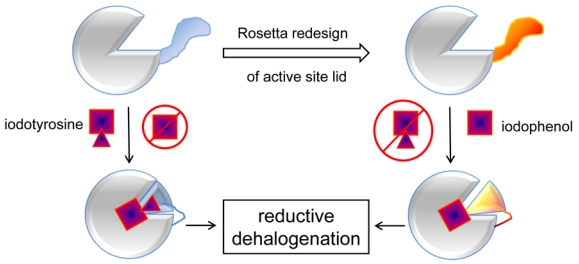
Towards a Halophenol Dehalogenase from Iodotyrosine Deiodinase via Computational Design

Zuodong Sun and Steven E. Rokita*

Department of Chemistry, Johns Hopkins University,
3400 N. Charles St., Baltimore, MD, 21218

ABSTRACT: Reductive dehalogenation offers an attractive approach for removing halogenated pollutants from the environment and iodotyrosine deiodinase (IYD) may contribute to this process after it can be engineered to accept a broad range of substrates. The selectivity of IYD is controlled in part by an active site loop of approximately 26 amino acids. In the absence of substrate, the loop is disordered and only folds into a compact helix-turn-helix upon halotyrosine association. The design algorithm of Rosetta was applied to redesign this loop for response to 2-iodophenol rather than iodotyrosine. One strategy using a restricted number of substitutions for increasing the inherent stability of the helical regions failed to generate variants with the desired properties. A series of point mutations identified strong epistatic interactions that impeded adaptation of IYD. This limitation was overcome by a second strategy that placed no restrictions on side chain substitution by Rosetta. Nine representative designs containing between 14-18 substitutions over 26 contiguous sites were evaluated experimentally. The top performing catalyst (UD08) supported a 4.5-fold increase in turnover of 2-iodophenol and suppressed turnover of iodotyrosine by 2000-fold relative to the native enzyme. The active site loop of UD08 appeared less disordered than the native sequence in the absence of substrate as evident from their relative sensitivity to proteolysis. Protection from proteolysis increased 9-fold for UD08 in the presence of 2-iodophenol and nearly rivaled the equivalent response of wild type IYD to iodotyrosine. Thus, the Rosetta designs achieved the goal of creating an active site sequence that gained structure in the presence of iodophenol. Although a limited number of point mutations was sufficient to increase the catalytic efficiency for 2-iodophenol dehalogenation, only Rosetta successfully created a loop structure responsive to this substrate.



Keywords: protein engineering; dehalogenation; Rosetta; flavin; oxidoreductase; active site lid, disorder-to-order transition

■ INTRODUCTION

Halogenated compounds are ubiquitous in our environment and many of those generated from industry resist degradation under standard methods of remediation. Contamination is so pervasive that even penguins in Antarctica were found to contain the insecticide dichlorodiphenyltrichloroethane (DDT) over 50 years ago.[1] The broad spectrum antimicrobial triclosan introduced over 40 years ago has now become one of the seven most commonly detected compounds in streams throughout the United States.[2] To date, treatments based alternatively on permanganate, persulfate, hydrogen peroxide, and ozone have been more successful than microbiological communities for degrading aryl halide pollutants, but these chemicals can also disrupt downstream processing of waste that relies on biological processes.[3] Enzymatic methods of degradation remain very attractive and should ultimately provide the most benign means to decontaminate the accumulating burden of halogenated compounds.[4-6] Hydrolytic methods based on haloalkane dehalogenase may ultimately be useful for aliphatic halides and is currently the target of engineering to increase the diversity of its substrates.[7,8] Efforts on comparable enzymes that act on aryl halides have been considerably more limited.[9] Oxidative process offer a complementary approach,[10] but this may be challenging since halogen substituents typically protect organic materials from oxidation.[11]

Reductive processes represent an appealing alternative to avoid the stabilizing effects of halogen substituents. An early candidate for this strategy of reduction is illustrated by tetrachlorohydroquinone dehalogenase and its general affiliation to the glutathione S-transferase superfamily.[12] More recently, attention has focused on anaerobic bacteria that support halorespiration for which organohalides act as terminal electron acceptors during oxidative phosphorylation.[13] The key enzyme in this process requires an unusual cobalamin derivative and is generally quite sensitive to molecular oxygen.[14] An oxygen-stable species has been recently reported but heterologous expression requires coordinated biosynthesis of its cobalamin cofactor.[15] While halorespiration has found use under anaerobic conditions, additional aerobic processes will be critical for general application on the oxygen-rich surface of the earth. The flavoenzyme iodotyrosine deiodinase (IYD) offers a feasible option that is readily available and oxygen stable.[16] This enzyme promotes reductive deiodination of mono- and diiodotyrosine (I-Tyr, I_2-Tyr, respectively) to salvage iodide for thyroid hormone biosynthesis in vertebrates.[17,18] However, IYD homologs have been discovered independent of an iodide requirement in most all metazoa as well as certain bacteria and archaea.[19,20] Examples of IYD derived from *Homo*

sapian, *Mus musculus*, and *Drosophila melanogaster* all promote reductive deiodination, debromination and dechlorination of the corresponding halotyrosines (Scheme 1).[21-23]

[insert Scheme 1]

Applications of IYD are currently limited by its specificity for halotyrosines to the near exclusion of simple halophenols that have been designated as priority pollutants.[24,25] The zwitterion of the tyrosine derivatives appears to help secure closure of an active site lid by extensive interactions with FMN, E157, Y161 and K182 (Figure 1).[26] In the absence of I-Tyr, residues 161-171 are disordered and do not generate electron density from X-ray diffraction but in the presence of I-Tyr these residues form a helix-turn-helix motif that covers the active site domain (Figure 1, residues illustrated in violet). Additionally, substrate binding induces a shift of a Thr residue into hydrogen bonding distance of the N5 position of FMN for promoting a mechanism involving single electron transfer.[26,27] Loss of the zwitterion significantly weakens binding as evident from the K_d of I-Tyr for human IYD (HsIYD) (0.09 μM, R= amino acid, Scheme 1) versus 2-iodophenol (1,410 μM, 2IP, R=H) and suppresses catalytic efficiency by 10^4-fold (k_{cat}/K_m of 1.4 x 10^4 $M^{-1}s^{-1}$ and 1.4 $M^{-1}s^{-1}$, respectively).[28] Although a bacterial IYD from *Haliscomenobacter hydrossis* (HhIYD) exhibited only an 8-fold discrimination in binding 2IP and I-Tyr, the lid region remained disordered after 2IP stacks above the FMN and turnover is still suppressed by $3x10^4$-fold relative to that of I-Tyr.[28] However, no evidence suggests that the zwitterion of I-Tyr is directly involved in the chemical mechanism of dehalogenation.[16]

When chemical rescue failed to accelerate 2IP dehalogenation as described below, the lid sequence was re-engineered to enhance catalytic degradation of halophenols. 2IP (a disinfection byproduct) was chosen as an initial substrate since the aryl iodide bond is the weakest of the common halogens and should be the most easily reduced. The ultimate goal is to generate an enzyme that will dehalogenate a range of halophenols with high efficiency. Rosetta was used to redesign the active site lid residues 157-182 to undergo a disordered to ordered transition in the presence of 2IP in analogy to that induced by I-Tyr in the native enzyme (Figure 1C). An unrestricted design (UD) strategy was found far superior to an alternative guided design (GD) strategy that was dictated by standard expectations of structural stability as described below. Variants derived from these trials contained between 9 and 18 substitutions within the 26 amino acid lid sequence. The top performing variant (UD08) switched substrate discrimination of HsIYD in favor of 2IP versus I-Tyr even though individual substitutions contributing to this variant demonstrated a high level of epistasis. Both design strategies yielded variants with a modest decrease in the disorder of the lid sequence relative to HsIYD but only UD generated a

variant (UD08) that exhibited a dramatic transition in the presence of 2IP as characterized by protection from limited proteolysis.

[Insert Figure 1]

■ RESULTS AND DISCUSSION

Chemical Rescue of 2IP Turnover by Addition of Gly and Ala to HsIYD. Often functional groups that have been deleted from an enzyme can be replaced by exogenous reagents to rescue catalytic activity. For example, turnover of an aminotransferase can be restored if a high concentration of an amine is added after mutation of an essential Lys,[29] and similarly imidazole can restore the activity of a tyrosine kinase after mutation of its essential Arg.[30] An analogous approach can be applied to the substrate of wild type HsIYDby combining Ala or Gly with 2IP to create a bimolecular version of I-Tyr. The influence of these zwitterions on the binding of 2IP to HsIYD was monitored by a standard assay based on quenching fluorescence of its active site FMN.[22] Neither Ala nor Gly at concentrations of 10 mM significantly enhanced the affinity of 2IP for HsIYD (< 15% deviation, Table S1). These conditions also did not promote the catalytic efficiency of 2IP. The presence of Gly (10 mM) increased V/[E] by only 20% and the presence of Ala (10 mM) had essentially no influence on V/[E] when using 2IP at concentrations equivalent to its K_m (4 mM)[28] and below (0.5 mM) (Table S1). Thus, chemical rescue did not offer a practical method of expanding the substrate scope of HsIYD. Consequently, computational protein engineering was used to evolve the lid sequence to accept alternative substrates lacking the zwitterion of I-Tyr.

