Exploring the Binding Sites of the Haloalkane Dehalogenase DhlA from Xanthobacter autotrophicus GJ10 †

ARTICLE *in* BIOCHEMISTRY · SEPTEMBER 2007

Impact Factor: 3.02 · DOI: 10.1021/bi700336y · Source: PubMed

3 AUTHORS:

Michael Silberstein

6 PUBLICATIONS **152** CITATIONS

SEE PROFILE

Jiří Damborský

Masaryk University

193 PUBLICATIONS **3,984** CITATIONS

SEE PROFILE

Sandor Vajda

Boston University

190 PUBLICATIONS **7,251** CITATIONS

SEE PROFILE

Exploring the Binding Sites of the Haloalkane Dehalogenase DhlA from *Xanthobacter autotrophicus* GJ10[†]

Michael Silberstein,[‡] Jiri Damborsky,[||] and Sandor Vajda*[,§]

Bioinformatics Program and the Departments of Biology and Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215, and Loschmidt Laboratories, Masaryk University, Faculty of Science, Brno, Czech Republic

Received February 16, 2007; Revised Manuscript Received May 28, 2007

ABSTRACT: The catalytic site of haloalkane dehalogenase DhlA is buried more than 10 Å from the protein surface. While potential access channels to this site have been reported, the precise mechanism of substrate import and product export is still unconfirmed. We used computational methods to examine surface pockets and their putative roles in ligand access to and from the catalytic site. Computational solvent mapping moves small organic molecule as probes over the protein surface in order to identify energetically favorable sites, that is, regions that tend to bind a variety of molecules. The mapping of three DhlA structures identifies seven such regions, some of which have been previously suggested to be involved in the binding and the import/export of substrates or products. These sites are the active site, the putative entrance of the channel leading to the active site, two pockets that bind Br^- ions, a pocket in the slot region, and two additional sites between the main domain and the cap of DhlA. We also performed mapping and free energy analysis of the DhlA structures using the substrate, 1,2-dichloroethane, and halide ions as probes. The findings were compared to crystallographic data and to results obtained by CAVER, a program developed for finding routes from protein clefts and cavities to the surface. Solvent mapping precisely reproduced all three Br^- binding sites identified by protein crystallography and the openings to four channels found by CAVER. The analyses suggest that (i) the active site has the highest affinity for the substrate molecule, (ii) the substrate initially binds at the entrance of the main tunnel, (iii) the site Br2, close to the entrance, is likely to serve as an intermediate binding site in product export, (iv) the site Br3, induced in the structure at high concentrations of Br^-, could be part of an auxiliary route for product release, and (v) three of the identified sites are likely to be entrances of water-access channels leading to the active site. For comparison, we also mapped haloalkane dehalogenases DhaA and LinB, both of which contain significantly larger and more solvent accessible binding sites than DhlA. The mapping of DhaA and LinB places the majority of probes in the active site, but most of the other six regions consistently identified in DhlA were not observed, suggesting that the more open active site eliminates the need for intermediate binding sites for the collision complex seen in DhlA.

Haloalkane dehalogenases (E.C.[1] 3.8.1.5) are microbial enzymes that hydrolyze a wide variety of haloalkyl substrates to form the corresponding alcohol and halide ion products. Because the halogenated substrates are often environmentally toxic industrial byproducts, these enzymes have been suggested to be useful catalysts for biodegradation and bioremediation (*1*, *2*). The haloalkane dehalogenase from the nitrogen-fixing bacterium *Xanthobacter autotrophicus* (DhlA) prefers 1,2-dichloroethane (DCE) as substrate and converts it to 2-chloroethanol and chloride (*3*). The structural architecture in haloalkane dehalogenases includes two interacting domains. The main domain contains a commonly observed α/β-hydrolase fold, whereas the more flexible cap domain is principally α-helical in structure and sits directly on the main domain (*4*). The active site lies at the interface of these two domains within a buried cavity and contains a catalytic triad consisting of a nucleophilic aspartate, an adjacent histidine that acts as the catalytic base, and either an aspartate or glutamate residue that serves as the catalytic acid. The four-step reaction mechanism of DhlA is well understood (*5*−*8*). The catalytic cycle begins with the formation of a Michaelis complex, followed by the S_N2 displacement of chloride by the nucleophilic attack of Asp124 to form a covalent alkyl-enzyme intermediate (*8*). During the reaction, the halide moiety is stabilized by the indole NH groups of Trp125 and Trp175. Afterward, His289 activates water in

[†] This investigation was supported by Grant P42 ES07381 from the National Institute of Environmental Health Sciences, Grant GM64700 from the National Institutes of Health, Grant LC06010 from the Czech Ministry of Education, and Grant DBI 0213832 from the National Science Foundation.

* To whom correspondence should be addressed. Tel: 617-353-4757. Fax: 617-353-6766. E-mail: vajda@bu.edu.

[‡] Bioinformatics Program, Boston University.

[§] Department of Biomedical Engineering, Boston University.

[||] Masaryk University, Faculty of Science.

[1] Abbreviations: E.C., Enzyme Comission number; DhlA, haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10; DCE, 1,2-dichloroethane; DhaA, haloalkane dehalogenase from *Rhodococcus* sp.; LinB, haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26; ACE, analytical continuum electrostatics; GRAMM, global range molecular matching; pdb, Protein Data Bank; rmsd, root mean square deviation.

10.1021/bi700336y CCC: $37.00
Published on Web 07/24/2007

order to hydrolyze the covalently bound intermediate and to form the alcohol product, which is rapidly released (*7*). In the rate-limiting last step, the halide ion leaves the cavity.

The hydrolytic cleavage of the carbon−halogen bond by the haloalkane dehalogenase DhlA has been the subject of many theoretical/computational studies (*9*−*29*), primarily using quantum mechanics, molecular dynamics, and quantum mechanics/molecular mechanics calculations for several reasons. First, the S_N2 reaction is one of the simplest reactions and has also been one of the best studied reactions (*27*). Second, the enzyme is monomeric, and the crystal structure has been solved at high resolution (*30*). Third, the natural substrate DCE is very small, and hence, simulations are computationally less demanding (*31*). Fourth, DhlA is ideal for the analysis of solvation in enzymatic reactions, as cleavage is carried out in buried active site cavity that is completely hydrophobic apart from the catalytic residues, and simulation can be carried out considering only a limited number of water molecules (*13*, *14*, *21*, *23*).

However, the restricted active site in DhlA also raises a number of questions. First, the substrate has to move almost 10 Å from the protein surface to the active site (*32*), and it would be interesting to know where exactly the substrate enters the putative tunnel. Second, it is not clear along what potential channels water gets to the active site. Notice that in addition to the activated water molecule participating in the reaction, water is needed to solvate the leaving halide ion (*33*). Third, the mechanism of halide dissociation also has uncertain elements. From stopped-flow fluorescence analysis of bromide import, Schanstra et al. (*33*) concluded that the enzyme employs two different routes that operate in parallel. At low bromide concentrations, the import is dominated by a route that consists of a slow unimolecular isomerization step, followed by a rapid bimolecular step, and a second enzyme isomerization. At higher bromide concentrations, a second route for halide binding can also be observed involving a fast bimolecular step preceding a slow enzyme isomerization. The bimolecular step has been explained as the rapid binding of a halide ion into an initial collision complex with the enzyme. Both routes assume that halide import/export requires some conformational change to occur because of relatively slow isomerization of the enzyme, but no further information on this step is available.

In addition to DhlA, two additional haloalkane dehalogenase variants exist in which there are pertinent crystallographic data: DhaA from *Rhodococcus* sp (*34*) and LinB from *Sphingomonas paucimobilis* UT26 (*35*). Although DhlA has been shown to dehalogenate haloalkane substrates with fewer than six carbon atoms, DhaA and LinB are able to dehalogenate larger haloalkanes ranging from 6 to 10 carbons. One primary difference in specificity between DhaA and LinB appears to be the enhanced capability of LinB to catalyze the dehalogenation of compounds with chlorines or bromines attached to secondary carbons versus that of DhaA. Thus, LinB is slightly less sterically restricted than DhaA (and is significantly less restricted than DhlA) (*37*). In agreement with their larger substrates, the volumes of the binding sites are 2 and 2.5 times larger in DhaA and LinB, respectively, than in DhlA (*37*). In addition, in DhaA and LinB the active site is more accessible than in DhlA (*37*).

