Structural Biology

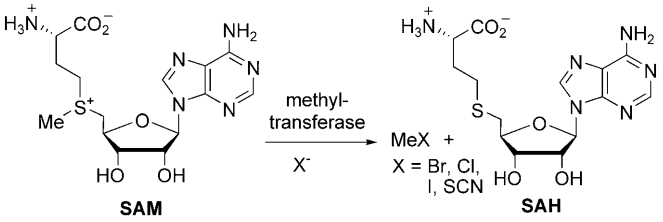
DOI: 10.1002/anie.201000119

Halomethane Biosynthesis: Structure of a SAM-Dependent Halide Methyltransferase from *Arabidopsis thaliana***

*Jason W. Schmidberger, Agata B. James, Robert Edwards, James H. Naismith, and David O'Hagan**

Of the gaseous halomethanes (MeCl, MeBr, MeI) found in the atmosphere, chloromethane is the major contributor. About 4×10^6 tonnes of chloromethane are estimated to be produced globally per year, which is contributed to by terrestrial plants (and fungi) and by photosynthetic microorganisms, algae, and marine plants.[1] Chloromethane is the most significant naturally produced volatile chlorocarbon, contributing up to 15% of stratospheric chlorine.[2] Bromomethane is significant too, but less abundant, and is estimated to be produced at about 1.8×10^5 tonnes per year[1] and contributing up to 55% of stratospheric bromine.[2] Iodomethane is generated at a higher level than bromomethane at about 8×10^5 tonnes per year,[1] but it appears to be the least significant with respect to atmospheric chemistry because of photolysis, resulting in a low stability and a short half-life.[3] Higher plants are estimated to account for 30–50% of the global production of chloromethane,[4] with its biogenesis receiving considerable attention owing to its role in ozone depletion. The genes responsible for halomethane biosynthesis in plants have been named the HOL (harmless to ozone layer) genes,[5,6] because of inactivation of the associated biosynthetic pathway when they are deleted. The halomethane gases almost certainly play some regulatory role too within the producing organisms, as methyl transfer vehicles, although their metabolic role is unclear.[7–9]

In plants, fungi, and bacteria, the enzyme product of the HOL gene is known to be an S-adenosyl-L-methionine (SAM)-dependant methyltransferase, which combines halide ion and SAM in a nucleophilic substitution reaction to generate halomethane.[10] Although this enzyme is responsible for halomethane production (Scheme 1), it was recently shown to have a preference for thiocyanate as a nucleophile.[11] The promiscuity shown by the enzyme for the halides (excluding fluoride) and thiocyanate renders its physiological role unclear, although its particular ability to methylate thiocyanate implies an intracellular role in glucosinolate metabolism.[12]



Scheme 1. The halomethane gases are generated in a nucleophilic reaction between halide ion and SAM to generate S-adenosyl-L-homocysteine (SAH). The reaction is catalyzed by the action of halide methyltransferases.

Over the last decade, there has been a substantial development in our understanding of biohalogenation and particularly enzymatic chlorination, and an intriguing range of different halogenation enzymes have been identified that are responsible for generating the C–Cl bond in secondary metabolites.[13–15] The earliest characterized are the haloperoxidases, found typically in marine plants, which oxidize chloride, bromide, and iodide ions at the expense of hydrogen peroxide to generate X^+ for reactions with aromatics and other unsaturated organics. These haloperoxidases fall into two categories, and are either vanadium- or heme-iron-dependent.[16] More recently, FAD-dependent chlorination enzymes have been characterized, which also mediate electrophilic reactions such as the chlorination of tryptophan.[17,18] These speocies are typically found in microbes rather than plants. Walsh has recently identified a class of iron–sulfur halogenation enzymes responsible for the selective chlorination of unactivated carbon atoms (for example, methyl groups).[19,20] These enzymes essentially generate chlorine radicals, and they are particularly relevant to the aliphatic chlorination of bacterial secondary metabolites in both terrestrial and marine environments. Nucleophilic chlorination beyond chloromethane production is rare, although Moore reported in 2008[21] the identification of the chlorinase, an enzyme responsible for the production of 5-chloro-5′-deoxyadenosine (ClDA), the first-formed intermediate in the biosynthesis of salinosporamide-A, a metabolite of the deep-sea soil bacterium *Salinospora tropic*a. This enzyme is chemically similar and almost structurally identical to the fluorinase, an enzyme from *Streptomyces cattleya* that utilizes both fluoride and to a lesser extent chloride ion in a similar nucleophilic attack to C5′ carbon of SAM.[22–24] X-ray structural data is in place for the nucleophilic fluorinase[23] and chlorinase[21] enzymes, as well as representatives of almost all of the other known halogenases.[15] Therefore, the