Computational Design Using Rosetta to Predict Sequences for Stabilizing a HsIYD•2IP Complex and Promoting Dehalogenation. Redesign of the lid seemed promising since its primary function appears to control substrate specificity. At least for vertebrates, the selectivity of Tyr derivatives is necessary to protect against a futile cycle of iodinating and deiodinating intermediates during thyroid hormone biosynthesis but such specificity is not inherent to catalysis. Chemical steps of reductive dehalogenation require only the halophenol (Scheme 1). Additionally, IYD belongs to the nitroreductase superfamily that shares a common FMN-binding core structure and an additional one to three insertions forming loops surrounding the active site to provide physical and chemical specificity.[20] Engineering protein loops has received considerable attention recently but remains a challenge.[31] Examining every variant within the lid region of HsIYD (26 amino acids) is not possible experimentally and thus Rosetta was used to create a limited library to test the feasibility of adapting the lid for non-physiological

substrates. Libraries based on natural variants of enzymes are often employed in a complementary approach but this was not applicable to IYD since all homologs examined to date share the same high selectivity for halotyrosines and do not extend catalytic efficiency to halophenols.[28]

Rosetta is a powerful modeling suite capable of handling a wide range of biological problems including predicting and designing protein structure.[32,33] This tool has been successfully applied to modulating catalytic activity and substrate specificity as well as generating enzymes de novo with novel activities.[34-37] Challenges associated with engineering HsIYD include the intrinsic disorder of its active site loop in the absence of a halotyrosine and the current availability of structural information on only a complex of I-Tyr with HsIYD containing FMN rather than its reduced $FMNH_2$ derivative required for turnover. As the first attempt to alter HsIYD substrate specificity, the fixed backbone manipulations in Rosetta[38] were used to generate lead structures for analysis. A model complex of 2IP and HsIYD was first built to provide an initial structure for the evolution of sequence. This began with an optimization of the co-crystal structure of I-Tyr and HsIYD containing the active site lid in its closed form covering the active site (pdb 4TTB, Figure 1)(See also, Figure S1A and Table S2).[26] The zwitterion ion and β-carbon of I-Tyr were then deleted to generate an equivalent complex of 2IP and HsIYD (Figure 1C).

Redesign of the lid region from residues 157-182 (violet and magenta in Figure 1) utilized an iterative process involving sequential side chain replacement and backbone relaxation as summarized in Table 2S and Figure S1B. Two variations of this strategy were pursued concurrently. One was based on an unguided trajectory in which any of the 20 canonical amino acids were considered by Rosetta (UD) and a guided trajectory in which substitutions were restricted to side chains that were expected to stabilize the helical structures of the ordered lid in the presence of 2IP instead of I-Tyr (GD, see Table S3 for specific restrictions at each site). During the relaxation stage, backbone conformation and side chain rotamer populations were optimized for each new sequence. Since Rosetta sampling is stochastic in nature, each strategy was independently repeated 100 times. The resulting sequences represent only a very small statistical sampling but still sufficient to explore local energy minima and gain backbone stability.[39,40] Additionally, convergence was noted in the final sequences (Figure S2). To ensure that the gain in stabilization could be attributed to side chain substitution, a wild type control was monitored in parallel for which no side chains were substituted but the native sequence was iteratively processed by the backbone relaxation protocols.

The final output confirmed little change to the wild type structure and variants generated by computation were predicted to have greater stability with 2IP by 8 and 13 Rosetta energy units (REU) from the GD and UD protocols, respectively (Table 1). Although these gains represent only a small fraction (10-15%) of the total REU calculated for each complex, the change was deemed to be significant and much greater than the corresponding value for the wild type enzyme. Previously, a small increase of only 0.8 REU still resulted in a 2-fold increase in binding of a peptide substrate to calpain.[41] Although both protocols successfully enhanced the total energy in REU, the gain was more limited when side chain substitution was restricted (GD). Similarly, the diversity of lid sequences was more limited for GD relative to UD protocols as summarized in Figure S2.

[Insert Table 1]

The variants from both UD and GD strategies were filtered by total energy, solvent accessibility of substrate and sequence redundancy prior to biochemical analysis. First, only the 20 most stable variants in complex with 2IP were considered further. The average stability of this subset generated by GD did not differ from the original pool of 100 species (Table 1). In contrast, the most stable subset generated by UD was significantly more stable than their original pool. The implied binding energy of 2IP remained relatively constant for the wild type control and the variants produced by GD and UD (Table 1) but lid stabilization was considerably more enhanced by an average of 4- and 8-fold after side chain replacement assigned by GD and UD, respectively. In contrast, lid stabilization associated with repacking of the wild type control resulted in only a minimal gain in energy (2.4 REU) and change of structure (all atom RMSD < 0.1 Å) relative the initial input defined by crystallography.

Solvent accessible area (SASA) was next used to gauge packing density around the 2IP ligand. Although Rosetta does not specifically penalize variants for possessing a cavity, the van der Waals attraction term in the energy function of Rosetta should favor more compact structures.[42] All but one of the top 20 GD variants had a lower SASA relative to the wild type control (Table 1). This was expected for such variants since its constraints favored bulky, hydrophobic residues near the void created by the absence of the I-Tyr zwitterion (Table S3). Of the 19 GD variants with improved SASA, 17 had redundant sequences and hence only two sequences were left for analysis (GD01, GD02). Each differed from the wild type sequence by 9 side chain substitutions (Figure 2). Of the 11 UD variants with improved SASA, 9 had unique sequences (UD01-UD09) and differed from the wild type by 14 to 18 out of a possible 26 residues. Not surprisingly, the three native residues (E157, Y161, and K182) that directly

interact with the zwitterion of I-Tyr were not retained in any of the final variants other than K182 in GD01 and GD02 (Figure 2).

[insert Figure 2]

Screen for Binding Affinity and Catalytic Efficiency of HsIYD Variants Designed by Rosetta. The 11 selected variants of HsIYD designed above were generated and expressed in *Escherichia coli* using standard methods of sequence- and ligation-independent cloning and site-directed mutagenesis. Each contained an N-terminal SUMO fusion and C-terminal His_6 sequence.[26] Samples were isolated after a single Ni-NTA column and the SUMO domain was released by addition of the protease UPL1 (Figure S3). Initial screening did not require removal of SUMO from the enzyme preparation since binding and activity assays of the dehalogenase were measured based on bound FMN, not total protein. Despite extensive mutation of the lid sequence, all variants generated soluble protein and expressed well. For example, UD08 was isolated in an unoptimized yield of ~ 8 mg/ L of culture. Dissociation constants for each were measured with 2IP using a standard assay based on FMN fluorescence as noted above.[22] Although the two proteins created by the GD strategy bound 2IP more weakly than wild type HsIYD, all 9 proteins created by the UD strategy bound 2IP more strongly than wild type (Figure 3A). The catalytic efficiency of each variant was assessed with 2IP at two concentrations of 0.5 mM and 1 mM to confirm that these conditions supported first-order kinetics of turnover (Table S6). Again, variants of GD performed poorly and expressed significantly lower activity than that of wild type HsIYD with 2IP (Figure 3B). Three of the variants generated by UD (UD02, UD05 and UD08) expressed a greater efficiency of dehalogenating 2IP than the wild type. Thus, this strategy yielded a 30% success rate despite the limited number of designs. The most active variant (UD08) also bound 2IP with the greatest affinity but otherwise, a correlation between binding and catalysis was not apparent. This is consistent with studies on HhIYD from *H. hydrossis* that revealed no trend between substrate affinity and catalytic efficiency.[28] Most importantly, the UD strategy provided enzymes with the desired characteristics and provided a stark contrast to the GD strategy even though both identified a number of common mutations (Figure 2). Restricting the pool of side chains was clearly not productive and suggests a naiveté about optimizing the lid of IYD based on helix stability.

[Insert Figure 3]

The top performing variant UD08 was subsequently purified to homogeneity for more detailed analysis (Figure S4). Dehalogenation was monitored as a function of 2IP concentration to identify the kinetic constants k_{cat} and K_m (Table 2, Figure S5). The K_m value for 2IP decreased

by only 1.4-fold relative to HsIYD despite a corresponding decrease in K_d of 10-fold. However, the k_{cat} value of UD08 for 2IP increased 3.5-fold over that of HsIYD to yield a total increase in catalytic efficiency of 4.5-fold (k_{cat}/K_m). While this represents only a modest improvement in turnover, the results are still significant for a multi-step enzyme promoting redox chemistry. This level of enhancement also fits within the common range of increases from 2- to 20-fold achieved by engineering and directed evolution of enzymes such as a Diels-Alderase,[43] laccase[44] and isomerase.[45] For comparison, UD08 was also examined with I-Tyr and found to have lost much of its native selectivity. Its K_d and K_m increased by over 3000-fold and almost 11,000-fold for I-Tyr, respectively (Table 2). The corresponding k_{cat} remained relatively constant to yield a significant decrease in the k_{cat}/K_m for I-Tyr in comparision to a modest increase for 2IP (Table 3). Therefore, the design strategy of Rosetta successfully generated a dehalogenase with an enhanced ability to dehalogenate 2IP at relatively little computational and experimental cost.

