Structural Basis for Substrate Recognition in the Salicylic Acid Carboxyl Methyltransferase Family

Chloe Zubieta,a,b,1 Jeannine R. Ross,a,c Paul Koscheski,a Yue Yang,c Eran Pichersky,c and Joseph P. Noela,b,2

a Structural Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037
b Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92037
c Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109-1048

Recently, a novel family of methyltransferases was identified in plants. Some members of this newly discovered and recently characterized methyltransferase family catalyze the formation of small-molecule methyl esters using S-adenosyl-L-Met (SAM) as a methyl donor and carboxylate-bearing substrates as methyl acceptors. These enzymes include SAMT (SAM:salicylic acid carboxyl methyltransferase), BAMT (SAM:benzoic acid carboxyl methyltransferase), and JMT (SAM:jasmonic acid carboxyl methyltransferase). Moreover, other members of this family of plant methyltransferases have been found to catalyze the N-methylation of caffeine precursors. The 3.0-Å crystal structure of Clarkia breweri SAMT in complex with the substrate salicylic acid and the demethylated product S-adenosyl-L-homocysteine reveals a protein structure that possesses a helical active site capping domain and a unique dimerization interface. In addition, the chemical determinants responsible for the selectivity of salicylic acid demonstrate the structural basis for facile variations of substrate selectivity among functionally characterized plant carboxyl-directed and nitrogen-directed methyltransferases and a growing set of related proteins that have yet to be examined biochemically. Using the three-dimensional structure of SAMT as a guide, we examined the substrate specificity of SAMT by site-directed mutagenesis and activity assays against 12 carboxyl-containing small molecules. Moreover, the utility of structural information for the functional characterization of this large family of plant methyltransferases was demonstrated by the discovery of an Arabidopsis methyltransferase that is specific for the carboxyl-bearing phytohormone indole-3-acetic acid.

INTRODUCTION

Carboxylate-bearing small molecules are an important class of compounds that plants produce to defend against herbivory (Karban, 1999; Van Poecke et al., 2001; Winz and Baldwin, 2001), to regulate the cell cycle (Staswick et al., 1992; Smirnoff and Wheeler, 2000), and to induce defense responses to pathogen attack (Ryals et al., 1995; Feys and Parker, 2000; Nurnberger and Scheel, 2001). Such compounds often are esterified and volatilized to serve as attractants for pollinators (Raguso and Pichersky, 1995; Dudareva et al., 2000) and for other types of plant–insect and plant–plant communications (Shulaev et al., 1997; Pare and Tumlinson, 1999; Kessler and Baldwin, 2001). Examples of volatile esters include methyl salicylate (MSA) and methyl jasmonate (MJA), both of which are found throughout the plant kingdom.

Jasmonic acid (JA) is intimately involved in cellular regulation and the control of such developmental processes as seed germination, flower development, fruit development, and senescence (Gundlach et al., 1992; Seo et al., 2001). Evidence supports the role of the methyl ester of JA, MJA, as an interplant signaling molecule produced by an infected plant that, upon release, subsequently induces defense responses in uninfected neighboring plants (Farmer and Ryan, 1990; Reymond and Farmer, 1998; Kessler and Baldwin, 2001).

Salicylic acid (SA) and possibly its volatile methyl ester, MSA, are involved in localized and systemic defense responses (Chen et al., 1995; Ryals et al., 1995; Seskar et al., 1998; Nurnberger and Scheel, 2001). For example, upon pathogen challenge, plants increase SA biosynthesis and mount an initial defense response. Production of SA can trigger localized acquired resistance to subsequent pathogen attack and induce systemic acquired resistance, in which pathogen resistance spreads beyond the initial site of challenge, systemically, throughout the plant (Chen et al., 1995; Hunt et al., 1996). Like MJA, the volatile methyl ester of SA, MSA, is hypothesized to act as an airborne signal that triggers defense responses in uninfected organs of the emitting plant or in adjacent, unchallenged plants (Shulaev et al., 1997).

Originally isolated from the petals of Clarkia breweri, SA carboxyl methyltransferase (SAMT) catalyzes the formation of MSA from SA (Figure 1A) (Ross et al., 1999). Apart from its putative roles in defense responses, MSA, with its characteristic wintergreen scent, is an important chemoattractant for moth-pollinated flowering plants, including Clarkia (Raguso and Pichersky, 1995; Raguso et al., 1996; Dudareva et al., 1998). When first characterized, the primary amino acid sequence of SAMT showed no significant similarity to any other known methyltransferases from any organism, including plants. Subsequently, two additional carboxyl methyltransferases with sequence similarity to SAMT were identified:
H₂O is depicted as the proton acceptor from the free acid or xanthine nitrogens forming the reactive methyl acceptors and a hydronium ion (H₃O⁺). Red letters highlight the atom undergoing methylation, and the transferred methyl groups are highlighted with blue letters.

(A) SAMT. (B) JMT. (C) 7-MXMT (alternatively theobromine synthase). Caffeine synthase from C. sinensis (tea) is reported to act as both a 7-MXMT and a 3,7-DMXMT (Kato et al., 2000). (D) 3,7-DMXMT (alternatively caffeine synthase).

BAMT (SAM:benzoic acid carboxyl methyltransferase) and JMT (JA carboxyl methyltransferase) (Figure 1B). In addition, it has been shown that this new family of methyltransferases also includes several alkaloid N-methyltransferases involved in the biosynthesis of caffeine (Kato et al., 2000; Ogawa et al., 2001) (Figures 1C and 1D).

