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. Smith[1,2]*

[1]Life Sciences Institute, University of Michigan, Ann Arbor, MI

[2]Department of Biological Chemistry, University of Michigan, Ann Arbor, MI

[3]Department of Medicinal Chemistry, University of Michigan, Ann Arbor, MI

[4]Center for Structural Biology, University of Michigan, Ann Arbor, MI 48109

[5]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA 92093

[6]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093

[7]Department of Chemistry, University of Michigan, Ann Arbor, MI

[8]Department 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 ± 0.05 mM, k_{cat} = 2.4 ± 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; α/β 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, NRB 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

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."[9] 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.[10] Two structures are known from evolutionary subfamily I (DhlA[11–13] and DppA[14]), and four from subfamily II (DhaA,[15,16] LinB,[17–20] DbjA,[21] and Rv2578[22]). 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.[9]

DmmA is a putative HLD of subfamily II.[10] This protein was originally annotated as CurN, and presumed to be the final gene product of the curacin A biosynthetic gene cluster[23] from the marine cyanobacterium *Lyngbya majuscula* (now designated *Moorea producta*[24]). However, resequencing revealed the authentic 3′ end of the gene cluster, which lacks *curN*,[25] suggesting that *dmmA* was cloned as part of a chimeric cosmid from the original metagenomic DNA library.[23] Herein, we present the crystal structure and HLD activity of DmmA.

Results

Biological source of DmmA (formerly CurN)

In the original curacin pathway sequencing,[23] 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 resequence the *curM-curN* region of the gene cluster, resulting in a corrected 3′ terminus that lacked *curN*.[25] In addition, no CurN coding sequence was identified by amplification from *M. producta (*formerly *L. majuscula)* genomic DNA, nor was CurN located in the recently sequenced *M. producta* 3L genome.[26] 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 *M. producta* genome ranged from 37 to 66% G + C content, with the great majority falling between 40 and 45%.[26] In addition, CurN (now referred to as DmmA (dehalogenase A from a marine microbe) in accord with other dehalogenases of unknown biological function[8]) 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 *M. 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 heterotroph bacteria that was growing on the *M. producta* filament at the time of cosmid preparation.

Table I. *DmmA Activity Toward Halogenated Substrates*

Construct	Substrate	% Activity[a]
$DmmA_{short}$	1,3-dibromopropane	100
$DmmA_{short}$	1,6-dibromohexane	3.5 ± 0.7
$DmmA_{short}$	bromocyclohexane	14.8 ± 0.6
$DmmA_{short}$	1,6-dichlorohexane	3.7 ± 1.4
$DmmA_{short}$	1,6-diiodohexane	3.2 ± 0.7
$DmmA_{long}$	1,3-dibromopropane	10 ± 2

[a] Normalized to activity of $DmmA_{short}$ with 1 mM 1,3-dibromopropane.

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.[10] 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 σ^{70} promoter regions and ribosome binding sequences (Supporting Information Fig. 1). In addition, several other annotated HLDs have N-termini of similar length to DmmA.[10] 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 II. *Steady State Kinetic Parameters of $DmmA_{short}$*

k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
2.4 ± 0.1	0.24 ± 0.05	1.0 ± 0.3 × 10^4

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 α/β 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 Cα 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)].

