Published in final edited form as:

Angew Chem Int Ed Engl. 2010 May 10; 49(21): 3646–3648. doi:10.1002/anie.201000119.

Halomethane production in plants: Structure of the biosynthetic SAM-dependent halide methyltransferase from *Arabidopsis thaliana***

Jason W. Schmidberger,

University of St Andrews, Department of Chemistry, Centre for Biomolecular Sciences, North Haugh, St Andrews, Fife, KY16 9ST, UK

Agata B. James, Robert Edwards,

University of Durham, School of Biological & Medical Sciences, South Road, Durham, DH1 3LE, UK Fax: (+44) 191 3341201 robert.edward@durham.ac.uk

James H. Naismith, and

University of St Andrews, Department of Chemistry, Centre for Biomolecular Sciences, North Haugh, St Andrews, Fife, KY16 9ST, UK jhn@st-andrews.ac.uk

David O'Hagan*

University of St Andrews, Department of Chemistry, Centre for Biomolecular Sciences, North Haugh, St Andrews, Fife, KY16 9ST, UK

Abstract

A product structure of the halomethane producing enzyme in plants (*Arabidopsis thaliana*) is reported and a model for presentation of chloride/bromide ion to the methyl group of S-adenosyl-L-methionine (SAM) is presented to rationalise nucleophilic halide attack for halomethane production, gaseous natural products that are produced globally.

Keywords

chloromethane; biohalogenation; methyltransferase; structural biology; S-adenosyl-L-methionine

Chloromethane is the major atmospheric gaseous halomethane (MeCl, MeBr, MeI) estimated to be produced in 4×10^6 tonnes per year by terrestrial plants (and fungi) as well as by photosynthetic microorganisms, algae and sea weeds in the oceans.¹ It is the most significant naturally produced volatile chlorocarbon contributing upto 15% of stratospheric chlorine.² Bromomethane is significant too, but less abundant, estimated to be produced at $\sim 1.8 \times 10^5$ tonnes per year¹ and contributing upto 55% of stratospheric bromine². Iodomethane is generated at about 8×10^5 tonnes per year¹ but it appears to be the least significant with respect to atmospheric chemistry due to photolysis resulting in a low stability and a short halflife.³ Higher plants are estimated to account for 30-50% of the global production of chloromethane⁴, with its biogenesis receiving considerable attention due 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} due to inactivation of

[**]We thank the BBSRC for a grant (BB/F007426/1) and JWS thanks the Wellcome Trust for the Value in People (VIP) Scheme for financial support.

[†]Fax: (+44) 1334 463808 do1@st-andrews.ac.uk.

Supporting Information for this article is available on the WWW under <http://www.angewandte.org> or from the authors.

the associated biogenic 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

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, in plants, fungi and bacteria.¹⁰ Although this enzyme is responsible for halomethane production as illustrated in Figure 1, it was recently shown to have a preference for thiocyanate, as a nucleophile.¹¹ The promiscuity shown by the enzyme for the range of 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

Over the last decade there has been a substantial development in our understanding of biohalogenation and particularly enzymatic chlorination, where a range of different halogenation enzymes have been identified, responsible for generating the C-Cl bond in secondary metabolites.¹³⁻¹⁵ The earliest characterised are the haloperoxidases, found typically in marine plants, which oxidise chloride, bromide and iodide ions at the expense of hydrogen peroxide, to generate X^+ for reactions with aromatics and other unsaturated organics. These fall into two categories and are either vanadium- or haem-Fe dependent haloperoxidases.¹⁶ More recently FAD-dependent chlorination enzymes have been characterised, which also mediate electrophilic reactions such as the chlorination of tryptophan.^{17,18} These are typically found in microbes rather than plants. Walsh has more recently identified a class of iron-sulfur halogenation enzymes responsible for the selective chlorination of unactivated carbon atoms (eg 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 the identification of the chlorinase, an enzyme responsible for the production of 5-chloro-5'-deoxyadenosine (CIDA), the first formed intermediate in the biosynthesis of salinosporamide-A, a metabolite of the deep sea soil bacterium *Salinospora tropica*.²¹ This enzyme is chemically similar and almost structurally identical to the fluorinase, an enzyme from *Streptomyces cattleya* that utilises both fluoride and to a lesser extent chloride ion in a similar nucleophilic attack to SAM.²²⁻²⁴ X-ray structural information is in place for the nucleophilic fluorinase and chlorinase enzymes, and structural data is available for representatives of all of the other known halogenases, however the plant halo/thiocyanate methyltransferases remain the last major group of halogenases without structural characterisation. 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 plants¹¹, we report in this *Communication* the first structural exemplar of this enzyme class.

