Biochemical and Biophysical Research Communications 696 (2024) 149471

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Foldseek reveals a CBGA prenylating enzyme GlyMa_02G168000 from *Glycine max*

Erin Noel Jordan, Christina Schmidt, Oliver Kayser *

Technical Biochemistry, TU Dortmund University, Emil-Figge-Straße 66, 44227, Dortmund, Germany

ARTICLE INFO

Keywords:
Cannabinoids
Cannabigerolic acid
Bioengineering
Saccharomyces cerevisiae
Alphafold
Aromatic prenyltransferase

ABSTRACT

The present research provides an application for an aromatic prenyltransferase from *Glycine max* for use in heterologous microorganism expression to generate cannabinoids. The known cannabinoid prenyltransferase CsPT04 was queried in FoldSeek. An enzyme derived from *Glycine max* known as GLYMA_02G168000, which is a predicted homogentisate solanyltransferase, was identified and found to have affinity for the prenylation of geranyldiphosphate (GPP) and olivetolic acid (OA) to produce cannabigerolic acid (CBGA) and cannabigerol (CBG). The *in vitro* production of CBGA was accomplished through the heterologous expression of this prenyltransferase in *Saccharomyces cerevisiae*. After growing the yeast cells, a purified microsomal fraction was harvested, which was rich in the membrane-bound prenyltransferase GlyMa_02G168000. Addition of purified microsomal fraction to a reaction matrix facilitated the successful prenylation of externally supplied OA with GPP, culminating in the production of CBGA. Structural comparisons revealed a notably closer similarity between GLYMA_02G168000 and CsPT04, compared to the similarity of other cannabinoid prenyltransferases with CsPT04. Herein, a novel application for a homogentisate solanyltransferase has been established towards the production of cannabinoids.

1. Introduction

Cannabinoids are the principal secondary metabolite of *Cannabis sativa*, and are defined as a molecule in which an isoprenyl moiety oriented towards a para-structure via a resorcinyl core [1]. While this occurs naturally in *C. sativa* and some other plant species, it can also be accomplished using heterologous cloning in microorganisms and plants [2–11]. This approach offers several advantages over plant cultivation, namely, speed of production, capacity of production, and precision of the target products. While cannabinoids are structurally diverse, all cannabinoids rely on the prenylation of an aromatic polyketide with an isoprenoid diphosphate via an aromatic prenyltransferase to form cannabigerolic acid (CBGA) or a structural homolog thereof (see Fig. 1a). CBGA then can be oxidized into other cannabinoids such as Δ^9-tetrahydrocannabinolic acid (THCA), cannabidiolic acid (CBDA), or cannabichromenic acid (CBCA). Once oxidized and rearranged enzymatically and decarboxylated into neutral cannabinoids, additional synthetic approaches can be pursued to derive the rare cannabinoids cannabinol (CBN), *iso*-tetrahydrocannabinol (*iso*-Δ^9-THC), cannabitriol (CBT) and cannabicyclol (CBL) from cannabidiol (CBD) and cannabichromene (CBC), as has been previously described [12]. CBGA and other cannabinoids are known to have significant commercial value for their potent pharmaceutical properties as well as value in the food and beverage markets for nutraceutical potential, and are also a valuable commodity in the cosmetics market. As CBGA is the central precursor for all cannabinoids, an eco-friendly, fast, and inexpensive system for its establishment proves useful in the context of a rapid expansion of cannabinoid availability through legal procurement.

In *C. sativa* cannabinoid biosynthesis, CsPT04 is responsible for CBGA production [3,6]. Former studies have demonstrated cannabinoid biosynthesis in *S. cerevisiae* by engineering the olivetolic acid and GPP pathways to higher yields, by implementing CsPT04 forming CBGA and further cannabinoid synthases to end up with the pharmacological interesting compounds like THCA. Other plant prenyltransferases have also been applied towards microbial cell factories, such as CsPT01 and HuCBGAS4. In addition, NphB, a soluble bacterial prenyltransferase from a *Streptomyces* species was found to catalyze the first step in cannabinoid biosynthesis [13]. In the present work, a new CBGA producing enzyme was identified, originated from *Glycine max,* commonly known as soy.

