JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

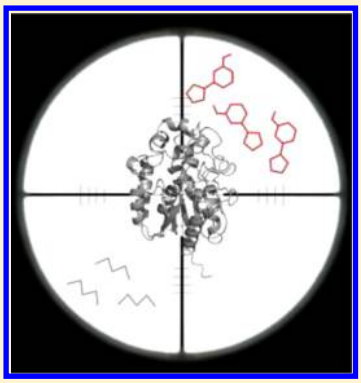
Mechanism-Based Discovery of Novel Substrates of Haloalkane Dehalogenases Using *in Silico* Screening

Lukas Daniel, Tomas Buryska, Zbynek Prokop, Jiri Damborsky, and Jan Brezovsky*

Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic

Supporting Information

ABSTRACT: Substrate specificity is a key feature of enzymes determining their applicability in biomaterials and biotechnologies. Experimental testing of activities with novel substrates is a time-consuming and inefficient process, typically resulting in many failures. Here, we present an experimentally validated *in silico* method for the discovery of novel substrates of enzymes with a known reaction mechanism. The method was developed for a model system of biotechnologically relevant enzymes, haloalkane dehalogenases. On the basis of the parametrization of six different haloalkane dehalogenases with 30 halogenated substrates, mechanism-based geometric criteria for reactivity approximation were defined. These criteria were subsequently applied to the previously experimentally uncharacterized haloalkane dehalogenase DmmA. The enzyme was computationally screened against 41,366 compounds, yielding 548 structurally unique compounds as potential substrates. Eight out of 16 experimentally tested top-ranking compounds were active with DmmA, indicating a 50% success rate for the prediction of substrates. The remaining eight compounds were able to bind to the active site and inhibit enzymatic activity. These results confirmed good applicability of the method for prioritizing active compounds—true substrates and binders—for experimental testing. All validated substrates were large compounds often containing polyaromatic moieties, which have never before been considered as potential substrates for this enzyme family. Whereas four of these novel substrates were specific to DmmA, two substrates showed activity with three other tested haloalkane dehalogenases, i.e., DhaA, DbjA, and LinB. Additional validation of the developed screening strategy with the data set of over 200 known substrates of *Candida antarctica* lipase B confirmed its applicability for the identification of novel substrates of other biotechnologically relevant enzymes with an available tertiary structure and known reaction mechanism.



INTRODUCTION

Enzymatic catalysis has matured into an important tool for various biotechnological applications. The high enantioselectivity and specificity of enzymes allow efficient production of many valuable chemicals.[1,2] However, engineering a successful catalyst remains a challenging task and often requires the construction of many variants to achieve the desired activity and stability.[2,3] Identification of conversions of novel substrates by natural enzymes is an important way for discovering practically useful reactions. *In silico* screening by molecular docking has been used to complement experimental *in vitro* screening of chemicals against biological targets in drug discovery.[4,5] Recently, *in silico* screening was also employed for functional assignment of enzymes with unknown function by prediction of their putative substrates.[6−8] This approach can be further extended for the discovery of novel substrates of enzymes used in biotechnology to broaden their scope. Although molecular docking has been used to predict substrates of short-chain dehydrogenase,[9] laccase,[10] lipase B,[11] and sulfotransferase,[12] those studies largely focused on small libraries of similar compounds, employed basic molecular docking for consideration of substrate binding only without addressing the reactivity, and sometimes lacked experimental validation.[11]

In a comprehensive study, Irwin et al.[13] employed molecular docking to identify new substrates of Zn-dependent phosphothioesterase, an enzyme hydrolyzing phosphate esters, including the insecticide paraoxon and nerve gases sarin, soman, and VX. The screening of 167,000 compounds revealed a known substrate ranked as the eighth hit. After visual inspection, seven compounds resembling known substrates were found among the 100 top-ranked hits and were subsequently experimentally verified as substrates. The authors did not explicitly address the reactivity of docked compounds. Reactivity was considered by docking substrates in high-energy forms in a study by Hermann et al.,[6] who docked 4,207 compounds and experimentally verified three of them as substrates. Xu et al.[11] addressed reactivity by considering distance-based screening criteria based on the sum of van der Waals radii of atoms involved in the nucleophilic attack. The authors correctly identified compounds reported in the literature, but no external validation was carried out. Moreover, the essential hydrogen bonds were kept

Received: August 6, 2014
Published: December 14, 2014

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

fixed during the docking procedure, which may have disfavored compounds showing different binding patterns.

Here, we present an *in silico* screening method calibrated on experimentally verified substrates for better prediction accuracy. We focused on biotechnologically interesting enzymes from the family of haloalkane dehalogenases (EC 3.8.1.5; HLDs) as model enzymes. HLDs are microbial enzymes structurally belonging to the α/β-hydrolase fold that catalyze the degradation of halogenated aliphatic hydrocarbons to the corresponding alcohol, a halide and a proton. They can be utilized in the bioremediation of environmental pollutants,[14] construction of biosensors,[15] decontamination of warfare agents,[16] synthesis of optically pure alcohols,[17] and imaging of cells and protein analysis.[18−20] The reaction mechanism of HLDs has been widely studied and is well understood.[21−24] Consequently, HLDs are often employed as benchmarks for testing of various molecular modeling protocols.[25−29] Dehalogenation by HLDs follows a two-step reaction mechanism involving a covalently bound ester intermediate. The ester is formed in the first step by S_N2 nucleophilic displacement of a halogen atom mediated by the catalytic triad. The transition state formed in this step is stabilized by two halide-stabilizing residues, which donate H-bonds to the leaving halide group to saturate the developing charge. The ester intermediate is then hydrolyzed in the second reaction step by a hydroxyl group, which is generated from water by the catalytic histidine. Eventually, an alcohol product and halide ion are released from the buried active site through the access tunnels.

