The Plant Cell, Vol. 31: 1633–1647, July 2019, www.plantcell.org © 2019 ASPB.

Molecular Basis of the Evolution of Methylthioalkylmalate Synthase and the Diversity of Methionine-Derived Glucosinolates

Roshan Kumar,[a,1] Soon Goo Lee,[b,1] Rehna Augustine,[a,1] Micheal Reichelt,[c] Daniel G. Vassão,[c] Manoj H. Palavalli,[b] Aron Allen,[b] Jonathan Gershenzon,[c] Joseph M. Jez,[b,2] and Naveen C. Bisht[a,2]

[a] National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi 110067, India
[b] Department of Biology, Washington University, St. Louis, Missouri 63130
[c] Department of Biochemistry, Max Planck Institute for Chemical Ecology, Jena D-07745, Germany

ORCID IDs: 0000-0002-0346-9125 (R.K.); 0000-0002-1541-8587 (S.G.L.); 0000-0002-6174-3372 (R.A.); 0000-0002-6691-6500 (M.R.); 0000-0001-8455-9298 (D.G.V.); 0000-0001-8020-8650 (M.H.P.); 0000-0003-3449-8380 (A.A.); 0000-0002-1812-1551 (J.G.); 0000-0001-7911-9108 (J.M.J.); 0000-0001-7817-2193 (N.C.B.)

The globally cultivated *Brassica* species possess diverse aliphatic glucosinolates, which are important for plant defense and animal nutrition. The committed step in the side chain elongation of methionine-derived aliphatic glucosinolates is catalyzed by methylthioalkylmalate synthase, which likely evolved from the isopropylmalate synthases of leucine biosynthesis. However, the molecular basis for the evolution of methylthioalkylmalate synthase and its generation of natural product diversity in *Brassica* is poorly understood. Here, we show that *Brassica* genomes encode multiple methylthioalkylmalate synthases that have differences in expression profiles and 2-oxo substrate preferences, which account for the diversity of aliphatic glucosinolates across *Brassica* accessions. Analysis of the 2.1 Å resolution x-ray crystal structure of *Brassica juncea* methylthioalkylmalate synthase identified key active site residues responsible for controlling the specificity for different 2-oxo substrates and the determinants of side chain length in aliphatic glucosinolates. Overall, these results provide the evolutionary and biochemical foundation for the diversification of glucosinolate profiles across globally cultivated *Brassica* species, which could be used with ongoing breeding strategies toward the manipulation of beneficial glucosinolate compounds for animal health and plant protection.

INTRODUCTION

The chemical diversity of specialized metabolites in plants helps them to better adapt and survive in a variety of environments (Milo and Last, 2012). The development of new biological functions for proteins from primary metabolism is a key step in the evolution of specialized biosynthesis (Khersonsky and Tawfik, 2010). For example, the close phylogenetic and chemical relationships between the enzymes of the Leu and glucosinolate biosynthesis pathways in plants highlight how new specialized biochemical capacity can evolve from existing pathways (Halkier and Gershenzon, 2006). Glucosinolates are specialized metabolites primarily found in plants of the order Brassicales, which includes agriculturally important *Brassica* crops and the model plant Arabidopsis (*Arabidopsis thaliana*) (Halkier and Gershenzon, 2006). The core glucosinolate structure consists of a β-D-glucosyl residue linked to a sulfonated aldoxime and a variable R-group derived from an α-amino acid (Figure 1A). Intact glucosinolates are inactive, but hydrolysis by β-thioglucoside glucohydrolases (also known as myrosinases) yields active molecules that help defend the plant against pests and pathogens and can contribute to human and animal nutrition (Cartea and Velasco, 2008; Hopkins et al., 2009). The molecular diversity of glucosinolates requires a set of enzymes that alter the length of the possible products of the pathway and modifications of the core chemical structure.

The biosynthesis of aliphatic glucosinolates derived from Met involves three major phases: side chain elongation, core structure formation, and side chain modification (Halkier and Gershenzon, 2006; Cartea and Velasco, 2008; Hopkins et al., 2009; Sønderby et al., 2010). Side chain elongation, the committed step of the process, involves three sequential reactions, in which acetyl-coenzyme-A (CoA) is condensed to a 2-oxo acid substrate derived from Met to form a substituted 2-malate derivative (Halkier and Gershenzon, 2006) (Figure 1B). Methylthioalkylmalate synthase (MAMS) catalyzes this aldol condensation, which lengthens the starting molecule (Kliebenstein et al., 2001; Benderoth et al., 2009). In the construction of these compounds, the number of times MAMS catalyzes elongation leads to products with varied aliphatic side chain lengths. Isomerization leads to a 3-malate derivative, which undergoes oxidative decarboxylation to an elongated 2-oxo acid (He et al., 2009, 2011; Sawada et al., 2009). The elongated 2-oxo acid can undergo up to six cycles of chain elongation or can be transaminated to enter synthesis of the core glucosinolate structure. The side chain length and secondary modifications contribute to the chemical diversity of the nearly 130 glucosinolates identified to date (Agerbirk and Olsen, 2012).

[1] These authors contributed equally to this work.
[2] Address correspondence to jjez@wustl.edu and ncbisht@nipgr.ac.in.
The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) are: Joseph M. Jez (jjez@wustl.edu) and Naveen C. Bisht (ncbisht@nipgr.ac.in).
www.plantcell.org/cgi/doi/10.1105/tpc.19.00046

IN A NUTSHELL

Background: The chemical diversity of specialized metabolites in plants helps them to better adapt and survive in a variety of environments. Glucosinolates are amino-acid derived molecules primarily found in plants of the order Brassicales, which includes agriculturally important Brassica crops and the model plant *Arabidopsis thaliana*. The biosynthesis of aliphatic glucosinolates derived from methionine involves three steps: side-chain elongation, core structure formation, and side-chain modification. Variation in these reactions leads to the chemical diversity of the nearly 130 glucosinolates identified to date. This diversity requires a set of enzymes that alter the length of the possible products of the pathway and modifications of the core chemical structure. Methylthioalkylmalate synthase (MAMS) catalyzes the side-chain elongation of methionine-derived aliphatic glucosinolates and is considered to be the 'gatekeeper' enzyme to the elongation process.

Question: Globally cultivated *Brassica* species possess diverse aliphatic glucosinolates important for plant defense and animal nutrition. The molecular basis for the evolution of MAMS and its generation of natural product diversity is poorly understood in Brassicaceae, owing to their polyploidization and neo-functionalization of paralogs during evolution.

Findings: Our results provide the evolutionary and mechanistic foundation for diversification of different glucosinolate profiles across globally cultivated Brassica crops. We show that the oilseed *Brassica juncea* genome encodes multiple MAMS that fall into two sub-clades with different expression profiles and substrate preferences. In vitro and in vivo work indicates that two related MAMS catalyze the extension of longer 2-oxoacid substrates and the other two enzymes are restricted to shorter 2-oxoacid substrates. X-ray crystallography of MAMs revealed the key amino acid changes that dictate specificity for different sized 2-oxoacid substrates. Biochemical studies show that a handful of changes in the MAMS active site can alter glucosinolate profiles in Brassica crops.

Next steps: The manipulation of beneficial glucosinolate compounds for animal health and plant protection is a major breeding objective in Brassica crops. Our work provides structural and biochemical insights into the key checkpoints shared in the committed steps of primary (i.e., leucine) and specialized (i.e., glucosinolate) metabolic pathways, which can be extrapolated to selectively manipulate nutritionally favored metabolites in crop plants.

Phylogenetic analyses suggest that MAMS are present specifically in the order Brassicales and are evolutionarily derived from isopropylmalate synthases (IPMSs), which catalyze the initial reaction of Leu biosynthesis (de Kraker et al., 2007; de Kraker and Gershenzon, 2011). In Arabidopsis, multiple *MAM* genes are found in the genome, with variations in gene number among accessions (Kroymann et al., 2003; Benderoth et al., 2006). Biochemical studies indicate that AtMAMS1 and AtMAMS2 catalyze the formation of short-chain (C_3–C_5) aliphatic glucosinolates and that AtMAMS3 catalyzes the formation of short- and long-chain (C_6–C_8) glucosinolates (Kroymann et al., 2003; Field et al., 2004; Textor et al., 2004, 2007; Benderoth et al., 2006; Halkier and Gershenzon, 2006). Thus, the substrate specificity of MAMS determines whether products continue through the cycle or enter the pathways that build and modify the core glucosinolate structure. The glucosinolate pathway gatekeeper function and broad substrate preference of MAMS differ from IPMS, which catalyzes only a single elongation step in Leu biosynthesis (Kroymann et al., 2003; Field et al., 2004; Textor et al., 2004, 2007; Benderoth et al., 2006; de Kraker et al., 2007; de Kraker and Gershenzon, 2011).

The molecular basis for MAMS function and how it controls the size of aliphatic glucosinolate is not well understood. Of the various glucosinolate biosynthesis pathways, the Met-derived glucosinolate pathways of Arabidopsis and other Brassicales are among the best studied and offer a model for analyzing the elongation reactions. Here, we examine the evolution of MAMS in agriculturally important *Brassica* species to understand the molecular-genetic determinants of the diversity of Met-derived glucosinolates, shedding light on these important natural plant products.

RESULTS

Molecular Evolution of *MAM* Genes in Cultivated *Brassica* Species

The globally cultivated *Brassica* species generate a wide range of glucosinolates with variable side chain lengths, but only limited information on the MAMS in these plants is available (Wang et al., 2011; Yang et al., 2016). As a first step in examining the product profiles of the MAMS encoded by these genes and the biochemical basis for varied aliphatic glucosinolates in these plants, we retrieved the *MAM*-like sequences from three diploid *Brassica* species based on sequence homology with genes from Arabidopsis.

Two *MAM* genes were identified from each diploid *Brassica* species: wild mustard (*Brassica rapa* [A genome; *BrMAM1* and *BrMAM2*]), black mustard (*Brassica nigra* [B genome; *BnMAM1* and *BnMAM2*]), and wild cabbage (*Brassica oleracea* [C genome; *BoMAM1* and *BoMAM2*]; Supplemental Table 1). The exon-intron organization of the *Brassica MAM* genes indicates that those in *B. nigra* differ from their highly similar *B. rapa* and *B. oleracea* counterparts (Supplemental Figure 1 and Supplemental Table 1). Four *MAM* sequences were identified from the allopolyploid *Brassica juncea* (AB genome; Supplemental Table 1). The *BjMAM* sequences share >99.5% sequence identity with the genes in the progenitor A and B genomes. Phylogenetic analysis suggested that the *MAM* genes from *Brassica* species share closer evolutionary ancestry with *AtMAM1* and *AtMAM2* than with *AtMAM3* and the Leu biosynthesis genes encoding AtIPMS1 and AtIPMS2 (Figure 2A; Supplemental File). The four *B. juncea MAM* genes grouped into two distinct orthologous clades, with *BjMAM1-A* and

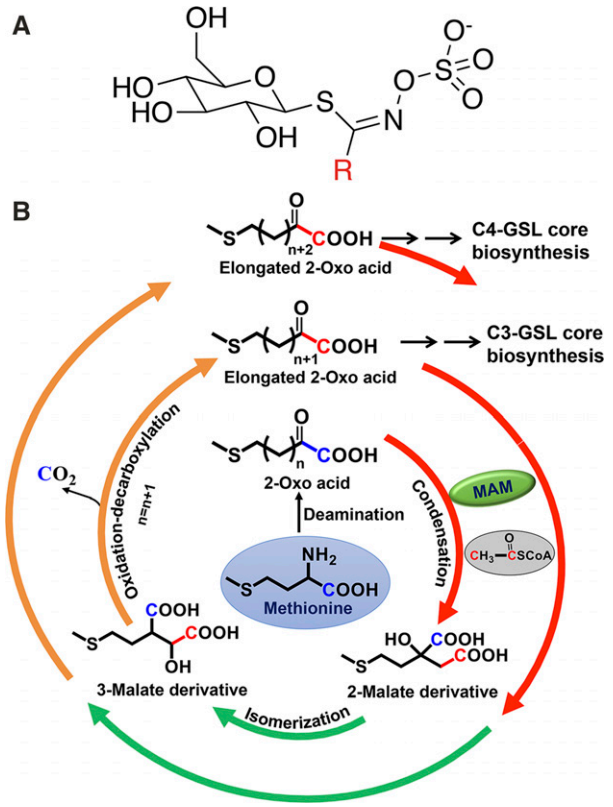

Figure 1. Chemical Diversity in Glucosinolate Biosynthesis.