In this article, we employ a relatively new method, computational solvent mapping (*24*, *38*, *39*), to explore the potential binding sites of three different DhlA structures. This method places small organic molecule probes over the protein surface, finds the most favorable positions, clusters the conformations, and ranks the clusters on the basis of average free energy (*38*). The idea of solvent mapping is based on the multiple solvent crystal structure experiments by Ringe and co-workers (*40*, *41*), who determined protein structures in aqueous solutions of organic solvents and in each case found only a limited number of solvent molecules bound to the protein. When five or six structures of a protein determined in different solvents were superimposed, the organic molecules tended to cluster in the active site, forming consensus sites that delineate important subsites of the binding pocket. All other bound solvent molecules are either in crystal contact, occur only at high ligand concentration, or are in small, buried pockets, where only a few types of solvent molecules cluster compared to that in the active site (*42*).

Table 1: List of Haloalkane Dehalogenase Structures Mapped

protein	pdb ID	binding state of the crystal structure
DhlA	2dhc	substrate-bound structure (DCE bound)
	2had	ligand-free structure
	1cij	halide-bound structure (three bromides bound)
DhaA	1cqw	halide-bound structure (two iodides bound)
LinB	1k6e	product-bound structure (1,2-propanediol and 1-bromopropane-2-ol bound)
	1iz7	ligand-free structure

The computational solvent mapping algorithm reproduced the available X-ray and NMR results on the binding of organic solvents to proteins (*38*). In all our previous applications to enzymes, the consensus sites found always corresponded to major subsites of the substrate-binding site (*24*, *38*). The method also applies to nonenzyme proteins (*43*). Mapping is particularly useful for comparing different structures of the same protein, as the differences in the number of probe clusters that bind to a particular surface region indicate even the very small conformational changes if those affect the binding properties of the pocket. This feature was fully utilized in our analysis of the ligand-binding domain of the peroxisome proliferator activated receptor-γ, where mapping identified 10 binding sites and revealed how conformational changes in the ligand-binding pocket correlate with changes in the co-activator-binding region (*43*). We emphasize that prior to mapping, all bound ligands and water molecules are removed. Thus, the results are based only on the structure of the protein, and hence, the method can be used to identify and characterize binding sites even without prior consideration of the chemical structure and properties of the bound ligand.

Here, we describe the application of computational solvent mapping to the three DhlA structures listed in Table 1. One of the structures has been cocrystallized with the substrate DCE (pdb code 2dhc), the second is a ligand-free structure (pdb code 2had), and the third is a halide-bound structure, crystallized in the presence of Br^- ions (pdb code 1cij). We show that the results indicate seven binding sites that consistently occur in all structures but with different binding properties. The first of these is the active site, denoted as A, which overlaps with the Br^- binding site Br1 in the halide-bound structure 1cij. The other six sites are the entrance E

of the tunnel leading to the active site, two Br^- binding sites (Br2 and Br3), a pocket (SL) in the slot region, and two additional sites (S1 and S2) between the main domain and the cap of DhlA. The biological roles of these sites were further studied by comparing the outcome of the mapping calculations to results obtained by CAVER, a program developed for finding routes from protein clefts and cavities to the surface (*44*). The three DhlA structures were also mapped using DCE and halide ions as probes. As expected, these probes also bind to most of the seven sites. To further explore the potential roles of these sites, we calculated the free energies of binding for DCE, Br^-, and Cl^- using the CHARMM potential with a continuum electrostatic model.

In addition to the analysis of the three DhlA structures, we have mapped a halide-bound DhaA structure (pdb code 1cqw) and two structures of LinB, the first with bound products 1,2-propanediol and 1-bromopropane-2-ol (pdb code 1k6e), and the second is a ligand-free structure (pdb code 1iz7). These structures are also listed in Table 1. As mentioned, DhaA and LinB are close homologues of DhlA. However, they have larger and more exposed active sites, and the mapping shows that this leads to a simpler binding site structure without the intermediate binding sites for the collision complex seen in DhlA.

MATERIALS AND METHODS

Haloalkane Dehalogenase Structures. The crystal structures used for computational solvent mapping and free-energy analyses are listed in Table 1. The protein structures are downloaded from the Protein Data Bank (pdb). All bound ligands and water molecules were removed prior to calculation (*38*). Protonation states for the ionizable side chains in the active site of DhlA were modeled as follows: E56 was protonated, D124 and D260 were deprotonated, and His289 was singly protonated on $N^{\delta 1}$. It was determined experimentally that this ionization state is suitable for characterizing the active site prior to any catalytic turnover (*32*). The ionization states for the structures of the haloalkane dehalogenases DhaA and LinB were modeled analogously.

Solvent Mapping

The five steps of computational solvent mapping were described previously (*24*, *38*, *39*), and only a summary is given here.

Step 1: Rigid Body Search. For each structure, seven small molecules (acetone, acetonitrile, *tert*-butanol, dimethylsulfoxide, phenol, isopropanol, and urea) and substrate DCE are used as probes (*24*). For each probe, 2000 docked conformations are generated by the rigid body docking algorithm GRAMM (*45*, *46*). GRAMM (global range molecular matching) requires only the atomic coordinates of the two molecules, that is, no *a priori* information on the binding site is used. The program places each molecule on a separate grid, and performs an exhaustive six-dimensional search through the relative intermolecular translations and rotations using a very efficient fast Fourier transform (FFT) correlation technique and a simple scoring function that measures shape complementarity and penalizes overlaps. We have used a 1.5 Å grid step for translations and 15° increments for rotations; 2000 docked conformations were retained.

Step 2: Minimization and Re-Scoring. The free energy of each of the 2000 complexes, generated in Step 1, is minimized using the free energy potential

$$\Delta G = \Delta E_{\mathrm{elec}} + \Delta E_{\mathrm{vdw}} + \Delta G_{\mathrm{des}} \qquad (1)$$

where ΔE_{elec}, ΔE_{vdw}, and ΔG_{des} denote the electrostatic, van der Waals, and desolvation contributions, respectively, to the protein-probe binding free energy (*38*). The sum $\Delta E_{\mathrm{elec}} + \Delta G_{\mathrm{des}}$ is obtained by the Analytic Continuum Electrostatic (ACE) model (*47*) as implemented in version 27 of CHARMM (*48*) using the parameter set from version 19 of the program. The model includes a surface area-dependent term to account for the solute−solvent van der Waals interactions. The minimization is performed using an adopted basis Newton−Raphson method as implemented in CHARMM. During minimization, the protein atoms are held fixed, while the atoms of the probe molecules are free to move.

Step 3: Clustering and Ranking. The minimized probe conformations from Step 2 are grouped into clusters on the basis of Cartesian coordinate information. The method creates an appropriate number of clusters such that the maximum distance between a cluster's hub and any of its members (the cluster radius) is smaller than half of the average distance between all the existing hubs. Clusters with less than 10 members are excluded from consideration. For each retained cluster, we calculated the probability $p_i = Q_i/Q$, where the partition function Q is the sum of the Boltzmann factors over all conformations, $Q = \Sigma_j \exp(-\Delta G_j/RT)$, and Q_i is obtained by summing the Boltzmann factors over the conformations in the ith cluster only. The clusters are ranked on the basis of their average free energies $\leq \Delta G >_i = \Sigma_j p_{ij} \Delta G_j$, where $p_{ij} = \exp(-\Delta G_j/RT)/Q_i$, and the sum is taken over the members of the ith cluster.

Step 4: Determination of Consensus Sites. Mapping is primarily used to find consensus sites at which many different probe molecules cluster. In order to find the consensus sites, we selected the minimum free energy conformation in each of the five lowest average free energy clusters for each solvent. The structures are superimposed, and the position at which most probes of different types overlap is defined as the main consensus site. An additional clustering of probes close to the main consensus site is likely to indicate another subsite of the active site. Ranking of the consensus sites is dependent on the number of different types of small molecules, with a maximum of 7, found within each consensus site, that is, a consensus site containing 6 of the 7 different probe molecules is ranked better than a site with only 4 of the 7 probe molecules. If two sites share the same number of different types, then duplicate types within the consensus site are considered in the count, that is, a consensus site containing clusters from all 7 different probe molecules and a second urea cluster is ranked better than a site containing a single cluster from all 7 different probe molecules. Any sites that are still equivalent by these comparisons are then ranked by average energy rank of the probe clusters included at the site, that is, a cluster where all 7 probe clusters were rank 1 by average Boltzmann-free energy for each probe type is ranked better than a cluster where all 7 probe clusters were rank 2 or lower.