In this study, the newly developed *in silico* method was tested by screening 41,366 halogenated compounds against the previously uncharacterized enzyme DmmA, revealing eight experimentally validated novel substrates and eight inhibitors. In the case of the top-ranking compounds, the success rate of the method was 50% for the prediction of catalysis and 100% for the prediction of binding. We note that these rates will most probably decline with an increasing number of tested compounds. The broader applicability of the method was further validated with 206 experimentally tested substrates of *Candida antarctica* lipase B (CALB). The method could be used to enrich the pool of converted substrates for other biotechnologically relevant enzymes with known tertiary structure and reaction mechanism.

METHODS

Preparation of Experimentally Verified Substrates. Thirty known substrates of HLDs[30] were prepared in both ground and high-energy states. The tertiary structures in the ground state were taken from our in house database.[31] For the preparation of high-energy states of substrates, the Antechamber module of AMBER 12[32] was employed to convert the PDB files to the z-matrix format compatible with the Gaussian 09 program revision D.01.[33] The structures were then minimized using the HF/6-31G(d) wave function for all atoms except iodine, for which the LANL2DZ basis set[34−36] was employed. The following variables were constrained during the minimization: (i) The length of the carbon−halogen bond increased by 0.5 Å in comparison to its respective ground state length. (ii) Four angles between the halogen, carbon, and three atoms bound to the carbon at 90°. Restrained electrostatic potential partial charges[37] were derived using the Antechamber module. In cases where multiple halogen atoms were present in the structure, the preparation of high-energy states was performed individually for each halogen atom. Input files were converted to an AutoDock4.0 compliant format by MGLTools.[38]

Preparation of Screening Database. The three-dimensional structures of 41,366 compounds were extracted from the EDULISS database.[39] The criteria for selection of these compounds were as follows: presence of a halogen atom at a sp^3 carbon atom, M log P < 8, molecular weight < 500 g mol^{-1}, atom types parametrized for AutoDock4.0, and less than 15 torsional degrees of freedom. Input files in the Sybyl mol2 format were converted into an AutoDock4.0 compliant format by MGLTools.

Preparation of Receptors. The coordinates of the crystal structures of seven HLDs were downloaded from the Protein Data Bank[40] (Table 1). Only chain A was used when screening proteins with multiple chains. All water molecules were deleted from the structures with the exception of "structured water", that is, molecules important for the catalytic mechanism of HLDs (Table 1). Hydrogen atoms were added to the receptor with the H++ server[41] at pH 7.5. The hydrogen atom bound to the NE2 atom of the catalytic histidine was deleted to reflect the catalytic mechanism of HLDs. Gasteiger charges were subsequently assigned by MGLTools. During the docking procedure, the receptor was represented by a set of atomic and electrostatic maps calculated by AutoGrid4.0.[42] The grid maps comprised 80 × 76 × 80 grid points with a spacing of 0.25 Å, which were centered at a position near the nucleophilic oxygen and catalytic base to cover the whole active site and main access tunnel (Figure 1, Supporting Information).

Table 1. PDB Codes of Crystal Structures of HLDs and IDs of "Structured Water" Molecules[a]

crystal structure	source	PDB code	ID of "structured water" molecules
DbjA	*Bradyrhizobium japonicum* USDA110	3A2M	544
DmbA	*Mycobacterium tuberculosis* H37Rv	2QVB	316, 341
LinB	*Sphingomonas paucimobilis* UT26	1MJ5	3050, 3508
DhaA	*Rhodococcus rhodochrous* NCIMB 13064	1CQW	565, 566
DhlA	*Xanthobacter autotrophicus* GJ10	2YXP	559, 610
DbeA	*Bradyrhizobium elkanii* USDA94	4K2A	NA
DmmA	metagenome of marine microbial consortium	3U1T	26, 622

[a]NA − not applicable, structured catalytic waters are missing, and the coordinates of water molecules were taken from LinB.

Docking Protocol. The energy of the unbound system was estimated as the internal energy of the unbound extended conformation determined from a Lamarckian Genetic Algorithm search. A total of 250 docking calculations were performed for each compound using the Lamarckian Genetic Algorithm with the following parameters: initial population size 300, maximum number of generations 27,000, elitism value 1, mutation rate 0.02, and crossover rate 0.8. The maximum number of energy evaluations was set to $(1.5 \times 10^5\ \mathrm{Ntor}^2) + 1.5 \times 10^6$, where Ntor is the number of torsional degrees of freedom of a docked compound. Each local search was based on the pseudo Solis and Wets algorithm with a maximum of 300 iterations per search.[43] Final conformations from all 250

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

docking calculations were clustered into groups with tolerance for their root-mean-square positional deviation at 2 Å.

Mechanism-Based Geometric Criteria for Reactivity Estimation. The largest cluster obtained from molecular docking of each compound was evaluated for proper positioning of the halogen atom between the two halogen-stabilizing residues using a 4.5 Å cutoff distance. Compounds with their halogens properly stabilized were investigated for a possible S_N2 displacement reaction. The —COO$^-$···C—Cl reactive distance (R_D) and —COO$^-$···C···Cl reactive angle (R_A) parameters were measured for a set of 30 known substrates. Six characterized HLDs were recorded, and values applicable to 70% of the known substrates were set as the cutoff for the description of reactivity (Figure 1).

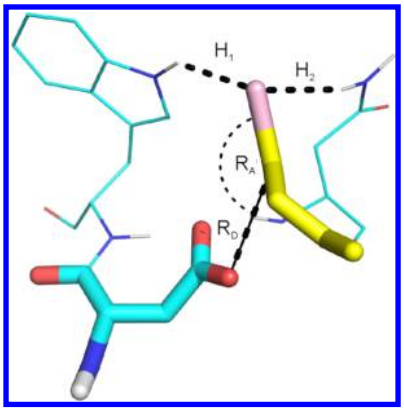


Figure 1. Geometric parameters used for estimation of reactivity. The reactive distance (R_D) between the catalytic aspartate (cyan sticks) and attacked carbon atom of the substrate molecule (yellow sticks) is represented by a thick dashed line. The reactive angle (R_A) between the catalytic aspartate, attacked carbon, and leaving halogen atoms is represented by a thin dashed line. The two hydrogen-bonding distances (H_1 and H_2) between the halide-stabilizing residues (cyan lines) and the leaving halogen atom of the substrate are represented by dashed lines.