Further characterization of UD08 indicated that it did not suffer from substrate or product inhibition. Turnover of 2IP and I-Tyr followed Michaelis-Menten kinetics even at high substrate concentrations of 8000 and 20,000 μM, respectively (Figure S5) and the presence of phenol at concentrations similar to that produced by UD08 did not affect dehalogenation of 2IP (Table S7). In contrast to the ability of HsIYD to dehalogenation I-Tyr and bromotyrosine with almost equal efficiency,[21] UD08 dehalogenated 2-bromophenol approximately 20-fold less efficiently than 2IP (Table S8). Substitution of 2-bromophenol at the *para* position decreased activity to levels below the detection threshold. Similarly, no dechlorination of 2-chlorophenol was observed and thus its turnover was at least 200-fold less efficient than 2IP with UD08. The aryl chloride bond is also sufficiently strong to suppress dehalogenation of chlorotyrosine relative to I-Tyr by 4- to 20-fold for various wild type IYDs.[21,23]

[insert Tables 2 and 3]

Amino Acids Responsible for the Enhanced Catalysis of UD08. Productive contributions were not expected from all 15 amino acids replaced in the active site lid of HsIYD to form UD08 (Figure 2). Some of the fifteen had the potential to stabilize the helical regions of the lid while others had the potential to contact 2IP directly (Figures S6 and S7). The model generated by Rosetta indicates that the structurally ordered lid packs with only a small void apparent (Figure S7D) and 2IP retains π-π stacking with the isoalloxazine ring of FMN (Figures S6 and S7). Additionally, aromatic side chains of the lid stack together with a Tyr of the second subunit of the native α_2-homodimer. Together, F165, W169 and L173 of the lid provide a relatively nonpolar environment for 2IP binding (Figure S7E and S7F) To identify the residues

most responsible for the improved activity with 2IP, each substitution used to convert HsIYD into UD08 was individually evaluated within the context of the wild type sequence. These new variants were generated by site-directed mutagenesis and characterized by their ability to dehalogenate 2IP in analogy to the screen above (Figure S7, Table S9). The effects of the mutations could be sorted into three groups. The first included beneficial mutations that resulted in a greater than 1.5-fold increase in V/[E] relative to wild type. This was achieved by only two species (E158Y and M162A, Figure 4). The remaining species were divided among the second group for which mutation had little effect on V/[E] (6 species) and the third group for which mutation was detrimental to activity and suffered a greater than 1.5-fold decrease in V/[E] (7 species). This distribution is actually quite favorable when compared to random mutagenesis that is typically assumed to generate favorable characteristics with a frequency of only 0.01-1%.[46]

[insert Figure 4]

By combining the individual effects of each mutation, UD08 could have been expected to exhibit a 25-fold decrease in activity with 2IP relative to the parent HsIYD rather than the 4.5-fold enhancement observed. This difference illustrates the importance of structural context surrounding each mutation, a phenomenon known as epistasis.[47] The origins of epistasis may be many fold but most usually arise from direct or even indirect interactions between residues.[47,48] The 15 mutations contributing to UD08 are all confined within a short region of 26 consecutive amino acids that form an active site lid and share numerous internal interactions. The strong epistasis encountered with the lid of HsIYD will likely be typical of most active site lids and produce a very rugged fitness landscape that may not be sufficiently responsive to a series of individual changes often generated during directed evolution. Thus, coordinated substitution of many residues such as that offered by Rosetta may prove most useful as evident from the ensemble enhancement in the catalysis supported by UD08.

The parent HsIYD was substituted with the two beneficial mutations identified above (E158Y and M162A) to create an additional variant, DM01. The effects of the combined mutations resulted in weak epistasis since 2IP turnover was enhanced by only 4.5-fold rather than the 8-fold expected from the total of the two individual effects (Figure S9, Table S10). Steady-state kinetic analysis of DM01 indicated that its K_m and k_{cat} values do not differ significantly from those of UD08 (Table 2). However, DM01 maintains an affinity for I-Tyr as predicted by the continued presence of the triad of residues necessary for interacting with the zwitterionic substrate (Figure 1B). The ability of DM01 to bind I-Tyr more strongly than its

HsIYD parent was not predicted nor was its 5-fold lower K_m and almost 4-fold higher k_{cat}/K_m with I-Tyr (Table 2). Thus, DM01 exhibits an enhanced ability to deiodinate both I-Tyr and 2IP without affecting the relative selectivity of HsIYD (Table 3). The performance of DM01 may result from stabilization gained from the ability of E158Y to participate in π-π stacking with W180 and the ability of M162A to decrease the hydrophobic surface exposed to solvent (Figure S9). These two substitutions are also evident in some of the natural variations of IYD (Figure S2).

Evolvability of Residue 157 is Diminished by Side Chain Interactions. Residue E157 is one of the three residues interacting with the zwitterion of I-Tyr in HsIYD (Figure 1B) and is strictly conserved in all IYD homologs (Figure S2).[16] The importance of this residue is further evident from the decrease in k_{cat}/K_m by almost three orders of magnitude for a Glu to Gln substitution in IYD from *Drosophila*.[23] This substitution also had a greater impact on catalysis than substitution of the corresponding active site Tyr to Phe and Lys to Gln that similarly contact the substrate zwitterion (Figure 1B).[23] For substrates lacking the zwitterion such as 2IP, a requirement for an active site Glu was not expected and might even be detrimental since its side chain would be proximal to the hydrophobic edge of 2IP according to the initial model structure (Figure 1C). Accordingly, residue 157 had the potential to be highly evolvable for expanding the substrate scope and represent a "hot spot" for engineering. We choose DM01 as the parent enzyme to evaluate the full range of substitutions at this position using saturation mutagenesis.

Nineteen variants of DM01 representing the full range of natural amino acids at position 157 were created by site-directed mutagenesis and isolated individually as described for the previous proteins designed by Rosetta (Figure S9). All expressed in *E. coli* with similar efficiency except for E157K which yielded 3- to 4-fold less enzyme as determined by SDS-PAGE and UV-vis spectroscopy. The V/[E] values of these species were again surveyed using 2IP at 0.5 mM and 1.0 mM (Figure 5, Table S10). In contrast to expectations, position 157 in DM01 exhibited very low evolvability. All but one substitution resulted in a decrease in the rate of 2IP deiodination by at least 2-fold. The hydrophobic and cationic side chains were particularly unfavorable. This result may explain in part why the GD protocol generated poor candidates since this site was restricted to nonpolar residues to complement the decrease in polarity when switching from I-Tyr to 2IP as a substrate (Table S3).[49] The only substitution supporting an increase in 2IP deiodination was the E157D variant of DM01. This result reinforced the necessity of an anionic side chain at this position in the helix since Asp is otherwise considered an α-helix breaker[50] and, in this example, proximal to the non-polar 2IP.

To test if Rosetta could recapitulate the importance of Glu or Asp in the context of DM01, the stability of each variant differing at only position 157 was computed (Figure S11). Qualitatively, the variants of low catalytic activity scored poorly and the most active scored best, but more detailed correlations were not apparent.

[Insert Figure 5]

The active final variant of DM01 represents a triple mutant (E157D/E158Y/M162A) of HsIYD and hence received the name TM01. This enzyme was subsequently purified to homogeneity and characterized by its substrate affinity and steady-state kinetics with I-Tyr and 2IP (Table 2, Figure S4). I-Tyr binds 240-fold more weakly to TM01 than to DM01 and exhibits a higher K_m by more than 90-fold. However, the k_{cat} for I-Tyr increased relative to that of DM01 by over 7-fold to hold the loss in catalytic efficiency of TM01 to 12-fold. This is a significantly smaller loss than the almost three orders of magnitude decrease observed for the single Glu to Gln mutation of *Drosophila* IYD as noted above.[23] For the target substrate 2IP, the E157D mutation in TM01 was moderately beneficial and increased its k_{cat}/K_m by 50% over that of DM01 to rival the efficiency of UD08. This was primarily achieved by a decrease in K_m since essentially no difference was detected in their k_{cat} values. The variable response to mutation of E175 once again likely reflects the rugged fitness landscape of the active site loop and the effects of local side chain interactions. Previous studies had confirmed the importance of the active site Glu to coordinate the ammonium group of I-Tyr and this interaction had overshadowed other potential contributions of Glu to stabilize the structure of the active site lid. The demand for an anionic side chain at position 157 is also likely dictated by its potential coordination to N160 and K182 (Figure S10B). Thus, context effects again dominate the results of single residue substitutions. Similar limitations caused by strong epistasis have already been implicated in slowing the evolution of natural enzymes by an order of magnitude when compared that anticipated without epistasis.[51] Rosetta overcomes this limitation by concurrent substitution of many residues. The UD protocol replaced E157 with neutral polar residues throughout UD01-UD09 based on the lower stringency of a K182H replacement preferred in all UD01-UD09 sequences (Figure 2). In contrast, the restricted approach used for generating GD01 and GD02 retained K182 and prevented its co-evolution with position 157. Approaches based on an adaptive walk may also become limited if they lack an ability for simultaneous co-evolution to overcome the constraints set by strong interactions between side chains.