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,

Table III. *Crystallographic Summary*

	$DmmA_{short}$ (SeMet)	$DmmA_{short}$ (SeMet) 1,5-dibromopentane soaked	$DmmA_{long}$
Diffraction Data			
Space group	$P6_2$	$P6_2$	$P3_121$
X-ray source	APS 23-ID-D	APS 23-ID-D	APS 23-ID-D
a, b, c (Å)	99.8, 99.8, 122.0	100.3, 100.3, 121.4	144.7, 144.7, 105.0
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Wavelength (Å)	0.97939	0.92003	0.97939
d_{min} (Å)	2.20 (2.28–2.2)[a]	2.10 (2.18–2.10)	2.50 (2.50–2.64)
Avg I/σ_I	13.2 (2.9)	16.7 (2.1)	10.4 (3.5)
R_{symm}[b]	0.122 (0.549)	0.118 (0.75)	0.171 (0.607)
Completeness	100.0 (99.9)	99.7 (97.6)	100 (100)
Avg. redundancy	5.8 (5.2)	10.5 (7.6)	9.6 (8.6)
Unique reflections	35,082	40,262	44,165
Refinement			
Data range (Å)	86.39–2.20		
No. reflections	34,919		
R_{work}/R_{free}[c]	0.139/0.177		
RMS deviations			
Bonds (Å)	0.009		
Angles (°)	1.195		
Avg *B*-factors (Å^2)			
Protein	23.4		
Ligand/Ion	53.3		
Water	35.0		
Ramachandran			
Allowed	99.66%		
Outliers	0.34%		
Number of Atoms			
Protein	4649		
Ligand/Ion	18		
Water	497		

[a] Outermost shell in parentheses.
[b] Including anomalous differences.
[c] The R_{free} data set included a random 5% of reflections.