(A) General glucosinolate structure with the R-group highlighted in red.
(B) Reaction cycle for elongation of Met-derived glucosinolates. Following deamination of Met (blue center) to a 2-oxo acid, MAMS catalyzes condensation with acetyl-CoA (red arrow). The 2-malate derivative isomerizes (green arrow) and undergoes oxidative decarboxylation (orange arrow) to the elongated product, which can either enter additional elongation cycles (red arrow) or be used for glucosinolate (GSL) synthesis.

BjMAM1-B in one group and *BjMAM2-A* and *BjMAM2-B* in the second group. Each clade also contains *Brassica* A, B, and C genome-specific orthologs (Figure 2A). The *B. juncea* MAMS protein sequences share 82% to 85% identity with the Arabidopsis homologs and 83% to 99% identity between each other (Figure 3).

Differential Expression of *BjMAM* Genes

Polyploidy events are associated with variable expression, subgenome expression bias, and functional divergence of the homologous gene pairs within the genome. To evaluate possible transcriptional subfunctionalization of the *MAM* genes in *B. juncea*, we analyzed the expression profile of each gene at different developmental stages in a high-glucosinolate cultivar of *B. juncea* (cv Varuna). The four *BjMAM* genes were all expressed, showing high transcript abundance in glucosinolate-synthesizing tissues such as seedling, leaf, and silique tissue (Figure 2B). Among the four genes, *BjMAM2-A* and *BjMAM2-B* showed similar expression profiles, with the other two genes exhibiting contrasting expression patterns. The abundance of the A subgenome-specific homologs was typically higher than that of the B subgenome-specific homologs in most of the tissue types tested, especially for *BjMAM1-A*, which indicates transcriptional dominance in allopolyploid *B. juncea*.

Biochemical Analysis of Product Profiles and Substrate Preferences of BjMAM Proteins

To investigate their biochemical function, we expressed each MAMS from *B. juncea* in *Escherichia coli* as His-tagged protein and purified the proteins using Ni^{2+}-affinity chromatography (Supplemental Figure 2). We performed enzyme activity assays using liquid chromatography-tandem mass spectrometry (LC-MS/MS) to monitor product formation after incubating each protein with acetyl-CoA (the carbon donor) and 4-methyl-thio-2-oxobutanoic acid (4MTOB), the 2-oxo acid substrate for the first step of Met-derived glucosinolate side chain elongation (Supplemental Figure 3 and Supplemental Table 2). The *B. juncea* MAMS showed maximal activity with acetyl-CoA and 4MTOB at pH 8.0 and 30°C and required either Mn^{2+} or Mg^{2+} for their function, but with specific activities lower than that of AtMAM1 (Table 1).

We examined the ability of the *B. juncea* MAMS to elongate glucosinolates of increasing side chain length using a series of 2-oxo-ω-methylthioalkanoic acids corresponding to C3- to C9-glucosinolate precursors. Each protein catalyzed the condensation of 2-oxo acids involved in the first two elongation cycles, namely 4MTOB and 5-methyl-thio-2-oxo-pentanoate (5MTOP). With 6-methyl-thio-2-oxo-hexanoate (6MTOH) and 2-oxo-nonanoate (a 7MTOH analog), the BjMAM displayed specific activities comparable to that of the Arabidopsis protein. BjMAM2-A elongated 6MTOH with an ~20-fold lower specific activity than either BjMAM1 isoform (Table 1). The use of longer chain 2-oxo acids corresponding to the C7- to C9-glucosinolate precursors did not yield products.

Kinetic analysis using 4MTOB, 5MTOP, and 6MTOH and acetyl-CoA as substrates revealed differences in the selectivity of the *B. juncea* MAMS (Table 2). Although each protein, along with AtMAM1, performed the first condensation reaction using 4MTOB with comparable catalytic efficiencies (k_{cat}/K_m), there were differences with the longer substrates. BjMAM1-A and BjMAM1-B were less than threefold more efficient than AtMAM1 with 5MTOP but were 85- to 1,140-fold more efficient with 5MTOP compared with BjMAM2-A and BjMAM2-B. The BjMAM1 proteins also displayed higher k_{cat}/K_m values for 6MTOH compared with AtMAM1 and the other two *B. juncea* enzymes. This analysis points to the differential evolution of the *B. juncea* MAMS, with all four catalyzing C3-glucosinolate formation but with the group 1 proteins forming the C4-glucosinolate pool.

Subfunctionalization of *BjMAM* Genes in *B. juncea*

To investigate the contribution of the divergent BjMAM proteins to the glucosinolate profile in *B. juncea*, we overexpressed *BjMAM1-A*

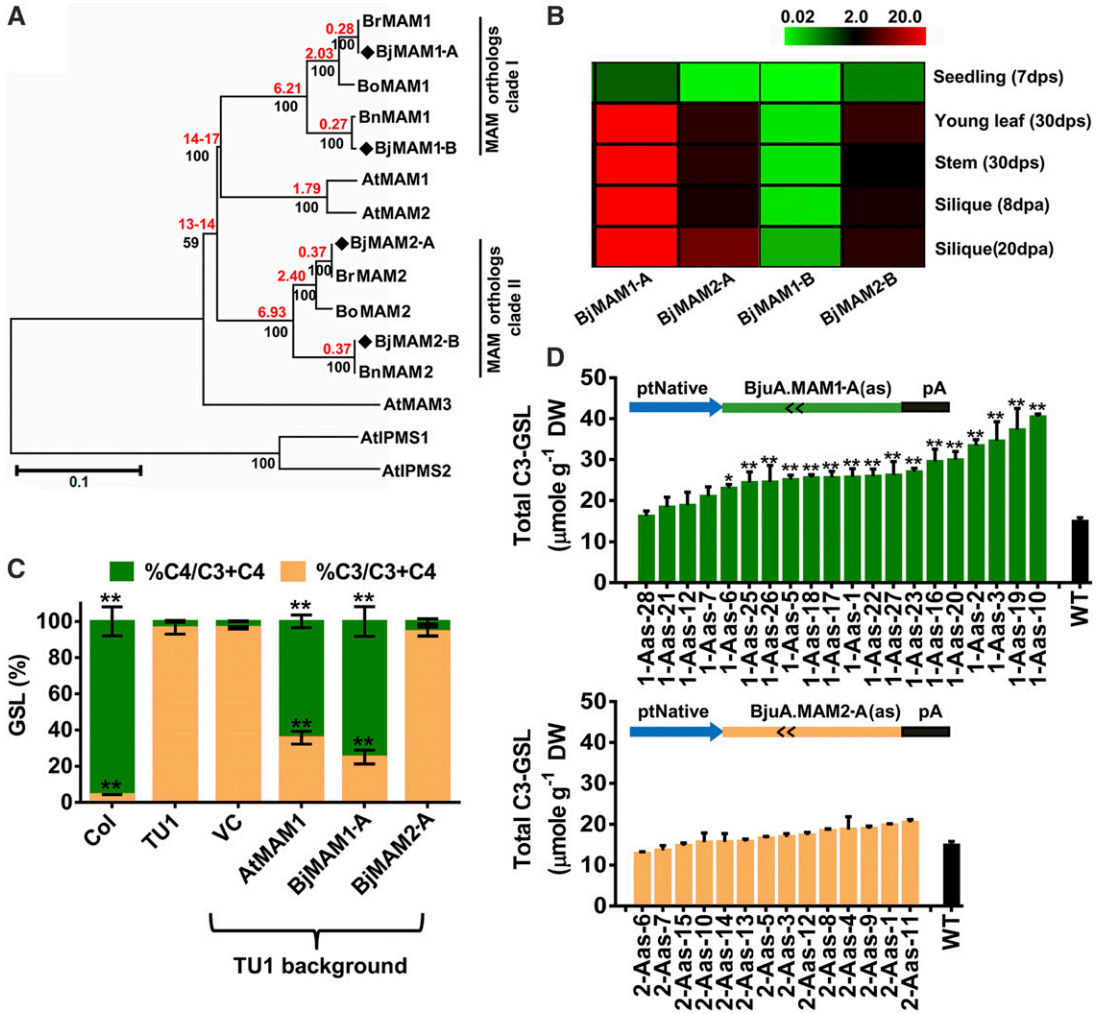

Figure 2. Expression Profile and Function of *B. juncea* MAMS.

(A) Phylogenetic and divergence time analysis of MAMS from *B. rapa* (*Br*), *B. nigra* (*Bn*), *B. oleracea* (*Bo*), *B. juncea* (*Bj*), and Arabidopsis (*At*). The AtIPMSs were the outgroup. Black numbers indicate the percentage of replicate trees in which the associated proteins clustered in the bootstrap test (1000 iterations). Red numbers indicate divergence time (million years ago) calculated from $T = Ks/2\lambda$, where λ is the synonymous mutation rate (1.5×10^{-8} substitutions per site per year for *Brassica* genes). The evolutionary history was inferred using the neighbor-joining method.

(B) Expression profile of *MAM* genes during *B. juncea* development. The heat map was constructed based on relative quantification from qRT-PCR and was normalized with expression of *BjACTIN* (set at 100). qRT-PCR was performed for three independent experiments for each gene and the data averaged. dpa, days post anthesis; dps, days post sowing.

(C) Comparison of C3- and C4-glucosinolate content in Arabidopsis Col-0 (wild type), the Arabidopsis TU1 mutant, and Arabidopsis TU1 mutant transformed with empty vector (VC) or constructs for overexpression of Arabidopsis and *B. juncea* MAM isoforms. The C3-glucosinolate (C3-GSL; peach) and C4-glucosinolate (C4-GSL; green) profiles (shown as percentages) were determined in T2 seeds. Data are summarized in Supplemental Table 3.

(D) Total C3-glucosinolate content in seeds of *B. juncea* transformed with antisense constructs. The BjMAM1-A(as) and BjMAM2-A(as) constructs were transformed in *B. juncea* (cv Varuna), and the glucosinolate profile was analyzed in T2 seeds of independent transgenic lines developed for each construct. Values represent means ± SE (μmol g^{-1} dry weight [DW]) from at least five independent T1 progeny. Data are summarized in Supplemental Table 4.