Status of the Active Site Lid as Determined by Limited Proteolysis. Since direct prediction of catalytic activity is beyond the scope of Rosetta, IYD variants were ranked by their

stability in the presence of 2IP. In particular, this ligand was expected to stabilize the ordered and compact form of the active site lid in the hope that this would enhance its turnover in analogy to the effect of I-Tyr on the structure and catalytic activity of wild type HsIYD. As a measure of lid structure and its response to 2IP, IYD variants were subjected to limited proteolysis. Proteases are often used to identify flexible and disordered regions of proteins based on their high accessibility and rapid hydrolysis.[52] At least for wild type HsIYD, the unstructured active site lid was predicted to undergo proteolysis most readily whereas compaction of the lid induced by association with I-Tyr was predicted to suppress proteolysis. Indeed, HsIYD was proteolyzed by trypsin in the absence of I-Tyr to form at least two species with molecular weights of ~16 kD and ~13 kD as separated by SDS-PAGE (Figure S12A). These were extracted from the gel and characterized by LC-MS. The 16 kD species was resolved into three similar but distinct fragments with m/z values consistent with hydrolysis at K163 and R164 (Figures S13 and S14, Table S13). These sites are located at the helix-turn region of the lid where the greatest change from disorder to compact structure was expected. Characterization of the ~13 kD digestion product was not successful due to a very low signal and the likely presence of many fragments as suggested by the rather diffuse band of proteins observed from SDS-PAGE analysis. Each polypeptide migrating in the region of 16 kD contained one extra oxygen per Met residue (Table S13). This was also detected for wild type HsIYD recovered after electrophoresis but not for the native control. Therefore, oxidation of the samples occurred either during gel separation or extraction but this did not interfere with the proteolytic assay to characterize lid stabilization and compaction.

Loss of the parent protein was monitored over time by quantifying its presence after staining with Coomassie Brilliant Blue. Data fit well to a first order decay and revealed a short half-life for HsIYD of ~1.5 min (Figures 6 and S14). Addition of I-Tyr increased the half-life by ~160-fold as expected for its ability to induce formation of a compact substrate-enzyme complex.[26] In contrast, HsIYD gained a rather negligible 1.2-fold protection from trypsin after alternative addition of 2IP. Control experiments indicated that the innate activity of trypsin was not affected by I-Tyr (100 μM) and was only inhibited ~1.3-fold by the high concentration of 2IP (10 mM) used in these experiments (Figure S16). Thus, the protection afforded to HsIYD by 2IP could be solely due to the general suppression of trypsin activity. The relative susceptibility of HsIYD to proteolysis is thus consistent with prior structural and catalytic analysis of IYD from multiple organisms[26,28,53] and thus validates limited proteolysis as a method to examine HsIYD variants.

The variant (UD08) supporting greatest activity with 2IP also demonstrated most protection from trypsin even in the absence of an active site ligand (Figure 6, half-life ~8.3 min). This protection represents more than a five-fold increase relative to the parent HsIYD and suggests a modest gain of structure for the lid as anticipated by the relative energies estimated by Rosetta (in REU, Table 1). Differences in the target sequences of HsIYD and UD08 for trypsin are not likely responsible for this result since the change from -KR- in the parent to -RR- in UD08 still provides two preferred sites for recognition (Figure 2). Additionally, the product fragments for these and all variants examined are comparable (Figure S12). Most importantly, much greater protection from proteolysis was established by the presence of 2IP and the half-life of UD08 increased by 9-fold to 74 min (Figure 6). Such a half-life is only 3-fold less than that observed for the native HsIYD·I-Tyr complex. The effect of 2IP on UD08 is unique to this example and well beyond its potential to act as a mild inhibitor of trypsin. Additionally, this effect cannot be attributed to stacking of 2IP over the active site FMN since prior studies with IYD from *H. hydrossis* demonstrated that such stacking was not sufficient to induce order in the active site lid.[28] Instead, the presence of 2IP is likely capable of inducing an order to the lid of UD08 as predicted by Rosetta.

[insert Figure 6]

The success of UD08 was based in part on sampling all natural amino acid substituents in the lid region since a representative of the restricted GD strategy exhibited little protection from trypsin. GD02 persisted longer than HsIYD but only 67% as long as UD08 in the absence of 2IP and addition of 2IP yielded negligible protection from proteolysis as expected for a lid sequence that remains disordered. These results are also consistent with the very low affinity and turnover of 2IP by GD02 (Table 2). The variant DM01 created by a double mutation of the wild type HsIYD exhibited even less resistance to proteolysis and little response to 2IP in analogy to its parent (Figure 6). The efficiency of 2IP turnover by DM01 relative to wild type HsIYD is consequently not related to an ordering of the active site lid and instead reflects a general increase in activity as observed for the native substrate I-Tyr as well (Table 3). Only the Rosetta design of UD08 yielded a change of substrate preference and conformational trigger capable of inducing order in an inherently disordered active site lid. However, further stabilization of the lid may not offer significant gains in catalytic efficiency since 2IP already provides nearly a wild type level of protection of UD08 from proteolysis. Instead, more subtle adaption of the active site will likely be required to enhance the desired k_{cat}/K_m once the mechanistic details and rate determining step(s) of reductive dehalogenation are identified.

■ CONCLUSION

The potential utility of IYD to eradicate halophenols from the aerobic surface of the earth depends on its ability to adapt to new substrates. This enzyme and its homologs are easily expressed, relatively stable and structurally well characterized but efficient substrates are currently limited to halotyrosines.[16] Halophenols bind only very weakly to wild type HsIYD and even when their affinity for a bacterial IYD rivals that of I-Tyr, the active site lid remains disordered and dehalogenation is inefficient.[28] IYD variants were consequently designed by the fixed backbone protocol of Rosetta with the goal to stabilize a compact and ordered lid structure in the presence of the model substrate 2IP. Intuition on forming stable helices was used to guide one set of designs but this failed to generate proteins with the desired properties. Success was achieved instead when Rosetta was allowed to sample all natural amino acids. While the gain in catalytic efficiency was modest, suppression of I-Tyr turnover was dramatic. Dissecting the source of this change and translating the results to additional sequences was limited by the epistasis originating from numerous internal interactions established by the active site lid. While increases in 2IP dehalogenation could be achieved by as few as two mutations of HsIYD, stabilization of the lid by 2IP was only observed for a variant designed by Rosetta in which 15 of a possible 26 residues were changed. Such a large number of substitutions likely overcame the potentials barriers to evolution created by strong local interactions between side chains that would have likely stifled an alternative approach relying on the sequential accumulation of numerous point mutations. Rosetta demonstrated greatest success at changing the responsiveness of a disordered polypeptide loop from I-Tyr to 2IP without need of multiple iterations or subsequent random mutagenesis and gene shuffling.

■ MATERIALS AND METHODS

Establishing the starting structure of HsIYD for computational studies. The co-crystal structure of HsIYD and I-Tyr (PDB ID: 4TTC) was used as the template for generating a model complex of 2IP and HsIYD with its active site lid in the closed form. The conformational energy of the original PDB file was first minimized with PyRosetta3 (an interactive Python-based interface to Rosetta in which hydrogens atoms are added to the pdb files automatically)[54] (Figure S1A and Table S2).[55,56*] A single optimization trajectory was applied consisting of an initial repacking of the side chains with conformations chosen from a backbone-dependent rotamer library.[57] Next, the torsion angles of the backbone and side chains and the positions of FMN and I-Tyr were optimized by gradient based minimization. The resulting structure was

either accepted for the next round of optimization or rejected based on the Monte Carlo Metropolis criterion.[58] This strategy was performed iteratively to generate a final output and the entire process was repeated 50 times to produce 50 independent structures that were ranked by their calculated energies (REU). I-Tyr was then transformed to 2IP in the best scored structure by deleting the coordinates of its zwitterion and β-carbon. Since IYD is a homodimer with two identical and independent active sites,[53] ligand substitution and subsequent computational substitution of the lid sequence were applied to only one of the two active sites. Parameterization of the small molecules I-Tyr, 2IP, and FMN, was performed as recommended by PyRosetta.[59] The geometries of I-Tyr and FMN from the original template were adopted without change. Hydrogens and partial charges of the small molecules were added by Chem3D Pro 13.0 (Cambridgesoft) and DiscoveryStudio 4.1 (Accelrys), respectively, to generate .mol2 files for conversion to .params files that were integrated into PyRosetta.[59]

Generating variants in the lid of IYD with a fixed backbone protocol. Similar to structure optimization, a two-stage protocol was used for the fixed backbone method following published procedures (RosettaDesign, Figure S1B and Table S2).[60] In the design stage, the side chains of the active site lid (residues 157-182, Figure 1C blue and magenta) were mutated in a random and combinatorial fashion in two parallel strategies labeled unguided (UD) and guided (GD). The unguided Rosetta design selected from all 20 common amino acids for each position varied. In contrast, the guided Rosetta design restricted amino acid selection in the α-helical region (Figure 1) to those favoring this conformation.[50] In addition, selections at specific sites were limited to those conserved in IYD homologs based on the expectation that these residues should yield the most stable structures.[61] A detailed library for this guidance and its rationale are summarized in Table S3. To reduce computational cost, torsion angles of the backbone and positions of the small molecules (FMN and 2IP) remained fixed during the design stage but side chain rotamers of residues within 10 Å of the lid were repacked to accommodate side chain substitutions. The backbone and side chain conformations of the lid (residues 157-182) and positions of small molecules were optimized by gradient-based minimization. No restrictions were imposed during global repacking and minimization. The resulting structures were then accepted or rejected by the Monte Carlo Metropolis criterion. The lowest energy of the trajectory was saved as the final output. This process was independently repeated 100 times each for the guided and unguided strategies. The parent model of the HsIYD•2IP complex was also subject to the same protocols without substitution of the side chains to confirm that stabilization gained in the variants was a result of their new amino acid sequences and not merely from

conformational changes of the native backbone and side chains.