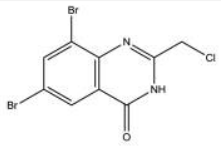
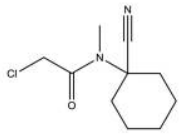
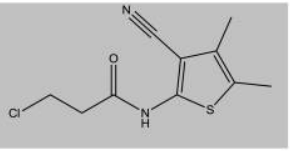
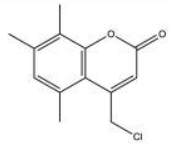
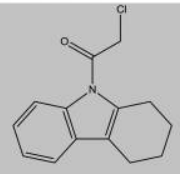
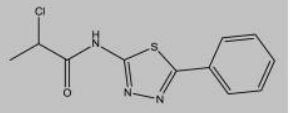
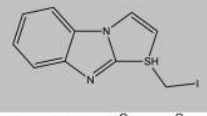
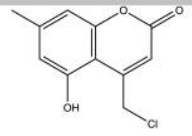
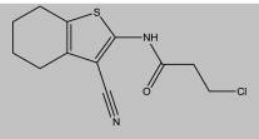
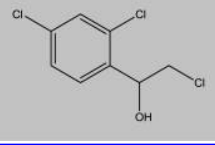
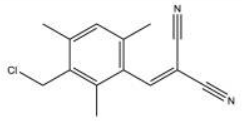
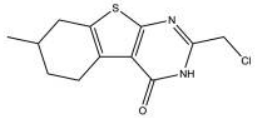
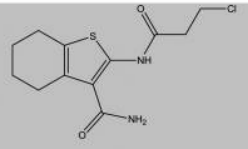
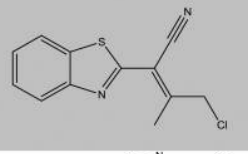
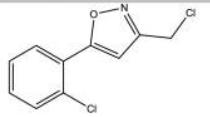
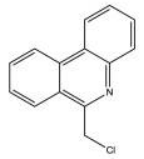
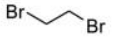
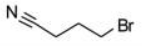
Rescoring of Binding Affinity. The scoring functions employed during the virtual screening are generally optimized for speed at the expense of accuracy. Therefore, the application of a more sophisticated scoring scheme is often advantageous to evaluate binding of docked compounds. The binding affinity of the docked compounds was therefore rescored using two principally different methods: (i) predicting the compound's binding affinity using the neural network-based scoring function NNScore 2.0 tool[44] and (ii) calculation of the free energy change between a bound and free state of the protein by the molecular mechanics/generalized Born surface area (MM/GBSA) method. In the last step, a consensus rank for each docked compound was calculated by averaging the data obtained by MM/GBSA and NNScore 2.0.

NNScore 2.0 is based on the scores obtained from the 20 best-performing neural networks that were developed and validated by authors of this tool.[44] To predict the binding affinity of the docked compound, the average value of the normalized scores from these 20 networks was employed for each compound according to the established protocol.[44] The binding free energies were calculated by combining the gas phase energies of the protein, compounds, and protein—compound complexes with their respective solvation free energies (both polar and nonpolar) calculated with an implicit solvent model. Force field parameters for the docked compounds were prepared with the Antechamber and prmchk modules of AmberTools 1.5.[32] AM1-BCC charges[45] were assigned to individual atoms of compounds using the Antechamber module of AmberTools 1.5. Molecular topologies of protein, compounds, and protein—compound complexes were prepared with the Leap module of AmberTools 1.5 using the ff03.r1 force field for proteins[46,47] and general Amber force field[48] for compounds. The PB radii were set to mbondi2 as required for generalized Born model 2.[49] To partially account for an induced fit upon the binding of a compound, the structures of the protein—compound complexes were minimized in two consequent rounds using the Sander module of AMBER 12. In the first round, the fast minimization consisted of 250 steps of the steepest descent method followed by 750 steps of the conjugate gradient energy method using a nonbonded cutoff of 50 Å and a distance-dependent dielectric multiplicative constant for the electrostatic interactions of 4rij to simulate solvation effects. The second round was carried out in actual implicit solvent with the following parameters: 100 steps of steepest descent followed by 400 steps of conjugate gradient energy minimization, nonbonded cutoff of 16 Å, interior dielectric constant of 2,[50] and generalized Born model 2.[49,51] The convergence criterion for the energy gradient was set to 0.1 kcal/mol Å for both minimization rounds. The nonpolar contribution to the solvation energy was computed using atomic surfaces calculated from the linear combinations of pairwise overlaps (LCPO) method.[52] Finally, the MM/GBSA calculation of the binding energy was performed on the minimized structure by Python script MMPBSA.py of AmberTools 1.5 using a setting consistent with the second level of minimization.

Protein Expression and Purification. The recombinant plasmid pET21 carrying corresponding His-tagged dehalogenase gene was introduced into *E. coli* BL21 DE3 (Zymo Research, Irvine, CA, U.S.A.) cells by heat shock transformation. Transformed cells were cultivated on agar plates containing 100 μg mL^{-1} of kanamycin as a selection marker. Plates were first incubated overnight at 37 °C. The next day, 10 mL of LB media containing 100 μg mL^{-1} of kanamycin was inoculated with a single colony of cells followed by overnight culture cultivation at 37 °C. Finally, the latter overnight culture was used to inoculate 1 L of LB medium with 100 μg mL^{-1} of kanamycin, which was cultivated at 37 °C until an OD_{600} of 0.6 was reached. Induction of protein expression was initiated by addition of IPTG to a final concentration of 0.5 mM and then cultivating the culture overnight at 20 °C. Cells were harvested by centrifugation at 14,000 *g* for 10 min at 4 °C, then washed with 20 mM phosphate buffer, resuspended in 20 mL of the same buffer, and frozen to −80 °C. Biomass was thawed on ice, whereupon DNase (2 μL of DNase for each 1 mL of sample) was added to the sample. Cells were disrupted by sonication on ice with a Hielscher UP200S ultrasonic processor (Hielscher, Germany). The cell lysate was subjected to centrifugation at 21,000 *g* for 1 h at 4 °C. Crude extracts were manually loaded onto a Ni-NTA Superflow cartridge (Qiagen, Germany) charged with Ni^{2+} ions and equilibrated with a purification buffer (16.4 mM K_2HPO_4, 3.6 mM KH_2PO_4, 0.5 M NaCl, and 10 mM imidazole) at pH 7.5. After washing out unbound and weakly bound fractions, histidine-tagged protein was eluted by increasing the imidazole concentration up to 250 mM. The collected protein was dialyzed against 50 mM phosphate buffer (pH 7.5). The homogeneity and purity of the prepared enzyme