[*] Dr. J. W. Schmidberger, Prof. Dr. J. H. Naismith, Prof. Dr. D. O'Hagan
University of St Andrews
Department of Chemistry, Centre for Biomolecular Sciences
North Haugh, St Andrews, Fife, KY16 9ST (UK)
Fax: (+44) 1334 463-808
E-mail: do1@st-andrews.ac.uk
A. B. James, Prof. Dr. R. Edwards
University of Durham, School of Biological & Medical Sciences
South Road, Durham, DH1 3LE (UK)

[**] We thank the BBSRC for a grant (BB/F007426/1) and J.W.S. thanks the Wellcome Trust for the Value in People (VIP) Scheme for financial support. SAM = S-adenosyl-L-methionine.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201000119.

plant halo/thiocyanate methyltransferases remain the last major group of halogenase enzymes that have not been structurally characterized. Given the importance of gaseous halomethane production to atmospheric chemistry,[1–3] and the developing recognition of the role of these enzymes in thiocyanate metabolism in plants,[11] we present herein the first structural study on this enzyme group.

Plant methyltransferases that are active towards halide/thiocyanate nucleophiles have been isolated and studied from *Batis maritima*,[5] *Brassica oleracea*,[25] and *Arabidopsis thaliana*.[5,6] Collectively, they have been termed halide methyltransferases (HMT), or halide/thiocyanate methyltransferases (HTMT), on the basis of their activity with halides alone, or additionally with thiol substrates, such as bisulfide or thiocyanate.[11] A phylogenetic analysis using the *A. thaliana* structural gene (*AtHOL1*) suggests a wide distribution of these enzymes amongst the plant kingdom, with two further family members encoded within the *A. thaliana* genome (*AtHOL2* and *AtHOL3*).[12]

Herein the coding sequence of *AtHOL1* (gene At2g43910; accession AY044314), now termed AtHTMT, was isolated by PCR[26] and the gene cloned and protein over-expressed in Rosetta II (DE3) cells with a C-terminal 6-His tag for ease of purification. For assay work (to remove the influence of the His tag), the *AtHTMT1* gene was cloned into a pEHISTEV vector[27] to give an N-terminal poly His tag with a TEV protease cleavage site. The two mutants (V23C and Y172F) were also constructed and cloned into the same vector. Purification of the over-expressed proteins then included a TEV digestion. The purified enzyme was concentrated (10 mg mL^{-1}), and was co-crystallized with S-adenosyl-L-homocysteine (SAH). The structure was solved to a resolution of 1.8 Å using the related (20% amino acid sequence identity) mouse thiopurine methyltransferase (PDBId: 2GB4) structure as a molecular replacement model. The low identity necessitated considerable model manipulation. The SAH ligand locates the active site of the enzyme, as illustrated in Figure 1, and the active site is shown more closely in Figure 2. The methyl group was modeled at the sulfonium stereogenic center to represent SAM (Figure 2b). It is well-established that SAM synthase only generates the (*S*)-SAM configuration at sulfur, and all enzymes appear to utilize this isomer of SAM.[28] The trajectory of the modeled methyl group projects into the active site. Furthermore, the Trp47 residue would prevent the diastereomeric (*R*)-SAM from binding, as there is an obvious clash with the methyl group. The modeled methyl group shown in Figure 2b occupies a rather open active site, which is consistent with the promiscuity of the enzyme, particularly for large nucleophiles.

