

# From Amino Acid to Glucosinolate Biosynthesis: Protein Sequence Changes in the Evolution of Methylthioalkylmalate Synthase in *Arabidopsis*

Jan-Willem de Kraker<sup>1</sup> and Jonathan Gershenzon<sup>2</sup>

Department of Biochemistry, Max-Planck Institute for Chemical Ecology, D-07745 Jena, Germany

**Methylthioalkylmalate synthase (MAM) catalyzes the committed step in the side chain elongation of Met, yielding important precursors for glucosinolate biosynthesis in *Arabidopsis thaliana* and other Brassicaceae species. MAM is believed to have evolved from isopropylmalate synthase (IPMS), an enzyme involved in Leu biosynthesis, based on phylogenetic analyses and an overlap of catalytic abilities. Here, we investigated the changes in protein structure that have occurred during the recruitment of IPMS from amino acid to glucosinolate metabolism. The major sequence difference between IPMS and MAM is the absence of 120 amino acids at the C-terminal end of MAM that constitute a regulatory domain for Leu-mediated feedback inhibition. Truncation of this domain in *Arabidopsis* IPMS2 results in loss of Leu feedback inhibition and quaternary structure, two features common to MAM enzymes, plus an 8.4-fold increase in the  $k_{cat}/K_m$  for a MAM substrate. Additional exchange of two amino acids in the active site resulted in a MAM-like enzyme that had little residual IPMS activity. Hence, combination of the loss of the regulatory domain and a few additional amino acid exchanges can explain the evolution of MAM from IPMS during its recruitment from primary to secondary metabolism.**

## INTRODUCTION

The enormous diversity of plant secondary metabolites arises from a complex array of biosynthetic pathways. To understand the origin of this diversity, it is necessary to learn something about the evolution of secondary metabolite pathways. Even a cursory inspection of plant secondary metabolism reveals close similarities between many reactions of secondary metabolite pathways and those of primary metabolism, as well as many similarities in enzyme properties (Modolo et al., 2009). However, there is still only scattered information about how enzymes of secondary metabolism have been recruited from primary metabolism.

Perhaps the best work in this area concerns the evolutionary origin of two enzymes of alkaloid metabolism. Homospermidine synthase, the committed step of pyrrolizidine alkaloid biosynthesis, has been demonstrated to have been recruited from deoxyhypusine synthase, an enzyme of primary metabolism involved in the activation of the translation factor 5A (Ober and Hartmann, 1999, 2000). The initial step in the formation of nicotine and tropane-type alkaloids, putrescine *N*-methyltransferase, was shown to have been recruited from spermidine synthase, an enzyme of polyamine synthesis (Teuber et al., 2007;


Biastoff et al., 2009). In the evolution of indole and benzoxazinone biosynthesis, modification of the TS<sub>A</sub> subunit of Trp synthase to prevent its interactions with TS<sub>B</sub> led to the formation of indole instead of Trp. TS<sub>A</sub> itself has only indole-3-glycerol phosphate lyase activity and will not add Ser to the indole reaction intermediate necessary for Trp formation (Frey et al., 1997, 2000; Gierl and Frey, 2001). In these three cases, the strongest evidence for enzyme recruitment comes from phylogenetic studies inferring duplication of the gene encoding the original enzyme and catalytic assays demonstrating the neofunctionalization of one of the duplicate gene products to enable it to carry out the new reaction. However, the actual changes on a protein level that cause the alterations in substrate specificity and enzyme function are not yet known.


Researchers of glucosinolates have long speculated about the evolution of this group of plant secondary metabolites. Glucosinolates occur throughout the Brassicales, including several important crops (oilseed rape [*Brassica napus*], cabbage [*B. oleracea*, Capitata group], and broccoli [*B. oleracea*, Italica group]) and the model plant *Arabidopsis thaliana*. These amino acid-derived thioglycosides interact with myrosinase (thioglycoside glucohydrolase) and associated proteins to create an activated plant defense system known as the mustard oil bomb. Tissue damage results in hydrolysis of glucosinolates and release of biologically active isothiocyanates, thiocyanates, and nitrile compounds (Halkier and Gershenzon, 2006). This multi-component defense mechanism shows a clear resemblance to that of the more widespread cyanogenic glycosides, amino acid-derived metabolites that are also activated through a glucohydrolase activity, and there is some evidence that the biosynthesis of the core glucosinolate skeleton evolved from the pathway for biosynthesis of cyanogenic glycosides (Halkier et al., 2002).

<sup>1</sup> Current address: ENZA Zaden Research and Development, PO Box 7, 1600 AA Enkhuizen, The Netherlands.

<sup>2</sup> Address correspondence to gershenzon@ice.mpg.de.

The author 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) is: Jonathan Gershenzon (gershenzon@ice.mpg.de).

 Online version contains Web-only data.

 Open Access articles can be viewed online without a subscription. www.plantcell.org/cgi/doi/10.1105/tpc.110.079269

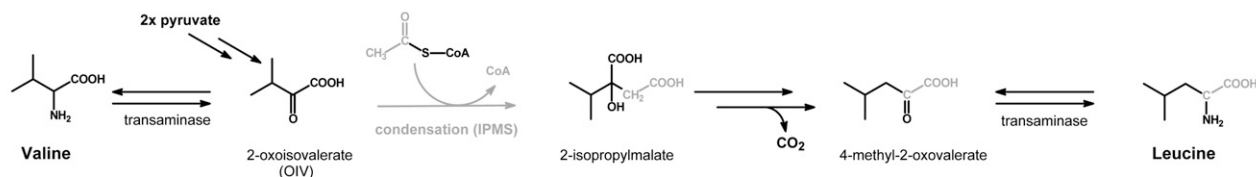
Prior to the formation of the core glucosinolate skeleton, the precursor amino acids, including Met, Phe, and others, often undergo side chain elongation in a series of reactions that resemble the late steps in Leu biosynthesis (Figure 1) (Kroymann et al., 2001; Textor et al., 2004, 2007; Benderoth et al., 2006; Halkier and Gershenzon, 2006; de Kraker et al., 2007). This elongation is an iterative three-step process that operates predominantly on Met in *Arabidopsis* and results in a net addition of up to six methylene groups and thus is an important determinant of glucosinolate variability in this plant. The committed step in Met side chain elongation is catalyzed by methylthioalkylmalate synthase (MAM) (Kroymann et al., 2001; Textor et al., 2004, 2007; Benderoth et al., 2006), and this enzyme appears to be evolutionarily derived from isopropylmalate synthase (IPMS) of Leu biosynthesis (de Kraker et al., 2007; Benderoth et al., 2009). A rigorous phylogenetic analysis of IPMS and MAM genes suggests that duplication of IPMS and subsequent changes in enzyme function were involved in the origin of MAM (Benderoth et al., 2006, 2009). Based on a tree constructed from analysis of 32 IPMS and MAM sequences from throughout the plant kingdom, these events appear to have occurred after the origin of the Brassicales, although many more taxa need to be sampled to determine the timing accurately.

*Arabidopsis* contains two functional IPMS-encoding genes (*IPMS1* [At1g18500] and *IPMS2* [At1g74040]). The MAM genes,

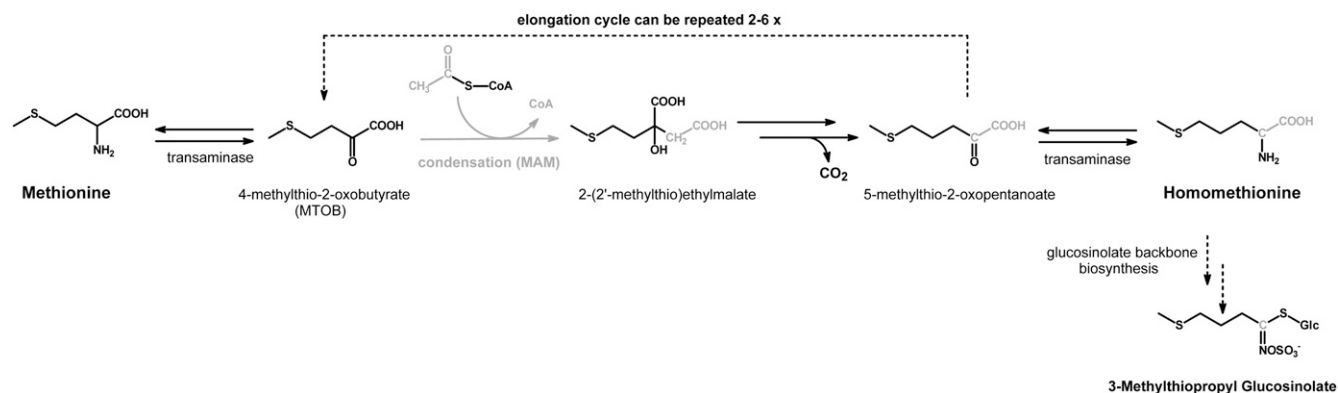
on the other hand, cluster at a completely different position in the genome that comprises up to three genes (*MAM1* [At5g2310], *MAM2* [not found in the Columbia-0 ecotype], and *MAM3* [At5g23020]) whose sequences vary among ecotypes because of gene deletion and conversion events (Benderoth et al., 2006). The *Arabidopsis* IPMS and MAM sequences share ~60% amino acid identity (Benderoth et al., 2009). Both catalyze aldol-type condensation reactions between acetyl-CoA and 2-oxo acids, but IPMS uses only 2-oxoisovalerate (OIV) *in vivo*, whereas MAM uses either the 2-oxo acid derived directly from Met, 4-methylthio-2-oxobutyrate (MTOB), or other elongated  $\omega$ -methylthio-2-oxoalkanoate homologs. Yet, *in vitro*, both IPMS and MAM can catalyze the reaction with the transaminated substrate of the other enzyme at a low rate (de Kraker et al., 2007; Textor et al., 2007).

Thus, a close evolutionary relationship of IPMS and MAM can be deduced from both phylogenetic analyses of their amino acid sequences and the overlap of their catalytic abilities. Yet, no information is available about the modifications in the IPMS protein that were necessary for the evolution of full MAM activity. In this study, we have begun to investigate the changes in IPMS that have allowed its recruitment from amino acid metabolism into glucosinolate biosynthesis as a MAM enzyme. We describe how loss of the C-terminal domain of the *Arabidopsis* IPMS, exchange of active-site residues, and other sequence changes

### Elongation in late steps of leucine biosynthesis (primary metabolism)



### Elongation of methionine in glucosinolate biosynthesis (secondary metabolism)



**Figure 1.** Metabolic Parallels between the Late Steps of Leu Biosynthesis and the Elongation of the Met Side Chain in Glucosinolate Biosynthesis.

Both reaction sequences are initiated by a condensation between a 2-oxoalkanoic acid and acetyl-CoA. This condensation is catalyzed by IPMS in Leu biosynthesis and by MAM in glucosinolate biosynthesis. Subsequent isomerization and oxidative decarboxylation results in the net addition of one methylene group. In the case of Met side chain elongation in *Arabidopsis*, this process can be repeated to give up to six methylene groups.

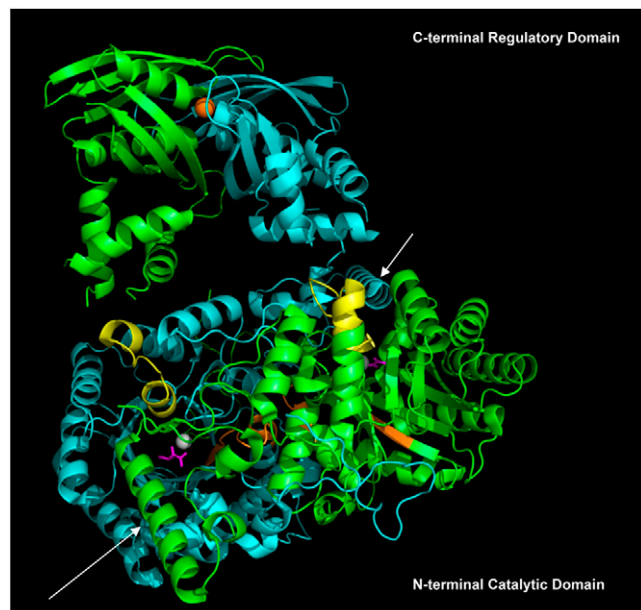
increase MAM activity, affecting not only substrate specificity but also allosteric inhibition and cofactor preference.

