Integrated Platform for Structural and Functional Analysis of Cannabis sativa Terpene Synthases

Danielle Wiles , James Roest , Bhuvana Shanbhag , Julian Vivian , Travis Beddoe *

Posted Date: 29 April 2025

doi: 10.20944/preprints202504.2364.v1

Keywords: Cannabis; terpene synthase; structure; recombinant expression; crystallisation

 doi:10.20944/preprints202504.2364.v1

Article

Integrated Platform for Structural and Functional Analysis of *Cannabis sativa* Terpene Synthases

Danielle Wiles [1,2], James Roest [3], Bhuvana Shanbhag [1,2], Julian Vivian [3,4,5] and Travis Beddoe [1,2,*]

[1] Research Hub for Medicinal Agriculture, La Trobe University, Bundoora, VIC, Australia; d.wiles@latrobe.edu.au (D.W.); bkshanbhag@mum.amity.edu (B.S.)
[2] Department of Animal, Plant and Soil Science, La Trobe University, Bundoora, VIC, Australia
[3] St. Vincents Institute Medical Research, Fitzroy, VIC, Australia; jroest@svi.edu.au (J.R.); jvivian@svi.edu.au (J.V.)
[4] Australian Catholic University, Fitzroy, Victoria, Australia
[5] Department of Medicine, The University of Melbourne, Melbourne, VIC, Australia
* Correspondence: t.beddoe@latrobe.edu.au

Abstract: Terpenoids are the largest and most diverse family of natural products. Volatile terpenes from *Cannabis sativa* are crucial in flavours, fragrances, and pharmaceuticals due to their unique odours and biological activities, including antimalarial, antibacterial, and insecticidal properties. Their synthesis is catalysed by terpene synthase (TPS) enzymes, which perform cyclisation and rearrangement reactions of over 55,000 distinct terpene compounds. However, low catalytic efficiency of *C. sativa* TPSs limits their use in large-scale commercial production. The complex biochemistry of these reactions is not well understood due to limited enzyme structure information. To address this, we developed a platform for recombinant expression, purification, enzymatic characterisation, and crystallisation of TPS enzymes from *C. sativa*. This includes comprehensive kinetic, thermostability, structural, and enzymatic analyses, along with a novel directed crystallisation screen informed by data mining existing conditions, facilitating diffraction-quality TPS crystals. Each TPS enzyme showed a distinct product profile, highlighting the need for systematic characterisation of *C. sativa* terpene biosynthesis. Our findings provide a framework for the structural and functional study of *C. sativa* TPSs. The developed platform sets the stage for future metabolic engineering aimed at optimising terpene production for pharmaceutical, pest management, and synthetic biology applications.

Keywords: Cannabis; terpene synthase; structure; recombinant expression; crystallisation

1. Introduction

Cannabis sativa (cannabis) has been utilised for thousands of years as a source of fibre, food, and oil, as well as a medicinal agent and a recreational intoxicant (Bonini et al., 2018; Hui-Lin, 1974; Kalant, 2001; Zuardi, 2006). Today, it is highly valued for its pharmacologically active specialised metabolites, including cannabinoids, monoterpenes, and sesquiterpenes. These bioactive compounds are predominantly concentrated in the resin of glandular trichomes found in female cannabis inflorescences (Andre et al., 2016). While research has focused on the bioactivity of cannabinoids, the terpenes are increasingly being examined for their bioactive properties and commercial value. Cannabis is a prolific producer of terpenoids, with over 230 distinct compounds identified across various tissues (Aizpurua-Olaizola et al., 2016; Downer, 2020; Hanuš and Hod, 2020; Roell, 2020).

Terpenes constitute the largest class of plant specialised metabolites, playing a crucial role in plants' aroma and flavour profiles (Croteau, 1998; Karunanithi and Zerbe, 2019; Theis and Lerdau, 2003). Terpenoids have a variety of commercial applications as therapeutics, cosmetics, flavouring agents, fragrances, agrochemicals, and disinfectants (Ajikumar et al. 2008; (Cox-Georgian et al., 2019; Masyita et al., 2022; Nuutinen, 2018; Paduch et al., 2007). In cannabis, the plant's characteristic scent and flavour are derived from the unique combination of structurally diverse terpenoids, influencing consumer preferences, particularly among recreational and medicinal cannabis users (Oswald et al., 2023). Additionally, there is growing evidence that terpenes elicit various medicinal properties that

 doi:10.20944/preprints202504.2364.v1

affect both humans and animals (Russo, 2011), prompting efforts to breed cannabis cultivars with specific terpene profiles (Barcaccia et al., 2020; Chandra et al., 2017; Grof, 2018; Rocha et al., 2020).

Terpenes are synthesised from the 5-carbon precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are produced via the mevalonate (MVA) pathway in the cytosol or the methylerythritol phosphate (MEP) pathway in plastids. These precursors are further condensed to form prenyl diphosphate intermediates, such as geranyl diphosphate (GPP) and farnesyl diphosphate (FPP), which are then converted into monoterpenes (C_{10}) and sesquiterpenes (C_{15}), respectively by terpene synthase (TPS) enzymes (Christianson, 2017; Lange et al., 2000). The remarkable structural and chemical diversity exhibited by terpenes is due to the vast array of reactions catalysed by TPSs, which can generate multiple products from a single precursor through cyclisation, rearrangement, and hydride shifts (Karunanithi and Zerbe, 2019). This diversity is further enhanced by the enzyme's active site architecture, which stabilises different carbocation intermediates, making TPSs essential for the biosynthesis of the complex and varied terpene landscape observed in nature (Christianson, 2017).

TPS enzymes are encoded by large gene families, typically comprising 20-150 genes in most plant species, enabling them to exhibit remarkable variability in both substrate preference and product specificity (Tholl., 2006; Zhou et al., 2020). This unique versatility allows them to generate multiple terpene products from a single substrate or utilise various substrates to produce a variety of different terpenes (Li and Tao., 2024; Kampranis et al., 2007). Such evolutionary plasticity means that even single amino acid changes within the active site can significantly alter the enzyme's product profile (Bohlmann and Gershenzon, 2009; Schilmiller et al., 2009; Zhou and Pichersky, 2020). However, given the complex evolutionary history of TPS, the product profile of a given enzyme cannot reliably be predicted based on sequence similarity alone and, therefore, requires functional characterisation of individual TPS enzymes to determine the specificities of their catalysis.

In cannabis, at least 55 TPS genes have been identified in the genome, reflecting the extensive diversity within this gene family (Allen et al., 2019; Booth et al., 2020, 2017; Xu et al., 2024). The overall gene structure of individual TPS genes is remarkably well conserved across cannabis TPS subfamilies; however, gene length can be quite variable (Allen et al., 2019; Xu et al., 2024). Despite the functional characterisation of several cannabis TPS genes (Booth et al., 2020, 2017; Günnewich et al., 2007; Livingston et al., 2020; Zager et al., 2019), the inability to precisely link specific TPS enzymes to their terpene products poses a major challenge. This limitation hinders efforts to generate targeted or novel terpene profiles and to optimise their production for industrial applications. Furthermore, the lack of comprehensive kinetic, thermostability, and structural analyses prevents a deeper understanding of TPS functionality and restricts their potential for metabolic engineering.

The ability to engineer TPSs for specific terpene production offers a transformative approach to overcoming the limitations of traditional terpene extraction methods, which are constrained by strict regulatory frameworks, inconsistent metabolite profiles due to environmental variability, and the inherent complexity of cannabis-specialised metabolites. This study directly addresses these challenges by leveraging heterologous expression systems to achieve consistent and scalable production of high-quality terpenes. Through the structural and functional characterisation of 10 cannabis TPS enzymes, we identify key determinants of enzyme specificity and catalysis, providing insights into their mechanisms of action. Furthermore, our research lays the groundwork for future engineering of TPS enzymes to enhance their catalytic efficiency and product specificity, thereby paving the way for optimised terpene production at industrial scales. By integrating detailed kinetic and thermostability analyses, we also establish a foundation for improving the robustness of TPS enzymes in diverse production environments. This work not only deepens our understanding of terpene biosynthesis but also highlights the potential of synthetic biology to harness these enzymes for sustainable, large-scale applications in pharmaceuticals, agriculture, and beyond.

2. Materials & Methods

2.1. Cloning of CsTPS

Expression constructs encoding three full-length sesqui-TPS (CsTPS9FN, CsTPS16CC and CsTPS20CT), two mono/sesqui-TPS (CsTPS5FN and CsTPS19BL), and five mono-TPS sequences (CsTPS12PK, CsTPS13PK, CsTPS3FN, CsTPS1SK and CsTPS37FN), which had their plastidial-targeting sequence motif truncated (Table 1), were synthesised (Twist Bioscience) as codon-harmonised

 doi:10.20944/preprints202504.2364.v1

(Codon Wizard) and codon-optimised for *Escherichia coli*. The genes were cloned into a pET28a+ expression vector (Luna-Vargas et al., 2011), using *NdeI* and *XhoI* restriction sites.

Table 1. Summary of CsTPS enzymes used in study.

Functional gene ID*	GenBank ID	Proposed functionality	*C. sativa* cultivar	TPS type	DNA seq. length (bp)	Protein seq. length (AA)	Predicted MW (kDa)	Theoretical pI
CsTPS3FN	KY014561	β-myrcene synthase	Finola	Mono	1692*	584	68.90	6.02
CsTPS1SK	ABI21837	(-)-limonene synthase	Skunk	Mono	1641*	567	66.32	6.11
CsTPS37FN	KY014554	terpinolene synthase	Finola	Mono	1692*	584	68.68	5.77
CsTPS12PK / CsTPS33PK	KY624371	α-terpinene, γ-terpinene synthase	Purple Kush	Mono	1854	633	73.75	5.46
CsTPS13PK	KY014558	(*Z*)-β-ocimene synthase	Purple Kush	Mono	1803	616	72.27	6.11
CsTPS9FN	KY014555	β-caryophyllene / α-humulene synthase	Finola	Sesqui	1704	587	68.88	6.00
CsTPS16CC	MK131289	germacrene-B synthase	Cherry Chem	Sesqui	1716	591	69.55	6.43
CsTPS20CT	MK801762	hedycaryol synthase	Canna Tsu	Sesqui	1656	571	66.84	6.29
CsTPS5FN	KY014560	β-myrcene / (-)-α-pinene synthase	Finola	Mono/Sesqui	1722	593	69.43	6.04
CsTPS19BL	MK801763	nerolidol /linalool synthase	Black Lime	Mono/Sesqui	1665	574	66.55	6.30

* Plastidial targeting sequence truncated.

2.2. Protein Expression and Purification

Recombinant TPS enzymes were expressed in *E. coli* BL21(DE3) *T7 Express lysY/Iq* competent cells (New England Biolabs) using the NEB high-efficiency transformation protocol. Briefly, 25 µl aliquots of competent cells were mixed with 2 µl of plasmid DNA by heat shock and incubated in Luria-Bertani (LB) medium (1.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 1.0% (w/v) NaCl) at 37°C for 1 hour with shaking at 200 rpm. Transformed cells were plated on LB-agar containing 50 µg ml^{-1} kanamycin and incubated overnight at 37°C.

A single colony was used to inoculate a 10 mL starter culture in Terrific Broth (TB; 1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4) with 50 µg ml^{-1} kanamycin and grown overnight at 37°C with shaking at 200 rpm. The starter culture (1:50 dilution) was used to inoculate 400 mL TB medium supplemented with kanamycin (50 µg ml^{-1}). Cultures were grown at 37°C for 16 hours, induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubated at 16°C for an additional 16 hours. Cells were harvested by centrifugation at 10,330 × g for 30 minutes at 4°C.