Solvent Mapping Using Halide Ion Product as the Probe. The mapping algorithm was additionally applied toward the mapping of chloride or bromide ions to the three DhlA structures in Table 1. Because GRAMM cannot dock single ions, the DOCK 4.0 program (*49*) was used, which we have already tested as an alternative sampling method in Step 1 of the solvent mapping algorithm (*39*). As with GRAMM, we restricted consideration to shape complementarity as the target function in rigid body docking and used the van der Waals terms of the AMBER force field (*50*). In DOCK, the ligands were placed by a random search within the search box using a 0.3 Å grid step and 0.75 as the bump condition, eliminating any conformation in which two atoms approach closer than 0.75 of the sum of their radii. We generated 50,000 positions for each ion and retained 6660 of these for further analysis. The conformations were minimized using the simplex method in the DOCK program. The parameters selected were 0.01 kcal/mol/Å for the energy convergence condition and 0.1 Å rmsd over a minimization cycle for the cycle convergence. The minimization was limited to 10 or 20 cycles. Using GRAMM, the rigid body docking was followed by CHARMM minimization in Step 2 of the mapping algorithm, where all charges, including those of the ions, were properly considered.

Binding Free Energy Analysis of Protein−Ligand Complexes. After applying the solvent mapping algorithm using either the substrate, DCE, or the chloride/bromide ions as the respective probes, we have calculated binding free energies for clusters in specific pockets. The cluster centers from the mapping were co-minimized with the protein for 200 steps in CHARMM with continuum electrostatics, considering both the protein and the substrate to be flexible in the minimization. Afterward, the protein and the ligand were separated and their energies calculated and subtracted from the energy of the protein−ligand complex. The contributions ΔE_{elec}, ΔE_{vdw}, and ΔG^*_{des} to the calculated binding free energy ΔG were also determined for each of the pockets considered. Because the entropy loss upon binding is not taken into account, the ΔG values are relative and can be used only to compare the binding of the same ligand to different sites. In the absence of any relevant cluster center at a selected pocket, no data was recorded.

RESULTS AND DISCUSSION

Mapping of DhlA. Solvent mapping was applied to the three DhlA structures shown in Table 1: a structure cocrystallized with the substrate DCE, a ligand-free structure, and a structure with three bound Br^- ions. The number of probe molecules clustered in each consensus site and their respective energy rankings are listed in Table 2. The distance from the center of each cluster to the bound ligand is also shown (values in parenthesis). For example, the first row of Table 2 shows that the largest consensus site (site E) found by the mapping of the substrate-bound structure includes the third lowest-free-energy cluster of acetone, located 9.84 Å from the bound DCE ligand, the third lowest-free-energy cluster of urea, at 7.80 Å from DCE, and so on. A consensus site may include several clusters of the same probe, for example, for DMSO we find the clusters ranked 1 and 5 at the second consensus site S1 (Table 2).

As shown in Table 2, the binding of DCE or the Br^- ions to DhlA affects the results of the mapping. Nevertheless, seven sites occur in all three DhlA structures, although with different numbers of clusters bound (Table 3). These sites, shown in Figure 1, are the active site denoted as A (which overlaps with the Br^- ion binding site Br1 in the bromide-bound structure), the putative entrance E of the channel leading to the active site, two pockets Br2 and Br3 that bind Br^- ions in the bromide-bound structure, and the so-called slot region that is likely to serve as a water-access channel (*17*), denoted SL. In at least two of the three structures, two additional sites occur between the main domain and the cap of DhlA that will be denoted S1 and S2. Finally, on each protein we find some sites that are unique (i.e., not seen in any other structure). These sites, denoted U1, U2, and U3 are listed in Table 2 but are not discussed further in the article. Table 4 is the summary of the number of different clusters seen at the seven consensus sites in the structures mapped.

Most of the sites above are located along the interface between the two domains, and may play a role in allowing substrates and products to enter and exit the active site, which is located in a largely hydrophobic pocket buried in the interface. In both the DCE-bound and ligand-free structures, the largest consensus site (E in Table 4) is the putative entrance to the active site (with 6 and 7 probe clusters bound to the DCE-bound and ligand-free structures, respectively). The mapping also finds the active site (A in Table 4) as expected; however, the active site cavity is so small that only the small probes are able to get in (acetone, urea, and isopropanol for the substrate-bound structure, and also DMSO for the ligand-free structure). Larger probes such as *tert*-butanol and phenol were simply too large to fit properly (*24*).

The probes also cluster in the three bromide binding sites that have been observed in the respective DhlA crystal structure (Table 1). As already mentioned, the site Br1 is in the active site between the known anion-stabilizing residues Trp125 and Trp175. The second site, Br2, is located along a short loop connecting helices α6 (residues 187−194) and α7 (residues 200−208) and is adjacent to the putative entrance region (E). The third site, Br3 is found in between helices α5 (residues 171−181) and α6 (*51*). It has been proposed that the Br2 site may play a direct biological role in halide import (*51*). The site is present in both the bromide-bound and substrate-bound structures, but was not found in the ligand-free structure, indicating that it is induced by the binding of the substrate and remains open when the halide ion is in the active site. In contrast, the Br3 site is well defined only in the presence of bromide. It has been shown that the Br2 site also binds a chloride ion, whereas this was not the case for Br3 (*51*).

Table 4 shows three additional sites, SL, S1, and S2. The so-called slot region (site SL) has been reported as a putative water-access channel (*17*). As mentioned in the Introduction section, after hydrolysis the alcohol product is rapidly released, but the halide ion remains in the cavity. The removal of the ion requires its solvation by a number of water molecules (*33*) and thus the existence of appropriate water-access channels. It has been observed that halide release is the rate-limiting last step of the mechanism, and in addition to solvating the ion, it is likely to require conformational changes in the cap domain that could contribute to further opening of water-access channels (*17*). The conformational

Table 2: Free Energy Ranking[a] of Probe Clusters for All Mapped Haloalkane Dehalogenase Structures[b]

		probe[d]						
protein	con. site[c]	acetone	urea	DMSO	isopropanol	*t*-butanol	phenol	acetonitrile
DhlA	E	3 (9.8)	3 (7.8)	3 (9.7)	3 (8.1)	1 (9.8)	2 (8.0)	
cocrystallized with DCE (1,2-dichloroethane)	S1	2 (10.8)	4 (10.3)	1 (10.2), 5 (10.3)	4 (9.8)	3 (10.4)		
pdb code 2dhc	Br2	4 (12.4)	2 (11.4)	4 (13.1)	5 (12.6)			
	A	1 (0.6)	1 (0.5)		1 (0.7)			
	U1		5 (22.7)		2 (23.2)			1 (22.4)
	Br3	5 (17.1)					5 (16.3)	
DhlA	E	1 (9.4)	2 (7.5)	1 (8.9)	1 (9.4)	1 (9.8)	1 (7.2)	2 (8.8)
ligand-free structure	U2	5 (29.3)		2 (29.2)	3 (29.3)	2 (29.36)	3 (29.6)	
pdb code 2had	A	2 (0.6)	1 (0.6)	3 (0.8)	2 (0.9)			
	S1	4 (10.8)	4 (10.4)	4 (10.3)	5 (10.8)			
	SL	3 (9.8)	3 (8.9)			4 (10.5)		
DhlA	Br3	3 (15.1)	5 (14.4)	1 (14.9), 2 (15.32)	2 (14.9)	1 (15.3)	4 (14.4)	2 (15.0)
crystallized in the presence of Br^- ions	E	1 (9.3)		5 (9.0)	3 (9.3)		2 (8.1)	
(3 bound Br^- ions)	SL		1 (7.7)	3 (8.9)			1 (7.6), 5 (9.8)	
pdb code 2cij	Br2	2 (12.9)	-	4 (12.5)	1 (12.2)			
	A		3 (0.7)					4 (0.4)
DhaA	A	1 (3.1)	2 (3.4)	2 (2.1), 4 (2.96)	2 (3.4)		3 (2.3), 4 (1.78)	
crystallized in the presence of I^- ions	U3	2 (12.6)	1 (11.6)	1 (11.7)	1 (12.0)	1 (11.7)	1 (12.5)	
pdb code 1cqw								
LinB	A	1 (3.6)	4 (2.3)	1 (3.0), 2 (2.8)	1 (2.2)	2 (2.5)	1 (3.0)	2 (3.0)
ligand-free structure	S1	3 (7.8)		3 (8.3)	2 (7.3)	4 (8.4)	4 (9.5)	4 (8.1)
pdb code 1iz7	U3	2 (11.3)				1 (12.0)	5 (12.7)	
LinB	A	1 (3.6)	2 (2.2)	1 (3.0)	1 (2.2), 2 (2.18)	1 (2.6)	1 (2.9), 2 (3.25)	2 (3.0)
cocrystallized with substrates	U3	2 (12.3)	5 (11.1)	2 (12.0)	3 (12.3)	2 (12.3)	3 (12.5)	
pdb code 1k6e								

[a] The numbers in the Table indicate the rank of the cluster, e.g., 1 indicates the lowest free energy cluster, 2 the second lowest free lowest free energy, etc. [b] The structures are listed in Table 1. [c] The sites are defined in Table 3. [d] The number in parenthesis indicates the distance (in Å) of each cluster center from the ligand in the X-ray structure of the corresponding bound enzyme after structural superposition.

changes may also involve some domain−domain motions (*17*). The SL site is best defined in the bromide-bound structure, which is in good agreement with the assumption that this location can serve as a water-access channel.