## RESULTS

### The Protein Sequence of MAM Differs from That of IPMS in Several Ways

As a basis for comparing the structures of the MAM and IPMS proteins, we used the only available crystal structure of an IPMS, that from *Mycobacterium tuberculosis* (Mt-IPMS; also referred to as 1SR9) (Koon et al., 2004). In this dimeric protein (Figure 2), each monomer is folded into two major domains, an N- and a C-terminal domain, which are separated by two small subdomains that have a flexible hinge in between them. The N-terminal domain is the catalytic domain and consists mainly of a  $(\beta/\alpha)_8$  barrel (TIM barrel) with a divalent metal cofactor necessary for substrate binding. The C-terminal domain contains an allosteric Leu binding site and is thought to act as a regulatory domain. The major domains, but not the subdomains, pack closely together in the homodimer.

The *Arabidopsis* MAM and IPMS sequences align closely with that of Mt-IPMS even though they show <25% amino acid identity (Figure 3). All these proteins share a common  $(\beta/\alpha)_8$  barrel as a catalytic domain. However, the MAM proteins lack



**Figure 2.** Dimer Structure of *M. tuberculosis* IPMS (Koon et al., 2004) Deposited in the Protein Data Bank (1SR9) and Drawn with PYMOL.

The individual monomers are colored blue and green. The Prosite motifs PS00815 and PS00816 mentioned in the text are colored in yellow and orange, respectively. The substrate OIV is depicted in magenta and positioned in the catalytic domain with a divalent cation (gray). The orange sphere marks one regulatory Leu binding site. Arrows point to the  $\alpha 10$  helix, which is positioned over the catalytic site of the opposite monomer.

~120 amino acids at their C-terminal ends that are present in the IPMSs and make up the regulatory domain in Mt-IPMS (Koon et al., 2004; Singh and Bhakuni, 2007). This could explain why the IPMS of *Arabidopsis* is known to be inhibited by Leu, whereas the MAMs have not been reported to be affected by Leu concentration (Falk et al., 2004; de Kraker et al., 2007). Allosteric inhibition of enzyme activity probably occurs through binding of Leu to the regulatory domain of one monomer, which results in small conformational changes influencing the position of residues found in or near helix  $\alpha 10$  (particularly His-379 and Trp-410) that surround the active site of the other monomer (Koon et al., 2004).

Less prominent differences between IPMS and MAM sequences are found in the conserved HxH[DN]D motif (Table 1) at the end of the seventh  $\beta$ -sheet in the  $(\beta/\alpha)_8$  barrel. This is part of a larger motif designated PS00816 in the Prosite database (Sigrist et al., 2002), that, together with a second motif PS00815 at the end of the first  $\beta$ -sheet, is involved in binding of the divalent metal cofactor (Wang et al., 1991; Koon et al., 2004). Motif PS00815 is well conserved in the IPMS/MAM enzyme family, but motif PS00816 is not. In contrast with microbial IPMSs and the MAMs, plant IPMSs have the second His replaced by a Gln in the HxH[DN]D sequence.

In a schematic view of ligand binding to the Mt-IPMS active site generated by LIGPLOT (Wallace et al., 1995), 11 amino acids are described that closely interact with the substrate OIV (Figure 4). Three of them, Asp-81, His-285, and His-287, also have a role in binding of the divalent metal cofactor and are part of the previously mentioned Prosite motifs. A fourth amino acid, Glu-218, also interacts with acetyl-CoA (Koon et al., 2004) and is conserved in the IPMS/MAM enzymes. Of the remaining seven amino acids, marked with a black circle in the alignment of Figure 3, two more appear to be conserved, Arg-80 and Thr-254. However, the other five amino acids, Leu-143, His-167, Ser-216, Asn-250, and Pro-252, are not only different between IPMS and MAM, but are also different among the MAM enzymes. A more detailed examination of these amino acids with the RCSB PDB Ligand Explorer 3.5 showed that each of them, except Pro-252, has a hydrophobic interaction with one or both of the two methyl groups of OIV, designated C4 and C5 in Figure 4. Leu-143 interacts with C4, while Ser-216, Asn-250 (weakly), and Pro-252 interact with C5, and His-167 has hydrophobic interactions with both methyl groups. For each of these five residues, a change in amino acid might cause a difference in substrate specificity.

### The C-Terminal Regulatory Domain of IPMS (Not Present in MAM) Is Responsible for Leu Feedback Inhibition

To verify experimentally whether the C-terminal sequence in *Arabidopsis* IPMS1 and IPMS2 comprises a Leu feedback regulatory domain, we made constructs for IPMS1 and IPMS2 that were truncated at the end of subdomain II ( $\alpha 13$ , after Asp-421 and Asp-433, respectively). These truncated recombinant proteins, designated IPMS1/-R1 and IPMS2/-R2, were expressed in *Escherichia coli*, purified over a Ni-NTA agarose affinity column, and assayed for enzyme activity. Removal of the C-terminal domain resulted in proteins whose enzymatic activity was not inhibited in the presence of Leu, unlike the corresponding full-length IPMS1 and IPMS2 that were inhibited to 30 to 50% (Figure

**Table 1.** Sequences of Two Regions in the IPMS/MAM Protein Family (Prosite Motifs PS00815 and PS00816) That Are Responsible for Divalent Metal Cofactor Binding in Mt-IPMS

Enzyme <sup>a</sup>	PS00815 LR[DE]GxQxxxx{L}xxxxxK <sup>b</sup>	PS00816 [LIVMF]xx <u>HxH</u> [DN]DxGx[GAS]x[GASLI] <sup>b</sup>
IPMS1/IPMS2 <i>Arabidopsis</i>	LRDGeQspgaTltskeK	Ist <u>Hc</u> QNDIGlStA
IPMSa/IPMSb <i>Solanum pennellii</i>	LRDGeQspgaTmttkeK	Ist <u>Hc</u> QNDIGlStA
IPMSa/IPMSb <i>Oryza sativa</i>	LRDGeQspgaTltsaeK	Ist <u>Hc</u> QNDIGlAtA
IPMS <i>Medicago truncatula</i> <sup>c</sup>	LRDGeQspgaSmtskeK	Ist <u>Hc</u> QNDIGlStA
IPMS <i>Ostreococcus tauri</i> (chlorophyte) <sup>c</sup>	LRDGeQspgaTltsreK	Ist <u>Hc</u> QNDIGlStA
IPMS ( <i>LeuA</i> ) <i>E. coli</i>	LRDGeQalqaSlsvkeK	Isv <u>Ht</u> HDDIGlAvG
IPMS <i>M. tuberculosis</i> (Mt-IPMS)	LRDGnQalidPmsparK	Ls <u>Hp</u> HNDrGtAvA
IPMS ( <i>LEU4</i> ) <i>Saccharomyces cerevisiae</i>	LRDGnQslpdPmsveqK	Ist <u>Hc</u> HNDrGcGvA
IPMS <i>Anabaena variabilis</i> (cyanobacterium) <sup>c</sup>	LRDGtQreglSistedK	Igi <u>Ht</u> HNDsDmAvA
IPMS <i>Methanocaldococcus jannaschii</i> (archaeon) <sup>c</sup>	LRDGeQtppgCftkeqK	Isv <u>Hc</u> HNDfGlAvA
MAM1 <i>Arabidopsis</i>	LRDGeQspggSltpqK	Vav <u>Hc</u> HNDIGlAtA
MAM2 <i>Arabidopsis</i>	LRDGeQapggSltpqK	Fsv <u>Hc</u> HNDIGlAtA
MAM3 <i>Arabidopsis</i>	LRDGeQspgaAltppqK	Fai <u>Hc</u> HNDIGvAtA
MAMa <i>Arabidopsis lyrata</i>	LRDGeQapggSltpqK	Fsv <u>Hc</u> HNDIGlAtA

The HxH[DN]D pattern contained within PS00816 is represented in bold. Underlined amino acids are proposed to act as ligands for the divalent metal cofactor.

<sup>a</sup>Phylogenetic relationships of these sequences have been proposed by Benderoth et al. (2009).

<sup>b</sup>Residues listed between brackets are ambiguous and exchangeable; residues listed between braces are not accepted at the given position.

<sup>c</sup>These proteins can be classified as IPMSs and not MAMs since the organism does not produce glucosinolates.

5). The truncated *Arabidopsis* IPMS2 also showed no reduction in enzyme activity, contrary to what was observed for the Mt-IPMS protein, which lost 88% of its enzyme activity when truncated in the same manner (in subdomain II- $\alpha$ 12, after Phe-457), (Singh and Bhakuni, 2007). These results confirm that the C-terminal parts of IPMS1 and IPMS2 do indeed contain a Leu regulatory domain that is absent in the MAM proteins.

### Removal of the Regulatory Domain Increases the MAM Activity of IPMS1 and Especially IPMS2

The lack of a regulatory domain did not influence the range of substrates accepted by the *Arabidopsis* IPMS enzymes. When evaluated in an assay with [<sup>14</sup>C]acetyl-CoA using radio-HPLC, the truncated enzymes IPMS1/-R1 and IPMS2/-R2 were able to condense the same 2-oxo acid substrates with acetyl-CoA as the full-length IPMS1 and IPMS2 did (de Kraker et al., 2007), including OIV (the natural substrate of IPMS) and MTOB (a MAM substrate) (Figure 6). However, IPMS2/-R2 showed an increase in activity with MTOB relative to the activity of the full-length IPMS2 with MTOB (see Supplemental Figure 1 online).

To obtain kinetic data, a spectrophotometric end-point assay with 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) was employed. The removal of the regulatory domain from IPMS1 and IPMS2 was found to increase the turnover number ( $k_{cat}$ ) for MTOB severalfold (Table 2) with increases greater for IPMS2/-R2 (nearly 4-fold versus IPMS2) than for IPMS1/-R1 (2-fold versus IPMS1). There were also increases in the  $k_{cat}$  for the native substrate, OIV, but of lesser magnitude. Taking into account changes in  $K_m$ , the resulting specificity constant ( $k_{cat}/K_m$ ) of the truncated enzymes for MTOB was raised by a factor of over 8 for IPMS2/-R2 and a factor of 3 for IPMS1/-R1. For OIV, the specificity constant of the truncated enzymes was increased by a factor of  $\sim 1.5$ .

### Removal of the Regulatory Domain Affects the Quaternary Structure of IPMS2 but Not of IPMS1

Although IPMS1 and IPMS2 have amino acid sequences that are 92% identical, removal of the regulatory domain had a much bigger impact on the MAM activity of IPMS2 than on that of IPMS1 (Table 2). The knowledge that IPMS1 and IPMS2 have different quaternary structures (de Kraker et al., 2007) and that MAM is monomeric (i.e., has no quaternary structure at all; Textor et al., 2004) prompted us to examine the quaternary structures of the truncated proteins IPMS1/-R1 and IPMS2/-R2 as a possible explanation for the difference in MAM activity. The molecular mass of these proteins in both full-length and regulatory domain-truncated forms was estimated by calibrated gel filtration chromatography (Table 3). The position of the eluting protein was measured by UV absorbance at 280 nm and by measuring the enzyme activity of eluted fractions with the DTNB assay. These two methods gave approximately the same results.