Cell pellets were resuspended in 30 mL lysis buffer (20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole) and lysed by sonication (Ultrasonics) at 40% amplitude for 30-second bursts with 30-second rest intervals repeated three times, on ice. Lysates were clarified by centrifugation at 10,330 x *g* at 4°C for 20 mins. The process was repeated to ensure maximal removal of cell debris. The clarified lysate was syringe-filtered (0.22 µm), prior to purification.

Cleared lysates were loaded onto a gravity column containing 2 mL of 50% (v/v) of Ni-NTA resin slurry (Thermo Fisher Scientific), pre-equilibrated with wash buffer (20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 10 mM imidazole). The column was washed twice with 30 mL wash buffer, and TPS proteins were eluted with 5 mL elution buffer (50 mM Tris-HCl (pH 7.5), 500 mM NaCl and 500 mM

 doi:10.20944/preprints202504.2364.v1

imidazole) in 1 mL fractions. Each step of the purification process was analysed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) and Coomassie blue staining.

Ni-affinity eluted fractions containing the proteins of interest were pooled and further purified by size-exclusion chromatography (SEC) using a HiPrep 16/60 Superdex 200 column (GE Healthcare Life Sciences) on a ÄKTA Basic Fast Protein Liquid Chromatography (FPLC) system. The gel filtration buffer comprised 25 mM 4-(2-Hydroxyethyl)piperazine-1-ethane-sulfonic acid (HEPES; pH 7.0), 10 mM $MgCl_2$, 100 mM KCl and 1 mM Dithiothreitol (DTT) and all steps were performed at a flow rate of 1 mL min^{-1}. Peak fractions containing the CsTPS enzymes [>95% pure as judged by SDS–PAGE] were combined and concentrated to 10 mg mL^{-1} using Amicon centrifugal filters (10 kDa molecular weight cut off, Millipore). Protein concentrations were measured using Bradford's method (Bradford, 1976) and stored at 4°C.

2.3. Enzyme Activity Assay

Enzymatic activity assays were performed as described by (Allen et al., 2019) with minor modifications. TPS activities were assayed in triplicate in a final volume of 500 µL assay buffer (20 mM HEPES (pH 7.5), 100 mM KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 10% (v/v) glycerol), 100 µM geranyl diphosphate (GPP), geranyl-geranyl diphosphate (GGPP) or farnesyl diphosphate (FPP) (Merk Sigma Inc.) and purified protein. 500 µL of hexane (Sigma-Aldrich) containing 2.5 µM isobutylbenzene as an internal standard was overlaid to trap the volatile products. Reactions were incubated at 30°C for 16 hrs and vortexed for 30 seconds. Volatile products were extracted by centrifugation at 1,000 x *g* for 30 min at 4°C. Boiled enzyme controls were included to determine the background noise of the assay.

2.4. Product Identification by Gas-Chromatography Mass Spectrometry (GC-MS)

GC-MS analysis was conducted using a Thermo Scientific system equipped with a Trace 1310 GC interfaced with a triple quadrupole MS TSQ 8000 Evo and a TriPlus Robotic Sample Handling autosampler. The column was a Thermo Scientific TG-5SILMS capillary column (60 m x 0.25 mm ID, 1.0 µm film thickness), and carrier gas was He at a constant flow rate of 1.5 mL min^{-1}. The inlet temperature was 280 °C with a split ratio of 5:1, and the injection volume was 1 µL. The initial oven temperature was set at 50 °C with 5 min hold time, then increased to 300 °C at a rate of 5 °C min^{-1}, and held at 300 °C for 5 min. The MS was set in full scan mode with a mass range of 35–400 amu, delay time of 10.5 min, and ionisation by electron impact with ionisation energy of 70 eV. The ion source temperature was 230 °C, and the MS transfer line temperature was 280 °C. Gerstel Maestro software v1.5 controlled the autosampler, and data were acquired using Thermo Scientific XCalibur Software v4.7. Data was processed using Thermo Scientific XCalibur Qual Browser v4.7 or FreeStyle v1.8.

To confirm the presence of most terpenes, authentic standards were used during the analysis. However, the sesquiterpenes δ-elemene, β-elemene, γ-elemene, epi-β-caryophyllene, alloaromadendrene, elemol, germacrene-D, guaiol, globulol, γ-eudesmol, and α-eudesmol were identified tentatively through comparisons with the NIST library due to the lack of available internal standards (Supplementary Figure S4). Similarly, several monoterpenes, including β-phellandrene, allo-ocimene, fenchol, β-terpineol, pinene hydrate, and geranyl methyl ether, were also putatively identified via NIST library comparisons. While geranyl methyl ether and pinene hydrate were detected as chromatographic peaks and labelled accordingly, these compounds are neither monoterpenes nor sesquiterpenes—the primary focus of this study—and were therefore excluded from the final quantification percentages.

2.5. Malachite Green Assay for Kinetic Measurements

Kinetic parameters were determined using the malachite green assay (Vardakou et al., 2014). Reactions were performed in triplicate in 96-well flat-bottomed plates (Greiner bio-one). Standard curves were established for both monophosphate (Pi) and pyrophosphate (PPi) using serial 2-fold dilutions ranging from 0.01 µM to 50 µM. For kcat determination, reactions (50 µL) contained malachite green assay buffer (25 mM 2-(N-morpholino)ethanesulfonic acid (MES), 25 mM 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS), 50 mM Tris, 5 mM MgCl2, pH 7.5), 25 mU of inorganic pyrophosphatase from Saccharomyces cerevisiae (Sigma), 100 µM substrate (FPP or GPP), and serial 2-fold dilutions of protein (0.003-0.2 µM). Reactions were incubated at 37°C for 30 minutes prior to

 doi:10.20944/preprints202504.2364.v1

being terminated by the addition of the malachite green development solution (prepared according to (Pegan et al., 2009) and further incubated for 15 minutes before measuring absorbance at 623 nm using a plate reader (Bio-Rad). Steady-state kinetic measurements were also conducted in triplicate in a 50 µL reaction volume containing malachite green assay buffer, a fixed concentration of 0.014 µM of the protein of interest, and serial 2-fold dilutions of the pyrophosphate substrate (FPP or GPP) ranging from 0.02-100 µM. Reactions were also incubated at 37°C for 30 minutes before termination and absorbance measurement at 623 nm, following the same procedure described above The kinetic parameters (V_{max}, K_M and k_{cat}) were obtained from non-linear regression analysis of the data using the Michaelis–Menten model in Graphpad Prism.

2.6. Thermofluor Protein Stability Assay

For downstream applications, purified CsTPS enzymes were concentrated to 10 mg ml^{-1}. To optimise storage conditions an 80-condition buffer screen (Supplementary Table S1) was developed, based on commonly used buffers for TPS enzymes. The temperature at which CsTPS unfolding occurred for each of the conditions (T_m values) was compared with the original storage buffer, and changes in unfolding temperature (ΔT_m) were calculated. Thermal denaturation was monitored using SYPRO Orange dye (Sigma) and a real-time PCR machine and melting temperatures (T_m) were calculated as described previously (Ericsson et al., 2006). Briefly, solutions of 7.5 µL of 300 x SYPRO Orange (Merk Sigma), 12.5 µL of test buffer conditions (Supplementary Table S1), and 5 µL of 2.5 mg mL^{-1} CsTPS proteins were added to the wells of a 384-well thin wall PCR plate (Bio-Rad). Water was added instead of a buffer in the control samples. The plates were sealed with Optical-Quality Sealing Tape (Bio-Rad) and heated in an ABI QuantStudio 5Dx Real-Time qPCR System from 20 to 95°C in increments of 1°C. Fluorescence changes in the wells of the late were monitored simultaneously with a charge-coupled device (CCD) camera. The wavelengths for excitation and emission were 490 and 575nm, respectively.

2.7. Creation of the TPS-Crystallisation Screen

A 48-condition screen, with conditions specifically directed towards the crystallisation of TPSs was developed. To create this screen, the Protein Data Bank (PDB; Berman et al., 2002), was searched to identify the crystallisation conditions of TPSs and related proteins. Crystallisation conditions for 120 deposited structures of proteins were returned. Individual searching of each entry led to the removal of 16 entries that had not published crystallisation conditions, leaving the final count of crystallisation conditions at 104. Various conditions from these entries were analysed, such as pH, buffer, precipitant, and salt to develop a TPS-specific sparse matrix screen containing 48 different conditions (Supplementary Table S1).

2.8. Crystallisation and Optimisation of TPS Crystals

After creating the TPS-crystallisation screen, crystallisation experiments were performed. Each of the 10 CsTPS proteins were screened through the 48 conditions of the newly created TPS-crystallisation screen. Crystallisation was performed using the hanging-drop vapour diffusion method at room temperature. Briefly, a 1 µL drop of protein solution (10 mg mL^{-1} CsTPS) was mixed with a 1 µL drop of precipitant solution (Supplementary Table S1) and equilibrated against a 500 µL reservoir of the precipitant solution.

3. Results

3.1. Development of a Pipeline for the Characterisation of CsTPS Enzymes

Existing protocols for TPSs overexpression in bacterial systems often suffer from challenges such as protein misfolding, formation of inclusion bodies, binding of chaperones to recombinant proteins, low yield and stability issues (Raman et al., 2014; Wiles et al., 2022). To overcome these limitations, we established a robust pipeline for the overexpression, functional and structural characterisation of TPS's (Figure 1). Currently, 55 different TPS have been identified in the Cannabis genome (Allen et al., 2019). We selected 10 CsTPS to act as candidates in our structural and functional pipeline (Table 1). The selection criteria included representing different TPS types (mono-TPS and sesqui-TPS), including TPSs that can utilise multiple substrates or produce multiple products, and encompassing enzymes from both drug and fibre-type cannabis from a broad selection of different cultivars.

