derivatives containing a second phenolic structural element. The hypothesis that structures **4f** - **4h** would not be converted as well as olivetolic acid **4a** confirms the proposed docking studies.

As only a low conversion has been identified for the molecules **4e - 4h**, 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.

Structure elucidation of enzymatic products

All assay samples were analyzed by HR-MS. Mass analysis allows confirmation of prenylated products and determination of the number of products formed by catalytic reaction with NphB (Figure S71, S79-83). 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 **4a**, **4b**, **4c**, **4d** and **4i**, the reaction producs were further analyzed by UPLC-ESI-MS/MS (Figure S72 - S75 and S84). The fragmentation pattern was similar to that of CBGA-C5, suggesting the formation of a C-prenylation. The two pattern distinguish in the formation of fragments smaller than *m/z* = 219.10. While the fragmentation of 2-*O*-GOA leads to smaller fragnents in a high distribution, CBGA-C5 does only have these fragments in traces. This was also observed for the beforementioned substrates. [8,27] 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 **4a** was not observed as a product in *in vitro* assays with NphB wild type or G286S/Y288A.[8,19,23] As the space of the group attached to C6 increases, the C5 position becomes even more blocked.

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 **4d** and **4i.** 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 RP-HPLC. ^{1}H and ^{13}C NMR, COSY, HSQC and HMBC were measured for **11i** (Table S7, Figure S85-89). Due to the HMBC correlation between H8 and C5, the isolated product has to be prenylated at C3 position. For **4i** the conversion towards a CBGA derivative **11i** was demonstrated. ^{1}H NMR, HMBC and HSQC were measured for **11d** (Figure S76, 77). With this we were able to confirm a C-prenylation 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.

For further identification of the prenylation site using **4c**, **4d** and **4i** assays employing the protein THCAS were conducted. As described before the THCAS is the last protein in the cannabinoid biosynthesis. It catalyzes the intramolecular cyclization of the monoterpene moiety in CBGA-C5 to THCA-C5 (Scheme S5). It has already been demonstrated, that this enzymatic reaction also occurs with CBGA derivatives that differ only in their side chain.[19] 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 THCAS. As substrates GPP and the most promising analogs **4a**, **4c**, **4d** and **4i** 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 **4a** as substrate. With comparsion of the fragmentation pattern and retention time of an analytic standard of THCA-C5 and CBCA-C5 the formation of these products using **4a** and GPP as substrates was confirmed (Figure S93). As described in former reports THCAS catalyzes the formation of THCAS-C5 and CBCA-C5 as main products using higher pH values.[28] Using the substrates **4c**, **4d** and **4i** 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 S93, S94). With this further proof, we can confirm the prenylation on C3 position for **4c**, **4d** and **4i** in the *in vitro* assays described above because only then the production of THCA and CBCA derivatives catalyzed by THCAS is possible.

Conclusion

Novel compounds containing the olivetolic acid resorcylate core and structural elements beyond the natural pentyl group were synthesized and evaluated as potential substrates for NphB. For all compounds, prenylated product masses were successfully identified after performing an *in vitro* assay. Alkyl chain modifications up to an octyl group attached to the core are accepted as substrates and show similar conversion to olivetolic acid. Similar conversion was also observed when a benzyl group completely replaced the alkyl group. A further analysis of the prenylation pattern was carried out for the most promising molecules. In an organism, like *S. cerevisiae*, NphB expression and olivetolic acid derivatives feeding could lead to CBGA derivatives formation.

First experiments using the following protein of the cannabinoid biosynthesis, the THCAS, were performed showing the potential for the production of further unnatural cannabinoids. The presented results show the possibility of predicting substrate acceptance and prenylation patterns of NphB and olivetolic acid derivatives *in silico* by docking. The docking experiment results provide a better understanding of the prenylation pattern and substrate binding.

In conclusion, due to its promiscuous substrate acceptance, designed NphB variants are attractive catalysts for the creation of a library of non-natural cannabinoids.