Site S2 is located in such a way that the removal of the V226 side chain opens a wide channel to the active site. It has been shown that the V226A and V226L mutations both increase the reaction rate (*52*). Because these mutations provide easier access to the active site, it is possible that the S2 pocket also serves as a water-access channel. However, no clusters were found at this site in the bromide-bound structure. Site S1 is very well defined in the DCE-bound structure (Table 4). It is also present in the ligand-free structure, but similar to Site S2, it is missing in the bromide-bound structure. Thus, again it is not clear, based solely on the mapping results, whether this site has any role in the solvation of the halide reaction product.

Analysis of Paths from the Active Site Using CAVER. For further investigation as to whether sites SL, S1, and S2 can serve as water-access channels, we have considered prior analyses using the CAVER methodology (*44*). CAVER is a graph-searching algorithm implemented to identify direct paths connecting the active site of the enzyme to the protein surface and has been applied to various haloalkane dehalogenase structures (*44*). Computational mapping and the CAVER algorithm explore two very different properties of protein structures. Although CAVER was implemented to locate putative access paths, computational mapping identifies regions on the protein surface that bind a variety of small organic molecules. Thus, CAVER assumes prior knowledge of the active site in order to determine paths leading to the surface, whereas mapping focuses on the identification of protein pockets and clefts that bind small organic molecules. This distinction is fundamental in order to compare the two sets of results, as the ability of mapping to identify distinct paths is limited at best. However, the two methods provide complementary information toward elucidating the biological function of protein clefts and pockets, and as we will show, for DhlA all results are in good agreement, thereby improving the reliability of our conclusions.

The CAVER method was applied to 14 DhlA structures, including the three structures that have been computationally mapped (*44*). In all 14 structures, the method identified paths connecting the active site with the putative tunnel entrance (E), the slot region (SL), and S2 (Figure 2). Additionally, in 8 of these 14 structures (including the bromide-bound

Table 3: Binding Sites of DhlA Structures Identified by Mapping

site	description	surrounding residues
A	active site; overlaps with Br^- binding site Br1	E56, D124, W125, F128, M152, F164, F172, W175, L179, F222, P223, V226, L262, L263, H289
E	putative entrance to the active site	R193, W194, A195, P196, K259, D260, K261, L262, A287, G288, H289, F290
Br2	site of Br^- binding close to the entrance of the channel; putative site for collision complex	W194, A195, P196, T197, K259, A287, G288, H289, F290, V291, E293, F294
Br3	site of Br^- binding	D184, L185, R186, L187, Y206, A207, F210, P211, D212, T213, Q216
SL	putative water entrance channel at the slot region	F164, Q167, P168, D260, K261, L262, L263, G264, P265, D266, V267
S1	pocket on the opposite end of the ligand entrance channel, separated from the active site by M225	Y97, T98, F99, K221, M225, Q228, R229, D230
S2	pocket separated from the active site by V226	T153, D154, P155, S162, V165, V226, A227, Q228, R229, I234
Ux	site with unidentified functionality	

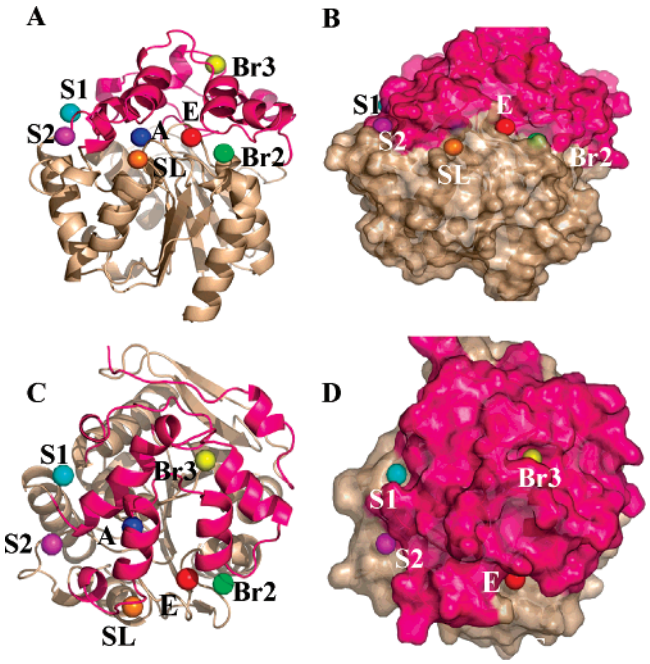

FIGURE 1: Depiction of the seven pertinent sites as observed in haloalkane dehalogenase DhlA structure (2dhc). In all four panels, the protein's cap domain is colored hot-pink, whereas the α/β hydrolase domain is colored wheat. Panels A and B represent ribbon and spacefill representations, respectively, of a side view of the structure (where the SL site, the tunnel entrance E, and the Br2 site are facing the viewer). Panels C and D represent ribbon and spacefill representations, respectively, where the cap domain has been placed in the front while the α/β hydrolase domain is toward the back. The coloring of the sites are as follows: active site (A), blue; tunnel entrance (E), red; Br2 site, green; Br3 site, yellow; slot region (SL), orange; S1 site, cyan; and S2 site, magenta.

structure previously mapped), a distinct path was also identified connecting the active site with S1 (Figure 2). Although this does not functionally verify the role of S1 and S2 as distinctive water-access channels, it does provide further evidence of distinctive paths, which potentially allow

Table 4: Number of Clusters in the Various Sites of the Haloalkane Dehalogenase Structures

	site[a]						
	A	E	Br2	Br3	SL	S1	S2
2dhc (DhlA)	3	6	4	2	1	6	1
2had (DhlA)	4	7	1		3	4	2
1cij (DhlA)	2	4	3	8	4		
1cqw (DhaA)	7				3[b]	1	
1iz7 (LinB)	8				1	6	
1k6e (LinB)	9				2		

[a] Each of these sites is defined in Table 3. [b] Three individual cluster centers were found at this site, which did not completely overlap, and hence are not considered as a single consensus site (see text for further discussion).

FIGURE 2: Superimposing the sites identified with solvent mapping with tunnels calculated by CAVER. The figures show the lowest free energy positions for the bound 1,2-dichloroethane (DCE) as determined by mapping. Although this lowest energy position is slightly shifted from the channel entrance in the relatively long slot region, the agreement between the results from the two methods is generally very good. (A) Side view, with channel entrance E in front. (B) Top view, with the cap domain facing the viewer.

water to penetrate the enzyme in order to hydrolyze the substrate and subsequently solvate the halide-ion product. As mentioned, the fact that the mutation V226A increases the reaction rate (*52*) also supports the assumption that S2 is the entrance of a water-access channel. Although no mutational information is available for S1, inspection of its location suggests a testable hypothesis, namely, that the mutation of residue M225 to one with a smaller side chain should increase the rate of the reaction. The observation that mapping finds the S1 and S2 sites both in the DCE-bound

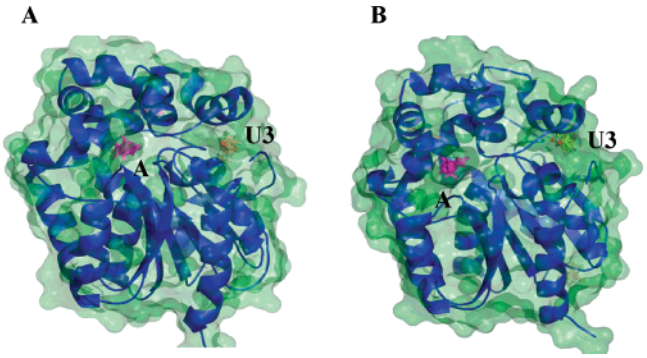

FIGURE 3: Surface (green) depiction of (A) haloalkane dehalogenase DhaA structure (1cqw) and (B) haloalkane dehalogenase LinB structure (1iz7) superimposed with the computationally mapped consensus sites. The active site is represented by the magenta colored consensus site. Notice the U3 site on the right side of the structures.

and ligand-free structures but not in the bromide-bound structures may indicate either that water molecules access the active site before the halide export step of the catalytic cycle or that there is no water uptake along these routes at the high Br^- concentration of crystallization.