 doi:10.20944/preprints202504.2364.v1

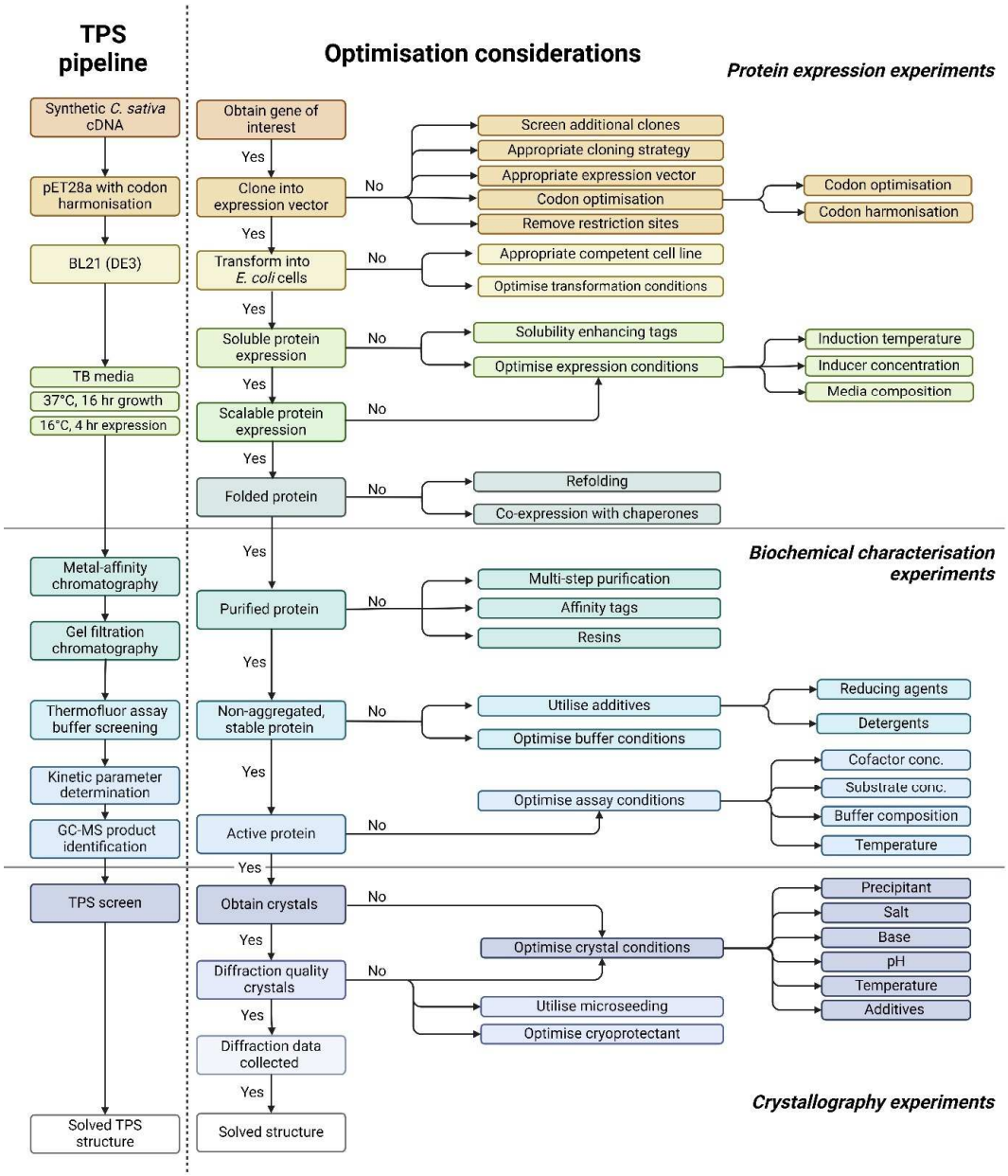

Figure 1. Pipeline for the functional and structural characterisation of *Cannabis sativa* terpene synthase enzymes. This flowchart outlines the comprehensive methodology for terpene synthase crystallography, highlighting the critical pre-crystallisation screening steps essential for successful crystal formation. The approach integrates key information from protein expression, biochemical characterisation, and secondary structure prediction, facilitating the design of optimised protein constructs and crystallisation conditions to enhance the likelihood of obtaining high-quality terpene synthase crystals.

3.2. Recombinant Expression and Purification of CsTPS Enzymes

The ten candidate CsTPS enzymes were successfully expressed recombinantly in *E. coli* and purified using immobilised metal-ion affinity chromatography (IMAC). SDS-PAGE analysis confirmed the presence of a single protein band with the expected molecular mass of 66.32-73.75 kDa for each enzyme, indicating that all ten CsTPS proteins were expressed and purified in their soluble forms (Supplementary Figure S1). After initial ion-affinity purification, yields of each CsTPS ranged from 2 to 25 mg L^{-1}. To further purify and characterise all CsTPS were subjected to size-exclusion chromatography (SEC). A typical trace chromatogram for each of the CsTPS (Supplementary Figure S2) revealed a single peak corresponding to a molecular weight of approximately 66.32-73.75 kDa, suggesting that all CsTPS are monomers in solution.

 doi:10.20944/preprints202504.2364.v1

3.3. Impact of Optimal Buffer Conditions for CsTPS Stability

The purified CsTPS enzymes were concentrated for structural studies. However, different levels of protein precipitation occurred after overnight storage. To overcome this problem, a thermofluor assay was used to determine the optimal storage buffer. Relatively large variations in stability could be observed when varying the buffers (Figure 2), and a few buffers appear to be generally more favourable for protein stabilisation (Figure 2b; Table 2). Tris pH 8.0 and 8.5, and HEPES pH 7.5 were the most stabilising buffers, whereas MES pH 5.5, Bis-Tris pH 6.0 and CAPS pH 9.0 significantly destabilised several of the CsTPS proteins (Table 2, Figure 2b). Several of the 56 buffer conditions investigated in the buffer screen gave no measurable transitions in combination with one or a few of the proteins, possibly caused by destabilisation or partial unfolding and potential aggregation of the proteins. For example, a clear thermal transition could be detected with only 6 of the 10 proteins in combination with CAPS pH 9.0. However, for some of the buffers, among them the overall most stabilising buffers, a measurable transition could be recorded together with all the proteins. The three most stabilising buffers were determined for each CsTPS protein, and Tris pH 8.0 was consistently found to improve the stability of the CsTPS enzymes. This contrasts with both the literature, where buffers with a pH of 7-7.5 are generally used for TPS proteins, and the theoretical PI, which is approximately 5-6. Given that the median ΔT m for Tris pH 8.0 with the addition of the salt additives 100 mM NaCl and 200 mM NaCl were found to be more than 4°C, these buffer conditions were considered to improve the stability of the CsTPS proteins significantly and were deemed suitable buffers for continuing functional characterisation of the enzymes.

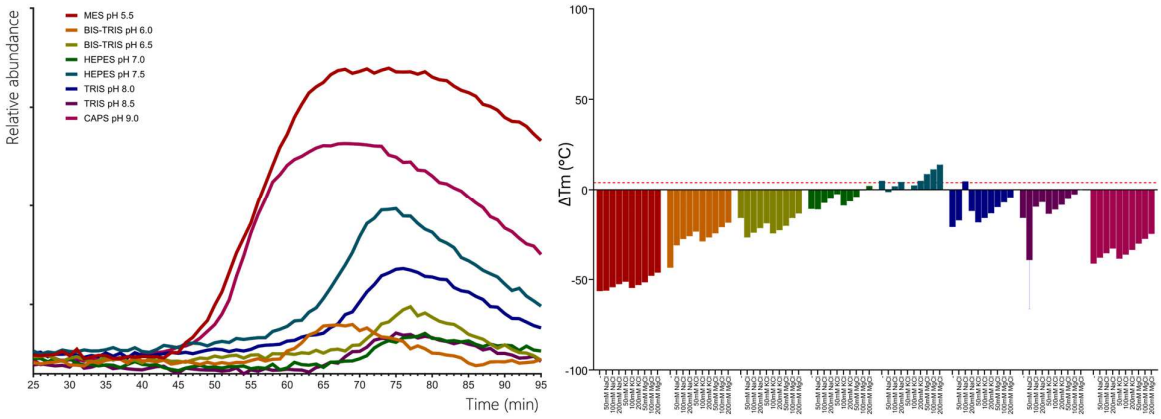

Figure 2. Thermal shift assay for improving protein stability of terpene synthases from Cannabis sativa **A.** Representative thermofluor melt curve for terpinolene synthases from C. sativa (CsTPS37FN). **B.** Changes in the unfolding transition temperature (ΔT_m) were calculated for each CsTPS protein in 80 buffer conditions. The bars represent the median ΔT_m values. A negative ΔT_m value signifies that the buffer destabilises the proteins, and a positive ΔT_m value indicates that the buffer has a stabilising effect. ΔT_m values more than 4°C are considered significantly stabilising, as indicated by the horizontal dashed red line.

Table 2. Summary of most stabilising buffer conditions resulting from the thermofluor melt curve for each of the Cannabis sativa terpene synthase (CsTPS) candidates.

TPS Name	*Best buffer condition*	T_m	ΔT_m
CsTPS3FN	0.2 M Tris pH 8.0 0.1 M NaCl	81.13°C	4.67°C
CsTPS1SK	0.2 M HEPES pH 7.5 0.1 M KCl	80.2°C	3.74°C
CsTPS5FN	0.2 M HEPES pH 7.0 0.1 M KCl	81.89°C	5.43°C
CsTPS37FN	0.2 M Tris pH 8.0 0.1 M NaCl	81.25°C	4.79°C
CsTPS9FN	0.2 M HEPES pH 7.5 0.1 M KCl	83.18°C	6.72°C

 doi:10.20944/preprints202504.2364.v1

CsTPS16CC	0.2 M HEPES pH 7.5 0.1 M KCl	79.36°C	2.9°C
CsTPS20CT	0.2 M HEPES pH 7.0 0.1 M NaCl	80.2°C	3.74°C
CsTPS19BL	0.2 M Tris pH 8.0 0.1 M NaCl	80.57°C	4.11°C
CsTPS12PK	0.2 M Tris pH 8.0 0.1 M KCl	81.03°C	4.57°C
CsTPS13PK	0.2 M HEPES pH 7.5 0.1 M KCl	81.67°C	5.21°C

3.4. Product Profile Analysis of CsTPS

To determine if the recombinant CsTPS were enzymatically active, all recombinant CsTPS were assayed as purified proteins with GGPP, GPP and FPP, and the reaction products were analysed by GC-MS. No products were seen for any CsTPS when assayed with GGPP, indicating that none of the CsTPS enzymes analysed form diterpenes.

3.5. Monoterpene Synthases (GPP Assay)

Five of the ten chosen CsTPS had previously been identified as proposed monoTPSs, while two were suggested to have dual activity with both GPP and FPP as substrates (Table 1). The five monoTPSs (CsTPS1SK, CsTPS3FN, CsTPS12PK, CsTPS13PK, and CsTPS37FN) exhibited enzymatic activity, producing a diverse array of monoterpenes (Figure 3, Table 3). CsTPS1SK predominantly produced limonene (74.72%), in line with previous studies (Günnewich et al., 2007), but we also detected the minor products β-pinene (5.18%), α-terpineol (4.83%), and β-terpineol (2.30%) (Figure 3; Table 3.). Interestingly, we also identified camphene (0.98%) and fenchol (3.55%), which were not previously reported for this enzyme. (Figure 3, Table 3). CsTPS3FN exhibited strict specificity for β-myrcene, producing this monoterpene at 100% when incubated with GPP, consistent with previous reports. (Figure 3, Table 3; (Booth et al., 2017)). CsTPS5FN also produced myrcene as its most abundant monoterpene product (37.09%) (Figure 3, Table 3). In addition to β-myrcene, several other monoterpenes were produced, including (-)-α-pinene (23.45%), (-)-limonene (17.33%), sabinene (15.26%), and (-)-β-pinene (8.87%), consistent with previous reports (Figure 3, Table 3; Booth et al., 2017b). CsTPS12PK predominantly produced α-terpinene (68.42%), a result consistent with previous studies (Booth et al., 2017), while also generating minor amounts of γ-terpinene (15.29%) and α-phellandrene (4.21%) (Figure 3, Table 3). Interestingly, our study also identified trace amounts of sabinene (3.87%), which had not been previously reported for this enzyme (Figure 3, Table 3). CsTPS13PK was highly specific for (Z)-β-ocimene, producing 94.11% of this product, which aligns with prior findings for this enzyme (Figure 3, Table 3; (Booth et al., 2017)). Small amounts of α-ocimene (3.49%) were also detected, though this was not previously noted (Figure 3, Table 3). For CsTPS37, the dominant product was terpinolene (81.62%), in agreement with past studies (Livingston et al., 2020b). Additionally, minor amounts of limonene (6.38%) and β-myrcene (3.89%) were observed, adding new insights to the product profile of this enzyme (Figure 3, Table 3). The TPS CsTPS19BL, when incubated with GPP, produced a mixture of monoterpenes, with the major products being (+)-linalool (54.12%) and (-)-linalool (40.87%) (Figure 3, Table 3). These results align with the previous studies that have suggested this enzyme's strong preference for linalool production when GPP is present in the reaction (Zager et al., 2019). The TPS CsTPS20CT was previously characterised as a sesquiterpene synthase, with its primary product being hedycaryol when incubated with FPP (Zager et al., 2019). However, when incubated with GPP in this study, CsTPS20CT displayed an unexpected monoterpene production profile, forming geraniol as the dominant product (87.64%) (Figure 3, Table 3).