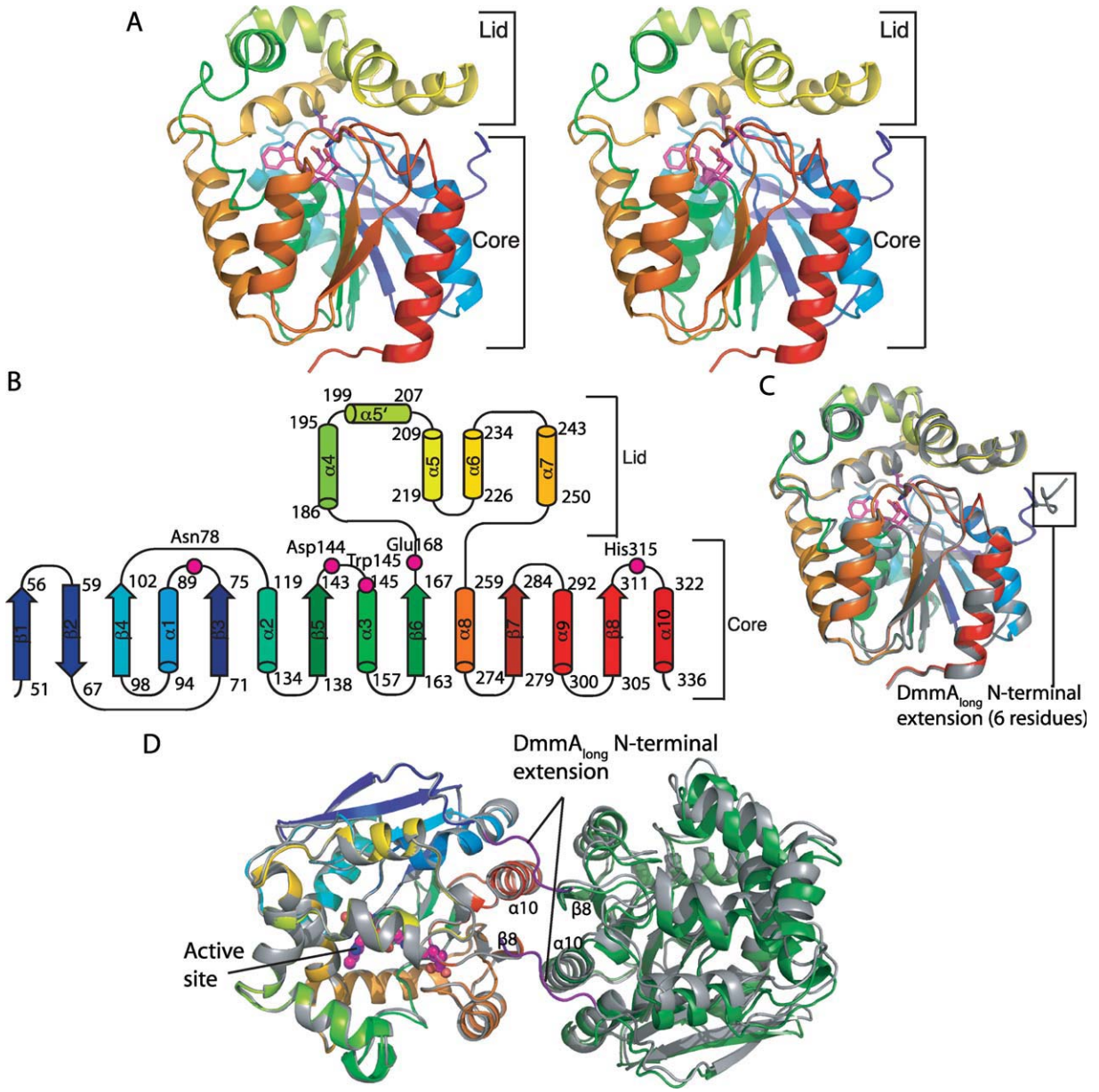

Figure 1. Structure of DmmA. A: DmmA polypeptide. The stereo ribbon diagram is colored as a rainbow from blue at the N-terminus to red at the C-terminus with catalytic pentad residues in stick form with magenta C. B: Topology diagram. The α/β hydrolase core is conserved among all α/β hydrolases and the lid is conserved among haloalkane dehalogenases. Residues of the catalytic pentad are labeled. C: Superposition of $DmmA_{long}$ (gray) and $DmmA_{short}$ (rainbow main-chain, magenta active site). Six additional residues on the N-terminus of $DmmA_{long}$ are visible. D: Protein-protein contact in the crystal structures. Only the left molecules of $DmmA_{short}$ (rainbow main-chain, magenta active site) and $DmmA_{long}$ (gray, purple N-terminal extension) are superimposed. The near superposition of the right molecules highlights the similar interfaces. Contacts of α10, β8, and N-terminal extension are labeled.

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 $Å^2$ vs. 610 $Å^2$) [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.[36] 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