The widely divergent chemical structures of newly identified substrates for this family of plant methyltransferases, as well as the large number of putative SAMT-like proteins found in many plant genomic and EST sequencing projects (including that for Arabidopsis), make structural studies an important tool to help identify potential substrate targets of uncharacterized methyltransferases of this family. A number of plant hormones and second messengers, including auxins, gibberellins, abscisic acid, and cytokinins, contain carboxyl- and/or nitrogen-containing groups that can serve as methyl acceptors. The identification of active site residues coupled with modeling studies provides a structure-based methodology valuable for the rapid functional characterization of these proteins. Notably, the 3.0-Å crystal structure of SAMT bound to SA and S-adenosyl-L-homocysteine (SAH) provides a template for deciphering the principles that underlie substrate recognition in SAMT and other members of this large family of plant methyltransferases.

RESULTS

Structural Elucidation by Protein X-Ray Crystallography

Recombinant SAMT from Clarkia was expressed in Escherichia coli as an N-terminal polyhistidine-tagged protein and purified by Ni²⁺ affinity and gel filtration chromatography. SAMT was crystallized from ammonium sulfate solutions in the presence of a 2-fold molar excess of SAH and a 1.5-fold molar excess of SA at 23°C. The structure of SAMT was determined with one native data set and two data sets obtained at two different wavelengths (absorption maximum and inflection point) for a Lu³⁺-derivatized crystal using a combination of multiple isomorphous replacement and multilength anomalous dispersion (MAD) phasing (Table 1). The final model was refined using the combined data sets used for phasing to 3.0-Å resolution. The model spanned the entire 359 residues of full-length

Table 1. Crystallographic Data, Phasing, and Refinement Statistics for Clarkia SAMT

<table>
<thead>
<tr>
<th>Variable</th>
<th>Native</th>
<th>Lutetium Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.6983</td>
<td>1.3407</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>99 to 3.2</td>
<td>99 to 2.9</td>
</tr>
<tr>
<td>Space group</td>
<td>P4₁,2,2</td>
<td>P4₁,2,2</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>12,433</td>
<td>27,353</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>8.3 (31.2)</td>
<td>6.6 (42.7)</td>
</tr>
<tr>
<td>No. of Lu sites</td>
<td>—</td>
<td>2</td>
</tr>
<tr>
<td>PPF (acentric/centric)</td>
<td>0.9/0.9</td>
<td>—</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>127</td>
<td>—</td>
</tr>
<tr>
<td>No. of SA molecules</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>No. of SA molecules</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Rcryst (3.0 Å)</td>
<td>23.6</td>
<td>—</td>
</tr>
<tr>
<td>Rfree (3.0 Å)</td>
<td>28.8</td>
<td>—</td>
</tr>
<tr>
<td>R.m.s.d. (Å)</td>
<td>0.012</td>
<td>—</td>
</tr>
<tr>
<td>R.m.s.d. angles (°)</td>
<td>1.8</td>
<td>—</td>
</tr>
</tbody>
</table>

Average B-factors

| Protein (Å²) | 71.2 |
| Water (nonprotein) (Å²) | 55.1 |
| SAH (Å²) | 57.7 |
| SA (Å²) | 86.7 |

Wavelength 1 (λ1) refers to data collected on a nonderivatized crystal of SAMT. Wavelengths 2 (λ2) and 3 (λ3) were collected on a second SAMT crystal derivatized with Lu³⁺ but isomorphous with native SAMT.

<table>
<thead>
<tr>
<th>Phase Determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
</tr>
<tr>
<td>No. of Lu sites</td>
</tr>
<tr>
<td>Rcryst (3.0 Å)</td>
</tr>
<tr>
<td>Rfree (3.0 Å)</td>
</tr>
<tr>
<td>R.m.s.d. (Å)</td>
</tr>
<tr>
<td>R.m.s.d. angles (°)</td>
</tr>
</tbody>
</table>

Refinement statistics

- Numbers in parentheses refer to the highest resolution shell.
- Rcryst = Σh(II - <H>)/Σh(II), where <H> is the average intensity over symmetry equivalent reflections.
- Rfree was calculated using 5% of data excluded from refinement.
- R.m.s.d. was calculated using 5% of data excluded from refinement.
- R.m.s.d. angles (°) is the average intensity over symmetry equivalent reflections.
Clarkia SAMT and included 2 Lu\(^{3+}\) ions, 12 water molecules, 1 SAH, and 1 SA molecule.

SAMT is a 41-kD protein of 359 residues that uses SAM as the methyl source for the transmethylation reaction, yielding SAH and the methyl ester of SA, MSA (Figure 1A). The overall structure of the SAMT monomer consists of a globular domain containing the extended \(\beta\)-sheet characteristic of all other SAM-dependent methyltransferases (Cheng and Blumenthal, 1999) and a unique \(\alpha\)-helical cap that forms the top one-third of the active site cavity. SAMT exists as a homodimer in solution, and this dimeric arrangement is preserved in the crystal lattice, with each monomer related by a crystallographic twofold axis (Figure 2A). Although it retains the core \(\alpha/\beta\)-fold typical of SAM-dependent methyltransferases (Figure 2B), the dimerization interface of SAMT is unique among structurally characterized plant methyltransferases.

Unlike previously reported plant O-methyltransferases, such as chalcone O-methyltransferase (ChOMT), isoflavone O-methyltransferase (IOMT) (Zubieta et al., 2001), and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase (COMT) (Zubieta et al., 2002), the dyad-related monomer of SAMT does not contribute to the active site of its partner molecule (Figure 2B). The dimerization interface also is much less extensive than that in ChOMT and its structural homologs, comprising only 1877 Å\(^2\) and burying \(\sim7\%\) of the available surface area. By contrast, other small-molecule plant methyltransferases characterized to date bury \(\sim30\%\) of the total surface area of the dimer at their interface. In SAMT, the dimer interface is composed predominantly of hydrophobic amino acid side chains, encompassing residues 100 to 115 and 119 to 128 donated from the two identical monomers, A and B. Hydrogen bonds form on the periphery of the interface between Arg-133(A) and Glu-91(B) and between Gln-5(A) and Asn-115(B), potentially stabilizing the hydrophobic core of the complementary inter-subunit surface (data not shown).