* Corresponding author.
E-mail address: oliver.kayser@tu-dortmund.de (O. Kayser).

https://doi.org/10.1016/j.bbrc.2024.149471
Received 21 December 2023; Accepted 3 January 2024
Available online 4 January 2024

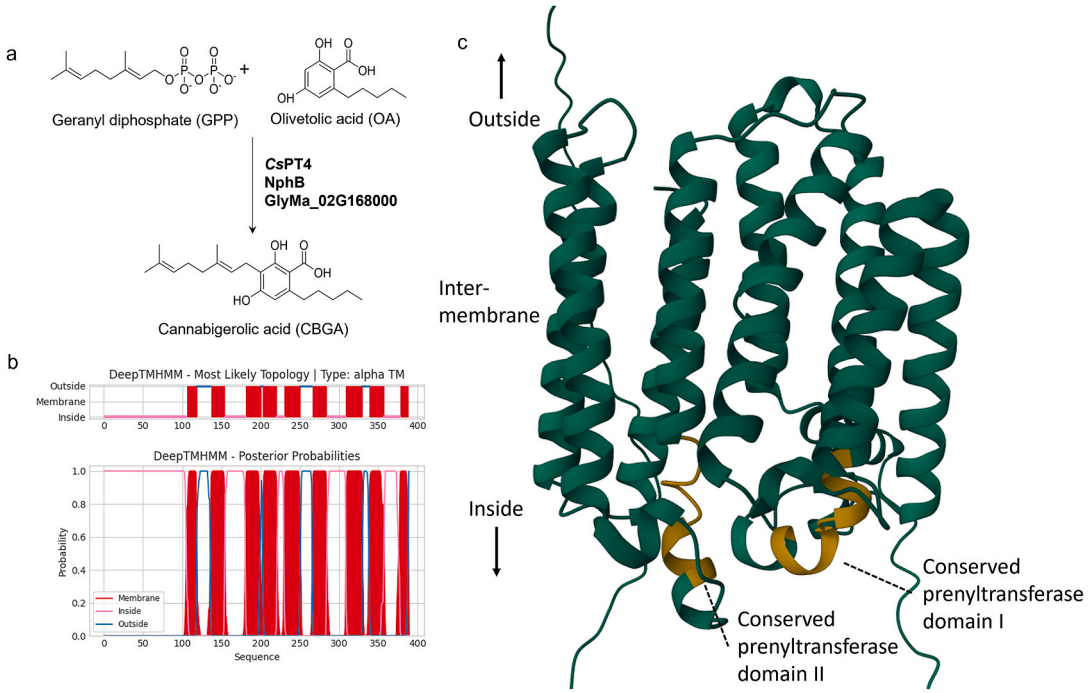

Fig. 1. Panel a: Generation of cannabigerolic acid (CBGA) using either CsPT04, NphB or GlyMa_02G168000 from geranyl pyrophosphate (GPP) and olivetolic acid (OA). Panel b: DeepTMHMM analysis of GlyMa_02G168000 for membrane associated residues. Panel c: Structural analysis of GlyMa_02G168000 (cropped) using AlphaFold and DeepTMHMM. The protein was also found to contain the conserved protein prenyltransferase domain I, NQIYDISID, and domain II, NQIYDISID, highlighted in yellow in the figure, left. These sequences correspond to the inner membrane domains of the beta barrels of the prenyltransferases and are adjacent to one another.

In this study, we identify an application for the prenyltransferase GlyMa_02G168000 to produce cannabinoids *in vitro* after heterologous expression in *S. cerevisiae*. We identified this enzyme by querying the CsPT04 structure using Foldseek [14], a recent algorithm that can identify structural homologs of proteins and enzymes, operating independently of sequence identity. We then analyze and characterize the functionality of this enzyme via an *in vitro* approach, demonstrating that it prenylates GPP and OA. These results suggest a promising new alternative to establish cannabinoid yeast-cell factories which are outside of the scope of intellectually-owned enzymatic reactions, and through further research may shed light on substrate specificity and reaction promiscuity in plant aromatic prenyltransferases.

2. Materials and methods

2.1. Identification of GLYMA_02G168000 and comparative analysis to other known prenyltransferases

The predicted model of the *Cannabis sativa CsPT04* was collected from the Alphafold database and was queried using Foldseek. The protein of interest, *GLYMA_02G168000,* was the most similar protein from all results. To confirm that the enzyme of interest was a homogentisate prenyltransferase, we analyzed the sequence in DeepTMHMM [15] and examined the locations of the conserved active site characteristic residues for this family of enzymes.

