nature chemical biology

Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate

Martina Pavlova[1,5], Martin Klvana[1,5], Zbynek Prokop[1], Radka Chaloupkova[1], Pavel Banas[2], Michal Otyepka[2], Rebecca C Wade[3], Masataka Tsuda[4], Yuji Nagata[4] & Jiri Damborsky[1]

Engineering enzymes to degrade anthropogenic compounds efficiently is challenging. We obtained *Rhodococcus rhodochrous* haloalkane dehalogenase mutants with up to 32-fold higher activity than wild type toward the toxic, recalcitrant anthropogenic compound 1,2,3-trichloropropane (TCP) using a new strategy. We identified key residues in access tunnels connecting the buried active site with bulk solvent by rational design and randomized them by directed evolution. The most active mutant has large aromatic residues at two out of three randomized positions and two positions modified by site-directed mutagenesis. These changes apparently enhance activity with TCP by decreasing accessibility of the active site for water molecules, thereby promoting activated complex formation. Kinetic analyses confirmed that the mutations improved carbon-halogen bond cleavage and shifted the rate-limiting step to the release of products. Engineering access tunnels by combining computer-assisted protein design with directed evolution may be a valuable strategy for refining catalytic properties of enzymes with buried active sites.

Combining rational protein design with directed evolution provides a very efficient two-step approach for engineering proteins with desired activities[1–7]. Computer modeling is initially used to identify residues participating in substrate binding and transition state stabilization, then directed evolution is used for systematic mutagenesis of these "hot spot" residues and for fine-tuning constructs prepared by computer-guided site-directed mutagenesis[8]. The amino acid residues that are most frequently predicted as hot spots for mutagenesis are located in the active site or its close vicinity[9]. However, in the study presented here we targeted residues located in the access tunnels connecting the active site of the engineered protein (DhaA haloalkane dehalogenase from *Rhodococcus rhodochrous* NCIMB13064) with bulk solvent. The results demonstrate the power of combining rational design with directed evolution and highlight the importance of access tunnels as both evolutionary loci and protein engineering targets.

Living organisms have typically had millions of years to evolve enzymes capable of metabolizing naturally occurring substances, including potential toxins. In contrast, anthropogenically generated substrates introduced into the biosphere by humans since the Industrial Revolution are often recalcitrant and persist in the environment because microorganisms have not had sufficient time to evolve enzymes capable of catalyzing their conversion[10,11]. One such toxic, non-natural compound is TCP (**1**), which is released into the environment as a result of its manufacture, formulation and use as a solvent and extractive agent. TCP is used as a chemical intermediate in the production of polysulfone liquid polymers and 1,3-dichloroprop-1-ene (**2**). In addition, it is produced in significant quantities as a byproduct in the manufacture of other chlorinated compounds—for example, 2-(chloromethyl)oxirane (also called epichlorohydrin, **3**)[12]. TCP has been detected in low concentrations in surface water, drinking water and groundwater, and is anticipated to be a human carcinogen. Further, given that its half-life under groundwater conditions is estimated to extend up to 100 years, TCP is likely to persist for long periods of time in the groundwater environment[13]. However, under laboratory conditions, TCP can be slowly (k_{cat} = 0.08 s^{-1}) converted to 2,3-dichloropropane-1-ol (DCL, **4**) by the action of microbial haloalkane dehalogenase enzymes[14], which cleave carbon-halogen bonds by a hydrolytic mechanism. The first dehalogenase reported to have detectable activity against TCP is the DhaA haloalkane dehalogenase[15], which has been isolated from various strains of *Rhodococcus* spp. worldwide[16–20]. The active site of DhaA is an occluded cavity located between two protein domains[21], with two major access tunnels (referred to as the main tunnel and the slot tunnel[22]) through which ligands can exchange between the active site and its surrounding environment. The dehalogenation reaction proceeds in the active site via nucleophilic attack of a nucleophile on a carbon atom of the halogenated substrate, leading to cleavage of the carbon-halogen bond, displacement of a halide and formation of a covalent alkyl-enzyme intermediate. This alkyl-enzyme intermediate is subsequently hydrolyzed by a water molecule activated by a catalytic base[23].

DNA shuffling and error-prone PCR have been applied to the *dhaA* gene to improve the kinetic properties of DhaA for TCP conversion[14,24]. Two evolved double point mutants (G3D C176F and C176Y Y273F (M2)) were obtained that were 4 times and 3.5 times

[1]Loschmidt Laboratories, Institute of Experimental Biology and National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Brno, Czech Republic. [2]Department of Physical Chemistry and Center for Biomolecules and Complex Molecular Systems, Palacký University, Olomouc, Czech Republic. [3]Molecular and Cellular Modeling Group, EML Research gGmbH, Heidelberg, Germany. [4]Department of Environmental Life Sciences, Graduate School of Life Sciences, Tohoku University, Sendai, Japan. [5]These authors contributed equally to this work. Correspondence should be addressed to J.D. (jiri@chemi.muni.cz).

Received 20 January; accepted 28 May; published online 23 August 2009; doi:10.1038/nchembio.205

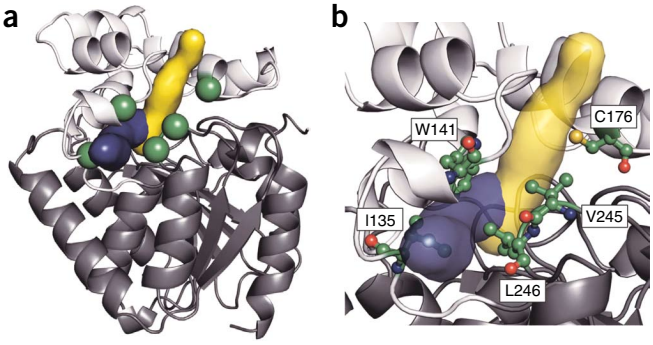

Figure 1 Hot spot residues lining the access tunnels selected for mutagenesis by computer simulations. (**a**) Cartoon model of wild-type DhaA. Gray, main domain; white, cap domain; yellow, product release pathway corresponding to the main tunnel; blue, product release pathway corresponding to the slot tunnel; green, residues selected for mutagenesis. (**b**) Ball and stick model of the residues selected for mutagenesis.

