

Synthesis of 4-methylvaleric acid, a precursor of pogostone, involves a 2-isobutylmalate synthase related to 2-isopropylmalate synthase of leucine biosynthesis

Chu Wang^{a,b,1}, Ying Wang^{a,b,1}, Jing Chen^{a,b,1}, Lang Liu^{a,b}, Mingxia Yang^c, Zhengguo Li^{a,b}, Chengyuan Wang^c, Eran Pichersky^d, Haiyang Xu^{a,b}

^aSchool of Life Sciences, Chongqing University, Chongqing 401331, China.

^bCenter of Plant Functional Genomics, Institute of Advanced Interdisciplinary Studies, Chongqing University, Chongqing 401331, China.

^cThe Center for Microbes, Development and Health, Institute Pasteur of Shanghai, Chinese Academy of Sciences, University of Chinese Academy of Sciences, Shanghai 200031, China.

^dDepartment of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA.

¹These authors contributed equally

Address correspondence to: hyxu@cqu.edu.cn

Received: 2 March 2022

Accepted: 19 April 2022

ORCID IDs: 0000-0002-8605-0716 (Chengyuan Wang); 0000-0002-4343-1535 (Eran Pichersky); 0000-0002-6954-0000 (Haiyang Xu)

Key words: Pogostone, 4-Methylvaleric acid, 2-Isobutylmalate, α -Ketoacid elongation

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/NPH.18186](https://doi.org/10.1111/NPH.18186)

This article is protected by copyright. All rights reserved

pathway, 2-Isobutylmalate synthase, 2-Isopropylmalate synthase, Branched chain fatty acid, *Pogostemon cablin*

SUMMARY

- We show here that the side chain of pogostone, one of the major components of patchouli oil obtained from *Pogostemon cablin* and possessing a variety of pharmacological activities, is derived from 4-methylvaleric acid.
- We also show that 4-methylvaleric acid is produced through the one-carbon α -ketoacid elongation pathway with the involvement of the key enzyme 2-isobutylmalate synthase (IBMS), a newly identified enzyme related to isopropylmalate synthase (IPMS) of Leu biosynthesis.
- Site-directed mutagenesis identified Met¹³² in the N-terminal catalytic region as affecting the substrate specificity of PcIBMS1. And even though PcIBMS1 possesses the C-terminal domain that in IPMS serves to mediate Leu inhibition, it is insensitive to Leu.
- The observation of the evolution of IBMS from IPMS, as well as previously reported examples of IPMS-related genes involved in making glucosinolates in Brassicaceae, acylsugars in Solanaceae, and flavor compounds in apple, indicate that IPMS genes represent an important pool for the independent evolution of genes for specialized metabolism.

INTRODUCTION

Pogostemon cablin is a traditional Chinese medicinal plant in the Lamiaceae family, from whose aerial parts *Pogostemonis Herba*, a herbal preparation, is obtained.

Pogostemonis Herba or its essential oil component (known as patchouli oil) have been widely used in many Asian countries for the treatment of many ailments since ancient time due to a variety of pharmacological activities (Li *et al.*, 2013; Wang *et al.*, 2016; Hu *et al.*, 2017; Chen, JR *et al.*, 2021). Pogostone, one of the major chemical components of patchouli oil, has recently been demonstrated to possess various bioactivities including antimicrobial activities as well as pharmacological activities (Li *et al.*, 2012; Huang *et al.*, 2014; Li *et al.*, 2014; Peng *et al.*, 2014; Chen *et al.*, 2015; Cao *et al.*, 2017; Luchesi *et al.*, 2020; Ye *et al.*, 2021).

The biosynthetic pathway of pogostone is still unclear, but an outline of the pathway has been proposed that involves the branched-chain 4-methylvaleric acid (Fig. 1; Chen, J *et al.*, 2021). Furthermore, PcAAE2, a cytosolic acyl-activating enzyme that catalyzes the conversion of 4-methylvaleric acid to 4-methylvaleryl-CoA as part of the proposed pogostone biosynthetic pathway, has been identified and characterized (Chen, J *et al.*, 2021).

The formation of the branched C6 fatty 4-methylvaleric acid has not yet been elucidated in any species, although evidence for the involvement of either α -ketoacid elongation (α KAE) pathway or the fatty acid synthase (FAS)-mediated elongation pathway in the elongation of other branched-chain fatty acids in various species has been presented (Ohlrogge & Browse, 1995; Kroumova & Wagner, 2003; Slocombe *et al.*, 2008; Li-Beisson *et al.*, 2013). In the FAS route, both carbons from acetyl-ACP are retained per elongation cycle. In the α KAE pathway, acetyl-CoA is added but only one carbon is retained per elongation cycle. The α KAE pathway plays a role in the biosynthesis of a number of primary and specialized metabolites such as leucine and glucosinolates (Melendez-Hevia *et al.*, 1996; de Quiros *et al.*, 2000; Halkier & Gershenzon, 2006; Binder, 2010).

The α KAE steps involved in leucine biosynthesis have been well documented in various biological organisms (Binder, 2010) and are illustrative of the α KAE route in

general (Fig. 1). First, 2-isopropylmalate synthase (IPMS; EC 2.3.3.13) catalyzes an aldol-type condensation between acetyl-CoA and 2-oxoisovalerate to yield 2-isopropylmalate. Next, 2-isopropylmalate is converted by isopropylmalate isomerase (IPMI; EC 4.2.1.33) to 3-isopropylmalate. The final reaction, the oxidative decarboxylation of 3-isopropylmalate to 4-methyl-2-oxovalerate, is catalyzed by the isopropylmalate dehydrogenase (IPMDH; EC 1.1.1.85). The one-carbon elongated product 4-methyl-2-oxovalerate then undergoes a reversible transamination catalyzed by branched-chain aminotransferase (BCAT; EC 2.6.1.42) to yield the branched chain amino acid (BCAA) leucine in plastid. 4-Methyl-2-oxovalerate can be regenerated by transamination reactions by BCAT from leucine in mitochondria and converted into isovaleryl-CoA through irreversible oxidative decarboxylation catalyzed by a branched chain keto acid dehydrogenase (BCKDH, EC 1.2.4.4) (Binder, 2010). The isovaleryl-CoA generated in mitochondria from 4-Methyl-2-oxovalerate might be transported into cytosol by a transporter along with losing the CoA group to give free 3-methylbutanoic (Xu *et al.*, 2013).

In the Solanaceae, many species produced specialized metabolites called acyl sugars, in which at least some of the acyls attached to the sugar moiety are elongated branched-chains (Kroumova & Wagner, 2003; Slocombe *et al.*, 2008; Ning *et al.*, 2015). Feeding studies with labeled amino acids showed that the branched short chain fatty acid (2-methylpropanoic acid, 3-methylbutanoic acid and 2-methylbutanoic acid) are directly derived from Val, Leu and Ile, respectively (Kroumova & Wagner, 2003), while the branched acyl chains of C6 to C12 are derived from these BCAAs by elongation through α KAE pathway in tobaccos and petunia, but FAS pathway in *L. pennellii* and *D. metel* (Kroumova & Wagner, 2003).

Since the branched C5 fatty 3-methylbutanoic acid is derived from leucine metabolism involving the one-carbon elongation α KAE pathway (Kroumova & Wagner, 2003; Binder, 2010), 4-methylvaleric acid, the branched C6 fatty acid with one more carbon than 3-methylbutanoic acid, is most likely derived from leucine metabolism by one

additional elongation cycle of α KAE involving three reactions catalyzed by IPMS, IPMI, and IPMDH, or analogous enzymes (Fig. 1). We show here that the first reaction, in which 4-methyl-2-oxovalerate is elongated to 2-isobutylmalate, is catalyzed by a novel enzyme, 2-isobutylmalate synthase (IBMS), that is closely related to IPMS.

MATERIALS AND METHODS

Plant materials and chemicals

The plants of *Pogostemon cablin* and *Nicotiana benthamiana* were grown in soil under the condition as described previously (Chen, J *et al.*, 2021). The harvested tissues were flash frozen in liquid nitrogen and stored at -80°C until use.

Since standard 2-isobutylmalate is not available commercially, the substituted 2-malate derivative generated by the condensation of acetyl-CoA to 4-methyl-2-oxovalerate by PcIBMS1 was identified as 2-isobutylmalate based on the exact mass and fragment pattern by LC-QTOF-MS. All other commercial chemicals were purchased from Sigma-Aldrich.

Isotope feeding study with deuterium-labeled [$^2\text{H}_{11}$] 4-methylvaleric acid (4MVA- d_{11})

The five-week-old top leaves (5W-TLeaf) of *P. cablin* plants were placed in a 10-ml feeding solution right before the onset of darkness and transferred to a vacuum desiccator equipped with a vacuum pump. After incubation for 4 mins at the atmospheric pressure of 0.04 Mpa, the leaves in the feeding solution were transferred to a growth chamber and incubated under darkness for 12 h at 25°C . Both 4-methylvaleric acid (4MVA) and 4MVA- d_{11} feeding were performed at a concentration of 1 mM. The detailed procedures for analysis of metabolites of *P. cablin* leaves fed with 4MVA or 4MVA- d_{11} can be found in Supporting Information Methods S1.

Gene identification and coexpression analysis

Identification of gene candidates involved in 4-methylvaleric acid biosynthesis from our previously constructed *P. cablin* transcriptome database (bioproject accession number, PRJNA713906) was performed by querying our local nucleotide database with Arabidopsis representative enzymes involved in BCAA Leu metabolism using the TBLASTN function as previously described (Chen, J *et al.*, 2021). Coexpression analysis was performed by comparing the transcript abundance of *PcAAE2* and the identified gene candidates involved in 4-methylvaleric acid biosynthesis using the Pearson correlation method to quantitate the similarity of their expression profiles as described previously (Li *et al.*, 2018).

Quantitative RT-PCR analysis of the identified gene candidates involved in 4-methylvaleric acid biosynthesis

For Quantitative real-time PCR (RT-qPCR) analysis of transcripts in different tissues, RNA was extracted using Total RNA Isolation Kit from Omega with a DNA digestion step. The reverse transcription reaction was performed using High-Capacity cDNA Reverse Transcription Kit from Thermo Fisher Scientific following the manufacturer's instructions using random primers. The RT-qPCR analysis was performed using Power SYBR Green PCR master mix as previously described (Chen, J *et al.*, 2021). The relative expressions of these genes relative to the three reference genes (*GAPDH*, *Actin7* and *Tubulin3*) in each tissue were calculated as previously described (Chen, J *et al.*, 2021). Primers used in this study are listed in the Table S1.

Isolation of His-tagged recombinant proteins of PcIBMS1, PcIPMS1, PcIBMS1-M132L and PcIPMS1-L135M

The constructs for expression of His-tagged recombinant proteins of PcIBMS1, PcIPMS1, PcIBMS1-M132L and PcIPMS1-L135M in *Escherichia coli* were constructed

based on expression vector pET-28a (+). The details of vector construction can be found in Supporting Information Methods S1. Recombinant proteins of PcIBMS1, PcIPMS1, PcIBMS1-M132L and PcIPMS1-L135M were expressed in *E. coli* strain BL21(BE3) and purified using Ni-NTA agarose chromatography (Qiagen) as previously described (Xu *et al.*, 2018). The primer pairs used for construction of these pET-28a (+) expression vectors fused with HIS₆ residues at the N-terminus for expression in *E. coli* are listed in Table S1.

Enzymatic assays of recombinant PcIBMS1, PcIPMS1, PcIBMS1-M132L and PcIPMS1-L135M

The enzyme assay for the condensation reaction between acetyl-CoA and different 2-oxo acids was performed with a spectrophotometric end-point assay with DTNB as previously described (de Kraker *et al.*, 2007) with minor modification and the enzymatic products of PcIBMS1 and PcIPMS1 towards 2-oxoisovalerate and 4-methyl-2-oxovalerate were further measured using an LC-QTOF-MS-based method. Details of these enzymatic assays for condensation reaction can be found in Supporting Information Methods S1.

Modeling structures of PcIBMS1, PcIPMS1, PcIPMS1-L135M and PcIBMS1-M1132L

The target sequences were first split into domains by FUPred based on the deep-learning predicted contact map (Zheng *et al.*, 2020). Then the model of each domain and the full-chain sequence were generated through AlphaFold2 (Jumper *et al.*, 2021). Finally, all individual domains were assembled together to create the final full-length model by DEMO (Zhou *et al.*, 2019) based on the domain orientations derived from the AlphaFold2 full-length model.