After identification, *Glycine max* enzyme *GLYMA_02G168000* and other known prenyltransferases were selected for comparative phylogenetic analysis. We constructed a protein sequence alignment of several aromatic prenyltransferases of interest using a MUSCLE PPP alignment [16] and built a phylogenetic tree using neighbor-joining with bootstrapping 1000 times. The *S. cerevisiae* protein GAL4 (Uniprot P04386 · GAL4_YEAST) was used for the outgroup. CsPT04 (Uniprot A0A455ZJC3 · GOT_CANSA), CsPT01 (Uniprot A0A455ZIK6 · A0A455ZIK6_CANSA), and HuCBGAS4 (GenBank: WFP21564.1) were included because they are known to produce cannabigerolic acid *in planta* and from microbial transformations. Prenyltransferases from the species *Humulus lupulus* (Uniprot A0A0B5A051 · PT1L_HUMLU, A0A0B4ZTQ2 · PT2_HUMLU, E5RP65 · PT1_HUMLU) were included because of its close relationship to *Cannabis sativa,* however these enzymes are not known to produce cannabigerolic acid either on their own or in microbial cells. *Artemisia capillaris* is a well-studied plant known for production of anti-malarial natural compounds through prenylation enzymes [17], one of which (Uniprot A0A5S9BNB0 · A0A5S9BNB0_9ASTR) has been included. Last, the enzyme GlyMa_02G168000 (Uniprot I1JFW4 · I1JFW4_SOYBN) has been previously identified to be a homogentisate solanyltransferase [18], and we have therefore included other known plant homogentisate solanyltransferases such as from *Chlamydomonas reinhardtii* (Uniprot A1JHN0 · HSTC_CHLRE), *Arabidopsis thaliana* (Uniprot Q8VWJ1 · HPT1_ARATH), and other enzymes known to carry our plastoquinone biosynthesis in *Glycine max*. For a multiple protein structural alignment and analysis, we used mTM-align [19] followed by PHYLIP to construct a neighbor-joining tree. All enzymes were the same as in the sequential alignment, except that *S. cerevisiae* GAL10 (Uniprot P04397 · GAL10_YEAST) was selected for the outgroup, because GAL4 was found to be a multimeric enzyme, which is treated as multiple proteins by mTM-align. We excluded one *Glycine max* prenyltransferase from this tree because no PDB was available.

2.2. Expression of GlyMa_02G168000 in yeast cells and microsomal purification

The *S. cerevisiae* codon-optimized ORF and translated amino acid sequence of the enzyme are given in the supplementary information (Page 4). The selected sequence was codon optimized and ordered from Thermo Fisher GeneArt and primers with corresponding overlaps to the expression vector pDionysos (pDIO) [20] were designed. This plasmid was then amplified in *E. coli* DH5α cells. The resulting construct (pDio-GlyMPT) were verified by Sanger sequencing and used to

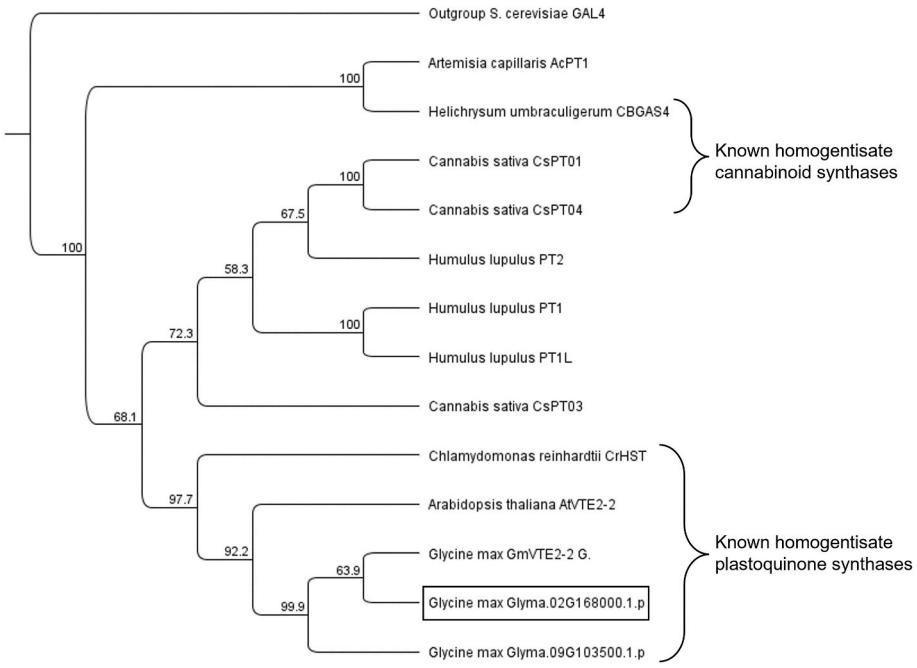

Fig. 2. Phylogenetic analysis of known homogentisate cannabinoid synthases, other plant prenyltransferases, and known homogentisate plastoquinone synthases from *Glycine max* and other species, including GlyMa_02G168000 (shown in the rectangle).

transform *S. cerevisiae* CEN. PK2-1 C (Δ*pep4* Δ*gal1* Δ*gal80 720a::HAC1s*) (obtained from Euroscarf) by the lithium acetate/single stranded carrier DNA/PEG method [21]. Transformed yeast cells were precultured in 10 ml leucine drop-out medium 24 h at 30 °C and 200 rpm. A second preculture was inoculated in 25 ml leucine drop-out medium at OD_{600} 0.25 and cultivated for 12 h at 30 °C and 200 rpm. The expression cultures were inoculated in 200 ml potassium phosphate (100 mM) buffered YPD medium (10 g/l yeast extract, 20 g/l peptone from meat, 2 % glucose) pH 6.7 at OD_{600} 0.5 and cultivated for 48 h at 30 °C and 200 rpm. All cultures were cultivated in Erlenmeyer flasks with three baffles.