Experimental Section

Molecular docking. Docking experiments were performed using GOLD.[22] The protein structure (PDB-ID: 1ZDW) were prepared using UCSF Chimera.[24] For the variant G286S/Y288A the amino acids on position 286 and 288 were exchanged. The ligands were prepared using KNIME.[29] 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 binding pocket was defined using the aromatic structure bound in the PDB structure and a 5 Å radius around this molecule. As scoring function ASP-score was used. For each ligand 50 GA runs were performed. The search efficiency for the GA runs were set to 200 %.

Protein expression and purification. The amino acid sequences of NphB fused with a N-terminal His-Tag are shown in the SI. For the recombinant expression, the gene construct was integrated into the expression vector pDionysos, which uses the pYES2 yeast expression vector (Fisher Scientific GmbH, Schwerte, Germany) with an additional Leu2d gene, under control of a Gal1 promotor.[30] The plasmids containing a construct for the synthesis of the recombinant NphB were transformed into *S. cerevisiae* CEN.PK2-1C Δpep4 Δgal1 Δgal80 through LiAc/SS-DNA/PEG method.[31] A single colony was picked to inoculate 10 mL synthetic mineral salt medium (6.7 g/L YNB without amino acids, 1.62 g/L drop out supplements without leucine, 20 g/L glucose) and incubated for 24 h at 30 °C and 200 rpm. The second pre culture was inoculated with the first one to an OD_{600} value of 0.1~0.2 and was incubated for 12 h at 30 °C and 200 rpm. For the main culture 100 mL complex medium (10 g/L yeast extract, 20 g/L-peptone, 20 g/L glucose) in 1 L baffled flasks were inoculated OD_{600} = 0.5 and cultivated for 48 h by 20 °C and 200 rpm**.**

After cultivation the cells were harvested by centrifugation (3000 *g*, 5 min, 4 °C) and the supernatant was discarded. The cells were re-suspended in 5 mL lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 30 mM Imidazole, pH 7,5) for each gram of cell dry pellet. The cells were lysated using glass beads (0.75 – 1.00 mm) and vortexing (20 min, 4 °C, 2.000 *g*). Afterwards cell debris were removed by centrifugation (15.000 rpm, 30 min, 4 °C). The proteins were purified using the ÄktaPure System (Ge Healthcare, Solingen, Germany) and a HisTrap™ HP column (Cytiva Europa, Freiburg im Breisgau, Germany). The preequilibration was performed using 10 mL of binding buffer (50 mM Tris-HCl, 500 mM NaCl, 30 mM Imidazole, pH 7.5). Afterwards the filtered cell solution was loaded onto the column (1 mL/min flow rate). To remove other proteins from the column, a washing step with 10 mL binding buffer was performed afterwards. The protein was eluted using a linear gradient of imidazole between 30 mM and 500 mM using 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 last step the desalting was performed using a PD10 column (Cytiva Europa, Freiburg im Breisgau, Germany) to replace the buffer to 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 before being stored at -80 °C until usage.

Synthesis of substrates 4. For a detailed description of strategies and experimental results see Supporting Information.

***In vitro* substrate conversion assay**. The enzyme assays were performed using a storage buffer. The potential prenyl acceptor substrates were dissolved in DMSO and added in a final concentration of 1 mM to a total volume of 120 µL containing 5 mM $MgCl_2$, 2 mM GPP and 100 µg purified protein. The reaction mixture was incubated for 18 h at 37 °C and 1100 rpm in the dark. The reaction was 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 minutes 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

Crude cell extract assay. The amino acid sequences of THCAS fused with a N-terminal vacuole tag are shown in the SI. The cultivation of cells containing NphB is similar to the protocol described in Protein expression and purification. For each assay cells with an OD_{600} = 125 were harvested. For the expression of THCAS, a pDionysus expression vector containing the gene for THCAS expression was used, the cultivation protocol was the same as described before. The harvested cells were each resuspended in 400 µL assay buffer (50 mM Tris-HCl, 100 mM NaCl pH = 7.5) and lysated using glass beads (0.75 – 1.00 mm) and vortexing (20 min, 4 °C, 2.000 *g*). Afterwards cell debris were removed by centrifugation (8.000 rpm, 30 min, 4 °C). The potential prenyl acceptor substrates were dissolved in DMSO and added in a final concentration of 1 mM to a total volume of 200 µL containing 5 mM $MgCl_2$, 2 mM GPP and 95 µL of each crude cell extract for the two different enzymes. The reaction mixture was incubated for 24 h at 37 °C and 1100 rpm in the dark. The reaction was 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 minutes 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.