Three crystallographically identifiable water molecules occupy the cavity in the SAH–enzyme co-crystal. One water molecule (W198) will be displaced by the methyl group of SAM and the central water molecule (W68) occupies the predicted location of the halide nucleophile. The remaining water molecule (W35) is hydrogen bonded to the side chain of Tyr172. Therefore, a model emerges in which this water moelcule is bridging, which helps to orient the nucleophile (bromide or chloride ion; Figure 2b). Site-directed mutagenesis (Tyr172Phe) of the tyrosine residue to phenylalanine led to a functional enzyme, but with a reduced efficiency with chloride ions (V_{max} drops from 2.43 to 0.92 nmol min^{-1} mg^{-1} protein), but otherwise a similar efficiency for bromide and thiocyanate (Table 1). This analysis suggests that the efficiency of the larger nucleophiles, bromide and thiocyanate, is not particularly compromised, and perhaps ordered hydrogen bonding, from tyrosine 172, through a bridging water molecule is important for orientating the smaller chloride nucleophile.

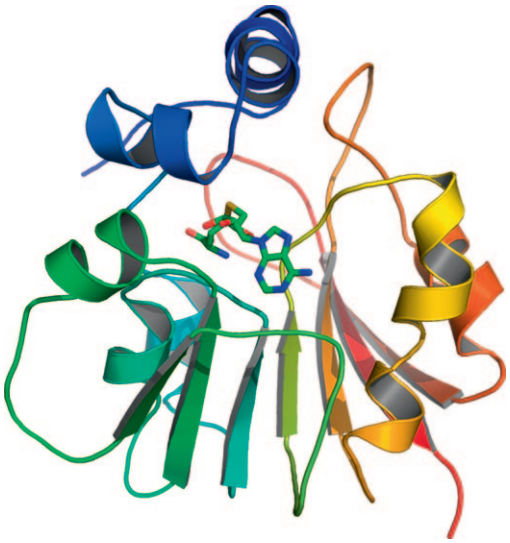


Figure 1. Stylized ribbon diagram of the structure of the *Aribidopsis thaliana* halomethyl transferase (AtHTMT1), solved crystallographically to a resolution of 1.8 Å and colored in a spectrum from N-terminus (blue) to C-terminus (red). S-Adenosyl-L-homocysteine (SAH) is bound to the putative active site of the enzyme.

The amino acid sequence homology is generally high between HMT/HTMT proteins, except for the first 30 or so residues of the N-terminus (blue helices in Figure 1). This region creates a cap over the active site, forming key contacts to the nucleophile during the reaction. When the residues lining the active site of AtHTMT1 are compared (through a sequence alignment) to those of the HTMT from *Batis maritima* (BmHTMT),[8] which features a much greater activity for Cl^-, Val-23 (of AtHTMT1) is the only active site residue that is not conserved. In BmHTMT, the equivalent residue is a cysteine. Accordingly this valine residue of AtHTMT1 was mutated to cysteine. Despite removing a hydrophobic active site residue in close contact with the nucleophile, the resultant V23C mutant remained functional, displaying a slightly improved activity for all of the nucleophiles explored, including chloride ions (Table 1). However, a lowered stability of this mutant was noticeable by decomposition on SDS-PAGE.