For microsomal fraction purification yeast cells were harvested by centrifugation at 3.000×*g* for 5 min at 4 °C. Cell pellets were then resuspended in 40 ml extraction buffer (50 mM Tris HCl pH 7.4, 1 mM EDTA, 0.1 M KCl, 1 mM PMSF) and lysed 3 times by French press® at a pressure of 7 MPa. To remove cell debris the lysates were centrifuged at 10.000×*g* 10 min at 4 °C. The supernatant was centrifuged again at 100,000×*g* for 1 h at 4 °C to collect the microsomal fraction. The remaining pellets were resuspended in 1.5 ml storage buffer (50 mM Tris HCl pH 8.0, 10 mM $MgCl_2$, 10 % glycerol) with a tissue grinder and stored at −20 °C.

2.3. Biochemical activity of GlyMa_02G168000 for CBGA production

The prenyltransferase was tested for its ability to form CBGA by using the corresponding microsomal fraction from yeast cultivations. The reaction buffer (50 mM Tris HCl pH 8.5, 10 mM $MgCl_2$), 10 % microsomal fraction, 2 mM GPP and 5 mM olivetolic acid with a final volume of 100 µl were incubated at 30 °C, 1100 rpm for 24 h. The reactions were terminated, and products extracted three times by supplementation of 300 µl ethyl acetate and 5 min vortexing. After evaporating the organic phase the remaining residues were dissolved in 100 µl acetonitrile:H_2O:formic acid:ethanol (77:19:3:1) mixture and analyzed by UPLC/MS.

2.4. UPLC/MS analysis

For analytical verification an Agilent 1290 Infinity II liquid chromatography system (with an Agilent EC-C18 100 mm × 2.1 mm x 1.9 µm column and an Agilent EC-C18 5 mm × 2.1 mm x 1.9 mm guard column) were used on a column temperature of 50 °C. Each sample was measured in negative mode on a Bruker Compact qTOF by injecting 5 µl sample volume. The solvent composition for separating compounds were conducted with Line A methanol (supplemented with 0.01 % formic acid) and line B water (supplemented with 5 mM ammonium formate and 0.1 % formic acid) with a flow rate of 0.5 ml/min. Starting with line A at 65 % for 2 min, the gradient was raised to 77 % An until 14 min and to 95 % An until 16.4 min. Afterward the method was equilibrated back to 65 % An until 20 min. Samples were acquired in high resolution MS (HRMS) mode and in multiple reaction monitoring (MRM) mode for MS/MS analysis, with selection of ions at 359.22 *m/z* and 315.22, with a collision energy of −36 eV. CBGA production was verified by retention time and mass comparison to a certified CBGA standard (Cerilliant®, Sigma Aldrich) (rt 10.6 min; *m/z* 359.2199).

3. Results

3.1. Identification of GLYMA_02G168000

The Alphafold [22] structure of Geranylpyrophosphate:olivetolate geranyltransferase, chloroplastic, also known as CsPT04, was collected from the Uniprot database and searched against the Alphafold and PDB databases using FoldSeek. A protein structure of a homogentisate solanyltransferase from *Glycine max* was identified with probable structural homology between amino acid residues 15 and 398. The overlay is shown in the supplementary information, Fig. 1. These structural alignments had a TM-score of 0.75648 and an RMSD of 23.03. The enzyme was then determined to have originated from its gene on chromosome 2 of the Williams 82 soy cultivar [23]. It was then determined that the splice variant KRH71769 was the source of the predicted protein structure with high structural homology to CsPT04. The sequence identity to CsPT04 was found to be 36.9 %. The protein was also found to contain the conserved protein prenyltransferase domain I, NQXXDXXXD, and domain II, KDXXDXXGD, and an analysis with DeepTMHMM identified that these residues were both present along the intercellular space in between intermembrane β-barrels (see Fig. 1b and c). This would allow the prenyltransferase to coordinate the phosphates