HPLC-UV analysis. The samples were analyzed by HPLC-DAD (Agilent Infinity II 1260, Waldbronn, Germany). The separation was carried out using a Poroshell 120 EC-C18 (2.1 × 100 mm, 2.7 µm; Agilent, Waldbronn, Germany) column. The mobile phases A and B were ddH_2O + 0.1 % formic acid and Acetonitrile respectively. The following method was used with a flow rate of 0.6 mL min^{-1} and an oven temperature of 40 °C: 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. The aromatic substrates and products were detected in the DAD at 225 nm.

UPLC-ESI-MS and UPLC-ESI-MS/MS. The samples were analyzed by chromatography (Agilent 1290 Infinity II UPLC, Waldbronn, Germany) with coupled electrospray ionization source (ESI) and Bruker Compact MS/qTOF. The separation was carried out using a Poroshell 120 EC-C18 (2.1 × 100 mm, 2.7 µm; Agilent, Waldbronn, Germany) column. The mobile

phase A and B were ddH_2O + 0.1 % formic acid and Acetonitrile. The following method for the analyzation of the *in vitro* assay samples was used with a flow rate of 0.6 mL min^{-1} and an oven temperature of 40 °C: 30 % B for 1 min, gradient from 30 % to 100 % B in 6 min, and gradient from 100 % to 30 % B in 0.5 min, 30 % B for 2.5 min. For the seperation of the crude cell extract assay samples the following method was used: 30 % B for 1 min, gradient from 30 % to 85,5 % B in 22,5 min, gradient from 85,5 % to 100 % in 1 min, 100 % B for 0.5 min, from 100 % to 70 % B in 1 min and 70 % B for 3 min. All MS acquisitions were performed in positive and negative mode. All MS/MS acquisition was performed with a collision energy of 36.0 V in the multireaction monitoring (MRM) mode.

Scale-up reaction. The enzymatic assays were performed using storage buffer. The potential prenyl acceptor substrates were dissolved in DMSO and added in a final concentration of 1 mM to a total volume of 1.5 mL containing 5 mM $MgCl_2$, 2 mM GPP and 1.5 mg purified protein. The reaction mixture was incubated for 24 h at 37 °C and 1100 rpm in the dark. The reaction was 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 minutes the samples were centrifuged (13,000 *g*, 4 °C, 30 min) and supernatant was stored by -20 °C until usage.

The products were purified by preparative HPLC (Shimadzu, Duisburg, Germany). The separation was carried out using Nucleodur EC-C18 (125 × 21 mm, 5 µm; Macherey-Nagel, Düren). The mobile phases A and B were ddH_2O and Acetonitrile 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.

Supporting Information

The authors have cited additional references within the Supporting Information.[32-45]

Acknowledgements

This work was financially supported by the German Federal Ministry of Education and Research in the program VIPplus for the project CannaCell (03VP06370). We gratefully acknowledge the state of North Rhine Westphalia (NRW) and the European Regional Development Fund (EFRE) for funding the project within the ‚CLIB-Kompetenzzentrum Biotechnologie', grant numbers 34-EFRE-0300096, 34-EFRE-0300097, and 34-EFRE-0300098. We would like to thank Christina Schmidt for the vector containing the gene for THCAS expression and Kevin Bochinsky and Till Steinmetz for the help with the measurement and analyzation of the 2D-NMR spectra.