For **(C)** and **(D)**, asterisks indicate significant differences (*, $P < 0.05$ and **, $P < 0.01$) calculated using one-way ANOVA following Fisher's least significant difference test.

and *BjMAM2-A* in the Arabidopsis *mam1* knockout line background, followed by the analysis of glucosinolate pools (Figure 2C; Supplemental Table 3). Each gene was expressed in the Arabidopsis TU1 homozygous mutant, harboring missense mutations in *AtMAM1* (Kroymann et al., 2001). Compared with the wild type, TU1 mutant seeds show a significant reduction in C4-glucosinolate levels with a concomitant increase in C3-glucosinolate levels. Transformation of the TU1 line with a vector control did not alter glucosinolate content, but complementation with AtMAM1 partially restored the wild-type glucosinolate profile. The expression of *BjMAM1-A* elevated C4-glucosinolate content to levels comparable to those of wild-type and *AtMAM1*-expressing Arabidopsis;

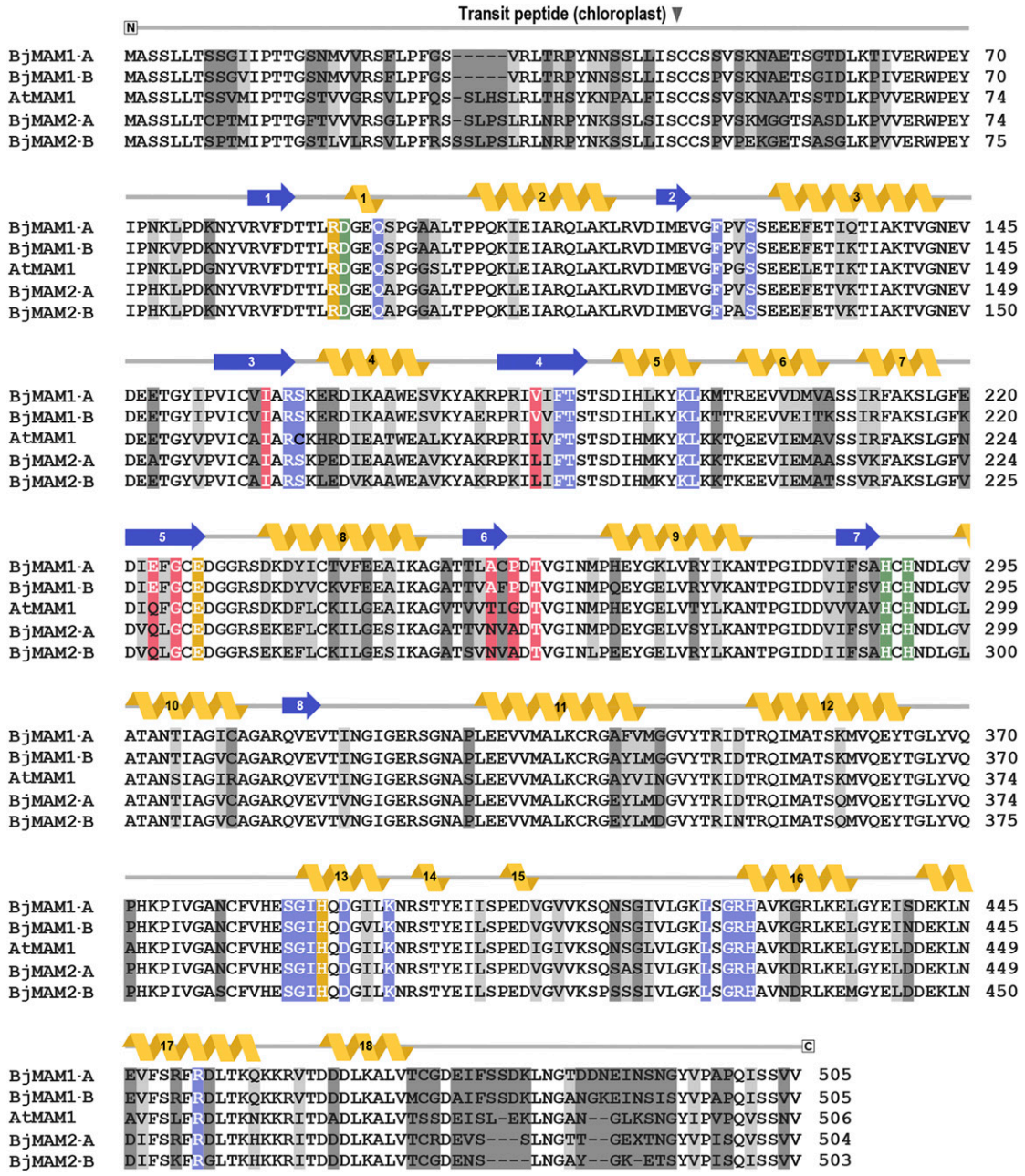

Figure 3. Multiple Sequence Alignment of *B. juncea* and Arabidopsis MAMS.

Multiple sequence alignment of the four *B. juncea* MAMS proteins and Arabidopsis MAMS1 is shown. Secondary structure features corresponding to the structure of BjMAM1-A (α-helices, gold; β-strands, blue) are shown above the alignment. Residues in the metal binding (green), catalytic (yellow), 2-oxo acid binding (red), and CoA binding (blue) sites are highlighted. Dark gray indicates regions of sequence difference and light gray indicates regions of sequence similarity.

however, overexpression of *BjMAM2-A did* not restore the C4-glucosinolate pool.

We also examined the effect of targeting *BjMAM1-A* and *BjMAM2-A* in *B. juncea* using an antisense strategy by expressing the antisense construct under the control of its respective native promoter. Multiple homozygous transgenic lines were generated using either the ptNative:BjMAM1-A(as) or ptNative:BjMAM2-A(as) construct in *B. juncea* (cv Varuna). T2 seeds of at least five progeny per event were analyzed to determine the C3- and C4-glucosinolate profiles (Figure 2D; Supplemental Table 4). The total glucosinolate content in the transgenic lines ranged from 62 to 110 μmol g^{-1} dry weight, suggesting that we achieved variable

Table 1. Comparison of Enzymatic Activities of AtMAM1 and BjMAM Proteins with 2-oxo Acid Substrates.

Protein	4MTOB	5MTOP	6MTOH	2ON
AtMAM1	192 ± 4	125 ± 2	27.9 ± 0.4	0.1 ± 0.01
BjMAM1-A	31.0 ± 0.3	69.6 ± 1.4	19.2 ± 0.4	0.8 ± 0.2
BjMAM1-B	40.1 ± 0.8	148 ± 2	19.5 ± 0.3	1.6 ± 0.1
BjMAM2-A	30.5 ± 0.9	12.9 ± 0.5	0.8 ± 0.1	–
BjMAM2-B	9.7 ± 0.2	1.1 ± 0.2	–	–

Specific activities (nmol min^{-1} mg^{-1} protein) are shown as means ± SE (n = 3). Assays were performed as described in Methods. 2ON, 2-Oxo-Nonanoate. Dashes indicate that activity was not detected.

levels of silencing efficiency. Changes in the C3-glucosinolate pool were gene-dependent. Targeted silencing of *BjMAM1-A* resulted in greater C3-glucosinolate content (i.e., up to 52% of total glucosinolate), but knockdown of *BjMAM2-A* led to C3-glucosinolate accumulation of only up to 24% of total glucosinolate, which was somewhat similar to the levels found in the transformation control (~16% C3-glucosinolates in seeds). Concomitantly, the level of the major C4-glucosinolate (gluconapin) in *B. juncea* was reduced to 36% in the *BjMAM1-A*(as) transgenic lines compared with 75% in the control (Supplemental Table 4).

We examined the expression levels of *BjMAM1* homologs in selected antisense lines harboring the BjMAM1-A(as) and BjMAM2-A(as) constructs by qRT-PCR (Supplemental Figure 4). The BjMAM1-A(as) lines showed pronounced silencing of both *BjMAM1-A* and *BjMAM1-B* homologs. By contrast, the BjMAM2-A(as) construct specifically silenced *BjMAM2-A* and *BjMAM2-B* homologs in most of the selected lines; however, the C3-glucosinolate pools in these lines were not significantly altered. Thus, the gene expression data correlate well with the predicted subfunctionalization of *BjMAM1* and *BjMAM2* genes. Thus, the divergent *BjMAM* genes encode functional proteins that have subfunctionalized roles toward the aliphatic glucosinolate pools in *B. juncea*, which correlates well with their biochemical attributes.

Structural and Functional Basis of MAMS Activity

Although the *B. juncea* MAM proteins share 83 to 99% sequence identity with each other and the homologs from Arabidopsis (Figure 3), the structural basis for the evolution of substrate specificity in Met-derived glucosinolate biosynthesis is unclear. To elucidate the determinants of the substrate preference and product profile of a MAMS, BjMAM1-A was crystallized as a dead-end complex with 4MTOB, Mn^{2+}, and CoA and the 2.1 Å resolution x-ray crystal structure was solved by molecular replacement (Table 3).

The overall structure of BjMAM1-A is dimeric, with each monomer folded into an N-terminal α/β-barrel domain (α1–α12 and β1–β8) and an α-helical C-terminal domain (α13–α18) (Figure 4A). The structure orients the C-terminal side of each α/β-barrel domain on the opposite face of the dimer. A flexible loop extends between α12 and α13 to position the α-helical C-terminal domain of one monomer along the CoA binding site of the adjacent monomer. Sequence and structural comparisons identified BjMAM1-A as a member of the DRE-TIM metallolyase superfamily of enzymes, which catalyze carbon-carbon bond-forming reactions between acetyl-CoA and α-ketoacids (Forouhar et al., 2006). A structural homology search using the Dali server (http://ekhidna2.biocenter.helsinki.fi/dali/) identified IPMS (Koon et al., 2004), citramalate synthase (Ma et al., 2008), and homocitrate synthase (Okada et al., 2010) as the closest structural relatives of BjMAM1-A. Comparison of the monomer structures of BjMAM1-A and the IPMS from the tuberculosis pathogen *Mycobacterium tuberculosis* (Koon et al., 2004) revealed the conservation of the N-terminal α/β-barrel domain and the C-terminal α-helical domain that forms part of the CoA binding site (Supplemental Figure 5A), but it also revealed key structural differences, including extension of the N terminus in IPMS and an additional C-terminal domain that allows for feedback inhibition by Leu in IPMS (Koon et al., 2004; Huisman et al., 2012).

Unambiguous electron density for CoA and 4MTOB (Supplemental Figure 5B) identified the location of the MAMS active site on the C-terminal face of the α/β-barrel domain (Figure 4A). In the structure, 4MTOB and Mn^{2+} bind in an interior pocket (Figure 4B; Supplemental Figure 5C). His-288, His-290, Asp-90, a water molecule, and 4MTOB form the divalent metal binding site. The metal ion helps orient 4MTOB to form additional hydrogen bonds with Arg-89 and Thr-257 and to position the extended side chain of the ligand toward the interior of the pocket and residues in β4, β5, and β6. The pantothenate arm of bound CoA extends into the active site from the surface of the protein (Supplemental Figure 5D), with the reactive thiol placed ~5.8 Å

Table 2. Kinetic Analysis of AtMAM1 and BjMAM Proteins

Protein	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
4MTOB			
AtMAM1	9.6 ± 0.2	987 ± 70	162
BjMAM1-A	1.6 ± 0.1	126 ± 8	212
BjMAM1-B	1.5 ± 0.1	198 ± 18	126
BjMAM2-A	2.1 ± 0.1	69 ± 11	507
BjMAM2-B	0.8 ± 0.1	38 ± 5	351
5MTOP			
AtMAM1	6.3 ± 0.1	123 ± 12	854
BjMAM1-A	3.5 ± 0.1	40 ± 6	1460
BjMAM1-B	7.4 ± 0.1	60 ± 4	2060
BjMAM2-A	0.65 ± 0.03	630 ± 90	17
BjMAM2-B	0.06 ± 0.01	570 ± 430	1.8
6MTOH			
AtMAM1	1.4 ± 0.1	816 ± 43	29
BjMAM1-A	1.0 ± 0.1	112 ± 13	149
BjMAM1-B	1.0 ± 0.1	83 ± 9	201
BjMAM2-A	0.04 ± 0.01	2,500 ± 780	0.3
BjMAM2-B	–	–	–
Acetyl-CoA			
AtMAM1	6.6 ± 0.1	107 ± 21	1030
BjMAM1-A	0.74 ± 0.03	136 ± 28	91
BjMAM1-B	2.6 ± 0.1	232 ± 37	190
BjMAM2-A	1.2 ± 0.01	52 ± 9	384
BjMAM2-B	0.47 ± 0.02	21 ± 10	373

Assays were performed with 0.5 mM acetyl-CoA and varied 2-oxo acid concentrations or varied acetyl-CoA concentrations and 3 mM 4MTOB. Fitted kinetic parameters are shown as means ± SE (n = 4). Dashes indicate that activity was not detected.