 doi:10.20944/preprints202504.2364.v1

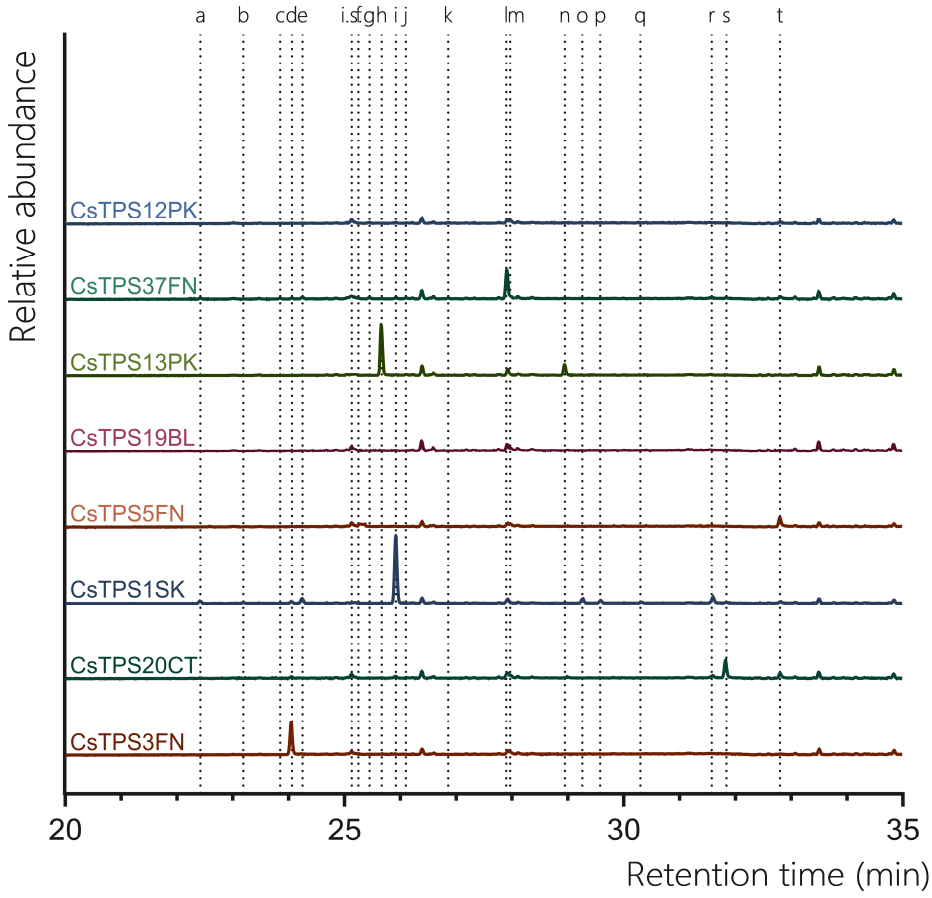

Figure 3. GC-MS traces showing the monoterpene products of CsTPS. Traces show GC-MS total ion chromatogram from CsTPS assays with GPP. CsTPS3FN: β-myrcene synthase; CsTPS16CC: Germacrene B synthase; CsTPS1SK: Limonene synthase; CsTPS5FN: Myrcene/ pinene synthase; CsTPS19BL: Nerolidol/ Linalool synthase; CsTPS13PK: Ocimene synthase; CsTPS37FN: Terpinolene synthase; CsTPS12PK: Terpinene synthase. Peaks: a) α-pinene, b) camphene c) β-phellandrene* d) β-myrcene, e) β -pinene, f) δ-carene, g) α-terpinene h) (E)-β-ocimene i) limonene, j) (Z)-β-ocimene, k) γ-terpinene, l) terpinolene, m) linalool, n) allo-ocimene*, o) fenchol*, p) β-terpineol*, q) pinene hydrate*, r) α-terpineol, s) geranyl methyl ether*, t) geraniol, i.s. = internal standard. *No reference standard available, putative identification of compound using National Institute of Standards and Technology (NIST) library.

Table 3. Terpene products formed by recombinant Cannabis sativa terpene synthase (CsTPS) enzymes.

TPS enzyme	*Substrate*	*Terpenes Produced*	*Percent Total (%)*
		Monoterpene Synthases	
CsTPS3FN	GPP	β-myrcene	100.00
CsTPS1SK	GPP	α-pinene	2.98
		camphene	0.98
		β-myrcene	2.25
		β-pinene	5.18
		limonene	74.72
		terpinolene	1.53
		fenchol*	3.55
		β-terpineol*	2.30
		α-terpineol	4.83
		geraniol	1.69
CsTPS12PK	GPP	α-terpinene	29.50
		limonene	33.36
		γ-terpinene	24.89
		β-myrcene	12.25
CsTPS13PK	GPP	(E)-β-ocimene	79.52
		allo-ocimene	1.80
		(Z)-β-ocimene	18.68

 doi:10.20944/preprints202504.2364.v1

CsTPS37FN	GPP	α-pinene	3.03
		β-phellandrene*	2.17
		β-myrcene	2.16
		β-pinene	4.77
		delta-3-Carene	3.67
		α-terpinene	2.86
		limonene	1.96
		y-terpinene	1.13
		terpinolene	70.69
		linalool	2.95
		geraniol	4.61
Sesquiterpene Synthases			
CsTP9FN	FPP	β-caryophyllene	2.76
		humulene	4.49
		epi-β-caryophyllene*	67.30
		germacrene D*	15.26
		globulol*	10.19
CsTPS16CC	FPP	β-elemene *	1.05
		γ-elemene*	3.82
		germacrene B*	92.73
		alloaromadendrene*	1.06
Mono/ Sesquiterpene Synthases			
CsTPS19BL	GPP	linalool	100.00
	FPP	nerolidol	100.00
CsTPS5FN	GPP	α-pinene	23.00
		β-myrcene	37.00
		β-pinene	8.00
		limonene	17.00
		sabinene*	15.00
	FPP	farnesol*	100.00
CsTPS20CT	GPP	β-myrcene	7.89
		limonene	12.72
		(Z)-β-ocimene	2.64
		terpinolene	7.74
		α-terpineol	26.06
		geraniol	42.95
	FPP	hedycaryol (elemol*)	31.42
		guaiol	19.96
		γ-eudesmol*	20.51
		α-eudesmol*	28.11

3.6. Sesquiterpene Synthases (FPP Assay)

All previously proposed sesquiTPSs analysed in this study were found to be active with FPP. CsTPS9FN produced a diverse range of sesquiterpenes, with epi-β-caryophyllene being the predominant product (67.30%) with β-caryophyllene (2.76%), humulene (4.49%) found to be present in low abundance (Figure 4, Table 3). Additionally, germacrene D (15.26%) and globulol (10.19%) were also detected, both of which were not previously reported for this enzyme (Figure 4, Table 3; (Booth et al., 2017). This contrasts previous studies where β-caryophyllene and α-humulene were the only products reported. Consistent with previous studies, when incubated with FPP, CsTPS20CT primarily produced hedycaryol (which was detected as elemol due to GC-MS thermal degradation) (31.42%) (Figure 4, Table 3; (Zager et al., 2019)). Other significant sesquiterpene products included γ-eudesmol (20.51%), α-eudesmol (28.11%), and guaiol (19.96%). This broad range of products indicates that CsTPS20CT is a multi-product sesquiterpene synthase, capable of producing a diverse array of eudesmol-type sesquiterpenes. The TPS CsTPS16CC produced γ-elemene (92.73%) as the dominant sesquiterpene when incubated with FPP, consistent with the enzyme's sesquiterpene synthase activity (Figure 3, Table 3). Other minor products included β-elemene (3.82%) and germacrene B (1.05%), further supporting its activity toward forming elemene-type sesquiterpenes. Interestingly, alloaromadendrene (1.06%) was also detected, which had not been previously associated with this enzyme. CsTPS5FN, when incubated with FPP, resulted in the exclusive production of farnesol (100%), a linear sesquiterpene alcohol (Figure 4, Table 3). Farnesol had previously been reported as the single sesquiterpene product of this enzyme, highlighting the dual activity of CsTPS5FN with GPP and FPP as substrates (Booth et al., 2017). This enzyme's ability to produce both monoterpenes and

 doi:10.20944/preprints202504.2364.v1

sesquiterpenes makes it a versatile target for biotechnological applications. Finally, the TPS CsTPS19BL produced 100% nerolidol when incubated with FPP, consistent with prior studies (Figure 4, Table 3; (Zager et al., 2019)).