Keywords: Cannabinoids • Enzyme catalysis • NphB • Prenytransferase • Protein design

[1] Y. Gaoni, R. Mechoulam, *J. Am. Chem. Soc.* **1964**, *86*, 1646–1647.
[2] W. A. Devane, F. A. Dysarz, M. R. Johnson, L. S. Melvin, A. C. Howlett, *Mol. Pharmacol.* **1988**, *34*, 605–613.
[3] a) C. Prandi, M. Blangetti, D. Namdar, H. Koltai, *Molecules* **2018**, *23*; b) R. G. Pertwee, *Br. J. Pharmacol.* **2006**, *147 Suppl 1*, S163-71;
[4] M. N. Hill, J. G. Tasker, *Neuroscience* **2012**, *204*, 5–16.
[5] K. J. H. Lim, Y. P. Lim, Y. D. Hartono, M. K. Go, H. Fan, W. S. Yew, *Molecules* **2021**, *26*.
[6] A. Zagzoog, K. A. Mohamed, H. J. J. Kim, E. D. Kim, C. S. Frank, T. Black, P. D. Jadhav, L. A. Holbrook, R. B. Laprairie, *Sci. Rep.* **2020**, *10*, 20405.
[7] Y. Shoyama, M. Yagi, I. Nishioka, T. Yamauchi, *Phytochemistry* **1975**, *14*, 2189–2192.
[8] B. Zirpel, F. Degenhardt, C. Martin, O. Kayser, F. Stehle, *J. Biotechnol.* **2017**, *259*, 204–212.
[9] Y. Yang, Y. Miao, B. Wang, G. Cui, K. M. Merz, *Biochemistry* **2012**, *51*, 2606–2618.
[10] X. Luo et al., *Nature* **2019**, *567*, 123–126.
[11] M. Tello, T. Kuzuyama, L. Heide, J. P. Noel, S. B. Richard, *Cell. Mol. Life Sci.* **2008**, *65*, 1459–1463.
[12] O. Saleh, Y. Haagen, K. Seeger, L. Heide, *Phytochemistry* **2009**, *70*, 1728–1738.
[13] T. Kuzuyama, J. P. Noel, S. B. Richard, *Nature* **2005**, *435*, 983–987.
[14] B. P. Johnson, E. M. Scull, D. A. Dimas, T. Bavineni, C. Bandari, A. L. Batchev, E. D. Gardner, S. L. Nimmo, S. Singh, *Appl. Microbiol. Biotechnol.* **2020**, *104*, 4383–4395.
[15] a) T. Kumano, S. B. Richard, J. P. Noel, M. Nishiyama, T. Kuzuyama, *Bioorg. Med. Chem.* **2008**, *16*, 8117–8126; b) K. Shindo, A. Tachibana, A. Tanaka, S. Toba, E. Yuki, T. Ozaki, T. Kumano, M. Nishiyama, N. Misawa, T. Kuzuyama, *Biosci., Biotechnol., Biochem.* **2011**, *75*, 505–510;
[16] K. W. George, J. Alonso-Gutierrez, J. D. Keasling, T. S. Lee, *Adv. Biochem. Eng./Biotechnol.* **2015**, *148*, 355–389.
[17] F. Degenhardt, PhD thesis, Technische Universität Dortmund, Dortmund, **2018**.
[18] S. Qian, J. M. Clomburg, R. Gonzalez, *Biotechnol. Bioeng.* **2019**, *116*, 1116–1127.
[19] M. A. Valliere, T. P. Korman, N. B. Woodall, G. A. Khitrov, R. E. Taylor, D. Baker, J. U. Bowie, *Nat. Commun.* **2019**, *10*, 565.
[20] O. Kayser, F. O. Stehle (UNIV DORTMUND TECH [DE]), WO2020016287 (A1), **2019**.
[21] T. Kuzuyama, J. P. Noel, S. B. Richard, *Co-crystal structure of Orf2 an aromatic prenyl transferase from Streptomyces sp. strain CL190 complexed with TAPS*, **2005**.
[22] G. Jones, P. Willett, R. C. Glen, A. R. Leach, R. Taylor, *J. Mol. Biol.* **1997**, *267*, 727–748.
[23] K. J. H. Lim, Y. D. Hartono, B. Xue, M. K. Go, H. Fan, W. S. Yew, *ACS Catal.* **2022**, *12*, 4628–4639.
[24] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[25] A. C. Wallace, R. A. Laskowski, J. M. Thornton, *Protein Eng.* **1995**, *8*, 127–134.
[26] Y.-E. Lee, T. Kodama, H. Morita, *J. Nat. Med.* **2022**, 1–8.
[27] G.-N. Nguyen, E. N. Jordan, O. Kayser, *J. Nat. Prod.* **2022**, *85*, 1555–1568.
[28] B. Zirpel, O. Kayser, F. Stehle, *J. Biotechnol.* **2018**, *284*, 17–26.
[29] M. R. Berthold, N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel in *Studies in Classification, Data Analysis, and Knowledge Organization* (Eds.: C. Preisach, H. Burkhardt, L. Schmidt-Thieme, R. Decker), Springer Berlin Heidelberg, Berlin, Heidelberg, **2008**.
[30] F. Stehle, M. T. Stubbs, D. Strack, C. Milkowski, *FEBS J.* **2008**, *275*, 775–787.
[31] a) K.-D. Entian, P. Kötter in *Methods in Microbiology*, Elsevier, **2007**; b) R. D. Gietz, R. H. Schiestl, *Nat Protoc* **2007**, *2*, 31–34;
[32] T. K. Kim, J. E. Kim, U. J. Youn, S. J. Han, I.-C. Kim, C.-G. Cho, J. H. Yim, *J. Nat. Prod.* **2018**, *81*, 1460-1467.
[33] H. Lin, T. Annamalai, P. Bansod, Y.-C. Tse-Dinh, D. Sun, *Med. Chem. Comm.* **2013**, *4*, 1613-1618.
[34] X.-J. Xu, T. Zeng, Z.-X. Huang, X.-F. Xu, J. Lin, W.-M. Chen, *J. Nat. Prod.* **2018**, *81*, 2621-2629.
[35] K. Nakata, C. Feng, T. Tojo, Y. Kobayashi, *Tetrahedron Lett.* **2014**, *55*, 5774-5777.
[36] M. Sisa, M. Dvorakova, T. Vanek, *Tetrahedron* **2017**, *73*, 5297-5301.
[37] M. H.-W. MULLER GEORGE W., ((CELGENE CORP)) in *US2006052596A1*, **2006**.
[38] M. Bös, R. Canesso, N. Inoue-Ohga, A. Nakano, Y. Takehana, A. J. Sleight, *Bioorg. Med. Chem.* **1997**, *5*, 2165