Table 3. Summary of Crystallographic Data Collection and Refinement Statistics

Parameter	BjMAM1-A-CoA-4MTOB-Mn^{2+}
Space group	$P2_1$
Cell dimensions	a = 71.75 Å, b = 90.28 Å, c = 93.98 Å; β = 107.7°
Data collection	
Wavelength (Å)	0.98
Resolution range (Å) (highest shell)	47.7–2.10 (2.17–2.10)
Reflections (total/unique)	165,473/59,376
Completeness (highest shell)	92.4% (89.9%)
I/σ (highest shell)	12.1 (1.4)
R_{sym} (highest shell)	6.8% (55.8%)
Model and refinement	
R_{cryst}/R_{free}	15.1%/18.7%
No. protein atoms	6,386
No. waters	430
No. ligand atoms	58
Root mean square deviation, bond lengths (Å)	0.007
Root mean square deviation, bond angles (°)	0.964
Average B-factor ($Å^2$): protein, ligand, water	34.8, 34.9, 41.7
Stereochemistry: favored, allowed, outlier	98.0, 2.0, 0.0%

from the carbon of the C2-carbonyl group of 4MTOB and in proximity to Gln-93 and His-388 (Figure 4B). Extensive charge-charge and ionic interactions extending from the enzyme surface along the length of the ligand lock CoA into the active site (Supplemental Figure 5E). These features define the core catalytic machinery that allows for extension of the Met-derived glucosinolate side chain and provide a view of how the active site architecture controls substrate preference.

The side chain of 4MTOB extends into a pocket with residues from β4 (Val-182, Ile-183, Phe-184), β5 (Glu-223, Phe-224, Glu-227), and β6 (Ala-253, Cys-254, Pro-255), along with Arg-89 and Glu-312, encompassing the substrate (Figure 4C). Sequence comparison of the two BjMAM1 proteins that catalyze extension of longer chain substrates (i.e., 4MTOB, 5MTOP, and 6MTOH) with the two BjMAM2 proteins that only accept 4MTOB highlights key differences in the residues oriented toward the bound substrate (Figure 4D). In particular, Val-182, Glu-223, Ala-253, and Pro-255 in the BjMAM1 proteins are replaced by a Leu, Gln, Asn, and Ala, respectively, in the BjMAM2 proteins. These four residues also form a major portion of the protein surface of the substrate binding site (Figure 4D, stereoview) and likely govern the substrate preference of MAMS.

Evolution of 2-Oxo Acid Substrate Preference

To test the above hypothesis, we mutated each of the four residues in BjMAM1-A to the corresponding residue in BjMAM2-A. We assayed the resulting BjMAM1-A V182L, E223Q, A253N, and P255A point mutants using the substrates 4MTOB and 5MTOP and determined their kinetic parameters (Table 4). The V182L and P255A mutants exhibited only modest differences in kinetic parameters for both substrates versus the wild type, but the E223Q and A253N mutants exhibited ~15-fold increases in K_m values for 4MTOB with a corresponding decrease in catalytic efficiency (k_{cat}/K_m). Using 5MTOP as a substrate, the E223Q and A253N mutants exhibited a drastically reduced catalytic efficiency, with reductions of 18- and 150-fold, respectively. A comparison of the k_{cat}/K_m of the wild type and point mutants of BjMAM1-A for 4MTOB and 5MTOP indicated that only the A253N mutation shifted the substrate profile from favoring 5MTOP toward 4MTOB (Figure 5; Table 4). Experiments using BjMAM2-A and a set of point mutants (L182V, Q223E, N253A, and A255P) in which each position was replaced with the residue found in BjMAM1-A yielded similar results, although the A255P mutant displayed a 530-fold reduction in catalytic efficiency with 4MTOB. Also, the k_{cat}/K_m of the BjMAM2-A N253A mutant was shifted to favor 5MTOP by 2.4-fold (Figure 5; Table 4). Of the four positions that differ between the two groups of *B. juncea* MAMS proteins, the identity of residue 253 has the largest effect on substrate preference.

We then combined all four substitutions to generate the BjMAM1-A V182L/E223Q/A253N/P255A and BjMAM2-A L128V/Q223E/N253A/A255P mutants for biochemical analysis (Figure 5; Table 4). Although neither "4x" mutant achieved the same kinetic properties as the target enzyme, the product specificity profiles of each 4x mutant indicate that all four substitutions are critical. For example, the product profile of BjMAM1-A, which prefers 5MTOP 4.8-fold over 4MTOB, shifted to a 25-fold higher catalytic efficiency for 4MTOB. This suggests that the substitutions likely reduce the size of the pocket to exclude the larger substrate. Similarly, the BjMAM2-A 4x mutant prefers 5MTOP by 6.7-fold versus 4MTOB. By contrast, the wild-type enzyme has a 57-fold higher k_{cat}/K_m for the shorter substrate, which likely results from an enlarged substrate binding pocket in the 4x mutant compared with BjMAM2-A. These results suggest a model for the diversification of aliphatic glucosinolate biosynthesis in a variety of plants, as described below.

DISCUSSION

The evolution of specialized metabolic pathways from core metabolism provides the basis for the chemical diversity of natural products and their resulting biological activities (Khersonsky and Tawfik, 2010; Milo and Last, 2012). In plants, the known glucosinolates share a common scaffold that allows for the incorporation of various groups derived from different amino acids,

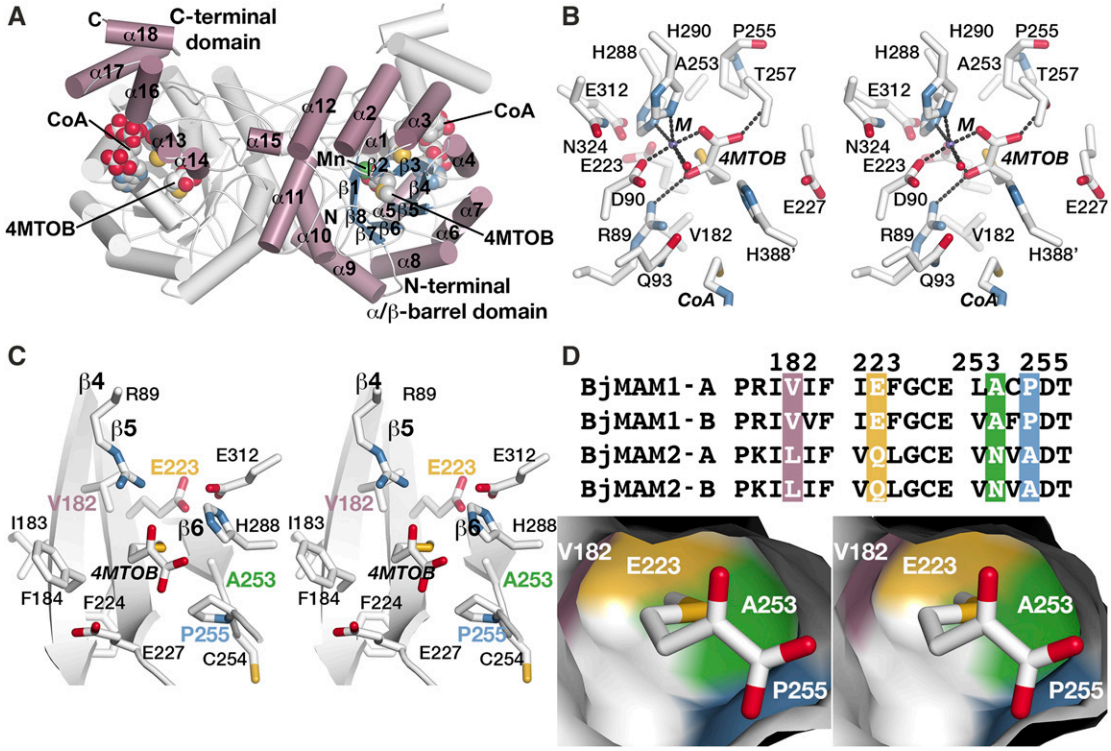

Figure 4. Crystal Structure of *B. juncea* MAMS in Complex with CoA and 4MTOB.

(A) Overall structure of BjMAM1-A. The dimer structure is shown as a ribbon diagram. In one monomer, the α-helices (rose) and β-strands (blue) of the structure are colored and labeled. The N-terminal α/β-barrel and the C-terminal extension domains are also indicated. Bound ligands are shown as space-filling models.

(B) Stereoview of the active site. Residues interacting with 4MTOB and Mn^{2+} (M) and surrounding the site are shown as stick models.

(C) Stereoview of the substrate binding pocket. Residues encompassing 4MTOB are shown as stick models with the three β-strands (β4–β6) forming the interior of the pocket shown as ribbons. Residues that differ in the *B. juncea* MAMS have their labels differentially colored to match panel **(D)**.

(D) Substrate selectivity residues. A targeted sequence alignment highlights residues in β4 (Val-182), β5 (Glu-223), and β6 (Ala-253 and Pro-255) that contact 4MTOB and differ between the two groups of *B. juncea* MAMS. The contour of the substrate binding pocket and the surface associated with each variable residue are shown in the stereoview.

such as Met, Trp, Phe, and Tyr, to generate that diversity (Figure 1A; Halkier and Gershenzon, 2006; Cartea and Velasco, 2008; Hopkins et al., 2009; Sønderby et al., 2010; Agerbirk and Olsen, 2012). The Met-derived glucosinolates are the most abundant form of these natural products in *A. thaliana* and many Brassicaceae crops (Kliebenstein et al., 2001; Kroymann et al., 2003; Field et al., 2004; Textor et al., 2004, 2007; Benderoth et al., 2006; Halkier and Gershenzon, 2006; de Kraker et al., 2007; de Kraker and Gershenzon, 2011; Agerbirk and Olsen, 2012). The elongation step of Met-derived glucosinolate biosynthesis (Figure 1B) is critical for generating the diversity of glucosinolates and for connecting primary and specialized metabolism. Although the evolution of the core features of Met-derived glucosinolate biosynthesis from Leu biosynthesis in Arabidopsis has been noted (Textor et al., 2004, 2007; Benderoth et al., 2006; de Kraker et al., 2007; He et al., 2009, 2011; Sawada et al., 2009; de Kraker and Gershenzon, 2011; Agerbirk and Olsen, 2012), the molecular basis for MAMS activity in plants has remained elusive and the contributions of different MAMS isoforms in *Brassica* species are poorly understood.