Selecting IYD variants for experimental characterization. The variants generated from the two strategies above were further ranked using two filters. First, a total Rosetta energy filter was applied to select the 20 most stable structures and second, SASA of 2IP within the active site were analyzed with Pymol (education version, Schrödinger). Variants with an SASA larger than that of the native control (6.4 $Å^2$) were discarded. After consolidating the redundant sequences, 9 variants from the unguided strategy and 2 variants from the guide strategy remained for characterization.

Cloning, expressing and purifying HsIYD variants. The plasmid pET28-SUMO-JH1 containing a gene fusion of SUMO, His_6 and HsIYD lacking its N-terminal 31 amino acids was described previously.[26] All variants were derived from this parent plasmid and generated by site-directed mutagenesis, PCR-based assembly and ligation-independent cloning as necessary (See Table S4). Proteins were expressed in Rosetta™ 2 DE3 cells at 16 ^oC after addition of 25 μM isopropyl-β-D-thiogalactoside. Cell lyste was treated with ULP1, a selective protease to release the SUMO fusion, and the variants were purified with a Hispur™ nickel column. See Supporting Information for complete details.

Ligand affinity. The affinity of HsIYD (4 uM) and its variants for I-Tyr and 2IP was determined by measuring FMN fluorescence as a function of ligand in the presence of 200 mM KCl, and 100 mM potassium phosphate pH 7.4 at 25 °C using a λ_{ex} of 450 nm and a λ_{em} of 516 nm as described previously.[22] Dissociation constants (K_d) were obtained from nonlinear fitting to equation 1 using Origin 2017.[62]

$$\frac{F}{F_o} = 1 + \frac{\Delta F}{F_o}\left(\frac{(K_d + [E]_t + [L]) - \sqrt{(K_d + [E]_t + [L])^2 - 4\,[E]_t[L]}}{2[E]_t}\right) \qquad \text{eq. 1}$$

Catalytic deiodination. Rates of dehalogenation were determined by formation of the deiodinated product (phenol, Tyr) as measured at A_{271} during reverse phase C-18 HPLC as described previously.[28] The indicated concentrations of enzyme and substrate were incubated in 900 uL containing 200 mM KCl and 100 mM potassium phosphate pH 7.4 at 25 °C for 5 min before reaction was initiated by addition of 100 uL 5% dithionite in 5% sodium bicarbonate. After 30 min, the reaction was quenched by 50 uL of 88% formic acid. An internal standard was added prior to HPLC separation as detailed in Table S11. The SUMO fusion was hydrolyzed but not removed for enzyme variants used to survey V/[E] (see Figures S3 and S8) but was removed prior to k_{cat} and K_m determinations (Table 2 and Figure S4).

Limited proteolysis of HsIYD and its variants. The deiodinases (1 μg/ml of HsIYD,

DM01, GD02 and UD08, alternatively) in the presence and absence of their corresponding substrate (100 μM I-Tyr or 10 mM 2IP) were incubated in a solution of 200 mM NaCl and 100 mM sodium phosphate pH 7.4 at 25 ± 1 °C for 5 min. Digestion was initiated by addition of TPCK-treated trypsin (Sigma-Aldrich) prepared freshly in 200 mM NaCl and 100 mM sodium phosphate pH 7.4. The final enzyme to trypsin ratio was 50:1 w/w. Aliquots (5 μL) were removed over time and mixed immediately with an equal volume of SDS-PAGE loading buffer. The mixture was then heated in a boiling water bath for 5 min and analyzed by 15% SDS-PAGE. The protein fragments were visualized with Coomassie Brilliant Blue and quantified by densitometry (ImageQuant TL 7.0, GE Healthcare) with aid of a calibration curve. Two independent sets of data quantifying consumption of the parent proteins were fit to a single exponential decay to calculate half-lives and their associated error (Origin 2017).

■ AUTHOR INFORMATION

Corresponding Author

*E-mail for S. E. R: rokita@jhu.edu

ORCID

Steven E. Rokita: 0000-0002-2292-2917

Notes

The authors declare no competing financial interest.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publication website.

Experimental methods, Tables S1-S13 and Figures S1-S16 describing the details of design protocols, enzyme mutation and characterization, structural coordinates of model HsIYD•2IP; Python script for PyRosetta simulations.

■ ACKNOWLEDGMENTS

We thank Professor Jeffrey J. Gray (Johns Hopkins University) and his laboratory for their advice and guidance on the application of Rosetta. This work was funded in part by the Johns Hopkins University Discovery Award (SER) and NSF CBET-1803771.

■ REFERENCES

1. Sladen, W. J. L.; Menzie, C. M.; Reichel, W. L. DDT Residues in Adelie Penguins and a Crabeater Seal from Antarctica. *Nature* **1966,** 210, 670-673.
2. Yueh, M. F.; Tukey, R. H. Triclosan: a Widespread Environmental Toxicant with many Biological Effects. *Annu. Rev. Pharmacol. Toxicol.* **2016,** 56, 251-272.
3. Sutton, N. B.; Atashgahi, S.; Saccenti, E.; Grotenhuis, T.; Smidt, H.; Rijnaarts, H. H. M. Microbial Community Response of an Organohalide Respiring Enrichment Culture to Permanganate Oxidation. *PLoS ONE* **2015,** 10, e0134615.
4. Häggblom, M. M.; Bossert, I. D., *Dehalogenation: Microbial Processes and Environmental Applications*. Kluwer Academic Publishers: Boston, 2003; p 520.
5. Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustáquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem. Rev.* **2017,** 117, 5619–5674.
6. Ang, T.-F.; Maiangwa, J.; Salleh, A. B.; Normi, Y. M.; Leow, T. C. Dehalogenases: from Improved Performance to Potential Microbial Dehalogenation Applications. *Molecules* **2018,** 23, 1100.
7. Sykora, J.; Brezovsky, J.; Koudelakova, T.; Lahoda, M.; Fortova, A.; Chernovets, T.; Chaloupkova, R.; Stepankova, V.; Prokop, Z.; Smatanova, I. K.; Hof, M.; J., D. Dynamics and Hydration Explain Failed Functional Transformation in Dehalogenase Design. *Nat. Chem. Biol.* **2014,** 10, 428-430.
8. Fibinger, M. P. C.; Davids, T.; Böttcher, D.; Bornscheuer, U. T. A Selection Assay for hHaloalkane Dehalogenase Activity Based on Toxic Substrates. *Appl. Microbiol. Biotechnol.* **2015,** 99, 8955-8962.
9. Wu, R.; Reger, A. S.; Cao, J.; Gulick, A. M.; Dunaway-Mariano, D. Rational Redesign of the 4-Chlorobenzoate Binding Site of 4-Chlorobenzoate Coenzyme A Ligase for Expanded Substrate Range. *Biochemistry* **2007,** 46, 14487-14499.
10. Pimviriyakul, P.; Thotsaporn, K.; Sucharitakul, J.; Chaiyen, P. Kinetic Mechanism of the Dechlorinating Flavin-dependent Monooxygenase HadA. *J. Biol. Chem.* **2017,** 292, 4818-4832.
11. Janssen, D. B.; Oppentocht, J. E.; Poelarends, G. J. Microbial Dehalogenation. *Curr. Opin. Biotech.* **2001,** 12, 254-258.
12. McCarthy, D. L.; Navarette, S.; Willett, W. S.; Babbitt, P. C.; Copley, S. D. Exploration of the Relationship between Tetrachlorohydroquinone Dehalogenase and the Glutathione S-Transferase Superfamily. *Biochemistry* **1996,** 35, 14634-14642.
13. Jugder, B.-E.; Ertan, H.; Lee, M.; Manefield, M.; Marquis, C. P. Reductive Dehalogenases Come of Age in Biological Destruction of Organohalides. *Trends Biotech* **2015,** 33, 595-610.
14. Bommer, M.; Kunze, C.; Fesseler, J.; Schubert, T.; Diekert, G.; Dobbek, H. Structural Basis for Organohalide Respiration. *Science* **2014,** 346, 455-458.
15. Payne, K. A.; Quezada, C. P.; Fisher, K.; Dunstan, M. S.; Collins, F. A.; Sjuts, H.; Levy, C.; Hay, S.; Rigby, S. E. J.; Leys, D. Reductive Dehalogenase Structure Suggests a Mechanism for B_{12}-dependent Dehalogenation. *Nature* **2015,** 517, 513-516.
16. Sun, Z.; Su, Q.; Rokita, S. E. The Distribution and Mechanism of Iodotyrosine Deiodinase Defied Expectations. *Arch. Biochem. Biophys.* **2017,** 632, 77-87.
17. Moreno, J. C.; Klootwijk, W.; van Toor, H.; Pinto, G.; D'Alessandro, M.; Lèger, A.; Goudie, D.; Polak, M.; Grüters, A.; Visser, T. J. Mutations in the Iodotryosine Deiodinase Gene and Hypothyroidism. *N. Engl. J. Med.* **2008,** 358, 1811-1818.
18. Afink, G.; Kulik, W.; Overmars, H.; de Randamie, J.; Veenboer, T.; van Cruchten, A.; Craen, M.; Ris-Stalpers, C., *J.* Molecular Characterization of Iodotyrosine Dehalogenase