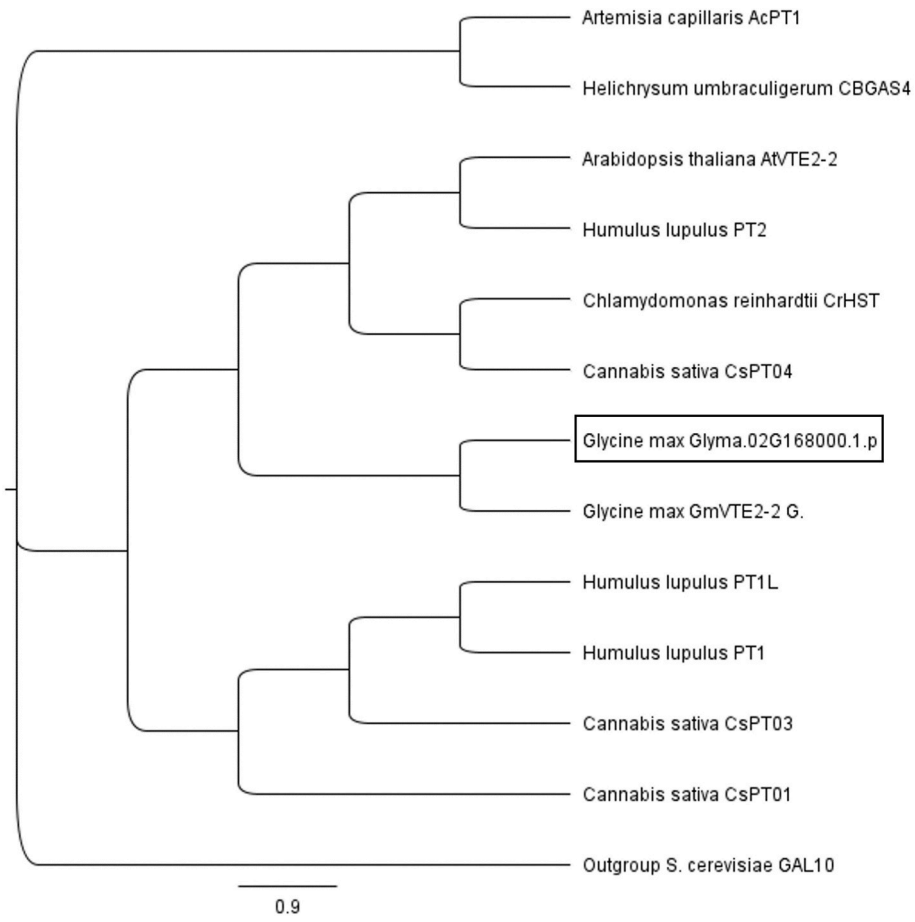

Fig. 3. Phylogenetic analysis of known homogentisate cannabinoid synthases, other plant prenyltransferases, and known homogentisate plastoquinone synthases from *Glycine max* and other species, including GlyMa_02G168000 (shown in the rectangle). Enzymes and corresponding Alpha-fold PDB's used: Artemesia capillaris AcPT1 (AF-A0A5S9BNB0–F1-model_v4), Helichrysum umbraculigerum CBGAS4 (self-produced Alphafold structure), *Arabidopsis thaliana* AtVTE2-2 (AF-Q8VWJ1-F1-model_v4), Humulus lupulus PT2 (AF-A0A0B4ZTQ2-F1-model_v4), Chlamydomonas reinhardtii CrHST (AF-A1JHN0-F1-model_v4), Cannabis sativa CsPT04 (AF-A0A455ZJC3-F1-model_v4), Glycine max GlyMa_02G168000 (AF-I1JFW4-F1-model_v4), Glycine max VTE2-2 G. (AF-Q1ACB2-F1-model_v4), Hulumus lupulus PT1L (AF-A0A0B5A051-F1-model_v4), Humulus lupulus PT1 (AF-E5RP65-F1-model_v4), Cannabis sativa CsPT03 (AF-A0A455ZLY7-F1-model_v4), Cannabis sativa PT01 (AF-A0A455ZIK6-F1-model_v4), Saccharomyces cerevisae GAL10 (AF-P04397-F1-model_v4).

from GPP to the OA substrate and Mg^{2+} to catalyze the prenylation.

3.2. Similarities to known C. sativa prenyltransferases 1 and 4 and Helichrysum umbraculigerum cannabigerolic acid synthase 4

The DNA and amino acid sequences of the *GLYMA_02G168000* KRH71769 splice variant were compared to *CSPT04*, *CSPT01* (both from *Cannabis sativa*) and *CBGAS4* from *Helichrysum umbraculigerum*. The DNA sequences encoding the newly identified polypeptide is remarkably different to former described gene sequences encoding olivetolic acid prenylating enzymes CsPT01 and CsPT04, with homologies of 55.5 % for *CSPT04* vs. *GLYMA_02G168000*, 55.3 % for *CSPT01* vs. *GLYMA_02G168000*, and 45.7 % for *CBGAS4* vs. *GLYMA_02G168000* at the sequence level of the CDS's. The *Glycine max* prenyltransferase shares homologies of 53.3 % for *CSPT04* vs. *GLYMA_02G168000*, 54.8 % for *CSPT01* vs. *GLYMA_02G168000*, and 44.8 % for *CBGAS04* vs. *GLYMA_02G168000* at the sequence level of the codon optimized CDS's. Finally, it shares homologies of 36.1 % for CsPT04 vs. GlyMa_02G168000, 38.3 % for CsPT01 vs. GlyMa_02G168000, and 27.6 % for CBGAS4 vs. GlyMa_02G168000 at the amino acid sequence level.