Mapping of DhaA and LinB. We have mentioned that the mapping of enzymes generally yields the largest number of overlapping probe clusters in a subsite of the enzyme active site (*24*, *38*). This was not the case for DhlA, and we hypothesized that the unique behavior is due to the size of the active site, which is not accessible to the larger probes. To test this hypothesis, we have mapped an apostructure of DhaA (pdb code 1cqw) and both apo and holo structures of LinB (pdb codes 1iz7 and 1k6e, respectively) as shown in Table 1. The haloalkane dehalogenases DhaA and LinB, although highly homologous to DhlA, hydrolyze larger haloalkanes and thus have substantially larger active sites.

The mapping results for DhaA and LinB are also listed in Tables 2 and 4, and are additionally depicted in Figure 3. As shown in Table 4, the largest consensus site is the active site in all three structures mapped, with at least 7 probe clusters. Sites Br1, Br2, and S2 do not exist. The probes cluster in the active site, and there is no binding at the entrance E. The exception is the DhaA structure, with a single *tert*-butanol cluster at the entrance in the upper tunnel (*17*, *34*). The slot region SL shows up in all three structures, but contains only a few probe clusters. Although the SL site binds two acetonitrile clusters and one acetone cluster in DhaA, the three clusters do not fully overlap. There is an additional consensus site in each structure, listed as U3 in Table 2. It is located about 12 Å from the active site and is shown on the right side of the proteins in Figure 3. Although the site is fairly large, nothing is known about it and hence will not be discussed further. We emphasize that unlike for DhlA, there were no obvious structural changes observed in the DhaA and LinB protein structures that could be attributed to the presence of ligands in the active site, and hence, we do not provide a detailed comparison of the mapping results for the LinB apo and holo structures.

We compared the information derived by the mapping to prior results from the CAVER method (*44*). In three crystal structures of DhaA, CAVER consistently identified the enzyme's upper tunnel as the highest-ranked path (*17*, *34*). The remaining paths found by the method span through the slot region (*44*). Notice that the three non-overlapping cluster centers found by mapping in the slot region of DhaA is in accordance with the results of the CAVER calculations, which showed a number of different tunnel openings (tunnel branching) from the main tunnel. For 10 out of 11 LinB structures, the CAVER method identified the lower tunnel (*17*, *35*, *53*−*55*) as the most predominant path. Another path found in all but one crystal structure was the upper tunnel (*17*, *35*, *53*−*55*). A number of sub-paths within the slot region were also observed, but no other paths were found. It appears that for DhaA and LinB, the CAVER program provides more information than mapping does because it identifies the substrate access path, whereas mapping places the probes directly in the active site rather than at the entrance, as it did for DhlA. However, the latter result implies that the need for intermediate collision complex locations that exist for DhlA may not be present in DhaA and LinB because their active site is more accessible.

Table 5: Calculated Binding Free Energy (in kcal/mol) of 1,2-Dichloroethane (DCE) Bound to Various Sites[a] of the Three DhlA Structures Studied[b]

protein	site	DCE cluster	ΔG	ΔE_{vdw}	ΔE_{elec}	ΔG_{des}
DhlA	A	1	−20.4	−18.6	−1.5	−0.3
cocrystallized with DCE	S1	2	−15.7	−14.7	−1.0	+0.0
(1,2-dichloroethane)	Br2	3	−15.8	−15.7	−0.2	+0.2
pdb code 2dhc	Br3	6	−18.5	−15.7	−1.3	−1.4
	S2	16	−11.9	−10.6	−0.3	−1.0
DhlA	A	1	−20.3	−18.6	−0.9	−0.8
ligand-free structure	Br2	2	−16.3	−16.1	−0.9	+0.8
pdb code 2had	E	3	−15.7	−15.4	+1.5	−1.8
	S1	5	−16.7	−15.6	−0.9	−0.3
	S2	9	−16.7	−13.0	+0.8	−4.4
DhlA	E	1	−16.0	−15.6	+1.5	−1.9
crystallized in the presence	Br3	2	−17.9	−16.7	−1.3	0.0
of Br^- ions (3 bound	Br2	3	−16.2	−16.1	−0.6	+0.5
Br^- ions} pdb code 2cij	SL	5	−14.7	−14.8	+1.2	−1.1
	A	6	−20.4	−18.8	−1.8	−0.5
	S2	12	−16.1	−13.1	+0.7	−3.7
	S1	16	−11.7	−11.2	−0.4	−0.1

[a] The sites are defined in Table 3. [b] The structures are listed in Table 1.

Binding of DCE to DhlA. The small natural ligand of DhlA, DCE, can be used as a probe in mapping. After mapping, we performed 200 steps of energy minimization using CHARMM for the clusters in the pockets shown in Table 3 (if such clusters exist in the mapping output). The results from this analysis are shown in Table 5. For both ligand-free and DCE-bound structures, mapping ranked the active site as the lowest free energy cluster. Additionally, after performing the minimization, the active site contained the lowest total free energy of all sites. Primarily, the van der Waals component of the free energy provided the greatest discrimination between the various sites, where the active site was 2.5−3.0 kcal/mol more favorable than all other sites. This difference is approximately the same for the total energy as well. As a result, despite the fact that the active site was not able to accommodate all probe types (Table 2), the site clearly has high affinity for the natural substrate DCE. These results contrast with the mapping of the bromide-bound structure, where the active site was ranked sixth overall, with a higher cluster free energy than that of the tunnel entrance, Br2 and Br3 sites, and the slot region (Table 2). However,

Table 6: Calculated Binding Free Energy (in kcal/mol) of Br^- Bound to Various Sites[a] of the Three DhlA Structures Studied[b]

protein	site	Br^- cluster	ΔG	ΔE_{vdw}	ΔE_{elec}	ΔG_{des}
DhlA	Br2	132	−7.5	−7.0	−0.6	+0.1
cocrystallized with DCE	S1	136	+0.6	−5.0	+8.0	−2.5
(1,2-dichloroethane)	A	139	+2.6	−6.2	+10.6	−1.8
pdb code 2dhc	SL	137	+0.2	−6.6	+17.9	−11.1
	Br3	135	+3.0	−4.3	+9.1	−1.9
DhlA	Br2	99	−8.0	−6.6	−2.8	+1.4
ligand-free structure	S1	105	−0.1	−5.4	+6.1	−0.8
pdb code 2had	A	106	+2.4	−6.1	+13.9	−5.3
	E	100	+0.2	−5.5	+0.3	+5.4
DhlA	Br2	182	−9.4	−6.9	−0.0	−2.4
crystallized in the presence	S1	189	+0.1	−5.5	+4.4	+1.2
of Br^- ions (3 bound	A	190	+1.4	−5.9	+11.4	−4.0
Br^- ions) pdb code 2cij	E	183	−0.3	−5.8	−1.6	+7.0
	SL	187	+1.3	−5.3	+23.6	−16.9
	Br3	188	−5.5	−6.2	+0.9	−0.3

[a] The sites are defined in Table 3. [b] The structures are listed in Table 1.

Table 7: Calculated Binding Free Energy (in kcal/mol) of Cl^- Bound to Various Sites[a] of the Three DhlA Structures Studied[b]

protein	site	Cl^- cluster	ΔG	ΔE_{vdw}	ΔE_{elec}	ΔG_{des}
DhlA	Br2	117	−5.8	−5.9	−0.4	+0.5
cocrystallized with DCE	S1	119	+1.8	−4.9	+8.1	−1.4
(1,2-dichloroethane)	A	122	+3.9	−5.2	+10.4	−1.3
pdb code 2dhc	SL	121	+1.8	−5.	+17.9	−10.6
	Br3	120	−2.1	−5.3	−2.2	+5.4
DhlA	Br2	152	−6.4	−5.6	−3.2	+2.4
ligand-free structure	S1	154	+1.2	−4.9	+6.0	+0.1
pdg code 2had	A	156	+3.5	−5.1	+14.7	−6.0
	E	153	+1.5	−4.4	−0.3	+6.2
DhlA	Br2	235	−5.0	−5.6	+0.4	+0.2
crystallized in the presence	S1	238	+1.5	−4.9	+4.8	+1.7
of Br^- ions (3 bound	A	239	+2.6	−5.0	+11.4	−3.7
Br^- ions) pdg code 2cij	E	236	+1.4	−5.4	+18.2	−11.4
	SL	237	−4.2	−5.3	+0.6	+0.5
	Br3	235	−5.0	−5.6	+0.5	+0.1

[a] The sites are defined in Table 3. [b] The structures are listed in Table 1.

after applying flexible minimization, the active site had the lowest total free energy out of all the analyzed sites, at −20.6 kcal/mol, and had an energy differential of 2.3 kcal/mol or greater versus all remaining sites. Thus, the binding of a bromide ion in the active site slightly alters some of the side chain conformations, thereby preventing DCE from optimally binding to this structure. However, a minor adjustment of the side chains allows this site to bind DCE with high affinity. This rearrangement of active site side chains is in agreement with quantum mechanics calculations (*9*) and molecular dynamics simulations (*13*), which have previously shown that active site residues W125 and W175 reposition themselves upon halide binding.