Deficiency in Patients with Hypothyroidism. *Clin. Endocinol. Metab.* **2008,** 93, 4894-4901.
19. Phatarphekar, A.; Buss, J. M.; Rokita, S. E. Iodotyrosine Deiodinase: a Unique Flavoprotein Present in Organisms of Diverse Phyla. *Mol. BioSyst.* **2014,** 10, 86-92.
20. Akiva, E.; Copp, J. N.; Tokuriki, N.; Babbit, P. C. Evolutionary and Molecular Foundations of Multiple Contemporary Functions of the Nitroreductase Superfamily. *Proc. Nat. Acad. Sci. (USA)* **2017,** 114, E9549–E9558.
21. Bobyk, K. D.; Ballou, D. P.; Rokita, S. E. Rapid Kinetics of Dehalogenation Promoted by Iodotyrosine Deiodinase from Human Thyroid. *Biochemistry* **2015,** 54, 4487–4494.
22. McTamney, P. M.; Rokita, S. E. A Mammalian Reductive Deiodinase has Broad Power to Dehalogenate Chlorinated and Brominated Substrates. *J. Am. Chem. Soc.* **2009,** 131, 14212–14213.
23. Phatarphekar, A.; Rokita, S. E. Functional Analysis of Iodotyrosine Deiodinase from *Drosophila melanogaster*. *Protein. Sci.* **2016,** 25, 2187-2195.
24. Martin Ruel, S.; Espernaza, M.; Choubert, J.-M.; Valor, I.; Budzinski, H.; Coquery, M. On-site Evaluation of the Efficiency of Conventional and Advanced Secondary Processes for the Removal of 60 Organic Mircopollutants. *Water Sci. Tech.* **2010,** 62, 2970-2978.
25. United States Environmental Protection Agency (2014) Priority pollutant list.
26. Hu, J.; Chuenchor, W.; Rokita, S. E. A Switch Between One- and Two-electron Chemistry of the Human Flavoprotein Iodotyrosine Deiodinase is Controlled by Substrate. *J. Biol. Chem.* **2015,** 290, 590-600.
27. Hu, J.; Su, Q.; Schlessman, J. L.; Rokita, S. E. Redox Control of Iodotyrosine Deiodinase. *Protein Sci.* **2018,** 27, in press. (https://doi.org/10.1002/pro.3479)
28. Ingavat, N.; Kavran, J. M.; Sun, Z.; Rokita, S. E. Active Site Binding is not Sufficient for Reductive Deiodination by Iodotyrosine Deiodinase. *Biochemistry* **2017,** 56, 1130-1139.
29. Toney, M. D.; Kirsch, J. F. Direct Brønsted Analysis of the Restoration of Activity to a Mutant Enzyme by Exogenous Amines. *Science* **1989,** 243, 1485-1488.
30. Qiao, Y.; Molina, H.; Pandey, A.; Zhang, J.; Cole, P. A. Chemical Rescue of a Mutant Enzyme in Living Cells. *Science* **2006,** 311, 1293-1297.
31. Nestl, B. M.; Hauer, B. Engineering of Flexible L in Enzymes. *ACS Catalysis* **2014,** 4, 3201-3211.
32. Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K. W.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y.-E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P.; Johnson, M. L.; Brand, L. Rosetta3: An Object-oriented Software Suite for the Simulation and Design of Macromolecules. *Meth. Enzymol.* **2011**, 487, 545-574.
33. Kaufmann, K. W.; Lemmon, G. H.; DeLuca, S. L.; Sheehan, J. H.; Meiler, J. Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You. *Biochemistry* **2010,** 49, 2987-2998.
34. Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008,** 453, 190-195.
35. Murphy, P. M.; Bolduc, J. M.; Gallaher, J. L.; Stoddard, B. L.; Baker, D. Alteration of Enzyme Specificity by Computational Loop Remodeling and Design. *Proc. Natl. Acad. Sci. U S A* **2009,** 106, 9215-20

36. Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St.Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010,** 329, 309-313.
37. Gordon, S. R.; Stanley, E. J.; Wolf, S.; Toland, A.; Wu, S. J.; Hadidi, D.; Mills, J. H.; Baker, D.; Pultz, I. S.; Siegel, J. B. Computational Design of an α-Gliadin Peptidase. *J. Am. Chem. Soc.* **2012,** 134, 20513-20520.
38. Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-level Accuracy. *Science* **2003,** 302, 1364-1367.
39. Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. De Novo Eenzyme Design using Rosetta3. *PLoS One* **2011,** 6, e19230.
40. Dantas, G.; Kuhlman, B.; Callender, D.; Wong, M.; Baker, D. A Large Scale Test of Computational Protein Design: Folding and Stability of Nine Completely Redesigned Globular Proteins. *J. Mol. Biol.* **2003,** 332, 449-460.
41. Renfrew, P. D.; Choi, E. J.; Bonneau, R.; Kuhlman, B. Incorporation of Noncanonical Amino Acids into Rosetta and Use in Computational Protein-peptide Interface Design. *PLoS One* **2012,** 7, e32637.
42. Alford, R. F.; Leaver-Fay, A.; Jeliazkov, J. R.; O'Meara, M. J.; DiMaio, F. P.; Park, H.; Shapovalov, M. V.; Renfrew, P. D.; Mulligan, V. K.; Kappel, K.; Labonte, J. W.; Pacella, M. S.; Bonneau, R.; Bradley, P.; Dunbrack, R. L.; Das, R.; Baker, D.; Kuhlman, B.; Kortemme, T.; Gray, J. J. The Rosetta All-atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Compu.* **2017,** 13, 3031-3048.
43. Eiben, C. B.; Siegel, J. B.; Bale, J. B.; Cooper, S.; Khatib, F.; Shen, B. W.; Players, F.; Stoddard, B. L.; Popovic, Z.; Baker, D. Increased Diels-Alderase Activity Through Backbone Remodeling Guided by Foldit Players. *Nat. Biotech.* **2012,** 30, 190-192.
44. Santiago, G.; de Salas, F.; Lucas, M. F. t.; Monza, E.; Acebes, S.; Martinez, Ã. n. T.; Camarero, S.; Guallar, V. Computer-aided Laccase Engineering: Toward Biological Oxidation of Arylamines. *ACS Catalysis* **2016,** 6, 5415-5423.
45. Currin, A.; Dunstan, M. S.; Johannissen, L. O.; Hollywood, K. A.; Vinaixa, M.; Jervis, A. J.; Swainston, N.; Rattray, N. J. W.; Gardiner, J. M.; Kell, D. B.; Takano, E.; Toogood, H. S.; Scrutton, N. S. Engineering the "Missing Link" in Biosynthetic (−)-Menthol Production: Bacterial Isopulegone Isomerase. *ACS Catalysis* **2018,** 8, 2012-2020.
46. Romero, P. A.; Arnold, F. H. Exploring Protein Fitness Landscapes by Directed Evolution. *Nat. Rev. Mol. Cell Biol.* **2009,** 10, 866-876.
47. Miton, C. M.; Tokuriki, N. How Mutational Epistasis Impairs Predictability in Protein Evolution and Design. *Protein Sci.* **2016,** 25, 1260-1272.
48. Avoigt, C.; Kauffman, S.; Wang, Z.-G. Rational Evolutionary Design: The Theory of in vitro Protein Evolution in *Advances in Protein Chemistry*, Academic Press: 2001; Vol. 55, pp 79-160.
49. Eriksson, A. E.; Baase, W. A.; Zhang, X. J.; Heinz, D. W.; Blaber, M.; Baldwin, E. P.; Matthews, B. W. Response of a Protein Structure to Cavity-creating Mutations and Its Relation to the Hydrophobic Effect. *Science* **1992,** 255, 178-183.
50. Fersht, A. R., *In Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*. W. H. Freeman and Company: New York, 1999.
51. Breen, M. S.; Kemena, C.; Vlasov, P. K.; Notredame, C.; Kondrashov, F. A. Epistasis as the Primary Factor in Molecular Evolution. *Nature* **2012,** 490, 535-538.
52. Fontana, A.; de Laureto, P. P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. Probing Protein Structure by Limited Proteolysis. *Acta Biochim. Pol.* **2004,** 51, 299-321.