more active, respectively, with TCP than the wild-type enzyme. It is noteworthy that both independently evolved mutants carried a substitution at position 176. We have recently used molecular dynamics simulations and quantum mechanical calculations to elucidate the structural basis for improved catalysis in the mutants obtained in these directed evolution experiments[25]. The results indicated that the mutations did not alter the positioning of the substrate in the active site or the stabilization of the transition state. Instead, the C176Y mutation was found to modify the mouth of the access tunnel connecting the buried active site with the surrounding solvent. Quantitative comparison of the access tunnels in the wild type and the mutants using CAVER[26] revealed that the substitution of Cys176 by a bulkier tyrosine or phenylalanine narrowed the main tunnel of DhaA[25].

In the present study, we combined several methods of rational design and directed evolution in an iterative manner (**Supplementary Fig. 1**) to obtain DhaA mutant enzymes with significantly improved TCP conversion. We applied random acceleration molecular dynamics (RAMD)[27] to simulate the release of the product from the enzyme's active site and identify the key residues for mutagenesis. These hot spots in the access tunnels were targeted by saturation mutagenesis, exhaustive libraries of mutant genes were constructed and mutants with improved activities were identified by screening. We obtained 25 unique protein variants with higher activities toward TCP, the most active of which showed 32-fold higher activity compared to the natural wild-type enzyme.

RESULTS

Saturation mutagenesis of hot spots in access tunnels

The release of the dehalogenation product DCL from the wild-type DhaA and the M2 mutant was studied using RAMD simulations to probe the access tunnels connecting the buried active site with bulk solvent and to identify possible bottlenecks. The calculations indicated that DCL was released more frequently via the main tunnel in the wild-type enzyme than in M2 (**Supplementary Table 1**). The release of the product through the main tunnel occurred in 17 out of 24 RAMD simulations (~70%) with the wild-type enzyme and in 4 out of 8 RAMD simulations (~50%) with M2. The simulations also indicated that the product could be released through the slot tunnel of both the wild-type enzyme and the M2 mutant (occurring in ~8% and ~13% of the simulations, respectively). We hypothesized that the slot tunnel has greater importance for the ligand's entry to or exit from the active site in the M2 mutant than in the wild-type enzyme, and that this increased importance is associated with the mutant's increased activity toward TCP. Thus, we selected residues located in the slot tunnel (Ile135), at the interface of the slot tunnel and the active site (Trp141 and Leu246) and at the interface of the slot tunnel, the active site and the main tunnel (Val245) as hot spots for mutagenesis (**Fig. 1**).

To assess the mutability of the sites identified as hot spots by the computer simulations, their amino acid propensities were analyzed[28] in a multiple sequence alignment of one of the three subfamilies of haloalkane dehalogenases: HLD-II (ref. 29). In this alignment, isoleucine in position 135 is substituted by methionine, valine, phenylalanine and threonine; tryptophan in position 141 is substituted by phenylalanine and serine; valine in position 245 is substituted by alanine, phenylalanine and valine; and leucine in position 246 is substituted by isoleucine. Trp141 is one of the key residues determining the diameter of the slot tunnel. However, mutations of tryptophan often lead to protein unfolding; therefore, a separate library containing the most conservative substitution (tryptophan to phenylalanine) was constructed. More specifically, W141F was inserted into a double mutant M2 described previously (C176Y Y273F)[14], resulting in the triple mutant M3 (W141F C176Y Y273F). The mutants M2 and M3 were used as starting proteins for directed evolution experiments leading to libraries A and B, respectively. Saturation mutagenesis with degenerate primer sets was used to introduce all possible point mutations at the desired positions 135, 245 and 246. Libraries A and B contained 2,568 and 2,705 clones, respectively.

Mutants with enhanced activity against TCP

A total of 5,273 clones from the final libraries were picked, and DhaA variants were expressed and screened for dehalogenating activity toward TCP. Most mutants in the library (96%) had little or no such activity, but a small proportion appeared to have higher activity than wild-type. The 51 clones exhibiting the highest improvement in activity relative to wild-type cultures were isolated and subjected to more detailed characterization. Sequencing revealed 25 unique mutant variants with higher activity than the template (**Supplementary Table 2**). Some of the mutations would probably never have been obtained by a routine directed evolution technique, such as error-prone PCR, due to the bias of the saturation mutagenesis method[30]. For instance, the mutation of valine to phenylalanine requires a two-base change in the codon triplet GTA to TTT. One clone from library A (variant 27) carried a spontaneous mutation, R254A.

Small-scale cultivation was used to compare the level of DhaA expression of the cultures carrying the 25 unique mutant variants with activity higher than wild-type and to obtain cell-free extracts for further analysis. The expression level was comparable, and no

a

Variant	Residue					
	135	141	176	245	246	273
WT	I	W	C	V	L	Y
M2	I	W	Y	V	L	F
M3	I	F	Y	V	L	F
17	F	F	Y	M	I	F
19	Y	F	Y	M	I	F
21	L	F	Y	F	I	F
27	V	W	Y	F	I	F
31	F	W	Y	F	I	F
33	C	W	Y	Y	L	F

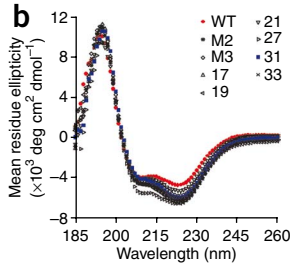

Figure 2 Characteristics of mutants with enhanced activity against TCP. (**a**) Summary of mutations in selected DhaA variants. (**b**) Far-UV circular dichroism spectra of the wild-type DhaA and its variants.

