Structure and activity of DmmA, a marine haloalkane dehalogenase

Jennifer J. Gehret,1,2 Liangcai Gu,1,3 Todd W. Geders,1 William Clay Brown,1,4 Lena Gerwick,5 William H. Gerwick,5,6 David H. Sherman,1,3,7,8 and Janet L. Smith1,2*

1Life Sciences Institute, University of Michigan, Ann Arbor, MI
2Department of Biological Chemistry, University of Michigan, Ann Arbor, MI
3Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI
4Center for Structural Biology, University of Michigan, Ann Arbor, MI 48109
5Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093
6Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093
7Department of Chemistry, University of Michigan, Ann Arbor, MI
8Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

Received 29 September 2011; Revised 16 November 2011; Accepted 23 November 2011
DOI: 10.1002/pro.2009
Published online 28 November 2011 proteinscience.org

Abstract: DmmA is a haloalkane dehalogenase (HLD) identified and characterized from the metagenomic DNA of a marine microbial consortium. Dehalogenase activity was detected with 1,3-dibromopropane as substrate, with steady-state kinetic parameters typical of HLDs ($K_m = 0.24 \pm 0.05$ mM, $k_{cat} = 2.4 \pm 0.1$ s$^{-1}$). The 2.2-Å crystal structure of DmmA revealed a fold and active site similar to other HLDs, but with a substantially larger active site binding pocket, suggestive of an ability to act on bulky substrates. This enhanced cavity was shown to accept a range of linear and cyclic substrates, suggesting that DmmA will contribute to the expanding industrial applications of HLDs.

Keywords: haloalkane dehalogenase; marine microbial consortium; $\alpha/\beta$ hydrolase; CurN; DmmA

Introduction

Haloalkane dehalogenases (HLDs) remove halogens from alkanes by hydrolysis, producing an alcohol, a halide ion, and a proton. These dehalogenating enzymes are of long-standing interest for industrial applications such as degradation of environmental pollutants$^{1,2}$ and biocatalysis.$^3$ Newer applications include remediation of chemical weapons,$^4$ biosensing,$^5,6$ and cellular imaging.$^7$ The applications for HLDs are growing with knowledge of their functional and structural properties.

Each HLD has a unique substrate selectivity, which can include chlorinated, brominated, and iodinated alkanes of varying length. Some HLDs also act on ring-containing substrates, or those with alcohol, ether, nitrile, or alkene functional groups.$^8$ Substrate selectivity is not easily predicted by sequence analysis. For HLDs of known structure, the size and shape of the active-site cavity are better predictors of substrate preference than is phylogenetic analysis.$^8$ The natural substrates are, in general, unknown.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: NIH; Grant number: R01 DK42303, R01 CA108874, U01 TW007404; Grant sponsor: Hans W. Vahlteich Professorship.

Liangcai Gu’s current address is Department of Genetics, Harvard Medical School, 77 Ave Louis Pasteur, NR 232, Boston, MA 02115. Todd W. Geders’s current address is R&D Systems, Inc., Minneapolis, MN.

*Correspondence to: Janet L. Smith, Life Sciences Institute, University of Michigan, 210 Washtenaw Ave., Ann Arbor, MI 48109. E-mail: JanetSmith@umich.edu

Published by Wiley-Blackwell © 2011 The Protein Society
HLDs are members of the α/β hydrolase superfamily, with the active site located between a conserved core subdomain and a more variable lid structure. The HLD active site departs from the typical α/β hydrolase catalytic triad by the replacement of the common Ser nucleophile with an Asp and the addition of two halide-stabilizing residues (Trp and Trp/Asn) to form a "catalytic pentad." The side chains of the catalytic pentad are identically positioned in the active site cleft of all HLDs of known structure, although some pentad residues vary in their locations in the primary sequence. This sequence-location difference of the catalytic pentad residues defines three evolutionary subfamilies of HLDs. Two structures are known from evolutionary subfamily I (DhLA and DppA), and four from subfamily II (DhaA, LinB, DhjA, and Rv25782). Several structures include bound substrates, products, or covalent intermediates, which together with many kinetic studies, led to a detailed reaction mechanism for this enzyme class.