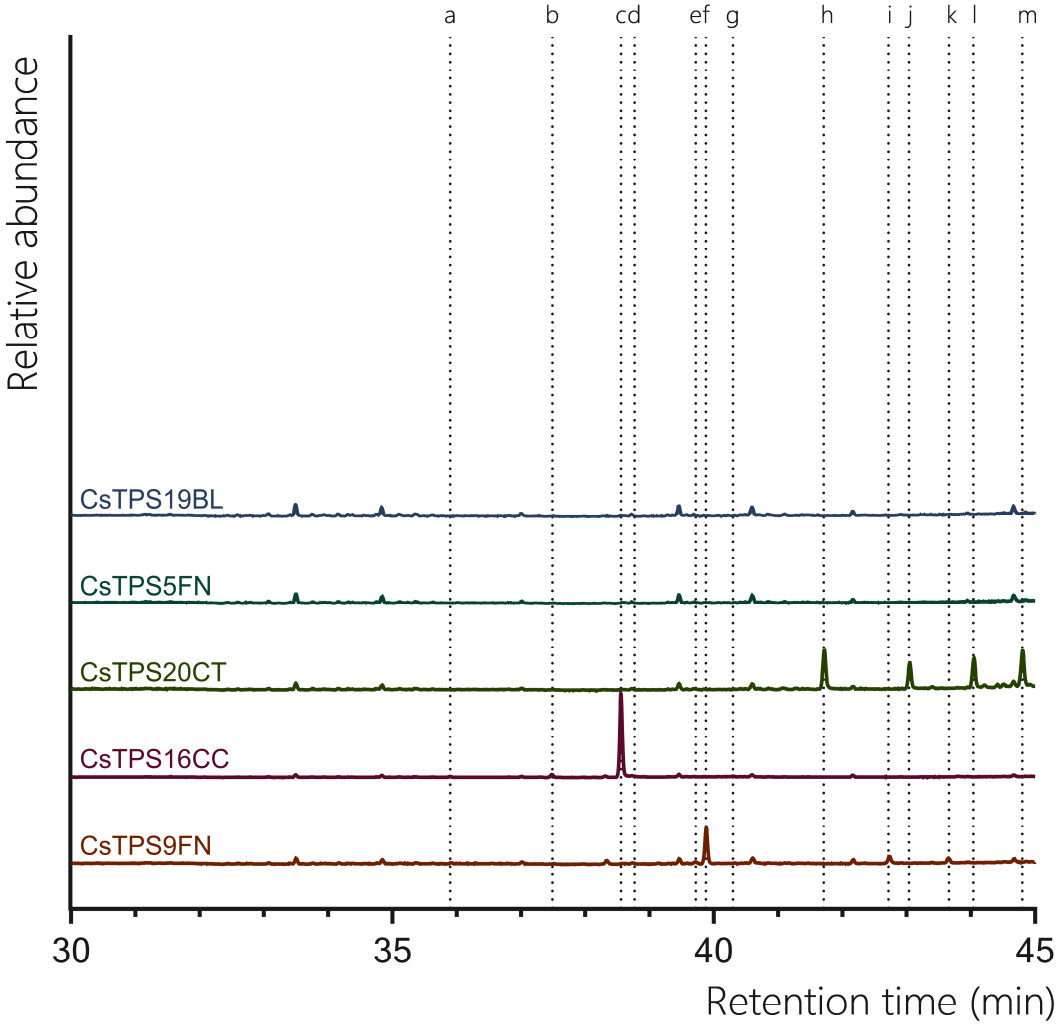

Figure 4. GC-MS traces showing the sesquiterpene products of CsTPS. Traces show GC-MS total ion chromatogram from CsTPS assays with FPP. CsTPS9FN: β-caryophyllene/ humulene synthase; CsTPS16CC: Germacrene B synthase; CsTPS20CT: Hedycaryol synthase; CsTPS5FN: Myrcene/ pinene synthase; CsTPS19BL: Nerolidol/ Linalool synthase. Peaks: a) β-elemene *, b) γ-elemene*, c) germacrene B*, d) β-caryophyllene, e) humulene, f) epi-β-caryophyllene*, g) alloaromadendrene*, h) nerolidol i) elemol*,j) germacrene-D*, k) guaiol, l) globulol*, m) γ-eudesmol*, n) α-eudesmol*. *No reference standard available, putative identification of compound using National Institute of Standards and Technology (NIST) library.

3.7. Kinetic Properties of CsTPS Enzymes

To better inform protein engineering strategies of TPSs, there is a need to understand the catalytic activity. We chose a high-throughput enzyme assay that can be performed without specialised equipment which is based on phosphate being produced. The kinetic profiles of the CsTPS are shown in Figure 5. Substrate concentrations (FPP, GPP) ranged from 0 to 100 μM. The calculated $k_{cat\ ap}$ for each CsTPS was found to be between 0.0011 and 0.0204 $s^{-1.}$ (Table 4).

 doi:10.20944/preprints202504.2364.v1

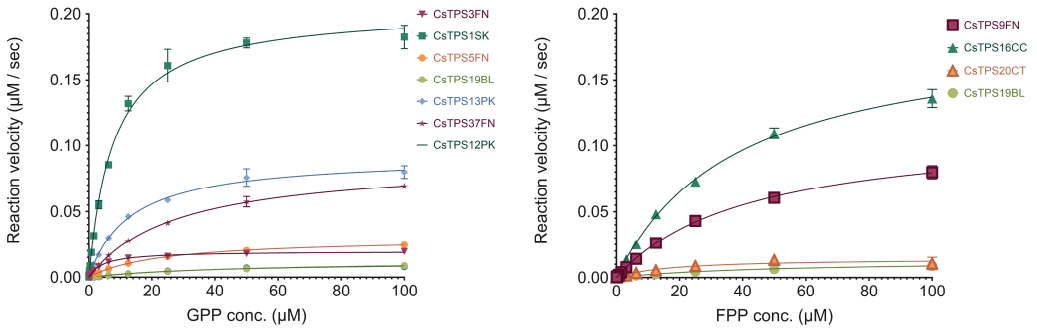

Figure 5. Michaelis-Menten kinetics of Cannabis sativa terpene synthases (CsTPS). (A) Non-linear regression analysis of steady-state kinetic assays for CsTPS enzymes using geranyl diphosphate (GPP) as the substrate, showing the rate of GPP catalysis (µM of GPP consumed per second). CsTPS3FN: β-myrcene synthase; CsTPS16CC: Germacrene B synthase; CsTPS1SK: Limonene synthase; CsTPS5FN: Myrcene/ pinene synthase; CsTPS19BL: Nerolidol/ Linalool synthase; CsTPS13PK: Ocimene synthase; CsTPS37FN: Terpinolene synthase; CsTPS12PK: Terpinene synthase (B) Non-linear regression analysis of steady-state kinetic assays for CsTPS enzymes using farnesyl diphosphate (FPP) as the substrate, showing the rate of FPP catalysis (µM of FPP consumed per second). Each curve represents the data fitted to the Michaelis-Menten equation to determine kinetic parameters. CsTPS9FN: β-caryophyllene/ humulene synthase; CsTPS16CC: Germacrene B synthase; CsTPS20CT: Hedycaryol synthase; CsTPS5FN: Myrcene/ pinene synthase; CsTPS19BL: Nerolidol/ Linalool synthase.

Table 4. Steady-state kinetic parameters for selected Cannabis sativa terpene synthases (CsTPS).

TPS Enzyme	*Substrate*	*K_m (µM)*	*V_{max} (µM^{-1} s^{-1})*	*K_{cat} (s^{-1})*
CsTPS3FN	GPP	4.569 ± 0.411	0.0196 ± 0.0005	0.0020
CsTPS9FN	FPP	41.7 ± 3.73	0.1127 ± 0.0047	0.0113
CsTPS16CC	FPP	38.43 ± 2.83	0.1895 ± 0.0063	0.0190
CsTPS20CT	FPP	16.86 ± 6.42	0.0144 ± 0.0022	0.0014
CsTPS1SK	GPP	7.809 ± 0.678	0.2038 ± 0.0053	0.0204
CsTPS5FN	GPP	23.3 ± 1.34	0.0300 ± 0.0007	0.0030
CsTPS19BL	GPP	48.45 ± 4.39	0.0129 ± 0.0006	0.0013
	FPP	17.32 ± 5.23	0.0102 ± 0.0002	0.0011
CsTPS12PK	GPP	41.85 ± 8.29	0.0119 ± 0.0011	0.0012
CsTPS13PK	GPP	12.96 ± 1.23	0.0918 ± 0.0029	0.0092
CsTPS37FN	GPP	27.71 ± 1.92	0.0884 ± 0.0025	0.0088

The kinetic analysis revealed a broad range of affinities and turnover rates across the different CsTPS variants, reflecting their diverse roles in terpene biosynthesis. For instance, CsTPS1SK exhibited a relatively low K_m for GPP (7.809 ± 0.678 µM) paired with a high V_{max} (0.2038 ± 0.0053 µM^{-1} s^{-1}) and K_{cat} (0.0204 s^{-1}), indicating a strong substrate affinity and rapid catalysis, which suggests its efficiency in monoterpene production. On the other hand, CsTPS9FN, which uses FPP as a substrate, demonstrated a higher K_m (41.7 ± 3.73 µM) but also a high V_{max} (0.1127 ± 0.0047 µM^{-1} s^{-1}), indicative of its function in sesquiterpene synthesis. Interestingly, CsTPS20CT had a much lower K_m for FPP (16.86 ± 6.42 µM) but exhibited the lowest V_{max} (0.0144 ± 0.0022 µM^{-1} s^{-1}) and K_{cat} (0.0014 s^{-1}), suggesting a distinct catalytic role or regulation mechanism within the terpene biosynthesis pathway. Additionally, CsTPS13PK showed a moderate K_m (12.96 ± 1.23 µM) for GPP and a V_{max} (0.0918 ± 0.0029 µM^{-1} s^{-1}), highlighting its efficient production of specific monoterpenes. This variability highlights the challenge of predicting enzyme function based solely on sequence similarity and the need for functional and kinetic characterisation. By elucidating these kinetic properties, our study provides a foundation for engineering CsTPS enzymes to produce specific terpenes more efficiently.

3.8. Development of a Targeted TPS Crystallisation Screen

The ability to make informed observations of the active site architecture of enzymes and thus engineer for productive or specified proteins of interest relies on finding appropriate crystallisation conditions. Currently, many macromolecular crystallisation screens are commercially available;

 doi:10.20944/preprints202504.2364.v1

however, relatively few are specifically designed for particular protein families. We created a 48-condition screen, called the TPS screen, with conditions specifically directed towards the crystallisation of the TPS protein family by analysing the crystallisation conditions for all TPS and TPS-related proteins identified in the PDB. Most TPS crystals were obtained between pH 6.0 and 6.9 (34.4%), with the next highest range from pH 7.0 to 7.9 (23.4%). The majority of TPS crystallised in Bis-Tris buffer (Figure 6A,B), with 24.6% of conditions ranging widely in pH from 3.8 to 10 (Figure 6A,B). Tris was the next most common buffer used to obtain TPS crystals (19.7%). Combining this data with the pH data resulted in the selection of six buffer/pH combinations for use in the TPS-directed screen. These buffers are Bis-Tris (pH 6.0), Bis-Tris (pH 6.5), Bis-Tris (pH 7.0), Tris (pH 7.5), Tris (pH 8.0) and Tris (pH 8.5). Of the 104 crystallisation conditions identified for TPS crystal production, 50 conditions reported the inclusion of salts to crystallise the molecule (Figure 6C). Although 22 different salts were used in these crystallisation conditions, a large majority (26%) of these conditions contained $MgCl_2$. NaCl was the next most common salt used and was seen in 12% of crystallisation conditions. This data led to the selection of both magnesium chloride and sodium chloride as counter ions in the TPS-directed screen.