Table 1 Steady state kinetic parameters of wild-type DhaA and its variants

Variant	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} M^{-1})	rel. k_{cat} (Mut/WT)	rel. K_m (Mut/WT)	rel. k_{cat}/K_m (Mut/WT)
WT[a]	0.08	2.2	36	1.0	1.0	1.0
M2[a]	0.28	1.0	280	3.5	0.5	7.8
WT[b]	0.04 ± 0.01	1.0 ± 0.2	40 ± 13	1.0	1.0	1.0
M2[b]	0.14 ± 0.04	0.9 ± 0.1	156 ± 48	3.5	0.9	3.9
M3[b]	0.07 ± 0.01	1.0 ± 0.1	70 ± 12	1.8	1.0	1.8
21[b]	0.55 ± 0.04	1.2 ± 0.2	458 ± 83	13.8	1.2	11.5
27[b]	1.02 ± 0.06	1.1 ± 0.1	927 ± 100	25.5	1.1	23.2
31[b]	1.26 ± 0.07	1.2 ± 0.1	1,050 ± 105	31.5	1.2	26.3

k_{cat}, catalytic constant; K_m, Michaelis constant; WT, wild type; Mut, mutant.
[a]Data published previously[14]. Kinetic constants were measured in 50 mM $NaHCO_3$-NaOH buffer at pH 9.4 and 30 °C. The s.d. for K_m and for k_{cat} were ≤25% and ≤15%, respectively. [b]Data from this study. Kinetic constants were measured in 100 mM glycine buffer at pH 8.6 and 37 °C.

significant difference among mutants was observed (data not shown). The relative activity of the mutants was up to 40-fold higher than that of the wild-type enzyme (**Supplementary Table 2**). The wild-type, M2, M3 and the clones with the highest activity toward TCP (17, 19, 21, 27, 31 and 33) were overexpressed and purified to homogeneity with an average purification yield of 20 mg l^{-1} (**Fig. 2a**).

Circular dichroism (CD) spectra in the far-UV spectral region were used to assess whether the mutations carried by the purified variants affected their folding and secondary structure. The spectra of all the tested enzyme variants show two negative peaks at 208 and 222 nm, and one positive peak at 195 nm (**Fig. 2b**). The spectra of the variants are nearly identical to that of the wild-type enzyme, suggesting that they fold compactly. Small differences between the mutants and the wild-type can be seen at the second negative peak, where the CD spectra of the variants exhibit an increase in $\Theta_{222\,nm}$ values (8 to 35%) compared to the wild type. This indicates that the inserted mutations have a small effect on the packing of amino acid residues in the secondary structure of the mutant enzymes, particularly on the overall number of amino acid residues contributing to the α-helical content. The improved catalytic properties of the constructed and evolved protein variants in terms of specific activities were analyzed (**Supplementary Fig. 2**).

The three variants with the highest activity (21, 27 and 31) clearly showed higher activity than the templates M2 and M3, and were characterized in detail. The steady state kinetic parameters were determined with TCP as a substrate (**Table 1**). The variants 21, 27 and 31 have higher k_{cat} values than the wild-type enzyme and both M2 and M3 mutants. The variant 31 has a k_{cat} value equal to 1.26 s^{-1}, which is 32 times higher than that of the wild-type. K_m is slightly increased for all tested clones, but the catalytic efficiency of the variants 21, 27 and 31, quantified by k_{cat}/K_m, is 12, 23 and 26 times higher, respectively, than that of the wild-type enzyme. The rates of dehalogenation catalyzed by the wild-type DhaA and the variant 31 with several other halogenated substrates were also measured (**Supplementary Fig. 3**), and the variant 31 showed a systematic increase in specific activity toward 1,2-disubstituted C2-C3 haloalkanes, and reduced activity toward longer (>C3) haloalkanes.

Molecular basis of enhanced activity against TCP

RAMD simulations of the egress of DCL were conducted with the wild-type enzyme, the template variant M2 and the most active variants obtained from mutagenesis experiments (21, 27 and 31) to gain better understanding of the molecular basis for enhanced activity with TCP. DCL release through the main tunnel was observed in simulations of all the variants (**Supplementary Fig. 4**), but with decreasing frequency from mutant 21 to mutant 31—that is, in the direction of increasing activity. The release of DCL through the slot tunnel appears to be completely prevented in mutant 31 by the I135F mutation. However, it can occur in mutants 21 and 27, although DCL preferentially exits via other pathways in these mutants: pathway 2c in mutant 21 as a result of the W141F substitution, and pathway 2b in mutant 27 due to the I135V substitution. The 31 mutant is the least amenable to DCL release because of three bulky substitutions: I135F, C176Y and V245F. Release of DCL from the active site of this mutant occurred in only one out of eight (~13%) simulations (**Supplementary Table 1**).

The 31 mutant, which shows the highest activity toward TCP, contains the large aromatic residues in two out of the three randomized positions as well as in the two positions modified by site-directed mutagenesis. Narrowing of the access tunnels lowered the frequency of product egress in simulations, suggesting that this product release is unlikely to be rate limiting for dehalogenation of TCP by wild-type DhaA. At the same time, the introduction of large hydrophobic residues into the access tunnels restricts the exchange of water molecules between the bulk solvent and the buried active site. The dynamics of water molecules were therefore analyzed together with the stability of the substrate TCP in the reactive orientation, the so-called near-attack configuration (NAC), during 2 ns of equilibration molecular dynamics (EMD) followed by 2 ns of production molecular dynamics (MD) simulation of the wild-type DhaA and of the 31 mutant. This analysis revealed that the active site of wild-type DhaA is easily accessible to water molecules from bulk solvent (**Fig. 3a**) and that increasing numbers of water molecules in the active site compete with TCP for interaction with the nucleophile, Asp106. The competition leads to reduced formation of the NAC since the water molecules sterically hinder access of TCP to the nucleophile and thus hinder formation of the activated complex (**Fig. 3b**,**c**). The population of the NAC in the wild type was 6.7-fold lower than in the 31 mutant. The narrowed tunnels in the 31 mutant shield the active site from the bulk solvent, and the activated complex is more stable than in the wild-type enzyme.