In conclusion, the structure of a plant halomethane-producing enzyme is presented and a model for substrate/nucelophile binding and reaction at the active site rationalized. The *Arabidopsis thaliana* enzyme presents the reactive sulfonium methyl group into a large cavity that accounts for its promiscuous nature with respect to a variety of nucleophiles. The enzyme promotes reaction most efficiently with

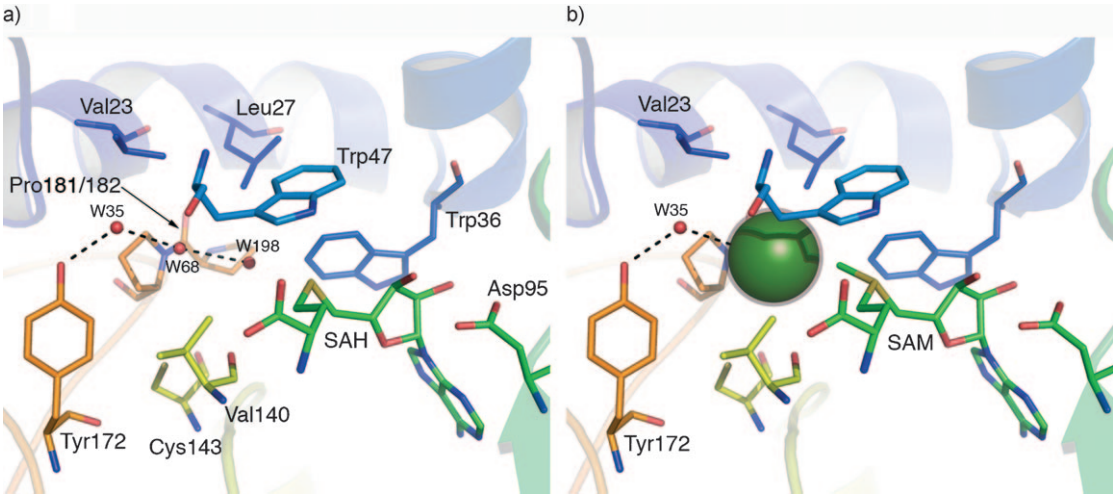


Figure 2. a) The *Aribidopsis thaliana* halomethyl transferase active site showing the amino acid side groups forming the putative nucleophile-binding site (shown). Three water molecules (W35, W68, and W198) occupy a channel from the bulk solvent to a proposed nucleophilic binding site occupied by W198. b) The missing methyl group of (*S*)-SAM is modeled into the structure with space-filling chloride (green sphere, 1.75 Å radius) and bromide ions (larger pale red sphere, eclipsed by green sphere, 1.85 Å) to simulate preorganization for reaction.

Table 1: Enzyme kinetic data for AtHTMT variants for the three nucleophiles NCS^-, Br^-, and Cl^-.

AtHTMT variant	Nucleophile	K_m [mM]	V_{max} [nmol min^{-1} mg protein^{-1}]
Native	NCS^-	0.099 ± 0.020	43.6 ± 2.52
V23C	NCS^-	0.103 ± 0.009	43.4 ± 0.93
Y172F	NCS^-	0.141 ± 0.009	46.0 ± 1.11
Native	Br^-	24.87 ± 2.785	11.4 ± 0.31
V23C	Br^-	21.06 ± 1.883	12.4 ± 0.25
Y172F	Br^-	30.16 ± 2.942	11.4 ± 0.33
Native	Cl^-	145.2 ± 26.56	2.43 ± 0.12
V23C	Cl^-	122.0 ± 25.31	2.78 ± 0.16
Y172F	Cl^-	137.2 ± 20.60	0.91 ± 0.04

thiocyanate and then the halides, with an efficiency order of $NCS^- > I^- > Br^- > Cl^-$ (but not F^-). The dual role of this enzyme class is intriguing and there remains much to learn regarding the relationship between halomethane biosynthesis and thiocyanate/methyl thiocyanate metabolism within a single organism.