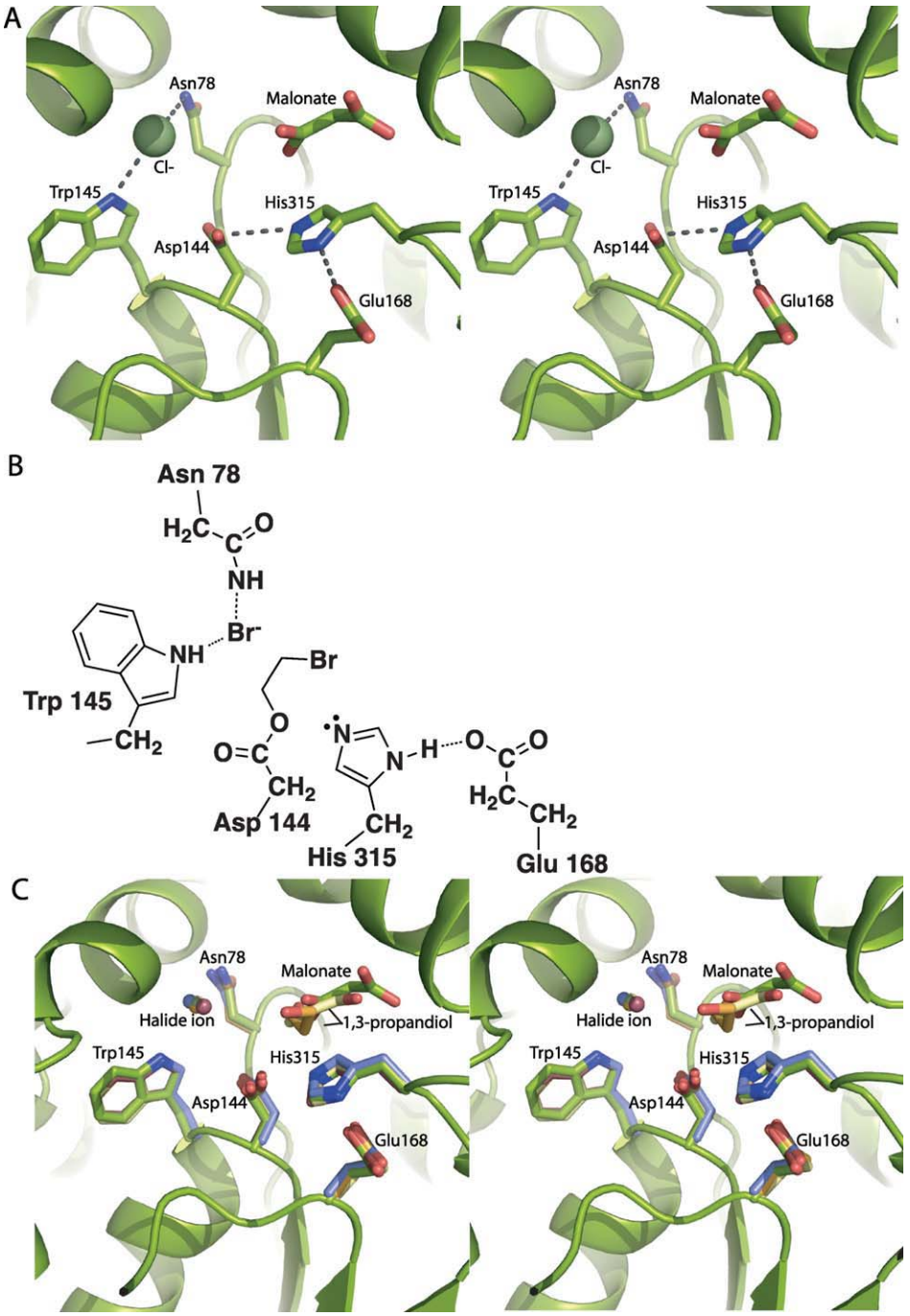

Figure 2. DmmA active site. A: DmmA active site showing catalytic pentad (sticks), Cl^- and malonate. B: Schematic of the DmmA active site showing the acyl-enzyme intermediate. The halide-stabilizing residues (Asn78 and Trp 145) bind the halide released when generating the acyl enzyme intermediate. The 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 Cα 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 2O2I[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-propandiol, for Rv2578 (yellow C) and LinB (orange C).