53. Thomas, S. R.; McTamney, P. M.; Adler, J. M.; LaRonde-LeBlanc, N.; Rokita, S. E. Crystal Structure of Iodotyrosine Deiodinase, a Novel Flavoprotein Responsible for Iodide Salvage in Thyroid Glands. *J. Biol. Chem.* **2009,** 284, 19659-19667.
54. Chaudhury, S.; Lyskov, S.; Gray, J. J. PyRosetta: a Script-based Interface for Implementing Molecular Modeling Algorithms Using Rosetta. *Bioinformatics* **2010,** 26, 689-691.
55. Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D. Protein-protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *J. Mol. Biol.* **2003,** 331, 281-299.
56. Wang, C.; Bradley, P.; Baker, D. Protein-protein Docking with Backbone Flexibility. *J. Mol. Biol.* **2007,** 373, 503-519.
57. Dunbrack, R. L., Jr.; Cohen, F. E. Bayesian Statistical Analysis of Protein Side-chain Rotamer Preferences. *Protein Sci.* **1997,** 6, 1661-1681.
58. Li, Z.; Scheraga, H. A. Monte Carlo-minimization Approach to the Multiple-minima Problem in Protein Folding. *Proc. Natl. Acad. Sci. (USA)* **1987,** 84, 6611-6615.
59. http://www.pyrosetta.org/obtaining-and-preparing-ligand-pdb-files
60. Kuhlman, B.; Baker, D. Native Rrotein Sequences are Close to Optimal for Their Structures. *Proc. Nat. Acad. Sci. (USD)* **2000,** 97, 10383-10388.
61. Yu, H.; Yan, Y.; Zhang, C.; Dalby, P. A. Two Strategies to Engineer Flexible Loops for Improved Enzyme Thermostability. *Sci. Rep.* **2017,** 7, 41212.
62. Warner, J. R.; Copley, S. D. Pre-steady-state Kinetic Studies of the Reductive Dehalogenation Catalyzed by Tetrachlorohydroquinone Dehalogenase. *Biochemistry* **2007,** 46, 13211-13222.

Table 1. Evaluation of Designs Generated by Rosetta

Average values of		Wild type control[a]	Guided Rosetta design[a]	Unguided Rosetta design[a]
Total energy	All 100 designs	-812.7 ± 0.2	-821 ± 1	-825 ± 2
	Top 20 designs	-812.9 ± 0.1	-821.0 ± 0.1	-829 ± 1
Binding energy of 2IP[b]	Top 20 designs	-9.4 ± 0.1	-9.6 ± 0.1	-9.8 ± 0.3
Lid stabilization energy[c]	Top 20 designs	-2.4 ± 0.1	-10.4 ± 0.2	-18 ± 1
SASA of 2IP ($Å^2$)	Top 20 designs	6.4 ± 0.3	3.2 ± 1.3	6 ± 3

[a]The energy calculated by Rosetta is expressed in Rosetta energy units (REUs) that provides a ranking of the variants and the uncertainties represent their standard deviation. [b]Binding energy of 2IP is defined as the energy of the enzyme/2IP complex minus that of the 2IP-free enzyme. [c]Lid stabilization energy is defined as the energy of the 2IP-free designed enzyme minus that of the 2IP-free starting model.

Table 2. Binding and Catalytic Constants for Enzyme Variants

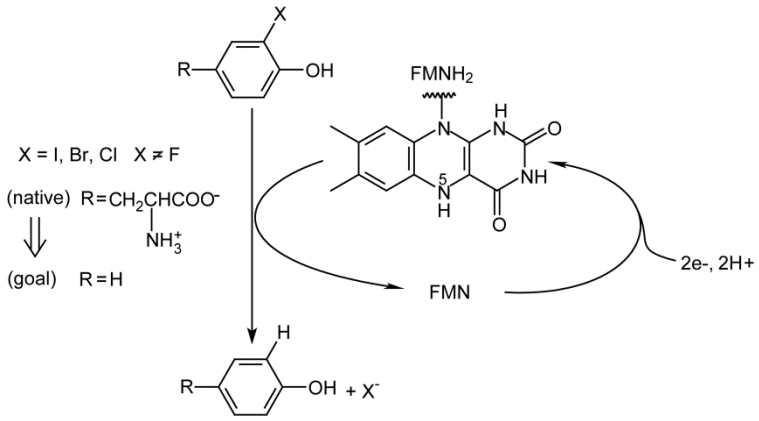
Substrate	Enzyme	K_d (mM)	k_{cat} (min^{-1})	K_m (mM)	k_{cat}/K_m ($min^{-1}\times\mu M^{-1}$)
2IP	HsIYD	2.4 ± 0.1	0.26 ± 0.01[a]	4.4 ± 0.4 [a]	$(6.0 \pm 0.6)\times10^{-5}$ [a]
	UD08	0.23 ± 0.06	0.92 ± 0.05	3.4 ± 0.4	$(2.7 \pm 0.4) \times10^{-4}$
	DM01	1.57 ± 0.03	1.03 ± 0.07	5.0 ± 0.7	$(2.1 \pm 0.3)\times10^{-4}$
	TM01	1.31 ± 0.04	0.96 ± 0.05	2.9 ± 0.4	$(3.3 \pm 0.5)\times10^{-4}$
I-Tyr	HsIYD	$(0.14 \pm 0.03)\times10^{-3}$	6.1 ± 0.4[a]	$(7.3 \pm 0.8)\times10^{-3}$[a]	0.8 ± 0.1[a]
	UD08	0.47 ± 0.01	5.7 ± 0.2	15 ± 1	$(3.9 \pm 0.4)\times10^{-4}$
	DM01	$(0.042 \pm 0.008)\times10^{-3}$	4.4 ± 0.2	$(1.4 \pm 0.4)\times10^{-3}$	3.1 ± 0.9
	TM01	0.010 ± 0.0003	32 ± 3	0.13 ± 0.003	0.25 ± 0.06

[a]Data from Ingavat et al.[28]

Table 3. Substrate Selectivity for 2IP

Enzyme	selectivity for 2IP vs I-Tyr[a]	Relative to HsIYD[b]
HsIYD	$(7.14 \pm 0.12)\times10^{-5}$	1
UD08	0.69 ± 0.11	1.4×10^{4}
DM01	$(7 \pm 2)\times10^{-5}$	1
TM01	$(1.3 \pm 0.4)\times10^{-3}$	18

[a]2IP selectivity is defined as the ratio of the k_{cat}/K_m for 2IP relative to that for I-Tyr.
[b]Relative selectivity is normalized to that of HsIYD.



Scheme 1. Catalytic Reductive Dehalogenation

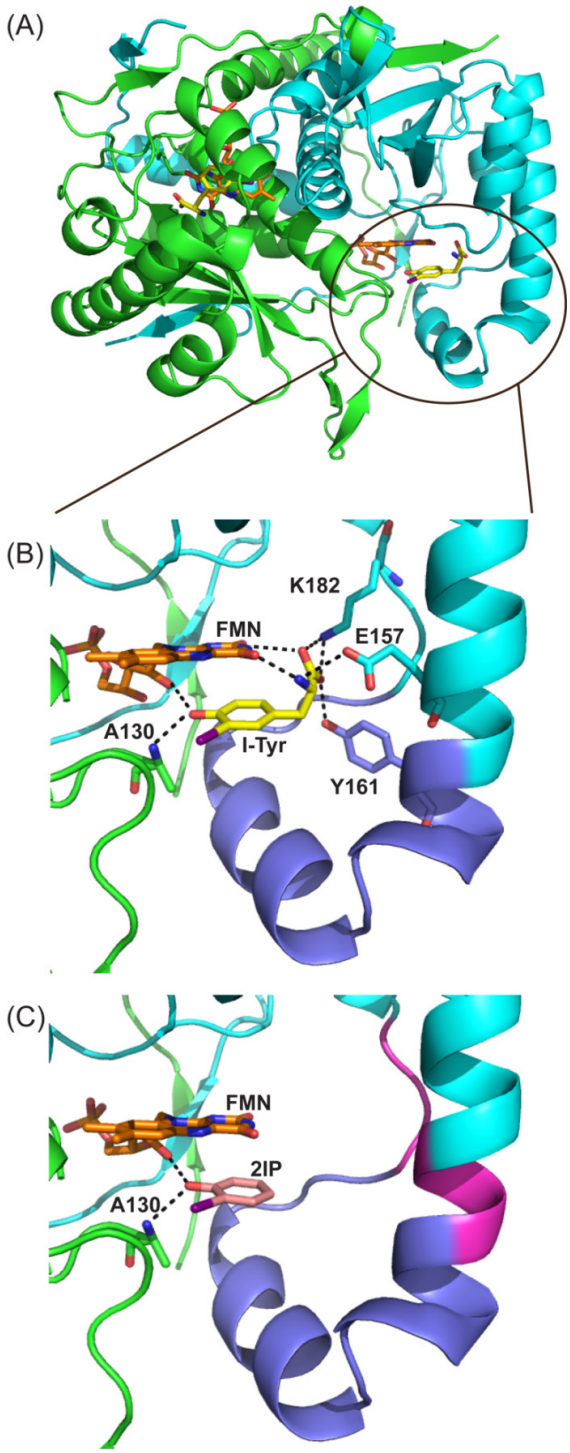


Figure 1. HsIYD in complex with its native substrate and desired target ligand. The two identical polypeptide of the native homodimer are indicated in green and cyan. Each active site contains FMN with carbon atoms depicted in orange. (A) Co-crystal structure of HsIYD and I-Tyr (carbon depicted in yellow) (PDB 4TTB). (B) An expanded view of one active site illustrating the multiple coordination between I-Tyr, FMN and protein. Residues (Y161–N179) colored in violet form the active site lid but are disordered in the absence of I-Tyr. (C) The computational model of the proposed complex of HsIYD and 2IP (carbon in pink) generated by PyRosetta.[54] The entire active site lid (purple) and three additional residues extending from both termini of the lid (magenta) were varied by RosettaDesign.[60] For clarity, residue 129 and residues 235-243 are hidden from view in (B) and (C).