Genetic studies have revealed that glucosinolate diversity in oilseed *Brassica* crops follows a complex inheritance pattern and is controlled by multiple loci (Sodhi et al., 2002; Mahmood et al., 2003; Lionneton et al., 2004; Ramchiary et al., 2007; Bisht et al., 2009). Elucidating the molecular basis of glucosinolate content in these crops, as compared with Arabidopsis, has also been complicated by the inherent polyploidy and complex genomic architecture of various *Brassica* species (Lysak et al., 2005; Wang et al., 2011). Comparative mapping and sequence-level studies confirmed the existence of three major whole-genome duplication events in the evolution of core Brassicaceae (Lysak et al., 2005; Franzke et al., 2011). The cultivated *Brassica* species experienced a lineage-specific whole-genome triplication event after their split from Arabidopsis ~13 to 17 million years ago, followed by gene duplication and large-scale chromosomal rearrangements during diploidization (Town et al., 2006; Lysak et al., 2007; Mun et al., 2009). Hybridization of the diploid *Brassica* species *B. rapa* (A genome), *B. nigra* (B), and *B. oleracea* (C) created allotetraploids such as *B. juncea* (AB) and *Brassica napus* (AC). Thus, the genomes of the later species contain multiple homologs of each

Table 4. Steady-state kinetic parameters of the BjMAM mutant proteins

Protein	4MTOB			5MTOP			k_{cat}/K_m 4MTOB:5MTOP
	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	
BjMAM1-A	2.2 ± 0.1	142 ± 37	260	3.1 ± 0.1	41.3 ± 11.1	1260	1:4.8
V182L (1-A)	1.6 ± 0.1	236 ± 50	113	0.8 ± 0.1	24.0 ± 6.5	556	1:4.9
E223Q (1-A)	2.8 ± 0.4	2390 ± 760	20	2.5 ± 0.2	571 ± 167	73	1:3.7
A253N (1-A)	2.1 ± 0.2	2440 ± 654	14	0.6 ± 0.1	1190 ± 173	8.4	1.7:1
P255A (1-A)	2.5 ± 0.2	305 ± 84	137	8.8 ± 0.2	116 ± 15	1260	1:9.2
4x mutant (1-A)	0.8 ± 0.1	258 ± 43	52	0.2 ± 0.1	1560 ± 650	2.1	25:1
BjMAM2-A	2.5 ± 0.1	60.5 ± 18.1	689	0.4 ± 0.1	535 ± 108	12	57:1
L182V (2-A)	0.8 ± 0.1	40.9 ± 4.6	325	0.1 ± 0.01	268 ± 93	6.2	52:1
Q223E (2-A)	3.8 ± 0.3	890 ± 115	71	0.8 ± 0.1	9100 ± 1,290	1.5	47:1
N253A (2-A)	4.6 ± 1.2	3,870 ± 2,200	20	1.0 ± 0.1	353 ± 84	47	1:2.4
A255P (2-A)	0.3 ± 0.1	3840 ± 560	1.3	0.03 ± 0.01	1280 ± 400	0.4	3.3:1
4x mutant (2-A)	0.3 ± 0.1	1150 ± 230	4.3	0.4 ± 0.1	229 ± 91	29	1:6.7

Wild-type and mutant proteins (template form indicated in parentheses) were assayed using either 4MTOB or 5MTOP and acetyl-CoA as substrates. Fitted kinetic parameters are shown as means ± SE ($n = 3$).

glucosinolate biosynthesis gene (Augustine et al., 2013a; Augustine and Bisht, 2015; Meenu et al., 2015), which suggests that genetic interactions of duplicated genes could shape glucosinolate diversity.

Divergence analysis further suggested that the *Brassica* MAMS separated from the AtMAM1/AtMAM2 proteins around the same time as the Arabidopsis-*Brassica* split (Figure 2A; Lysak et al., 2007; Mun et al., 2009). As a consequence of whole-genome duplication and genome fractionation events, each of the mesohexaploid *Brassica* genomes has retained duplicate *MAM* genes (Figure 2A; Supplemental Figure 1 and Supplemental Table 1). Our molecular analysis indicated that the four *MAM* genes in *B. juncea* are conserved and likely evolved through gene duplication and hybridization of two relatively simple *Brassica* genomes while retaining sequence conservation following allopolyploidization of the *B. rapa* and *B. nigra* genomes.

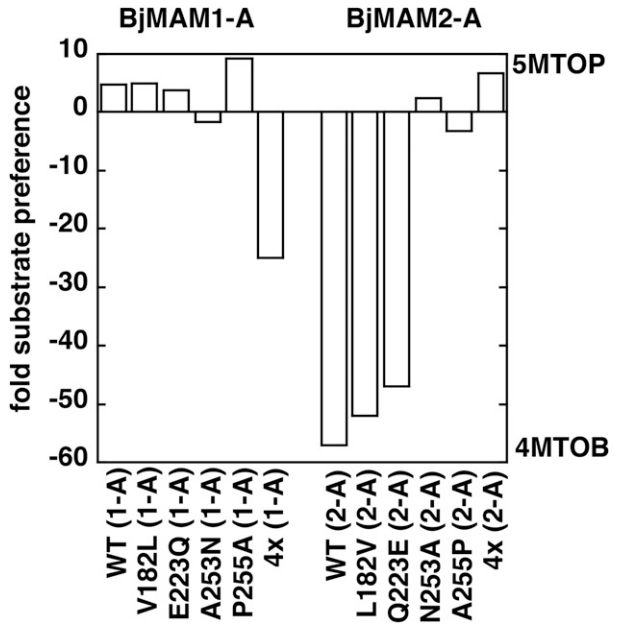

Figure 5. Comparison of Wild-Type and Mutant *B. juncea* MAMS Substrate Preferences.

The fold preferences of BjMAM1-A and BjMAM2-A wild type and mutants for either 5MTOP or 4MTOB, as determined by steady-state kinetic assays (see Table 4), are shown graphically. Fold preference for 5MTOP is plotted on the positive *y* axis, and fold preference for 4MTOB is plotted on the negative *y* axis. The 4x protein combines all of the point mutations. In each case, combination of the four point mutants flips the substrate preference of each MAMS.

The Met-derived glucosinolate profiles differ between Arabidopsis (Supplemental Table 3) and *B. juncea* (Supplemental Table 4). The cultivated *Brassica* crops, such as *B. juncea*, *B. napus*, *B. rapa*, *B. nigra*, and *B. oleracea*, accumulate short-chain (C3–C5) glucosinolates (Wang et al., 2011; Yang et al., 2016). By contrast, Arabidopsis accumulates both short-chain (C3–C5) glucosinolates, which is attributed to AtMAM1 and AtMAM2 isoforms, and long-chain (C6–C9) glucosinolate pools from the activity of the AtMAM3 isoform (Kroymann et al., 2003; Field et al., 2004; Textor et al., 2004, 2007; Benderoth et al., 2006; Halkier and Gershenzon, 2006). Comparison of BjMAM1-A and AtMAM1 showed that the *Brassica* enzyme displays a 10-fold higher activity with 2-oxononanoate (an analog of the C7-glucosinolate substrate) than AtMAM1, although its activity is substantially lower than that for the preferred substrates (Table 1). Thus, the BjMAM1 and AtMAM1 proteins do not prefer long-chain precursors. This was also evident in the BjMAM1-overexpressing lines in the Arabidopsis TU1 mutant background, in which there was no change in the long-chain glucosinolate pools compared with either wild-type or TU1 mutant plants (Supplemental Table 3). The overexpression of BjMAM1-A in the TU1 background altered the profiles of C3- and C4-glucosinolates without affecting the long-chain glucosinolate pools. Similarly, no long-chain glucosinolates were detected in the BjMAM antisense lines (Supplemental Table 4).

The clear differences in the expression profiles (Figure 2B), biochemical activities (Tables 1 and 2), and effect of altered expression in Arabidopsis and *B. juncea* on glucosinolate profiles (Figures 2C and 2D; Supplemental Tables 3 and 4) indicate distinct

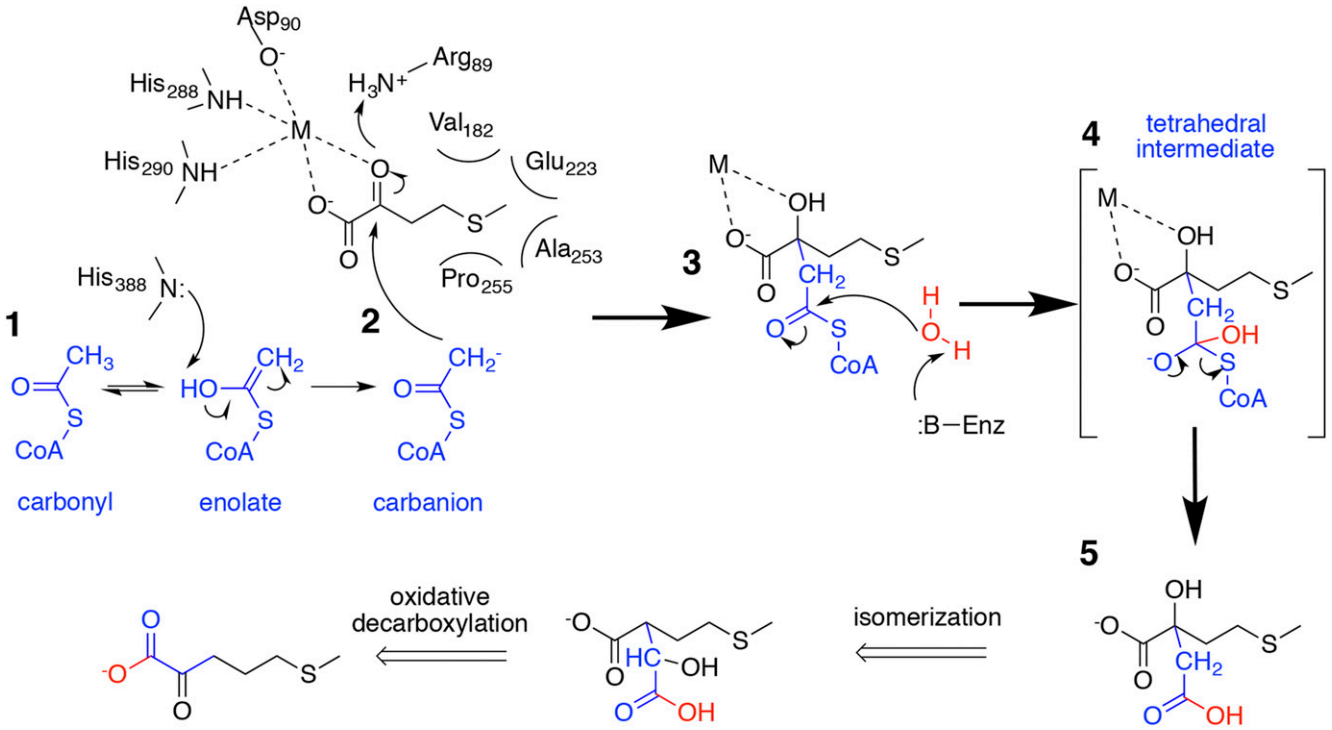

Figure 6. Proposed MAMS Reaction Mechanism.