Subcellular Localization of PcIBMS1 and PcIPMS1

The vector used for subcellular localization of analysis was pTF486. The construction

of these vectors expressing fusion proteins of PcIBMS1 and PcIPMS1 with GFP fused at C-terminus was performed as previously described (Chen, J *et al.*, 2021). The full-length open reading frames of PcIBMS1 and PcIPMS1 were cloned into pTF486 between Sall and BamHI sites using In-Fusion® HD Cloning Kit (TaKaRa) following the protocol. Protoplasts were prepared from Arabidopsis leaves, and transformation and confocal microscopy were performed as described previously (Yoo *et al.*, 2007). Primers used in the construction of these vectors are listed in Table S1.

Transient expression in *N. benthamiana* leaves

The pEAQ-HT constructs carrying each of *PcIBMS1* and *PcIPMS1* used for transformation were constructed as described previously (Xu *et al.*, 2019). The procedure used for transient expression in *N. benthamiana* leaves was performed as previously described (Xu *et al.*, 2019). *N. benthamiana* plants transformed with empty pEAQ-HT vector were used as control. The infiltrated leaves were collected 7 days after infiltration. The detailed procedures for analysis of metabolites in *N. benthamiana* leaves can be found in the Supporting Information Methods S1.

GC-MS analysis and LC-QTOF-MS analysis

For GC-MS analysis, 1 µl aliquot of the sample was injected into Agilent GC-MSD (Agilent 7890B-5977B) system equipped with the Rxi-5Sil MS column (30 m × 0.25 mm × 0.25 µm film thickness, RESTEK, USA). The oven temperature was programmed as previously described (Chen, J *et al.*, 2021). For LC-QTOF-MS analysis, 2 µl aliquot of the sample was injected into LC-QTOF-MS system (Agilent, 6545 LC/QTOF-MS) coupled with a C₁₈ column (ZORBAX RRHD Plus C₁₈, Φ2.1 × 50mm, 1.8 µm) at 35°C. The gradient (solvent A, water + 0.1% formic acid; solvent B, acetonitrile) program and operating parameters were set as previously described (Chen, J *et al.*, 2021).

RESULTS

4-Methylvaleric acid is a precursor of pogostone

Chen, J et al (2021) hypothesized that the side chain of pogostone is derived from 4-methylvaleric acid, and furthered showed that the expression of PcAAE2, a cytosolic acyl-activating enzyme that catalyzes the conversion of 4-methylvaleric acid to 4-methylvaleryl-CoA, is correlated with pogostone biosynthesis. However, a direct proof that 4-methylvaleric acid is a precursor of pogostone has not yet been presented. To determine if the side chain of pogostone is derived from 4-methylvaleric acid, five-week-old top leaves (5W-TLeaf) of *P. cablin* plants were placed in a solution containing the deuterium-labeled [$^2\text{H}_{11}$] 4-methylvaleric acid (4-methylvaleric- d_{11} acid or 4MVA- d_{11}) at the concentration of 1 mM. Following an overnight treatment with the labeled 4MVA- d_{11} , volatiles were extracted with MTBE and aliquots were analyzed by gas chromatography-mass spectrometry (GC-MS) (Fig. 2). Another portion of each MTBE extract was further dried and dissolved in 90% methanol for liquid chromatography-quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) (Fig. S1). We observed that 7.8% of the pogostone molecules extracted had an 11-D shift (from 223.0976 for pogostone to 234.1666 for pogostone- d_{11} in negative mode in LC-QTOF-MS analysis and from 224.1 for pogostone to 235.2 for pogostone- d_{11} in GC-MS analysis) in the parent ion mass (Fig. 2 and Fig. S1), indicating that 4-methylvaleric acid is a bona fide precursor of pogostone biosynthesis.

Identification of the gene candidates involved in 4-methylvaleric acid biosynthesis by coexpression analysis

To identify the candidate genes involved in 4-methylvaleric acid biosynthesis, we performed homology searches on our previously constructed *P. cablin* RNAseq database (Chen, J et al., 2021), using genes involved in leucine metabolism in *Arabidopsis thaliana* using TBLASTN software (Fig. 1). We obtained a total 13 unique genes (Table S2): two

genes with similarity to AtIPMS (named *PcIBMS1* and *PcIPMS1*, see below); two genes (*PcIPMIL SSU* and *PcIPMIL LSU*; “L” for “-like”) related to the small subunit and large subunit of Arabidopsis IPMI, respectively; one gene (*PcIPMDHL*) related to AtIPMDH; three genes putatively encoding branched-chain aminotransferases (*PcBCATL1*, *PcBCATL2* and *PcBCATL3*); and five genes putatively encoding the three subunits of branched chain α -keto acid dehydrogenase (*PcBCKDHL E1 α 1*, *PcBCKDHL E1 α 2*, *PcBCKDHL E1 β* , *PcBCKDHL E2* and *PcBCKDHL E3*).

We next performed coexpression analysis of these 13 unique genes through Pearson correlation analysis to look for genes whose expression was best correlated with that of *PcAAE2*, whose expression was previously shown to be correlated with pogostone biosynthesis (Chen, J *et al.*, 2021), in all six *P. cablin* tissues including Seedling, 5W-Root, 5W-Stem, 5W-TLeaf, 8W-TLeaf and 8W-SLeaf (Table S3). Two of the 13 genes, *PcIBMS1* and *PcIPMDHL*, showed statistically significance correlation ($P < 0.01$) (Table 1). In addition, in the coexpression analysis of all the unique genes in the *P. cablin* RNAseq database with the *PcAAE2* gene through Pearson correlation analysis, *PcIBMS1* ranked the top third unique gene (Table S4).

To more accurately test whether the transcript profiles of these 13 genes correlated with that of *PcAAE2* in different *P. cablin* tissues at different developmental stages as previously described (Chen, J *et al.*, 2021), their expression patterns were further confirmed by RT-qPCR in the same 13 tissue samples used for *PcAAE2* (Fig. S2). Again, the expression patterns of *PcIBMS1* in the tested tissue samples showed a strong positive correlation with that of *PcAAE2*, displaying a similar developmental pattern as that of *PcAAE2* and the distribution pattern of pogostone (Chen, J *et al.*, 2021).

Subcellular localization of PcIBMS1 and PcIPMS1

To investigate the subcellular compartmentation of PcIBMS1, subcellular localization prediction software TargetP 2.0 was first utilized for signal peptide prediction of PcIBMS1

as well as for PcIPMS1. Both PcIBMS1 and PcIPMS1 were predicted to harbor chloroplast transit peptides in their N-termini with a high score (Table S2). The subcellular localizations of these two proteins were further analyzed experimentally by transiently expressing in Arabidopsis leaf protoplasts PcIBMS1 and PcIPMS1 each fused to the green fluorescent protein (GFP) at their C-termini. In these experiments, complete co-localization of both PcIBMS1-GFP and PcIPMS1-GFP signals with red chlorophyll autofluorescence was observed (Fig. 3), indicating that both proteins are targeted to the plastids.

***In vitro* characterization of isopropylmalate and isobutylmalate synthase activities of PcIBMS1**

The combination of the positive correlation of expression patterns of *PcIBMS1* with that of *PcAAE2* and its further confirmation by RT-qPCR analysis made it the most likely candidate gene encoding the key enzyme in the biosynthetic pathway of 4-methylvaleric acid compared with the other twelve unique genes. To test this hypothesis, we performed *in vitro* biochemical assays with PcIBMS1 and, for comparison, PcIPMS1. Truncated PcIBMS1 and PcIPMS1 proteins, both lacking the first 40 amino acids (the estimated length of the transit peptide), were produced in *Escherichia coli* with a fused N-terminal His6 tag, and purified as soluble proteins. Both purified recombinant proteins were initially tested in a spectrophotometric end-point enzyme assay (DTNB, 5,5' -dithiobis (2-nitrobenzoic acid)) with a variety of 2-oxo acid substrates at a 0.3 mM concentration of the 2-oxo acid substrate and 1mM of acetyl-CoA.

Recombinant PcIBMS1 showed strong activity with 4-methyl-2-oxovalerate as a substrate, and also had significant activity with 2-oxoisovalerate (55% of its activity with 4-methyl-2-oxovalerate), but lower or no activity with the other substrates tested (Table 2). Recombinant PcIPMS1 utilized 2-oxoisovalerate as the best substrate, with only low levels of activity or no activity with all other substrates tested (Table 2). In particular, PcIPMS1 showed no detectable activity with 4-methyl-2-oxovalerate (Table 2).

To further characterize PcIBMS1 and PcIPMS1, we first tested the dependency of their enzymatic activities on pH and cations. The highest levels of activities of PcIBMS1 and PcIPMS1 with their respective preferred substrates were observed at pH 8.0 (Fig. S3). The activities of PcIBMS1 and PcIPMS1 were dependent on divalent cations with Mg²⁺ being the most preferred cation at optimal concentration of 10 mM (Fig. S3). The kinetic properties of PcIBMS1 with 4-methyl-2-oxovalerate and 2-oxoisovalerate as substrates, and PcIPMS1 with 2-oxoisovalerate as the substrate were measured in a spectrophotometric end-point enzyme assay. The kinetic analysis revealed that PcIBMS1 had a K_m value of $139.8 \pm 39.4 \mu\text{M}$ and a catalytic efficiency of $1.72 \times 10^4 \text{ s}^{-1}\text{M}^{-1}$ for 4-methyl-2-oxovalerate and a K_m value of $1303.3 \pm 127.7 \mu\text{M}$ and a catalytic efficiency of $3.337 \times 10^3 \text{ s}^{-1}\text{M}^{-1}$ for 2-oxoisovalerate, and that PcIPMS1 had a K_m value of $288.3 \pm 8.2 \mu\text{M}$ and a catalytic efficiency of $9.94 \times 10^3 \text{ s}^{-1}\text{M}^{-1}$ for 2-oxoisovalerate (Table 3). The enzymatic products of PcIBMS1 and PcIPMS1 with 2-oxoisovalerate and 4-methyl-2-oxovalerate as substrates were further verified with LC-QTOF-MS (Fig. S4).

Heterologous expression of *PcIBMS1* and *PcIPMS1* in *Nicotiana benthamiana*

To further test the activities of PcIBMS1 and PcIPMS1 *in planta*, *N. benthamiana* leaves were infiltrated with *Agrobacterium tumefaciens* strains harboring plasmids containing either *PcIBMS1* or *PcIPMS1*. For control, *Agrobacterium tumefaciens* strains harboring empty vector (pEAQ-HT) were infiltrated into *N. benthamiana* leaves. In these experiments, the complete open reading frames of *PcIBMS1* and *PcIPMS1* were used, including the corresponding transit peptides that were shown to direct the protein to the plastid. Transformed leaves were harvested seven days after infiltration and the products were extracted and analyzed by LC-QTOF-MS. As expected, transient expression of *PcIBMS1* in *N. benthamiana* leaves resulted in the overproduction of 2-isobutylmalate and 2-isopropylmalate, which were verified by the comparison of their retention times and mass spectrums with corresponding standards (Fig. 4a). In contrast, transient expression of

PcIPMS1 in *N. benthamiana* leaves resulted in negligible amounts of 2-isopropylmalate and 2-isobutylmalate compared with that produced in *N. benthamiana* leaves transiently expressing *PcIBMS1* (Fig. 4a).

To test whether the introduction of *PcIBMS1* or *PcIPMS1* can lead to the production of free 4-methylvaleric acid or an increase in its content, *N. benthamiana* leaf samples were extracted with MTBE, and the MTBE extracts were further analyzed by GC-MS. No free 4-methylvaleric acid was detected in control *N. benthamiana* leaves and in *N. benthamiana* leaves transiently expressing *PcIBMS1* or *PcIPMS1* (Fig. 4b). Since it had been previously observed that compounds with carboxylic groups generated in *N. benthamiana* leaves by the introduction of heterologous genes tend to be modified by endogenous glycosylases (Yang *et al.*, 2011; Xu *et al.*, 2018), we searched for glycosylated 4-methylvaleric acid by performing hydrolysis on leaf samples with 0.4N NaOH at 80 °C for 20 min. Surprisingly, 4-Methylvaleric acid and 5-methylhexanoic acid (containing one more carbon than 4-methylvaleric acid) were detected at low levels in the hydrolyzed MTBE extracts of control *N. benthamiana* leaves. However, these two compounds in *N. benthamiana* leaves expressing *PcIBMS1* showed 6.9-fold and 2.7-fold increases in their contents, respectively, compared with that in control *N. benthamiana* leaves (Fig. 4b and c). Transient expression of *PcIPMS1* did not result in any observable changes in the contents of these two branched chain fatty acids in the hydrolyzed MTBE extracts compared with that in the control *N. benthamiana* leaves (Fig. 4b and c). In addition, the relative content of Leu in *N. benthamiana* leaves expressing *PcIBMS1* declined to about 58% compared with that in control *N. benthamiana* leaves (Fig. S5), while no change in Leu concentrations was observed in *N. benthamiana* leaves expressing *PcIPMS1*, which yielded almost the same content of Leu as control *N. benthamiana* leaves (Fig. S5).