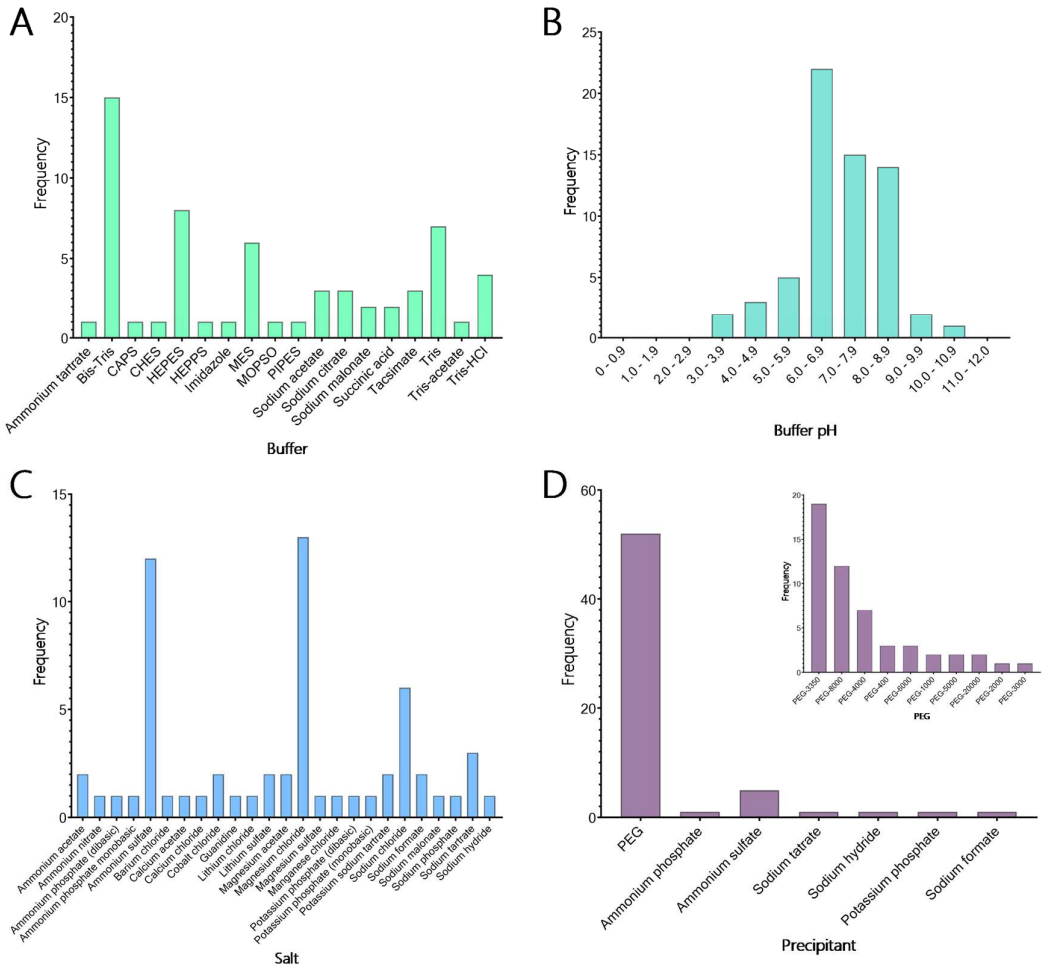

Figure 6. Crystallisation conditions for terpene synthase proteins. Data mining of the Protein Data Bank (PDB; Berman et al., 2002) yielded conditions used in the crystallisation of many terpene synthases. **A)** The buffers used in 104 crystallisation conditions shows that the majority of terpene synthases crystallise with Bis-Tris. The next most common buffers were Tris, HEPES and MES. **B)** The predominant pH used shows that the majority of terpene synthases crystallise between pH 6.0 and 6.9, with an overall preference for slightly basic pH. **C)** In the conditions that reported salts a total of 26 different salts were used, with most conditions containing magnesium chloride, ammonium sulphate and sodium chloride in the crystallisation of terpene synthase proteins. **D)** Of the crystallisation conditions that reported the precipitant composition, 84% contained some variation of PEG (inset). Of these conditions, there is a strong preference for PEG-3350 and PEG-8000.

 doi:10.20944/preprints202504.2364.v1

The final components of the analysed crystallisation conditions were precipitants, which were provided for 62 of the 104 conditions. Of the 16 different precipitants used, various molecular weights of polyethylene glycol (PEG) comprised 83.9% of the conditions (Figure 6D (Inset)). The next most used precipitant was ammonium sulphate, which was used in 8% of the conditions, indicating a strong preference for PEG in the crystallisation of TPS proteins. When examining the 52 conditions containing PEG in more depth (Figure 6D), PEG-3350 and PEG-8000 were most often used (30.6 and 19.4% of conditions, respectively). As a result, PEG-3350 and PEG-8000 were chosen to be included in the TPS-directed screen.

To create the final TPS screen, four concentrations (5, 15, 25 and 35%) of each molecular weight PEG (3350 and 8000) were chosen, along with 100 mM of each of the buffers (Bis-Tris pH 6.0, Bis-Tris pH 6.5, Bis-Tris pH 7.0, Tris pH 7.5, Tris pH 8.0 and Tris pH 8.5). Each condition additionally contains 200 mM $MgCl_2$ and 200 mM NaCl (Supplementary Table S2). This systematic approach to screen design provides a balanced matrix of variables aimed at identifying conditions for TPS crystal formation. By incorporating a diverse yet targeted range of parameters, this screen facilitates the formation of crystals, addressing the challenges of obtaining high-quality crystals essential for structural characterisation and subsequent protein engineering efforts.

3.9. Crystallisation and Optimisation of CsTPS Crystals

Each of the 10 CsTPS proteins were screened against the 48 conditions of the TPS-screen, as well as several other commercially available screens, all at room temperature. Initial crystal formation was observed as early as 24 hours post-screen setup, in many conditions of the TPS-screen. All 10 proteins demonstrated varying degrees of crystallisation, with at least some form of precipitate forming in every case. Well-defined crystals were observed for CsTPS37FN, CsTPS1Sk, CsTPS3FN, CsTPS13PK, and CsTP12PK, while several of the other proteins produced spherical aggregates. Interestingly, with the exception of CsTPS1SK, all CsTPS formed crystals with either star or cubic morphologies. CsTPS1SK, on the other hand, consistently produced rod-shaped crystals (Figure 7). The best crystals in terms of size and morphology were found in condition 16 of the TPS-screen (0.1 M Bis-Tris (pH 7.0), 25% PEG-3350, 0.2 M NaCl, 0.2M $MgCl_2$), using a 2:1 ratio of CsTPS protein (10 mg ml^{-1}) to reservoir solution. This condition was further optimised by varying the PEG, salt and protein: reservoir solution ratio. Diffraction-quality crystals were successfully formed for CsTPS37FN and CsTPS1SK. The crystals were taken to the Australian Synchrotron, where they diffracted to approximately 2.5 Å and 3.2 Å resolution, respectively (Figure 8).

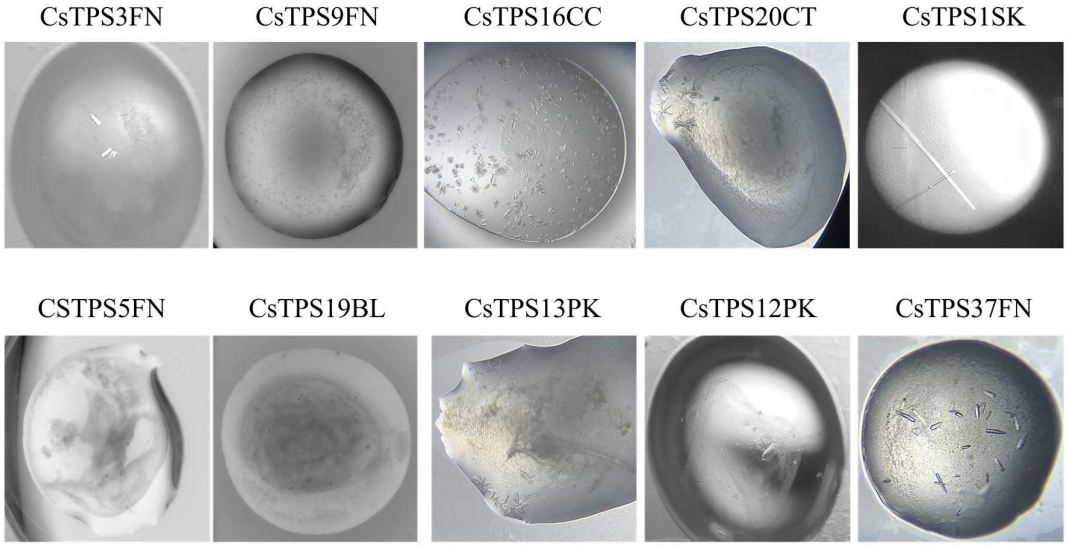

Figure 7. Crystals of recombinant terpene synthases from Cannabis sativa (CsTPS). The recombinant proteins were purified from E. coli and crystals were produced using the targeted terpene synthase screen.

 doi:10.20944/preprints202504.2364.v1

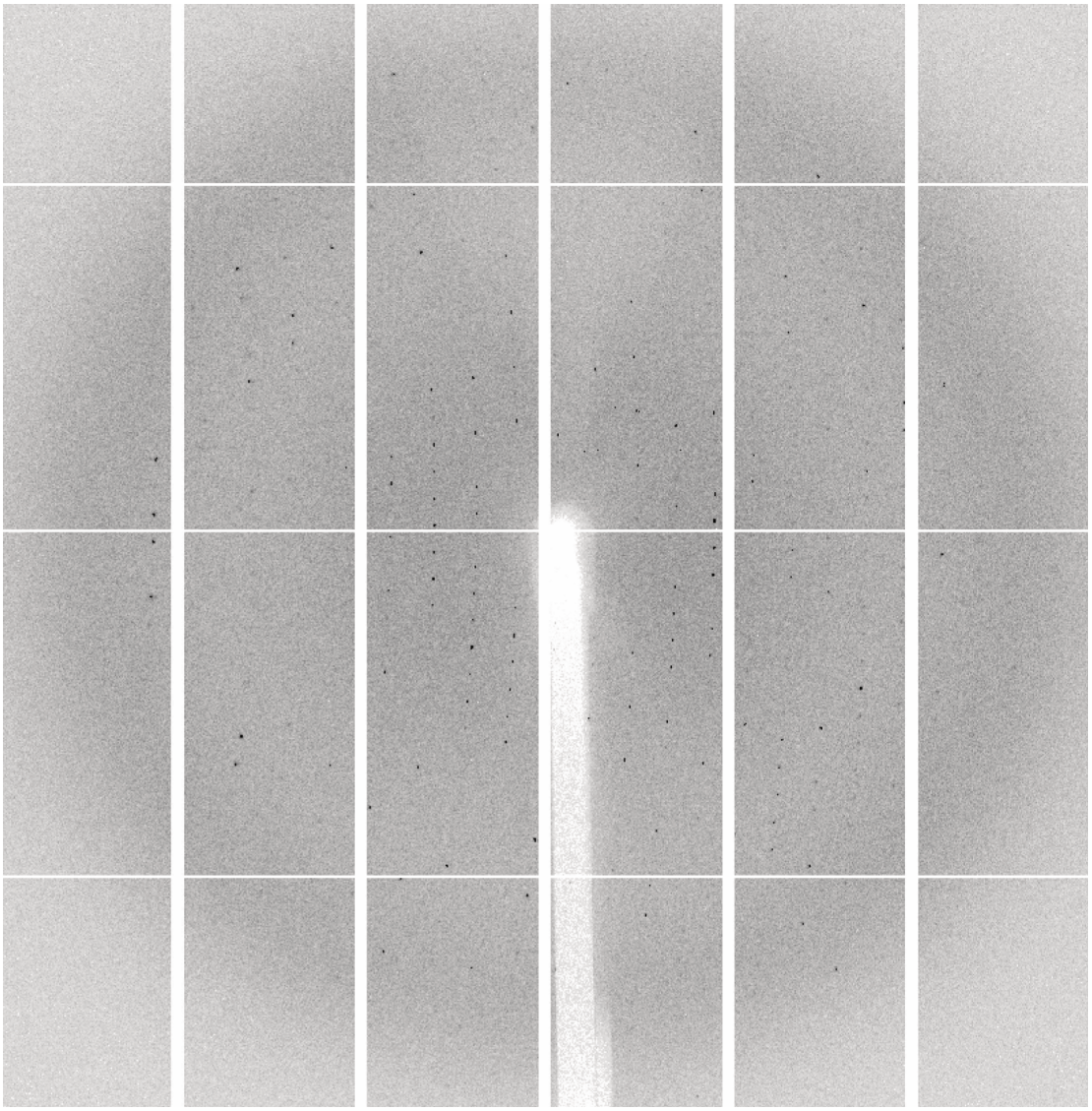

Figure 8. Data collection image from a crystal containing CsTPS37FN. Data was collected on the MX2 beamline at the ANSTO Australian Synchrotron, Clayton, VIC. Crystals diffracted to 2.5 Å resolution and were indexed in the space group $P2_1$.