As previously reported (de Kraker et al., 2007), IPMS1 eluted predominantly as a dimer and IPMS2 as a tetramer. After removal of the regulatory domain, the truncated IPMS1/-R1 still behaved as a dimer, but IPMS2/-R2 behaved as a monomer. Interestingly, IPMS2/-R2, which showed greater MAM activity than IPMS1/-R1, appeared to have the same three-dimensional properties in solution as the native MAMs (Table 3, last column). Meanwhile, IPMS1/-R1 had three-dimensional properties more similar to those of the other IPMS proteins, and loss of the regulatory domain resulted in elution behavior (expected  $n$ /rounded  $n = 1.0$ ) similar to those of the globular proteins used for calibration. Thus, removal of the regulatory domain of IPMS2 has a major impact on protein quaternary structure as well as catalysis, but this was not the case for IPMS1, which remained a dimer.



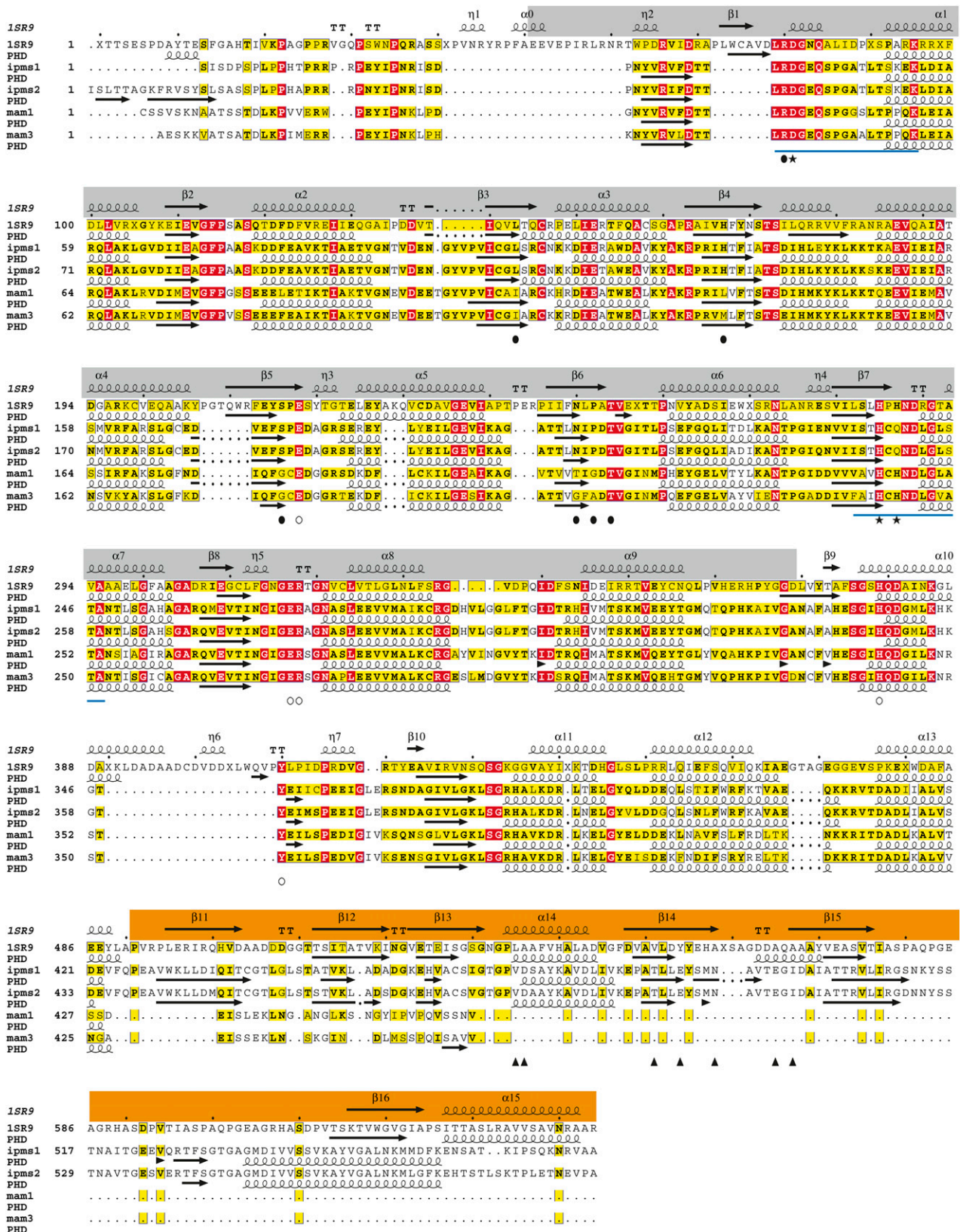
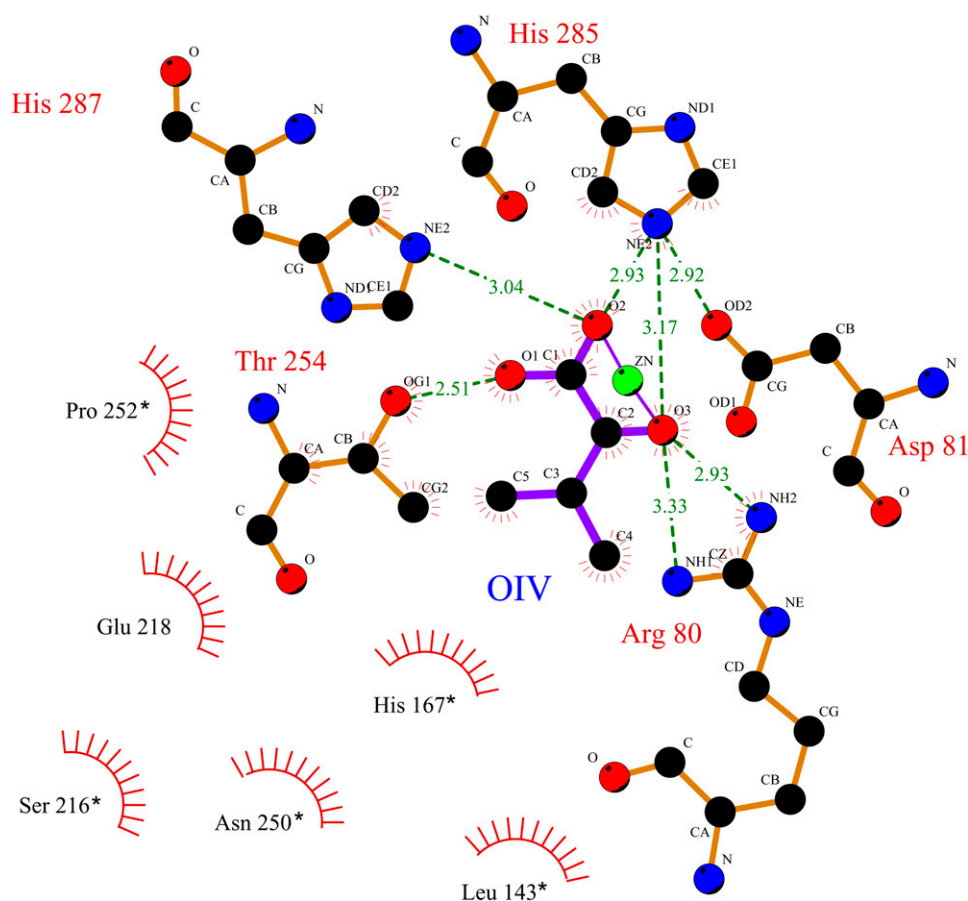


Figure 3. Alignment of Deduced Amino Acid Sequences for the IPMS and MAM Enzymes of *Arabidopsis* and the IPMS of *M. tuberculosis* (PDB 1SR9).



**Figure 4.** LIGPLOT Diagram of One of the Active Sites of Mt-IPMS Showing Interactions between Amino Acid Residues and the Substrate OIV.

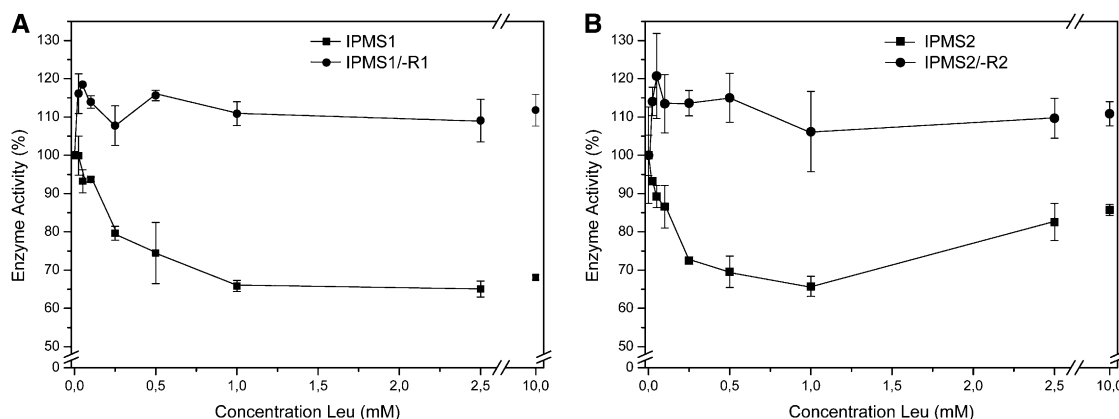
The plot of the second active site of this functional dimeric protein is the mirror image of this diagram showing the same residues from the opposite protein chain. Amino acids are represented by brown covalent bonds, whereas blue bonds represent OIV; carbon atoms are in black, nitrogen in blue, oxygen in red, and zinc in green. Dashed lines in green represent hydrogen bonds between substrate and protein. Spoked arcs represent hydrophobic contacts between enzyme and substrate and indicate the direction of such interactions. Amino acids that differ between the IPMS and MAM enzymes are labeled with an asterisk.

To explore further the effect of the regulatory domain of IPMS2 on quaternary structure and catalysis, we fused this domain (R2) to both MAM1 and MAM3, which naturally lack regulatory domains, to observe whether the resulting proteins might behave as multimers and have reduced MAM activity. The regulatory domain was added at a position similar to that in IPMS2 (after Glu-430 and Glu-428 in MAM1 and MAM3, respectively) resulting in proteins designated MAM1/+R2 and MAM3/+R2. These eluted as mixtures of

monomers, dimers, and tetramers from the gel filtration column with dimers predominating, compared with the exclusive presence of monomers for the native MAM proteins (Table 3). However, there was no change in substrate specificity, though enzyme activity was reduced by 20 to 40%. Addition of 2 mM Leu to the enzyme assay with MAM1/+R2 or MAM3/+R2 had no measurable effect on enzyme activity, indicating that we had not been able to construct a functional allosteric binding site in these chimeric proteins.

**Figure 3.** (continued).

The first line shows the exact secondary structure for *M. tuberculosis* IPMS ( $\alpha$  for  $\alpha$ -helix,  $\beta$  for  $\beta$ -sheet,  $\eta$  for  $3_{10}$ -helix, and TT for  $\beta$ -turn). The secondary structures of the other protein sequences given are those predicted by the program PHD. Numbering of secondary structure elements is the same as that in Figure 1 of Koon et al. (2004). The catalytic domain (gray rectangle) consists of the  $(\beta/\alpha)_8$ -barrel and two additional helices  $\alpha 0$  and  $\alpha 9$ ; the regulatory domain (orange rectangle) present only in IPMS includes two  $\beta\beta\alpha$  units and contains an allosteric Leu binding site (arrowheads). The catalytic domain and regulatory domain are linked by subdomain I that includes  $\beta 9$ - $\beta 10$  plus  $\alpha 10$  and subdomain II that includes  $\alpha 11$ - $\alpha 13$ . In *M. tuberculosis* IPMS, amino acid residues that act as ligands for the divalent metal cofactor are marked with stars, residues that interact with OIV with black circles, and residues that interact with acetyl-CoA with open circles. The positions of the motifs identified in the Prosite database as PS00815 and PS00816 are underlined in blue. For simplification, proteins are shown without the ChloroP-predicted transit peptide.



**Figure 5.** Effect of Leu on the Enzyme Activity of *Arabidopsis* IPMS Proteins with and without the C-Terminal Regulatory Domain.