DmmA is a putative HLD of subfamily II. This protein was originally annotated as CurN, and presumed to be the final gene product of the curacin A biosynthetic gene cluster from the marine cyanobacterium Lyngbya majuscula (now designated Moorea producta). However, resequencing revealed the authentic 3' end of the gene cluster, which lacks curN, suggesting that dmmA was cloned as part of a chimeric cosmid from the original metagenomic DNA library. Herein, we present the crystal structure and HLD activity of DmmA.

Results

Biological source of DmmA (formerly CurN)

In the original curacin pathway sequencing, DNA encoding 64 amino acids at the CurN N-terminus overlapped with DNA encoding the CurM C-terminus in a different reading frame. This unusually long overlap and peculiarities of the CurM protein sequence led us to isolate a new cosmid, and to sequence the curM-curN region of the gene cluster, resulting in a corrected 3' terminus that lacked curN. In addition, no CurN coding sequence was identified from genomic DNA, nor was CurN located in the recently sequenced Moorea producta genome. Furthermore, the G + C content of curN (61%) contrasts with the rest of the curacin biosynthetic gene cluster (45.5%). The scaffolds used to assemble the Moorea producta genome ranged from 37 to 66% G + C content, with the great majority falling between 40 and 45%. In addition, CurN (now referred to as DmmA (dehalogenase A from a marine microbe) in accord with other dehalogenases of unknown biological function) has high similarity (∼50%) to HLDs from other marine bacteria, although none are from the cyanobacterial phyla. From these data, we conclude that curN (dmmA) is most likely a product of one of the organisms that grew in close association with the Moorea producta field isolate whose metagenomic DNA cosmid library was used to sequence the cur gene cluster. It is evident that the 3' end of the curacin gene cluster concatenated with curN (dmmA) genetic material from one of the heterotrophic bacteria that was growing on the Moorea producta filament at the time of cosmid preparation.

Because the biological source and the natural termini of DmmA are uncertain, we turned to sequence analysis to design dmmA constructs for protein expression. HLD homologs have C-termini of very similar length and sequence to the DmmA C-terminus, but the N-termini are poorly conserved and of variable length. In addition, DmmA amino acids 1-40 are predicted to be unstructured. Strong sequence conservation among subfamily-II HLDs begins at DmmA residue 48. This evidence could suggest that the N-terminus of DmmA is not in the original gene, however several considerations led us to consider the full open reading frame as the natural length of DmmA. First, the DNA sequence upstream of the putative translation start site contains sequence motifs and spacers similar to prokaryotic promoter regions and ribosome binding sequences (Supporting Information Fig. 1). In addition, several other annotated HLDs have N-termini of similar length to DmmA. Thus, we conclude that the natural protein likely includes all the residues encoded in the curN (dmmA) open reading frame. Two dmmA constructs were made, one encoding the full-length protein (DmmA_long, residues 1-341) and one with a 43-residue N-terminal truncation (DmmA_short, residues 44-341) to produce a potentially more structured N-terminus (Supporting Information Fig. 2). Both DmmA_short and DmmA_long were produced in high yield, as stable and soluble proteins in an E. coli expression system.

HLD activity

As for many other HLDs, the natural substrate for DmmA is unknown. Small di-halogenated alkanes are typical substrates to assess the activity of HLDs. We chose 1,3-dibromopropane, a standard substrate.