Proton inventory was performed to investigate rate limitation in the conversion of TCP by the 31 mutant (**Fig. 3d**, inset). The proton inventory was linear, suggesting that the solvent kinetic isotope effect is due to a single site in a transition state. Increasing the fraction of ^2H_2O results in increased enzymatic activity of the 31 mutant. This result indicates a shift of the rate-limiting step in the conversion of TCP in the 31 mutant, since no effect of ^2H_2O has been detected in experiments with the wild-type DhaA[31]. The increase of the catalytic rate in the presence of ^2H_2O further suggests that the step determining the rate of TCP conversion by the 31 mutant is neither hydrolysis nor slow conformational change[31]. Transient kinetic experiments were conducted with the wild-type DhaA and the 31 mutant to gain mechanistic insight into the catalytic performance of these two enzymes and to identify steps affected by the mutagenesis. No sign of burst was observed for either chloride ion or DCL formation after rapid mixing of the wild-type DhaA with excess TCP (**Fig. 3d**). Instead, linear formation of both products was observed with a rate constant of 0.07 ± 0.01 s^{-1}, which corresponds well with the steady state turnover number of 0.04 ± 0.01 s^{-1}. The absence of burst for both chloride ion and DCL formation suggests that the slowest step of TCP conversion by the wild-type DhaA lies at the beginning of the reaction cycle, before or in the step forming the chloride ion product (**Fig. 3d**). Based on previous pre-steady state kinetic studies—that is, on DhlA with

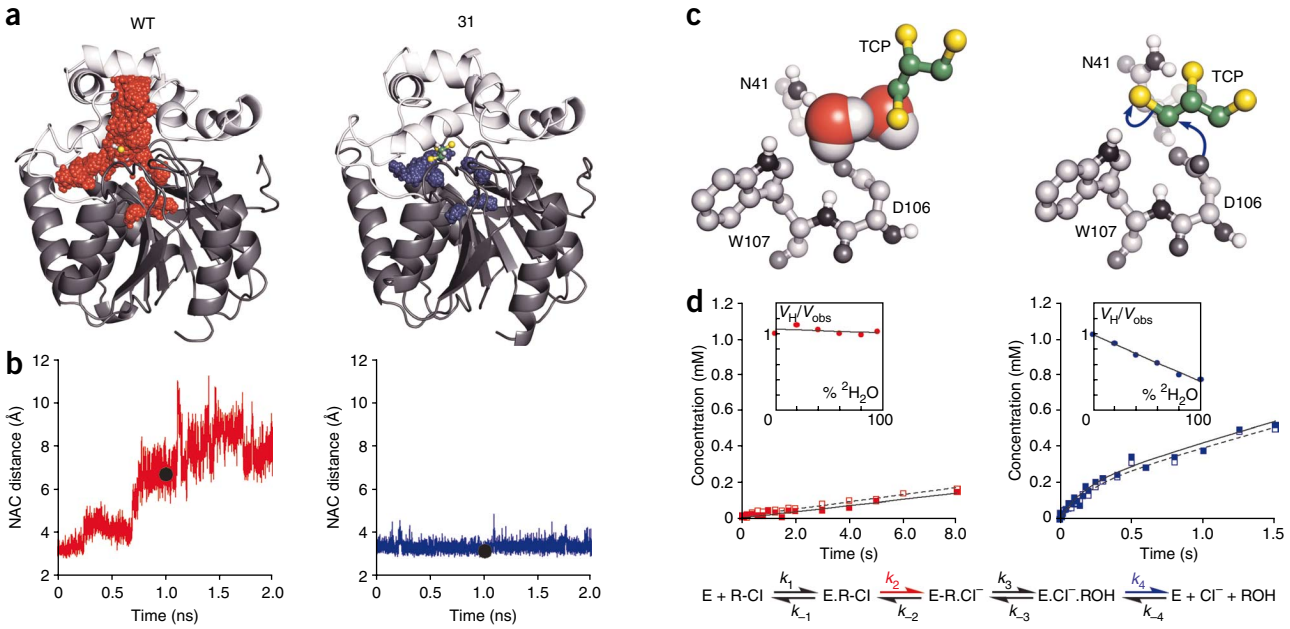

Figure 3 Structural basis of enhanced activity against TCP. Comparison of the wild-type DhaA (left) and the 31 mutant (right). (**a**) Accessibility of active sites for bulk solvent during MD simulations. The active site cavity of the wild-type DhaA is accessible to bulk solvent (red); the cavity of the 31 mutant is closed, and only discrete internal positions are occupied by structural waters (blue). Spheres representing waters have a radius of 0.5 Å centered on the oxygen atoms. (**b**) Effect of water molecules on the NAC distance of two reacting atoms in the active site during MD simulations. Snapshots shown in **c** are indicated by (•). (**c**) Snapshots from MD simulations of the nonreactive configuration of TCP competing with waters for the nucleophile in the wild-type DhaA and the NAC for TCP in the 31 mutant. Only polar hydrogens are shown for clarity. (**d**) Rate-limiting step from the pre-steady state kinetics. Solid and dashed lines represent the best fits to the chloride ion (■) and DCL (□) kinetic data. Reaction scheme shows rate-limiting steps for the reactions catalyzed by the wild-type DhaA (red) and the 31 mutant (blue). E, free enzyme; R-Cl, substrate; E.R-Cl, enzyme-substrate complex; E-R.Cl⁻, alkyl-enzyme intermediate; E.Cl⁻.ROH, enzyme-product complex; Cl⁻ and ROH, products. The inset illustrates the solvent kinetic isotope effect on the dehalogenase activity. The ratio between the activity in H_2O (V_H) and in D_2O (V_{obs}) was plotted versus various $^1H_2O/^2H_2O$ ratios. The data for wild-type DhaA are from a previous study[31].