Leu feedback inhibition assay of *PcIBMS1* and *PcIPMS1*

IPMSs from plants and bacteria generally form a homodimeric protein (Koon *et al.*,

2004; de Carvalho *et al.*, 2005; de Kraker *et al.*, 2007; de Kraker & Gershenzon, 2011). Each monomer contains two core domains including an N-terminal catalytic region consisting mainly of a (β/α)₈ barrel (TIM barrel) with a divalent metal cofactor necessary for substrate binding (Koon *et al.*, 2004), and a C-terminal allosteric regulatory domain responsible for Leu feedback inhibition in microbes and plants (de Carvalho *et al.*, 2005; de Kraker & Gershenzon, 2011). The two major domains are separated by two small subdomains that have a flexible hinge in between them (Koon *et al.*, 2004). The full-length *PcIBMS1* gene encodes a protein containing 591 amino acids, showing 74%, 66% and 64% amino acid identities with PcIPMS1, AtIPMS1 and OsIPMS1, respectively (Fig. S6), the latter two proteins functionally identified as typical IPMSs in *Arabidopsis thaliana* and *Oryza sativa* (de Kraker *et al.*, 2007; Xing & Last, 2017; He *et al.*, 2019). Sequence comparison with other typical IPMSs showed that the protein encoded by *PcIBMS1* gene contains the N-terminal catalytic region and likely C-terminal allosteric regulatory domain (Fig. S6). However, many changes in residues occur in the PcIBMS1 protein at positions corresponding to the conserved sites of the other typical IPMS proteins in both N-terminal catalytic region (Fig. 5) and C-terminal allosteric regulatory domain, notably the loss of five amino acids in the C-terminal domain of the PcIBMS1 protein (Fig. S7). The sequence divergence in the C-terminal allosteric regulatory domain of PcIBMS1 protein compared with other typical IPMSs that are sensitive to Leu feedback inhibition led us to hypothesize that *PcIBMS1* encodes an IPMS family protein that is not subject to Leu feedback inhibition.

To test this hypothesis, both recombinant PcIBMS1 and PcIPMS1 were assayed for aldol-type condensation activities between acetyl-CoA and 4-methyl-2-oxovalerate and 2-oxoisovalerate, in the end-point enzyme assay (DTNB) in the presence of Leu at a series of concentrations from 0.05 mM to 10 mM. PcIBMS1 activity was insensitive to Leu concentrations up to the highest concentration measured (10 mM) no matter whether the 2-oxo acid substrate is 4-methyl-2-oxovalerate or 2-oxoisovalerate (Fig. 6a and b). On the

other hand, the activity of PcIPMS1 with the substrate 2-oxoisovalerate declined to 57.8% at 50 μ M Leu and was reduced to 34.7% at the highest Leu concentration measured (10 mM) (Fig. 6b). The observations of substrate specificity and kinetic properties of PcIPMS1 enzyme towards 2-oxoisovalerate and acetyl-CoA and its sensitivity to Leu feedback inhibition indicate that PcIPMS1 encodes a bona fide IPMS enzyme responsible for Leu biosynthesis.

Determination of amino acid sites in the N-terminal domain that may affect the substrate specificity of PcIBMS1 and PcIPMS1

Our data indicates that PcIBMS1 is an IPMS-like protein with an altered preferred substrate (4-methyl-2-oxovalerate rather than 2-oxoisovalerate). Previous studies on the 3D structure of MtIPMS from *Mycobacterium tuberculosis* and BjMAM1-A from *Brassica juncea* identified seven amino acid sites in the N-terminal catalytic domain that may affect the 2-oxo acid substrate chain length specificities of IPMS and methylthioalkylmalate synthase (MAMS) (de Kraker & Gershenzon, 2011; Kumar *et al.*, 2019), the latter a neofunctionalized form of IPMS involved in the biosynthesis of Methionine-Derived glucosinolates (de Kraker *et al.*, 2007; de Kraker & Gershenzon, 2011). The overall fold of MAMs shares with IPMSs the N-terminal catalytic α/β -barrel domain and the C-terminal α -helical region that forms part of the CoA binding site (Kumar *et al.*, 2019). The two major structural differences between MAMs and IPMSs are the loss of the N-terminal extension and the C-terminal Leu binding regulatory domain found in the Leu biosynthesis enzyme. In both MAMs and IPMSs, the catalytic machinery and reaction chemistry are conserved (Kumar *et al.*, 2019).

Sequence comparison of the N-terminal domain of PcIBMS1 that catalyzes the extension of the longer chain substrate 4-methyl-2-oxovalerate with those of the five IPMS proteins, including PcIPMS1, AtIPMS1, AtIPMS2, MdIPMS1 and MdIPMS2, all of which show no or negligible activities towards the longer chain substrate 4-methyl-2-oxovalerate

compared with that towards the classic IPMS substrate 2-oxoisovalerate (de Kraker *et al.*, 2007; Sugimoto *et al.*, 2021), highlights the sequence divergence in PcIBMS1 (Fig. 5). Twenty conserved amino acid sites of N-terminal catalytic domains of these five typical IPMS proteins diverged from the corresponding sites in PcIBMS1 protein (Fig. 5). In particular, the conserved Leu located in one of the seven sites in these IPMSs that were identified as accounting for the differences in substrate binding between IPMSs and MAMs (de Kraker & Gershenzon, 2011; Kumar *et al.*, 2019), is replaced by Met¹³² in the same position in PcIBMS1 (Fig. 5), suggesting that the substituted Met¹³² may play a vital role in the alteration of substrate specificity of PcIBMS1 from typical IPMSs. To further explore the implication of the Leu-to-Met substitution, we performed structure modeling of PcIBMS1 and PcIPMS1. The overall structures of PcIBMS1 and PcIPMS1 obtained by modeling are similar and share a conserved substrate binding pocket (Fig. S8). The Leu-to-Met substitution in this pocket results in the distance between 4-Methyl-2-oxovalerate and the residues changing from 4.9Å to 3.1Å (Fig. S8 and Fig. S9). Thus, the Met¹³² in PcIBMS1 has a stronger binding affinity to 4-Methyl-2-oxovalerate than Leu¹³⁵ in PcIPMS1.

To test the effect of Met¹³² in determining substrate specificity of PcIBMS1, we mutated this residue in PcIBMS1 to the corresponding residue Leu found in the typical IPMS, and we did the reciprocal experiment by changing the corresponding Leu¹³⁵ in PcIPMS1 to Met. We assayed the resulting PcIBMS1-M132L and PcIPMS1-L135M point mutants using 4-methyl-2-oxovalerate and 2-oxoisovalerate and determined their kinetic parameters (Table 3). The PcIBMS1-M132L mutant exhibited about 3-fold decrease in K_m value for 2-oxoisovalerate with a corresponding approximate 2.4-fold increase in catalytic efficiency (k_{cat}/K_m), while showing only a modest difference in the K_m value for 4-methyl-2-oxovalerate and about 2-fold decrease in catalytic efficiency versus the wild type. The PcIPMS1-L135M point mutant exhibited about 2.2-fold increase in K_m value for 2-oxoisovalerate with a corresponding approximate 5-fold decrease in catalytic efficiency

(k_{cat}/K_m) versus the wild type, while still showing no detectable activities with 4-methyl-2-oxovalerate. These results suggest that this amino acid site indeed contributes to the ability of the enzyme to accept the longer chain 2-oxo acid 4-methyl-2-oxovalerate but that other amino acids play a crucial role as well.

Phylogenetic relationships of PcIBMS1 homologs

To gain further insight into the evolution of PcIBMS1, we expanded the phylogenetic analysis of PcIBMS1 and PcIPMS1 to include their closest homologs from plant species and microorganisms, including bona fide IPMSs, MAMs from *Arabidopsis thaliana* and *Brassica juncea*, and citramalate synthase (CMS) from *Malus domestica* (Fig. 7). PcIBMS1, MdCMS1 and MAMs, as well as SIIPMS3, a tomato true IPMS that is nonetheless lacking the regulatory C-terminal regulatory domain and is involved in synthesizing the C5 acyl precursor for acylsugar biosynthesis rather than leucine, split off from typical plant IPMSs prior to the split between monocot and dicot IPMSs (Fig. 7a). However, these leucine-insensitive enzymes represent neofunctionalized forms of typical plant IPMS (Ning *et al.*, 2015; Kumar *et al.*, 2019; Sugimoto *et al.*, 2021), that either lack the regulatory C-terminus domain altogether or possess the C-terminal domain but show high rate of change in it. We therefore redid the phylogenetic analysis with the protein sequences from which the C-terminal regulatory domain was removed (Fig. 7b). The phylogenetic tree obtained in this analysis showed that PcIBMS1 is most closely related to PcIPMS1.

DISCUSSION

4-Methylvaleric acid is a precursor of pogostone

Pogostone possesses many pharmaceutical activities and thus is considered as a potential therapeutic agent for treatment of many diseases. In our previous study, we

constructed a *P. cablin* RNAseq database, from which we isolated and characterized PcAAE2, an enzyme that catalyzes the formation of 4-methylvaleryl-CoA from 4-methylvaleric acid, a branched-chain fatty acid (Chen, J *et al.*, 2021). The expression of *PcAAE2* correlates spatially and temporally with pogostone biosynthesis, and based on this data as well as the structure of pogostone, 4-Methylvaleric acid was hypothesized as a precursor of pogostone (via 4-methylvaleryl-CoA). Here, by feeding *P. cablin* young leaves with deuterium-labeled [²H₁₁] 4-methylvaleric acid (4MVA-d₁₁) and following the incorporation of this label into pogostone, we were able to show that 4-methylvalerate is indeed a precursor in pogostone biosynthesis (Fig. 2 and Fig. S1).

PcIBMS1 is a key enzyme involved in the formation of 4-methylvaleric acid

Branched and straight fatty acids have been hypothesized as the precursors of many specialized metabolites such as bitter acids in *Humulus lupulus*, acylsugars in Solanaceae and cannabinoids in *Cannabis sativa* (Kroumova & Wagner, 2003; Stout *et al.*, 2012; Xu *et al.*, 2013; Ning *et al.*, 2015), although direct evidence for the activity of key enzymes in their biosynthesis has seldom been presented. To further elucidate the route leading to biosynthesis of pogostone, we set to characterize the key enzyme involved in the biosynthesis of 4-methylvaleric acid. This branched C₆ fatty acid is longer by one carbon than 3-methylbutanoic acid, a derivative of Leu whose own synthesis involves the α -ketoacid elongation (α KAE) pathway (Kroumova & Wagner, 2003; Binder, 2010). Therefore, 4-methylvaleric acid is most likely derived from leucine metabolism by one additional elongation cycle through α KAE pathway.

Among the 13 gene candidates for 4-methylvaleric acid biosynthesis that were identified by utilizing Arabidopsis representative genes involved in leucine metabolism to query the assembled *P. cablin* transcriptome data through homology search by TBLASTN software, PcIBMS1 showed the highest correlation with pogostone biosynthesis markers (Table 1, Table S4 and Fig. S2) and was homologous to IPMS, the key enzyme in the one-

carbon elongation of 2-oxoisovalerate to 4-methyl-2-oxovalerate (via 2-isopropylmalate) in Leu biosynthesis, making it the most likely candidate encoding the key enzyme for 4-methylvaleric acid biosynthesis. *In vitro* biochemical assays of the recombinant PcIBMS1 protein showed that PcIBMS1 not only effectively catalyzed the formation of 2-isobutylmalate from 4-methyl-2-oxovalerate and acetyl-CoA but was also able to catalyze the formation of 2-isopropylmalate from 2-oxoisovalerate and acetyl-CoA (Table 2, Table 3 and Fig. S4). A second IPMS-like gene found in the *P. cablin* transcriptome but whose expression pattern did not correlate with pogostone biosynthesis did not have activity with 4-methyl-2-oxovalerate (Table 3). This gene, which was designated as PcIPMS1, was shown to be a *bona fide* isopropylmalate synthase that effectively uses 2-oxoisovalerate as a substrate with kinetic properties similar to other plant typical IPMSs (Table 3; de Kraker *et al.*, 2007; Ning *et al.*, 2015). The K_m value of PcIBMS1 for 2-oxoisovalerate was almost 5-fold higher than that of PcIPMS1, but its K_m value for 4-methyl-2-oxovalerate was 3-fold lower than the K_m value of PcIPMS1 with 2-oxoisovalerate. However, PcIBMS1 has a somewhat high K_m value for acetyl-CoA when acting on its preferred substrate, compared with IPMSs (Table 3; de Kraker *et al.*, 2007; Ning *et al.*, 2015). These enzymatic properties of PcIBMS1 suggest it may perform distinct biological functions from typical plant IPMSs that are responsible for Leu biosynthesis.