Extension of the 2-oxo acid substrate requires the activation of the acetate group donated from acetyl-CoA (1). Formation of a carbanion allows for nucleophilic attack on the substrate's C2 carbonyl group (2). Activation of a water molecule (3) leads to the formation of a tetrahedral intermediate (4) and the subsequent release of CoA and the reaction product (5). In the Met-derived glucosinolate biosynthesis cycle, isomerization and oxidative decarboxylation reactions result in the formation of the elongated product, which can either be further elongated or transaminated for the synthesis of the final glucosinolate.

functionalization of the four *B. juncea* MAMS, which are divided into two clades. Although each of the four *B. juncea* MAMS elongated 4MTOB, the two BjMAM1 proteins displayed a clear preference for the C4-glucosinolate precursor 5MTOP (Tables 1 and 2). The distinct activities of the BjMAM1 versus BjMAM2 isoforms were also observed in plants, as transformation of the Arabidopsis TU1 mutant line showed that the expression of BjMAM1-A restored the levels of C4-glucosinolate to nearly those observed in wild-type Arabidopsis, whereas overexpression of BjMAM2-A did not (Figure 2C; Supplemental Table 3). Similarly, the altered glucosinolate profiles of the antisense knockdown lines of BjMAM1-A and BjMAM2-A in *B. juncea* also corresponded to the biochemical activities of these proteins (Figure 2D; Supplemental Table 4). We conclude that divergent *BjMAM* genes encode functional proteins that have subfunctionalized roles that contribute to the broadening of the Met-derived glucosinolate pools in *B. juncea*.

Earlier studies that identified the genes and proteins involved in Met-derived glucosinolate biosynthesis in Arabidopsis highlighted the evolution of not just MAMS but also other pathway enzymes involved in Leu metabolism (Field et al., 2004; Textor et al., 2004, 2007; de Kraker et al., 2007; Benderoth et al., 2009; He et al., 2009, 2011; Sawada et al., 2009; de Kraker and Gershenzon, 2011; Agerbirk and Olsen, 2012). The three-dimensional structure of a MAMS in complex with 4MTOB, Mn^{2+}, and CoA derived in this study (Figure 4; Table 3; Supplemental Figure 5) provides insights into the evolution of glucosinolate biosynthesis and reveals key structural alterations from IPMS.

The overall fold of MAMS (Figure 4A) shares the N-terminal catalytic α/β-barrel domain and the C-terminal α-helical region that forms part of the CoA binding site with IPMS (Supplemental Figure 5A). The two major structural differences between MAMS and IPMS are the loss of the N-terminal extension and the C-terminal Leu binding regulatory domain found in the Leu biosynthesis enzyme. Previous sequence comparisons also noted these changes and identified two amino acid changes in the active site critical for the evolution of MAMS function from IPMS (de Kraker and Gershenzon, 2011). The first substitution is a Ser in IPMS for the Gly that corresponds to Gly-225 of β5 in BjMAM1-A. Removal of the Ser side chain would increase the size of the substrate binding pocket to fit the longer 4MTOB side chain in MAMS compared with the smaller α-ketoisovalerate substrate of IPMS (Koon et al., 2004; Huisman et al., 2012). The second substitution identified was a Pro in IPMS to a Gly in AtMAMS. This residue corresponds to Pro-255 of BjMAM1-A, which is an Ala in BjMAM2-A. It should be noted that the Ser-to-Gly substitution had a greater effect on introducing MAMS activity into IPMS than the second change (de Kraker and Gershenzon, 2011). Although only general product profiles were examined, key point mutations, along with changes in the N- and C-terminal regions, helped shift IPMS toward the evolution of MAMS function.

In both MAMS and IPMS, the catalytic machinery and reaction chemistry are conserved. The BjMAM1-A active site residues that interact with the reactive groups of 4MTOB, the divalent metal, and CoA (Figure 4B; Supplemental Figures 5C and 5E) are invariant across the MAMS from Arabidopsis and *B. juncea* (Figure 3). The core chemistry of MAMS, like other DRE-TIM metallolyases (Forouhar et al., 2006), centers on the metal site and how it orients the substrate for catalysis. In MAMS, Arg-89 and Asp-90 are part of the active site signature of the DRE-TIM metallolyases, in which the Asp coordinates the divalent cation and the Arg contributes to catalysis. The MAMS crystal structure suggests a chemical mechanism for the extension of 2-oxo acids in glucosinolate biosynthesis (Figure 6).

The binding of 4MTOB and Mn^{2+} in a cavity deep in the MAMS fold suggests that these ligands bind to the enzyme first, as binding of CoA and the positioning of the pantothenate arm extending from the protein surface into the pocket would block substrate access without inducing substantial conformational changes. Asp-90, His-288, and His-290 coordinate the interaction of the divalent cation with the carboxylate and C2-carbonyl of 4MTOB and orient the aliphatic side chain toward Val-182, Glu-223, Ala-253, and Pro-255 (Figure 6, step 1). The acetyl group of CoA needs to be activated as a reactant for the extension of the 2-oxo acid. Tautomerization of the acetyl group carbonyl to an enolate allows His-388, which is positioned to act as a general base, to abstract a proton and allow for subsequent carbanion formation. Nucleophilic attack of the carbanion on the C2-carbonyl of the substrate and Arg-89 acting as a general acid (Figure 6, step 2) results in condensation of the acetyl group with 4MTOB. As suggested for other DRE-TIM metallolyases (Forouhar et al., 2006), a water molecule activated by a general base in the active site likely serves as a nucleophile to react with the thioester carbonyl (Figure 6, step 3). The identity of the general base is unclear, although Glu-227 and His-388 are possible candidates in the MAMS active site. Collapse of the resulting tetrahedral intermediate (Figure 6, step 4) leads to the release of free CoA and the formation of the extended 2-malate derivative (Figure 6, step 5).

To continue the elongation cycle (Figures 1B and 6), isomerization to the 3-malate derivative by either isopropylmalate isomerase or a related homolog occurs next (He et al., 2010; Imhof et al., 2014). Oxidative decarboxylation of the 3-malate derivative by a specialized isopropylmalate dehydrogenase yields the elongated 2-oxo acid (He et al., 2009, 2011; Lee et al., 2016). The extended 2-oxo acid can reenter the elongation cycle or be modified by a branched-chain amino acid aminotransferase for entry into the core pathway of glucosinolate synthesis (Schuster et al., 2006). Although MAMS performs the elongation reaction in the pathway, each of the downstream enzymes needs to accommodate substrates of varied length to allow for a complete turn of the synthesis cycle (Figure 1B). This raises an interesting question about the biosynthetic capacity of the other enzymes in the pathway, as each protein of Met-derived glucosinolate biosynthesis could potentially limit the number of complete elongations. For example, if MAMS makes an elongated substrate and any of the subsequence enzymes do not accept a substrate of that length, then the cycle ends. More biochemical data on the biochemical properties of all the steps in the elongation pathway are needed.

The MAMS proteins share key catalytic and ligand binding residues, but the two *B. juncea* subtypes display clear substrate preferences for precursors of C3-glucosinolate (BjMAM2) and C4-glucosinolate (BjMAM1; Figures 2B and 2D; Tables 1 and 2; Supplemental Tables 3 and 4). Analysis of the MAMS crystal structure and sequence comparisons of the four *B. juncea* MAMS identified four positions (Val-182, Glu-223, Ala-253, Pro-255) that varied between the two sets of isoforms (Figures 4C and 4D). Point mutations that interconvert each residue did not drastically alter substrate preference (Figure 5; Table 4). Only the combination of all four changes yielded proteins with 4MTOB:5MTOP preferences comparable to that of the target enzyme. Although substrate preference was altered, the k_{cat}/K_m values of each 4x mutant did not match the values of the respective wild-type enzymes. This suggests that additional changes are likely necessary to optimize catalytic efficiency. Nonetheless, the combined structural and biochemical analyses of *B. juncea* MAMS identified the combination of changes needed to alter elongation substrate preference in the gatekeeper enzyme of Met-derived glucosinolate biosynthesis.

In summary, this study provides important molecular insights into the evolution of MAMS and how specific amino acid changes lead to the diversification of aliphatic glucosinolates. MAMS activity in the Brassiceae evolved through a whole-genome duplication event and natural hybridization, which resulted in isoforms that exhibit diverse expression patterns (subfunctionalization) and biochemical specificities (neofunctionalization). These processes shaped the chemical diversity of Met-derived glucosinolate structures in the extant *Brassica* species. The MAMS crystal structure reveals the biochemical basis for the extension of 2-oxo acids during glucosinolate biosynthesis. The structural and biochemical insights of the key checkpoints shared in the committed steps of primary (i.e., Leu) and specialized (i.e., glucosinolate) metabolic pathways could be stacked with ongoing breeding strategies toward the manipulation of beneficial glucosinolate compounds for animal health and plant protection.

METHODS

Plant Materials and Growth Conditions

The *Brassica* species and cultivars used in this study were grown in a growth chamber under a 10-h-light (400 μmol m^{-2} s^{-1} balanced light spectrum using fluorescent [70-W] and halogen incandescent [100-W] lamps at 24°C) and 14-h-dark (18°C) cycle with 70% relative humidity. For gene isolation and qRT-PCR analysis, different developmental stages (i.e., seedlings, roots, stems, leaves, and siliques at 8 and 20 d post anthesis) were collected, frozen in liquid nitrogen, and stored at −80°C. The Arabidopsis (*Arabidopsis thaliana*) wild type (Col-0) and the *AtMAM1* (At5g23010) missense mutant *TU1* were grown in a growth room set at 22°C under a 16-h/8-h light/dark cycle with 250 μmol m^{-2} s^{-1} balanced light spectrum using fluorescent lamps (40 W) at 40% relative humidity.

Isolation, Cloning, Sequence Analysis, and Expression Analysis of *BjMAM* Homologs

The coding sequences of the *MAM* genes were isolated from *Brassica juncea* (cv Varuna), *Brassica nigra* (cv IC257), and *Brassica rapa* (cv Pusa Gold). Total RNA was isolated using a Spectrum Total RNA Isolation kit (Sigma-Aldrich). RNA (2 μg) was reverse transcribed into cDNA with oligo(dT) primers using a First Strand cDNA Synthesis kit (Applied

Biosystems). PCR amplification of the coding regions was performed using gene-specific primers designed based on the *MAM* (*GSL-ELONG*) gene sequence reported from *Brassica* species (Bisht et al., 2009). The PCR products were cloned into pGEM-T-Easy cloning vector (Promega), sequenced, and analyzed using DNASTAR software (Lasergene).

Multiple sequence alignment of the deduced amino acid sequences of MAMS from Arabidopsis and *Brassica* species was performed using Multalin. Phylogenetic analysis of the deduced MAMS protein sequences from Arabidopsis, *B. rapa*, *B. nigra*, and *Brassica oleracea* (https://phytozome.jgi.doe.gov/pz/portal.html) was performed using the neighbor-joining method in MEGA5.1 with 1000 bootstrap iterations (Supplemental File; Tamura et al., 2011). The AtIPMSs were used as the outgroup. To estimate the divergence time, pairwise alignments of coding DNA sequences of *Brassica*-specific MAMS genes with AtMAM1 was performed, and Ks (synonymous substitution rate) and Ka (nonsynonymous substitution rate) were calculated using the DnaSPv5 program. The divergence time (T) was calculated using the equation $T = Ks/2\lambda$, where λ is the synonymous mutation rate, reported as 1.5×10^{-8} substitutions per site per year for *Brassica* genes (Koch et al., 2000).

Relative expression of the candidate genes was analyzed by qRT-PCR with an ABI-7900HT real-time PCR machine (Applied Biosystems) using the SYBR Green protocol and universal cycling conditions (95°C for 5 min, 40 cycles of 15 s at 95°C and 60°C for 60 s) in a final volume of 20 μL. The *BjACTIN2* gene was used as an endogenous control (Chandna et al., 2012). Relative expression values for each target gene were calculated using the 2^{-ddct} method, in three independent experiments with two technical replicates each. All experiments were performed using samples harvested from homozygous T2 lines.