**SAM/SAH and SA Binding Pockets**

Hydrogen-bonding interactions secure the SAH molecule in the active site of SAMT. In the crystal structure, Asp-98 forms hydrogen bonds with both of the ribose hydroxyls of SAH. SAH forms a hydrogen bond through its carboxyl tail with the side chain amide of Asn-65 and a water-mediated link through its amino tail with Asp-57. Additionally, the exocyclic amino group of the adenine ring of SAH hydrogen bonds to the hydroxyl group of Ser-129 (Figures 3A and 3B). These interactions constitute a conserved hydrogen-bonding motif characteristic of SAM-dependent methyltransferases (Cheng and Blumenthal, 1999). Van der Waals interactions also contribute to the stabilization of the SAH molecule within the binding cavity. The adenine ring of SAH is involved in a pi-stacking arrangement with Phe-130, whereas Leu-99 brackets the adenine ring on the opposite side, providing a favorable hydrophobic clamp for the aromatic portion of SAH (Figure 3B).

The thioether moieties of Met-150 and Met-308 create a second molecular clamp that encompasses both faces of the benzyl ring of SA (Figures 3A and 3B). This Met-rich active site is similar to previously characterized plant methyltransferases, which also must position aromatic rich substrates near the reactive methyl group of SAM. Moreover, the reactive carboxylate moiety of SA is secured by hydrogen-bonding interactions that tether the carboxylate tail of SA to both the indole nitrogen of Trp-151 and the amide nitrogen of Gln-25. In this structure, there are no amino acid residues or water molecules within hy-
hydrogen-bonding distance of the 2-hydroxyl group of SA. However, the 2-hydroxyl moiety of SA forms an intramolecular hydrogen bond with its own carboxylate, thus constraining the carboxyl group to remain coplanar with the aromatic ring of SA. The rest of the active site provides a complementary hydrophobic pocket that sterically restrains the SA molecule in a favorable orientation for catalytic transmethylation to occur (Figure 3B).

**Putative Reaction Mechanism Accompanying Methyl Group Transfer**

Within the transmethylation pocket of the SAMT active site, there are no residues that would likely participate in the preparation of the SA carboxyl group for the methyl transfer reaction using general acid/base catalysis. By contrast, general acid/base catalysis is used in ChOMT, IOMT (Zubieta et al., 2001), and COMT (Zubieta et al., 2002) during the formation of the reactive phenolate anions before SAM-mediated methyl transfer. Instead, SAMT relies on the proper positioning of the SA substrate’s ionized and desolvated carboxyl group near the reactive methyl group of SAM to facilitate methyl transfer (Figures 2B, 3A, and 3B). The use of proximity and orientation effects as well as the need for desolvation before methyl transfer have been summarized by Takusagawa et al. (1998) for other SAM-dependent methyltransferases. At neutral pH, SA is found predominantly as a solvated and negatively charged carboxylate-bearing anion before SAM-mediated methyl transfer. Instead, SAMT relies on the proper positioning of the SA substrate’s ionized and desolvated carboxyl group near the reactive methyl group of SAM to facilitate methyl transfer (Figures 2B, 3A, and 3B). The use of proximity and orientation effects as well as the need for desolvation before methyl transfer have been summarized by Takusagawa et al. (1998) for other SAM-dependent methyltransferases. At neutral pH, SA is found predominantly as a solvated and negatively charged carboxylate-bearing anion that can readily undergo SAM-dependent transmethylation in an energetically favorable direction as long as the carboxylate moiety is desolvated before its sequestration near the reactive methyl group of SAM. With its positively charged sulfur, SAM easily donates its methyl group to the SA anion to form the more stable, uncharged compounds SAH and MSA.

The hydrogen bonds formed between the carboxylate moiety of SA and Trp-151 and Gln-25 ensure that the position of the carboxylate acceptor is within 3 Å of the expected location of the methyl group of SAM, whereas the overall active site provides the binding energy necessary to facilitate the desolvation of the reactive oxyanion of SA (Figure 3B). These two residues, which are involved in the formation of critical hydrogen-bonding interactions with the carboxyl-bearing SA substrate, are highly, although not absolutely, conserved within the family of SAMT-like plant methyltransferases (Figure 4). As discussed below, in cases in which Gln is replaced by other hydrogen bond donors, structural models predict that the change is functionally relevant, because it provides the necessary geometric constraints to ensure the selective binding of chemically different methyl acceptors such as 7-methylxanthine and theobromine.

The other characterized members of the SAMT/JMT family use either carboxylate-containing small molecules, such as SA and JA (Figures 1A and 1B), or nitrogen-containing heterocycles, such as 7-methylxanthine and theobromine (Figures 1C and 1D), as substrates. In all cases, the reactive groups are expected to be fully or predominantly deprotonated at cellular pH values; therefore, these groups are poised chemically and structurally for transmethylation to occur without enzyme-mediated general acid/base catalysis. In the case of 7-methylxanthine and theobromine, the methyl-accepting nitrogen resides next to a keto group, which affords the formation of an enol tautomer, facilitating solvent-assisted deprotonation of the reactive nitrogen.

**Structurally Guided Modulation of SAMT Substrate Preferences**

The crystal structure of SAMT serves as a model for other carboxylate-specific and related methyltransferases and provides a useful starting point for the functional analysis of residues thought to be important for substrate recognition and catalysis. To test both the significance of the SAMT-SA complex and the utility of the SAMT structure for understanding substrate diver-
sification in the SAMT family of plant methyltransferases, a close relative of SAMT, JMT from Arabidopsis, was selected for modeling to provide a starting point for mutagenic analysis of Clarkia SAMT. JMT converts the nonvolatile substrate JA to the volatile product MJA (Figure 1B).