4. Discussion

The resin of *Cannabis sativa* is rich in mono- and sesquiterpenes, which are believed to contribute to its pharmacological effects. While much of the research on cannabis terpenes has focused on phytochemical composition for forensics and breeding, the molecular biology of terpene formation in cannabis has received less attention. Understanding the enzyme functions and active site architecture of terpene biosynthesis in *C. sativa* is key for genetic improvements aimed at optimising terpene profiles. The functional diversity of TPSs, driven by subtle amino acid changes in the active site, plays critical roles in plant defence and scent profiles and holds significant biotechnological potential for industries such as cosmetics, food, agrochemicals, and pharmaceuticals. Here we describe a systematic pipeline that begins with the recombinant over-expression and characterisation of *C. sativa* TPS proteins and aims at targeted terpene production. Existing protocols for overexpressing TPSs in bacterial systems often face challenges such as protein misfolding, low yield, and instability. Our pipeline addresses these challenges by optimising factors like protein solubility, stability, and activity. By successfully expressing and purifying ten CsTPS enzymes in *E. coli*, this study advances our understanding of *C. sativa* TPS functionalities and offers a promising solution for high-yield production suitable for industrial-scale terpene synthesis. This work builds on previous work by (Allen et al., 2019; Booth et al., 2020, 2017; Booth and Bohlmann, 2019; Günnewich et al., 2007; Xu et al., 2024; Zager

 doi:10.20944/preprints202504.2364.v1

et al., 2019), which identified at least 55 distinct TPS genes in the *C. sativa* genome, functionally validating 38 of them (Wiles et al., 2022). Key insights, such as optimising buffer conditions to prevent protein aggregation, address the stability issues often encountered in the expression of plant-derived enzymes. Although this expression system offers a promising pathway for scaling up terpene production, further optimisation may be necessary to improve yields and stability, as seen in similar studies of other plant-derived enzymes. Understanding the molecular factors that govern enzyme expression in heterologous systems will be crucial for maximising production yields and integrating these enzymes into biotechnological workflows.

The functional characterisation of CsTPS enzymes presented in this work offers valuable insights into the catalytic properties, substrate specificities and product variability of distinct CsTPSs, all of which are critical for applications in synthetic biology and bio-based production platforms. Our findings demonstrate the considerable diversity in product profiles of CsTPS enzymes, as exemplified by limonene synthase (CsTPS1SK) and terpinolene synthase (CsTPS37FN) both of which utilise a common substrate (GPP) to produce a distinct dominant terpene product (limonene or terpinolene, respectively) alongside a myriad of other shared and distinct monoterpene products. In contrast, CsTPS19BL exhibited broader versatility, synthesising both the monoterpene alcohol linalool and the sesquiterpene alcohol nerolidol by catalysing two different substrates (GPP and FPP, respectively.). This versatility highlights the complexity of terpene biosynthesis, a theme also observed in other studies (Booth et al., 2017), where closely related TPS enzymes exhibit distinct functional behaviours due to minute differences in active site residues.

Importantly, our results reinforce that genetic sequence alone is insufficient to predict TPS function, necessitating biochemical validation. Furthermore, despite high sequence homology, CsTPS homologs from different plant species display distinct product profiles, as exemplified by CsTPS1SK, which, despite its similarity to limonene synthases from *Abies grandis* and *Mentha spicata* (Bohlmann et al., 1997; Hyatt et al., 2007; Srividya et al., 2020) , produces a different terpene spectrum. This divergence likely reflects evolutionary adaptations to ecological niches, emphasising the evolutionary plasticity of TPS enzymes. Beyond genetic variation, post-translational modifications and assay conditions may further modulate catalytic activity, as demonstrated by our characterisation of β-caryophyllene/α-humulene synthase (CsTPS9FN), which exhibited a broader product range than previously reported. This finding aligns with prior studies (Zager et al., 2019) who also reported condition-dependent differences in CsTPS enzyme activity, particularly CsTPS19BL, responsible for producing linalool and nerolidol. Our results confirmed nerolidol production when incubated with FPP and linalool with GPP, consistent with (Zager et al., 2019) who documented condition-dependent shifts in CsTPS enzyme activity, particularly CsTPS19BL, whose linalool and nerolidol production varied with assay conditions. Our results confirm that CsTPS19BL produces nerolidol from FPP and linalool from GPP, yet differences in product ratios suggest an interplay between genetic background and experimental parameters.

The catalytic turnover rates (*kcat*) of CsTPS enzymes, ranging from 0.0011 to 0.0204 s^{-1}, are consistent with the generally slow kinetics of TPS enzymes (Günnewich et al., 2007), but variations in substrate affinity indicate functional specialisation. Compared to thermostable TPSs with higher catalytic rates (Styles et al., 2017), CsTPS enzymes appear optimised for physiological conditions rather than extreme stability or activity, highlighting the balance between enzyme specificity and efficiency in terpene biosynthesis. Structural comparisons to other TPSs, such as *Artemisia annua* β-farnesene synthase (AaFS) and *Nicotiana tabacum* 5-epi-aristolochene synthase (TEAS), analysed using the malachite green assay (Vardakou et al., 2014), suggest that variations in *kcat* and *Km* may be attributed to structural adaptations. These insights reinforce the necessity of optimising expression systems and assay conditions for precise TPS characterisation. Overall, this study adds to the growing body of work on TPS enzymes, providing essential insights into the factors that influence terpene production in *C. sativa* and offering guidance for future efforts to optimise these enzymes for large-scale production of specialised metabolites.

The development of the TPS-crystallisation screen represents a significant advancement in the structural characterisation of CsTPS enzymes, facilitating the understanding of their catalytic mechanisms and enabling the future design of engineered variants with enhanced activity of altered product specificity. This work builds on previous structural studies of plant TPS enzymes, such as limonene synthases from *Mentha spicata* and *Citrus sinensis* (Hyatt et al., 2007; Kumar et al., 2017; Morehouse et al., 2017), 1,8-cineole synthase from *Salvia fruticose (Kampranis* et al., *2007),* taxadiene

 doi:10.20944/preprints202504.2364.v1

synthase from *Taxus brevifolia* (Köksal et al., 2011), abietadiene synthase from *Abies grandis* (Zhou et al., 2012), α-bisabolol synthase from *Artemisia annua* (Li et al., 2013) and 5-epi-aristolochene synthase from *Nicotiana tabacum* (Starks et al., 1997). Building on these studies, our research focuses on optimising a streamlined method to unveil the active sites of CsTPS. We optimised crystallisation conditions for CsTPS, yielding crystals and high-resolution diffraction data that will allow us to reveal key active site features, providing new insights into conserved motifs critical for catalysis and product specificity. These structural insights will assist in deepening our understanding of the conserved motifs that play critical roles in ionisation and protonation during catalysis and uncover the impact of variable plasticity residues on product specificity, suggesting which regions may be as influential as conserved motifs in determining product outcomes. This study is notable for its novel approach to crystallisation, which reveals conserved structural motifs and introduces the identification of variable plasticity residues, which may influence product outcomes, suggesting that these regions are as influential as conserved motifs in determining enzyme specificity. These findings align with previous studies (Kumari et al., 2013) that demonstrated the role of structural flexibility in TPS catalysis, further underscoring the complexity of terpene biosynthesis. To achieve this, we systematically developed a pipeline for obtaining diffraction-quality CsTPS crystals, optimising every stage of the process—from protein construct design to crystallisation conditions. This systematic approach, akin to that of (Pryor et al., 2012) highlights the adaptability of developing a narrow range of crystallisation conditions for specific protein families. The crystallisation pipeline developed here not only aids in the study of terpene synthases but also sets a precedent for the structural analysis of other enzyme families. By focusing on the crucial steps of crystallisation—such as pH balance, salt concentration, and precipitants, we were able to obtain high-resolution diffraction data for two CsTPS proteins and improved starting conditions for several other CsTPS. By refining the crystallisation conditions and optimising protein constructs, our study provides a valuable framework for the structural analysis of CsTPS enzymes, paving the way for future mutagenesis and kinetic studies. The insights gained from this work hold significant implications for the rational design of more efficient and specific CsTPS enzymes, with potential applications in synthetic biology and industrial biotechnology. The development of this crystallisation pipeline not only improves our understanding of terpene biosynthesis but also offers a model for the study of other complex enzymes, contributing to the broader field of enzyme engineering.

5. Conclusions

In this study, we have developed a TPS Screen for targeted crystallisation of CsTPS enzymes that allows for comprehensive characterisation, providing valuable insights into their biochemical properties, stability, and structural features. The findings emphasise the importance of optimising expression, purification, and assay conditions to maximise enzyme functionality and stability. The structural data obtained pave the way for future studies aimed at elucidating the mechanisms underlying terpene synthesis and exploring potential industrial applications of these versatile enzymes. Such endeavours not only enhance our understanding of plant secondary metabolism but also hold promise for developing novel biotechnological applications harnessing the power of terpenes.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Acknowledgments: We would like to acknowledge Myrna Deseo for her assistance with the Gas-Chromatography Mass Spectrometry.

References

1. Aizpurua-Olaizola, O., Soydaner, U., Öztürk, E., Schibano, D., Simsir, Y., Navarro, P., Etxebarria, N., Usobiaga, A., 2016. Evolution of the Cannabinoid and Terpene Content during the Growth of Cannabis sativa Plants from Different Chemotypes. J Nat Prod 79, 324–331. https://doi.org/10.1021/acs.jnatprod.5b00949
2. Allen, K.D., McKernan, K., Pauli, C., Roe, J., Torres, A., Gaudino, R., 2019. Genomic characterization of the complete terpene synthase gene family from Cannabis sativa. PLoS One 14, e0222363. https://doi.org/10.1371/journal.pone.0222363