Attempts to establish a reliable *P. cablin* transformation system, including gene silencing (VIGS) using tobacco rattle virus, were not successful, preventing us from testing the biological functions of PcIBMS1 directly in *P. cablin* plants. Therefore, the activities of PcIBMS1 *in planta* were tested in *N. benthamiana* leaves by transient expression. Analysis of compounds produced in *N. benthamiana* leaves when *PcIBMS1* was transiently expressed lent additional evidence for its function in 4-methylvaleric acid biosynthesis. We were able to directly detect free 2-isopropylmalate and 2-isobutylmalate, which were not detected in the control, as well as a large increase in the content of glycosylated 4-methylvaleric acid and glycosylated 5-methylhexanoic acid (Fig. 4). These results indicate

that when *PcIBMS1* is overexpressed in *N. benthamiana*, the resulting protein can also act with 2-oxoisovalerate, despite its high K_m value for this substrate, and also with 5-methyl-2-oxohexanoate. It further indicates that endogenous *N. benthamiana* proteins can complete the elongation cycle with 2-isobutylmalate and longer intermediates. *N. benthamiana* is known to synthesize such longer branched chain fatty acids as part of its acylsugar biosynthesis in glandular trichomes (Kroumova & Wagner, 2003; Slocombe *et al.*, 2008), which is a possible explanation why we detected low levels of glycosylated 4-methylvaleric acid and glycosylated 5-methylhexanoic acid in control *N. benthamiana* leaves.

Met¹³² amino acid site in PcIBMS1 affects its substrate preference

MAM structure of a MAMS from *Brassica juncea*, designated as BjMAM1-A, was determined by X-ray crystallography (Kumar *et al.*, 2019). Seven amino acid sites that may affect substrate preference of IPMS and MAMS were identified. Our sequence comparisons showed that one of the seven corresponding sites is substituted by Met¹³² in PcIBMS1 (Fig. 5). This site corresponds to Leu¹⁴³ in Mt-IPMS that was shown to have hydrophobic interaction with one of the two methyl groups of 2-oxoisovalerate by RCSB PDB Ligand Explorer 3.5 (Koon *et al.*, 2004; de Kraker & Gershenzon, 2011), suggesting that Met¹³² may play a vital role in substrate preference of PcIBMS1. PcIBMS1-M132L mutant obtained by site-directed mutagenesis showed much higher affinity and catalytic efficiency towards 2-oxoisovalerate and lower catalytic efficiency towards 4-methyl-2-oxoalate than wild-type PcIBMS1 (Table 3), which is consistent with the prediction of the effect of the Leu-to-Met substitution in the binding pocket in the structural modeling of PcIBMS1 (Fig. S8 and Fig. S9). The reciprocal PcIPMS1 mutant, PcIPMS1-L135M, showed a strong decrease in affinity and catalytic efficiency towards 2-oxoisovalerate compared with the wild-type enzyme, but it still did not have the capability of utilizing the larger substrate 4-methyl-2-oxoalate. Therefore, while Met¹³² plays a critical role in

PcIBMS1 substrate specificity, simply substituting Leu in this position in IPMS, is clearly not sufficient to convert it into IBMS.

PcIBMS1 has altered feedback regulation of enzyme activity

PcIPMS1 is a typical IPMS enzyme which exhibits Leu feedback inhibition both *in vitro* (Fig. 6) and *in planta* (overexpression of *PcIPMS1* in *N. benthamiana* did not lead to an increase in Leu concentrations, Fig. S5). In contrast, our sequence comparisons showed that PcIBMS1 has a wide sequence divergence from typical plant IPMS proteins in both the catalytic N-terminal (Fig. 5) and the C-terminal regulatory domain (Fig. S7). In the latter domain, a deletion of five amino acids in PcIBMS1 was particularly notable (Fig. S7). Feedback inhibition assay indicated that *PcIBMS1* encodes an enzyme that is not subject to Leu inhibition (Fig. 6). The loss of Leu regulation by the sequence divergence and the loss of five amino acids in PcIBMS1 were thus obtained by a different mechanism of the complete loss of the C-terminal domain, which is how Brassica MAMS and Solanaceae SIIPMS3 acquired their insensitivity to Leu inhibition. In addition, the loss of Leu feedback inhibition in PcIBMS1 is consistent with the observation that transient expression of *PcIBMS1* in *N. benthamiana* led to the overproduction of 2-isopropylmalate but transient expression of *PcIPMS1* did not. Further structure analysis and gain-of-function mutations in the C-terminal regulatory domain of PcIBMS1 might provide a better understanding of the mechanism leading to the insensitivity to feedback inhibition of this enzyme.

IPMSs have been recruited and neofunctionalized to specialized metabolism pathways multiple times throughout plant evolution

Several instances of IPMS neofunctionalization and recruitment to specialized metabolism within eudicots have been documented, including Brassicaceae MAMSs involved in biosynthesis of Met-related glucosinolate compounds, apple CMS involved in

flavor esters, and wild tomato *IPMS3* involved in acylsugar biosynthesis (Ning *et al.*, 2015; Kumar *et al.*, 2019; Sugimoto *et al.*, 2021). In these three cases, a loss of the C-terminal allosteric regulatory domain has occurred, which freed these enzymes from Leu inhibition. When the phylogeny of these enzymes is analyzed by standard methodology, they appear to have split from the bona fide IPMS clade before the split of monocot and dicot IPMSs (Fig. 7a). However, this analysis is probably misleading, since the genes encoding these enzymes have undergone a major change involving the loss of a domain and likely accelerated substitutions elsewhere. When only the catalytic domain is used in the phylogenetic analysis, MAMs and CMS still split early from the IPMS clade, but PcIBMS1 now clusters with PcIPMS1, and *wild tomato* IPMS3 groups with other Solanaceae IPMSs (Fig. 7b). This analysis indicates that at least in the case of PcIBMS1, its recruitment and neofunctionalization into the pogostone pathway occurred relatively recently.

Besides the changes in the catalytic domain that are obviously necessary to evolve a new substrate specificity, the uniform loss of Leu inhibition is of interest. Since the end product of the pathways that these enzymes participate in is not Leu, a Leu sensitivity will serve no clear function. But perhaps more importantly, specialized metabolites such as glucosinolates, flavor esters, acylsugars, and pogostone are typically made in large quantities in specific tissues or organs, and being inhibited by related metabolites would be detrimental to achieving such high levels of production.

Based on our results with PcIBMS and the previous research on the origin of MAMs and CMS, it is likely that many more IPMS-derived enzymes responsible for synthesizing substrates for the biosynthesis of specialized metabolism are still to be discovered. In particular, one anticipates such genes in the Solanaceae, where many species synthesize acyl sugars using acyls up to 12 carbons long, and where the α KAE pathway was shown to be involved in *Nicotiana* and *Petunia*. In fact, our detection in *N. benthamiana* leaves (not expressing PcIBMS) of basal levels of 4-methylvaleric acid and 5-methylhexanoic acid (Fig. 4) suggests that *N. benthamiana* has endogenous enzymes related to IBMS.

However, the three IPMS-like genes that can be presently identified in the *N. benthamiana* genome all cluster together with *bona fide* IPMS sequences, whereas PcIBMS1 is in a separate clade (Fig. 7). Furthermore, all three share the motif that is responsible for Leu feedback inhibition (Fig. S10), and the IPMS-specific Leu residue corresponding to Met¹³² in PcIBMS1 (Fig. S11). Therefore, at this point it cannot be determined if any of these three IPMS-like genes from *N. benthamiana* encode an enzyme with IBMS activity, or whether the observed IBMS activity in *N. benthamiana* is due to an enzyme encoded by another gene.

PcIBMS1 catalyzes the committed step in pogostone biosynthesis

The initial elongation step of 4-methylvaleric acid biosynthesis catalyzed by PcIBMS1 is critical for connecting primary and specialized metabolism (Fig. 1). It is of interest that PcIBMS1 has evolved to use 4-methyl-2-oxovalerate as a more efficient substrate than 2-oxoisovalerate, the preferred substrate of typical plant IPMSs, but still retains high enzymatic activities towards the latter substrate (Table 3). Both 2-isobutylmalate and 2-isopropylmalate were detected in *N. benthamiana* leaves transiently expressing PcIBMS1 (Fig. 4), suggesting that 2-oxoisovalerate can be also a natural substrate *in planta*. It is plausible that the initial elongation steps from both 2-oxoisovalerate and 4-methyl-2-oxovalerate in the α KAE route for 4-methylvaleric acid biosynthesis are catalyzed by only PcIBMS1, allowing ultimately high levels of pogostone production from Leu inhibition. Since only one homolog of each of IPMI and IPMDH was screened from *P. cablin* transcriptome database (Table S2), the elongation cycles of α KAE route for 4-methylvaleric acid biosynthesis probably share both enzymes with Leu biosynthetic pathway in plastid.

Four of five screened BCKDH subunits are predicted to locate in mitochondria (Table S2), suggesting that the proposed BCKDH-like reactions in 4-methylvaleric acid formation occurs in mitochondria like leucine catabolism (Fig. 1; Binder, 2010). A recent

characterization of the intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding cassette transporter that is required for transport and metabolism of fatty acids (Lousa *et al.*, 2013) raises the possibility of the involvement of such transporter in transport of 4-methylvaleryl-CoA generated by the action of BCKDH-like reactions in mitochondria into cytosol along with losing the CoA group to yield free 4-methylvaleric acid (Fig. 1). Although the complete pathway leading to pogostone biosynthesis is still unresolved, the characterization of PcIBMS1 will allow for the breeding of high-yield-pogostone *P. cablin* varieties and reconstitution of the pathway leading to biosynthesis of the branched side of pogostone combined with PcAAE2 in a heterologous system (e.g., *N. benthamiana*).

Data availability

Data supporting the finding of this study are available within the article and its supplemental files. The sequence data used in this study can be obtained from NCBI with the following GenBank accession numbers: PcIBMS1, MW413958; PcIPMS1, MW413959.

Acknowledgements

We would like to thank Dr. Deng at Analytical and Testing Center of Chongqing University for their assistance with LC-QTOF-MS analysis and the laser scanning confocal microscopy analysis. This work was supported by the National Natural Sciences Foundation of China (Grant 31970319 to H.X.), Fundamental Research Funds for the Central Universities (2020CDJQY-A077), the National Key Research and Development Program (2018YFD1000800), 100 Talent Program (Chongqing University, 0304001104433 to H.X.). The authors have no conflicts of interest to declare.

Author contributions

This article is protected by copyright. All rights reserved

Haiyang Xu designed the experiments; Chu Wang, Ying Wang, Jing Chen, Lang Liu and Zhengguo Li conducted the experiments, analyzed data or provided material; Mingxia Yang and Chengyuan Wang conducted and analyzed the protein modeling. Eran Pichersky and Haiyang Xu wrote the article; all authors edited the article.

References

- Binder S, Knill T, Schuster J. 2007.** Branched-chain amino acid metabolism in higher plants. *Physiol Plant* **129**(1): 68-78.
- Binder S. 2010.** Branched-Chain Amino Acid Metabolism in *Arabidopsis thaliana*. *Arabidopsis Book* **8**: e0137.
- Cao ZX, Yang YT, Yu S, Li YZ, Wang WW, Huang J, Xie XF, Xiong L, Lei S, Peng C. 2017.** Pogostone induces autophagy and apoptosis involving PI3K/Akt/mTOR axis in human colorectal carcinoma HCT116 cells. *J Ethnopharmacol* **202**: 20-27.
- Chen HM, Liao HJ, Liu YH, Zheng YF, Wu XL, Su ZQ, Zhang X, Lai ZQ, Lai XP, Lin ZX, et al. 2015.** Protective effects of pogostone from *Pogostemonis Herba* against ethanol-induced gastric ulcer in rats. *Fitoterapia* **100**: 110-117.
- Chen J, Liu L, Wang Y, Li Z, Wang G, Kraus GA, Pichersky E, Xu H. 2021.** Characterization of a Cytosolic Acyl-Activating Enzyme Catalyzing the Formation of 4-Methylvaleryl-CoA for Pogostone Biosynthesis in *Pogostemon cablin*. *Plant Cell Physiol* **62**(10): 1556-1571.
- Chen JR, Xie XF, Li MT, Xiong QY, Li GM, Zhang HQ, Chen GR, Xin X, Yin YP, Fu P, et al. 2021.** Pharmacological activities and mechanisms of action of *Pogostemon cablin* Benth: a review. *Chinese Med* **16**(1): 5.
- de Carvalho LP, Argyrou A, Blanchard JS. 2005.** Slow-onset feedback inhibition: inhibition of *Mycobacterium tuberculosis* alpha-isopropylmalate synthase by L-leucine. *J Am Chem Soc* **127**(28): 10004-10005.