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

Table 2. Structural Representation of Experimentally Tested Compounds[a]

Compound ID	Molecular Structure
C01	
C02	
C03	
C04	
C05	
C06	
C07	
C08	
C09	
C10	
C11	
C12	
C13	
C14	
C15	
C16	
BRB	
IOP	
IOB	
DBE	
BBN	

[a]All compounds without exception were experimentally confirmed as either substrates (unshaded) or inhibitors (shaded). Five universal substrates of HLDs used as positive controls are shown at the end of the table.

was verified by SDS-PAGE on 15% polyacrylamide gels. Staining was performed by Coomassie brilliant blue R-250 dye (Fluka, Switzerland), and the approximate molecular weight was determined on the basis of the protein molecular weight marker (Fermentas, Canada).

Activity and Inhibition Assay. Stock solutions were prepared by dissolving 1 mg of powder in 200 μL of 99% DMSO, giving a concentration of 5 mg mL^{-1}. The reaction progress was monitored by using a modified Holloway assay.[53] Briefly, phenol red serving as a pH indicator was added to 1.0 mM HEPES buffer (pH 8.0) to a final concentration of 60 μM. The reaction mixture was prepared in a transparent 96 well microtiter plate, which contained 120 μL of buffer with indicator, 15 μL of stock compound in DMSO, and 15 μL of

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

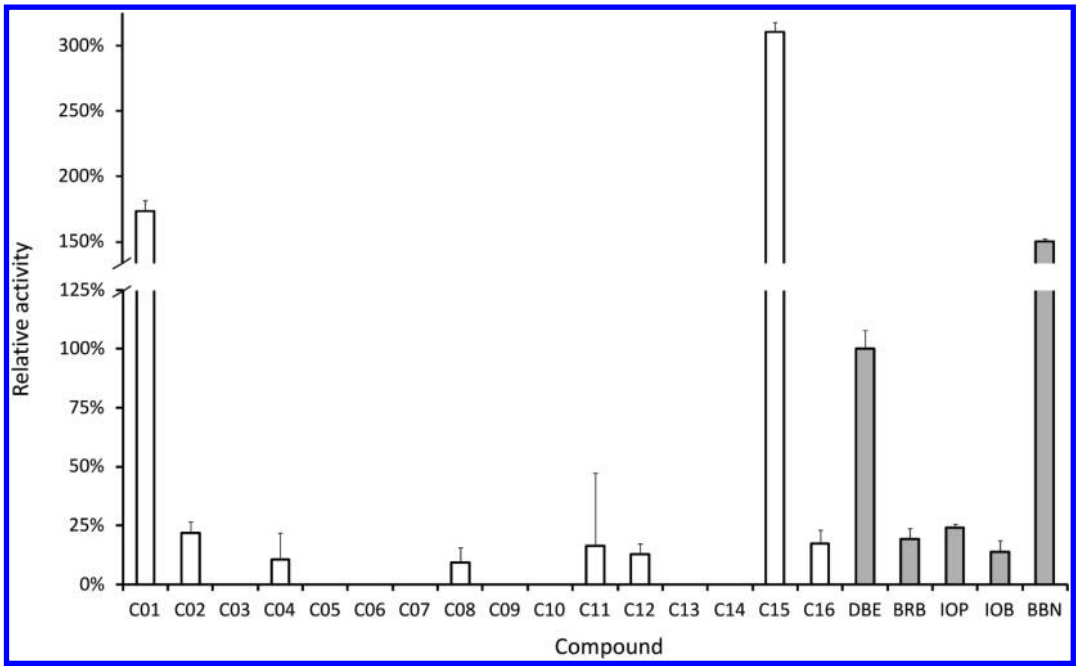


Figure 2. Experimentally determined activities of DmmA with 16 compounds identified by virtual screening as potential substrates. Each bar represents an average of three independent experiments. Activities with five universal substrates of HLDs are shown for comparison: 1,2-dibromoethane (DBE), 1-bromobutane (BRB), 1-iodopropane (IOP), 1-iodobutane (IOB), and 4-bromobutyronitrile (BBN). The activity with 1,2-dibromoethane was set to 100%. The same substrate is commonly used for kinetic characterization of HLDs.

enzyme. The enzymatic reaction was monitored by measuring the absorbance change at 550 nm associated with a decrease in pH using a FLUOstar Optima spectrophotometer (BMG Labtech, U.S.A.). The final concentration of DMSO was 30% to achieve comparable concentrations even for less soluble substrates. The inhibition assay was performed with 1,2-dibromoethane as substrate at a concentration of 9.8 mM and tested compounds at concentrations from 270 to 490 μM. The results were related to data for the reaction without addition of inhibitor.

Validation Using External Data Set. Structures of 233 compounds extracted from a data set published by Xu et al.[11] were converted to an AutoDock4.0 compliant format by Open Babel 2.3[54] and MGLTools. The coordinates of the crystal structure of CALB were downloaded from the Protein Data Bank (PDB ID: 1TCA). Hydrogen atoms were added to the receptor with the H++ server at pH 7.5. Gasteiger charges were subsequently assigned by MGLTools. During the docking procedure, the receptor was represented by a set of atomic and electrostatic maps calculated by AutoGrid4.0. The grid maps comprised 70 × 80 × 70 grid points with a spacing of 0.25 Å, which were centered at a position near the nucleophilic oxygen and catalytic base to cover the whole active site. The docking protocol was the same as described previously for HLDs.