To further interrogate the relationships between GlyMa_02G168000 and known cannabinoid synthases and/or other prenyltransferases, we constructed a phylogenetic tree based on the sequence-based alignments. To contextualize the relationship between the prenyltransferases we expanded the selected prenyltransferases, as described in the methods (Fig. 2). We find that GlyMa_02G168000 is most closely related to other *Glycine max* enzymes at the sequence-level, and then to enzymes involved in plastoquinone-9 biosynthesis from other species (see supplementary material, Fig. 4 for pathway diagram). As expected, enzymes from *C. sativa* were most closely related to prenyltransferases from *Humulus lupulus,* a close relative of *C. sativa.* HuCBGAS4 was most closely related to AcPT1, a prenyltransferase from the plant *Artemesia capillaris*, known for generating prenylated natural products that are effective in the treatment of malaria.

Next, we examined the relationship between these prenyltransferases using structural alignment in mTM-align and built a structure-based neighbor-joining phylogenetic tree, shown in Fig. 4. The relationships between the prenyltransferases differed substantially between the sequence-based and the structure-based approaches. Whereas in sequence-based alignment, the enzymes cladded based on their characterized physiologies (except for HuCBGAS4) and evolutionary relationships, structurally we observed that GlyMa_02G168000 and CsPT04 were more similar than CsPT04 and other cannabinoid prenyltransferases. A snapshot of the protein alignment is shown in the supplementary material, Fig. 5.

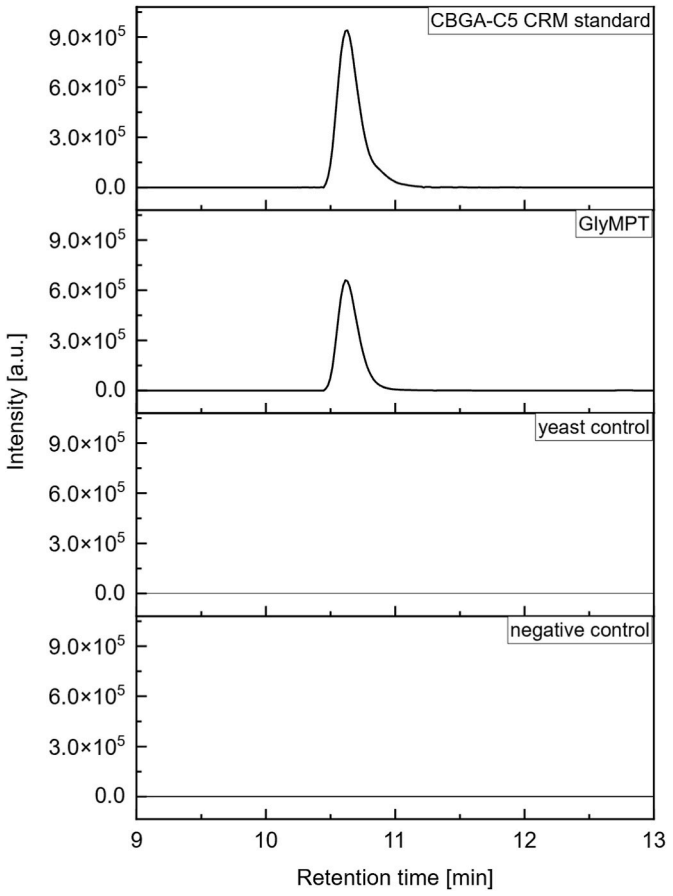

Fig. 4. UPLC/MS extracted ion chromatogram (EIC) negative selected ion monitoring theoretical *m/z* value of (a) certified CBGA standard (Cerilliant®, Sigma Aldrich). (rt 10.6 min, *m/z* 359.2199); (b) EIC at *m/z* 359.2199 of extracted material of purified microsomal fractions from *GLYMA_02G168000* transformed *S. cerevisiae* CENPK2-1C *(Δpep4 Δgal1 Δgal80 HAC1s)* incubated with 2 mM GPP and 5 mM OA for 24 h in reaction buffer; (c) EIC at *m/z* 359.2199 of extracted material of purified microsomal fractions from non-*GLYMA_02G168000* transformed *S. cerevisiae* CENPK2-1C *(Δpep4 Δgal1 Δgal80 HAC1s)* incubated with 2 mM GPP and 5 mM OA for 24 h in reaction buffer; (d) EIC at *m/z* 359.2199 of extracted material of 2 mM GPP and 5 mM OA for 24 h in reaction buffer.