Binding of Halides to DhlA. As shown in Table 4, the mapping of DhlA structures found two pockets, denoted Br2 and Br3, that are Br^- binding sites in the respective crystal structure. Pikkemaat et al. (*51*) proposed that the binding of bromide to the Br2 site may represent a collision complex intermediate formed during halide import. As mentioned, the Br2 site is close to entrance E of the tunnel, which is likely to provide a direct path from the protein surface to the internal active site. This importance of the Br2 site was further supported by the previously performed mutational and kinetic experiments involving key residues in the Br2 site: T197 and F294 both converted to alanines (*51*). The subsequent conversion from a parabolic increase in the catalytic rate constant to a linear increase upon these mutations supports the assumption that the Br2 site provides an intermediate binding location in the pathway of halide import. The Br2 site also binds a chloride ion, which further supports that the site may serve as the collision complex with halides. All these properties of the Br2 pocket suggest that the site may also play a role in the export of halide from the active site which has been established as the rate-limiting step in the catalytic mechanism for DhlA (*33*).

We mapped the product chloride and bromide ions and subsequently performed a comparative analysis of the binding free energies at each of the aforementioned sites. The results are shown in Tables 6 (for bromide) and 7 (for chloride). In contrast to the mapping results using the substrate (Table 5), where the active site of the enzyme had the most favorable binding free energies with DCE, for the two halide ions the binding free energy at the active site is relatively unfavorable because of the electrostatic repulsion between the ionic species and the active site residues E56 and D124. In fact, the carboxyl functional group of the D124 side chain is pointed directly toward the ion, resulting in a highly unfavorable conformation, and thereby destabilizing the enzyme−ion complex at the active site. Previous quantum mechanical calculations have shown that this repulsion is partially compensated by the rearrangement of halide-stabilizing Trp residues in the active site (*9*, *13*).

For all three DhlA structures, irrespective of which ion was mapped, the lowest binding free energy at the analyzed sites was found to be at the Br2 site. For bromide binding, as shown in Table 6, the calculated binding free energy ΔG at the Br2 site ranged from −9.4 to −7.5 kcal/mol, depending on the structure. Likewise, for chloride binding (Table 7), this value at the Br2 site ranged from −6.4 to −5.0 kcal/mol. This site is the only location where the free energy for ion binding was found to be consistently favorable. The only other site which has been found to bind the ions with favorable ΔG values was the Br3 site. However, binding at Br3 is highly dependent on crystal structure. For example, the free energies of bromide binding at Br3 are −5.5, +0.2, and +3.0 kcal/mol in the bromide-bound, free, and substrate-bound structures, respectively, confirming that the Br3 pocket is induced by bromide binding. This seems to be in reasonable agreement with kinetic experiments that demonstrated that a slowly reacting path was kinetically observed only at high bromide concentration (*51*). In contrast, the Br2 site is able to bind the ion products with favorable binding free energy in all structures.

CONCLUSIONS

We have mapped three DhlA structures for sites that tend to bind small organic molecules: a structure cocrystallized with the substrate DCE, a ligand-free structure, and a structure with three bound Br^- ions. Mapping located seven distinct regions on the protein. These sites include the enzyme's active site A, which overlaps with the Br^- binding site Br1, the putative tunnel entrance E, the crystallographi-

cally determined bromide binding pockets Br2 and Br3, a pocket SL in the slot region, which is likely to be the entrance of the water-access channel, and two additional sites S1 and S2 that occur in both the DCE-bound and ligand-free structures and may serve as additional water-access channels. The importance of these regions was strongly supported by the results obtained by CAVER, a program developed for finding routes from protein clefts and cavities to the surface. We also performed mapping and free energy analysis of DhlA structures using the substrate, DCE, and halide ions as probes. The similarities and differences among the results for the three different DhlA structures provide information on the mechanistic roles of the seven binding sites, particularly their contributions to substrate import and product export.

The comparison of ligand-free, DCE-bound, and Br^- ion-bound DhlA structures show that ligand binding to this protein does not involve large conformational changes. However, computational solvent mapping shows that even these relatively minor changes can substantially alter the ligand-binding properties of various pockets. As emphasized in an earlier report on the method (*43*), one of the major advantages of mapping is that it can be regarded as a molecular microscope, revealing information that is contained in the X-ray structure but difficult to see without a special tool. In fact, some of the results presented here may seem to be unexpected in view of the small differences among the various structures, but such differences are real, and their analysis provides mechanistic insight.

Active Site. As discussed previously, solvent mapping of enzymes generally places the largest consensus site (i.e., the largest number of different probe clusters) in the active site, but this was not the case for DhlA (*24*), where the active site is so small that it can accommodate only the smallest probes. For comparison, we also mapped haloalkane dehalogenases DhaA and LinB, both of which contain significantly larger and more solvent-accessible active sites, and found the largest consensus site in the active-site pocket as expected. This result confirms that the unusual behavior of the DhlA active site is due to its limited accessibility. In particular, in the bromide-bound structure the active site binds only urea and acetonitrile. The most likely explanation is that bromide binding changes some of the side-chain conformations in the site (*9*, *13*). However, mapping with DCE and calculating the free energy of binding show that the active site is the most important binding pocket as far as the substrate is concerned and that the substrate has the lowest free energy at this site in all structures.

Halide Ion Export from the Active Site. As shown by halide mapping of DhlA and by free energy analysis, the Br2 site is the most likely intermediate binding location for the halide product upon export from the active site. In all three DhlA structures, irrespective of whether chloride or bromide was used in the analysis, the binding free energy is at least 2−3 kcal/mol more favorable at Br2 than at any other site, including both the active site and the Br3 site. Because Br2 is approximately only 5 Å away from the tunnel entrance, it is conceivable that Br2 plays a distinct role in halide export via the substrate-access tunnel. Indeed, Pikkemaat et al. (*51*) proposed that the binding of bromide to the Br2 site may represent a collision complex formed during halide import. The Br3 site can also be part of halide export. The site is well defined in the bromide bound structure, but in the other two structures, it binds fewer probes and shows diminished affinity for Br^-, suggesting that this pocket is formed in the process of Br^- binding. This is in good agreement with prior kinetic data implying a slower auxiliary route, induced exclusively in the presence of high Br^- concentration.

Mapping and free energy analysis show that the binding of halides within the binding site is unstable. Nevertheless, multiple DhlA structures (such as the bromide-bound structure used in this study) exist where an ion has been found within the active site, and it has been reported that the release of the halide ion is the rate-limiting step. The existence of a long residence time bromide or chloride ion in the negatively charged active site pocket suggests that the ion is kinetically trapped within the site and cannot leave without significant conformational change. This conclusion is in good agreement with the predictions that halide release requires some conformational change (*51*).

Water-Access Channels. Water molecules are needed both for the reaction and the solvation of the halide ions before their release from the active site. The site SL in the slot region has been reported as a putative water-access channel (*17*). Because the SL site binds the largest number of probe clusters in the bromide-bound structure, the mapping results support this assumption. The sites S1 and S2 could also serve as entrance sites for water channels. Both sites are at the perimeter of the interface between the two domains of DhlA, and each is separated from the active site essentially by a single side chain, S1 by M225 and S2 by V226. In addition, the mutation V226A increases the reaction rate (*52*). Analysis using CAVER has demonstrated the presence of well-defined paths leading from the active site to both S1 and S2 in multiple DhlA structures (*51*). Thus, it is conceivable that these channels may be used for water access. Because neither S1 nor S2 are present in the bromide-bound structure, solvation through these channels is likely to occur at earlier stages of the reaction cycle.