Protein Expression and Purification

The coding sequences of *AtMAM1* and the four *BjMAM* genes lacking the putative N-terminal signal sequence were amplified and cloned within *Nhe*I/*Not*I sites of the pET-28a expression vector (Novagen). The resulting constructs were transformed and expressed in *Escherichia coli* BL21(DE3) cells grown in Terrific broth with 50 μg mL^{-1} kanamycin until A_{600} was ~0.6. Following the addition of isopropyl-1-thio-β-D-galactopyranoside (1 mM final), the cells were grown overnight at 18°C. After centrifugation, the cell pellet was dissolved in lysis buffer (50 mM Tris-HCl, pH 8.0, 20 mM imidazole, 500 mM NaCl, 1% [v/v] Tween 20, and 10% [v/v] glycerol) with cell lysis by sonication. After clarification, the supernatant was passed over an Ni^{2+}-NTA-agarose column (4°C). The column was washed with lysis buffer lacking Tween 20, and bound protein was eluted in elution buffer (50 mM Tris-HCl, pH 8.0, 250 mM imidazole, 500 mM NaCl, and 10% [v/v] glycerol). Protein concentration was determined using the Bradford method with bovine serum albumin as the standard.

Enzyme Assays

The enzyme assay for the condensation reaction between acetyl-CoA and different 2-oxo acids was performed as previously described (Textor et al., 2007). Steady-state kinetic assays (50 μL) of MAMS used a range of 0.05 to 8 mM 2-oxo acid substrate with 1 mM acetyl-CoA with 100 nM (~250 ng) protein at 30°C for 10 min in Tris-buffered conditions (pH 8.0), which provides for a linear range of product formation. Reactions were quenched by the addition of 150 μL of ethanol. Kinetic parameters were determined by fitting the data to the Michaelis-Menten equation using GraphPad Prism (version 5.0; GraphPad Software).

LC-MS/MS Analysis

A 1-μL aliquot of the final diluted assays was used for LC-MS/MS analysis to quantify malate derivatives produced in the MAMS assays. Chromatography was performed on an Agilent 1200 HPLC system. Separation was achieved on a Zorbax Eclipse XDB-C18 column (50 × 4.6 mm, 1.8 μm; Agilent). Formic acid (0.05%, v/v) in water and acetonitrile were employed as mobile phases A and B, respectively. The elution profile was as follows: 0 to 0.5 min, 5% B; 0.5 to 3.0 min, 5 to 57.8% B; 3.0 to 3.1 min, 57.8 to 100% B; 3.1 to 4 min, 100% B; and 4.1 to 6.5 min, 5% B with a flow rate of 1.1 mL min^{-1} at 25°C. An API3200 tandem mass spectrometer (Applied Biosystems) equipped with a Turbospray ion source was operated in negative ionization mode. The ion spray voltage was maintained at −4500 eV. The turbo gas temperature was set at 700°C. Nebulizing gas was set at 60 p.s.i., curtain gas at 25 p.s.i., heating gas at 60 p.s.i., and collision gas at 7 p.s.i. Multiple reaction monitoring was used to monitor analyte parent ion-to-product ion transition (Supplemental Table 2). Both Q1 and Q3 quadrupoles were maintained at unit resolution. Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing. Malate derivatives were quantified based on an external standard curve of 2-hexyl-malate, applying a theoretical response factor of 1 for all malate derivatives with 2-hexyl-malate synthesized as described previously (Textor et al., 2007).

BjMAM Overexpression Constructs and Complementation of Arabidopsis *AtMAM1* (TU1) Mutant

Full-length coding DNA sequences of *AtMAM1*, *BjMAM1-A*, and *BjMAM2-A* were cloned into the pPZP200:lox(bar) binary vector (Hajdukiewicz et al., 1994) with the *Cauliflower mosaic virus* 35S promoter and *bar* gene as a plant selection marker. Each construct was transferred into *Agrobacterium tumefaciens* strain GV3101 and subsequently into the homozygous missense mutant *AtMAM1* (TU1) genetic background by the floral dip method (Clough and Bent, 1998). Transformants were selected using the herbicide Basta (120 mg L^{-1}; Agrevo) with at least three independent T2 lines for each construct analyzed for glucosinolate content.

BjMAM Knockdown Constructs and *B. juncea* Antisense Lines

Antisense constructs were developed by cloning the partial coding sequences of the target *BjMAM* genes in antisense orientation (within *Nru*I/*Spe*I sites of the binary vector) with expression under the control of either the *BjMAM1-A* or *BjMAM2-A* native promoter (directionally cloned at *Xho*I/*Pst*I sites) in the pPZP200:lox(bar) binary vector (Augustine et al., 2013b). The transformation vectors were mobilized into *A. tumefaciens* strain GV3101 using the freeze-thaw method. Genetic transformation of a high-glucosinolate-containing *B. juncea* (cv Varuna) was performed (Augustine et al., 2013b). The rooted transgenic plants were transferred directly to soil and grown in a containment net-house facility according to the guidelines of the Department of Biotechnology, Government of India. Transgenic plants were selected after spraying with 200 mg mL^{-1} Basta (Agrevo). For each event, Basta segregation analysis of T1 progeny was performed and the plant was selfed and maintained as a homozygous stock.

Analysis of Glucosinolate Content

The transgenic events (T2 homozygous seeds) were analyzed for seed/leaf glucosinolate profiles (Meenu et al., 2015). Briefly, 20 mg of freeze-dried sample with sinalbin added as an internal standard in the extraction solution (70% [v/v] methanol) was used for analysis. Desulfation of glucosinolates was performed overnight using purified sulfatase (25 mg mL^{-1}; Sigma-Aldrich) on a DEAE Sephadex-A25 column. Desulfo-glucosinolates were eluted with 1 mL of HPLC-grade water and 30 μL of eluent was analyzed using a Shimadzu CLASS-VP V6.14 HPLC device and a Luna C18 reverse-phased column (150 × 4.6 mm; 0.5 μm i.d.). A gradient of water (solvent A) and 1% to 19% (v/v) acetonitrile (solvent B) over a period of 25 min for *Brassica* and 30 min for Arabidopsis was used with a flow rate

of 1 mL min^{-1}. Glucosinolates were detected with a UV light detector at A_{229}, and concentrations of individual glucosinolates were calculated in micromoles per gram dry weight relative to the area of the internal standard peak applying their respective response factors (Augustine et al., 2013a). All data are averages of at least three independent experiments ± SE. Statistical analyses were conducted using one-way ANOVA following Fisher's least significant difference test (Supplemental Table 5).

Protein Crystallography

For protein crystallization, BjMAM1-A was further purified by size-exclusion chromatography using an Akta fast protein liquid chromatography device with a Sephadex-200 column equilibrated with 25 mM HEPES, pH 7.5, 100 mM NaCl, and 1 mM DTT. Crystallization was performed at 12°C using the vapor diffusion method in hanging drops of a 1:1 mixture of protein (10 mg mL^{-1}) and crystallization buffer. Crystals of the BjMAM1A-4MTOB-CoA-Mn^{2+} complex were obtained in 100 mM imidazole, 100 mM MES monohydrate (pH 6.5), 30 mM manganese chloride tetrahydrate, 30 mM calcium chloride dihydrate, 20% (v/v) ethylene glycol, and 10% (w/v) PEG-8000. Crystals were stabilized in cryoprotectant (i.e., crystallization solution plus 30% [v/v] glycerol) before flash freezing in liquid nitrogen for data collection at 100 K. Diffraction images were collected at beamline 19ID of the Advanced Photon Source at Argonne National Laboratory. Diffraction data were indexed, integrated, and scaled using HKL3000 (Minor et al., 2006). The MAMS structure was solved by molecular replacement using MOLREP (Vagin and Teplyakov, 1997) implemented in CCP4 (Collaborative Computational Project, Number 4, 1994) using the x-ray crystal structure of the truncated IPMS from *Neisseria meningitidis* (PDB: 3RMJ; Huisman et al., 2012) as a search model. For iterative rounds of model building and refinement, COOT (Emsley et al., 2010) and PHENIX (Adams et al., 2010) were used, respectively. Data collection and refinement statistics are summarized in Table 3.

Site-Directed Mutagenesis and Mutant Protein Analysis

Site-directed mutants of BjMAM1-A and BjMAM2-A were generated using the QuikChange PCR protocol (Agilent Genomics). The fidelity of each pET-28a plasmid was confirmed by sequencing, and each plasmid was transformed into *E. coli* BL21(DE3) for protein expression, purification, and biochemical analysis as described above.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL libraries under the following accession numbers: *BjMAM1-A* (FM161920), *BjMAM2-A* (FM161923), *BjMAM1-B* (FM161916), and *BjMAM2-B* (FM161918). Atomic coordinates and structure factors were deposited in the RCSB Protein Data Bank (PDB: 6E1J).

Supplemental Data

Supplemental Figure 1. Gene structures of *MAM* genes identified in diploid *brassica* species.

Supplemental Figure 2. Expression and purification of recombinant MAMS from *A. thaliana* and *B. juncea*.

Supplemental Figure 3. Analysis of 2-malate derivatives by LC-MS/MS.

Supplemental Figure 4. Expression of *B. juncea MAM* genes in 5-d-old T2 seedlings of selected BjMAM1-A(as) and BjMAM2-A(as) lines.

Supplemental Figure 5. BjMAM1-A domain comparison and ligand binding.

Supplemental Table 1. Summary of *MAM* genes from *A. thaliana* and the cultivated *brassica* species.

Supplemental Table 2. Details of the analysis of malate derivatives by LC-MS/MS using an agilent HPLC 1200/API-3200 instrument in negative ionization mode.

Supplemental Table 3. Average glucosinolate content and profile in T2 seeds of various lines harboring *MAM* constructs in the *A. thaliana MAM1* knockout line (*TU1*) background.

Supplemental Table 4. Average glucosinolate content and profile in seeds of wild-type *B. juncea* (cv varuna) and BjMAM antisense transgenic *B. juncea* lines.

Supplemental Table 5. ANOVA Statistical Analysis.

Supplemental File. Alignment used to produce the phylogenetic tree.

ACKNOWLEDGMENTS

We thank the Central Instrumentation and Plant Growth Facilities at the National Institute for Plant Genome Research (NIPGR). Portions of this research were carried out at the Argonne National Laboratory Structural Biology Center of the Advanced Photon Source, a national use facility operated by the University of Chicago for the U.S. Department of Energy Office of Biological and Environmental Research (grant DE-AC02-06CH11357). The work was supported by the Department of Biotechnology, Ministry of Science and Technology, India (grants BT/PR271/AGR/36/687/2011 and BT/06/IYBA/2012 to N.C.B.); the National Science Foundation (grant NSF-MCB-1614539 to J.M.J.), and a grant from the Max Planck Society to J.G. Support was also provided by a NIPGR Short Term Overseas Fellowship to N.C.B.; by a Max Planck India Fellowship to N.C.B.; by the University Grants Commission (India) to R.K.; and by NIPGR to R.A.

AUTHOR CONTRIBUTIONS

N.C.B. conceived and designed the study; R.K., R.A., and N.C.B. performed and collected all in planta data; S.G.L., M.H.P., A.A., and J.M.J. conducted the structural biology studies; R.K., N.C.B., M.R., and D.G.V. performed the biochemical work; N.C.B., J.M.J., J.G., S.G.L., and R.K. drafted the article; and all the authors edited and approved the article.

Received January 22, 2019; revised March 29, 2019; accepted April 19, 2019; published April 25, 2019.

REFERENCES

Adams, P.D., et al. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66:** 213–221.

Agerbirk, N., and Olsen, C.E. (2012). Glucosinolate structures in evolution. Phytochemistry **77:** 16–45.