- de Kraker JW, Gershenzon J. 2011.** From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in *Arabidopsis*. *Plant Cell* **23**(1): 38-53.
- de Kraker JW, Luck K, Textor S, Tokuhisa JG, Gershenzon J. 2007.** Two *Arabidopsis* genes (*IPMS1* and *IPMS2*) encode isopropylmalate synthase, the branchpoint step in the biosynthesis of leucine. *Plant Physiol* **143**(2): 970-986.
- de Quiros HC, Magrath R, McCallum D, Kroymann J, Scnabelrauch D, Mitchell-Olds T, Mithen R. 2000.** alpha-keto acid elongation and glucosinolate biosynthesis in *Arabidopsis thaliana*. *Theor Appl Genet* **101**(3): 429-437.
- Edgar RC. 2004.** MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform* **5**: 113.
- Halkier BA, Gershenzon J. 2006.** Biology and biochemistry of glucosinolates. *Annu Rev Plant Biol* **57**: 303-333.
- He Y, Cheng J, He Y, Yang B, Cheng Y, Yang C, Zhang H, Wang Z. 2019.** Influence of isopropylmalate synthase OsIPMS1 on seed vigour associated with amino acid and energy metabolism in rice. *Plant Biotechnol J* **17**(2): 322-337.
- Hu G, Peng C, Xie X, Zhang S, Cao X. 2017.** Availability, Pharmaceuticals, Security, Pharmacokinetics, and Pharmacological Activities of Patchouli Alcohol. *Evid Based Complement Alternat Med* **2017**: 4850612.
- Huang SH, Xian JD, Kong SZ, Li YC, Xie JH, Lin J, Chen JN, Wang HF, Su ZR. 2014.** Insecticidal activity of pogostone against *Spodoptera litura* and *Spodoptera exigua* (Lepidoptera: Noctuidae). *Pest Manag Sci* **70**(3): 510-516.
- Jones DT, Taylor WR, Thornton JM. 1992.** The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**(3): 275-282.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, et al. 2021.** Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873): 583-589.

- Koon N, Squire CJ, Baker EN. 2004.** Crystal structure of LeuA from *Mycobacterium tuberculosis*, a key enzyme in leucine biosynthesis. *Proc Natl Acad Sci U S A* **101**(22): 8295-8300.
- Kroumova AB, Wagner GJ. 2003.** Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants. ■ *Planta* **216**(6): 1013-1021.
- Kumar R, Lee SG, Augustine R, Reichelt M, Vassao DG, Palavalli MH, Allen A, Gershenzon J, Jez JM, Bisht NC. 2019.** Molecular Basis of the Evolution of Methylthioalkylmalate Synthase and the Diversity of Methionine-Derived Glucosinolates. *Plant Cell* **31**(7): 1633-1647.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018.** MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**(6): 1547-1549.
- Li-Beisson Y, Shorrosh B, Beisson F, Andersson MX, Arondel V, Bates PD, Baud S, Bird D, Debono A, Durrett TP, et al. 2013.** Acyl-lipid metabolism. *Arabidopsis Book* **11**: e0161.
- Li CW, Wu XL, Zhao XN, Su ZQ, Chen HM, Wang XF, Zhang XJ, Zeng HF, Chen JN, Li YC, et al. 2013.** Anti-Inflammatory Property of the Ethanol Extract of the Root and Rhizome of *Pogostemon cablin* (Blanco) Benth. *Sci World J* **2013**: 434151.
- Li W, Zhou F, Pichersky E. 2018.** Jasmone Hydroxylase, a Key Enzyme in the Synthesis of the Alcohol Moiety of Pyrethrin Insecticides. *Plant Physiol* **177**(4): 1498-1509.
- Li YC, Liang HC, Chen HM, Tan LR, Yi YY, Qin Z, Zhang WM, Wu DW, Li CW, Lin RF, et al. 2012.** Anti-Candida albicans activity and pharmacokinetics of pogostone isolated from *Pogostemonis Herba*. *Phytomedicine* **20**(1): 77-83.
- Li YC, Xian YF, Su ZR, Ip SP, Xie JH, Liao JB, Wu DW, Li CW, Chen JN, Lin ZX, et al. 2014.** Pogostone suppresses proinflammatory mediator production and

- protects against endotoxic shock in mice. *J Ethnopharmacol* **157**: 212-221.
- Lousa CD, van Roermund CWT, Postis VLG, Dietrich D, Kerr ID, Wanders RJA, Baldwin SA, Baker A, Theodoulou FL. 2013.** Intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding cassette transporter is required for transport and metabolism of fatty acids. *Proc Natl Acad Sci U S A* **110**(4): 1279-1284.
- Luchesi LA, Paulus D, Busso C, Frata MT, Oliveira JB. 2020.** Chemical composition, antifungal and antioxidant activity of essential oils from *Baccharis dracunculifolia* and *Pogostemon cablin* against *Fusarium graminearum*. *Nat Prod Res* **34**: 1-4.
- Melendez-Hevia E, Waddell TG, Cascante M. 1996.** The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J Mol Evol* **43**(3): 293-303.
- Ning J, Moghe GD, Leong B, Kim J, Ofner I, Wang Z, Adams C, Jones AD, Zamir D, Last RL. 2015.** A Feedback-Insensitive Isopropylmalate Synthase Affects Acylsugar Composition in Cultivated and Wild Tomato. *Plant Physiol* **169**(3): 1821-1835.
- Ohlrogge J, Browse J. 1995.** Lipid biosynthesis. *Plant Cell* **7**(7): 957-970.
- Peng F, Wan F, Xiong L, Peng C, Dai M, Chen JP. 2014.** *In vitro* and *in vivo* antibacterial activity of Pogostone. *Chin Med J* **127**(23): 4001-4005.
- Slocombe SP, Schauvinhold I, McQuinn RP, Besser K, Welsby NA, Harper A, Aziz N, Li Y, Larson TR, Giovannoni J, et al. 2008.** Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiol* **148**(4): 1830-1846.
- Stout JM, Boubakir Z, Ambrose SJ, Purves RW, Page JE. 2012.** The hexanoyl-CoA precursor for cannabinoid biosynthesis is formed by an acyl-activating enzyme in *Cannabis sativa* trichomes. *Plant J* **71**(3): 353-365.
- Sugimoto N, Engelgau P, Jones AD, Song J, Beaudry R. 2021.** Citramalate synthase yields a biosynthetic pathway for isoleucine and straight- and branched-chain ester

- formation in ripening apple fruit. *Proc Natl Acad Sci U S A* **118**(3): e2009988118.
- Wang XF, Huang YF, Wang L, Xu LQ, Yu XT, Liu YH, Li CL, Zhan JYX, Su ZR, Chen JN, et al. 2016.** Photo-protective activity of pogostone against UV-induced skin premature aging in mice. *Exp Gerontol* **77**: 76-86.
- Xing A, Last RL. 2017.** A Regulatory Hierarchy of the Arabidopsis Branched-Chain Amino Acid Metabolic Network. *Plant Cell* **29**(6): 1480-1499.
- Xu H, Moghe GD, Wiegert-Rininger K, Schillmiller AL, Barry CS, Last RL, Pichersky E. 2018.** Coexpression Analysis Identifies Two Oxidoreductases Involved in the Biosynthesis of the Monoterpene Acid Moiety of Natural Pyrethrin Insecticides in *Tanacetum cinerariifolium*. *Plant Physiol* **176**(1): 524-537.
- Xu H, Li W, Schillmiller AL, van Eekelen H, de Vos RCH, Jongsma MA, Pichersky E. 2019.** Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in *Nicotiana benthamiana*. *New Phytol* **223**(2): 751-765.
- Xu H, Zhang F, Liu B, Huhman DV, Sumner LW, Dixon RA, Wang G. 2013.** Characterization of the formation of branched short-chain fatty acid:CoAs for bitter acid biosynthesis in hop glandular trichomes. *Mol Plant* **6**(4): 1301-1317.
- Yang T, Stoopen G, Yalpani N, Vervoort J, de Vos R, Voster A, Verstappen FWA, Bouwmeester HJ, Jongsma MA. 2011.** Metabolic engineering of geranic acid in maize to achieve fungal resistance is compromised by novel glycosylation patterns. *Metab Eng* **13**(4): 414-425.
- Ye Q, Ling Q, Shen J, Shi L, Chen J, Yang T, Hou Z, Zhao J, Zhou H. 2021.** Protective effect of pogostone on murine norovirus infected-RAW264.7 macrophages through inhibition of NF-kappaB/NLRP3-dependent pyroptosis. *J Ethnopharmacol* **278**: 114250.
- Yoo SD, Cho YH, Sheen J. 2007.** Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* **2**(7): 1565-1572.

Zheng W, Zhou XG, Wuyun QG, Pearce R, Li Y, Zhang Y. 2020. FUpred: detecting protein domains through deep-learning-based contact map prediction. *Bioinformatics* 36(12): 3749-3757.

Zhou XG, Hu J, Zhang CX, Zhang GJ, Zhang Y. 2019. Assembling multidomain protein structures through analogous global structural alignments. *Proc Natl Acad Sci U S A* 116(32): 15930-15938.

Figure Legends:

Fig. 1 The proposed biosynthetic pathway of 4-methylvaleric acid for pogostone biosynthesis in *Pogostemon cablin*. The anabolism and catabolism of branched chain amino acid leucine are well documented and occur in plastid and mitochondria, respectively, in land plants (Binder *et al.*, 2007; Binder, 2010). The anabolism and catabolism of 2-amino-5-oxohexanoate, the leucine analogue with one more carbon than leucine, most likely occurs in the same way as that of leucine. 5-Methyl-2-oxohexanoate, the precursor of 2-amino-5-oxohexanoate biosynthesis, is elongated by one carbon atom in plastid through a α -keto acid elongation (α KAE) pathway from 4-Methyl-2-oxovalerate, which is derived from anabolism of leucine through the α KAE pathway (Binder, 2010). The one-carbon elongation process for 5-Methyl-2-oxohexanoate through α KAE route is performed in three sequential chemical reactions: 2-oxo acid substrate 4-methyl-2-oxovalerate first undergoes an aldol-type condensation with acetyl-CoA by IBMS to give 2-isobutylmalate, which then undergoes isomerization and oxidative decarboxylation by IPMI and IPMDH or their homologies to yield 5-Methyl-2-oxohexanoate in plastid. 5-Methyl-2-oxohexanoate is regenerated from 2-amino-5-oxohexanoate through transamination catalyzed by AT in mitochondria. 4-Methylvaleryl-CoA is formed from 5-Methyl-2-oxohexanoate in a reaction catalyzed by BCKDH in mitochondria, and then was transported into cytosol by a transporter along with losing the CoA group to give free 4-

methylvaleric acid, which is subsequently activated again to CoA derivative by the cytosol PcAAE2 for pogostone biosynthesis (Chen, J *et al.*, 2021). The dashed arrows represent proposed steps remaining to be elucidated. IPMS, 2-Isopropylmalate synthase; IBMS, 2-Isobutylmalate synthase; IPMI, Isopropylmalate isomerase; IPMDH, 3-Isopropylmalate dehydrogenase; AT, Aminotransferase; BCKDH, Branched-chain keto acid dehydrogenase; PcAAE2, Acyl-activating enzyme 2 from *Pogostemon cablin*.

Fig. 2 GC-MS analysis of stable isotope incorporation into pogostone from *Pogostemon cablin* leaves fed with 4-methylvaleric-d₁₁ acid. (a) The chemical structures of pogostone from 4-methylvaleric acid (4MVA) and 4-methylvaleric-d₁₁ acid (4MVA-d₁₁). (b) GC-MS analysis of *P. cablin* five-week-old top leaf (5W-TLeaf) fed with 4MVA or 4MVA-d₁₁. The total ion chromatograms (TIC) (left) and extracted ion chromatograms of *m/z* 224 for pogostone (middle) and of *m/z* 235 for pogostone-d₁₁ (right) are shown. Tetradecane was used as internal standard. The retention times of pogostone and pogostone-d₁₁ are slightly different, most likely due to the slight differences in physical and chemical properties between these two compounds. (c) The mass spectrums in GC-MS analysis of standard pogostone (upper left), pogostone from *P. cablin* 5W-TLeaf fed with 4MVA (upper right) and pogostone-d₁₁ from *P. cablin* 5W-TLeaf fed with 4MVA-d₁₁ (lower left). Note the 11-mass-unit shift of the parent ions of pogostone (224) to pogostone-d₁₁ (235) upon feeding with 4MVA-d₁₁. (d) The relative abundances of pogostone and pogostone-d₁₁ in *P. cablin* 5W-TLeaf fed with 4MVA-d₁₁. The relative abundances in GC-MS analysis of pogostone and pogostone-d₁₁ in *P. cablin* 5W-TLeaf fed with 4MVA-d₁₁ were calculated by peak area normalization to internal standard tetradecane and the relative abundance of pogostone was set as 100%. Data are represented as mean ± SD of three independent biological replicates.