Received: January 8, 2010
Revised: February 22, 2010
Published online: April 7, 2010

Keywords: atmospheric organohalogens · biohalogenation · chloromethane · methyltransferases · structural biology

[1] D. B. Harper, *Nat. Prod. Rep.* **2000**, *17*, 337 – 348.
[2] M. L. Cox, P. J. Fraser, G. A. Sturrock, S. T. Siems, L. W. Porter, *Atmos. Environ.* **2004**, *38*, 3839 – 3852.
[3] M. L. Cox, G. A. Sturrock, P. J. Fraser, S. T. Siems, P. B. Krummell, *J. Atmos. Chem.* **2005**, *50*, 59 – 77.
[4] T. Saito, Y. Yokouchi, Y. Yosugi, M. Tani, E. Philip, T. Okuda, *Geophys. Res. Lett.* **2008**, *35*, L19812.
[5] Y. Nagatoshi, T. Nakamura, *Plant Biotechnol.* **2007**, *24*, 503 – 506.
[6] R. C. Rhew, L. Ostergaard, E. S. Saltzman, M. F. Yanofsky, *Curr. Biol.* **2003**, *13*, 1809 – 1813.
[7] X. H. Ni, L. P. Hager, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12866 – 12871.
[8] X. H. Ni, L. P. Hager, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3611 – 3615.
[9] R. C. Rhew, B. R. Miller, R. F. Weiss, *Nature* **2000**, *403*, 292 – 295.
[10] T. S. Bayer, D. M. Widmaier, K. Temme, E. A. Mirsky, D. V. Santi, C. A. Voigt, *J. Am. Chem. Soc.* **2009**, *131*, 6508 – 6515.
[11] N. Itoh, H. Toda, M. Matsuda, T. Negishi, T. Taniguchi, N. Ohsawa, *BMC Plant Biol.* **2009**, *9*, 116.
[12] Y. Nagatoshi, T. Nakamura, *J. Biol. Chem.* **2009**, *284*, 19301 – 19309.
[13] C. D. Murphy, *Nat. Prod. Rep.* **2006**, *23*, 147 – 152.
[14] C. S. Neumann, D. C. Fujimori, C. T. Walsh, *Chem. Biol.* **2008**, *15*, 99 – 109.
[15] L. C. Blasiak, C. L. Drennan, *Acc. Chem. Res.* **2009**, *42*, 147 – 155.
[16] J. Littlechild, *Curr. Opin. Chem. Biol.* **1999**, *3*, 28 – 34.
[17] C. Dong, S. Flecks, S. Unversucht, C. Haupt, K. H. van Pee, J. H. Naismith, *Science* **2005**, *309*, 2216 – 2219.
[18] E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan, C. T. Walsh, *Biochemistry* **2007**, *46*, 1284 – 1292.
[19] F. H. Vaillancourt, J. Yin, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10111 – 10116.
[20] D. P. Galonic, E. W. Barr, C. T. Walsh, J. M. Bollinger, C. Krebs, *Nat. Chem. Biol.* **2007**, *3*, 113 – 116.
[21] A. Eustáquio, F. Pojer, J. P. Noel, B. S. Moore, *Nat. Chem. Biol.* **2008**, *4*, 69 – 74.
[22] D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. G. Murphy, C. D. Murphy, *Nature* **2002**, *416*, 279.
[23] C. Dong, F. L. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan, J. H. Naismith, *Nature* **2004**, *427*, 561 – 565.
[24] H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. D. O'Hagan, D. A. Robinson, J. B. Spencer, *Angew. Chem.* **2006**, *118*, 773 – 776; *Angew. Chem. Int. Ed.* **2006**, *45*, 759 – 762.
[25] J. M. Attieh, A. D. Hanson. H. S. Saini, *J. Biol. Chem.* **1995**, *270*, 9250 – 9257.
[26] D. P. Dixon, B. G. Davis, R. Edwards, *J. Biol. Chem.* **2002**, *277*, 30859 – 30869.
[27] H. Liu, J. H. Naismith, *Protein Expression Purif.* **2009**, *63*, 102 – 111.
[28] J. Komoto, T. Yamada, Y. Takata, G. D. Markham, F. Takusagawa, *Biochemistry* **2004**, *43*, 1821 – 1831.