Examples of the halo/thiocyanate methyltransferases plant enzymes 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 thiocyanate¹¹ a phylogenetic analysis using the *A. thaliana* structural gene (AtHOL1) suggests a wide distribution of these enzymes amongst the plant kingdom, with two other homologs within the *A. thaliana* genome itself (AtHOL2, and AtHOL3).¹²

In this study the coding sequence of AtHOL1 (gene At2g43910; accession AY044314) was isolated from the cDNA, of a mixture of *Arabidopsis* plants and cultures, by PCR

amplification.²⁶ The oligonucleotides HOL1F 5' GCGCGCCATATGGCTGAAGAACAACAAACTC 3' and HOL1R PCR 5' GCGCGCCTCGAGATTGATCTTCTTCCACCTTCCC 3' were used as the forward and reverse primers respectively containing the NdeI and XhoI restriction sites (underlined). Amplification was performed through 25 cycles using Phusion DNA polymerase (Finnzymes), with cycling parameters of 98°C for 15min, 60°C for 15min and 72°C for 30min. The NdeI and XhoI restriction sites then used for the direct cloning of the insert into the pET24 (Novagen) plasmid, with the addition of a C-terminal 6-His tag. The plasmid was transformed into Rosetta II (DE3) cells, over expressed O/N at 20°C, and was purified using Ni-Sepharose affinity beads and size exclusion chromatography. For assay work (to remove the influence of the His tag), the AtHOL1 gene was cloned into a pEHISTEV vector²⁷ to give an N-terminal poly His tag with a TEV protease cleavage site. Two mutants (V23C, and Y172F) were also constructed and cloned into the same pEHISTEV vector. Over-expression of the pEHISTEV constructs was the same as for the pET28a construct, though purification included a TEV digestion at 4°C for 3 h (with ~60 µg His-tagged TEV/mg of HTMT), immediately after the Ni beads. The mix was then applied to SE column, before being passed back over Ni beads to remove uncleaved material and any remaining TEV. Pure HTMT was concentrated to ~2 mg/ml and frozen (-80°C) for subsequent use. The purified enzyme was concentrated to 10 mg/ml, and was co-crystallised with S-adenosyl-L-homocysteine (SAH - 330 µM final), and a 1.8 Å structure was solved using the homologous mouse thiopurine methyltransferase (PDBId: 2GB4) structure as a molecular replacement model. The final model features SAH located in the active site of the enzyme as illustrated in Figure 2. Figure 3 illustrates the active site more closely. In Figure 3B The methyl group was modeled at the sulfonium stereogenic centre to represent SAM. It is well established that SAM synthase only generates the (S)-SAM stereogenicity at sulfur. This is universal for all enzymology involving SAM. With this stereochemistry 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 3B, occupies a rather open active site, consistent with the promiscuity of the enzyme, particularly for large nucleophiles.

Three crystallographically identifiable water molecules occupy the cavity in the SAH 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. Thus a model emerges where this is a bridging water which helps to orient the nucleophile (bromide or chloride ion) as illustrated in Figure 2B. Consistent with this, site directed mutagenesis (Tyr172Phe) of the tyrosine residue to phenylalanine led to a functional enzyme, however with a reduced affinity for chloride ion ($V_{\max}$ drops from 2.43 to 0.92 nmol min⁻¹ mg⁻¹ protein), but otherwise a similar affinity for bromide and isocyanate as illustrated in Table 1. This analysis suggests that the nucleophilic efficiency of the larger nucleophiles, bromide and thiocyanate, are not particularly compromised and perhaps ordered hydrogen bonding, from tyrosine 172, through a bridging water molecule is particularly important for orientating the smaller chloride nucleophile.

The amino acid sequence homology is generally high between these proteins except for the first 30 or so residues of the N-terminus (blue helices in Figure 2). 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), 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 the active site residue of AtHTMT1 was mutated to cysteine. Despite being an 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 Cl^- (Table 2), however the stability of this mutant to purification was noticeably reduced on SDS Page gels.