Summary of the Reaction Cycle. Because the active site in DhlA is located almost 10 Å from the protein surface, the reaction is accompanied by a complex sequence of binding and unbinding events. Here, we have studied both the sites that exhibit high affinity for the binding of small ligands and the potential channels connecting the active site to the protein surface. According to mapping results, putative entrance E of the substrate-access channel is already well formed in the apostructure and has high affinity for most probes. However, the lowest free energy clusters of the smallest probes are directly formed in the active site, indicating a very short residence time at site E for small substrates. Substrates such as DCE quickly transfer to the active site along a strong negative gradient of the binding free energy. The negative gradient is due to electrostatic and van der Waals interactions that are both very favorable for substrate binding at the active site.

The reaction mechanism of DhlA after substrate binding has been extensively studied and is well understood (*5*−*8*). The catalytic cycle begins with the formation of a Michaelis complex, followed by the S_N2 displacement of chloride by the nucleophilic attack of Asp124 to form a covalent alkyl-enzyme intermediate (*8*). During the reaction, the halide moiety is stabilized by the indole NH groups of Trp125 and Trp175. The next steps of the mechanism require the

presence of water molecules. Both mapping and CAVER analysis show three different pathways that can serve as water-access channels to the active sites. On the basis of mapping, all three corresponding surface pockets are present in the ligand-free structure. When the substrate is bound, site S1 on the opposite end of the substrate-access channel shows the highest affinity for binding, and hence, it appears to be the most likely water-access channel at this stage of the reaction. Because both S1 and S2 pockets are separated from the active site by the side chains of M225 and V226, respectively, actual water access requires conformational changes. These changes may be restricted to the particular side chains but, on the basis of short molecular dynamics simulations, are more likely to involve loop regions or even some interdomain motion (*17*).

The side chain of His289 activates water to hydrolyze the covalently bound intermediate and to form the alcohol product, which is rapidly released (*7*). The halide ion remains bound between the side chains of Trp125 and Trp175. According to the energy calculations, these interactions rearrange the active site geometry, but the binding of the halide at the active site is unfavorable. However, a number of X-ray structures show bound halides, and on the basis of kinetic studies, halide release is the rate-limiting step of the entire reaction (*33*). As mentioned, this contradiction suggests a kinetically trapped halide in the active site and the need for conformational changes to permit halide release (*56*). Because the active site is in the interface of two domains, the conformational change may involve the repositioning of the domains relative to each other (*17*). Indeed, molecular dynamics simulations as short as 1 ns show functionally relevant motions spread over the entire cap domain in DhlA (*17*). The simulations also revealed that the most flexible part of the haloalkane dehalogenases is the random coil interconnecting the two domains, suggesting that substantial domain−domain motions may occur at somewhat longer time scales. Because these enzymes perform catalytic cycles on the time scale of a second, such slow, large scale motions can be functionally important and are indicated by kinetic methods (*33*, *56*).

In addition to the conformational change, halide release requires the solvation of the ion. In the halide-bound structure, the only water-access channel with a well-defined entrance site is SL in the slot region (Table 4). As discussed, the leaving halide most likely binds to the Br2 site that is approximately 5 Å away from the tunnel entrance and shows highest affinity for the ion. Because the Br2 site is partially exposed, and the desolvation contribution ΔG_{des} to the free energy is unfavorable, the halide is then released to the solvent. The bromide bound crystal structure indicates that a second site (Br3) also can bind ions. However, on the basis of the mapping results, the Br3 pocket is formed only in the process of Br^- binding. This is in good agreement with prior kinetic data implying a slower, auxiliary route induced exclusively in the presence of high Br^- concentration.

Buried Active Site Makes DhlA Unique among Haloalkane Dehalogenases. As discussed above, ligand access to the DhlA acive site requires elaborate tunnels and intermediate pockets (*24*), in contrast to the significantly larger active sites of DhaA and LinB. The discrepancy in mapping results for DhlA versus DhaA and LinB at these sites indicates that a well-defined, open active site in the latter two structures substantially simplifies the mechanism of substrate binding and product release.

REFERENCES

1. van Agteren, M. H., Keuning, S., and Janssen, D. B. (1998) *Biodegradation and Biological Treatment of Hazardous Organic Compounds*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
2. Copley, S. D. (1998) Microbial dehalogenases: enzymes recruited to convert xenobiotic substrates, *Curr. Opin. Chem. Biol. 2*, 613−617.
3. Schanstra, J. P., Kingma, J., and Janssen, D. B. (1996) Specificity and kinetics of haloalkane dehalogenase, *J. Biol. Chem. 271*, 14747−14753.
4. de Jong, R. M., and Dijkstra, B. W. (2003) Structure and mechanism of bacterial dehalogenases: different ways to cleave a carbon-halogen bond, *Curr. Opin. Struct. Biol. 13*, 722−730.
5. Verschueren, K. H. G., Seljee, F., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1993) Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase, *Nature 363*, 693−698.
6. Pries, F., Kingma, J., Pentenga, M., van Pouderoyen, G., Jeronimus-Stratingh, C. M., Bruins, A. P., and Janssen, D. B. (1994) Site-directed mutagenesis and oxygen isotope incorporation studies of the nucleophilic aspartate of haloalkane dehalogenase, *Biochemistry 33*, 1242−1247.
7. Pries, F., Kingma, J., Krooshof, G. H., Jeronimus-Stratingh, C. M., Bruins, A. P., and Janssen, D. B. (1995) Histidine 289 is essential for hydrolysis of the alkyl-enzyme intermediate of haloalkane dehalogenase, *J. Biol. Chem. 270*, 10405−10411.
8. Pries, F., Kingma, J., and Janssen, D. B. (1995) Activation of an Asp-124→Asn mutant of haloalkane dehalogenase by hydrolytic deamidation of asparagines, *FEBS Lett. 358*, 171−174.
9. Damborsky, J., Kuty, M., Nemec, M., and Koca, J. (1997) A molecular modeling study of the catalytic mechanism of haloalkane dehalogenase: 1. quantum chemical study of the first reaction step, *J. Chem. Inf. Comput. Sci. 37*, 562−568.
10. Damborsky, J., Kuty, M., Nemec, M., and Koca, J. (12997) Molecular Modelling to Understand the Mechanisms of Microbial Degradation: Application to Hydrolytic Dehalogenation with Haloalkane Dehalogenases, in *Quantitative Structure-Activity Relationships in Environmental Sciences, VII* (Chen, F. and Schüürmann, G., Eds.) pp 5−20, SETAC Press, Pensacola, FL.
11. Maulitz, A. H., Lightstone, F. C., Zheng, Y. J., and Bruice, T. C. (1997) Nonenzymatic and enzymatic hydrolysis of alkyl halides: a theoretical study of the S(N)2 reactions of acetate and hydroxide ions with alkyl chlorides, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6591−6595.
12. Damborsky, J., Bohac, M., Prokop, M., Kuty, M., and Koca, J. (1998) Computational site-directed mutagenesis of haloalkane dehalogenase in position 172, *Protein Eng. 11*, 901−907.
13. Lightstone, F. C., Zheng, Y. J., and Bruice, T. C. (1998) Molecular dynamics simulations of ground and transition states for the S(N)2 displacement of Cl^- from 1,2-dichloroethane at the active site of *Xanthobacter autotrophicus* haloalkane dehalogenase, *J. Am. Chem. Soc. 120*, 5611−5621.
14. Lau, E. Y., Kahn, K., Bash, P. A., and Bruice, T. C. (2000) The importance of reactant positioning in enzyme catalysis: a hybrid quantum mechanics/molecular mechanics study of a haloalkane dehalogenase, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9937−9942.
15. Kmunicek, J., Luengo, S., Gago, F., Ortiz, A. R., Wade, R. C., and Damborsky, J. (2001) Comparative binding energy analysis of the substrate specificity of haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10, *Biochemistry 40*, 8905−8917.
16. Bohac, M., Nagata, Y., Prokop, Z., Prokop, M., Monincova, M., Koca, J., Tsuda, M., and Damborsky, J. (2002) Halide-stabilizing residues of haloalkane dehalogenases studied by quantum mechanic calculations and site-directed mutagenesis, *Biochemistry 41*, 14272−14280.
17. Otyepka, M., and Damborsky, J. (2002) Functionally relevant motions of haloalkane dehalogenases occur in the specificity-modulating cap domains, *Protein Sci. 11*, 1206−1217.
18. Shurki, A., Strajbl, M., Villa, J., and Warshel, A. (2002) How much do enzymes really gain by restraining their reacting fragments? *J. Am. Chem. Soc. 124*, 4097−4107.