Table I. DmmA Activity Toward Halogenated Substrates

<table>
<thead>
<tr>
<th>Construct</th>
<th>Substrate</th>
<th>% Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmmA_short</td>
<td>1,3-dibromopropane</td>
<td>100</td>
</tr>
<tr>
<td>DmmA_short</td>
<td>1,6-dibromohexane</td>
<td>3.5 ± 0.7</td>
</tr>
<tr>
<td>DmmA_short</td>
<td>bromocyclohexane</td>
<td>14.8 ± 0.6</td>
</tr>
<tr>
<td>DmmA_short</td>
<td>1,6-dichlorohexane</td>
<td>3.7 ± 1.4</td>
</tr>
<tr>
<td>DmmA_short</td>
<td>1,6-diiodohexane</td>
<td>3.2 ± 0.7</td>
</tr>
<tr>
<td>DmmA_long</td>
<td>1,3-dibromopropane</td>
<td>10 ± 2</td>
</tr>
</tbody>
</table>

*Normalized to activity of DmmA_short with 1 mM 1,3-dibromopropane.
for DmmA assays. HLD activity was tested in a colorimetric assay to detect the production of protons.27,28 Both DmmA short and DmmA long were active in this assay, however DmmA short exhibited tenfold higher activity than the longer form of the protein (Table I), suggesting that the N-terminus may interfere with catalysis. Steady-state kinetic constants were determined for DmmA short (Table II), which revealed that these parameters are comparable to other HLDs.15,27,29 DhlA has threefold greater catalytic efficiency than DmmA for 1,3-dibromopropane (fourfold lower $k_{cat}$, 12-fold lower $K_m$),27 whereas DhaA has a nearly equal $k_{cat}$ and a 50-fold lower $K_m$ for 1,3-dibromopropane resulting in a 50-fold greater catalytic efficiency.29

**Structure of DmmA**

We solved crystal structures for both DmmA short and DmmA long (Table III). The first structure was determined from selenomethionyl DmmA short (2.2 Å). DmmA long (2.9 Å) was solved from the DmmA short structure. The short and long variants crystallized under different conditions in different crystal forms, each with two polypeptides per asymmetric unit. The two structures thus provide four independent views of the DmmA dehalogenase.

Due to the higher-resolution of DmmA short, our efforts focused on analysis of that form of the protein. DmmA adopts the $\alpha/\beta$ hydrolase fold as expected [Fig. 1(A)], consisting of core and lid sub-domains with the active site at the top of the core, which is covered by the lid [Fig. 1(A,B)]. DmmA short and DmmA long are identical within experimental error (root-mean-square deviation (RMSD) = 0.32 Å for 298 Cx atoms). The DmmA short structure is complete with no disordered residues. In DmmA long, six additional N-terminal residues are visible compared with DmmA short, but 37 residues at the N-terminus are disordered [Fig. 1(C)].

### Table II. Steady State Kinetic Parameters of DmmA short

<table>
<thead>
<tr>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}/K_m$ (M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4 ± 0.1</td>
<td>0.24 ± 0.05</td>
<td>1.0 ± 0.3 × 10$^7$</td>
</tr>
</tbody>
</table>

Gehret et al.

### Table III. Crystallographic Summary

<table>
<thead>
<tr>
<th>Diffraction Data</th>
<th>DmmA short (SeMet)</th>
<th>DmmA short (SeMet) 1,5-dibromopentane soaked</th>
<th>DmmA long</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>$P6_2$</td>
<td>$P6_2$</td>
<td>$P3_21$</td>
</tr>
<tr>
<td>X-ray source</td>
<td>APS 23-ID-D</td>
<td>APS 23-ID-D</td>
<td>APS 23-ID-D</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>99.8, 99.8, 122.0</td>
<td>100.3, 100.3, 121.4</td>
<td>144.7, 144.7, 105.0</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97939</td>
<td>0.92003</td>
<td>0.97939</td>
</tr>
<tr>
<td>$d_{min}$ (Å)</td>
<td>2.20 (2.28–2.2)</td>
<td>2.10 (2.18–2.10)</td>
<td>2.50 (2.50–2.64)</td>
</tr>
<tr>
<td>Avg $R_{cryst}$</td>
<td>13.2 (2.9)</td>
<td>16.7 (2.1)</td>
<td>10.4 (3.5)</td>
</tr>
<tr>
<td>$R_{sym}$</td>
<td>0.122 (0.549)</td>
<td>0.118 (0.75)</td>
<td>0.171 (0.607)</td>
</tr>
<tr>
<td>Completeness</td>
<td>100.0 (99.9)</td>
<td>99.7 (97.6)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Avg. redundancy</td>
<td>5.8 (5.2)</td>
<td>10.5 (7.6)</td>
<td>9.6 (8.6)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>35,082</td>
<td>40,262</td>
<td>44,165</td>
</tr>
</tbody>
</table>