2171.
[39] G. Navarro, A. Gonzalez, A. Sanchez-Morales, N. Casajuana Martin, M. Gómez-Ventura, A. Cordomí, F. Busqué, R. Alibés, L. Pardo, R. Franco, *J. Med. Chem.* **2021**, *64*, 9354-9364.
[40] S. C. Dakdouki, D. Villemin, N. Bar, J. Org. Chem., **2010**, *2*, 333-337.
[41] C. R. Chen, S. Zhou, D. B. Biradar, H. M. Gau, *Adv. Synth. Catal.* **2010**, *352*, 1718-1727.
[42] B. M. Kemkuignou, A. Y. Moussa, C. Decock, M. Stadler, *J. Nat. Prod.* **2022**, *85*, 846-856.
[43] B. Lü, C. Fu, S. Ma, *Tetrahedron Lett.* **2010**, *51*, 1284-1286.
[44] C. C. McAtee, P. S. Riehl, C. S. Schindler, *J. Am. Chem. Soc.* **2017**, *139*, 2960-2963.
[45] Z. Hu, F. Belitz, G. Zhang, F. Papp, L. J. Gooßen, *Org. Lett.* **2021**, *23*, 3541-3545.

10.1002/cbic.202300441

WILEY-VCH

RESEARCH ARTICLE

Entry for the Table of Contents

The conversion of novel olivetolic acid derivatives with the highly promiscuous prenyltransferase NphB is analyzed as a tool for the creation of synthetic cannabinoid libraries. With *in silico* and *in vitro* experiments CBGA derivatives were synthesized and characterized as products of enzyme catalysis.