The largest cluster obtained from the molecular docking of each compound was evaluated for proper positioning of the carbonyl between the two stabilizing residues (Thr40 and Gln106) using 4.5 Å as the cutoff distance. Properly stabilized compounds had at least one of the distances to each of the stabilizing residues within this cutoff. The —O···C═O reactive distance one (D_1) and —H···O reactive distance two (D_2) parameters measured for the set of 22 known substrates of CALB were recorded, and values applicable to 64% of the known substrates were set as the cutoff for the description of reactivity (Figure 2, Supporting Information).

RESULTS AND DISCUSSION

Development of Screening Protocol. Enzymes preferably recognize compounds in their transition (or so-called high-energy) states over the ground state structures.[55,56] Therefore, high-energy intermediates of substrates should have better affinity toward active sites than nonsubstrates and should be easily recognized by their lower binding energies. However, this is largely influenced by the accuracy of binding energies calculated from the docking. On the other hand, compounds in the ground state could be subjected to further minimization and rescoring procedures, allowing more precise prediction of binding. Moreover, they do not have to be exhaustively prepared for molecular docking because many databases of compounds in ready-to-dock formats are available.[57−60] The drawback of using ground states is a need for evaluation of their reactivity in addition to their binding.

To compare both approaches and select the most beneficial one, we initially docked 32 ground states and 43 corresponding high-energy states of known substrates to the DbjA, DmbA, LinB, DhaA, DhlA, and DbeA enzymes. Substrates with a halogen atom stabilized within 4.5 Å from the halogen-stabilizing residues were assessed for possible S_N2 displacement reaction, and the —COO^-···C−Cl reactive distance (R_D) and —COO^-···C−Cl reactive angle (R_A) parameters were recorded (Tables 1−4, Supporting Information).

Substrates docked in their ground state showed better predicted affinity and better correspondence between their predicted and experimental reactivity for six evaluated enzymes (Tables 5 and 6, Supporting Information). For these reasons and the aforementioned practical benefits of using ground states, we employed them for further screening. Using the data on the docked geometry of the known substrates in their ground states, strict cutoff values for reactivity applying to 70% of known substrates were identified as follows: R_D < 3.3 Å and R_A > 140° (Figure 3, Supporting Information).

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

Identification and Characterization of Novel Substrates. Altogether, 41,366 compounds were docked to the active site of DmmA. A total of 11,273 compounds had the halogen positioned within 4.5 Å from the halide-stabilizing residues. Their binding energies predicted by AutoDock ranged from −12.6 to −2.1 kcal mol^{-1} (Figure 4, Supporting Information). A total of 548 structurally unique compounds fulfilling the cutoff values for reactivity were selected for rescoring by NNScore 2.0 and MM/GBSA. Their binding energies calculated by MM/GBSA ranged from −51.4 to −12.4 kcal mol^{-1} confirming the potential of these compounds to bind the active site of DmmA (Figure 5, Supporting Information). Fifty top-ranked compounds were assessed for their commercial availability in the PubChem database and visually inspected in PyMOL[61] to verify proper binding according to the assumed catalytic mechanism (Table 7, Supporting Information). Sixteen compounds were selected (Table 2) and experimentally assayed for their activities with DmmA. Eight of these compounds were confirmed as substrates of DmmA. The new substrates comprised aromatic moieties with on average 50% higher molecular weight than common substrates of HLDs. The conversion of such bulky aromatic compounds has not previously been reported for HLDs but is consistent with the large active site of DmmA.[30,62] The new substrates also exhibited new types of interactions within the mostly hydrophobic active site. For example, compounds C01, C08, and C12 created a hydrogen bond with the backbone oxygen of N78. The catalytic activities of DmmA with these novel substrates were compared to the activity of this enzyme with five universal substrates of HLDs[30] (Figure 2), that is, 1-bromobutane (BRB), 1-iodopropane (IOP), 1-iodobutane (IOB), 1,2-dibromoethane (DBE), and 4-bromobutyronitrile (BBN). DmmA showed a 2-fold higher specific activity with C15 in comparison to BBN, which is the best universal substrate of HLDs, whereas the specific activity with C01 was comparable to BBN. The experimentally determined specific activity of DmmA toward the remaining six new substrates was within the activity range of BRB, IOB, and IOP.

Eight novel substrates of DmmA were assayed against three further enzyme family members (Figure 6, Supporting Information). DhaA was found to be active with C11, C15, and C16. LinB was active with C02, C11, C15, and C16. DbjA was active with C15. The activity of C11, C15, and C16 with three out of the four tested HLDs might be a consequence of the catalytic robustness of these enzymes, which belong to the same substrate specificity group.[30] Members of this group are generally regarded as the most active HLDs and are widely used for mechanistic studies,[22] as well as biotechnological applications.[16−20,63,64] The activity of DmmA with the substrates unique to this enzyme may be due to the unusual size and shape of its active site or architecture of the access tunnels.[30,65]

Characterization of Inactive Compounds. To understand the origin of unsuccessful prediction for the eight inactive compounds, we probed their ability to bind the active site of DmmA by measuring their inhibitory effect on the activity of DmmA with DBE. The determined residual activities were in the range of 21−93% (Figure 3). The observed inhibitory effects confirmed that all eight inactive compounds were able to bind to the active site of DmmA. Taking the results for both the experimentally confirmed substrates and inhibitors together, the prediction method achieved 100% success rate in the prediction of binding for the evaluated set of 16 compounds. Visual analysis of the bound inhibitors revealed that five of them were able to form a hydrogen bond with the reactive oxygen of the catalytic aspartate, hindering nucleophilic attack (Figure 4). We were not able to discern any obvious pattern for the three remaining compounds. Nonetheless, the knowledge of interactions between inhibitors and the active site can be applied for subsequent screening of new substrates or inhibitors of HLDs. The compounds showing no activity with DmmA were also assayed for activity with other three HLDs, DhaA, LinB, and DbjA (Figure 6, Supporting Information). However, no activity was detected with these enzymes, suggesting that hydrogen bonding with a nucleophile is a common inhibitory motif preventing conversion of these potential substrates by different HLDs.