19. Damborsky, J., Kmunicek, J., Jedlicka, T., Luengo, S., Gago, F., Ortiz, A. R., and Wade, R. C. (2003) Rational Re-Design of Haloalkane Dehalogenases Guided by Comparative Binding Energy Analysis, in *Enzyme Functionality: Design, Engineering and Screening* (Svendsen, A., Ed.) pp 79−96, Marcel Dekker, New York.
20. Devi-Kesavan, L. S., and Gao, J. (2003) Combined QM/MM study of the mechanism and kinetic isotope effect of the nucleophilic substitution reaction in haloalkane dehalogenase, *J. Am. Chem. Soc. 125*, 1532−1540.
21. Hur, S., Kahn, K., and Bruice, T. C. (2003) Comparison of formation of reactive conformers for the SN2 displacements by $CH_3CO_2^-$ in water and by Asp124-CO_2^- in a haloalkane dehalogenase, *Proc. Natl. Acad. Sci. U.S.A. 100*, 2215−2219.
22. Kahn, K., and Bruice, T. C. (2003) Comparison of reaction energetics and leaving group interactions during the enzyme-catalyzed and uncatalyzed displacement of chloride from haloalkanes, *J. Phys. Chem. 107*, 6876−6885.
23. Kmunicek, J., Bohac, M., Luengo, S., Gago, F., Wade, R. C., and Damborsky, J. (2003) Comparative binding energy analysis of haloalkane dehalogenase substrates: modelling of enzyme-substrate complexes by molecular docking and quantum mechanic calculations, *J. Comput.-Aided Mol. Des. 17*, 299−311.
24. Silberstein, M., Dennis, S., Brown, III, L., Kortvelyesi, T., Clodfelter, K., and Vajda, S. (2003) Identification of substrate binding sites in enzymes by computational solvent mapping, *J. Mol. Biol. 332*, 1095−1113.
25. Soriano, A., Silla, E., and Tunon, I. (2003) Internal rotation of 1,2-dichloroethane in haloalkane dehalogenase. A test case for analyzing electrostatic effects in enzymes, *J. Phys. Chem. 107*, 6234−6238.
26. Nam, K., Prat-Resina, X., Garcia-Viloca, M., Devi-Kesavan, L. S., and Gao, J. (2004) Dynamics of an enzymatic substitution reaction in haloalkanen dehalogenase, *J. Am. Chem. Soc. 126*, 1369−1376.
27. Olsson, M. H., and Warshel, A. (2004) Solute solvent dynamics and energetics in enzyme catalysis: the SN2 reaction of dehalogenase as a general benchmark, *J. Am. Chem. Soc. 126*, 15167−15179.
28. Kmunicek, J., Hynkova, K., Jedlicka, T., Nagata, Y., Negri, A., Gago, F., Wade, R. C., and Damborsky, J. (2005) Quantitative analysis of substrate specificity of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26, *Biochemistry 44*, 3390−3401.
29. Soriano, A., Silla, E., Tunon, I., and Ruiz-Lopez, M. F. (2005) Dynamic and electrostatic effects in enzymatic processes. An analysis of the nucleophilic substitution reaction in haloalkane dehalogenase, *J. Am. Chem. Soc. 127*, 1946−1957.
30. Ridder, I. S., Rozeboom, H. J., and Dijkstra, B. W. (1999) Haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10 refined at 1.15 Å resolution, *Acta Crystallogr. 55*, 1273−1290.
31. Damborsky, J., Rorije, E., Jesenska, A., Nagata, Y., Klopman, G., and Peijnenburg, W. J. (2001) Structure-specificity relationships for haloalkane dehalogenases, *Environ. Toxicol. Chem. 20*, 2681−2689.
32. Verschueren, K. H., Franken, S. M., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1993) Refined X-ray structures of haloalkane dehalogenase at pH 6.2 and pH 8.2 and implications for the reaction mechanism, *J. Mol. Biol. 232*, 856−872.
33. Schanstra, J. P., and Janssen, D. B. (1996) Kinetics of halide release of haloalkane dehalogenase: evidence for a slow conformational change, *Biochemistry 35*, 5624−5632.
34. Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J. F., Unkefer, C. J., and Terwilliger, T. C. (1999) Haloalkane dehalogenase: structure of a *Rhodococcus enzyme*, *Biochemistry 138*, 16105−16114.
35. Marek, J., Vevodova, J., Kuta-Smatanova, I., Nagata, Y., Svensson, L. A., Newman, J., Takagi, M., and Damborsky, J. (2000) Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26, *Biochemistry 39*, 14082−14086.
36. Nagata, Y., Miyauchi, K., Damborsky, J., Manova, K., Ansorgova, A., and Takagi, M. (1997) Purification and characterization of a haloalkane dehalogenase of a new substrate class from a gamma-hexachlorocyclohexane-degrading bacterium, *Sphingomonas paucimobilis* UT26, *Appl. Environ. Microbiol. 63*, 3707−3710.
37. Damborsky, J., and Koca, J. (1999) Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons, *Protein Eng. 12*, 989−998.
38. Dennis, S., Kortvelyesi, T., and Vajda, S. (2002) Computational mapping identifies the binding sites of organic solvents on proteins, *Proc. Natl. Acad. Sci. U.S.A. 99*, 4290−4295.
39. Kortvelyesi, T., Dennis, S., Silberstein, M., Brown III, L., and Vajda, S. (2003) Algorithms for computational solvent mapping of proteins, *Proteins 51*, 340−351.
40. Mattos, C., and Ringe, D. (1996) Locating and characterizing binding sites on proteins, *Nat. Biotechnol. 14*, 595−599.
41. Mattos, C., and Ringe, D. (2001) Proteins in organic solvents, *Curr. Opin. Struct. Biol. 11*, 761−764.
42. Mattos, C., Bellamacina, C. R., Peisach, E., Pereira, A., Vitkup, D., Petsko, G. A., and Ringe, D. (2005) Multiple solvent crystal structures: probing binding sites, plasticity, and hydration, *J. Mol. Biol. 357*, 1471−1482.
43. Sheu, S., Kaya, T., Waxman, D. J., and Vajda, S. (2005) Exploring the binding site structure of the PPARγ ligand binding domain by computational solvent mapping, *Biochemistry 44*, 1193−1209.
44. Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., and Damborsky, J. (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities, *BMC Bioinf. 7*, 316−324.
45. Vakser, I. A., Matar, O. G., and Lam, C. F. (1999) A systematic study of low-resolution recognition in protein-protein complexes, *Proc. Natl. Acad. Sci. U.S.A. 96*, 8477−8482.
46. Vakser, I. A. (1995) Protein docking for low-resolution structures, *Protein Eng. 8*, 371−377.
47. Schaefer, M., and Karplus, M. A. (1996) Comprehensive analytical treatment of continuum electrostatics, *J. Phys. Chem. 100*, 1578−1599.
48. Brooks, B. R., Bruccoleri, R. E., Olafson, B., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 197−214.
49. Gschwend, D. A., Good, A. C., and Kuntz, I. D. (1996) Molecular docking towards drug discovery, *J. Mol. Recognit. 9*, 175−816.
50. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, P., Profeta, Jr., S., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *J. Am. Chem. Soc. 106*, 765−784.
51. Pikkemaat, M. G., Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W., and Janssen, D. B. (1999) Crystallographic and kinetic evidence of a collision complex formed during halide import in haloalkane dehalogenase, *Biochemistry 38*, 12052−12061.
52. Schanstra, J. P., Ridder, A., Kingma, J., and Janssen, D. B. (1997) Influence of mutations of Val226 on the catalytic rate of haloalkane dehalogenase, *Protein Eng. 10*, 53−61.
53. Oakley, A. J., Prokop, Z., Bohac, M., Kmunicek, J., Jedlicka, T., Monincova, M., Kuta-Smatanova, I., Nagata, Y., Damborsky, J., and Wilce, M. C. J. (2002) Exploring the structure and activity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26: Evidence for product- and water-mediated inhibition, *Biochemistry 41*, 4847−4855.
54. Oakley, A. J., Klvana, M., Otyepka, M., Nagata, Y., Wilce, M. C. J., and Damborsky, J. (2004) Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 Å resolution: dynamics of catalytic residues, *Biochemistry 43*, 870−878.
55. Streltsov, V. A., Prokop, Z., Damborsky, J., Nagata, Y., Oakley, A., and Wilce, M. C. J. (2003) Haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26: X-ray crystallographic studies of dehalogenation of brominated substrates, *Biochemistry 42*, 10104−10112.
56. Krooshof, G. H., Floris, R., Tepper, A. W., and Janssen, D. B. (1999) Thermodynamic analysis of halide binding to haloalkane dehalogenase suggests the occurrence of large conformational changes, *Protein Sci. 8*, 355−360.

BI700336Y