Fig. 3 Subcellular localization of PcIBMS1 and PcIPMS1 in Arabidopsis leaf mesophyll protoplasts revealed by laser confocal microscopy. The signal of free GFP was used as

control and chloroplasts are revealed by red chlorophyll autofluorescence. Scale bar = 10 μm .

Fig. 4 LC-QTOF-MS and GC-MS analyses of *Nicotiana benthamiana* leaves expressing *PcIBMS1* and *PcIPMS1*. (a) LC-QTOF-MS analyses of *N. benthamiana* leaves expressing the two enzymes *PcIBMS1* and *PcIPMS1*. *N. benthamiana* leaves transformed with empty vector (EV) were used as negative control. Extracted ion chromatograms in negative mode of both m/z 175.0612 for 2-isopropylmalate and 189.0768 for 2-isobutylmalate are shown. Standard 2-isopropylmalate is available commercially while standard 2-butylmalate was generated by the *in vitro* enzymatic reactions of purified *PcIBMS1* with 4-methyl-2-oxovalerate as the substrate in the presence of acetyl-CoA. (b) GC-MS analyses of *N. benthamiana* leaves expressing the two enzymes *PcIBMS1* and *PcIPMS1* with or without base treatment. The MTBE extracts of *N. benthamiana* leaf samples with or without base treatment were analyzed by GC-MS and total ion chromatograms are shown. (c) Concentration of 4-methylvaleric acid and 5-methylhexanoic acid in *N. benthamiana* leaves expressing the two enzymes *PcIBMS1* and *PcIPMS1* with or without base treatment. The quantifications were achieved by normalization of the peaks to the internal standard geraniol and comparison with the standard curves of authentic 4-methylvaleric acid and 5-methylhexanoic acid ($n = 3$; means \pm SD). EV, empty vector; FW, fresh weight; N.D., “not detected”.

Fig. 5 Multiple sequence alignment of N-terminal catalytic regions of *Pogostemon cablin*, *Malus domestica* and Arabidopsis IPMSs and *Brassica juncea* MAM1-A. Multiple sequence alignment of the N-terminal catalytic regions of *P. cablin*, *M. domestica* and Arabidopsis IPMSs and BjMAM1-A is shown. Sequence features corresponding to the characterized BjMAM1-A crystal structure are shown (Kumar *et al.*, 2019). Residues in catalytic sites, metal binding sites, CoA binding sites and α -ketoacid binding sites are

marked by asterisk, solid circle symbol, hollow circle symbol and solid diamond symbol, respectively. The black solid diamond symbol depicts α -ketoacid binding sites implicated in substrate selectivity, while red solid diamond symbol are those shown to impact substrate size discrimination (de Kraker & Gershenzon, 2011; Kumar *et al.*, 2019). Protein domain designations are derived from BjMAM1-A crystal structure (Kumar *et al.*, 2019). Black background shows perfectly conserved sequences across the IPMS, while the light blue background shows amino acids in the conserved sites in the selected five typical IPMSs diverge from that in the corresponding sites in PcIBMS1. PcIBMS1 and PcIPMS1 were directly obtained from *P. cablin* transcriptome database. The information of other proteins used in this sequence alignment is shown in Supporting Information Table S5.

Fig. 6 Effect of Leu on the enzymatic activity of PcIBMS1 and PcIPMS1. (a) The effect of Leu on the enzymatic activities of PcIBMS1 towards 4-methyl-2-oxovalerate. (b) The effect of Leu on the enzymatic activities of both PcIBMS1 and PcIPMS1 towards 2-oxoisovalerate. Activities are expressed as a percentage of the activity. The 100% activity was set in the absence of leu. Data represent means \pm SD of three independent biological replicates.

Fig. 7 Phylogenetic analyses of PcIBMS1 homologs, methylthioalkymalate synthases (MAMs) and citramalate synthase (CMS) from selected species. (a) Phylogenetic analyses of whole PcIBMS1 homologs, MAMs and CMS from selected species. The whole protein sequences of PcIBMS1 homologs, MAMs and CMS were used in this phylogenetic analysis. (b) Phylogenetic analyses of the N-terminal domains of PcIBMS1 homologs, MAMs and CMS from selected species. PcIBMS1 and all the IPMSs except SIIPMS3, from which the predicted C-terminal allosteric regulatory domains were removed, were used in this phylogenetic analysis. The proteins naturally lacking the C-terminal allosteric regulatory domain that is presented in typical IPMSs were marked in

red color. For the construction of phylogenetic trees in (a) and (b), all proteins except PcIBMS1 and PcIPMS1 were obtained from the PHYTOZONE v12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>), NCBI website (<https://www.ncbi.nlm.nih.gov/>), TAIR website (<https://www.arabidopsis.org/>) and UniProt website (<https://www.uniprot.org/>) (Table S5). PcIBMS1 and PcIPMS1, the protein sequences of which were directly obtained from *Pogostemon cablin* transcriptome database, were marked in blue color. Sequences were aligned using MUSCLE (Edgar, 2004), and the trees were conducted using the Maximum likelihood algorithm and JTT matrix-based model in MEGA X (Jones *et al.*, 1992; Kumar *et al.*, 2018). Branch point bootstrap values were calculated with 1000 replicates. The percentage of trees in which the associated taxa clustered together is shown next to the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The species were listed along with the protein names.

Supporting Information

Fig. S1 LC-QTOF-MS analysis of stable isotope incorporation into pogostone from *Pogostemon cablin* leaves fed with 4-methylvaleric-d₁₁ acid.

Fig. S2 RT-qPCR analysis of transcript levels of gene candidates involved in 4-methylvaleric acid biosynthesis in different tissues of *Pogostemon cablin* at different stages of development.

Fig. S3 Dependency of PcIBMS1 and PcIPMS1 enzymatic activity on incubation conditions and components.

Fig. S4 *In vitro* analyses of PcIBMS1 and PcIPMS1 activities.

Fig. S5 LC-QTOF-MS analyses of relative abundances of leucine in *N. benthamiana* leaves expressing *PcIBMS1* and *PcIPMS1*.

Fig. S6 Amino acid sequence alignment of PcIBMS1, PcIPMS1, AtIPMS1, OsIPMS1,

SlIPMS3 and MtIPMS.

Fig. S7 Sequence alignment of C-terminal domain of PcIBMS1, PcIPMS1, MtIPMS and other typical plant IPMSs that are subject to Leu feedback inhibition.

Fig. S8 Modeling structures of PcIBMS1 and PcIPMS1.

Fig. S9 Structure modeling and substrate docking of PcIPMS1-L135M and PcIBMS1-M1132L.

Fig. S10 Sequence alignment of C-terminal domains of PcIBMS1, PcIPMS1, MdIPMS1, MdIPMS2, AtIPMS1, AtIPMS2 and the three IPMS-like proteins from *N. benthamiana*.

Fig. S11 Sequence alignment of N-terminal catalytic regions of PcIBMS1, PcIPMS1, MdIPMS1, MdIPMS2, AtIPMS1, AtIPMS2 and the three IPMS-like proteins from *N. benthamiana*.

Table S1 All primers used in this present study.

Table S2 Bioinformatic analysis of gene candidates involved in 4-methylvaleric acid biosynthesis screened from *P. cablin* RNAseq database.

Table S3 Average normalized counts of gene candidates involved in 4-methylvaleric acid biosynthesis and *PcAAE2* in *P. cablin* RNA-seq database.

Table S4 Ranking of the top 10 unique genes in the *P. cablin* RNAseq database by coexpression analysis with the *PcAAE2* gene.

Table S5 GenBank accession or locus numbers of functional IPMSs, MAMs and MdCMS1 from NCBI (plants), TAIR and UniProt (bacteria and yeast) sites used for phylogenetic reconstruction and sequence comparison.

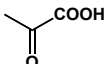
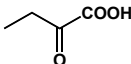
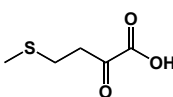
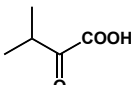
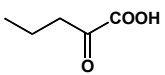
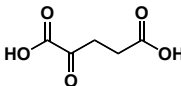
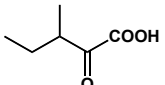
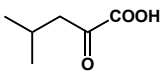
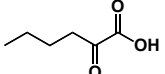
Methods S1 Details on experimental procedures.

Table 1 Ranking of 4-methylvaleric acid biosynthetic pathway gene candidates by coexpression correlation analysis with the *PcAAE2* gene.

Genes	Pearson Correlation Coefficient	<i>P</i> (Two Tailed)
-------	---------------------------------	-----------------------

PcIBMS1	0.969	0.001
PcIPMDHL	0.95	0.004
PcIPMIL SSU	0.732	0.098
PcBCATL3	0.528	0.281
PcIPMIL LSU	0.455	0.365
PcIPMS1	0.451	0.369
PcBCATL2	0.387	0.449
PcBCKDHL E3	0.228	0.664
PcBCKDHL E1 β	-0.066	0.9
PcBCKDHL E1 α 2	-0.143	0.788
PcBCKDHL E2	-0.203	0.699
PcBCKDHL E1 α 1	-0.385	0.452
PcBCATL1	-0.605	0.203

Table 2 Substrate Specificity of Recombinant PcIBMS1 and PcIPMS1.

Substrate	Structure	PcIBMS1	PcIPMS1
Pyruvate		N.D.	N.D.
2-Oxobutyrate		18	10
4-Methylthio-2-oxobutanoic acid		15	N.D.
2-Oxoisovalerate		55	100 ^b
2-Oxovalerate		34	20
2-Oxoglutaric acid		N.D.	N.D.
3-Methyl-2-oxovalerate		8	N.D.
4-Methyl-2-oxovalerate		100 ^a	N.D.
2-Oxohexanoic acid		7	N.D.

These activities were measured with substrate and acetyl CoA at concentration of 0.3 mM and 1 Mm, respectively, in the end-point enzyme assay (DTNB). Data are expressed as relative mean percentages from triplicate independent assays. N.D. represents “not detected or relative activities < 5%”.

^a 100% relative activity of PcIBMS1 corresponds to 1.60 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ on 4-methyl-2-oxovalerate.

^b 100% relative activity of PcIPMS1 corresponds to 1.42 $\mu\text{mol min}^{-1} \text{mg}^{-1}$ on 2-oxoisovalerate.

Table 3 Kinetic Properties of Recombinant PcIBMS1, PcIPMS1, PcIBMS1-M132L and PcIPMS1-L135M.

Enzyme	Substrate	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m ($\text{s}^{-1} \text{M}^{-1}$)
PcIBMS1	2-Oxoisovalerate ^a	1303.3 \pm 127.7	4.5 \pm 0.21	3337 \pm 156
	Acetyl CoA ^b	42.0 \pm 5.8	2.3 \pm 0.07	26361 \pm 761
	4-Methyl-2-oxovalerate ^c	139.8 \pm 39.4	2.4 \pm 0.09	17246 \pm 653
	Acetyl CoA ^d	145.4 \pm 21.2	2.5 \pm 0.02	17124 \pm 106
PcIBMS1-M132L	2-Oxoisovalerate ^a	443.5 \pm 44.4	3.6 \pm 0.16	8007 \pm 352
	Acetyl CoA ^b	57.6 \pm 11.7	3.1 \pm 0.01	53336 \pm 220
	4-Methyl-2-oxovalerate ^c	95.0 \pm 17.9	0.95 \pm 0.01	9978 \pm 87
PcIPMS1	Acetyl CoA ^d	131.9 \pm 6.5	0.95 \pm 0.03	7185 \pm 189
	2-Oxoisovalerate ^a	288.3 \pm 8.2	2.9 \pm 0.04	9940 \pm 126
PcIPMS1-L135M	Acetyl CoA ^b	48.1 \pm 2.8	2.7 \pm 0.02	56094 \pm 429
	2-Oxoisovalerate ^a	621.5 \pm 58.3	1.2 \pm 0.03	1985 \pm 49
L135M	Acetyl CoA ^b	22.7 \pm 1.9	1.1 \pm 0.02	48828 \pm 1072