Sequence alignments of SAMT, JMT, and other related methyltransferases predict the conservation of the SAH/SAM binding residues, whereas residues observed to be in contact with SA in Clarkia SAMT vary widely among members of this methyltransferase family (Figure 4). With respect to Arabidopsis JMT, Clarkia SAMT differs at six active-site residues at positions 147 (Tyr to Ser), 150 (Met to His), 225 (Ile to Gln), 308 (Met to Ile), 347 (Phe to Tyr), and 349 (Asn to Ile). Five of these positions (excluding the semiconservative substitution of Ile for Met at position 308) were chosen for the first set of functional tests (Figure 5A). The specific activities of five mutants differing at one or more of these positions (Figure 5) were determined using 12 different carboxyl-bearing organic acids, including SA and JA (Figures 5A to 5C). Although each of the mutants possessed attenuated specific activities against a partial subset of the examined substrates, the mutations resulted in altered substrate specificity preferences, thus functionally linking the structure of the active site to the selection of substrates in the SAMT family of methyltransferases (Figure 5C).

The presence of Tyr-147 in SAMT occludes much of the active site near the SAM/SAH binding site, making the binding of JA or any substrate larger than SA sterically unfavorable. Sequence comparisons of putative SAMT-like proteins showed a high conservation of the SSYS motif corresponding to residues 145 to 148 of SAMT from Clarkia. JMT from Arabidopsis presents an SSSS motif at these analogous positions. The replacement of the conserved Tyr at position 147 of Clarkia SAMT by a much smaller Ser side chain in JMT appears to provide the extra space necessary to accommodate the alkenyl portion of JA in the modeled JMT active site (Figure 5A). Hypothesizing that the identity of the amino acid at this position (residue 147 in Clarkia SAMT) plays a critical role in the selection of substrates, we made the Y147S mutation in SAMT. Wild-type Clarkia SAMT was assayed using concentrations of JA up to 5 mM, with no measurable methylation of JA detected.
**Figure 5.** Structurally Guided Functional Analysis of Clarkia SAMT.

(A) Schemes comparing the SAMT active site (left) and the putative active site of JMT (right). Residues are numbered according to the Clarkia SAMT sequence shown in Figure 4. SA and JA are shown as black bonds, whereas side chains are depicted as gray bonds with atoms color coded according to type. Hydrogen bonds are shown as dashed blue lines. Residue substitutions are connected by arrows colored by mutation. Green indicates Y147S, orange indicates N349I, blue indicates F347Y, red indicates I225Q, and yellow indicates M150H.

(B) Twelve carboxyl-containing small molecules used in assays against wild-type and mutant Clarkia SAMTs. The structures are arranged from small (acetic acid/acetate) to large (caffeic acid/caffeate). The two demonstrated physiological substrates for the SAMT-like family, SA and JA, are colored brown. The 3-hydroxy group of 3-hydroxybenzoic acid and the keto group on the five-membered ring of JA are colored red to emphasize their structural relationship when viewed in the context of the Y147S-M150H-I225Q-F347Y quadruple mutant that potentially hydrogen bonds with each of these moieties through the Gln side chain replacement at position 225. The blue arrows highlight small-molecule substrates for which one or more of the SAMT mutants exhibits 25% or greater activity.

(C) Activities of wild-type (wt) Clarkia SAMT and five mutant SAMTs against 12 carboxyl-containing small molecules. Relative activity is expressed as a percentage of the maximal specific activity measured against a particular substrate. Values are derived from specific activities measured in triplicate. The maximal specific activity for a particular mutant/substrate combination is shown in brackets and expressed as picomoles of product per second per milligram of enzyme. Mutant and substrate combinations displaying 25% or greater relative activity are highlighted in tan blocks.
(Figure 5C). The combination of the Y147S and M150H mutations led to a significant increase in the ability of the double mutant to turn over JA while preserving substantial SA-methylating activity (Figure 5C). An additional mutation of Phe-347 to Tyr, resulting in the triple mutant Y147S-M150H-F347Y, did not significantly alter the properties of the mutant enzyme compared with the double mutant, but when combined separately with mutations at two additional positions, 225 or 349, it resulted in enzymes capable of methylating additional carboxyl-bearing substrates.

Notably, the quadruple mutant Y147S-M150H-I225Q-F347Y demonstrated a striking switch in substrate specificity. In particular, the substitution of Ile-225 by the hydrogen bond donor/acceptor Gln resulted in a quadruple mutant displaying the greatest specific activities against 3-hydroxybenzoic acid and JA (Figure 5C). The structure of Clarkia SAMT suggests that the presence of a Gln at position 225 is ideally situated to form substrate-specifying hydrogen bonds with the 3-hydroxyl group or the keto oxygen of 3-hydroxybenzoic acid or JA, respectively (Figures 5A and 5B).

Collectively, the mutations described above and in Figure 5 generally result in more promiscuous activity rather than in clear specificity switches, implying that substrate recognition and catalytic efficiency are more subtle biochemical characteristics of this family of methyltransferases. Indeed, these two biochemical properties likely correlate with the identities of residues that contact the substrate directly and outer tier residues that only indirectly affect the architecture and catalytic properties of the substrate binding surface. Although the current set of mutations did not completely switch substrate specificity, the amino acid replacements examined to date support the active-site model derived from the crystal structure of Clarkia SAMT presented here. Additional changes in residues that either line the SA/JA binding pocket directly or modulate the shape and chemical reactivity of the active-site surface through second-tier interactions may be necessary to quantitatively and completely switch the specificity of SAMT to an authentic JMT. In addition, point mutations and other combinations of mutations will be necessary to more quantitatively assess the contribution of particular active-site surfaces to substrate recognition and catalytic efficiency. These experiments are in progress.