3.3. Confirmation of CBGA production by GlyMa_02G168000

The function of this enzyme was confirmed with UPLC-HRMS and MS/MS analysis, where at 10.6 min the fragmentation pattern of the spectra matched to the fragmentation pattern of the Cerriliant CBGA standard (supplementary material, Fig. 2). Further, CBG was identified as a byproduct of the reaction or its preparation, as shown in the supplementary data, Fig. 3. GlyMa_02G168000 expressed in *S. cerevisiae* catalyzed CBGA production by external supplementation of the substrates to the purified microsomal fraction, while substrate supplementation to microsomal fractions of yeast without expression of GlyMa_02G168000 did not lead to detection of CBGA formation.

4. Discussion

Deep Mind's Alphafold algorithm, alongside dramatic improvements in the field of genomics, has created an unprecedented amount of enzyme structural data. Foldseek has enabled users to query PDB's created by Alphafold and identify structural homologs for any enzyme or protein of interest. These technological advances have made genome mining for applications in the field of biotechnology more accessible than ever before and have eased the burden of restrictions due to sequence-based intellectual property protection. In this short report, we have identified a structural homolog of CsPT04 and determined that it demonstrates promiscuity towards cannabinoid prenylation. Recent advances in protein-structure prediction, coupled with an efficient algorithm for querying these structures, have enabled broad, non-targeted, and rapid data mining strategies which were not previously available.

We have identified that structural homology to CsPT04 did not correlate to relatedness based on sequence identity, and have determined that structurally, GlyMa_02G168000 and CsPT04 are more similar than even CsPT04 is to CsPT01, a known cannabinoid prenyltransferase from *Cannabis sativa*. Furthermore, it is also closer in structure to CsPT04 than HuCBGAS4, another characterized cannabinoid prenyltransferases. We hypothesize that the structural similarity of GlyMa_02G168000 to CsPT04 enabled promiscuity to accept the target substrates of GPP and OA. These results warrant further investigation into the promiscuity of aromatic prenyltransferases and the mechanistic level and their application in microbial organisms for cannabinoid production. However, the results suggest that other prenyltransferases could potentially be more suitable to cannabinoid prenylation, such as CrHST, HuPT2, and AtVTE2-2. CrHST was CsPT04's closest sibling in the structure-based phylogenetic tree (Fig. 3). Nevertheless, in this paper we have limited the scope based on the Foldseek results that led us to identification of CsPT04 structural homologs to begin with. We hope that further enzyme characterization of homogentisate prenyltransferases will reveal insights about how enzyme structures and ligand binding influence the activity of this enzyme family.

In this study, we have identified a candidate enzyme that can be employed in yeast cell factories that is not currently held in patent or restricted for use based on existing intellectual property. This result is significant within the field of cannabinoid yeast cell factory research because of the competitive nature of the research application and the drive to monetize biotechnological processes towards the production of cannabinoids. We hope that this will aid bioengineers of cannabinoid yeast cell factories to either directly use *GlyMa_02G168000* in their cell strains, or to use the pipeline we have established in this paper to identify other structural homologs for their own strain development. Finally, we hope that bioengineers of other biosynthetic processes can borrow from these methods to apply to their own work flows.

CRediT authorship contribution statement

Erin Noel Jordan: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Christina Schmidt:** Investigation, Methodology, Writing – review & editing. **Oliver Kayser:** Funding acquisition, Project administration, Supervision.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

We acknowledge the financial support from the Ministry of Economic Affairs, Industry, Climate Protection and Energy of the State of North Rhine-Westphalia in the program ZukunftBIO.NRW for the project CannaSyn (BI-1-09B).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrc.2024.149471.

References

[1] L.O. Hanuś, S.M. Meyer, E. Muñoz, O. Taglialatela-Scafati, G. Appendino, Phytocannabinoids: a unified critical inventory, Nat. Prod. Rep. 33 (12) (2016) 1357–1392.

[2] S. Dusséaux, W.T. Wajn, Y. Liu, C. Ignea, S.C. Kampranis, Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids, Proc. Natl. Acad. Sci. U.S.A. 117 (50) (2020 Dec 15) 31789–31799.

[3] M. Geissler, J. Volk, F. Stehle, O. Kayser, H. Warzecha, Subcellular localization defines modification and production of Δ9-tetrahydrocannabinolic acid synthase in transiently transformed Nicotiana benthamiana, Biotechnol. Lett. 40 (6) (2018 Jun) 981–987.