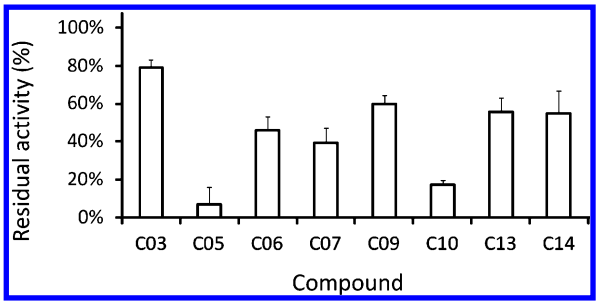


Figure 3. Experimentally determined inhibitory effect of eight compounds identified by virtual screening and showing no activity with DmmA. The inhibition effect was apparent by decreased conversion of 1,2-dibromoethane. The initial activity of DmmA with 1,2-dibromoethane in the absence of inhibitor was set to 100%.

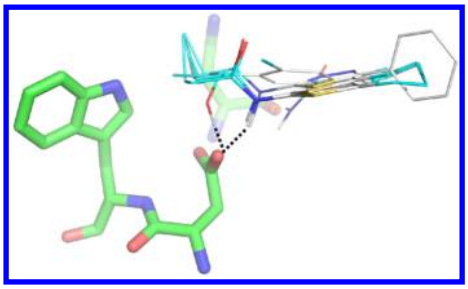


Figure 4. Binding modes of inhibitors in the active site of DmmA. Ligands are represented by lines, whereas the catalytic residues are represented by green sticks. The dashed lines represent the hydrogen bonds between the ligands and nucleophile.

Validation Using External Data Set. To evaluate the applicability of our method with other systems, we tried to identify potential substrates of CALB using the experimental data set[11] containing 206 substrates and 27 nonsubstrates (Tables 8 and 9, Supporting Information). Twenty-two substrates were used for development of mechanism-based geometric criteria to estimate the reactivity (Figure 2, Supporting Information). The selected cutoff values for the reactivity, applying to 64% of selected substrates, were $D_1 < 4.1$ Å and $D_2 < 5.1$ Å (Table 8, Supporting Information). The remaining 211 compounds were evaluated by the method. In this experiment, 158 compounds were predicted to be correctly stabilized, and 80 compounds were predicted as potential substrates (Table 9, Supporting Information). Depending on the number of compounds selected for experimental testing, the portion of successfully identified substrates would range

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

from 96% for the 50 top-ranked potential substrates to 89% for all 80 predicted substrates (Figure 7, Supporting Information). While the high success rates obtained for the CALB data set are partially due to the bias in the data set toward real substrates, they clearly confirm ability of our approach to enrich true substrates among selected compounds.

CONCLUSIONS

Substrate specificity is one of the most important features of enzymes. Here, we show that the substrate scope of an enzyme can be efficiently explored by *in silico* screening, requiring only knowledge of the tertiary structure, reaction mechanism, and several experimentally verified substrates.

The newly developed method applied to DmmA correctly predicted 50% of substrates and 100% of binders from 16 compounds proposed for experimental testing, confirming the possible utilization of this strategy for prioritizing active compounds. These statistics should be interpreted with caution because selection of more compounds for experimental testing would reduce these success rates.

The activities of eight novel substrates were found to be comparable to those of five universal substrates of HLDs; two of these substrates (C01 and C15) were found to be more active with DmmA than the best universal substrate of this enzyme family BBN. Three novel substrates, C11, C15, and C16, were additionally converted by three different HLD enzymes, DhaA, DbjA, and LinB.

The developed method was successfully evaluated with an external data set of 206 substrates of CALB, which confirmed its broader applicability for selection of true substrates. Our study demonstrates the potential of *in silico* screening for finding novel substrates of enzymes outside the chemical space of experimentally characterized compounds.

ASSOCIATED CONTENT

Supporting Information

Tables listing the reactive parameters and binding energies of docked substrates in the ground states and high-energy states; 50 top-ranked compounds obtained from the virtual screening; figures showing a region of DmmA enzyme selected for the molecular docking; cumulative distribution functions of reactive distances and reactive angles in HLDs; distribution of binding energies of properly stabilized and rescored compounds; activity assay of HLDs LinB, DhaA, and DbjA toward compounds identified by the virtual screening; and tables and figures related to validation on CALB substrates. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +420 5 4949 2616. E-mail: brezovsky@mail.muni.cz.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The work was supported by the Grant Agency of the Czech Republic (P503/12/0572) and the Czech Ministry of Education of the Czech Republic (LO1214 and LH14027). J.B. was supported by the "Employment of Best Young Scientists for International Cooperation Empowerment" (CZ.1.07/2.3.00/30.0037) project cofinanced by the European Social Fund and the state budget of the Czech Republic. MetaCentrum and CERIT-SC are acknowledged for providing access to computing facilities (LM2010005 and CZ.1.05/3.2.00/08.0144).

REFERENCES

(1) Schoemaker, H. E.; Mink, D.; Wubbolts, M. G. Dispelling the Myths—Biocatalysis in Industrial Synthesis. *Science* **2003**, *299*, 1694−1697.

(2) Reetz, M. T. Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **2013**, *135*, 12480−12496.

(3) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485*, 185−194.

(4) Lavecchia, A.; Di Giovanni, C. Virtual Screening Strategies in Drug Discovery: A Critical Review. *Curr. Med. Chem.* **2013**, *20*, 2839−2860.

(5) Shoichet, B. K. Virtual Screening of Chemical Libraries. *Nature* **2004**, *432*, 862−865.

(6) Hermann, J. C.; Marti-Arbona, R.; Fedorov, A. A.; Fedorov, E.; Almo, S. C.; Shoichet, B. K.; Raushel, F. M. Structure-Based Activity Prediction for an Enzyme of Unknown Function. *Nature* **2007**, *448*, 775−779.