IPMS1 and the truncated IPMS1/-R1 (**A**) and IPMS2 and the truncated IPMS2/-R2 (**B**). Activities are expressed as a percentage of the activity of the full-length protein with regulatory domain in the absence of Leu. Data represent means  $\pm$  SD ( $n = 3$ ).

### Amino Acid Exchanges in the Active Site of IPMS Increase Activity with MAM Substrates

Whereas removal of the regulatory domain from IPMS2 increased the MAM activity, it did not affect the actual range of accepted substrates (Figure 6). Hence, we examined more closely the active site amino acid residues in the *Arabidopsis* IPMSs likely to interact with the substrate (marked with black

circles in Figure 3) that differed from the corresponding residues found in the MAMs. These five residues were changed in IPMS2/-R2 to those present in MAM1 or MAM3 by making the exchanges L143I, H167L, S216G, N250G, and P252G (numbering follows that for the *M. tuberculosis* IPMS). The last three changes should increase the space in the IPMS active site, which might be important in accommodating the larger MAM substrate. Several of the mutated enzymes were able to convert not only the MAM

Substrate	Structure	Enzyme using substrate in planta	Substrate acceptance <i>in vitro</i>		
			IPMS1	IPMS1/-R1	IPMS2/-R2 IPMS2
Pyruvate			+	+	+
2-Oxobutyrate			+	+	+
2-Oxoisovalerate (OIV)		IPMS1 IPMS2	+	+	+
2-Oxovalerate			+	+	+
3-Methyl-2-oxovalerate			+	±	±
4-Methyl-2-oxovalerate			±	±	±
2-Oxohexanoate			±	±	±
4-Methylthio-2-oxobutyrate (MTOB)		MAM1 MAM2 MAM3	±	±	+
2-Oxoheptanoate (OHP)			-	-	-
5-Methylthio-2-oxopentanoate (MTOB)		MAM1 MAM3	-	-	-
6-Methylthio-2-oxohexanoate (MTOH)		MAM3 (MAM1) <sup>a</sup>	-	-	-

**Figure 6.** Substrate Specificity of IPMS1, IPMS2, IPMS1/-R1, and IPMS2/-R2 *In Vitro*.

When substrates are employed as principal substrates by one of the IPMS or MAM enzymes in planta, this is also mentioned. Data for IPMS1 and IPMS2 are from de Kraker et al. (2007). <sup>a</sup>Very small *in vitro* activity of uncertain role in planta (S. Textor and J. Gershenzon, unpublished data; assays performed as described in Textor et al. 2007). +, accepted substrate; ±, enzyme activity <2% of that with OIV; -, no detectable conversion.

**Table 2.** Influence of the IPMS Regulatory Domain on Kinetic Data for the Native IPMS Substrate OIV and the MAM Substrate MTOB

Enzyme <sup>a</sup>	$k_{cat}$ (s <sup>-1</sup> ) ± SE		$K_m$ (μM) ± SE		$k_{cat}/K_m$ (M <sup>-1</sup> s <sup>-1</sup> )	
	OIV	MTOB	OIV	MTOB	OIV	MTOB
IPMS1	3.44 ± 0.15	0.034 ± 0.002	234 ± 41	1657 ± 332	14.7 × 10 <sup>3</sup>	20.6
IPMS2	2.23 ± 0.06	0.035 ± 0.002	138 ± 18	1161 ± 170	16.2 × 10 <sup>3</sup>	30.1
IPMS1/-R1	4.05 ± 0.15	0.065 ± 0.003	184 ± 28	1090 ± 164	22.0 × 10 <sup>3</sup>	59.2
IPMS2/-R2	6.14 ± 0.19	0.127 ± 0.005	237 ± 29	500 ± 75	25.9 × 10 <sup>3</sup>	253.2

<sup>a</sup>Incubations ( $n = 4$ ) were done at saturating concentrations of 500 μM acetyl-CoA at 30°C and stopped after 10 min. The data for IPMS1 and IPMS2 are similar to those previously reported using a less sensitive continuous spectrophotometric assay with *N*-ethylmaleimide (de Kraker et al., 2007).

substrate, MTOB, but also larger MAM substrates, including 5-methylthio-2-oxopentanoate (MTOB) and its aliphatic analog 2-oxoheptanoate (OHP). Native IPMSs have no activity with these larger MAM substrates (de Kraker et al. 2007). However, it should be noted that these mutant enzymes had reduced overall activity (10 to 25% of the activity present in IPMS2), and the IPMS substrate, OIV, was still most preferred. The exchange S216G was also made on untruncated IPMS2, but this did not lead to any enhancement of MAM activity, and both IPMS and MAM activities were even more reduced. Hence, no further point mutations were tested on the untruncated enzyme.

Combinations of the amino acid exchanges made in IPMS2/-R2 usually resulted in inactive enzymes (Table 4) except the pairing of S216G with P252G, which resulted in an enzyme that

like native MAM enzymes actually preferred the MAM substrate MTOB to OIV by a factor of 10, as shown with the radio-HPLC assay (Figure 7). This level of activity was sufficient to allow the use of the DTNB endpoint assay to obtain kinetic data for MTOB (at 1 mM of acetyl-CoA) as follows:  $K_m = 4.12 \pm 0.51$  mM,  $k_{cat} = 0.057 \pm 0.003$  s<sup>-1</sup>, and  $k_{cat}/K_m = 13.9$  M<sup>-1</sup> s<sup>-1</sup>.

#### The HxH[DN]D Motif Defines the Divalent Metal Cofactor Preferences for IPMS and MAM

The *Arabidopsis* MAM enzymes contain an HxHND sequence conserved among many IPMS proteins that is involved in binding of the divalent metal cofactor (Koon et al., 2004). However, in all known plant IPMS sequences this is substituted by a HxQND

**Table 3.** Estimation of molecular mass (MM) and number of subunits (N) of Native and Truncated IPMS Proteins and of Native and Extended MAM Proteins by Calibrated Gel Filtration

Enzyme	MM of Monomer (kD) <sup>a</sup>	Experimental MM (kD) <sup>b</sup>	Experimental $n^c$	Rounded $n^d$	Exp $n$ /Rounded $n^e$
IPMS1	65.5				
Major peak		158, 153	2.4, 2.3	2	1.2, 1.2
Minor peak <sup>f</sup>		(342, 309)	(5.2, 4.7)	(4)	(1.3, 1.2)
IPMS2	66.7				
Major peak		322, 286	4.8, 4.3	4	1.2, 1.1
Minor peak <sup>f</sup>		(149, 153)	(2.2, 2.3)	(2)	(1.1, 1.1)
MAM1	53.2	85	1.6	1	1.6
MAM3	53.0	84	1.6	1	1.6
IPMS1/-R1	49.5	101, 96	2.0, 1.9	2	1.0, 1.0
IPMS2/-R2	50.6	83, 76	1.6, 1.5	1	1.6, 1.5
MAM1/+R2	66.6				
Major peak		168	2.5	2	1.3
Major peak		111 <sup>g</sup>	1.7	1	1.7
Minor peak <sup>f</sup>		(344)	(5.2)	(4)	(1.3)
MAM3/+R2	66.6				
Major peak		181	2.7	2	1.4
Minor peak <sup>f</sup>		(327)	(4.9)	(4)	(1.2)
Minor peak <sup>f</sup>		(85) <sup>g</sup>	(1.3)	(1)	(1.3)

<sup>a</sup>Calculated from the protein sequence including the C-terminal His-tag of 3.3 kD. MM, molecular mass.

<sup>b</sup>Position of peak was determined by the UV trace of the HPLC. Values in italics show position of the same peak as determined by the DTNB assay. This assay was performed only for IPMS activity. MAM enzyme activity in the peaks of the UV trace was confirmed by radio-HPLC assay. However, exact positioning of the peak by enzymatic assay was not possible due to the necessary larger volume of assayed fractions.

<sup>c</sup>Calculated from ratio of experimental molecular mass to molecular mass of monomer.

<sup>d</sup>Integral number of subunits rounded from experimental  $n$  based on comparison with all other enzymes of the same type.

<sup>e</sup>Ratio is dependent upon the three-dimensional arrangement of the protein in solution in relation to the globular proteins used for calibration.

<sup>f</sup>Values for minor peaks <30% of the major peak are given in parentheses.

<sup>g</sup>Exact position of this monomer peak was hard to determine as it coelutes in the shoulder of the dimer peak.



**Table 4.** Substrate Specificity of IPMS2/-R2 after Various Amino Acid Exchanges in the Active Site<sup>a</sup>

Amino Acid Exchange	Native IPMS Substrate				
	OIV	MTOB	MTOP	OHP	MTOH
No mutation	++	+	–	–	–
L143I	++	+	–	–	–
H167L	–	–	–	–	–
S216G	++	+	±	–	–
N250G	++	+	+	+	–
P252G	++	+	± <sup>b</sup>	± <sup>b</sup>	–
H167L, L143I	–	–	–	–	–
H167L, N250G	± <sup>c</sup>	–	–	–	–
S216G, N250G	–	–	–	–	–
S216G, P252G	+	±	++	+	–
S216G, N250G, P252G	–	–	–	–	–

++, Preferred substrate; +, other accepted substrate; ± activity <2% of that with preferred substrate; –, no detectable conversion. Numbering of the residues is the same as in Mt-IPMS. MTOH, 6-methylthio-2-oxohexanoate

<sup>a</sup>Incubations were done in the presence of 6 mM 2-oxoalkanoic acid substrate, 1 mM acetyl-CoA, and 30 µg of purified protein for 4 h at 30°C, and products were analyzed by radio-HPLC.

<sup>b</sup>Minor activity detected only if an elevated amount (>80 µg) of purified protein was used.

<sup>c</sup>Minor activity detected only in the presence of 2 mM DTT.

sequence (Table 1). We suspected that this difference might be responsible for the fact that the preferred cofactor of the MAMs is Mn<sup>2+</sup> (Falk et al., 2004; Textor et al., 2004, 2007), whereas the preferred cofactor of IPMS1 and IPMS2 is Mg<sup>2+</sup> (de Kraker et al., 2007).