**Modeling and Biochemical Characterization of an SAMT-Like Enzyme Capable of Methylating Indole-3-Acetic Acid**

We examined the role of additional residues in determining substrate specificity in the SAMT-like methyltransferase family by modeling a particularly interesting Arabidopsis SAMT-like sequence encoded by the gene *At5g55250* (Figure 4). The Arabidopsis genome contains 23 such genes in addition to JMT. A multifaceted and long-range project aimed at elucidating the substrate specificity of all of these gene products was launched recently (D’Auria et al., 2003). *At5g55250* encodes a protein that contains an active site divergent from that of SAMT or JMT (Figure 4). A cDNA of *At5g55250* was expressed in *E. coli* and purified using conventional column chromatography. In vitro kinetic analysis of the protein product of *At5g55250* with several potential carboxyl- and nitrogen-containing substrates demonstrated that the encoded protein preferentially methylated the carboxyl group of the phytohormone indole-3-acetic acid (IAA). Steady state kinetic analysis of Clarkia SAMT (Figure 6A) and *At5g55250* (Figure 6B) against SA and IAA, respectively, demonstrated that each enzyme possessed comparable kinetic properties. SAMT possesses a $K_m$ value of 23.0 $\mu$M and a $k_{cat}$ of 0.092 s$^{-1}$ ($k_{cat}/K_m = 4000$ s$^{-1}$M$^{-1}$) using SA as a substrate, whereas *At5g55250* possesses a $K_m$ value of 13.0 $\mu$M and a $k_{cat}$ of 0.028 s$^{-1}$ ($k_{cat}/K_m = 2200$ s$^{-1}$M$^{-1}$) for IAA. Given *At5g55250*’s kinetic efficiency against IAA, it was named IAA carboxyl methyltransferase (IAMT).

Notably, the active-site model of IAMT serves as a useful starting point from which to develop a hypothesis concerning the in vitro specificity of IAMT. The most striking change in the active-site cavity of IAMT relative to SAMT occurs at position 226, where Trp-226 of Clarkia SAMT, which resides along the back wall of the active site (Figure 6C), is replaced by a Gly residue in Arabidopsis IAMT (Figure 6D). The loss of the Trp residue at this active-site location creates a large and spacious pocket for the recognition and binding of the indole ring of IAA (Figure 6D). This facile mode of adapting the methyltransferase scaffold for the biochemical modification of particular metabolites suggests that alterations in the substrate specificity of enzymes comprising specialized metabolic pathways can occur rapidly, because single amino acid changes can give rise to distinct substrate preferences. The biological role and physiological significance of this newly discovered methyltransferase involved in chemically modifying an important plant growth regulator, auxin (IAA), are under investigation.

**Structural Hypothesis Underpinning N-Methylation in Caffeine Biosynthesis**

Although the overall chemical structures and sites of methylation of the substrates of 7-methylxanthine N-methyltransferase (MXMT) and 3,7-dimethylxanthine N-methyltransferase (DMXMT [TCS1]) are distinct from those of SAMT, JMT, and IAMT (Figures 7A to 7C), the experimentally determined Clarkia SAMT structure (Figure 7A) nonetheless serves as an effective template for modeling the active sites of the N-methyltransferases involved in caffeine biosynthesis (Figures 7D and 7E). The positioning of the 7-methylxanthine and 3,7-dimethylxanthine (theobromine) molecules in the modeled active sites of MXMT and DMXMT (TCS1), respectively, was accomplished by preserving the hydrogen-bonding pattern that brackets the reactive carboxyl group in SAMT, JMT, and IAMT (Figures 7A to 7C). These two hydrogen bonds can be modeled in MXMT and DMXMT such that the nitrogen atom undergoing methylation is positioned next to the methyl-bearing sulfur on SAM. In SAMT, JMT, and IAMT, conserved Trp and Gln residues provide the hydrogen-bonding arrangement necessary to accurately position the methyl-accepting carboxyl group of the substrate near the reactive methyl group residing on SAM.

In the case of both *Coffea arabica* 7-MXMT and *Camellia sinensis* 3,7-DMXMT (TCS1), the Trp residue is conserved and the Gln residue is substituted by an Asn and a Thr residue, respectively. The replacement of this Gln residue is notable in that the
shortened side chain length of Asn and Thr in 7-MXMT and 3,7-
DMXMT, respectively, may be a key element for providing the
added volume necessary to firmly sequester the nitrogen-bear-
ing substrates near SAM. Hypothetically, 7-MXMT could use
the side chain amide group of Asn to provide a hydrogen bond
to the five-membered ring nitrogen of 7-methylxanthine (Figure
7D). In the case of 3,7-DMXMT (TCS1), a Thr at the analogous
position hypothetically fulfills this role by providing a hydrogen
bond to the second carbonyl oxygen that brackets the nitrogen
methyl acceptor on the six-membered ring of 3,7-dimethylxan-
thine (Figure 7E).

**DISCUSSION**

A significant issue that remains to be addressed in greater de-
tail and more globally across a large number of SAMT-like en-
zymes in plants concerns the role of primary and secondary in-
teractions in modulating the three-dimensional architecture and
reactivity of the substrate binding cavities in this and other
methyltransferase families. These additional levels of architect-
ural plasticity may play important roles in the mechanistic di-
versity exhibited by this widespread family of plant methyl-
transferases, because each enzyme chemically tailors its core
catalytic scaffold to accommodate and methylate numerous
specialized metabolites in plant cells. An experimental ap-
proach using a combination of atomic resolution structure de-
terminations, homology modeling, site-directed mutagenesis,
and kinetic analysis provides a formidable arsenal of methodol-
gies for elucidating and manipulating the enzymological diver-
sity in secondary metabolism.