(7) Gerlt, J. A.; Allen, K. N.; Almo, S. C.; Armstrong, R. N.; Babbitt, P. C.; Cronan, J. E.; Dunaway-Mariano, D.; Imker, H. J.; Jacobson, M. P.; Minor, W.; Poulter, C. D.; Raushel, F. M.; Sali, A.; Shoichet, B. K.; Sweedler, J. V. The Enzyme Function Initiative. *Biochemistry* **2011**, *50*, 9950−9962.

(8) Zhao, S.; Kumar, R.; Sakai, A.; Vetting, M. W.; Wood, B. M.; Brown, S.; Bonanno, J. B.; Hillerich, B. S.; Seidel, R. D.; Babbitt, P. C.; Almo, S. C.; Sweedler, J. V.; Gerlt, J. A.; Cronan, J. E.; Jacobson, M. P. Discovery of New Enzymes and Metabolic Pathways by Using Structure and Genome Context. *Nature* **2013**, *502*, 698−702.

(9) Favia, A. D.; Nobeli, I.; Glaser, F.; Thornton, J. M. Molecular Docking for Substrate Identification: The Short-Chain Dehydrogenases/reductases. *J. Mol. Biol.* **2008**, *375*, 855−874.

(10) Suresh, P. S.; Kumar, A.; Kumar, R.; Singh, V. P. An *in Silico* Approach to Bioremediation: Laccase as a Case Study. *J. Mol. Graph. Model.* **2008**, *26*, 845−849.

(11) Xu, T.; Zhang, L.; Wang, X.; Wei, D.; Li, T. Structure-Based Substrate Screening for an Enzyme. *BMC Bioinf.* **2009**, *10*, 257.

(12) Yalcin, E. B.; Stangl, H.; Pichu, S.; Mather, T. N.; King, R. S. Monoamine Neurotransmitters as Substrates for Novel Tick Sulfotransferases, Homology Modeling, Molecular Docking, and Enzyme Kinetics. *ACS Chem. Biol.* **2011**, *6*, 176−184.

(13) Irwin, J. J.; Raushel, F. M.; Shoichet, B. K. Virtual Screening against Metalloenzymes for Inhibitors and Substrates. *Biochemistry* **2005**, *44*, 12316−12328.

(14) Erable, B.; Goubet, I.; Lamare, S.; Legoy, M.-D.; Maugard, T. Bioremediation of Halogenated Compounds: Comparison of Dehalogenating Bacteria and Improvement of Catalyst Stability. *Chemosphere* **2006**, *65*, 1146−1152.

(15) Campbell, D. W.; Müller, C.; Reardon, K. F. Development of a Fiber Optic Enzymatic Biosensor for 1,2-Dichloroethane. *Biotechnol. Lett.* **2006**, *28*, 883−887.

(16) Prokop, Z.; Oplustil, F.; DeFrank, J.; Damborský, J. Enzymes Fight Chemical Weapons. *Biotechnol. J.* **2006**, *1*, 1370−1380.

(17) Prokop, Z.; Sato, Y.; Brezovsky, J.; Mozga, T.; Chaloupkova, R.; Koudelakova, T.; Jerabek, P.; Stepankova, V.; Natsume, R.; van Leeuwen, J. G. E.; Janssen, D. B.; Florian, J.; Nagata, Y.; Senda, T.; Damborsky, J. Enantioselectivity of Haloalkane Dehalogenases and Its Modulation by Surface Loop Engineering. *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 6111−6115.

(18) Ohana, R. F.; Hurst, R.; Vidugiriene, J.; Slater, M. R.; Wood, K. V.; Urh, M. HaloTag-Based Purification of Functional Human Kinases from Mammalian Cells. *Protein Expr. Purif.* **2011**, *76*, 154−164.

(19) Mazzucchelli, S.; Colombo, M.; Verderio, P.; Rozek, E.; Andreata, F.; Galbiati, E.; Tortora, P.; Corsi, F.; Prosperi, D.