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

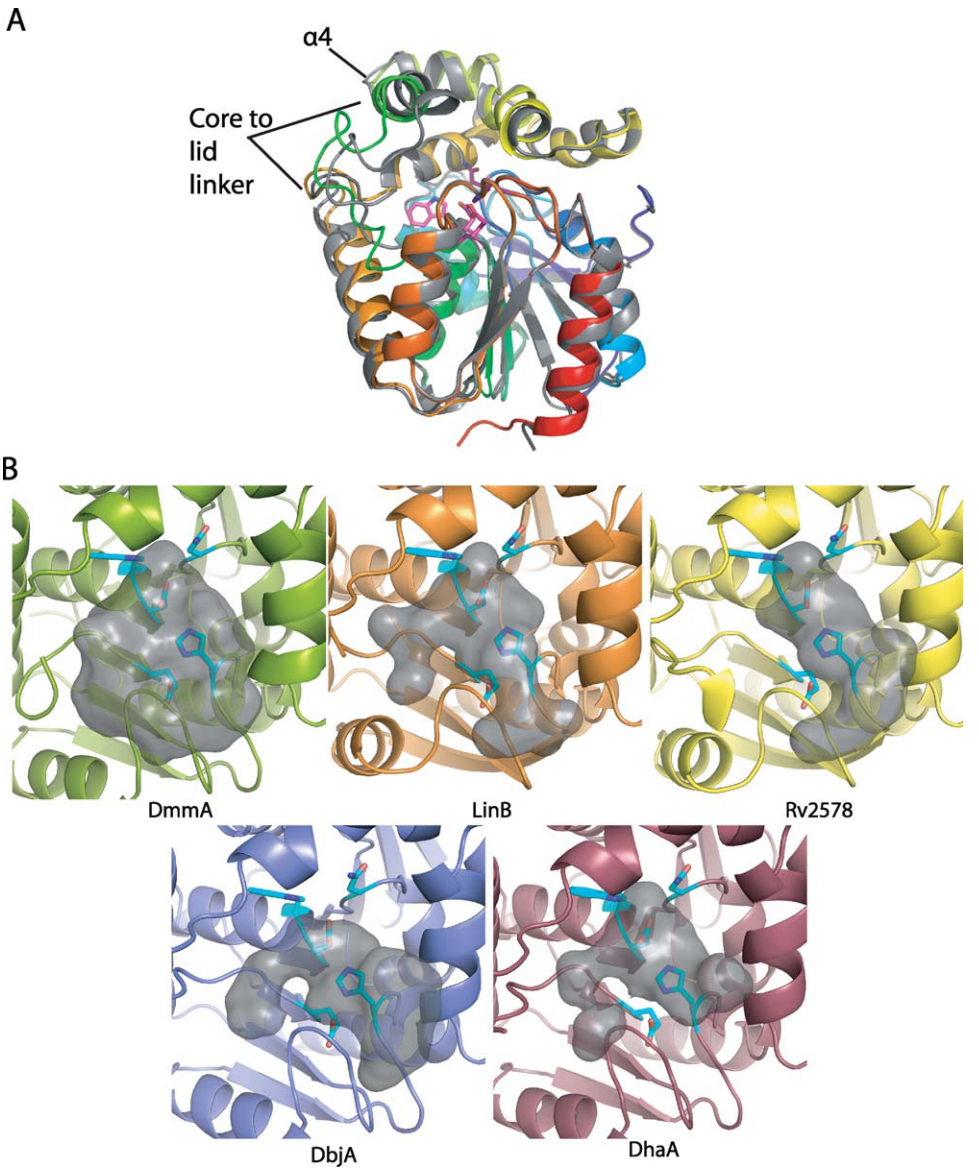

Figure 3. DmmA and other HLDs. A: Superposition of DmmA (rainbow main-chain, magenta active site) and the HLD DhaA (1BN7[15]). Differences in the lid-to-core linker and α4 are labeled. B: Active site cavity of DmmA and other type II haloalkane dehalogenases: DmmA (green), LinB 1IZ8[19] (orange), Rv2578 2O2I[22] (yellow), DbjA 3A2M[21] (purple), and DhaA 1CQW[15] (mauve). The gray surface represents the internal volume of the cavity. Residues that form the forward surface are omitted. Active site residues are shown in sticks (cyan C, blue N, red O).

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 are 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 $DmmA_{short}$ 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 II 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[8] and chiral selectivity.[21] 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[17,21] have very broad substrate specificity and react well with larger substrates.[8] 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, 1,6-diiodohexane, and bromocyclohexane) (Table I). In contrast to those tested previously,[8] DmmA has greater activity for the bromocyclohexane compared to the larger straight-chain substrates. This observation is consistent with the wide cavity of DmmA and suggests that it has robust activity toward other bulky substrates.

Discussion

This work establishes that DmmA is a HLD. The kinetic properties of DmmA with a standard substrate, 1,3-dibromopropane, are within the range reported for other HLDs. However, as for most HLDs, the natural substrates are unknown. Of the two protein variants studied, $DmmA_{short}$ was significantly faster than $DmmA_{long}$, exhibiting tenfold greater activity. Although most of the additional 43 residues on the N-terminus of $DmmA_{long}$ are disordered, they may interfere with substrate access to the active site, resulting in the reduced activity observed for the longer form of the protein.