Several families of methyltransferases involved in specialized
metabolism in plants have been described (for review, see
Ibrahim et al., 1998). The three-dimensional structures of IOMT
and ChOMT (Zubieta et al., 2001) and COMT (Zubieta et al., 2002),

![Figure 6. Kinetic and Structural Characterization of a New SAMT-Like Enzyme from Arabidopsis.](image)
all of which belong to the same family of plant methyltransferases, have been elucidated and described previously. The structural information generated from these enzymes has proven useful in understanding the function and evolution of additional methyltransferases from this structurally and functionally distinct family of plant small-molecule methyltransferases. For example, Gang et al. (2002) used the IOMT structure to model the active site of eugenol $O$-methyltransferase (EOMT) and chavicol $O$-methyltransferase (CVOMT) from basil. This study showed that the active sites of both enzymes are very similar to that of IOMT, even though the overall sequence identity between IOMT and either EOMT or CVOMT is only 65%. The modeling further demonstrated that EOMT and CVOMT, although 10% divergent in sequence identity overall, differ by a single amino acid in the modeled active sites. A change of this amino acid in EOMT to the residue found in CVOMT fully converted the substrate specificity and kinetic properties of the mutant enzyme to that of CVOMT, and an identical result was obtained in the reciprocal experiment:

CVOMT was converted to EOMT by a single amino acid substitution of the equivalent position in the active site.

The family of methyltransferases to which SAMT belongs was identified only recently (D’Auria et al., 2003). Members of this methyltransferase family appear to share no primary sequence similarity outside of the core SAM/SAH binding motif with other families of plant methyltransferases, such as the IOMT-like group or the group responsible for the methylation of CoA-derivatized phenylpropanoids, including caffeoyl-CoA $O$-methyltransferase. Furthermore, this newly discovered methyltransferase family includes those enzymes that catalyze the transfer of a methyl group to a nitrogen rather than to a carboxyl group oxygen (Figures 1C and 1D). The elucidation of the three-dimensional structure of SAMT, the founding member of this family, has made it possible to accurately depict and experimentally examine the active site surface that is in contact with the carboxyl-bearing substrate and the methyl donor SAM, as demonstrated in structural and biochemical comparative analyses of SAMT and JMT. This level of three-dimensional de-
tall provides an understanding of the catalytic mechanism that underlies substrate recognition and methyl group transfer.

The approach of structurally guided functional characterization was demonstrated through the in vitro mutagenesis and biochemical analysis performed on mutants of Clarkia SAMT designed on the basis of sequence comparison with Arabidopsis JMT. Additional structure/function studies conducted on many more SAMT-like family members will include the elucidation of the contribution of specific residues to the recognition of particular specialized metabolites in plants and should provide a reliable means of predicting the possible substrate specificities of newly identified but functionally uncharacterized members of this growing methyltransferase family.

Moreover, the structural information gleaned from Clarkia SAMT also has proven useful in deciphering the molecular interactions that accompany the recognition of the plant hormone IAA by a newly characterized protein encoded by the Arabidopsis gene At5g55250. Given the number of other important plant growth regulators that contain carboxyl- or nitrogen-containing functional groups, including gibberellins, cytokinins, and asaccic acid, it is reasonable to postulate that one or more of the currently uncharacterized SAMT-like accessions in Arabidopsis may, in a manner similar to IAMT, modify both the chemical scaffold and physiochemical properties of other vital plant hormones.

Because of the large number of functionally unannotated gene products that bear a high degree of similarity to Clarkia SAMT in sequence databases, the identification of active-site residues using structural models and mutagenesis will greatly facilitate the determination of possible substrates for these proteins before biochemical characterization. Currently, the only methyltransferases of this family whose characterized substrates have been reported, in addition to Clarkia SAMT, are the SAMTs from Stephanotis floribunda (Pott et al., 2002), Atropa belladonna (Fukami et al., 2002), and Antirrhinum majus (snapdragon) (Negre et al., 2002), BAMT from A. majus (Dudareva et al., 2000), JMT from Arabidopsis (Seo et al., 2001), and the N-methyltransferases theobromine synthase (7-MXMT) from C. arabica and caffeine synthase (both a 7-MXMT and theobromine XMT) from C. sinensis (Kato et al., 2000; Ogawa et al., 2001) (Figure 4).

Database searches using the Clarkia SAMT and Arabidopsis JMT sequences identified numerous putative gene products, including 23 members in Arabidopsis alone, few of which have been examined experimentally. Some of the SAMT-like and JMT-like proteins found in these searches revealed active-site motifs characteristic of SA and JA binding and in one case IAA binding (Figure 4). Nevertheless, when viewed with reference to the Clarkia SAMT structure, most of these sequences revealed highly divergent active-site surfaces suggestive of the recognition of many novel and specialized metabolites in plants. Additional structural elucidations, computer modeling studies, automated small-molecule docking procedures, mutagenesis experiments, and biochemical screening assays will enable the rapid identification of the likely in vivo substrates of these putative methyltransferases.

With the vastly accelerated pace of plant genomics and the availability of fully sequenced genomes, including those of Arabidopsis and rice, the functional characterization of gene products is becoming a critical obstacle for the accurate and complete annotations of genomes in a timely manner. The ability to identify or limit the choice of the most likely substrates and catalytic mechanisms of newly discovered enzymes before in-depth biochemical characterization will facilitate the understanding of complex biosynthetic pathways in plants and other organisms. Structural studies of protein families, such as the SAMT family of carboxyl-specific methyltransferases, are well suited to this approach. Structures of canonical member proteins provide an important tool in the often laborious process of protein characterization by suggesting probable reactivities and substrates of putative gene products. In addition, the identification of the substrate binding domains and the amino acid residues responsible for these specificities in enzymes will contribute to the direct modulation of biosynthetic pathways in plants via the introduction of rationally designed mutations into the enzymes that constitute these metabolic pathways. MSA, MJA, and other methyl esters are important components of floral scent, and they also contribute to the distinct flavors of plants. Many flavoring and aroma compounds are produced commercially at relatively high costs. Thus, the optimization and diversification of industrially significant biosynthetic pathways in plants are of great commercial interest and should benefit from the critical insight into substrate specificity gained from the elucidation of the three-dimensional structures of enzymes of specialized metabolism.