1,2-dichloroethane (**5**) and 1,2-dibromoethane (**6**)[32], on LinB with 1-chlorohexane (**7**) and bromocyclohexane (**8**)[33], and on DhaA with 1,3-dibromopropane (**9**)[31], in which substrate binding was found to be significantly faster than cleavage of the carbon-halogen bond—we expected the TCP conversion catalyzed by the wild-type DhaA to be limited by cleavage of the carbon-chlorine bond. In contrast, upon rapid mixing of the 31 mutant with excess TCP, a clear pre-steady state burst was identified for both chloride ions and DCL, with observed rates of 5.8 ± 2.3 and 5.2 ± 1.5 s^{-1}, respectively (**Fig. 3d**). The burst was followed by a linear steady state phase, with rate constants of 1.42 ± 0.31 and 1.36 ± 0.18 s^{-1} for chloride ions and DCL, respectively, which correspond well with the k_{cat} value observed in steady state kinetics (1.26 ± 0.07 s^{-1}). These findings suggest that none of the chemical steps is limiting for the 31 mutant; instead, the rate-limiting step for the reaction with TCP catalyzed by this mutant is the release of a product (**Fig. 3d**).

DISCUSSION

There are many documented cases of organisms acquiring the ability to degrade anthropogenic chemicals within a few years (or less) following contact with them[34]. It is unclear how the proteins involved can evolve new functions so rapidly, but the ability generally depends on the introduction of new phenotypic traits by a small number of mutations. Two major approaches can be applied to mimic such rapid evolution and enhance the catalytic performance of enzymes toward non-natural substrates in the laboratory: rational protein design and directed evolution. The former is the most promising and versatile approach for generating new activities, while the latter seems to be the best way to optimize catalytic properties of existing enzymes[8].

In this study, we constructed a DhaA variant with 32-fold higher catalytic activity and 26-fold higher catalytic efficiency for TCP than wild-type DhaA by 'focused directed evolution' combining computer-assisted molecular design and directed evolution. Identification of specific sites in the protein for combinatorial mutagenesis substantially reduced the sequence space required to isolate viable mutants. A mutant library of ~5,000 clones yielded 51 positive clones and 25 unique sequences with improved TCP conversion—significantly more than in a previous study in which a random library of DhaA mutants created by DNA shuffling[14] yielded 30 potentially positive variants out of 10,000 clones, and only one (M1, C176Y) showing improved TCP conversion. A second-generation library created by error-prone PCR provided 14 potentially positive variants and one (M2, C176Y Y273F) with further improved catalytic performance[14]. However, neither the wild-type enzyme nor M2 have sufficient catalytic activity and catalytic efficiency with TCP (k_{cat} = 0.04 and 0.14 s^{-1}, k_{cat}/K_m = 40 s^{-1} M^{-1} and 156 s^{-1} M^{-1}, respectively) for biodegradation in a continuous-flow system, which would be essential for cost-effective, full-scale bioremedial processes. The catalytic constants of our mutant with highest activity toward TCP (k_{cat} = 1.26 s^{-1}, k_{cat}/K_m = 1,050 s^{-1} M^{-1}) are of the same order of magnitude as those of the haloalkane dehalogenase DhlA for 1,2-dichloroethane (k_{cat} = 3.3 s^{-1}, k_{cat}/K_m = 6,200 s^{-1} M^{-1}), which has proven utility in 1,2-dichloroethane–mineralizing microorganisms used in a full-scale groundwater purification plant[35]. The haloalkane dehalogenase DhlA-producing microorganisms have survived, grown and shown

potential to degrade 1,2-dichloroethane efficiently for at least five years after the startup of the plant. Construction of a TCP-degrading bacterium carrying the mutant dehalogenase identified in this work is currently in progress.

In the present study, focusing on the apparent importance of narrowing the main access tunnel[25], the most active mutant 31 contains the large aromatic residues in four out of five positions targeted by mutagenesis, leading to the most occluded active site. Product release was a rare event in RAMD simulations of variant 31, and MD simulations of its enzyme-substrate complex showed that the modification of its entrance pathways restricts access of water molecules to the occluded active site cavity. One water molecule must be present in the active site as a co-substrate for the hydrolytic step, but additional water molecules inhibit carbon-halogen bond cleavage[36–38]. Rapid quench-flow analysis confirmed indications from MD simulations that carbon-halogen bond cleavage is the rate-limiting step for TCP conversion by the wild-type DhaA—as reported previously[31]—and that this step is significantly improved in the 31 mutant. The observed switch of the rate-limiting step from the chemical reaction to the product release in the 31 mutant documents the trade-off that favors catalysis in the occluded active site and disfavors product release. Substrate mapping for the 31 mutant showed that it also has increased activity toward other 1,2-halogenated ethanes and propanes, in accordance with the proposed limitation of carbon-halogen bond cleavage for these substrates. Precise formation of the activated complex and stabilization of the transition state are important for the bimolecular nucleophilic substitution–mediated conversion of 1,2-halogenated substrates by the haloalkane dehalogenases, since halogen substitutents in the 1,2-position strongly sterically hinder the nucleophilic attack.

Following previous comparative analysis of crystal structures, we proposed that access tunnels substantially influence haloalkane dehalogenases' substrate specificity and activity[39] and verified this experimentally by exhaustive site-directed mutagenesis of the tunnel residue Leu177 in the haloalkane dehalogenase LinB[40]. Access tunnels have also been shown to affect the substrate specificity of cytochrome P450s[27,41,42]. Papers describing variants of several enzymes carrying single-point mutations in access tunnels with modified specificities, enantioselectivities or activities[40,43–49] support the general applicability of redesigning access tunnels to alter enzymes' activities. Here we have demonstrated the utility of simultaneously randomizing several residues that influence ligand and water passage through access tunnels for optimizing tunnel anatomy, physicochemical properties and dynamics.