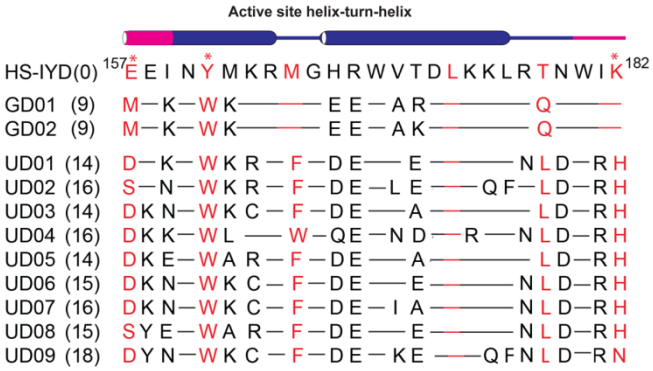


Figure 2. Sequences of HsIYD and its variants in the lid region. Conserved sequences are indicated by the horizontal lines. Residues with side chains within 5 Å of the I-Tyr in PDB 4TTB are colored in red and those forming polar interactions with the zwitterion are highlighted by a star*. The total number of amino acid substitutions relative to the native parent are shown in parenthesis for each sequence. Color coding of the helix-turn-helix cartoon above is equivalent to that introduced in Figure 1.

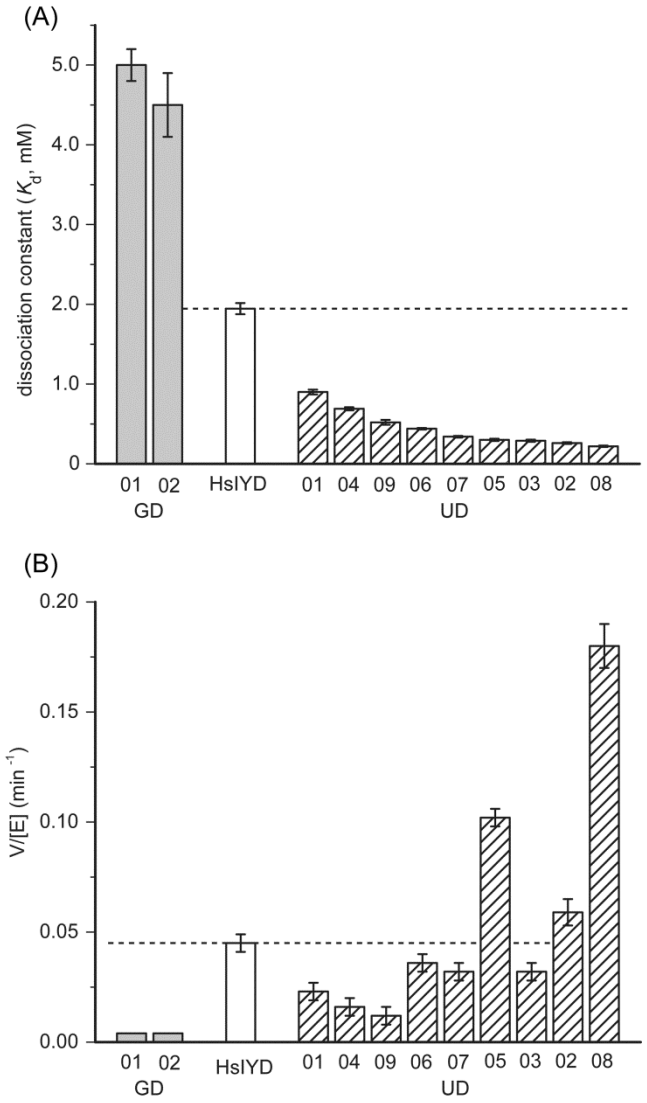


Figure 3. Characterization of IYD variants with 2IP. (A) Affinity for 2IP was measured by ligand-dependent quenching of FMN fluorescence in HsIYD variants. Two independent measurements were fit to a single binding curve to obtain K_d values and their associated error. (B) Deiodination activity (V/[E]) represents the average of two independent measurements for the variants (5 μM) with 2IP (0.5 mM). The error represents the larger value of either the range of the two measurements or three times the background of the assay (See Table S6). V/[E] for GD01 and GD02 are no larger than background signals and represent the threshold for detection. The dashed line represents the activity of WT HsIYD.

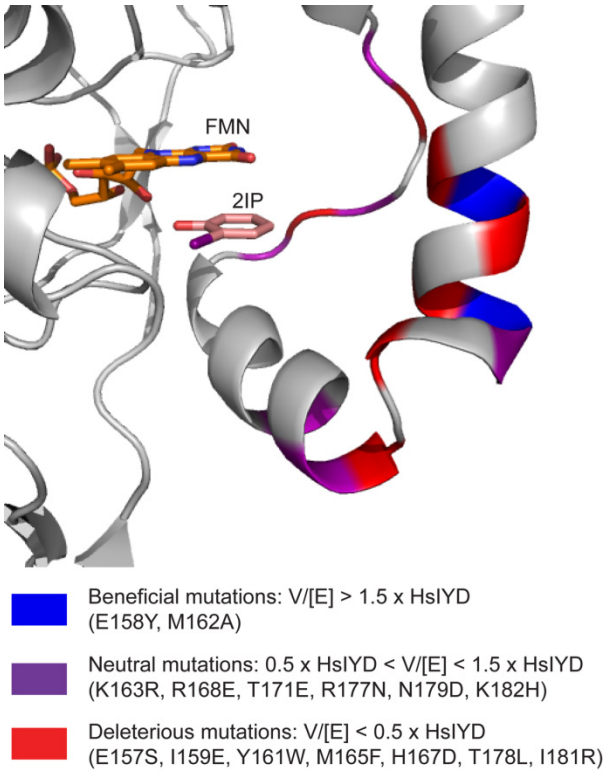


Figure 4. The effects of individual substitutions on HsIYD that combine to form UD08 are mapped onto the parent HsIYD•2IP complex. The color code illustrates the influence of these substitutions on V/[E] (see Table S9 for details).

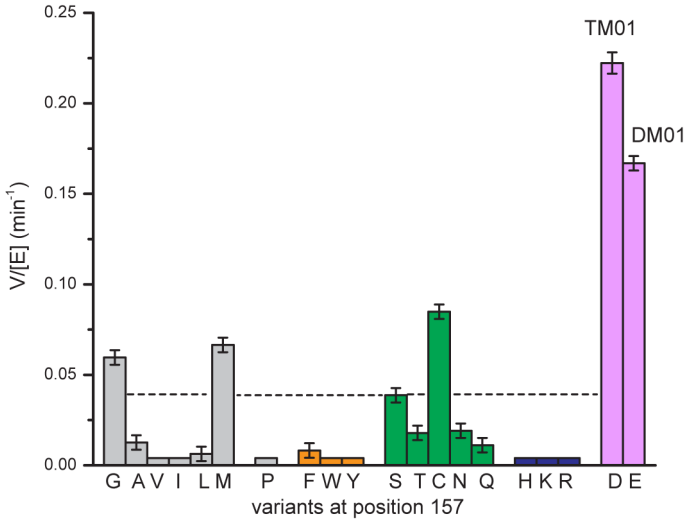


Figure 5. Deiodination of 2IP by variants of DM01 created by site-saturation mutagenesis at residue E157. Values of V/[E] represent the average of two independent measurements and the error represents the larger of either the range of two independent measurements or three times the background signal of the assay. Coloring indicates the type of amino acid side chain (gray, non-polar; orange, aromatic; green, polar; blue, positively charged; and violet, negatively charged). The dashed line represents the activity of WT HsIYD. For details, see Table S10).

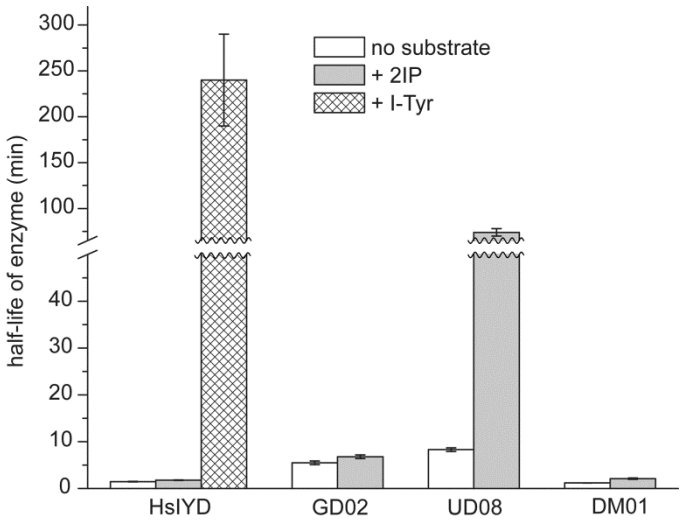


Figure 6. Half-life of HsIYD and its variants during limited proteolysis with trypsin. These values and their error derive from fitting two independent sets of data to a single exponential decay. (For details, see Table S12.)