[4] T. Gülck, J.K. Booth, Â. Carvalho, B. Khakimov, C. Crocoll, M.S. Motawia, et al., Synthetic biology of cannabinoids and cannabinoid glucosides in *Nicotiana benthamiana* and *Saccharomyces cerevisiae*, J. Nat. Prod. 83 (10) (2020 Oct 23) 2877–2893.

[5] Keasling JD, D'Espaux L, Wang J, Luo X, Reiter M, Denby C, et al. MICROORGANISMS AND METHODS FOR PRODUCING CANNABINOIDS AND CANNABINOID DERIVATIVES. EP 3998336 A1..

[6] X. Luo, M.A. Reiter, L. d'Espaux, J. Wong, C.M. Denby, A. Lechner, et al., Complete biosynthesis of cannabinoids and their unnatural analogues in yeast, Nature 567 (7746) (2019 Mar) 123–126.

[7] Y. Zhang, J. Guo, P. Gao, W. Yan, J. Shen, X. Luo, et al., Development of an efficient yeast platform for cannabigerolic acid biosynthesis, Metab. Eng. 80 (2023 Nov) 232–240.

[8] J.E. Page, Z. Boubakir, Aromatic Prenyltransferase from Cannabis, WO, 2011 017798. A1.

[9] P. Berman, L.A. De Haro, A. Jozwiak, S. Panda, Z. Pinkas, Y. Dong, et al., Parallel evolution of cannabinoid biosynthesis, Nat. Plants 9 (5) (2023 May 1) 817–831.

[10] B. Zirpel, F. Degenhardt, C. Martin, O. Kayser, F. Stehle, Engineering yeasts as platform organisms for cannabinoid biosynthesis, J. Biotechnol. 259 (2017 Oct) 204–212.

[11] B. Zirpel, F. Degenhardt, C. Zammarelli, D. Wibberg, J. Kalinowski, F. Stehle, et al., Optimization of Δ 9 -tetrahydrocannabinolic acid synthase production in Komagataella phaffii via post-translational bottleneck identification, J. Biotechnol. 272–273 (2018 Apr) 40–47.

[12] G.N. Nguyen, E.N. Jordan, O. Kayser, Synthetic strategies for rare cannabinoids derived from *Cannabis sativa*, J. Nat. Prod. 85 (6) (2022 Jun 24) 1555–1568.

[13] Â. Carvalho, E.H. Hansen, O. Kayser, S. Carlsen, F. Stehle, Designing microorganisms for heterologous biosynthesis of cannabinoids [Internet], FEMS Yeast Res. 17 (4) (2017 Jun 1). cited 2020 Sep. 25, https://academic.oup.com/femsyr/article/doi/10.1093/femsyr/fox037/3861260.

[14] M. Van Kempen, S.S. Kim, C. Tumescheit, M. Mirdita, J. Lee, C.L.M. Gilchrist, et al., Fast and accurate protein structure search with Foldseek [Internet], Nat. Biotechnol. (2023 May 8) [cited 2023 Jul 21], https://www.nature.com/articles/s41587-023-01773-0.

[15] J. Hallgren, K.D. Tsirigos, M.D. Pedersen, J.J. Almagro Armenteros, P. Marcatili, H. Nielsen, et al., DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks [Internet], Bioinformatics (2022 Apr) [cited 2023 Jul 21], http://biorxiv.org/lookup/doi/10.1101/2022.04.08.487609.

[16] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (5) (2004 Mar 8) 1792–1797.

[17] R. Munakata, T. Takemura, K. Tatsumi, E. Moriyoshi, K. Yanagihara, A. Sugiyama, et al., Isolation of Artemisia capillaris membrane-bound di-prenyltransferase for phenylpropanoids and redesign of artepillin C in yeast, Commun. Biol. 2 (1) (2019 Oct 18) 384.

[18] A. Sukumaran, T. McDowell, L. Chen, J. Renaud, S. Dhaubhadel, Isoflavonoid-specific *prenyltransferase* gene family in soybean: GmPT01, a pterocarpan 2-dimethylallyltransferase involved in glyceollin biosynthesis, Plant J. 96 (5) (2018 Dec) 966–981.

[19] R. Dong, Z. Peng, Y. Zhang, J. Yang, mTM-align: an algorithm for fast and accurate multiple protein structure alignment. Valencia A, editor, Bioinformatics 34 (10) (2018 May 15) 1719–1725.

[20] F. Stehle, M.T. Stubbs, D. Strack, C. Milkowski, Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism: heterologous expression and kinetic mechanism of AtSMT, FEBS J. 275 (4) (2008 Feb) 775–787.

[21] R.D. Gietz, R.H. Schiestl, High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method, Nat. Protoc. 2 (1) (2007 Jan) 31–34.

[22] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021 Aug 26) 583–589.

[23] J. Schmutz, S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, et al., Genome sequence of the palaeopolyploid soybean, Nature 463 (7278) (2010 Jan) 178–183.