In summary, we provide experimental evidence that the activity of enzymes toward anthropogenic substrates can be improved by modifying access routes connecting the buried active site cavity with the surrounding solvent. The results indicate that identifying hot spot residues lining access tunnels by computer modeling and simulations combined with exhaustive mutagenesis by directed evolution may be a valuable, generally applicable methodology for engineering new biocatalysts with buried active sites. The tunnel residues located outside the active site represent good targets for mutagenesis since their replacement does not lead to loss of functionality by disruption of the active site architecture.

METHODS

Molecular modeling and simulation. The RAMD simulations of DhaA and its mutants were based on the crystal structure of a variant of the free enzyme (Protein Data Bank code 1CQW). Three point substitutions were made computationally in the crystal structure to recover the wild-type DhaA sequence. The structures of the mutant proteins were modeled using the mutagenesis module of PyMol 0.99 (DeLano Scientific). Details of the RAMD simulations of product release are provided in the **Supplementary Methods**. The effects of the active site water molecules on reactivity were studied by classical MD. The simulations consisted of 2 ns of EMD followed by 2 ns of production MD. The presence of NACs along the calculated MD trajectories was evaluated every 0.5 ps using the two-parameter geometric conditions proposed previously for this reaction class: (i) the distance between the nucleophilic oxygen and the attacked carbon atom ≤ 3.41 Å and (ii) the angle formed by the nucleophilic oxygen, the attacked carbon and the leaving chlorine > 157° (ref. 36).

Mutagenesis, seeding and propagation of the libraries. The His-tagged DhaA wild type and two mutants denoted M2 and M3 (**Fig. 2a**) were constructed by site-directed mutagenesis using the principle of inverted PCR with KOD Plus polymerase (Toyobo). The oligonucleotides used for mutagenesis are listed in **Supplementary Table 3**. Saturation mutagenesis at amino acid positions 135, 245 and 246 was carried out using a set of degenerate synthetic oligonucleotides. Two oligonucleotide primer pairs were designed to create two separate sublibraries: DhaA135 (diversity 32) and DhaA245-6 (diversity 1,056). Details of saturation mutagenesis are described in the **Supplementary Methods**. The individual colonies were picked with sterile toothpicks and resuspended in separate wells of a 96-well plate containing LB medium and 100 μg ml^{-1} ampicillin (**10**). Each 96-well plate contained eight positive controls of M2, four negative controls with the pAQN vector and four wild-type controls. The plates were shaken at 120 r.p.m. at 37 °C for at least 16 h, and a replica of the 96 wells was made on LB agar plates. The microcultures were induced with IPTG (**11**) at a final concentration 0.5 mM and cultivated for a further 5 h at 30 °C, pelleted at 4 °C by centrifugation at 1,800 r.p.m. for 35 min and stored at −80 °C.

High-throughput microplate screening. The microplates with the propagated mutant library were defrosted, and cell pellets were lysed in 1 mM HEPES buffer (pH 8.2). 10 μl of cell free extract was transferred into 140 μl of phenol red (**12**) assay buffer (1 mM HEPES buffer, 20 mM Na_2SO_4, 1 mM EDTA, pH 8.2, and phenol red at final concentration 25 μg ml^{-1}). The TCP substrate was added to an apparent concentration of 10 mM. Haloalkane dehalogenases catalyze the hydrolytic displacement of a chloride ion from the substrate. One water molecule is consumed per substrate molecule, and the reaction products are a primary alcohol, a chloride ion and a proton. The pH of the weakly buffered medium is reduced rapidly by the production of protons. Dehalogenating activity was therefore detected by adding a phenol red dye that changes color upon a change in pH. The microplates were sealed with parafilm and incubated at 25 °C for 16 h. Active colonies produced a bright yellow color, while inactive colonies and negative controls formed a deep purple color. Wells containing microcultures exhibiting the most significant improvement in activity were identified, and corresponding replica colonies were used to inoculate 3 ml of LB media. The plasmids were subsequently isolated and used for further analysis.

Overexpression and purification. The *dhaA* gene variants were transcribed in a pAQN vector by the *tac* promotor under the control of *lacI*^{q}. *Escherichia coli* BL21 cells with plasmids were cultured in 10 ml (for cell-free extract) and in 2 l (for purification) of LB at 37 °C. The induction of enzyme expression at 30 °C was initiated by the addition of IPTG to a final concentration of 0.5 mM. The cells were harvested and disrupted by sonication using a Soniprep 150 (Sanyo Gallenkamp PLC). The supernatant was used after centrifugation at 21,000 g for 1 h. The cell-free extract was further purified using a HiTrap affinity column packed with high performance chelating sepharose (Amersham Pharmacia Biotech). The His-tagged enzymes were bound to the resin in the equilibrating buffer (50 mM phosphate buffer, pH 7.5, containing 500 mM sodium chloride and 10 mM imidazole). Unbound and weakly bound proteins were washed out with buffer containing 500 mM imidazole. The active fractions were pooled and dialyzed overnight against 50 mM phosphate buffer (pH 7.5). The enzymes were sterilized and stored in the same buffer. The entire purification process and storage was performed at 4 °C. CD spectra of protein samples were recorded as described in the **Supplementary Methods**.