**DmmA oligomeric state**

A common protein-protein contact exists in the unrelated crystal forms of DmmA long and DmmA short. The contact has twofold (non-crystallographic) symmetry,
suggesting that the protein may be dimeric [Fig. 1(D)]. The primary interaction occurs within the HLD core between α10 and β8, and is comprised of both hydrophobic and hydrophilic contacts and buries only four water molecules. The additional six ordered residues in DmmA_long (residues 38–43) associate with the partner subunit at this interface, resulting in a larger buried surface area in DmmA_long compared to DmmA_short (980 Å² vs. 610 Å²) [Fig. 1(D)]. The small size of the interface led us to question whether it exists in solution. Other HLDs have been characterized as monomers, dimers, or dimers only under low-salt conditions. Among HLDs of known structure, only DbjA is dimeric in solution. The helix α10 is central to the subunit contacts in both DmmA and DbjA, but the interfaces otherwise differ (Supporting Information Fig. 3).

The quaternary structure in solution was probed by analytical gel filtration chromatography. Both DmmA_short and DmmA_long eluted with apparent
molecular weights intermediate between calculated values for the monomer and the dimer (48 kDa observed vs. 34.6 kDa for a DmmA_short monomer, and 59 kDa vs. 38.8 kDa for DmmA_long) (Supporting Information Fig. 4). This result implies a dynamic equilibrium between monomeric and dimeric forms, however the retention volume did not change over a 20-fold concentration range of either DmmA_short or DmmA_long. In addition, the larger buried surface in the putative DmmA_long dimer compared to DmmA_short does not result in a more dimeric retention volume. Thus, we conclude that the interface observed in the crystal structures does not represent an interface in solution and that DmmA is monomeric. The

Figure 2. DmmA active site. A: DmmA active site showing catalytic pentad (sticks), Cl\(^-\) and malonate. B: Schematic of the DmmA active site showing the acyl-enzyme intermediate. The halide-stabilizing residues (Asn78 and Trp 145) bind the halide released when generating the acyl enzyme intermediate. The acyl enzyme on Asp144 will subsequently be hydrolyzed by a water activated by His315 C. Active site of subfamily II haloalkane dehalogenases, based on superposition of core Cx atoms. The catalytic pentad is shown in sticks and bound halide ion as a sphere for DmmA (green C), LinB 1IZ8\(^{19}\) (orange C), Rv2578 2OJF\(^{22}\) (yellow C), DbjA 3A2M\(^{21}\) (purple C), and DhaA 1CQW\(^{15}\) (mauve C). Bound malonate is shown for DmmA and bound product, 1,3-propanediol, for Rv2578 (yellow C) and LinB (orange C).
aberrant elution volumes may result from the dynamic behavior of a monomeric DmmA.

**DmmA active site**

The active site has an intact and well ordered catalytic pentad, consisting of the nucleophile (Asp144), base (His315), acid (Glu168), and two halide-stabilizing residues (Trp145 and Asn78) [Fig. 2(A,B)]. The nucleophile (Asp144), base (His315), and Trp145 are conserved in all HLDs. Asn78 is conserved in subfamilies II and III, and Glu168 is present only in subfamily II. The catalytic pentad resides in an otherwise hydrophobic chamber that is accessed through an entrance tunnel. This feature and the location of the pentad are well conserved in structures of all subfamily II HLDs [Fig. 2(C)]. Density resembling malonate, a component of the crystallization solution, was found in the active site cleft [Fig. 2(A), Supporting Information Fig. 5]. A product alcohol is bound at this position in structures of several other HLDs [Fig. 2(C)]. Difference density for a stronger scatterer than water (peak height 7.7 σ) appeared in the putative halide-binding site within hydrogen bonding distance of Trp145 and Asn78, the halide-stabilizing residues [Fig. 2(A), Supporting Information Fig. 5]. A halide occupies this position in the structures of several other HLDs [Fig. 2(C)]. The halide-binding site was confirmed by Br anomalous scattering from a crystal of DmmA soaked in 1,5-dibromopentane prior to data collection at the energy of the bromine K absorption edge. Br−
binding resulted in an anomalous difference peak of height 9 σ; no density was observed for the hydroxyl product (Supporting Information Fig. 6).