The DmmA structure has the same fold and active site position as other HLDs, as expected. Nevertheless, DmmA has unique structural features that set it apart from other enzymes within this class. The active site cavity of DmmA is significantly larger than the cavity of other homologs, which was not predictable from straightforward sequence comparisons to other HLDs of known structure. Instead, sequence differences led to subtle structural variation in the position of helix α4 and the core-to-lid linker, which resulted in a large cavity volume. The size of amino acid side chains lining the active site pocket makes a secondary contribution to the cavity size. The larger cavity volume may confer an ability to accept a wide range of large, bulky substrates compared to other HLDs. Among the larger substrates tested, DmmA had a preference for the bulky bromocyclohexane over corresponding linear molecules. 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*[23]) 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, barbamide, which possesses a trichloromethyl group,[37] and in another strain, jamaicamide, which possesses vinyl chloride and alkynyl bromide groups.[38] 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,[23] inserted into the pET-24b vector, and verified by sequencing to form $pDmmA_{long}$ (encoding residues 1-341) and $pDmmA_{short}$ (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_{600} of 0.5, cooled to 18°C, induced with IPTG (final concentration 0.4 mM) and grown for an additional 18 h. Selenomethionyl (SeMet) $DmmA_{short}$ was produced in the same strain in SeMet minimal media.[39]

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 Ni NTA 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) prequilibrated with Buffer A. DmmA was concentrated to 14 mg/mL, flash frozen in liquid N_2, and stored at −80°C. SeMet $DmmA_{short}$ was purified as the wild type with addition of 2 mM DDT 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 within 24–48 h from a 1:1 mix of protein stock (14 mg/mL $DmmA_{long}$ in Buffer A) and well solution (0.6 M lithium sulfate, 27.5% PEG3350, and 0.1 M BisTris pH 5.5). Crystals of $DmmA_{long}$ were cyroprotected 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.[40] Determination of selenium atomic positions, experimental phasing, density modification phase refinement, and initial model building were performed using the programs SOLVE and RESOLVE.[41,42] And 22 of the 24 expected selenium positions were identified. The model was finished manually in COOT.[43] REFMAC5[30] from the CCP4 suite[31] was used for refinement. The $DmmA_{long}$ structure was solved by molecular replacement in Phaser[32] using $DmmA_{short}$ as a search model. Model quality was evaluated with MolProbity.[33] Figures were made using PyMOL[34] and hollow.[35] 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.[27,28] Substrate [0.1–6.0 mM 1,3-dibromopropane (Aldrich)] was prepared in an indicator solution (20 μg/mL phenol red, 20 mM Na_2SO_4, 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 ($v_{initial} = ([enzyme]*[substrate]*k_{cat})/(K_m+[substrate])$ to obtain steady state kinetic constants. Comparison with other substrates (1,6-dibromohexane (Acros), 1,6-dichlorohexane (Aldrich), 1,6-diiodohexane (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

1. Stucki G, Thuer M (1995) Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. Environ Sci Technol 29:2339–2345.
2. Alcalde M, Ferrer M, Plou FJ, Ballesteros A (2006) Environmental biocatalysis: from remediation with enzymes to novel green processes. Trends Biotechnol 24:281–287.
3. Swanson PE (1999) Dehalogenases applied to industrial-scale biocatalysis. Curr Opin Biotechnol 10: 365–369.
4. Prokop Z, Oplustil F, DeFrank J, Damborsky J (2006) Enzymes fight chemical weapons. Biotechnol J 1: 1370–1380.
5. Campbell DW, Muller C, Reardon KF (2006) Development of a fiber optic enzymatic biosensor for 1,2-dichloroethane. Biotechnol Lett 28:883–887.
6. Bidmanova S, Chaloupkova R, Damborsky J, Prokop Z (2010) Development of an enzymatic fiber-optic biosensor for detection of halogenated hydrocarbons. Anal Bioanal Chem 398:1891–1898.
7. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, Wood MG, Learish R, Ohana RF, Urh M, Simpson D, Mendez J, Zimmerman K, Otto P, Vidugiris G, Zhu J, Darzins A, Klaubert DH, Bulleit