Activity assay, steady state and pre-steady state kinetics. Haloalkane dehalogenase activity was assayed by a previously described method[50], by measuring

released halide ions spectrophotometrically at 460 nm with mercuric thiocyanate and ferric ammonium sulfate. One unit of enzymatic activity is defined as the release of 1 µmol of halide ion per minute. Solvent kinetic isotope effects were determined by performing activity assays in glycine buffer (pH 8.6) containing TCP and increasing concentrations of ^2H_2O. Michaelis-Menten kinetic constants of DhaA wild type and mutants with TCP were determined by initial velocity measurements. Rapid quench-flow experiments were performed at 37 °C in a glycine buffer (pH 8.6) using a QFM 400 rapid quench-flow instrument (BioLogic). Details of the experimental design and data analysis are described in the **Supplementary Methods**.

Accession codes. Entrez Nucleotide: the nucleotide sequence of *dhaA* (AF060871) was deposited as part of a previous study. Protein Data Bank: the crystal structure of DhaA (1CQW) was deposited as part of a previous study.

Note: Supplementary information and chemical compound information is available on the Nature Chemical Biology website.

ACKNOWLEDGMENTS

We thank V. de Lorenzo (Centro Nacional de Biotecnologia, Spain) and U. Bornscheuer (University Greifswald, Germany) for critical reading of this manuscript. R.C.W. gratefully acknowledges the support of the Klaus Tschira Foundation and M.P. acknowledges a scholarship from the Japan Society for Promotion of Science. We acknowledge financial support from the Ministry of Education of the Czech Republic (LC06010 and MSM0021622412), the Grant Agency of the Czech Republic (201/07/0927 and 203/08/0114), the Grant Agency of the Czech Academy of Sciences (IAA401630901), the North Atlantic Treaty Organization (EST.CLG.980504), and Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology, Japan and the Ministry of Agriculture, Forestry, and Fisheries, Japan. We acknowledge the Supercomputing Centre Brno for computational resources.

AUTHOR CONTRIBUTIONS

M.P. performed mutagenesis and activity measurements; M.K. performed molecular modeling and designed mutants; Z.P. performed pre-steady state kinetics measurements; R.C. performed CD spectroscopy measurements and solvent kinetic isotopic effect measurements; P.B. and M.O. designed mutants; R.C.W. contributed the RAMD modeling tool; M.T. and Y.N. contributed molecular biology tools; J.D. interpreted data and designed mutants. All authors contributed to the writing of the paper.

Published online at http://www.nature.com/naturechemicalbiology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Doyle, S.A., Fung, S.Y. & Koshland, D.E. Jr. Redesigning the substrate specificity of an enzyme: isocitrate dehydrogenase. *Biochemistry* **39**, 14348–14355 (2000).
2. Whittle, E. & Shanklin, J. Engineering delta 9–16:0-acyl carrier protein (ACP) desaturase specificity based on combinatorial saturation mutagenesis and logical redesign of the castor delta 9–18:0-ACP desaturase. *J. Biol. Chem.* **276**, 21500–21505 (2001).
3. Wymer, N. *et al.* Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from *Escherichia coli*. *Structure* **9**, 1–9 (2001).
4. Reetz, M.T., Wang, L.W. & Bocola, M. Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. *Angew. Chem. Int. Edn Engl.* **45**, 1236–1241 (2006).
5. Fox, R.J. *et al.* Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344 (2007).
6. Bartsch, S., Kourist, R. & Bornscheuer, U.T. Complete inversion of enantioselectivity towards acetylated tertiary alcohols by a double mutant of a *Bacillus subtilis* esterase. *Angew. Chem. Int. Edn Engl.* **47**, 1508–1511 (2008).
7. Rothlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
8. Woycechowsky, K.J., Vamvaca, K. & Hilvert, D. Novel enzymes through design and evolution. *Adv. Enzymol.* **75**, 241–294 (2007).
9. Morley, K.L. & Kazlauskas, R.J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **23**, 231–237 (2005).
10. van der Meer, J.R., Devos, W.M., Harayama, S. & Zehnder, A.J.B. Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol. Rev.* **56**, 677–694 (1992).
11. Janssen, D.B., Dinkla, I.J., Poelarends, G.J. & Terpstra, P. Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities. *Environ. Microbiol.* **7**, 1868–1882 (2005).
12. Swanson, P.E. Dehalogenases applied to industrial-scale biocatalysis. *Curr. Opin. Biotechnol.* **10**, 365–369 (1999).
13. Yujing, M. & Mellouki, A. Rate constants for the reactions of OH with chlorinated propanes. *Phys. Chem. Chem. Phys.* **3**, 2614–2617 (2001).
14. Bosma, T., Damborsky, J., Stucki, G. & Janssen, D.B. Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol.* **68**, 3582–3587 (2002).
15. Bosma, T., Kruizinga, E., de Bruin, E.J., Poelarends, G.J. & Janssen, D.B. Utilization of trihalogenated propanes by *Agrobacterium radiobacter* AD1 through heterologous expression of the haloalkane dehalogenase from *Rhodococcus* sp. strain m15–3. *Appl. Environ. Microbiol.* **65**, 4575–4581 (1999).
16. Yokota, T., Omori, T. & Kodama, T. Purification and properties of haloalkane dehalogenase from *Corynebacterium* sp. strain m15–3. *J. Bacteriol.* **169**, 4049–4054 (1987).
17. Janssen, D.B. *et al.* Purification and characterization of a bacterial dehalogenase with activity toward halogenated alkanes, alcohols and ethers. *Eur. J. Biochem.* **171**, 67–72 (1988).
18. Sallis, P.J., Armfield, S.J., Bull, A.T. & Hardman, D.J. Isolation and characterization of a haloalkane halidohydrolase from *Rhodococcus erythropolis* Y2. *J. Gen. Microbiol.* **136**, 115–120 (1990).
19. Kulakova, A.N., Larkin, M.J. & Kulakov, L.A. The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* **143**, 109–115 (1997).
20. Poelarends, G.J., Wilkens, M., Larkin, M.J., van Elsas, J.D. & Janssen, D.B. Degradation of 1,3-dichloropropene by *Pseudomonas cichorii* 170. *Appl. Environ. Microbiol.* **64**, 2931–2936 (1998).
21. Newman, J. *et al.* Haloalkane dehalogenase: structure of a *Rhodococcus* enzyme. *Biochemistry* **38**, 16105–16114 (1999).
22. Otyepka, M. & Damborsky, J. Functionally relevant motions of haloalkane dehalogenases occur in the specificity-modulating cap domains. *Protein Sci.* **11**, 1206–1217 (2002).
23. Verschueren, K.H.G., Seljee, F., Rozeboom, H.J., Kalk, K.H. & Dijkstra, B.W. Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* **363**, 693–698 (1993).
24. Gray, K.A. *et al.* Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Adv. Synth. Catal.* **343**, 607–617 (2001).
25. Banas, P., Otyepka, M., Jerabek, P., Petrek, M. & Damborsky, J. Mechanism of enhanced conversion of 1,2,3-trichloropropane by mutant haloalkane dehalogenase revealed by molecular modeling. *J. Comput. Aided Mol. Des.* **20**, 375–383 (2006).
26. Petrek, M. *et al.* CAVER: A new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **7**, 316 (2006).
27. Winn, P.J., Ludemann, S.K., Gauges, R., Lounnas, V. & Wade, R.C. Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. USA* **99**, 5361–5366 (2002).
28. Pavelka, A., Chovancova, E. & Damborsky, J. HotSpot Wizard: a web server for identification of hot spots in protein engineering. *Nucleic Acids Res.* **37** (web server issue): W376–W383 (2009).
29. Chovancova, E., Kosinski, J., Bujnicki, J.M. & Damborsky, J. Phylogenetic analysis of haloalkane dehalogenases. *Proteins Struct. Funct. Bioinf.* **67**, 305–316 (2007).
30. Neylon, C. Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res.* **32**, 1448–1459 (2004).
31. Bosma, T., Pikkemaat, M.G., Kingma, J., Dijk, J. & Janssen, D.B. Steady-state and pre-steady-state kinetic analysis of halopropane conversion by a *Rhodococcus* haloalkane dehalogenase. *Biochemistry* **42**, 8047–8053 (2003).
32. Schanstra, J.P. & Janssen, D.B. Kinetics of halide release of haloalkane dehalogenase: evidence for a slow conformational change. *Biochemistry* **35**, 5624–5632 (1996).
33. Prokop, Z. *et al.* Catalytic mechamism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *J. Biol. Chem.* **278**, 45094–45100 (2003).
34. Aharoni, A. *et al.* The "evolvability" of promiscuous protein functions. *Nat. Genet.* **37**, 73–76 (2005).
35. Stucki, G. & Thuer, M. Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. *Environ. Sci. Technol.* **29**, 2339–2345 (1995).
36. 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. USA* **100**, 2215–2219 (2003).
37. Devi-Kesavan, L.S. & Gao, J. 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 (2003).
38. Olsson, M.H. & Warshel, A. Solute solvent dynamics and energetics in enzyme catalysis: the S(N)2 reaction of dehalogenase as a general benchmark. *J. Am. Chem. Soc.* **126**, 15167–15179 (2004).
39. Marek, J. *et al.* Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. *Biochemistry* **39**, 14082–14086 (2000).