**METHODS**

**Materials**

Nickel-nitrilotriacetic acid agarose (Ni²⁺-NTA) resin was purchased from Qiagen (Valencia, CA). Benzamidine-Sepharose and Superdex-200 fast protein liquid chromatography columns were obtained from Amersham Biosciences (Piscataway, NJ). Lutetium acetate, salicylic acid (SA), jasmonic acid (JA), S-adenosyl-L-Met (SAM), and S-adenosyl-L-homocysteine (SAH) were obtained from Sigma (St. Louis, MO). Cryoloops were purchased from Hampton Research (Laguna Niguel, CA).

**Expression, Purification, and Mutagenesis**

Clarkia breweri SA carboxyl methyltransferase (SAMT) was cloned into the *Escherichia coli* expression vector pET28a(+). The SAMT construct was transformed into *E. coli* BL21(DE3). Transformed *E. coli* was grown at 37°C in terrific broth (Tartof and Hobbs, 1987) containing 50 µg/mL kanamycin until an A₅₀₀ value of 1.0 was obtained. After induction with 0.5 mM isopropyl 1-thio-β-D-galactopyranoside, the cultures were grown for 8 h at 25°C. Cells were pelleted, harvested, and resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, pH 8.0, 20 mM β-mercaptoethanol, 10% [v/v] glycerol, and 1% [v/v] Tween 20) and stirred at 4°C for 1 h with lysozyme (0.5 mg/mL).

After sonication and centrifugation, the supernatant was passed over a Ni²⁺-NTA column, washed with 10 bed volumes of lysis buffer and 10 bed volumes of wash buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 20 mM imidazole, pH 8.0, 20 mM β-mercaptoethanol, 10% [v/v] glycerol), and the His-tagged protein was eluted with elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole, pH 8.0, 20 mM β-mercaptoethanol, and 10% [v/v] glycerol). Incubation with thrombin during dialysis for 24 h at 4°C against 25 mM Hepes, pH 7.5, 100 mM NaCl,
and 1 mM DTT removed the N-terminal His tag. Dialyzed protein was re-
loaded onto a Ni²⁺-NTA column to remove the cleaved His tag followed by
thrombin depletion using a 0.5-mL benzamidine-Sepharose column.

Gel filtration performed using a Superdex-75 column equilibrated with
25 mM Hepes, pH 7.5, 100 mM NaCl, and 1 mM DTT resulted in protein
of >99% purity based on Coomassie Brilliant Blue staining of pooled
fractions by SDS-PAGE analysis. Fractions containing the protein of in-
terest were pooled and concentrated to ~8 to 10 mg/mL and stored at
~80°C. All SAMT mutants were generated using the QuikChange (Strat-
agene, San Diego, CA) PCR-based method. Arabidopsis thaliana (Co-
lumbia strain) At5g55250 cDNA was obtained by reverse transcriptase-
mediated PCR using mRNA from floral tissue (D’Auria et al., 2003), li-
gated into the expression vector pCRT7/CT-TOPO (Invitrogen, Carls-
bad, CA), transformed into E. coli BL21(DE3), and expressed and puri-
ified under conditions substantially similar to those described above for
the Clarkia SAMT construct.

Enzyme Activity Assays

Wild-type Clarkia SAMT and all SAMT mutants were purified by Ni²⁺ af-
finity chromatography as described above, dialyzed against 50 mM Tris-
HCl, pH 7.5, 100 mM KCl, and 5 mM β-mercaptoethanol, and concen-
trated to ~1 to 8 mg/mL. Potassium salts were used in place of NaCl
because of the greater activity exhibited by native SAMT from Clarkia
(Ross et al., 1999) and jasmonic acid carboxyl methyltransferase from
Arabidopsis (Seo et al., 2001) in the presence of potassium ions. Specific
activities against 12 organic acids were measured using a standard ra-
diochemical assay mixture consisting of a 10-μL enzyme solution (2.5 to
13 μM final concentration), 1.0 μL of 50 mM substrate solution in etha-
nol, 10 μL of 5× assay buffer (250 mM Tris-HCl, pH 7.5, 500 mM KCl, and
14 mM β-mercaptoethanol), and 10 μL of a 1:1 mixture of 10 mM cold
SAM and 0.385 mM 14C-SAM (0.1 to 0.2 Ci per reaction) in a total vol-
ume of 50 μL. The reactions were performed at 25°C and stopped after
30 min by extraction with 200 μL of ethyl acetate. After extraction, 20 to
100 μL of the ethyl acetate phase was placed in a scintillation vial con-
taining 2 mL of Econo-safe scintillation fluid (RPI, Mount Prospect, IL)
and counted in a Beckmann LS 6500 scintillation counter. All activities
were normalized to account for enzyme concentration, and extract vol-
ume was counted.

Steady state kinetic constants for wild-type Clarkia SAMT and wild-
type Arabidopsis indole-3-acetic acid carboxyl methyltransferase were
determined by fitting initial velocity versus SA or indole-3-acetic acid
concentrations to the hyperbolic Michaelis-Menten equation using
GraphPad Prism version 4.0a for Macintosh (GraphPad Software, San
Diego, CA) and normalized to account for enzyme concentration, and extract
volume. Kinetic parameters were determined by curve fitting using
GraphPad Prism. Reactions were incubated at 23°C, and velocity was deter-
mined by thin-layer chromatography and quantitated as described previously (Ross et al., 1999).