RF, Wood KV (2008) HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem Biol 3:373–382.
8. Koudelakova T, Chovancova E, Brezovsky J, Monincova M, Fortova A, Jarkovsky J, Damborsky J (2011) Substrate specificity of haloalkane dehalogenases. Biochem J 435:345–354.
9. Janssen DB (2004) Evolving haloalkane dehalogenases. Curr Opin Chem Biol 8:150–159.
10. Chovancova E, Kosinski J, Bujnicki JM, Damborsky J (2007) Phylogenetic analysis of haloalkane dehalogenases. Proteins 67:305–316.
11. Verschueren KH, Kingma J, Rozeboom HJ, Kalk KH, Janssen DB, Dijkstra BW (1993) Crystallographic and fluorescence studies of the interaction of haloalkane dehalogenase with halide ions. Studies with halide compounds reveal a halide binding site in the active site. Biochemistry 32:9031–9037.
12. Verschueren KH, Seljee F, Rozeboom HJ, Kalk KH, Dijkstra BW (1993) Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. Nature 363:693–698.
13. Pikkemaat MG, Ridder IS, Rozeboom HJ, Kalk KH, Dijkstra BW, Janssen DB (1999) Crystallographic and kinetic evidence of a collision complex formed during halide import in haloalkane dehalogenase. Biochemistry 38:12052–12061.
14. Hesseler M, Bogdanovic X, Hidalgo A, Berenguer J, Palm GJ, Hinrichs W, Bornscheuer UT (2011) Cloning, functional expression, biochemical characterization, and structural analysis of a haloalkane dehalogenase from *Plesiocystis pacifica* SIR-1. Appl Microbiol Biotechnol 91:1049–1060.
15. Newman J, Peat TS, Richard R, Kan L, Swanson PE, Affholter JA, Holmes IH, Schindler JF, Unkefer CJ, Terwilliger TC (1999) Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. Biochemistry 38: 16105–16114.
16. Stsiapanava A, Dohnalek J, Gavira JA, Kuty M, Koudelakova T, Damborsky J, Kuta Smatanova I (2010) Atomic resolution studies of haloalkane dehalogenases DhaA04, DhaA14 and DhaA15 with engineered access tunnels. Acta Crystallogr D Biol Crystallogr 66: 962–969.
17. Marek J, Vevodova J, Smatanova IK, Nagata Y, Svensson LA, Newman J, Takagi M, Damborsky J (2000) Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. Biochemistry 39: 14082–14086.
18. Oakley AJ, Prokop Z, Bohac M, Kmunicek J, Jedlicka T, Monincova M, Kuta-Smatanova I, Nagata Y, Damborsky J, Wilce MC (2002) Exploring the structure and activity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26: evidence for product- and water-mediated inhibition. Biochemistry 41:4847–4855.
19. Streltsov VA, Prokop Z, Damborsky J, Nagata Y, Oakley A, Wilce MC (2003) Haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26: X-ray crystallographic studies of dehalogenation of brominated substrates. Biochemistry 42:10104–10112.
20. Oakley AJ, Klvana M, Otyepka M, Nagata Y, Wilce MC, Damborsky J (2004) Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 A resolution: dynamics of catalytic residues. Biochemistry 43:870–878.
21. Prokop Z, Sato Y, Brezovsky J, Mozga T, Chaloupkova R, Koudelakova T, Jerabek P, Stepankova V, Natsume R, van Leeuwen JG, Janssen DB, Florian J, Nagata Y, Senda T, Damborsky J (2010) Enantioselectivity of haloalkane dehalogenases and its modulation by surface loop engineering. Angew Chem Int Ed Engl 49: 6111–6115.
22. Mazumdar PA, Hulecki JC, Cherney MM, Garen CR, James MN (2008) X-ray crystal structure of *Mycobacterium tuberculosis* haloalkane dehalogenase Rv2579. Biochim Biophys Acta 1784:351–362.
23. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, Sherman DH, Gerwick WH (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. J Nat Prod 67:1356–1367.
24. Engene N, Rottacker EC, Kastovsky J, Byrum T, Choi H, Ellisman MH, Komarek J, Gerwick WH (in press) *Moorea producta* gen. nov., sp. nov. and *Moorea bouillonii* comb. nov., tropical marine cyanobacteria rich in bioactive secondary metabolites. Int J Syst Evol Microbiol.
25. Gu L, Wang B, Kulkarni A, Gehret JJ, Lloyd KR, Gerwick L, Gerwick WH, Wipf P, Hakansson K, Smith JL, Sherman DH (2009) Polyketide decarboxylative chain termination preceded by o-sulfonation in curacin a biosynthesis. J Am Chem Soc 131:16033–16035.
26. Jones AC, Monroe EA, Podell S, Hess WR, Klages S, Esquenazi E, Niessen S, Hoover H, Rothmann M, Lasken RS, Yates JR, III, Reinhardt R, Kube M, Burkart MD, Allen EE, Dorrestein PC, Gerwick WH, Gerwick L (2011) Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. Proc Natl Acad Sci U S A 108: 8815–8820.
27. Schindler JF, Naranjo PA, Honaberger DA, Chang CH, Brainard JR, Vanderberg LA, Unkefer CJ (1999) Haloalkane dehalogenases: steady-state kinetics and halide inhibition. Biochemistry 38:5772–5778.
28. Marvanova S, Nagata Y, Wimmerova M, Sykorova J, Hynkova K, Damborsky J (2001) Biochemical characterization of broad-specificity enzymes using multivariate experimental design and a colorimetric microplate assay: characterization of the haloalkane dehalogenase mutants. J Microbiol Methods 44:149–157.
29. Bosma T, Pikkemaat MG, Kingma J, Dijk J, Janssen DB (2003) Steady-state and pre-steady-state kinetic analysis of halopropane conversion by a *rhodococcus* haloalkane dehalogenase. Biochemistry 42:8047–8053.
30. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–255.
31. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760–763.
32. Zwart PH, Afonine PV, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, McKee E, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Storoni LC, Terwilliger TC, Adams PD (2008) Automated structure solution with the PHENIX suite. Methods Mol Biol 426: 419–435.
33. Chen VB, Arendall WB, III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66:12–21.
34. DeLano WL (2002) The PyMOL Molecular graphics system, Version 1.3, Schrödinger, LLC. Palo Alto, CA: DeLano Scientific.
35. Ho BK, Gruswitz F (2008) HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. BMC Struct Biol 8:49–54.

36. Hasan K, Fortova A, Koudelakova T, Chaloupkova R, Ishitsuka M, Nagata Y, Damborsky J, Prokop Z (2011) Biochemical characteristics of the novel haloalkane dehalogenase DatA, isolated from the plant pathogen *Agrobacterium tumefaciens* C58. Appl Environ Microbiol 77:1881–1884.
37. Chang Z, Flatt P, Gerwick WH, Nguyen VA, Willis CL, Sherman DH (2002) The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. Gene 296:235–247.
38. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. Chem Biol 11:817–833.
39. Guerrero SA, Hecht HJ, Hofmann B, Biebl H, Singh M (2001) Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. Appl Microbiol Biotechnol 56:718–723.
40. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode methods in enzymology 276:307–326.
41. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr 55:849–861.
42. Terwilliger TC (2003) Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallogr D Biol Crystallogr 59: 38–44.
43. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132.