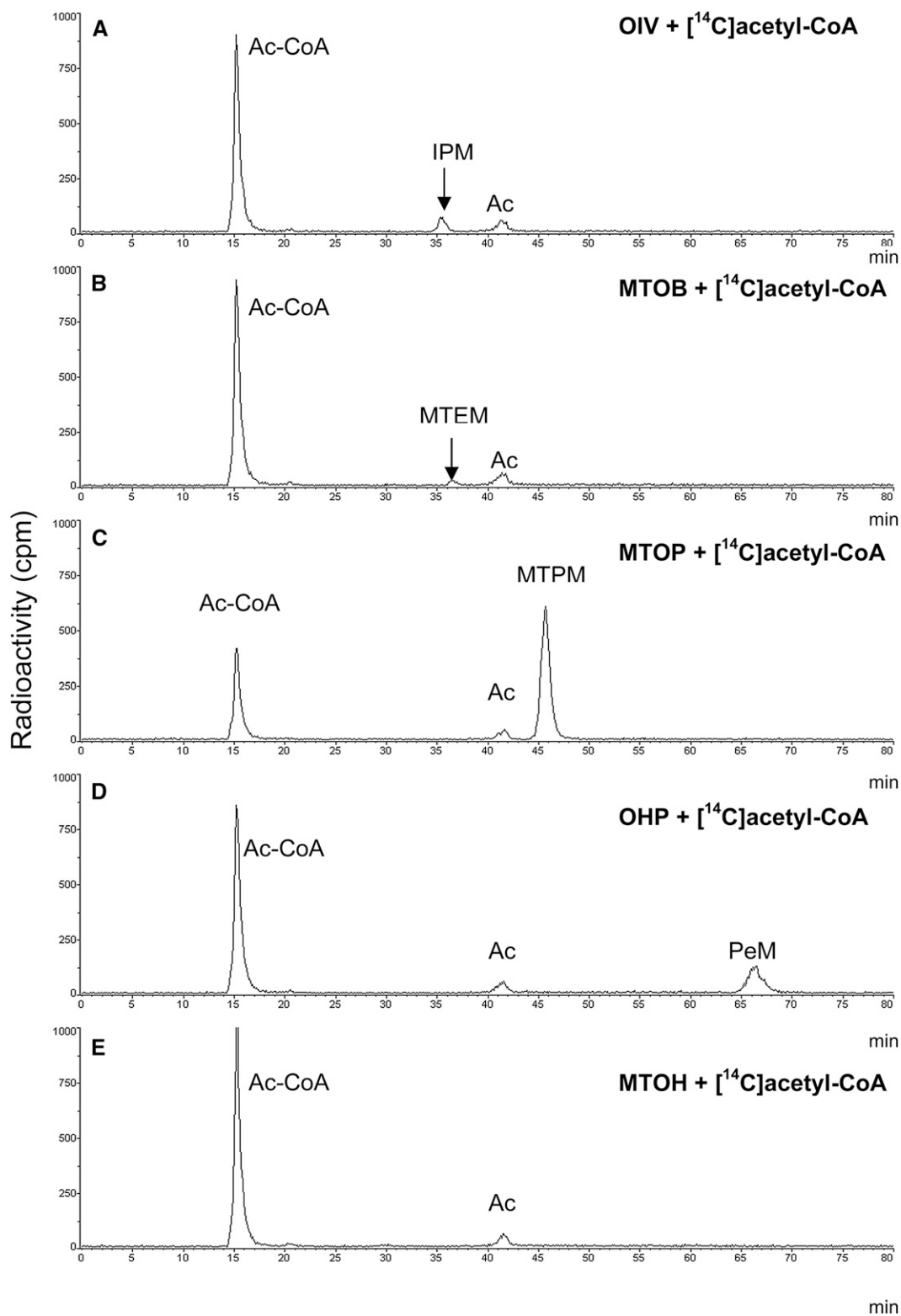
To test the effect of this sequence on cofactor preference, we replaced the HxQND sequence of IPMS2 along with three adjacent residues with the corresponding sequence of MAM1. This altered what is designated as Prosite sequence motif PS00816 (Wang et al., 1991; Sigrist et al., 2002) with the exception of the first residue. The effect of various divalent metal ions on activity of the resulting enzyme, named IPMS2/PS00816<sub>MAM</sub>, was investigated and compared with the effects on IPMS2, MAM3, and the IPMS of *E. coli* (Ec-IPMS) (Table 5). Before the assay, the enzymes were desalted to a buffer without any metal cations. Some residual activity was present that was completely inhibited by the addition of 10 mM EDTA in all but one case. In accordance with previous results, the preferred cofactor of IPMS2 was found to be Mg<sup>2+</sup>, whereas MAM3 preferred Mn<sup>2+</sup> (Falk et al., 2004; Textor et al., 2004). IPMS and MAM3 responded similarly to most of the other cations tested except that millimolar concentrations of Ca<sup>2+</sup> supported reaction only by MAM3.

By contrast, the mutated versions of IPMS2 had a cofactor preference like MAM3, with Mn<sup>2+</sup> being preferred over Mg<sup>2+</sup> (Table 5). In addition, Ca<sup>2+</sup> supported activity as it did with MAM3. Nevertheless, activity was only 30% of that of wild-type IPMS2, and the maximum enzyme activity of IPMS2/PS00816<sub>MAM</sub> was reached only at millimolar instead of micromolar concentration of Mn<sup>2+</sup>, indicating that other more subtle changes beyond the HxH[DN]D motif have an influence on cofactor binding as well. The altered cation preferences of IPMS2/PS00816<sub>MAM</sub> and IPMS2/-R2/PS00816<sub>MAM</sub> were not accompanied by any changes in specificity for the substrates

shown in Figure 6, irrespective of the cofactor used (data not shown). The cofactor preference of Ec-IPMS was very different from those of IPMS2 and MAM3. Zn<sup>2+</sup> was preferred and the metal cofactor was more tightly bound on desalting or the addition of EDTA. All other divalent cations with the exception of Co<sup>2+</sup> were inhibitory to Ec-IPMS, probably because they displace the Zn<sup>2+</sup> ion.

## DISCUSSION

The amino acid sequence of the glucosinolate biosynthetic enzyme MAM is very similar to that of IPMS, an enzyme involved in Leu biosynthesis (de Kraker et al., 2007; Benderoth et al., 2009). Based on phylogenetic analyses and an overlap of catalytic abilities, MAM most likely evolved from IPMS through a process of gene duplication and change in enzyme function (de Kraker et al., 2007; Benderoth et al., 2009). Using site-directed mutagenesis and domain-scale modifications, we demonstrated major alterations at the protein level that may have contributed to the functional changes from IPMS to MAM, including altered substrate specificity, loss of Leu feedback inhibition, and modification in divalent metal cofactor preference. Whereas this change can be formally considered a subfunctionalization (sensu Ohno, 1970) since some MAM activity was already present in IPMS, neofunctionalization may be a better term to describe the evolution of MAM since the MAM activity of IPMS is seen in vitro only at low rates (de Kraker et al., 2007; Textor et al., 2007) and is not detectable in vivo (de Kraker et al., 2007). Moreover, the original MAM activity of IPMS is able to catalyze the reaction with only the smallest MAM substrate. Nevertheless, due to the widespread debate about the distinctions between subfunctionalization, neofunctionalization, and other terms for designating the maintenance of duplicated genes (Hahn, 2009), it is perhaps



**Figure 7.** Radio-HPLC Traces of Assays of C-Terminal Truncated IPMS2 (IPMS2/-R2) with the Amino Acid Exchanges S216G and P252G in Its Active Site. The resulting protein shows a greater preference for one of the MAM substrates than for the IPMS substrate OIV in condensation with [ $^{14}\text{C}$ ]acetyl-CoA. Depicted are assays with the native IPMS substrate OIV (**A**), the MAM substrate MTOB (**B**), the elongated MAM substrate MTOP (**C**), the elongated

**Table 5.** Effect of Altering HxH[DN]D (Motif PS00816) on Divalent Cation Preference of IPMS2

Added Cofactor	Relative Substrate Conversion (%) <sup>a</sup>				
	IPMS2 HxQND	MAM3 HxHND	IPMS2/PS00816 <sub>MAM</sub> HxHND	IPMS2/-R2/PS00816 <sub>MAM</sub> HxHND	Ec-IPMS HxHDD
None	4	32	27	13	78
EDTA (10 mM)	0	0	0	0	30
Mg <sup>2+</sup> (4 mM) <sup>b</sup>	<b>100</b>	<b>73</b>	<b>79</b>	<b>83</b>	62
(40 μM)	<b>21</b>	<b>65</b>	29		
Mn <sup>2+</sup> (4 mM)	<b>56</b>	<b>100</b>	<b>100</b>	<b>100</b>	37
(40 μM)	<b>50</b>	<b>117</b>	<b>53</b>		
(4 μM)	<b>37</b>	<b>103</b>	<b>42</b>	<b>37</b>	
Ca <sup>2+</sup> (4 mM) <sup>b</sup>	0	<b>67</b>	<b>64</b>	<b>71</b>	58
Co <sup>2+</sup> (4 μM) <sup>c</sup>	<b>94</b>	<b>72</b>	<b>49</b>	<b>24</b>	<b>102</b>
Cu <sup>2+</sup> (4 μM) <sup>d</sup>	6	0	14	<b>28</b>	74
Fe <sup>2+</sup> (4 μM) <sup>d</sup>	0	27	4	15	65
Ni <sup>2+</sup> (4 μM) <sup>d</sup>	<b>27</b>	11	13	12	63
Zn <sup>2+</sup> (4 μM) <sup>d</sup>	<b>44</b>	39	32	13	<b>100</b>

Both full-length IPMS2 protein (IPMS2/PS00816<sub>MAM</sub>) and IPMS2 without its regulatory domain (IPMS2/-R2/PS00816<sub>MAM</sub>) were tested. Data are compared with the divalent cation preferences of wild-type IPMS2, MAM3, and *E. coli* IPMS. Enzyme activities that exceed the no-cofactor control assay by >10% are represented in bold.

<sup>a</sup>Incubations were done after desalting in presence of 6 mM OIV for IPMS enzymes (MTOB for MAM3), 1 mM acetyl-CoA, and 4.1, 7.4, 7.5, 4.1, and 18.1 μg of purified protein, respectively. The absolute activities corresponding to the 100% values are 132, 57, 51, 105, and 39 nmol 2-oxo acid product μg protein<sup>-1</sup> h<sup>-1</sup>, respectively.

<sup>b</sup>At 4 μM, there was no effect on activity of IPMS enzymes, but MAM3 activity was similar to assay with 40 μM Mg<sup>2+</sup>.

<sup>c</sup>At 4 mM, concentration was inhibitory.

<sup>d</sup>At 4 mM, concentration completely inhibited enzyme activity.

better to avoid their use in favor of more precise descriptions of enzyme function.

### Removal of the Regulatory Domain of IPMS2 Increases MAM Activity through Changes in Quaternary Structure

The absence in MAM proteins of ~120-amino acid residues at the C-terminal end corresponds to the loss of one exon in the *IPMS* gene structure (Benderoth et al., 2009). This region includes an allosteric Leu binding site and probably acts in all IPMSs as a regulatory domain in a way such as that described for Mt-IPMS (Koon et al., 2004). Accordingly, removal of this domain from *Arabidopsis* IPMS1 and IPMS2 resulted in a complete loss of Leu feedback inhibition (Figure 5) and a limited increase in enzyme activity, indicating a small inhibitory effect of this domain even in the absence of Leu (Table 2). Moreover, removal of the regulatory domain raised the enzyme activity with MTOB, a MAM substrate that is converted only at a very low rate by wild-type IPMS (see Supplemental Figure 1 online; Table 2). This increase in MAM activity is stronger for truncated IPMS2 without a regulatory domain (IPMS2/-R2) than for truncated IPMS1

(IPMS1/-R1) and is accompanied by a loss of quaternary structure (Table 3). In truncated IPMS1, the quaternary structure is retained and the increase in MAM activity is less evident.

The loss of quaternary structure in truncated IPMS2 is accompanied by an increased specificity constant for MTOB ( $k_{cat}/K_m$ ) of 253 M<sup>-1</sup> s<sup>-1</sup> that begins to approach the 834 M<sup>-1</sup> s<sup>-1</sup> of MAM1 (S. Textor and J. Gershenzon, unpublished data; assays performed as in Textor et al., 2007) and the 1380 M<sup>-1</sup> s<sup>-1</sup> of MAM3 (Textor et al., 2007), which also lack quaternary structure. Such a pronounced influence of quaternary structure on substrate specificity is not often reported. However, truncation of the N-terminal domain of *Thermus* sp maltogenic amylase also resulted in complete monomerization of the enzyme and increased activity toward a larger substrate, bulky starch, at the expense of activity toward small β-cyclodextrin (Kim et al., 2001). In the dimeric structure of Mt-IPMS, the active sites are covered by helix α10 of the opposite monomer, which stacks with His-397 and Tyr-410 in the active site. This feature is an important component of the allosteric inhibition mechanism and also places some restraints on the size of the catalytic site. The loss of quaternary structure in the truncated IPMS2/-R2 and the MAM proteins removes helix

**Figure 7.** (continued).

MAM substrate analog OHP (**D**), and the further elongated MAM substrate 6-methylthio-2-oxohexanoate (MTOH) (**E**). Assays were performed with 1 mM [<sup>14</sup>C]acetyl-CoA (Ac-CoA), 6 mM 2-oxoalkanoate, and 30 μg protein for 4 h at 30°C. The minor peak of [<sup>14</sup>C]acetate (Ac) that occurs in all incubations was also observed when no 2-oxoalkanoic acid substrate was added (data not shown). Kinetic data for the major activity (**D**) are given in the text. IPM, isopropylmalate; MTEM, 2-(2'-methylthio)ethylmalate; MTPM, 2-(3'-methylthio)propylmalate; PeM, 2-pentylmalate.