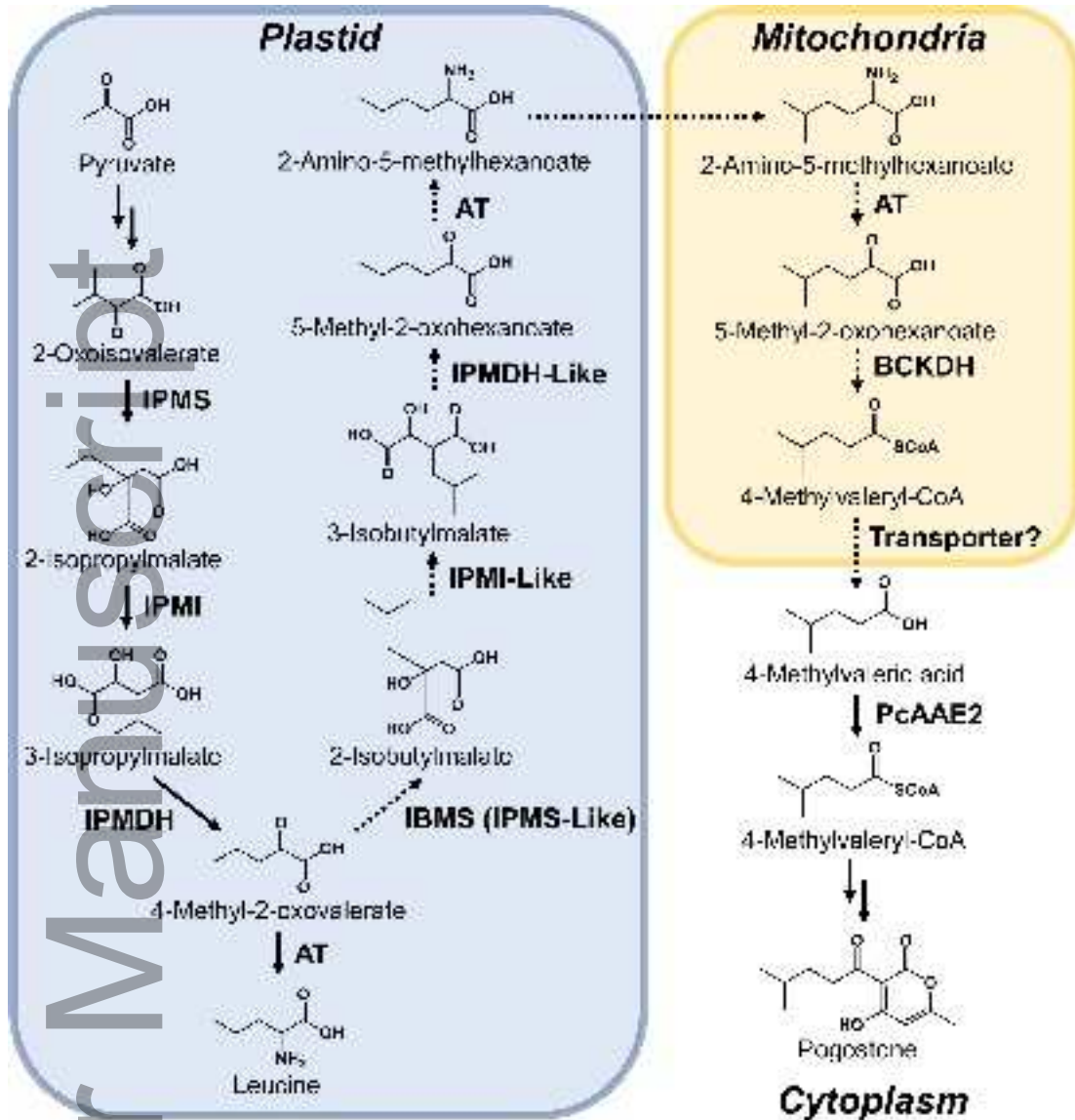
Data are presented as mean \pm SD from triplicate independent assays using end-point enzyme assay (DTNB) method.

^a Kinetic parameters were determined with 1 mM acetyl CoA.

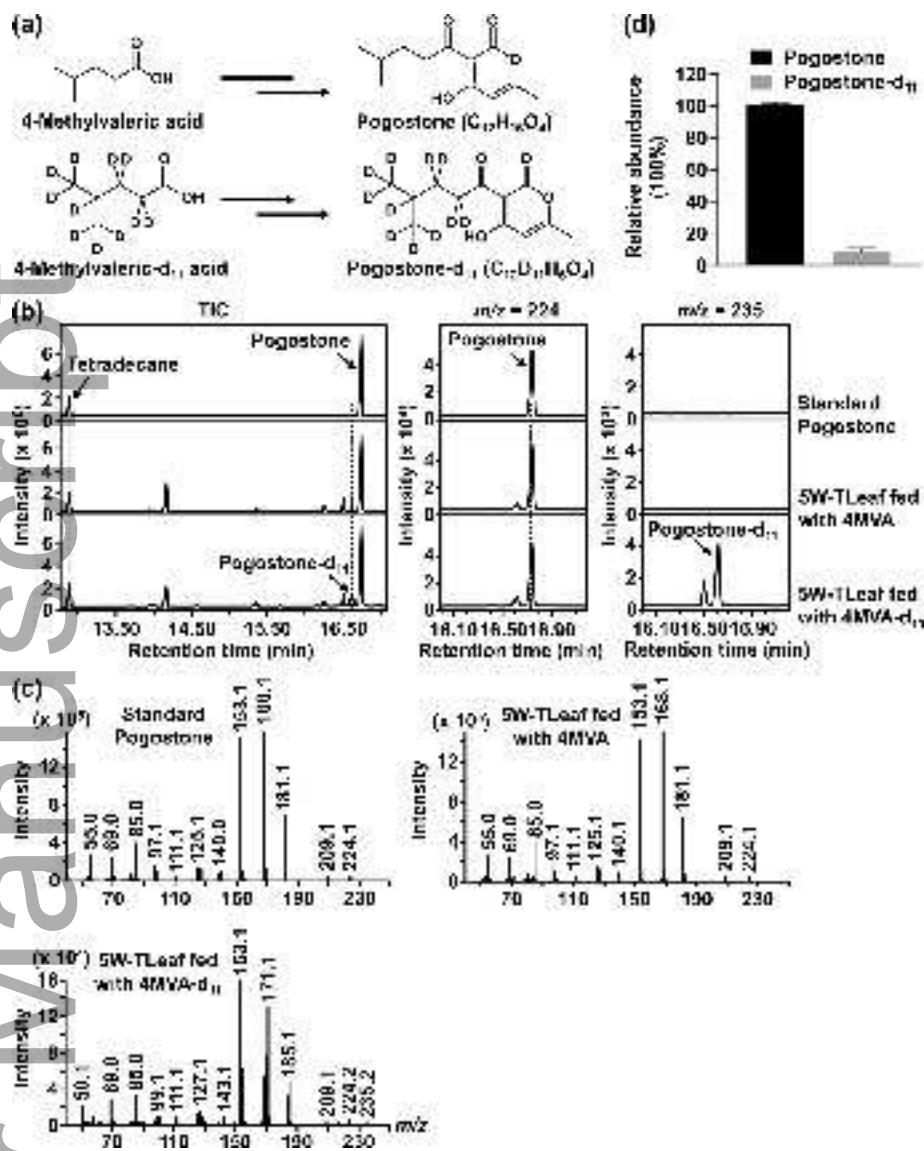
^b Kinetic parameters were determined with 2mM 2-oxoisovalerate.

^c Kinetic parameters were determined with 1mM acetyl CoA.

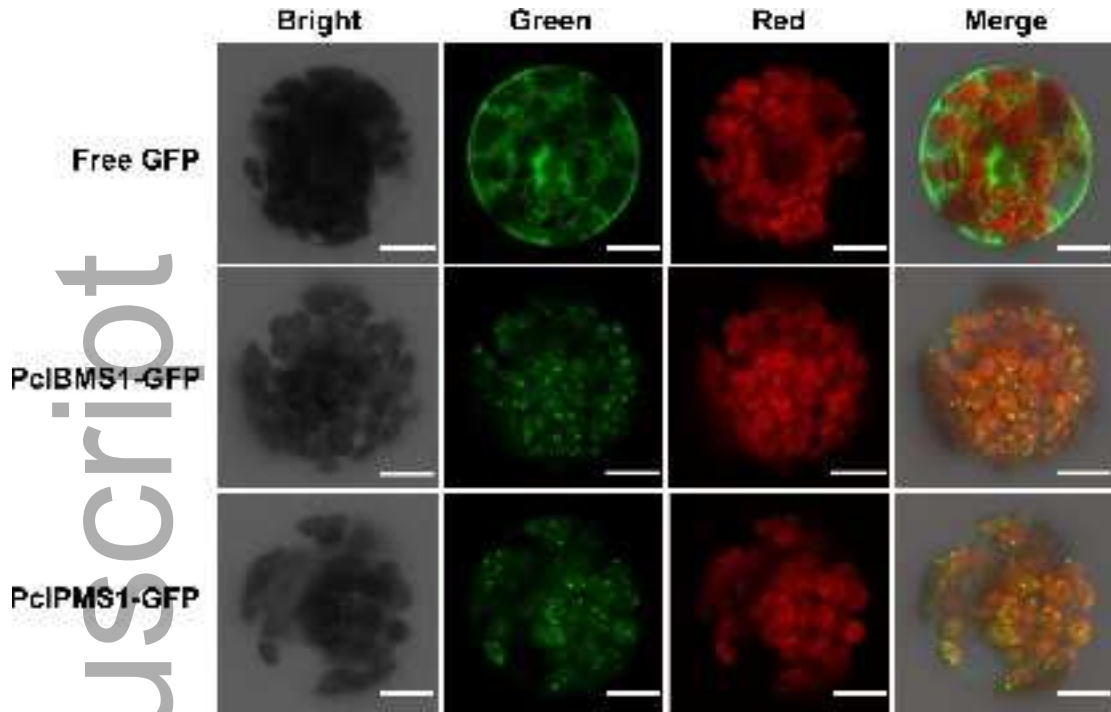
^d Kinetic parameters were determined with 2mM 4-methyl-2-oxovalerate.



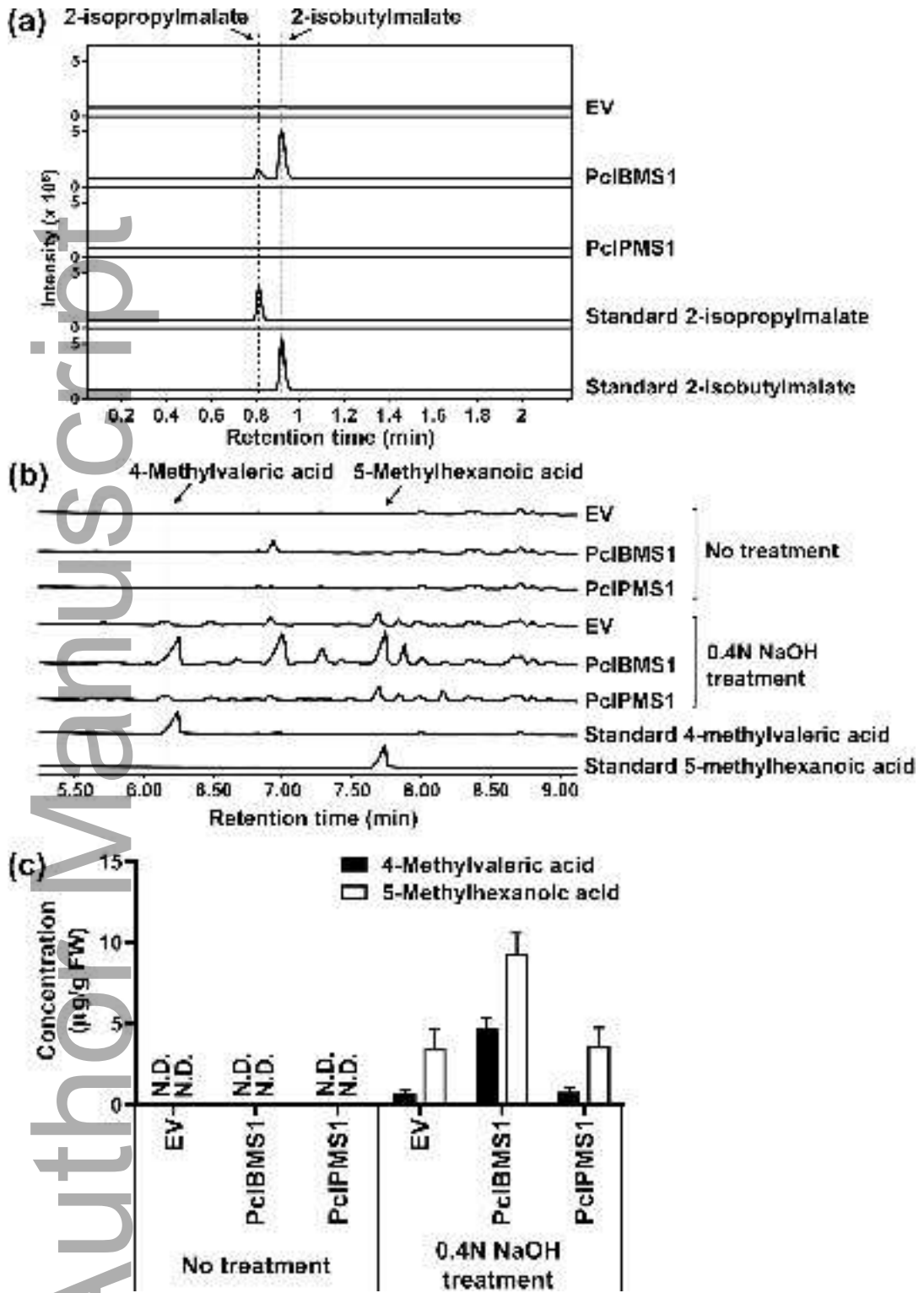
nph_18186_f1.jpg



nph_18186_f2.jpg



nph_18186_f3.jpg



nph_18186_f4.jpg

*8

```

PcIBMS1: ---PSAVIRC SLTRFHY IPNHIPNPNVRI FDTLLRDGEQSPGATM TTKERLITARQLAKLGVDI : 95
PcIPMS1: ---PATAVRC SIARRDY IPNHIPDPKVVRI FDTLLRDGEQSPGATM TTKERLITARQLAKLGVDI : 98
AiiPMS1: DPSLDI DPHTRPRDREY IPNRISDPNIVRV FDTLLRDGEQSPGATM TTKERLITARQLAKLGVDI : 126
AiiPMS2: ASSPLPFAFRR-RPNMIPNRI SDPNVRI FDTLLRDGEQSPGATM TTKERLITARQLAKLGVDI : 124
MiiPMS1: ILCSTQDNPKPTFRDY IPNRISDPNIVRI FDTLLRDGEQSPGASLT TTKERLITARQLAKLGVDI : 124
MiiPMS2: ILCSTQDNPKPTFRDY IPNRISDPNIVRI FDTLLRDGEQSPGASLT TTKERLITARQLAKLGVDI : 129
BjMAM1-A: ETSQTDLKTIVERWDEY IPEKLPDQNVVRFDTLLRDGEQSPGASLT TTKERLITARQLAKLGVDI : 118