**Comparison to other HLDs**

The DmmA structure is similar to structures of other subfamily I HLDs (0.5–0.7 Å RMSD for 219–226 Cα atoms of proteins that are 38–48% identical overall to DmmA). The largest difference between the most similar structure, DhaA, and DmmA is in lid helix α4 and the core-to-lid linker from β6 to α4 [Fig. 3(A)]. As expected, DmmA is less similar to the subfamily I HLDs, DhlA (1.4 Å RMSD for 203 Cα) and DppA (1.9 Å for 209 Cα), which have lower overall sequence identity to DmmA.

Structure-function studies of other HLDs have demonstrated that the size of the active site cavity and entry tunnel play a role in substrate specificity and chiral selectivity. DmmA has a larger entry tunnel and wider substrate cleft than other HLDs [Fig. 3(B)]. This difference is due primarily to the position of lid helix α4 and the preceding loop, and to a lesser extent to amino acid substitutions within the cavity. The loop and helix α4 are the parts of the DmmA structure that differ most from other HLDs. Other HLDs with large active sites have very broad substrate specificity and react well with larger substrates. DmmA, with an even larger active site, should accommodate substrates of greater size and shape diversity. To test this, we assayed DmmA for dehalogenase activity with several halogenated molecules larger than the 1,3-dibromopropane standard substrate (1,6-dibromohexane, 1,6-dichlorohexane, and bromocyclohexane) (Table I). This suggests that the natural DmmA substrate may bear a halogenated ring system. The large active site also suggests the value of DmmA as a potential chemoenzymatic tool for large substrates.

The ultimate biological source of DmmA is unknown, as attempts to amplify dmmA from *M. producta* (formerly *L. majuscula*) DNA or to find dmmA in the genome sequence have been unsuccessful. *dmmA* (originally annotated as *curN*) was originally found in the cosmid library developed from an *M. producta* field isolate, which also contained bacterial symbionts or associants as a microbial consortium. To our knowledge, no other HLD has been found in a cyanobacterium, which also suggests an origin outside *M. producta*, in a symbiont or associated bacterium. Presumably, the natural function of DmmA is to dehalogenate molecules present in the native marine environment of *M. producta*. In this respect, *M. producta* has been a rich source of halogenated metabolites, including in this strain, barbamidine, which possesses a trichloromethyl group, and in another strain, jamai-camidine, which possesses vinyl chloride and alkynyl bromide groups. Despite its orphan status, the large active-site cavity poises DmmA for development as a new biotechnology tool.

**Materials and Methods**

**Cloning, protein expression, and purification**

Gene fragments encoding DmmA were generated from cosmid pLM17, inserted into the pET-24b vector, and verified by sequencing to form pDmmA<sub>long</sub> (encoding residues 1-341) and pDmmA<sub>short</sub> (encoding residues 44-341). *E. coli* strain BL21(DE3) was transformed with the expression plasmid, grown at 37°C in 1 L 2xYT to an OD<sub>600</sub> of 0.5, cooled to 18°C, induced with IPTG (final concentration 0.4 mM) and grown for an additional 18 h. Selenomethionyl (SeMet) DmmA<sub>short</sub> was produced in the same strain in SeMet minimal media.