$\alpha$ 10 and can be assumed to increase the size of the active site, which is probably responsible for the improved enzyme activity with MTOB. An x-ray structure of the monomeric protein is necessary to confirm this conclusion.

### A Change of Two Amino Acids in the Active Site Gives a Preference for MAM Instead of IPMS Substrates

Although removal of the regulatory domain from IPMS2 results in a considerable increase of activity with MTOB, it does not explain why MAM enzymes have strongly reduced activities with OIV and accept substrates larger than MTOB. Thus, the evolution of MAM from IPMS must have involved other changes. In fact, the catalytic domain of the IPMS/MAM protein family contains a  $(\beta/\alpha)_8$  barrel, a very common fold said to be optimized for the evolution of new activities (Gerlt and Raushel, 2003).

We focused on the five amino acids of the catalytic domain that differ between IPMS and MAM and that in Mt-IPMS interact directly with the OIV substrate (Leu-143, His-167, Ser-216, Asn-250, and Pro-252; marked with black circles in Figure 3). Except for His-167, these amino acid residues are located at the C-terminal ends of the  $\beta$ -strands or within the loops that link  $\beta$ -strands with  $\alpha$ -helices. These are typical locations for amino acids in the active site of a  $(\beta/\alpha)_8$  barrel, a structure that is best described as a closed cylindrical parallel  $\beta$ -sheet formed from eight parallel  $\beta$ -strands surrounded by eight  $\alpha$ -helices. Because the R group residues of the amino acids surrounding the active site are structurally independent and stability is maintained by the opposite side of the barrel, alteration of functional groups at the ends of some  $\beta$ -strands, but not others, allows the evolution of new activities. Some examples from the literature show that single or multiple changes in amino acid residues (not necessarily near the catalytic center) can have a drastic effect on the substrate specificity of  $(\beta/\alpha)_8$  barrel enzymes (Gerlt and Raushel, 2003; Wise and Rayment, 2004; Höcker, 2005).

The combined exchange of two amino acid residues in truncated IPMS2, S216G plus P252G, resulted in a strong reduction of activity with OIV and a 10 times higher activity with MTOP (Figure 7), a MAM substrate that is not accepted at all by wild-type IPMS. By contrast, the individual exchanges of these two amino acids resulted in only a very low activity for MTOP that remained far below the activity for OIV (Table 4). The replacement of amino acid residues in the active site with Gly might be expected to accommodate larger substrates such as MTOP, as observed, but does not necessarily explain the loss of activity with the smaller substrate OIV. With a larger active site, smaller substrates may not be properly bound in the active site or correctly positioned with respect to the cosubstrate, acetyl-CoA. In Mt-IPMS, Ser-216 and Pro-252 both interact with one methyl group of OIV. The other methyl group has mainly hydrophobic interactions with residues Leu-143 and His-167. These two residues also differ in MAM, but exchanging the residues in IPMS for those found in MAM (Ile and Leu, respectively) did not increase activity with larger substrates. The H167L exchange led to a complete inactivation of the enzyme, probably because this residue has a critical position in the middle of a  $\beta$ -strand, while the L143I exchange had no effect on substrate preference at all.

The observed  $k_{\text{cat}}$  of the truncated S216G/P252G IPMS construct for MTOP was  $0.057 \text{ s}^{-1}$ , similar to those reported for *Arabidopsis* MAM3 with its substrates, which range from  $12.5 \text{ s}^{-1}$  for 6-methylthio-2-oxohexanoic acid to  $0.03 \text{ s}^{-1}$  for 9-methylthio-2-oxononanoic acid (Textor et al., 2007). However, the measured  $K_m$  value of this modified IPMS for MTOP was 4.12 mM, 4- to 10-fold higher than those of MAM3 for its substrates, so these two mutations did not create a typical MAM enzyme. In fact, the observed substrate specificity was different from all those observed among the MAMs so far. Of the natural MAM substrates, the truncated S216G/P252G IPMS2 prefers MTOP and has only low activity with MTOB (Table 4). MAM1 converts MTOP also but has its major activity with MTOB (Textor et al., 2004), a poor substrate for the modified IPMS2. MAM3 accepts a whole range of 2-oxo acids, including MTOB and MTOP (Textor et al., 2007). Only MAM2 activity is restricted to one substrate, but this is MTOB and not MTOP (Benderoth et al., 2006). However, the combination of amino acid residues in the active site of our construct (Leu-143, His-167, Gly-216, Gly-252, and Pro-252) is different from that in the characterized MAMs. Other combinations of the five residues might give a substrate specificity nearer to those of previously described MAM enzymes.

### Alterations in the HxH[DN]D Motif Shift Preference for the Divalent Metal Cofactor

The HxH[DN]D sequence motif (Prosite motif PS00816; Wang et al., 1991; Sigrist et al., 2002) determines the cofactor that is necessary for enzyme activity of the IPMS/MAM family. Exchanging the HcQND of IPMS2 for the HcHND of MAM1 (plus three adjacent residues) shifts the cofactor preference of IPMS2 from  $\text{Mg}^{2+}$  to  $\text{Mn}^{2+}$  (Table 5). This change replaces the sequence present in all known plant IPMSs with one that is conserved among most other IPMSs (Table 1). In the protein structure of Mt-IPMS (Koon et al., 2004), in which this motif appears as HpHND, it appears that the two His residues act as ligands for the metal cofactor together with a conserved Asp present in the adjacent PS80015 (Table 1). The presence of His (a nitrogen-containing ligand) instead of Gln (a mixed oxygen- and nitrogen-containing ligand) in this motif causes a preference for  $\text{Mn}^{2+}$  in MAM enzymes instead of for  $\text{Mg}^{2+}$  as in plant IPMS probably because  $\text{Mn}^{2+}$  has a slightly greater affinity for nitrogen-containing ligands than  $\text{Mg}^{2+}$  (Bock et al., 1999; Glusker et al., 1999).

Nonetheless, the amino acids of the HxH[DN]D motif do not exclusively determine divalent metal cofactor binding, since *E. coli* IPMS (HtHDD) and yeast IPMS (HcHND) employ  $\text{Zn}^{2+}$  preferentially (Table 5) (Roeder and Kohlhaw, 1980) despite having the same cofactor binding ligands as most other non-plant IPMSs. The existence of additional factors that determine the preferred metal cofactor is also demonstrated by the behavior of *Arabidopsis* IPMS2 in which the HxH[DN]D motif was exchanged for that of the MAM enzymes (IPMS2/PS00816<sub>MAM</sub>). This modified enzyme preferred  $\text{Mn}^{2+}$  instead of  $\text{Mg}^{2+}$  but required much higher concentrations of this cofactor for saturation than MAM3 does (4 mM versus 4  $\mu\text{M}$ ; Table 5). Possible reasons for this might be that the distance and/or angle between the cofactor binding amino acid residues in the modified enzyme are not optimal. These factors have been shown to play a role in the

effects of metal cofactors on the catalysis of other enzymes (e.g., *EcoRV* restriction endonuclease) (Vipond et al., 1996; Bock et al., 1999). In addition to  $Mg^{2+}$ ,  $Mn^{2+}$ , and  $Zn^{2+}$ , other divalent cations support catalysis to some extent, including  $Ca^{2+}$  and  $Co^{2+}$ . The ability of  $Ca^{2+}$  (ionic radius = 0.99 Å) to replace  $Mn^{2+}$  (ionic radius = 0.78 Å), but not  $Mg^{2+}$  (ionic radius = 0.62 Å), can be rationalized by their size differences (Fraústo da Silva and Williams, 2000). However, the differences in size between  $Mg^{2+}$  and  $Mn^{2+}$  are relatively small with respect to the length of the carbon-carbon single bond (1.20 to 1.54 Å). Thus, it is unlikely that switching cofactors would cause the differences in substrate specificity between IPMS and MAM, a premise substantiated in this investigation (see last section of Results). In agreement with previous literature (Fraústo da Silva and Williams, 2000), the different cofactors were observed to bind at different strengths to the proteins ( $Mg^{2+} < Mn^{2+} < Zn^{2+}$ ) as shown by the amount of residual activity with no added cofactor (after desalting) or after addition of EDTA (Table 5). The strengths of these conclusions must be tempered by the fact that the recombinant proteins investigated here all had a His-tag present to aid in purification. This modification may have altered cofactor binding compared with the respective native proteins. Nevertheless, the His-tag probably did not have a large influence or all the proteins would have been much more similarly influenced by changes in the type and concentration of divalent metal ion cofactors.

It is unclear why the known plant IPMSs, including those of chlorophytes, contain a HxQND sequence and function with  $Mg^{2+}$ , whereas the IPMSs of microorganisms have a HxH[DN]D sequence and function with  $Zn^{2+}$  or  $Mn^{2+}$  (de Kraker et al., 2007). A possible explanation is the localization of plant IPMSs in the chloroplast (Hagelstein and Schultz, 1993; Falk et al., 2004). Under illumination, free  $Mg^{2+}$  increases from ~1 to 3 mM to ~3 to 6 mM in the stroma of the chloroplast, and this change is an important regulator of enzymes of the Calvin cycle (Fraústo da Silva and Williams, 2000; Malkin and Niyogi, 2000). Changes in  $Mg^{2+}$  concentration within such a millimolar range can also affect IPMS activity (de Kraker et al., 2007), and in this way, Leu biosynthesis might be adjusted to coincide with photosynthetic activity. Although the MAM enzymes evolved from plant IPMSs, they have curiously returned to having  $Mn^{2+}$  as their preferred cofactor despite MAM being localized in the chloroplast as well (Textor et al., 2007). Micromolar concentrations of free  $Mn^{2+}$  are possible in chloroplasts also, as this cation is present in such concentrations everywhere in the cell, except for the cytoplasm where it is maintained at  $10^{-7}$  M (Fraústo da Silva and Williams, 2000). However, a change to  $Mn^{2+}$  as cofactor instead of  $Mg^{2+}$  could allow MAM activity to be regulated in a different way, independent from fluctuating  $Mg^{2+}$  concentrations. The association of MAM (and other glucosinolate biosynthetic enzymes) with vascular tissue, in particular the phloem parenchyma (Schuster et al., 2006), may also necessitate a different type of regulation than that for IPMS. Based on our *in vitro* measurements (Table 5),  $Ca^{2+}$  and  $Co^{2+}$  could also support MAM activity at least partially *in vivo*. However, millimolar concentrations of free  $Ca^{2+}$  do not occur in chloroplasts, and  $Co^{2+}$  is generally not a cofactor for plant enzymes (Johnson et al., 1995; Fraústo da Silva and Williams, 2000).

### A Combination of Various Changes Has Converted IPMS to MAM

By removing the C-terminal domain of IPMS and making a few amino acid exchanges in the active site, we altered its substrate preference from OIV to MTOP, a substrate for the MAM enzymes of glucosinolate chain elongation. These changes reflect sequence alterations that appear to have occurred in the evolutionary recruitment of IPMS from amino acid to glucosinolate metabolism. Our results suggest that the loss of the regulatory domain was one of the early events in this recruitment because the absence of the C terminus leads to changes in relative substrate preference and to the loss of Leu feedback inhibition, both features of MAM. After the loss of the regulatory domain, a strong selection against any remaining IPMS activity could be expected because of the negative effects of unregulated Leu formation. Overexpression of a *Brassica* IPMS gene in *Arabidopsis* resulted in aberrant morphology and perturbed amino acid metabolism (Field et al., 2006).

The timing of other events in the evolution of MAM from IPMS is uncertain. Changes in the active site that increase MAM activity relative to IPMS activity may have preceded or followed the loss of the regulatory domain. The timing of the change in cofactor preference from  $Mg^{2+}$  to  $Mn^{2+}$  is also not known. However, as the nature of the cofactor has no effect on substrate preference, we assume that this transition occurred at a late stage in the evolution of MAM from IPMS.