In conclusion the structure of a plant halomethane producing enzyme is presented and a model for nucleophile binding and reaction at the active site is rationalised. The *Arabidopsis thaliana* enzyme has a large hydrophobic cavity which accounts for its promiscuous nature with respect to a variety of nucleophiles. The enzyme reacts most efficiently with thiocyanate and then the halides with a efficiency order of $\text{NCS}^- > \text{I}^- > \text{Br}^- > \text{Cl}^-$ (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/methylthiocyanate metabolism within a single organism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- [1]. Harper DB. Nat. Prod. Rep. 2000; 17:337–348. [PubMed: 11014336]
- [2]. Cox ML, Fraser PJ, Sturrock GA, Siems ST, Porter LW. Atmospheric Environment. 2004; 38:3839–3852.
- [3]. Cox ML, Sturrock GA, Fraser PJ, Siems ST, Krummell PB. J. Atmos. Chem. 2005; 50:59–77.
- [4]. Saito T, Yokouchi Y, Yosugi Y, Tani M, Philip E, Okuda T. Geophys. Res. Lett. 2008; 35:L19812.
- [5]. Nagatoshi Y, Nakamura T. Plant Biotechnol. 2007; 24:503–506.
- [6]. Rhew RC, Ostergaard L, Saltzman ES, Yanofsky MF. Current Biol. 2003; 13:1809–1813.
- [7]. Ni XH, Hager LP. Proc. Natl. Acad. Sci. 1998; 95:12866–12871. [PubMed: 9789006]
- [8]. Ni XH, Hager LP. Proc. Natl. Acad. Sci. 1999; 96:3611–3615. [PubMed: 10097085]
- [9]. Rhew RC, Miller BR, Weiss RF. Nature. 2000; 403:292–295. [PubMed: 10659844]
- [10]. Bayer TS, Widmaier DM, Temme K, Mirsky EA, Santi DV, Voigt CA. J Am Chem Soc. 2009; 131:6508–6515. [PubMed: 19378995]
- [11]. Itoh N, Toda H, Matsuda M, Negishi T, Taniguchi T, Ohsawa N. BMC Plant Biology. 2009; 9:116. [PubMed: 19723322]
- [12]. Nagatoshi Y, Nakamura T. J. Biol. Chem. 2009; 284:19301–19309. [PubMed: 19419967]
- [13]. Murphy CD. Nat. Prod. Rep. 2006; 23:147–152. [PubMed: 16572225]
- [14]. Neumann CS, Fujimori DC, Walsh CT. Chem.Biol. 2008; 15:99–109. [PubMed: 18291314]
- [15]. Blasiak LC, Drennan CL. Acc. Chem Res. 2009; 42:147–155. [PubMed: 18774824]
- [16]. Littlechild J. Curr. Op., Chem. Biol. 1999; 3:28–34.
- [17]. Dong C, Flecks S, Unversucht S, Haupt C, van Pee KH, Naismith JH. Science. 2005; 309:2216–2219. [PubMed: 16195462]
- [18]. Yeh E, Blasiak LC, Koglin A, Drennan CL, Walsh CT. Biochem. 2007; 46:1284–1292. [PubMed: 17260957]
- [19]. Vaillancourt FH, Yin J, Walsh CT. Proc. Natl. Acad. Sci. USA. 2005; 102:10111–10116. [PubMed: 16002467]
- [20]. Galonic DP, Barr EW, Walsh CT, Bollinger JM, Krebs C. Nature. Chem. Biol. 2007; 3:113–116. [PubMed: 17220900]
- [21]. Eustáquio A, Pojer F, Noel JP, Moore BS. Nature Chem. Biol. 2008; 4:69–74. [PubMed: 18059261]
- [22]. O'Hagan D, Schaffrath C, Cobb SL, Murphy JTG, Murphy CD. Nature. 2002; 416:279.
- [23]. Dong C, Huang FL, Deng H, Schaffrath C, Spencer JB, O'Hagan D, Naismith JH. Nature. 2004; 427:561–565. [PubMed: 14765200]

- [24]. Deng H, Cobb SL, McEwan AR, McGlinchey RP, Naismith JH, O'Hagan DD, Robinson DA, Spencer JB. *Angew. Chem. Int. Ed. Engl.* 2006; 45:759–762. [PubMed: 16370017]
- [25]. Attieh JM, Hanson AD, Saini HS. *J. Biol. Chem.* 1995; 270:9250–9257. [PubMed: 7721844]
- [26]. Dixon DP, Davis BG, Edwards R. *J. Biol. Chem.* 2002; 277:30859–30869. [PubMed: 12077129]
- [27]. Liu H, Naismith JH. *Protein Expr Purif.* 2009; 63:102–111. [PubMed: 18845260]

Figure 1.

The halomethane gases are generated in a nucleophilic reaction between halide ion and SAM catalysed by the action of halide methyltransferases. These enzymes also utilise thiocyanate as a substrate.

Figure 2. Stylized ribbon diagram of the crystallographic structure of the *Arabidopsis thaliana* halomethyl transferase (AtHTMT1), solved to a resolution of 1.8 Å and coloured in a spectrum from N-terminus (blue) to C-terminus (red). S-Adenosyl-homocysteine (SAH) is bound to the putative active site of the enzyme.

Figure 3A.

The *Aribidopsis thaliana* halomethyl transferase active site showing the amino acid side groups forming the putative nucleophile-binding site shown. Three waters (W35, W68, and W198) were found to occupy a channel from the bulk solvent to a proposed nucleophilic-binding site occupied by W198.

Figure 3B.

The missing methyl group of (*S*)-SAM is modelled into the structure with chloride (green sphere) and bromide (larger pale red sphere) ions to simulate preorganisation for reaction. The halides are represented as “space-filled” spheres at 1.75 Å, and 1.85 Å radii for chloride and bromide ions respectively.

Table 1Enzyme 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