All purification steps were performed at 4°C. The cell pellet from 1 L of cell culture was resuspended in 40-mL Buffer A (20 mM Tris pH 7.9, 500 mM NaCl, 10% glycerol) plus 20-mM imidazole,
lysed by sonication, and the soluble fraction loaded onto a 5-mL HisTrap NiNTA column (GE Healthcare). DmmA was eluted with Buffer A with a linear gradient of 20–650 mM imidazole. DmmA was further purified by size exclusion chromatography with a HiPrep 16/60 Sephacryl S100 HR column (GE Healthcare) prequillibrated with Buffer A. DmmA was concentrated to 14 mg/mL, flash frozen in liquid N2, and stored at approx. –80°C. SeMet DmmA_short was purified as the wild type with addition of 2 mM DTT to Buffer A during the size exclusion step. Yields per 1 L culture were 150 mg for DmmA_short, 80 mg for DmmA_long, and 15 mg for SeMet DmmA_short.

**Crystallization**

Crystals of DmmA_short (native and SeMet) grew by vapor diffusion at 20°C within 24 h from a 1:1 mix of protein stock (6 mg/mL DmmA_short in Buffer A) and well solution (2.2 M sodium malonate pH 7.0, 5% glycerol). Crystals of DmmA_long formed at 20°C and well solution (2.2 M sodium malonate pH 7.0, 5% PEG3350, and 0.1 M BisTris pH 5.5). Crystals of DmmA_long were cryoprotected by a 5–10 s transfer to a mixture of well solution also containing 25% PEG400.

**Data collection and structure determination**

Data were collected at the GM/CA-CAT beamline 23ID-D at the advanced photon source (APS) at Argonne National Lab (Argonne, IL). A 2.2 Å single-wavelength anomalous diffraction dataset was collected at the wavelength of peak absorption at the selenium edge from a DmmA_short, SeMet crystal. A 2.9 Å dataset was collected from a DmmA_long crystal. All data were processed using the HKL2000 suite. Determination of selenium atomic positions, experimental phasing, density modification phase refinement, and initial model building were performed using the programs SOLVE and RESOLVE. And 22 of the 24 expected selenium positions were identified. The model was finished manually in COOT using DmmA_short as a search model. Model quality was evaluated with MolProbity. Figures were made using PyMOL and PyMOL 2.4. Electron density was complete throughout the polypeptide chain for both polypeptides for DmmA_short (residues 44–341) and DmmA_long (residues 38–341). No density was observed for the C-terminal His tag of either DmmA variant or for the first 37 residues of DmmA_long.

**Enzyme assay**

Activity was measured using a pH indicator dye-based colorimetric method. Substrate [0.1–6.0 mM 1,3-dibromopropane (Aldrich)] was prepared in an indicator solution (20 µg/mL phenol red, 20 mM Na2SO4, 1 mM EDTA, 1 mM Hepes pH 8.2). Reactions were initiated by addition of DmmA_short (1 µM final concentration, 100–µL final volume). Absorbance at 550 nm was monitored at 5-second intervals during the reaction to detect the decrease in pH generated by the release of H+. Parallel negative controls lacking the enzyme or substrate were used to correct for nonenzymatic dehalogenation or pH change. A standard curve relating absorbance at 550 nm to [H+] was used to convert absorbance to [H+]. The initial rates were fit using KaleidaGraph to the Michaelis-Menten equation to obtain steady state kinetic constants. Comparison with other substrates (1,6-dibromohexane (Acros), 1,6-dichlorohexane (Aldrich), 1,6-diodohexane (Aldrich), bromocyclohexane (Aldrich)) were performed at 1 mM substrate concentration in duplicate and normalized to the percent of 1,3-dibromopropane activity.

**PDB Coordinates**

Coordinate and structure factors for DmmA_short have been deposited in the Protein Data Bank with accession number 3U1T.

**Acknowledgments**

Beamline 23ID-D is supported by the National Institutes of Health, National Institute of General Medical Sciences Grant Y1-GM-1104 and National Cancer Institute Grant Y1-CO-1020 through the GM/CA Collaborative Access Team at the APS, which is supported by the United States Department of Energy.

**References**