The creation of a MAM activity may have driven the evolution of the entire Met chain elongation pathway since products of the MAM reaction are likely to be converted successively by the later enzymes in Leu biosynthesis. The ability of Leu biosynthetic enzymes to accept alternative substrates has already been demonstrated in *E. coli* for formation of norleucine (Bogosian et al., 1989) and in *Serratia marcescens* for norvaline (Kisumi et al., 1976), both of which are side chain-elongated nonprotein amino acids also occurring in plants (Nikiforova et al., 2005; Urbanczyk-Wochniak and Fernie, 2005). In the case of MAM products arising by mutation of IPMS, it is likely that they would be converted at an appreciable rate to elongated, Met-derived 2-oxo acids by the remaining two enzymes of Leu biosynthesis. These two enzymes have a very high similarity to the last two enzymes of Met chain elongation and their activities overlap at least partially. The heteromeric isopropylmalate isomerase of *Arabidopsis* has a large subunit that functions in either Leu or glucosinolate biosynthesis, whereas the small subunit appears to be specialized for each pathway (Knill et al., 2009). The isopropylmalate dehydrogenase of *Arabidopsis* may even serve in both Met chain elongation and Leu biosynthesis (He et al., 2009), although a direct comparison between Met elongation and Leu substrates *in vitro* is necessary to confirm this supposition.

In summary, we engineered a new enzyme from IPMS2 that prefers MTOP as a substrate over OIV, uses  $Mn^{2+}$  as cofactor instead of  $Mg^{2+}$ , and is not under control of Leu feedback inhibition. Although the enzyme does not display the exact substrate specificity and kinetic behavior of previously described MAM proteins, the changes we made closely mimic those in the evolution of MAM from IPMS. Thus, this research has been able to identify alterations at the protein level that lead to the

recruitment of an enzyme from primary to secondary metabolism. The results also provide a foundation for making further changes in substrate specificity for all members of the IPMS/MAM family. This approach should prove valuable in modifying the chain length of aliphatic glucosinolates in crop plants to improve insect resistance or increase their value as functional foods for cancer prevention.

## METHODS

### Protein Sequence Alignment and Structure Prediction

The alignment and secondary structure prediction shown in Figure 3 was made on the NPS@web server (Combet et al., 2000) in ClustalW using default settings. Secondary structure was predicted with the program PHD (Rost and Sander, 1993), and the representation was made with ESPript (Gouet et al., 1999). Structural data for Mt-IPMS (PDB: 1sr9) were accessed through the RCSB Protein Databank (<http://www.rcsb.org>). The molecular structure viewer used was an educational version of PYMOL 1.3 (DeLano Scientific). The LIGPLOT of Mt-IPMS was accessed via PDBsum at [www.ebi.ac.uk/pdbsum](http://www.ebi.ac.uk/pdbsum) (Wallace et al., 1995; Laskowski et al., 2005), and the output was edited for improved readability with an Acrobat PDF file editor.

### Assembly of IPMS/MAM cDNA Constructs

The cloning of *IPMS1*, *IPMS2*, *MAM1*, and *MAM3* cDNA from *Arabidopsis thaliana* Columbia-0 in vector pCR-T7/CT-TOPO (Invitrogen) has been previously described (Textor et al., 2004, 2007; de Kraker et al., 2007); all constructs have a truncated open reading frame that lacks the putative chloroplast transit peptide (ChloroP; Emanuelsson et al., 1999). Minipreps (Invisorb spin plasmid mini kit; Invitex) of these constructs were used to make *IPMS1/-R1*, *IPMS2/-R2*, *MAM1/+R2*, *MAM3/+R2*, *IPMS2/PS00816<sub>MAM</sub>*, and *IPMS2/-R2/PS00816<sub>MAM</sub>*. cDNAs of *IPMS1/-R1* and *IPMS2/-R2* were amplified from *IPMS1* and *IPMS2* with the Expand High Fidelity System (Roche) using primer pair 1ipms1i+atg/ipms1-short and 1ipms2m+atg/ipms2-short, respectively (for all primers used, see Supplemental Table 1 online). The PCR product was directly cloned into pCR-T7/CT-TOPO according to the manufacturer's instructions. The new constructs, like all other constructs described in this section, were transformed into Top10 *Escherichia coli* cells (Invitrogen), and the DNA of transformed colonies was purified by miniprep and sequenced on an ABI 3700 DNA sequencer with Big Dye terminators (Applied Biosystems).

The regulatory domain of *IPMS2*, *R2*, was added to *MAM1* with the help of overlapping primers by a method called Splicing by Overlap Extension or gene SOEing (Horton et al., 1989). First, in two different PCR reactions, the primer pair mam1xf/mam1+sttipms2 was used to amplify the *MAM1* fragment (1289 bp) and the primer pair 2ipms2n/ipms2+sttmam1 to amplify the *R2* encoding fragment (459 kb). PCR reactions were performed with 0.5  $\mu$ L of Miniprep ( $\pm$  70 ng DNA of *MAM1* and *IPMS2*, respectively) in a total volume of 50  $\mu$ L containing 5  $\mu$ L of each primer (50 pmol/mL), 5  $\mu$ L of supplied 10 $\times$  buffer, 5  $\mu$ L deoxynucleotide triphosphate (2 mM), and 1  $\mu$ L High Fidelity taq polymerase. To a third new PCR, 0.5  $\mu$ L of each previous PCR mix (*MAM1* plus *R2* fragment) was added to the primer pair mam1xf/2ipms2n. This PCR resulted in a fusion product *MAM1/+R2* of 1749 bp (confirmed on agarose gel electrophoresis) that was directly cloned in pCR-T7/CT-TOPO by adding 0.5  $\mu$ L of the PCR mixture to 1  $\mu$ L of vector according to the manufacturer's instructions. In the same way, but using different primers, *R2* was added to *MAM3*, yielding *MAM3/+R2*.

The constructs *IPMS2/PS00816<sub>MAM</sub>* and *IPMS2/-R2/PS00816<sub>MAM</sub>* were made in a very similar way from *IPMS2* and *IPMS2/-R2* through

gene SOEing. Each gene was amplified as two separate cDNA fragments: one fragment before the region encoding for cofactor binding, PS00816, and a second fragment after this region. The SOE-primers for the first two reactions had overhangs that overlapped the PS00816 region, which during the third PCR resulted in a fusion product of both fragments containing the PS00816 region as present in *MAM1* with exception of the first residue (Ile was not changed into Leu). Point mutations (L1431, H167L, S216G, N250G, and P252G) and combinations of these were introduced using the Quick Change site-directed mutagenesis kit of Stratagene (primers given in Supplemental Table 1 online). The cloning of the *EcIPMS* has been described before (de Kraker et al., 2007).

### Heterologous Expression and Enzyme Purification

The pCR-T7/CT-TOPO constructs were expressed in *E. coli* strain BL21 (DE3) pLysS (Invitrogen) at 18°C using 2 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside. A bacterial pellet originating from 50 mL of induced culture was homogenized in 1.5 mL of lysis buffer containing 50 mM Tris, pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM MgCl<sub>2</sub>, and cell debris was precipitated by centrifugation (7 min, 20,500g, 4°C). The expressed protein was purified from the bacterial lysate over a small column filled with 1.5 mL of 50% Ni-NTA agarose (Qiagen). The column was washed with 12 mL of lysis buffer containing 20 mM imidazole, and after that the His-tagged protein was eluted with 3 mL of lysis buffer containing 250 mM imidazole. The eluent was immediately transferred to an Econo-Pac 10 DG column (Bio-Rad) and desalted into 4 mL of a 50 mM Tris buffer, pH 8.0, containing 10 mM MgCl<sub>2</sub> and 10% glycerol. For experiments determining the cofactor preferences of the enzyme, the eluent was desalted into buffer without MgCl<sub>2</sub>, and only the first 3.5 mL were collected. Protein purity was checked on SDS-PAGE gels stained with Coomassie Brilliant Blue and was usually above 90%. The protein concentration of each preparation was determined with the BCA protein assay kit (Pierce) using BSA as a standard and usually ranged from 400 to 800  $\mu$ g/mL. More details of the protein expression and purification have been described elsewhere (de Kraker et al., 2007).

### Enzyme Assays and Enzyme Characterization

The substrate specificity of *IPMS1/-R1*, *IPMS2/-R2*, *IPMS2/PS00816<sub>MAM</sub>*, *IPMS2/-R2/PS00816<sub>MAM</sub>*, *MAM1/+R2*, and *MAM3/+R2* were tested with a qualitative enzyme assay with [<sup>14</sup>C]acetyl-CoA and radio-HPLC detection as previously described (de Kraker et al., 2007; Textor et al., 2007) using the following substrates (all purchased from Sigma-Aldrich, Fluka, or Merck unless stated): glyoxylate, pyruvate, 2-oxobutyrates, OIV, 2-oxovalerate, 3-methyl-2-oxovalerate, 2-oxohexanoate, and MTOB, OHP (Falk et al., 2004), MTOP (Applichem), 2-oxooctanoate, and 6-methylthio-2-oxohexanoate (Applichem). If not stated otherwise, the standard incubation at 30°C contained 100  $\mu$ L of enzyme preparation (diluted to 275  $\mu$ g mL<sup>-1</sup> in desalting buffer), 6 mM of 2-oxoacid substrate, 1 mM [<sup>14</sup>C]acetyl-CoA (0.4 mCi mmol<sup>-1</sup>), and 4 mM MgCl<sub>2</sub> in a final volume of 250  $\mu$ L 100 mM Tris buffer, pH 8.0, and was stopped after 1 h by adding 750  $\mu$ L ethanol.

Kinetic data for *IPMS1*, *IPMS2*, *IPMS1/-R1*, *IPMS2/-R2*, and *IPMS2/-R2* with S216G/P252G were determined with a spectrophotometric endpoint assay with DTNB (Sigma-Aldrich) (de Kraker et al., 2007). Each incubation contained 50  $\mu$ L of diluted enzyme preparation, 500  $\mu$ M acetyl-CoA, 10 mM MgCl<sub>2</sub>, and 0.06 to 10.0 mM OIV or 0.2 to 10 mM MTOB, in a final volume of 250  $\mu$ L 100 mM Tris, pH 8.0. The amount of purified enzyme protein per assay was 0.4 to 0.6  $\mu$ g in determinations with OIV and 20 to 40  $\mu$ g in determinations with MTOB; for *IPMS2/-R2*, 0.1 to 0.2  $\mu$ g and 10 to 20  $\mu$ g were used, respectively. In case of *IPMS2/-R2* with S216G/P252G,  $\sim$ 30  $\mu$ g of enzyme protein were used in the assay with 1 mM acetyl-CoA and 0.2 to 20 mM MTOP. The incubations at 30°C



were stopped after 10 min by freezing in liquid nitrogen. After addition of 200  $\mu\text{L}$  ethanol and 100  $\mu\text{L}$  of 2 mM DTNB (fresh solution in 100 mM Tris, pH 8.0), the absorbance of the assays was measured against water at 412 nm ( $\epsilon_{412}$  is 14,140  $\text{M}^{-1} \text{cm}^{-1}$ ). Incubations without substrate were used as a blank, and corrections were made for the small color development in the presence of millimolar concentrations of MTOB. The enzyme assay under these conditions was linear as determined by assessing activity at numerous different protein concentrations and time intervals. The activity at each substrate concentration was assayed in triplicate, and the  $K_m$  and  $V_{\text{max}}$  values were calculated from four independent experiments using the Enzyme Kinetic Module (version 1.1) of Sigmaplot (version 8.0).

The spectrophotometric assay was also used to test the effect of Leu on enzyme activity in the range from 25  $\mu\text{M}$  to 10 mM. The cofactor preference of the recombinant enzymes was tested with the qualitative radio-HPLC assay that, unlike the spectrophotometric assay, is compatible with all the divalent ions tested. The chloride salts of the cations were tested in concentrations and enzyme quantities mentioned in Table 5, and incubations were stopped after 20 min. Each cation and concentration was assayed twice in each experiment, and enzyme activity was expressed as a percentage of the activity obtained by incubation with the preferred cation. Each experiment was repeated at least twice.