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

Orientation-Controlled Conjugation of Haloalkane Dehalogenase Fused Homing Peptides to Multifunctional Nanoparticles for the Specific Recognition of Cancer Cells. *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 3121−3125.
(20) Neklesa, T. K.; Noblin, D. J.; Kuzin, A.; Lew, S.; Seetharaman, J.; Acton, T. B.; Kornhaber, G.; Xiao, R.; Montelione, G. T.; Tong, L.; Crews, C. M. A Bidirectional System for the Dynamic Small Molecule Control of Intracellular Fusion Proteins. *ACS Chem. Biol.* **2013**, *8*, 2293−2300.
(21) Verschueren, K. H.; Seljée, F.; Rozeboom, H. J.; Kalk, K. H.; Dijkstra, B. W. Crystallographic Analysis of the Catalytic Mechanism of Haloalkane Dehalogenase. *Nature* **1993**, *363*, 693−698.
(22) Prokop, Z.; Monincová, M.; Chaloupková, R.; Klvana, M.; Nagata, Y.; Janssen, D. B.; Damborský, J. Catalytic Mechanism of the Maloalkane Dehalogenase LinB from *Sphingomonas Paucimobilis* UT26. *J. Biol. Chem.* **2003**, *278*, 45094−45100.
(23) Schanstra, J. P.; Kingma, J.; Janssen, D. B. Specificity and Kinetics of Haloalkane Dehalogenase. *J. Biol. Chem.* **1996**, *271*, 14747−14753.
(24) Lau, E. Y.; Kahn, K.; Bash, P. A.; Bruice, T. C. 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.* **2000**, *97*, 9937−9942.
(25) Lightstone, F. C.; Zheng, Y.-J.; Bruice, T. C. Molecular Dynamics Simulations of Ground and Transition States for the S_N2 Displacement of Cl- from 1,2-Dichloroethane at the Active Site of *Xanthobacter Autotrophicus* Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **1998**, *120*, 5611−5621.
(26) Devi-Kesavan, L. S.; Garcia-Viloca, M.; Gao, J. Semiempirical QM/MM Potential with Simple Valence Bond (SVB) for Enzyme Reactions. Application to the Nucleophilic Addition Reaction in Haloalkane Dehalogenase. *Theor. Chem. Acc.* **2003**, *109*, 133−139.
(27) Hur, S.; Kahn, K.; Bruice, T. C. Comparison of Formation of Reactive Conformers for the SN2 Displacements by CH3CO2- in Water and by Asp124-CO2- in a Haloalkane Dehalogenase. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2215−2219.
(28) Soriano, A.; Silla, E.; Tuñón, I.; Ruiz-López, M. F. Dynamic and Electrostatic Effects in Enzymatic Processes. An Analysis of the Nucleophilic Substitution Reaction in Haloalkane Dehalogenase. *J. Am. Chem. Soc.* **2005**, *127*, 1946−1957.
(29) Shurki, A.; Strajbl, M.; Villà, J.; Warshel, A. How Much Do Enzymes Really Gain by Restraining Their Reacting Fragments? *J. Am. Chem. Soc.* **2002**, *124*, 4097−4107.
(30) Koudelakova, T.; Chovancova, E.; Brezovsky, J.; Monincova, M.; Fortova, A.; Jarkovsky, J.; Damborsky, J. Substrate Specificity of Haloalkane Dehalogenases. *Biochem. J.* **2011**, *435*, 345−354.
(31) Kmunícek, J.; Luengo, S.; Gago, F.; Ortiz, A. R.; Wade, R. C.; Damborský, J. Comparative Binding Energy Analysis of the Substrate Specificity of Haloalkane Dehalogenase from *Xanthobacter Autotrophicus* GJ10. *Biochemistry* **2001**, *40*, 8905−8917.
(32) Case, D.A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; I. Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; R. Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California: San Francisco, 2012.
(33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, P. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01 ; Gaussian, Inc.: Wallingford, CT, 2009.
(34) Hay, P. J.; Wadt, W. R. *Ab Initio* Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82*, 270−283.
(35) Hay, P. J.; Wadt, W. R. *Ab Initio* Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82*, 299−310.
(36) Wadt, W. R.; Hay, P. J. *Ab Initio* Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284−298.
(37) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(38) Sanner, M. F. Python: A Programming Language for Software Integration and Development. *J. Mol. Graph. Model.* **1999**, *17*, 57−61.
(39) Hsin, K.-Y.; Morgan, H. P.; Shave, S. R.; Hinton, A. C.; Taylor, P.; Walkinshaw, M. D. EDULISS: A Small-Molecule Database with Data-Mining and Pharmacophore Searching Capabilities. *Nucleic Acids Res.* **2011**, *39*, D1042−D1048.
(40) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C. The Protein Data Bank. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 899−907.
(41) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368−W371.
(42) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(43) Solis, F. J.; Wets, R. J.-B. Minimization by Random Search Techniques. *Math. Oper. Res.* **1981**, *6*, 19−30.
(44) Durrant, J. D.; McCammon, J. A. NNScore 2.0: A Neural-Network Receptor-Ligand Scoring Function. *J. Chem. Inf. Model.* **2011**, *51*, 2897−2903.
(45) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.
(46) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(47) Lee, M. C.; Duan, Y. Distinguish Protein Decoys by Using a Scoring Function Based on a New AMBER Force Field, Short Molecular Dynamics Simulations, and the Generalized Born Solvent Model. *Proteins* **2004**, *55*, 620−634.
(48) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(49) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383−394.
(50) Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the Molecular mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *J. Comput. Chem.* **2011**, *32*, 866−877.
(51) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L. Performance Comparison of Generalized Born and Poisson

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62

Methods in the Calculation of Electrostatic Solvation Energies for Protein Structures. *J. Comput. Chem.* **2004**, *25*, 265−284.
(52) Weiser, J.; Shenkin, P. S.; Still, W. C. Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). *J. Comput. Chem.* **1999**, *20*, 217−230.
(53) Holloway, P.; Trevors, J. T.; Lee, H. A Colorimetric Assay for Detecting Haloalkane Dehalogenase Activity. *J. Microbiol. Methods* **1998**, *32*, 31−36.
(54) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminformatics* **2011**, *3*, 33.
(55) Schramm, V. L. Enzymatic Transition States and Transition State Analogues. *Curr. Opin. Struct. Biol.* **2005**, *15*, 604−613.
(56) Warshel, A.; Florián, J. Computer Simulations of Enzyme Catalysis: Finding out What Has Been Optimized by Evolution. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950−5955.
(57) Irwin, J. J.; Shoichet, B. K. ZINC−a Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−182.
(58) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−D1107.
(59) Del Rio, A.; Barbosa, A. J. M.; Caporuscio, F.; Mangiatordi, G. F. CoCoCo: A Free Suite of Multiconformational Chemical Databases for High-Throughput Virtual Screening Purposes. *Mol. Biosyst.* **2010**, *6*, 2122−2128.
(60) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; Tang, A.; Gabriel, G.; Ly, C.; Adamjee, S.; Dame, Z. T.; Han, B.; Zhou, Y.; Wishart, D. S. DrugBank 4.0: Shedding New Light on Drug Metabolism. *Nucleic Acids Res.* **2014**, *42*, D1091−D1097.
(61) Delano, W. T. *The PyMOL Molecular Graphics System*, Version 1.5; Schrödinger, LLC: New York, 2009.
(62) Gehret, J. J.; Gu, L.; Geders, T. W.; Brown, W. C.; Gerwick, L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Structure and Activity of DmmA, a Marine Haloalkane Dehalogenase. *Protein Sci.* **2012**, *21*, 239−248.
(63) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nat. Chem. Biol.* **2009**, *5*, 727−733.
(64) Kurumbang, N. P.; Dvorak, P.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Computer-Assisted Engineering of the Synthetic Pathway for Biodegradation of a Toxic Persistent Pollutant. *ACS Synth. Biol.* **2014**, *3*, 172−181.
(65) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* **2009**, *392*, 1339−1356.

DOI: 10.1021/ci500486y
J. Chem. Inf. Model. 2015, 55, 54−62