```

* 30

```

PcIBMS1: IBACIFGAEENADFAAVSLIARVCGADG--D-HYVVICCLMARCNKRI IRTAWEAWEACAKPRRIET : 157
PcIPMS1: IBACIFGAESEADFEAVLIATEVGNVDS--D-HIVVICCLLARCNKRI IOKSWEAWEAFKPRRIET : 160
AiiPMS1: IEGGFAASKDQFEAVKTLAETVGNVVDENG-YVPIVIGGLSACNKKI IERAWDAWEYAKPRRIET : 190
AiiPMS2: IEGGFAASKDQFEAVRTI AETVGNVVDENG YVPIVIGGLSACNKKI IRTAWEAWEYAKPRRIET : 188
MiiPMS1: IBACIFGAESEDDAEAVMLAKVGNVVDKDC-YVPIVIGGLSACNRSI IQTAWDAWEYAKPRRIET : 188
MiiPMS2: IEGGFAASEKDDAEAVMLAKVGNVVDKDC-YVPIVIGGLSACNKKI IQTAWDAWEYAKPRRIET : 193
BjMAM1-A: MEGGFRVSSSEEFETIQTI AKTVGNVDEETGYIPIVIGVIARSKERI IKAASWEYAKPRRIET : 183

```

* * *

```

PcIBMS1: FFLPSEIHKYETKNNAEIVLRTAKSMVAYARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 222
PcIPMS1: FFLPSEIHKYETKNTTIEVIEKARSMVAYARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 225
AiiPMS1: FLTPSEIHLKYELKTKAEVIEIARSMVRFARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 255
AiiPMS2: FLTPSEIHLKYELKTKAEVIEIARSMVRFARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 253
MiiPMS1: FFLPSEIHLKYELKTKAEVIEIARSMVRFARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 253
MiiPMS2: FFLPSEIHLKYELKTKAEVIEIARSMVRFARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 258
BjMAM1-A: FFLPSEIHLKYELKTKAEVIEIARSMVRFARS LCCNVVFS PRDGRS DREFLYRILGQVIEAC : 248

```

* * *

```

PcIBMS1: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 287
PcIPMS1: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 290
AiiPMS1: RTILNIEDTVGYNLSEDFGQLITDPRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 320
AiiPMS2: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 318
MiiPMS1: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 318
MiiPMS2: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 323
BjMAM1-A: ATPLNIEDTVGYNLSEDFGQLIADIRANTPGIENWII STHCNDLGLSTANTLASACAGARQEV : 313

```

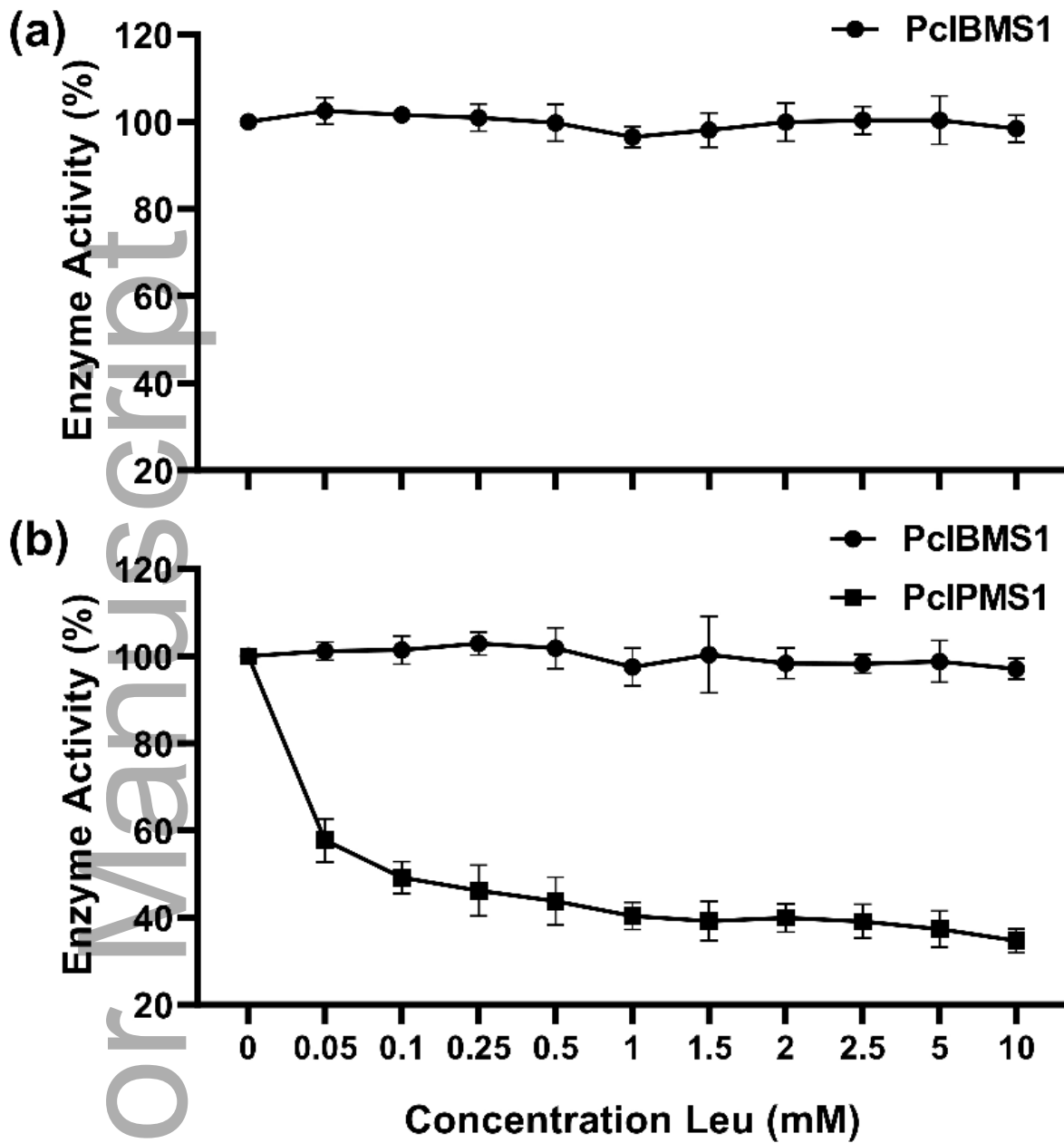
```

PcIBMS1: TINGIGERAGNASLEEVVMTLKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 352
PcIPMS1: TINGIGERAGNASLEEVVMTLKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGT : 355
AiiPMS1: TINGIGERAGNASLEEVVMAIKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 385
AiiPMS2: TINGIGERAGNASLEEVVMAIKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 383
MiiPMS1: TINGIGERAGNASLEEVVMTLKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 383
MiiPMS2: TINGIGERAGNASLEEVVMTLKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 388
BjMAM1-A: TINGIGERAGNASLEEVVMTLKRCEBVLCCGLYPSLNKHIYVTERMVELYTGLAVQBNKAIIVGA : 378

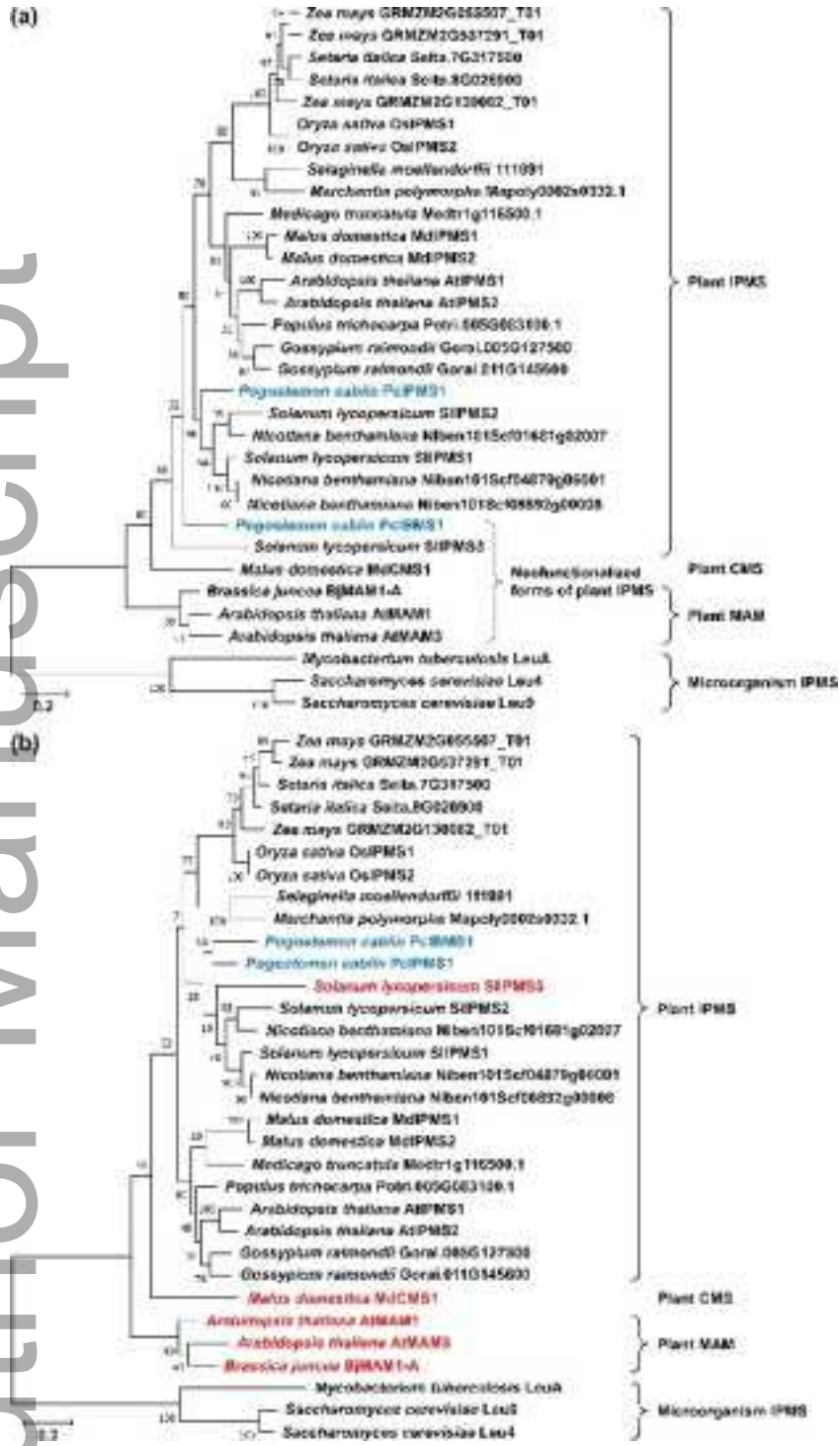
```

* Catalytic sites • Metal binding sites ○ CoA binding sites ◆ β-ketocoid binding sites — TIM barrel

nph_18186_f5.jpg



nph_18186_f6.jpg



nph_18186_f7.jpg