The molecular masses of purified recombinant proteins were estimated by exclusion chromatography on a Superdex 200 column (Hiload 16/60) and run at 0.75  $\text{mL min}^{-1}$  with a buffer consisting of 50 mM Tris, pH 8.0, 150 mM NaCl, and 1 mM  $\text{MgCl}_2$ . The column was calibrated with proteins from a molecular mass marker kit for gel filtration chromatography (Sigma-Aldrich) employing proteins known to be nearly spherical and exhibiting regular, globular elution behavior (Whitaker, 1963):  $\beta$ -amylase (200 kD), alcohol dehydrogenase (150 kD), BSA (66 kD), carbonic anhydrase (29 kD), and cytochrome C (12.4 kD). Before size exclusion, proteins were purified on the Ni-NTA column (eluted in 1.5 mL) and then desalted through a HiTrap desalting column (GE Healthcare) in 2 mL of 50 mM Tris buffer, pH 8.0, with 1 mM  $\text{MgCl}_2$ . After loading 1 mL, fractions of 1 mL were collected and protein peaks recorded at 280 nm. Individual fractions were tested for IPMS activity as previously described (de Kraker et al., 2007) and analyzed for purity by SDS gel electrophoresis. Enzyme activity of active fractions from MAM1, MAM3, MAM1/+R2, and MAM3/+R2 chromatography was verified with the qualitative radio-HPLC assay in the presence of MTOB, and an additional 2 mM DTT. In the case of MAM1, 2 mM ATP was added (Textor et al., 2004). Each exclusion chromatography run was repeated at least twice.

#### Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: IPMS1, At1g 18500; IPMS2, At1g74040; MAM1, At5g23010; MAM2, Q8VX04; MAM3, At5g23020; *S. pennellii* IPMSa, O04973; *S. pennellii* IPMSb, O04974; *O. sativa* IPMSa, Q0IUQ0; *O. sativa* IPMSb, Q2QXY9; *O. tauri* IPMS, Q01FR2; *E. coli* IPMS (*LeuA*), P09151; *M. tuberculosis* IPMS, P96420; *S. cerevisiae* IPMS (*LEU4*), C8ZGC6; *A. variabilis* IPMS, Q3MBA3; *M. jannaschii* IPMS, Q58595; *A. lyrata* MAMa, Q1JRZ2; Mt-IPMS structure used is listed in PDB as 1sr9.

#### Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Radio-HPLC Traces of Assays of *Arabidopsis* IPMS2 and IPMS2 Modified to Remove the C-Terminal Regulatory Domain (IPMS2/-R2).

**Supplemental Table 1.** Oligonucleotide Primers Used in This Study.

#### ACKNOWLEDGMENTS

We thank Susanne Textor and Meike Burow for useful suggestions and discussions, Beate Rothe for excellent technical assistance, and the Max Planck Society for financial support.

Received September 1, 2010; revised December 2, 2010; accepted December 16, 2010; published January 4, 2011.

#### REFERENCES

- Benderoth, M., Pfalz, M., and Kroymann, J. (2009). Methylthioalkylmalate synthases: Genetics, ecology, and evolution. *Phytochem. Rev.* **8**: 255–268.
- 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.
- Biaostoff, S., Reinhardt, N., Reva, V., Brandt, W., and Dräger, B. (2009). Evolution of putrescine N-methyltransferase from spermidine synthase demanded alterations in substrate binding. *FEBS Lett.* **583**: 3367–3374.
- Bock, C.W., Katz, A.K., Markham, G.D., and Glusker, J.P. (1999). Manganese as a replacement for magnesium and zinc: Functional comparison of the divalent ions. *J. Am. Chem. Soc.* **121**: 7360–7372.
- Bogosian, G., Violand, B.N., Dorward-King, E.J., Workman, W.E., Jung, P.E., and Kane, J.F. (1989). Biosynthesis and incorporation into protein of norleucine by *Escherichia coli*. *J. Biol. Chem.* **264**: 531–539.
- Combet, C., Blanchet, C., Geourjon, C., and Deléage, G. (2000). NPS@: Network protein sequence analysis. *Trends Biochem. Sci.* **25**: 147–150.
- 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.
- Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**: 978–984.
- Falk, K.L., Vogel, C., Textor, S., Bartram, S., Hick, A., Pickett, J.A., and Gershenzon, J. (2004). Glucosinolate biosynthesis: Demonstration and characterization of the condensing enzyme of the chain elongation cycle in *Eruca sativa*. *Phytochemistry* **65**: 1073–1084.
- Field, B., Furniss, C., Wilkinson, A., and Mithen, R. (2006). Expression of a Brassica isopropylmalate synthase gene in Arabidopsis perturbs both glucosinolate and amino acid metabolism. *Plant Mol. Biol.* **60**: 717–727.
- Fraústo da Silva, J.J.R., and Williams, R.J.P. (2000). *The Biological Chemistry of the Elements*. (Oxford, UK: Oxford University Press).
- Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., and Gierl, A. (1997). Analysis of a chemical plant defense mechanism in grasses. *Science* **277**: 696–699.
- Frey, M., Stettner, C., Pare, P.W., Schmelz, E.A., Tumlinson, J.H., and Gierl, A. (2000). An herbivore elicitor activates the gene for indole emission in maize. *Proc. Natl. Acad. Sci. USA* **97**: 14801–14806.
- Gerlt, J.A., and Raushel, F.M. (2003). Evolution of function in ( $\beta/\alpha$ )<sub>8</sub>-barrel enzymes. *Curr. Opin. Chem. Biol.* **7**: 252–264.
- Gierl, A., and Frey, M. (2001). Evolution of benzoxazinone biosynthesis and indole production in maize. *Planta* **213**: 493–498.
- Glusker, J.P., Katz, A.K., and Bock, C.W. (1999). Metal ions in biological systems. *Rigaku J.* **16**: 8–16.

- Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999). ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305–308.
- Hagelstein, P., and Schultz, G. (1993). Leucine synthesis in spinach chloroplasts: Partial characterization of 2-isopropylmalate synthase. *Biol. Chem. Hoppe Seyler* **374**: 1105–1108.
- Hahn, M.W. (2009). Distinguishing among evolutionary models for the maintenance of gene duplicates. *J. Hered.* **100**: 605–617.
- Halkier, B.A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol.* **57**: 303–333.
- Halkier, B.A., Hansen, C.H., Mikkelsen, M.D., Naur, P., and Wittstock, U. (2002). The role of cytochromes P450 in biosynthesis and evolution of glucosinolates. In *Phytochemistry in the Genomics and Post-Genomics Eras*, Vol. 36, Recent Advances in Phytochemistry, J.T. Romeo and R.A. Dixon, eds (Amsterdam: Pergamon Press), pp. 223–248.
- 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.
- Höcker, B. (2005). Directed evolution of (betaalpha)(8)-barrel enzymes. *Biomol. Eng.* **22**: 31–38.
- Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**: 61–68.
- Johnson, C.H., Knight, M.R., Kondo, T., Masson, P., Sedbrook, J., Haley, A., and Trewavas, A. (1995). Circadian oscillations of cytosolic and chloroplastic free calcium in plants. *Science* **269**: 1863–1865.
- Kim, T.J., et al. (2001). Modulation of the multisubstrate specificity of *Thermus* maltogenic amylase by truncation of the N-terminal domain and by a salt-induced shift of the monomer/dimer equilibrium. *Biochemistry* **40**: 14182–14190.
- Kisumi, M., Sugiura, M., and Chibata, I. (1976). Biosynthesis of norvaline, norleucine, and homoisoleucine in *Serratia marcescens*. *J. Biochem.* **80**: 333–339.
- Knill, T., Reichelt, M., Paetz, C., Gershenzon, J., and Binder, S. (2009). *Arabidopsis thaliana* encodes a bacterial-type heterodimeric isopropylmalate isomerase involved in both Leu biosynthesis and the Met chain elongation pathway of glucosinolate formation. *Plant Mol. Biol.* **71**: 227–239.
- 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., Tokuhsa, 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.
- Laskowski, R.A., Chistyakov, V.V., and Thornton, J.M. (2005). PDBsum more: New summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.* **33** (Database issue): D266–D268.
- Malkin, R., and Niyogi, K. (2000). Photosynthesis. In *Biochemistry and Molecular Biology of Plants*, B. Buchanan, W. Gruissem, and R. Jones, eds (Rockville, MD: American Society of Plant Physiologists), pp. 568–628.
- Modolo, L.V., Reichert, A.I., and Dixon, R.A. (2009). Introduction to the different classes of biosynthetic enzymes. In *Plant-Derived Natural Products*, A.E. Osbourn and V. Lanzotti, eds (Dordrecht, The Netherlands: Springer), pp. 143–163.
- Nikiforova, V.J., Kopka, J., Tolstikov, V., Fiehn, O., Hopkins, L., Hawkesford, M.J., Hesse, H., and Hoefgen, R. (2005). Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of *Arabidopsis* plants. *Plant Physiol.* **138**: 304–318.
- Ober, D., and Hartmann, T. (1999). Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. *Proc. Natl. Acad. Sci. USA* **96**: 14777–14782.
- Ober, D., and Hartmann, T. (2000). Phylogenetic origin of a secondary pathway: The case of pyrrolizidine alkaloids. *Plant Mol. Biol.* **44**: 445–450.
- Ohno, S. (1970). *Evolution by Gene Duplication*. (Berlin: Springer Verlag).
- Roeder, P.R., and Kohlhaw, G.B. (1980). alpha-Isopropylmalate synthase from yeast. A zinc metalloenzyme. *Biochim. Biophys. Acta* **613**: 482–487.
- Rost, B., and Sander, C. (1993). Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proc. Natl. Acad. Sci. USA* **90**: 7558–7562.
- 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.
- Sigrist, C.J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., and Bucher, P. (2002). PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief. Bioinform.* **3**: 265–274.
- Singh, K., and Bhakuni, V. (2007). Cation induced differential effect on structural and functional properties of *Mycobacterium tuberculosis* alpha-isopropylmalate synthase. *BMC Struct. Biol.* **7**: 39.
- Teuber, M., Azemi, M.E., Namjoyan, F., Meier, A.C., Wodak, A., Brandt, W., and Dräger, B. (2007). Putrescine *N*-methyltransferases—a structure-function analysis. *Plant Mol. Biol.* **63**: 787–801.
- 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 Tokuhsa, J.G. (2007). MAM3 catalyzes the formation of all aliphatic glucosinolate chain lengths in *Arabidopsis*. *Plant Physiol.* **144**: 60–71.
- Urbanczyk-Wochniak, E., and Fernie, A.R. (2005). Metabolic profiling reveals altered nitrogen nutrient regimes have diverse effects on the metabolism of hydroponically-grown tomato (*Solanum lycopersicum*) plants. *J. Exp. Bot.* **56**: 309–321.
- Vipond, I.B., Moon, B.J., and Halford, S.E. (1996). An isoleucine to leucine mutation that switches the cofactor requirement of the EcoRV restriction endonuclease from magnesium to manganese. *Biochemistry* **35**: 1712–1721.
- Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**: 127–134.
- Wang, S.Z., Dean, D.R., Chen, J.S., and Johnson, J.L. (1991). The N-terminal and C-terminal portions of NifV are encoded by two different genes in *Clostridium pasteurianum*. *J. Bacteriol.* **173**: 3041–3046.
- Whitaker, J.R. (1963). Determination of molecular weights of proteins by gel filtration on Sephadex. *Anal. Chem.* **35**: 1950–1953.
- Wise, E.L., and Rayment, I. (2004). Understanding the importance of protein structure to nature's routes for divergent evolution in TIM barrel enzymes. *Acc. Chem. Res.* **37**: 149–158.