Protein X-Ray Crystallography

Crystals of wild-type Clarkia SAMT were grown by vapor diffusion in
hanging drops consisting of 1:1 mixtures of concentrated enzyme (see
above) containing 1 mM SAH and crystallization buffer [2.0 to 2.2 M
(NH₄)₂SO₄, 0.5 mM SA, and 0.05 M 3-(N-morpholino)-propanesulfonic acid–NaCl; pH 7.5]. Crystals formed by incubating crystallization trials for
3 days at 15°C and for 1 to 3 days at 23°C. Using these crystals as a
source of microseeds, subsequent crystallizations were performed at
23°C after the introduction of diluted crystalline microseeds into the con-
centrated enzyme/crystallization mixture described above. Crystals
grew in space group P4₂½,2 with one molecule per asymmetric unit. The
correct space group enantiomorph (P4₂½,2 or P4₂,2,2) was determined by
MAD phasing as described below. Unit cell dimensions were a,b =
141.74 Å, c = 63.983 Å, and α,β,γ = 90°.

Lutetium acetate soaks were performed on stabilized crystals [2.2 M
(NH₄)₂SO₄, 0.05 M 3-(N-morpholino)-propanesulfonic acid, pH 7.5, and
10% (v/v) glycerol]. Lutetium acetate was added to the stabilization
buffer in three steps reaching final concentrations of 1, 2, and 5 mM.
Crystals were equilibrated at each concentration for ~1 h. The 5 mM
soaks were allowed to proceed overnight, after which time the crystals
were harvested and flash-frozen in liquid N₂. Diffraction data were col-
llected at 105 K from single crystals mounted in a cryoloop.

All diffraction data were collected at the European Synchrotron Radi-
ation Facility in Grenoble, France, beamline FIP, on a Quantum 4 charge-
coupled device detector. All images were indexed and scaled using au-
tomatic indexing, refinement, and scaling based on DENZO and SCALEPACK (Otwinowski and Minor, 1997). Three wavelengths for the
absorption maxima, the inflection point, and a remote wavelength were
collected on Lu³⁺-derivatized crystals. As a result of radiation damage, the remote wavelength was not usable. In subsequent pseudo-MAD phasing experiments, a native data set was substituted for the remote
wavelength. Initial heavy-atom sites were found with CNS (Bruenger et al., 1998). SHARP (de la Fourtelle and Bricogne, 1997) was used to refine ini-
tial sites, and phases were improved with SOLOMON (Abrahams and
Leslie, 1996).

The initial model was constructed using O (Jones et al., 1991), and ex-
perimental electron density maps were constructed using composite
structure-factor amplitudes calculated in SHARP from the native data
collected at 1.698 Å and the two lutetium acetate data sets collected at
1.3407 and 1.3411 Å (Table 1). Refinements were performed with RE-
MAC and CNS (Bruenger et al., 1998) using the combined structure-factor
amplitudes determined in SHARP. The quality of all models was deter-
mained with the program PROCHECK (Laskowski et al., 1996). Sixty-
seven percent of the residues were in the most favored region of the
Ramachandran plot, with 26% in the allowed, 6% in the generously
allowed, and 1% in the disallowed regions (Tyr-18, Ser-61, Asn-115, and
Thr-260). Of the latter residues, Tyr-18 and Ser-61 appeared to be mod-
ified correctly because the electron density for each was unambiguous
and 1% in the disallowed regions (Tyr-18, Ser-61, Asn-115, and
Thr-260). Of the latter residues, Tyr-18 and Ser-61 appeared to be mod-
nicated in this study to specifically methylate IAA.

ACKNOWLEDGMENTS

We thank the staff of the European Synchrotron Radiation Facility, beamline FIP (BM30A), especially Jean-Luc Ferrer and members of the
Structural Biology Laboratory, for technical assistance. We also thank
John C. D’Auria and Feng Chen for the initial cloning of the IAMT cDNA.
This work was supported by funds awarded to J.P.N. by the Salk Insti-
tute for Biological Studies. Work conducted by J.R.R., Y.Y., and E.P. was

Accession Numbers

The accession numbers for the sequences shown in Figure 4 are as
follows: SAMT from Clarkia, AF133053.1; JMT from Arabidopsis,
AA233434; MXMT from C. arabica, BAB39216; and TCS1 from C. sinen-
sis, BAB12278. At5g55250 encodes an Arabidopsis protein character-
ized in this study to specifically methylate IAA.
Plant Carboxyl Methyltransferases 1715


Received June 10, 2003; accepted June 12, 2003.

REFERENCES


Structural Basis for Substrate Recognition in the Salicylic Acid Carboxyl Methyltransferase Family
Chloe Zubieta, Jeannine R. Ross, Paul Koscheski, Yue Yang, Eran Pichersky and Joseph P. Noel
Plant Cell 2003;15:1704-1716; originally published online July 14, 2003;
DOI 10.1105/tpc.014548

This information is current as of October 28, 2017

References
This article cites 40 articles, 15 of which can be accessed free at:
/content/15/8/1704.full.html#ref-list-1

Permissions

eTOCs
Sign up for eTOCs at:
http://www.plantcell.org/cgi/alerts/ctmain

CiteTrack Alerts
Sign up for CiteTrack Alerts at:
http://www.plantcell.org/cgi/alerts/ctmain

Subscription Information
Subscription Information for The Plant Cell and Plant Physiology is available at:
http://www.aspbo.org/publications/subscriptions.cfm

© American Society of Plant Biologists
ADVANCING THE SCIENCE OF PLANT BIOLOGY