Augustine, R., and Bisht, N.C. (2015). Biofortification of oilseed Brassica juncea with the anti-cancer compound glucoraphanin by suppressing GSL-ALK gene family. Sci. Rep. **5:** 18005.

Augustine, R., Majee, M., Gershenzon, J., and Bisht, N.C. (2013a). Four genes encoding MYB28, a major transcriptional regulator of the aliphatic glucosinolate pathway, are differentially expressed in the allopolyploid *Brassica juncea*. J. Exp. Bot. **64:** 4907–4921.

Augustine, R., Mukhopadhyay, A., and Bisht, N.C. (2013b). Targeted silencing of *BjMYB28* transcription factor gene directs development of low glucosinolate lines in oilseed *Brassica juncea*. Plant Biotechnol. J. **11:** 855–866.

Benderoth, M., Textor, S., Windsor, A.J., Mitchell-Olds, T., Gershenzon, J., and Kroymann, J. (2006). Positive selection driving diversification in plant secondary metabolism. Proc. Natl. Acad. Sci. USA **103:** 9118–9123.

Benderoth, M., Pfalz, M., and Kroymann, J. (2009). Methylthioalkylmalate synthases: Genetics, ecology and evolution. Phytochem. Rev. **8:** 255–268.

Bisht, N.C., Gupta, V., Ramchiary, N., Sodhi, Y.S., Mukhopadhyay, A., Arumugam, N., Pental, D., and Pradhan, A.K. (2009). Fine mapping of loci involved with glucosinolate biosynthesis in oilseed mustard (*Brassica juncea*) using genomic information from allied species. Theor. Appl. Genet. **118:** 413–421.

Cartea, M.E., and Velasco, P. (2008). Glucosinolates in *Brassica* foods: Bioavailability in food and significance for human health. Phytochem. Rev. **7:** 213–229.

Chandna, R., Augustine, R., and Bisht, N.C. (2012). Evaluation of candidate reference genes for gene expression normalization in *Brassica juncea* using real time quantitative RT-PCR. PLoS One **7:** e36918.

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735–743.

Collaborative Computational Project, Number 4. (1994). The *CCP4* suite: Programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:** 760–763.

de Kraker, J.W., and Gershenzon, J. (2011). From amino acid to glucosinolate biosynthesis: Protein sequence changes in the evolution of methylthioalkylmalate synthase in Arabidopsis. Plant Cell **23:** 38–53.

de Kraker, J.W., Luck, K., Textor, S., Tokuhisa, J.G., and Gershenzon, J. (2007). Two Arabidopsis genes (IPMS1 and IPMS2) encode isopropylmalate synthase, the branchpoint step in the biosynthesis of leucine. Plant Physiol. **143:** 970–986.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66:** 486–501.

Field, B., Cardon, G., Traka, M., Botterman, J., Vancanneyt, G., and Mithen, R. (2004). Glucosinolate and amino acid biosynthesis in Arabidopsis. Plant Physiol. **135:** 828–839.

Forouhar, F., et al. (2006). Crystal structures of two bacterial 3-hydroxy-3-methylglutaryl-CoA lyases suggest a common catalytic mechanism among a family of TIM barrel metalloenzymes cleaving carbon-carbon bonds. J. Biol. Chem. **281:** 7533–7545.

Franzke, A., Lysak, M.A., Al-Shehbaz, I.A., Koch, M.A., and Mummenhoff, K. (2011). Cabbage family affairs: The evolutionary history of Brassicaceae. Trends Plant Sci. **16:** 108–116.

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994). The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol. Biol. **25:** 989–994.

Halkier, B.A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. Annu. Rev. Plant Biol. **57:** 303–333.

He, Y., Mawhinney, T.P., Preuss, M.L., Schroeder, A.C., Chen, B., Abraham, L., Jez, J.M., and Chen, S. (2009). A redox-active isopropylmalate dehydrogenase functions in the biosynthesis of glucosinolates and leucine in Arabidopsis. Plant J. **60:** 679–690.

He, Y., Chen, B., Pang, Q., Strul, J.M., and Chen, S. (2010). Functional specification of Arabidopsis isopropylmalate isomerases in glucosinolate and leucine biosynthesis. Plant Cell Physiol. **51:** 1480–1487.

He, Y., Galant, A., Pang, Q., Strul, J.M., Balogun, S.F., Jez, J.M., and Chen, S. (2011). Structural and functional evolution of isopropylmalate dehydrogenases in the leucine and glucosinolate pathways of *Arabidopsis thaliana*. J. Biol. Chem. **286:** 28794–28801.

Hopkins, R.J., van Dam, N.M., and van Loon, J.J.A. (2009). Role of glucosinolates in insect-plant relationships and multitrophic interactions. Annu. Rev. Entomol. **54:** 57–83.

Huisman, F.H.A., Koon, N., Bulloch, E.M., Baker, H.M., Baker, E.N., Squire, C.J., and Parker, E.J. (2012). Removal of the C-terminal regulatory domain of α-isopropylmalate synthase disrupts functional substrate binding. Biochemistry **51:** 2289–2297.

Imhof, J., Huber, F., Reichelt, M., Gershenzon, J., Wiegreffe, C., Lächler, K., and Binder, S. (2014). The small subunit 1 of the Arabidopsis isopropylmalate isomerase is required for normal growth and development and the early stages of glucosinolate formation. PLoS One **9:** e91071.

Khersonsky, O., and Tawfik, D.S. (2010). Enzyme promiscuity: A mechanistic and evolutionary perspective. Annu. Rev. Biochem. **79:** 471–505.

Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., and Mitchell-Olds, T. (2001). Genetic control of natural variation in Arabidopsis glucosinolate accumulation. Plant Physiol. **126:** 811–825.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae). Mol. Biol. Evol. **17:** 1483–1498.

Koon, N., Squire, C.J., and Baker, E.N. (2004). Crystal structure of LeuA from *Mycobacterium tuberculosis*, a key enzyme in leucine biosynthesis. Proc. Natl. Acad. Sci. USA **101:** 8295–8300.

Kroymann, J., Textor, S., Tokuhisa, J.G., Falk, K.L., Bartram, S., Gershenzon, J., and Mitchell-Olds, T. (2001). A gene controlling variation in Arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol. **127:** 1077–1088.

Kroymann, J., Donnerhacke, S., Schnabelrauch, D., and Mitchell-Olds, T. (2003). Evolutionary dynamics of an Arabidopsis insect resistance quantitative trait locus. Proc. Natl. Acad. Sci. USA **100** (suppl. 2): 14587–14592.

Lee, S.G., Nwumeh, R., and Jez, J.M. (2016). Structure and mechanism of isopropylmalate dehydrogenase from *Arabidopsis thaliana*: Insights on leucine and aliphatic glucosinolate biosynthesis. J. Biol. Chem. **291:** 13421–13430.

Lionneton, E., Aubert, G., Ochatt, S., and Merah, O. (2004). Genetic analysis of agronomic and quality traits in mustard (*Brassica juncea*). Theor. Appl. Genet. **109:** 792–799.

Lysak, M.A., Koch, M.A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe *Brassiceae*. Genome Res. **15:** 516–525.

Lysak, M.A., Cheung, K., Kitschke, M., and Bures, P. (2007). Ancestral chromosomal blocks are triplicated in Brassiceae species with varying chromosome number and genome size. Plant Physiol. **145:** 402–410.

Ma, J., Zhang, P., Zhang, Z., Zha, M., Xu, H., Zhao, G., and Ding, J. (2008). Molecular basis of the substrate specificity and the catalytic mechanism of citramalate synthase from *Leptospira interrogans*. Biochem. J. **415:** 45–56.

Mahmood, T., Ekuere, U., Yeh, F., Good, A.G., and Stringam, G.R. (2003). Molecular mapping of seed aliphatic glucosinolates in *Brassica juncea*. Genome **46:** 753–760.

Meenu, A.R., Augustine, R., Majee, M., Pradhan, A.K., and Bisht, N.C. (2015). Genomic origin, expression differentiation and regulation of multiple genes encoding CYP83A1, a key enzyme for core

glucosinolate biosynthesis, from the allotetraploid *Brassica juncea*. Planta **241:** 651–665.

Milo, R., and Last, R.L. (2012). Achieving diversity in the face of constraints: Lessons from metabolism. Science **336:** 1663–1667.

Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006). HKL-3000: The integration of data reduction and structure solution—From diffraction images to an initial model in minutes. Acta Crystallogr. D Biol. Crystallogr. **62:** 859–866.

Mun, J.H., et al. (2009). Genome-wide comparative analysis of the *Brassica rapa* gene space reveals genome shrinkage and differential loss of duplicated genes after whole genome triplication. Genome Biol. **10:** R111.

Okada, T., Tomita, T., Wulandari, A.P., Kuzuyama, T., and Nishiyama, M. (2010). Mechanism of substrate recognition and insight into feedback inhibition of homocitrate synthase from *Thermus thermophilus*. J. Biol. Chem. **285:** 4195–4205.

Ramchiary, N., Bisht, N.C., Gupta, V., Mukhopadhyay, A., Arumugam, N., Sodhi, Y.S., Pental, D., and Pradhan, A.K. (2007). QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed Brassica juncea: Importance of recurrent selection backcross scheme for the identification of 'true' QTL. Theor. Appl. Genet. **116:** 77–85.

Sawada, Y., Kuwahara, A., Nagano, M., Narisawa, T., Sakata, A., Saito, K., and Hirai, M.Y. (2009). Omics-based approaches to methionine side-chain elongation in Arabidopsis: Characterization of gene encoding methylthioalkylmalate isomerase and methylthioalkylmalate dehydrogenase. Plant Cell Physiol. **50:** 1180–1190.

Schuster, J., Knill, T., Reichelt, M., Gershenzon, J., and Binder, S. (2006). Branched-chain aminotransferase4 is part of the chain elongation pathway in the biosynthesis of methionine-derived glucosinolates in Arabidopsis. Plant Cell **18:** 2664–2679.

Sodhi, Y.S., Mukhopadhyay, A., Arumugam, N., Verma, J.K., Gupta, K., Pental, D., and Pradhan, A.K. (2002). Genetic analysis of total glucosinolate in crosses involving a high glucosinolate Indian variety and a low glucosinolate line of *Brassica juncea*. Plant Breed. **121:** 508–511.

Sønderby, I.E., Geu-Flores, F., and Halkier, B.A. (2010). Biosynthesis of glucosinolates: Gene discovery and beyond. Trends Plant Sci. **15:** 283–290.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:** 2731–2739.

Textor, S., Bartram, S., Kroymann, J., Falk, K.L., Hick, A., Pickett, J.A., and Gershenzon, J. (2004). Biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*: Recombinant expression and characterization of methylthioalkylmalate synthase, the condensing enzyme of the chain-elongation cycle. Planta **218:** 1026–1035.

Textor, S., de Kraker, J.W., Hause, B., Gershenzon, J., and Tokuhisa, J.G. (2007). MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in Arabidopsis. Plant Physiol. **144:** 60–71.

Town, C.D., et al. (2006). Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell **18:** 1348–1359.

Vagin, A., and Teplyakov, A. (1997). MOLREP: An automated program for molecular replacement. J. Appl. Crystallogr. **30:** 1022–1025.

Wang, H., Wu, J., Sun, S., Liu, B., Cheng, F., Sun, R., and Wang, X. (2011). Glucosinolate biosynthetic genes in *Brassica rapa*. Gene **487:** 135–142.

Yang, J., et al. (2016). The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. Nat. Genet. **48:** 1225–1232.