40. Chaloupkova, R. *et al.* Modification of activity and specificity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26 by engineering of its entrance tunnel. *J. Biol. Chem.* **278**, 52622–52628 (2003).
41. Schleinkofer, K., Sudarko, Winn, P.J., Ludemann, S.K. & Wade, R.C. Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling? *EMBO Rep.* **6**, 584–589 (2005).
42. Cojocaru, V., Winn, P.J. & Wade, R.C. The ins and outs of cytochrome P450s. *Biochim. Biophys. Acta* **1770**, 390–401 (2007).
43. Carmichael, A.B. & Wong, L.L. Protein engineering of *Bacillus megaterium* CYP102. The oxidation of polycyclic aromatic hydrocarbons. *Eur. J. Biochem.* **268**, 3117–3125 (2001).
44. Schmitt, J., Brocca, S., Schmid, R.D. & Pleiss, J. Blocking the tunnel: engineering of *Candida rugosa* lipase mutants with short chain length specificity. *Prot. Eng.* **15**, 595–601 (2002).
45. Lauble, H. *et al.* Structure determinants of substrate specificity of hydroxynitrile lyase from *Manihot esculenta*. *Protein Sci.* **11**, 65–71 (2002).
46. Fishman, A., Tao, Y., Bentley, W.E. & Wood, T.K. Protein engineering of toluene 4-monooxygenase of Pseudomonas mendocina KR1 for synthesizing 4-nitrocatechol from nitrobenzene. *Biotechnol. Bioeng.* **87**, 779–790 (2004).
47. Fedorov, R., Vasan, R., Ghosh, D.K. & Schlichting, I. Structures of nitric oxide synthase isoforms complexed with the inhibitor AR-R17477 suggest a rational basis for specificity and inhibitor design. *Proc. Natl. Acad. Sci. USA* **101**, 5892–5897 (2004).
48. Kotik, M., Stepanek, V., Kyslik, P. & Maresova, H. Cloning of an epoxide hydrolase-encoding gene from *Aspergillus niger* M200, overexpression in *E. coli*, and modification of activity and enantioselectivity of the enzyme by protein engineering. *J. Biotechnol.* **132**, 8–15 (2007).
49. Feingersch, R., Shainsky, J., Wood, T.K. & Fishman, A. Protein engineering of toluene monooxygenases for synthesis of chiral sulfoxides. *Appl. Environ. Microbiol.* **74**, 1555–1566 (2008).
50. Iwasaki, I., Utsumi, S. & Ozawa, T. New colorimetric determination of chloride using mercuric thiocyanate and ferric ion. *Bull. Chem. Soc. Jpn.* **25**, 226 (1952).