 doi:10.20944/preprints202504.2364.v1

3. Andre, C.M., Hausman, J.F., Guerriero, G., 2016. Cannabis sativa: The plant of the thousand and one molecules. Front Plant Sci 7. https://doi.org/10.3389/fpls.2016.00019
4. Barcaccia, G., Palumbo, F., Scariolo, F., Vannozzi, A., Borin, M., Bona, S., 2020. Potentials and Challenges of Genomics for Breeding Cannabis Cultivars. Front Plant Sci 11. https://doi.org/10.3389/fpls.2020.573299
5. Bohlmann, J., Gershenzon, J., 2009. Old substrates for new enzymes of terpenoid biosynthesis. Proc Natl Acad Sci U S A 106, 10402–10403. https://doi.org/10.1073/pnas.0905226106
6. Bohlmann, J., Steele, C.L., Croteau, R., 1997. Monoterpene synthases from grand fir (Abies grandis): cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)- limonene synthase, and (-)-(1S,5S)-pinene synthase. Journal of Biological Chemistry 272, 21784–21792. https://doi.org/10.1074/jbc.272.35.21784
7. Bonini, S.A., Premoli, M., Tambaro, S., Kumar, A., Maccarinelli, G., Memo, M., Mastinu, A., 2018. Cannabis sativa: A comprehensive ethnopharmacological review of a medicinal plant with a long history. J Ethnopharmacol 227, 300–315. https://doi.org/10.1016/j.jep.2018.09.004
8. Booth, J.K., Bohlmann, J., 2019. Terpenes in Cannabis sativa—From plant genome to humans. Plant Science 284, 67–72. https://doi.org/10.1016/j.plantsci.2019.03.022
9. Booth, J.K., Page, J.E., Bohlmann, J., 2017. Terpene synthases from Cannabis sativa. PLoS One 12. https://doi.org/10.1371/journal.pone.0173911
10. Booth, J.K., Yuen, M.M.S., Jancsik, S., Madilao, L.L., Page, A.J.E., 2020. Terpene synthases and terpene variation in cannabis sativa. Plant Physiol 184, 130–147. https://doi.org/10.1104/PP.20.00593
11. Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3
12. Chandra, S., Lata, H., ElSohly, M.A., 2017. Cannabis sativa L.—botany and biotechnology. Cannabis sativa L.—Botany and Biotechnology 1–474. https://doi.org/10.1007/978-3-319-54564-6
13. Christianson, D.W., 2017. Structural and Chemical Biology of Terpenoid Cyclases. Chem Rev. https://doi.org/10.1021/acs.chemrev.7b00287
14. Cox-Georgian, D., Ramadoss, N., Dona, C., Basu, C., 2019. Therapeutic and medicinal uses of terpenes. Medicinal Plants: From Farm to Pharmacy 333–359. https://doi.org/10.1007/978-3-030-31269-5_15
15. Croteau, R., 1998. The Discovery of Terpenes. Discoveries in Plant Biology 329–343. https://doi.org/10.1142/9789812817563_0020
16. Downer, E.J., 2020. Anti-inflammatory Potential of Terpenes Present in Cannabis sativa L. ACS Chem Neurosci 11, 659–662. https://doi.org/10.1021/acschemneuro.0c00075
17. Ericsson, U.B., Hallberg, B.M., DeTitta, G.T., Dekker, N., Nordlund, P., 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal Biochem 357, 289–298. https://doi.org/10.1016/j.ab.2006.07.027
18. Grof, C.P.L., 2018. Cannabis, from plant to pill. Br J Clin Pharmacol 84, 2463–2467. https://doi.org/10.1111/bcp.13618
19. Günnewich, N., Page, J.E., Köllner, T.G., Degenhardt, J., Kutchan, T.M., 2007. Functional expression and characterization of trichome-specific (-)-limonene synthase and (+)-α-pinene synthase from Cannabis sativa. Nat Prod Commun 2, 223–232. https://doi.org/10.1177/1934578x0700200301
20. Hanuš, L.O., Hod, Y., 2020. Terpenes/Terpenoids in Cannabis: Are They Important? Med Cannabis Cannabinoids 3, 25–60. https://doi.org/10.1159/000509733
21. Hui-Lin, L., 1974. An archaeological and historical account of Cannabis in China. Econ Bot 28, 437–448.
22. Hyatt, D.C., Youn, B., Zhao, Y., Santhamma, B., Coates, R.M., Croteau, R.B., Kang, C., 2007. Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. Proc Natl Acad Sci U S A 104, 5360–5365. https://doi.org/10.1073/pnas.0700915104
23. Kalant, H., 2001. Medicinal use of cannabis: History and current status. Pain Res Manag 6, 80–91. https://doi.org/10.1155/2001/469629
24. Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, E., Katerelos, N.A., Anssour, S., Dunwell, J.M., Degenhardt, J., Makris, A.M., Goodenough, P.W., Johnsona, C.B., 2007. Rational conversion of

 doi:10.20944/preprints202504.2364.v1

substrate and product specificity in a Salvia monoterpene synthase: Structural insights into the evolution of terpene synthase function. Plant Cell 19, 1994–2005. https://doi.org/10.1105/tpc.106.047779

25. Karunanithi, P.S., Zerbe, P., 2019. Terpene Synthases as Metabolic Gatekeepers in the Evolution of Plant Terpenoid Chemical Diversity. Front Plant Sci 10. https://doi.org/10.3389/fpls.2019.01166
26. Köksal, M., Jin, Y., Coates, R.M., Croteau, R., Christianson, D.W., 2011. Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. Nature 469, 116–122. https://doi.org/10.1038/nature09628
27. Kumar, R.P., Morehouse, B.R., Matos, J.O., Malik, K., Lin, H., Krauss, I.J., Oprian, D.D., 2017. Structural Characterization of Early Michaelis Complexes in the Reaction Catalyzed by (+)-Limonene Synthase from Citrus sinensis Using Fluorinated Substrate Analogues. Biochemistry 56, 1716–1725. https://doi.org/10.1021/acs.biochem.7b00144
28. Kumari, S., Priya, P., Misra, G., Yadav, G., 2013. Structural and biochemical perspectives in plant isoprenoid biosynthesis. Phytochemistry Reviews. https://doi.org/10.1007/s11101-013-9284-6
29. Lange, B.M., Rujan, T., Martin, W., Croteau, R., 2000. Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. Proc Natl Acad Sci U S A 97, 13172–13177. https://doi.org/10.1073/pnas.240454797
30. Li, J.X., Fang, X., Zhao, Q., Ruan, J.X., Yang, C.Q., Wang, L.J., Miller, D.J., Faraldos, J.A., Allemann, R.K., Chen, X.Y., Zhang, P., 2013. Rational engineering of plasticity residues of sesquiterpene synthases from Artemisia annua: Product specificity and catalytic efficiency. Biochemical Journal 451, 417–426. https://doi.org/10.1042/BJ20130041
31. Livingston, S.J., Quilichini, T.D., Booth, J.K., Wong, D.C.J., Rensing, K.H., Laflamme-Yonkman, J., Castellarin, S.D., Bohlmann, J., Page, J.E., Samuels, A.L., 2020. Cannabis glandular trichomes alter morphology and metabolite content during flower maturation. Plant Journal 101, 37–56. https://doi.org/10.1111/tpj.14516
32. Luna-Vargas, M.P.A., Christodoulou, E., Alfieri, A., van Dijk, W.J., Stadnik, M., Hibbert, R.G., Sahtoe, D.D., Clerici, M., Marco, V. De, Littler, D., Celie, P.H.N., Sixma, T.K., Perrakis, A., 2011. Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. J Struct Biol 175, 113–119. https://doi.org/10.1016/j.jsb.2011.03.017
33. Masyita, A., Mustika Sari, R., Dwi Astuti, A., Yasir, B., Rahma Rumata, N., Emran, T. Bin, Nainu, F., Simal-Gandara, J., 2022. Terpenes and terpenoids as main bioactive compounds of essential oils, their roles in human health and potential application as natural food preservatives. Food Chem X 13, 100217. https://doi.org/10.1016/J.FOCHX.2022.100217
34. Morehouse, B.R., Kumar, R.P., Matos, J.O., Olsen, S.N., Entova, S., Oprian, D.D., 2017. Functional and Structural Characterization of a (+)-Limonene Synthase from Citrus sinensis. Biochemistry 56, 1706–1715. https://doi.org/10.1021/acs.biochem.7b00143
35. Nuutinen, T., 2018. Medicinal properties of terpenes found in Cannabis sativa and Humulus lupulus. Eur J Med Chem 157, 198–228. https://doi.org/10.1016/j.ejmech.2018.07.076
36. Oswald, I.W.H., Paryani, T.R., Sosa, M.E., Ojeda, M.A., Altenbernd, M.R., Grandy, J.J., Shafer, N.S., Ngo, K., Peat, J.R., Melshenker, B.G., Skelly, I., Koby, K.A., Page, M.F.Z., Martin, T.J., 2023. Minor, Nonterpenoid Volatile Compounds Drive the Aroma Differences of Exotic Cannabis. ACS Omega 8, 39203–39216. https://doi.org/10.1021/acsomega.3c04496
37. Paduch, R., Kandefer-Szerszeń, M., Trytek, M., Fiedurek, J., 2007. Terpenes: Substances useful in human healthcare. Arch Immunol Ther Exp (Warsz) 55, 315–327. https://doi.org/10.1007/s00005-007-0039-1
38. Pegan, S., Tian, Y., Sershon, V., Mesecar, A., 2009. A Universal, Fully Automated High Throughput Screening Assay for Pyrophosphate and Phosphate Release from Enzymatic Reactions. Comb Chem High Throughput Screen 13, 27–38. https://doi.org/10.2174/138620710790218203
39. Pryor, E.E., Wozniak, D.J., Hollis, T., 2012. Crystallization of Pseudomonas aeruginosa AmrZ protein: Development of a comprehensive method for obtaining and optimization of protein-DNA crystals. Acta Crystallogr Sect F Struct Biol Cryst Commun 68, 985–993. https://doi.org/10.1107/S1744309112025316
40. Raman, S., Rogers, J.K., Taylor, N.D., Church, G.M., 2014. Evolution-guided optimization of biosynthetic pathways. Proc Natl Acad Sci U S A 111, 17803–17808. https://doi.org/10.1073/pnas.1409523111

 doi:10.20944/preprints202504.2364.v1

41. Rocha, E.D., Silva, V.E.A., Pereira, F.C.S., Jean, V.M., Costa Souza, F.L., Baratto, L.C., Vieira, A.C.M., Carvalho, V.M., 2020. Qualitative terpene profiling of Cannabis varieties cultivated for medical purposes [Perfil de terpenos de variedades de Cannabis cultivadas para uso medicinal]. Rodriguesia 7.
42. Roell, M.S., 2020. Terpenes in Cannabis: Solving the Puzzle of How to Predict Taste and Smell. Plant Physiol. https://doi.org/10.1104/PP.20.00919
43. Russo, E.B., 2011. Taming THC: Potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. Br J Pharmacol 163, 1344–1364. https://doi.org/10.1111/j.1476-5381.2011.01238.x
44. Schilmiller, A.L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A.L., Schmidt, A., Wilkerson, C., Last, R.L., Pichersky, E., 2009. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc Natl Acad Sci U S A 106, 10865–10870. https://doi.org/10.1073/pnas.0904113106
45. Srividya, N., Lange, I., Lange, B.M., 2020. Determinants of Enantiospecificity in Limonene Synthases. Biochemistry 59, 1661–1664. https://doi.org/10.1021/acs.biochem.0c00206
46. Starks, C.M., Back, K., Chappell, J., Noel, J.P., 1997. Structural basis for cyclic terpene biosynthesis by tobacco 5-epi- aristolochene synthase. Science (1979) 277, 1815–1820. https://doi.org/10.1126/SCIENCE.277.5333.1815
47. Theis, N., Lerdau, M., 2003. The evolution of function in plant secondary metabolites. Int J Plant Sci 164. https://doi.org/10.1086/374190
48. Vardakou, M., Salmon, M., Faraldos, J.A., O'Maille, P.E., 2014. Comparative analysis and validation of the malachite green assay for the high throughput biochemical characterization of terpene synthases. MethodsX 1, e187–e196. https://doi.org/10.1016/j.mex.2014.08.007
49. Wiles, D., Shanbhag, B.K., O'Brien, M., Doblin, M.S., Bacic, A., Beddoe, T., 2022. Heterologous production of Cannabis sativa-derived specialised metabolites of medicinal significance—Insights into engineering strategies. Phytochemistry 203, 113380. https://doi.org/10.1016/j.phytochem.2022.113380
50. Xu, J., Kong, L., Ren, W., Wang, Z., Tang, L., Wu, W., Liu, X., Ma, W., Zhang, S., 2024. Identification and expression analysis of TPS family gene in Cannabis sativa L. Heliyon 10, e27817. https://doi.org/10.1016/j.heliyon.2024.e27817
51. Zager, J.J., Lange, I., Srividya, N., Smith, A., Markus Lange, B., 2019. Gene networks underlying cannabinoid and terpenoid accumulation in cannabis. Plant Physiol 180, 1877–1897. https://doi.org/10.1104/pp.18.01506
52. Zhou, F., Pichersky, E., 2020. The complete functional characterisation of the terpene synthase family in tomato. New Phytologist 226, 1341–1360. https://doi.org/10.1111/nph.16431
53. Zhou, K., Gao, Y., Hoy, J.A., Mann, F.M., Honzatko, R.B., Peters, R.J., 2012. Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from Abies grandis. Journal of Biological Chemistry 287, 6840–6850. https://doi.org/10.1074/jbc.M111.337592
54. Zuardi, A.W., 2006. History of cannabis as a medicine: A review. Revista Brasileira de Psiquiatria. https://doi.org/10.1590/S1516-44462006000200015